# Rebuilding macroeconomic theory

Edited by David Vines and Samuel Wills

## Contents

The rebuilding macroeconomic theory project: an analytical assessment  
*David Vines and Samuel Wills*  
1

On the future of macroeconomic models  
*Olivier Blanchard*  
43

Ending the microfoundations hegemony  
*Simon Wren-Lewis*  
55

Where modern macroeconomics went wrong  
*Joseph E. Stiglitz*  
70

On the future of macroeconomics: a New Monetarist perspective  
*Randall Wright*  
107

Is something really wrong with macroeconomics?  
*Ricardo Reis*  
132

Good enough for government work? Macroeconomics since the crisis  
*Paul Krugman*  
156

Stagnant productivity and low unemployment: stuck in a Keynesian equilibrium  
*Wendy Carlin and David Soskice*  
169

Macro needs micro  
*Fabio Ghironi*  
195
An interdisciplinary model for macroeconomics  
_A. G. Haldane and A. E. Turrell_  
219

The financial system and the natural real interest rate: towards a 'new benchmark theory model'  
_David Vines and Samuel Wills_  
252

DSGE models: still useful in policy analysis?  
_Jesper Lindé_  
269

The future of macroeconomics: macro theory and models at the Bank of England  
_David F. Hendry and John N. J. Muellbauer_  
287

Modelling a complex world: improving macro-models  
_Warwick J. McKibbin and Andrew Stoeckel_  
329
The rebuilding macroeconomic theory project: an analytical assessment

David Vines* and Samuel Wills**

Abstract: In this paper we review the Rebuilding Macroeconomic Theory Project, in which we asked a number of leading macroeconomists to describe how the benchmark New Keynesian model might be rebuilt, in the wake of the 2008 crisis. The need to change macroeconomic theory is similar to the situation in the 1930s, at the time of the Great Depression, and in the 1970s, when inflationary pressures were unsustainable. Four main changes to the core model are recommended: to emphasize financial frictions, to place a limit on the operation of rational expectations, to include heterogeneous agents, and to devise more appropriate microfoundations. Achieving these objectives requires changes to all of the behavioural equations in the model governing consumption, investment, and price setting, and also the insertion of a wedge between the interest rate set by policy-makers and that facing consumers and investors. In our view, the result will not be a paradigm shift, but an evolution towards a more pluralist discipline.

Keywords: benchmark model, New Keynesian, financial frictions, rational expectations, heterogeneous agents, microfoundations

JEL classification: A23, A31, B22, B41, E00

The study of economics does not seem to require any specialized gifts of an unusually high order. Is it not, intellectually regarded, a very easy subject compared with the higher branches of philosophy and pure science? Yet good, or even competent, economists are the rarest of birds. An easy subject, at which very few excel! The paradox finds its explanation, perhaps, in that the master-economist must possess a rare combination of gifts. He must reach a high standard in several different directions and must combine talents not often found together. He must be mathematician, historian, statesman, philosopher—in some degree. He must understand symbols and speak in words. He must contemplate the

*Balliol College, Political Economy of Financial Markets Programme at St Antony’s College, and Institute for New Economic Thinking (INET) in the Oxford Martin School at Oxford University; and Centre for Economic Policy Research, e-mail: david.vines@economics.ox.ac.uk
**School of Economics, University of Sydney; Oxford Centre for the Analysis of Resource Rich Economies, Oxford University; Centre for Applied Macroeconomic Analysis, Australian National University, e-mail: samuel.wills@sydney.edu.au

Many thanks to Paul Luk for much help in preparing Appendix II. We thank the authors of this double issue of the Oxford Review of Economic Policy for many useful discussions. And we are grateful to the following people for many detailed comments on an earlier draft of this paper: Christopher Adam, Christopher Allsopp, Olivier Blanchard, Wendy Carlin, James Forder, Fabio Ghironi, Dieter Helm, Kevin Hoover, Colin Mayer, Ken Mayhew, Warwick McKibbin, John Muellbauer, Adrian Pagan, Alex Teytelboym, Arthur Turrell, Annie Williamson, Randall Wright, and Simon Wren-Lewis.

doi:10.1093/oxrep/grx062
© The Authors 2018. Published by Oxford University Press.
For permissions please e-mail: journals.permissions@oup.com
particular in terms of the general, and touch abstract and concrete in the same flight of thought. He must study the present in the light of the past for the purposes of the future.

Quotation from the obituary of Alfred Marshall by John Maynard Keynes (Keynes, 1924, pp. 322–3)

[T]he economic problem . . . should be a matter for specialists—like dentistry. If economists could manage to get themselves thought of as humble, competent people on a level with dentists, that would be splendid.

Quotation from an essay called ‘The Economic Possibilities for our Grandchildren’, published in Essays in Persuasion by John Maynard Keynes (Keynes, 1930, p. 373)

I. The Rebuilding Macroeconomic Theory Project

In his obituary for Alfred Marshall, published in 1924, John Maynard Keynes remarked that competent economists are rare. Then, in a few short sentences, he suggested why. Nevertheless, some 6 years later, in 1930, Keynes said that economic management should be delegated to technical experts; he hoped that they might become humble.

Why this change of view? As we discuss below, Keynes was already working on the material which would lead to his General Theory. That book was his response to the crisis of the 1930s—the Great Depression—and his realization that Marshallian economics was not enough. It would lead to an interventionist approach to macroeconomic policy, managed by specialists. Friedrich Hayek, by contrast, never believed that it would be necessary, or possible, to achieve the kind of policy competence that Keynes sought (see Hayek, 1931, 1944).1

Nearly a century later, the Great Moderation collapsed into another crisis: the global financial crisis, or GFC. When this happened, the macroeconomic experts—who were by now in charge—appeared to lack both competence and humility. As a result of the GFC we are no longer clear what macroeconomic theory should look like, or what to teach the next generation of students. We are still looking for the kind of constructive response to this crisis that Keynes produced in the 1930s.

That is why the Oxford Review of Economic Policy set up a ‘Rebuilding Macroeconomic Theory Project’. Of course, since the GFC, there have been many discussions about how to fix macro. Why do it all again?

We see a clear reason for another discussion. During the Great Moderation, the New Keynesian Dynamic Stochastic General Equilibrium (DSGE) model had become the ‘benchmark model’: the one taught to students at the start of the first-year graduate macro course. Many of us—although not all—were proud of what had been achieved.2 But the benchmark model has let us down; it explained neither why the GFC happened, nor what to do about it. What new ideas are needed? What needs to be thrown away? What might a new benchmark model look like? Will there be a ‘paradigm shift’?

1 As we also discuss below, similar views were later set out by Milton Friedman, and are still maintained by RBC theorists and freshwater economists.

2 Olivier Blanchard famously said ‘the state of macro is good’ (Blanchard, 2009). But he went on to criticize DSGE models; his paper was not Panglossian.
And how should the new model to be used in our teaching—let us call it the ‘new core model’—relate to the evidence-based professional work that macroeconomists do when giving policy advice? There has not yet been a satisfactory discussion of these questions. So we invited a group of macroeconomists to examine them with us.

To provoke initial discussion, we presented our colleagues with the following six examination questions.

(i) Is the benchmark DSGE model fit for purpose?
(ii) What additions to this model are necessary to help us understand growth from now on?
(iii) What are the important inter-temporal complications?
(iv) What are the important intra-temporal complications?
(v) Should the above questions be discussed with big or small models?
(vi) How should our models relate to data?

Appendix I provides some detail on each of the questions.

In response to these questions, Olivier Blanchard posted a piece called ‘Do DSGE Models Have a Future?’ on the Peterson Institute website. This blog-post provoked considerable interest; some of those who responded joined our project. The responses also led Blanchard to make further postings; his postings are collected together in the article which follows this one (Blanchard, 2018). To focus discussion among our authors, we organized two conferences in Oxford. The first took place in October 2016, before any papers had been written. The second, at which drafts were presented, was held in February 2017. Discussion concentrated not on particular pieces of analysis but on how a new core model might be constructed, what it might look like, and how it might behave as a whole. Very quickly, the discussion also came to focus on the relationship between the existing benchmark model, used in teaching, and the professional practice of macroeconomists, providing policy advice. Should that relationship be altered? And how might a new core model facilitate such a changed relationship?

All of the papers have been greatly re-written after that second meeting, and many of them have been discussed among the authors by email. It will be apparent to any reader of this issue of the Oxford Review that we invited people with a broad range of views to join our project. It was important to do this. The conversations have been sympathetic and wide-ranging; there has been much good humour and considerable tolerance. We think that much has been learned.

In this initial article we set out our view of what the project has achieved, framed within an account of what macroeconomists had already learned before we began.

Any understanding of what had been learned necessarily depends on being clear about how we got here. As Margaret Macmillan, the Warden of St Antony’s College, Oxford, has said recently: ‘[learning from history] is like looking in a rear-view mirror: if you only look back, you will land in the ditch, but it helps to know where you have come from and who else is on the road.’

So, in section II of the paper we provide an analytical history of two key times in the past when there was a paradigm shift: the 1930s, when Keynes invented macroeconomics, and the 1970s, when the microfoundations revolution happened. These two periods can help one to understand what a paradigm shift is, and whether we now need another one. The 1970s also shows what happens when a paradigm shift is contested and—ultimately—only partly successful.
Any understanding of what had been learned also requires some agreement on what the benchmark model actually was in 2008. To clarify the starting point we ourselves identified the 2008 benchmark as the model of Smets and Wouters (2007). Our contributors—by and large—agreed with this starting point but—when pressed—did not agree on the implications of such a starting point. We therefore provide a brief description of this model, and how it works, in section III and Appendix II.

Then, in section IV we describe the response by our contributors to the first of our examination questions: is the new-Keynesian DSGE model fit for purpose? Most of authors agree that the answer is no. Nevertheless, the wide range of responses has already led to much comment and discussion among the authors. We think that this will be of interest to our readers.

In section V we provide an account, in four parts, of what our authors think needs to be done. First, we summarize their views of what is required in a new core model. These can be grouped under four headings:

(i) incorporating financial frictions rather than assuming that financial intermediation is costless;
(ii) relaxing the requirement of rational expectations;
(iii) introducing heterogeneous agents; and
(iv) underpinning the model—and each of these three new additions—with more appropriate microfoundations.

Second, we briefly describe what the authors say about how the new core model might actually be built. We also collect these ideas together and present our own idea of what we think needs to be done. This will—we think—involves amending or replacing the main behavioural equations of the model for consumption, investment, and price-setting, and also incorporating a wedge between the interest rate set by monetary policy and the rate which affects the private sector’s decisions about consumption and investment. The outcome will be a model that is rather different from the benchmark model that was in place in 2008. Nevertheless, we characterize the likely outcome as progressive evolution; we do not think that there needs to be a paradigm shift.

Third, we outline how the new core model should relate to data. In summary, we think that it might best be estimated using Bayesian techniques, provided that much closer attention is paid to evidence from single-equation estimation informed by microeconomic ideas.

Fourth, we describe how most of our contributors have rallied around a much more general proposal: that the macroeconomics profession should delegitimize what Simon Wren-Lewis has called the microfoundations hegemony. If this happens—as we think it should—we think that the outcome will be a more pluralist approach to the subject.

In section VI we offer a brief conclusion describing what we think has been learned.

II. What we can learn from two previous paradigm shifts in macroeconomics

The purpose of this section is not to give a full run-down of the history of macroeconomic theorizing over the last 100 years. Instead, we wish to describe the way in which the economics profession has been in its current situation—a crisis exposing flaws in its
models—twice before: in the 1930s and in the 1970s. In the first case there was a clear paradigm shift. In the second case something much more confused emerged.

(i) Lessons from the 1930s: changes in content and a change in method

The punchline of the 1930s is that, prior to that time, economists only had Alfred Marshall's partial equilibrium method of analysing macroeconomic problems. Then the Great Depression came along. To explain the Depression Keynes took the Marshallian model and added nominal rigidities. This meant that, in response to a fall in investment, the economy did not rapidly return to full employment. To understand what followed, Keynes needed to invent the consumption function, the multiplier, and liquidity preference. We call these changes in content. However, to understand the implications of these changes one also needed a change in method: the kind of general-equilibrium analysis provided by the IS–LM system. This change in both content and method was a clear paradigm shift.

Keynes came to his new position only gradually. We can see what he needed to learn by examining what he said at the Macmillan Committee in 1930, which was convened by a new Labour government at the onset of the Great Depression. Keynes's task was a difficult one. Montagu Norman, the Governor of the Bank of England, said to the Committee, ‘I have never been able to see myself why for the last few years it should have been impossible for industry, starting from within, to have readjusted its own position.’ It has to be said that Keynes failed in his task: he did not know how to deal with Montagu Norman's objection; he did not yet have the necessary tools.

Here is why. Keynes was, at that stage, still a prisoner of his training in Cambridge as an expositor of Marshallian economics. The Marshallian model analysed three markets—the labour market, the goods market, and the money market—and did so separately, one market at a time.

To someone trained in the Marshallian tradition the problem of unemployment seems simple: it is caused by trade unions and other institutions keeping the wage above the market clearing level. At that wage the supply of labour exceeds the demand for labour. If wages are cut the supply of labour will fall, as some workers will no longer want a job at the lower wage. If wages are cut the demand for labour will also rise, because firms will find it profitable to hire more workers. (The demand for labour is, we now know, also influenced by the level of income and output, but this general-equilibrium idea is not present in the analysis. Income and output are both assumed to be exogenous, for reasons explained below.) A wage cut can therefore restore full employment. Such reasoning explains how Montagu Norman saw the situation: employment is determined by the wages set within each industry, which has nothing to do with the Bank of England's monetary policy. Wage adjustment, he believed, was a matter for industry and its workers. Only after writing The General Theory could Keynes see how to object to such an analysis. According to The General Theory, if wages are cut but there is no increase in aggregate demand at the same time, then firms find themselves

---

3 As Kevin Hoover has emphasized to us, Walras's analysis of general equilibrium predates Marshall. But the interactions between markets, of the kind studied in the Keynesian multiplier process and described below, were not analysed by Walras.

4 For further details about the pages which follow, see Temin and Vines (2014, 2016).
unable to sell any increase in output, and so do not increase their demand for labour. A cut in wages simply leads firms to cut their prices.

The problem of an imbalance between savings and investment was analysed in a similar way in the Marshallian tradition. The price for savings and investment was taken to be the interest rate. At a high interest rate, firms have to pay more to borrow, so the demand curve for investment slopes down. At a high interest rate, people are eager to save, so the savings supply curve slopes up. (Savings, as we now know, are also influenced by the level of income and output, but—as already noted—this general-equilibrium idea is not present in the analysis.) An interest rate above the equilibrium level will produce a supply of savings above the demand for investment. A decrease in the interest rate will lead to an equilibrium where the supply of savings equals the demand for investment. In this Marshallian framework, there is no reason ever to think that an excess supply of savings could lead to a fall in production, as Keynes wanted to argue.

It was also only after Keynes had written the *General Theory* that he saw how to object to such analysis. According to the *General Theory* the interest rate is determined in the money market by liquidity preference. It will adjust to make the demand for money equal to the supply of money, rather than adjusting to make savings equal to investment. If people decide to save more, then the interest rate need not fall by much and there might emerge a ‘paradox of thrift’: where savings are not brought into line with investment by a fall in the interest rate, but by a fall in output. This was the kind of analysis that Keynes tried to deploy at the Macmillan Committee. But such a line of argument was not yet available to him. That is because he still believed that the interest rate adjusts to ensure that investment and savings are brought into line with each other, and that resources remain fully employed.

Indeed, the very possibility that Keynes was trying to understand—low output of goods and generalized unemployment—seemed completely impossible to comprehend for people trained in the Marshallian tradition. The quantity theory of money told them that prices would adjust to make sure that this could not happen. With a fixed quantity of money, prices would adjust to ensure that the demand for money equalled this fixed supply, and that all goods produced would actually be purchased. Of course, prices needed to be flexible to make this happen. Keynes did not yet understand why it did not happen. Only by writing the *General Theory* did he come to understand that, if wages do not adjust in the labour market, the flexible prices in the goods market that Marshall had relied on would not ensure that all production would necessarily be purchased.5

The key methodological change in the *General Theory* is that Keynes began to think about the general-equilibrium implications, acknowledging that markets interact with each other. It is now generally understood that once Keynes had assumed sticky nominal wages, he needed to make four more analytical moves to create the model in the *General Theory*. First, if wages do not adjust after a decline in investment, as Montagu Norman believed they would, then there will be a shortage of aggregate demand in the goods market. Second, this fall in aggregate demand will cause consumption and savings to fall, which Keynes analysed using his new piece of equipment: the consumption function. Third, the consumption function can be used to calculate the size of the multiplier, which is needed to show how much output must fall after the decline in investment, to make savings again equal to investment. This is then needed to calculate

---

5 Keynes thought that wages did not move fast enough to ensure that the labour market continuously clears; but he also explained why this was so. See Hoover (1995).
the associated fall in employment, which is solely the result of low investment, rather than too high a level of wages. Fourth, and finally, liquidity preference is needed to show that the interest rate will not fall enough to prevent any shortage of aggregate demand from actually emerging. These four moves enabled Keynes to demonstrate that equilibrium unemployment was a genuine possibility.

Meade was the first to set out Keynes’s system formally, in a complex system of nine equations (Meade, 1937). But it was not until Hicks exogenized the price level, and so extracted the two-equation IS–LM model from Meade’s nine-equation system, that the general-equilibrium properties of the General Theory were properly understood (Hicks, 1937). The full workings of the model in the General Theory, incorporating exogenous wages but an endogenous price level (because of diminishing return to labour), were not fully explained until much later by Samuelson in his neoclassical synthesis, a much simpler system than Meade’s (Samuelson, 1951, 1955). As we have already noted, the General Theory was clearly a paradigm shift from Marshallian economics: there was both a change in content (exogenous nominal wages; consumption function; multiplier; liquidity preference) and a change in method (a move to general equilibrium). There was a period of 25 years, after the Second World War, when this new paradigm was used to guide macroeconomic policy-making. The outcomes were good: it became known as a Golden Age.

Two papers on stabilization policy published by Phillips at the time show what policy-makers were aiming to do (Phillips, 1954, 1957). Phillips showed that, in the face of shocks, a well-designed policy could help to produce good outcomes. In particular, fiscal policies might be designed, making use of PID (proportional, integral, and differential) feedbacks, which would ensure that economic activity converged rapidly to its desired level, without cyclical overshoot, even after allowing for realistic lags in private-sector behaviour. There was a flourishing of empirical macroeconomics at the time, under the influence of Frisch, Goldberger, Haavelmo, Klein, Stone, and Tinbergen. This led to the first economy-wide macroeconomic models being built, models which were used to provide empirical support for the kind of macroeconomic policy-making process described by Phillips.

(ii) Lessons from the 1970s and 1980s: two different responses, many changes in content, and a (contested) change in method

The punchline of the 1970s is that, when the Great Inflation came along, economists were no longer able to use the fixed-price IS–LM system, or the models based on it, to

---

7 Hicks saw that there were four markets in play—goods, money, bonds, and labour—but, because of nominal wage rigidity and thus a non-clearing labour market, it was only necessary to analyse equilibrium in the first three of these markets. He already had much experience, coming from writing his book Value and Capital, in showing that equilibrium in three markets could be analysed using a two-dimensional diagram, illustrating the market-clearing outcome in just two markets, as a function of the relative prices of just two goods, because of Walras’s Law. The interest rate was clearly a relative price and so belonged on the vertical axis of his diagram. It was the work of genius to realize that, because he had exogenized the price level, he could make the level of output the variable on the horizontal axis, and so obtain the IS–LM diagram.
8 The resulting model—incorporating all of the goods market, the money market, the bond market, and the labour market—is well set out in Scarth (2014).
9 See Goodfriend and King (1997, p. 234) for a brief description of the Federal Reserve System’s MPS model, perhaps the best representative of the models of that era.
give adequate policy advice. However, compared with what had happened in the 1930s, the response was not a decisive paradigm shift. Instead, there was a much more contested outcome, the consequences of which are still being felt.

The first set of responses to the Great Inflation were put forward by ‘saltwater economists’ from the US East Coast and those working in the UK, who wanted existing models to evolve. Their approach remained interventionist, but full-employment Keynesianism gave way to a regime of inflation targeting, and active fiscal policy made way for an active monetary policy. These were significant changes, but were an evolution rather than a paradigm shift. They led directly to the New Keynesian approach described in section III below.

The second response was much more of a revolution. ‘Freshwater economists’ in the US thought that the emergence of inflation had discredited active Keynesianism. Their response involved a striking change in modelling approach. First, they required that models be microfounded, optimizing, and forward-looking, with expectations of the future being model-consistent (at least up to a random error). Second, they required that the economy be treated as if it was in constant equilibrium and therefore did not require policy intervention. The first of these requirements has been largely accepted and the second comprehensively rejected. But both of these outcomes—the acceptance and the rejection—have had significant implications.

The **evolutionary approach—adaptation**

There were four steps to the evolutionary approach: incorporating a Phillips curve, allowing for adaptive inflation expectations, creating an explicit nominal anchor, and endogenizing the supply side of the model.

The first of these four steps involved adding a Phillips curve to the IS–LM model (Phillips, 1958). As Goodfriend and King (1997, pp. 235–6) say:

In the early years of the neoclassical synthesis, macroeconometric models were constructed and practical policy analysis was undertaken assuming that nominal wages and prices evolved independently from real activity and its determinants. In fact, in the 1950s, there was relatively little variability in inflation. By the mid-1960s this premise could no longer be maintained—inflation became a serious policy concern and it was plain to see that inflation was related to developments in the economy.

Adding a Phillips-curve equation to the IS–LM model greatly changed the way that macroeconomists thought about policy. Phillips (1954) had already noted that aggregate demand would need to be reduced if inflation was positive, and vice versa; 2 years after the publication of the Phillips curve paper, Samuelson and Solow (1960) argued that demand would need to be stabilized around a level of output at which inflation would be stable.10 This involved recognizing what we now call the ‘natural level of output’.

The second step involved modelling what would happen if inflation was not stabilized in this way. Both Friedman (1968) and Phelps (1968) argued that a sustained increase in aggregate demand would cause inflation to rise and that this would gradually be reflected in higher inflation expectations. That would (slowly) shift up the Phillips curve.

---

creating a ‘wage–price spiral’, something which would continue until output returned to its natural level. This argument led to the development of an expectations-augmented Phillips curve, and to a second new equation being added to the IS-LM model: describing how inflation expectations evolve gradually in response to any changes in the rate of inflation. Including this extra equation led to the Phillips curve becoming vertical at the natural level of output.

The inflationary experience of the 1960s led Friedman to argue that Keynesian policymakers had been discredited: they would inevitably cause a wage–price spiral unless there was some institutional constraint that prevented them from doing this. Kydland and Prescott (1977) and Barro and Gordon (1983) carried this argument further, suggesting that policy-makers would promise low inflation, but actually stimulate demand enough to cause inflation, thereby showing themselves to be untrustworthy.

Friedman's response to this problem was to advocate rule-based, non-interventionist monetary policy instead of active Keynesianism. For him no new theory was needed. Simply fixing the money supply was enough; an idea already embedded in the IS–LM system. With a fixed money supply the economy would—he argued—converge to the natural rate. This made short-run, interventionist macroeconomic management unnecessary. Policy-makers should just ‘fix the money supply and go home’. Such a strategy was tried in the early 1980s in Britain and (briefly) in the US, but could not be made to work. It quickly become apparent that the demand for money is unstable and that the supply of money cannot be controlled. Furthermore, it also became apparent that such an approach might, of itself, lead to macroeconomic instability. Thus monetarism, as a non-interventionist macroeconomic policy, was a dead horse.

11 Wright makes this claim in his paper in this issue.
12 Meade (1978) argued that in the UK a wage–price spiral had arisen for a rather different reason: Keynesian policies made it impossible to resist union militancy, and the resulting wage ‘explosions’. Layard et al. (1991) developed this argument by showing that union monopoly power would lower the natural level of output—as would producer market power in the product market—thus making it likely that policy-makers would overestimate the natural level of output. Orphanides (2004) provided a related explanation for what had happened in the 1970s in the US. Policy-makers had, he argued, repeatedly overestimated the natural level of output, which had fallen because of oil-price shocks and a slowdown in productivity.
13 Friedman presented his arguments using the Quantity Theory of Money, but we can encompass what he said within an IS–LM framework, even though Friedman resisted such encompassing. See the discussion by Friedman (1975) and Tobin (1975b).
14 The reason for this is that if the money supply remains unaltered in the face of a positive demand shock, cumulative inflation might emerge, rather than inflation being controlled and the economy returning to the natural level of output. This is because a demand shock will raise inflation, raise the price level, and in turn raise the nominal interest rate. That will reduce aggregate demand, and so bring output back towards its natural level. But the real interest rate interest rate will only rise if the nominal interest rate rises by more than the increase in inflation. With a fixed money supply, this might not happen if the elasticity of demand for money is sufficiently high. Of course a higher price level will also reduce the real value of assets fixed in monetary terms, in turn reducing aggregate demand. But higher prices will also redistribute wealth from creditors to debtors whose propensity to spend is high. As a result, the overall outcome might well be cumulative inflation and instability.

Something similar might happen, in reverse, in the case of a negative demand shock, resulting in the kind of debt deflation described by Fisher (1933).

Keynes had set out these arguments (informally) in ch. 19 of the General Theory, and they were first set out formally in the second part of Phillips (1954). However, these arguments were more or less completely overlooked; they were not widely recognized until they were set out again by Tobin (1975a). Even so, many macroeconomics textbooks have argued—at least until recently—that if only wages (and prices) can be made more flexible but the money supply is held constant, then any shock to aggregate demand would be rapidly corrected. Something like Marshallian economics might even re-emerge!
But if macroeconomic policy was to remain interventionist while resolving the inflation problem, then its objective needed to change. This was achieved by shifting the purpose of macroeconomic policy from ensuring full employment by managing aggregate demand, to actively anchoring nominal prices (rather than passively trying to anchor prices by fixing the money supply). This was the third step in the evolutionary approach to reforming macroeconomic policy from the 1970s onwards. It was only taken gradually.

Initially, Meade (1978, 1981) suggested that a nominal anchor could be achieved by targeting nominal income or nominal GDP; in response to inflation the policy authority would reduce aggregate demand to keep price-times-quantity constant. Weale et al. (1989) showed how an interest rate rule could be used to do this. Meade opted for this, rather than an inflation target, because he thought that the latter would be too inflexible.

It was another decade before John Taylor (1992) advocated his eponymous interest rate rule for stabilizing inflation. A Taylor rule can be a nominal anchor if it satisfies the ‘Taylor principle’: when inflation rises the nominal interest rate must rise by more, so that the real interest rate also rises. Such Taylor-rule macro only brings inflation back on target gradually, and so involves the kind of flexible inflation targeting that would have satisfied Meade. As this view was accepted it replaced the idea of fixed money supply in the IS–LM system: the requirement that the interest rate equate the demand for money to a fixed money supply was replaced by the introduction of a third new equation: the Taylor-rule (or something like it).

Taylor had initially offered his rule as a positive characterization of the way the Fed had actually implemented policy as it moved away from a Keynesian focus on stabilizing employment. But it has since become a normative recommendation for how monetary policy should be pursued in the face of shocks to output and inflation. We have learned that a Taylor rule that targets inflation will ensure that both inflation and the output gap will return to their equilibrium levels. This two-for-one property, described colourfully by Blanchard and Galí (2005) as a ‘divine coincidence’, depends on the accelerationist nature of the Phillips curve: inflation can only be on target, and unchanging, if output is equal to its natural level. Furthermore, we have learned that

15 Poole (1970) had already discussed the use of the interest rate, rather than the money supply, as the instrument of monetary policy. But he had done this within a framework in which the objective of policy still remained that of stabilizing real output, rather than controlling inflation. He showed that a good reason for choosing the interest rate, rather than the money supply, as the policy instrument might be that the variance of shocks hitting the LM curve is greater than the variance of shocks hitting the IS curve. But that paper did not address the problem being analysed by Meade, or that would later be analysed by Taylor.

16 Early proposals for the use of interest rate to control inflation were also put forward in 1992 by Henderson and McKibbin. See Henderson and McKibbin (1993).

17 See Bean (1998) and Clarida et al. (1999).

18 The output gap is the gap between the level of output and the level output would need to take for inflation to remain constant.

19 This is the case even if—as in the model set out in Appendix 2—the output gap does not explicitly feature in the monetary authority’s reaction function, as it does in a fully specified Taylor Rule.

20 This new monetary-policy regime made use of, and indeed came to require, the floating exchange rate regime that emerged after the collapse of the Bretton Woods system. A floating exchange rate enables a country to separately use its interest rate as a policy instrument in the presence of open international capital markets. In such a system it is possible, at least in principle, for policy-makers to use movements in the interest rate to insulate a country from shocks, both domestic and foreign. See the paper by McKibbin and Stoeckel in this issue of the Oxford Review of Economic Policy.
if inflation is stable, monetary policy can ensure that the resources of the economy
remain fully employed in the face of long-lasting shifts to aggregate demand, effectively
by shifting the constant term in the Taylor rule (see Allsopp and Vines, 2015).

The fourth and final step of the post-1970s evolution of macroeconomic theory
involved incorporating an endogenous supply-side. The supply side had been studied
in detail since the 1950s by the growth theory literature, coming from the models of
Solow–Swan and Ramsey.21 Incorporating an endogenous supply side into the main-
stream macroeconomic model was given an impetus by the oil shocks of the 1970s,
effectively beginning with the book by Bruno and Sachs (1985) on worldwide stagfla-
tion. The work which followed involved recognizing that output depends not just on
labour inputs but also on the stock of capital, the level of technology, and the cost
of raw-material inputs. The supply side is therefore endogenous, not least because the
capital stock depends on the level of investment and the short-run developments that
influence investment, including the interest rate. The equation for the evolution of capi-
tal added a fourth new equation to the IS–LM model.

The evolution in the structure of macroeconomic theory brought about by the
four steps which we have just described was significant. It implied that any long-run
reduction in unemployment could only be brought about by supply-side reforms that
increased investment, raised technical progress, or improved labour-market practices,
rather than by stimulating aggregate demand.22 The new macroeconomic model which
emerged is the benchmark model which we describe in the next section of the paper.
Such a model represented a major change in the way macroeconomic theory was under-
stood. But the innovations that caused the change—our four steps—were evolutionary
not revolutionary.

The revolutionary approach—a partial change in paradigm

The second response to the inflation problem of the 1960s and 1970s was much more
radical. Our discussion of it will be rather brief since Simon Wren-Lewis examines this

Supporters of this approach—freshwater economists in the US—proposed a funda-
mental change in the way in which macroeconomic models are constructed. Like
Friedman, they argued that Keynesian economics had been discredited by the inflation
of the 1960s and 1970s. But, unlike him, they thought that a whole new approach was
needed. Their new approach had two components.

First came the ‘Lucas critique’ (Lucas, 1976). This involved criticizing the use of
existing models for evaluating the effects of economic policy changes. Such models
would not—said Lucas—have a stable structure. They would be estimated on data
taken from a period before the policies were implemented, but used to predict the
behaviour of a private sector which would adapt its behaviour after the policies had
been put in place; such adaption would render the predictions invalid. But—said
Lucas—if the models allowed the private sector to form expectations about the future,
and to change these expectations in response to new policies, then the models being

21 The Klein–Goldberger model contained such a feature (Klein and Goldberger, 1955).
22 This view was first elaborated for the UK by Rowthorn (1977), by Meade (1982), and in subsequent
papers by Meade. It was taken up and developed by Layard et al. (1991) and has become an accepted part of
conventional wisdom.
used (and their evaluation of policy) could become acceptable; providing that the private sector’s expectations were consistent with the outcomes predicted by the model (‘rational expectations’).

Second, Lucas and Sargent (1979) built on this criticism to outline how models should be built to make this rational expectations analysis internally coherent. The new models should—it was said—not only incorporate rational expectations, but should also describe behaviour that was explicitly derived from optimization by economic agents in the light of such expectations. Only then could one be confident that the model would appropriately capture how the private sector would respond to any changes in external circumstances, including changes in economic policy.

Putting these two components together appeared to have a revolutionary effect: rendering macroeconomic policy both ineffective and unnecessary (Sargent and Wallace, 1976). We have already described how the evolutionary approach responded to the inflation of the 1970s by introducing an expectations augmented Phillips curve that was vertical in the long run. In that Phillips curve, inflation at time $t$ depends on expected inflation at time $t$, on the deviation of output from its natural level, and on the effects of any shocks. This formulation was widely used, not just in large-scale econometric models but also in the smaller analytical models used by macroeconomic theorists. Before the Lucas critique, expected inflation was modelled as some combination of past inflation rates. If, however, that assumption was replaced with rational expectations, then any deviations from the natural level of output must only be due to random shocks. This follows from the fact that, if expected inflation equals actual inflation, then a vertical long-run Phillips curve will imply that any deviation of output from its natural rate must be only driven by shocks. It is no surprise that many Keynesian macroeconomists at the time saw rational expectations as an existential threat. For a while, it appeared that this attack by the freshwater revolutionaries on the traditional Keynesian ideas had been fatal.

While basking in their supposed glory, freshwater economists developed real business cycle (RBC) theory, a set of ideas which attributed economic cycles to technology shocks, rather than to the aggregate demand fluctuations that had been analysed by (discredited) Keynesian theorists. In doing this they used the same insights from the Solow–Swan and Ramsey models that the more traditional economists had used to incorporate an endogenous supply side into their models (in the way which we discussed above). They showed that in a set-up with these features, productivity shocks could give rise to business cycles, in a way which now seems rather obvious.23

We now know that the freshwater attack on Keynesian economics failed because it depended not only on rational expectations and optimizing agents, but also on an inadequate formulation of the expectations-augmented Phillips curve. This was

---

23 This capital accumulation equation creates a single first-order difference equation for the level of aggregate supply. If productivity shocks are assumed to be autoregressive, which is what RBC analysts assume, this adds another first-order difference equation. The result is a second-order difference equation system for aggregate supply which can produce cycles.

In fact, RBC analysts turned out to engage rather little with the Solow–Swan–Ramsey growth literature. That is because, in most empirical work on RBC models, underlying growth is filtered out of the data using an HP filter. So what is of interest here is simply that RBC models made use of the growth-theory ideas which had already been used by the more traditional economists, rather than doing something completely different. (This point is made by Goodfriend and King, 1997.)
demonstrated in an important set of papers by Fischer (1977), Taylor (1980), and Calvo (1983). All of these papers showed that if wages or prices are not all set simultaneously, then the optimal response to a demand shock of those who set prices in the current period will depend on what inflation is expected to be in the next period, when others can adjust their prices. The same will be true next period, and so on. This kind of ‘friction’ means that changes in aggregate demand will cause changes in output as well as changes in prices. As a result, monetary and fiscal policy are able to influence output.

As Wren-Lewis describes in his article, this freshwater attempt at revolution thus had two strands. The first, which attempted to show that Keynesian policy was unnecessary and ineffective, failed. The second, which aimed to change the way academic macroeconomics is done, was successful. This success can be seen from the fact that it came to be required that all theoretical models be based on an optimizing framework with model-consistent expectations. Even those who followed the evolutionary Keynesian approach described in the previous section were now required to employ optimizing agents with model-consistent expectations. To proceed in this way, macroeconomists needed to do two things.

First, they needed to microfound the IS curve by splitting it into two components: one for consumption and one for investment. The benchmark model therefore now has an Euler equation for consumption which is based on intertemporal optimization by a representative consumer. This equation gives rise to consumption-smoothing unless interest rates vary over time. The benchmark model also now has an equation for investment which is based on profit maximization by a representative firm. Such a firm chooses investment based on the production function in the supply side of the model, considering its anticipated need for capital in the future, and the costs of adjusting the capital stock.

Second, macroeconomists also needed to microfound their model of inflation. This meant replacing the expectations-augmented Phillips curve (described above) with an equation describing the behaviour of optimizing price-setters, who adjust their prices in the knowledge that not all prices will be adjusted at the same time. This equation follows the work of Fischer (1977), Taylor (1980), and Calvo (1983), and also of Rotemberg (1982).

As a result of these changes, the New Keynesian benchmark model that evolved out of the old IS–LM system, in the way which we have described above, also came to incorporate the kind of microfounded features that had been advocated by the freshwater revolutionaries. One of the earliest models of this kind was constructed in the UK by the Meade group (Weale et al., 1989). That group, to which David Vines belonged, thought that the task of policy design had become the application of ideas from control theory, including ideas about PID control coming from Phillips (1954), to a model with a forward-looking, optimizing private sector. The central new idea was that explicit policy rules are necessary for the control of such an economy, since what people do now will depend on what they expect policy to do in the future.24 The recent, best-practice versions of models with these features include those constructed by Smets.

---

24 Another early model of this kind was that produced by Warwick McKibbin, working initially with Jeffrey Sachs (see McKibbin and Sachs, 1991, and McKibbin and Vines, 2000). John Taylor produced a similar sort of model (Taylor, 1993), although his model did not include a study of the capital accumulation process, which—we argue in this paper—it is essential to include.
These last two models form the basis of the benchmark model presented in the next section.

This change of approach happened partly because those building policy models came to realize that taking a microfounded approach would greatly assist them in their work. Microfounded models, in which the representative agents have rational expectations, make it possible to show how a policy regime can become more effective when the private sector understands the nature of the policy, and can be relied on to react optimally, in the light of this understanding. Much work in the 1990s and early 2000s showed that this effect is all the stronger if the private sector also comes to believe that the policy-maker is acting with commitment.25

Nevertheless, this dominance of microfounded methods in macroeconomics may well have been all too pervasive. Wren-Lewis (2016) describes the very large effect that the microfoundations requirement has had on those building macroeconomic models to be used for policy purposes. As he says, before the attempted revolution, policy models were empirical, using the developments in theory and in econometrics which had followed the publication of Keynes's *General Theory*. But after the attempted revolution, with its emphasis on microfoundations, even those doing policy-related work became much more deductive in their approach. Even those building policy models now see the foundations of their work as coming from basic microeconomic theory, rather than from empirical knowledge about the functioning of the macroeconomic system.

Whether this pervasiveness has been too great is one of the key questions to which we will turn. But first, we set out the New Keynesian benchmark model which emerged from all of the changes to macroeconomic theory which we have discussed in this section of the paper.

### III. The New Keynesian benchmark DSGE model

If the task is to improve on the benchmark model that was in place at the time of the 2008 crisis, we must first agree on what that benchmark was. In the interests of clarity we now provide a verbal account of the model: a New Keynesian DSGE model with investment and an endogenous capital stock, one following Christiano *et al.* (2005) and Smets and Wouters (2007).26 In Appendix II we also provide an algebraic account of the model; we do this because it is in fact hard to find a simple straightforward exposition of this model.

The benchmark model is a microfounded representative-agent model. It is clearly a general equilibrium model since there is an analysis of demand and supply in the goods market and the labour market, and also in the money market and the equity market. It includes equations for consumption, investment, and price-setting that are derived from inter-temporal optimization. Inter-temporal budget constraints are critical in determining asset prices. There is short-term stickiness in wages, and adjustment costs influence investment.

---

25 See Woodford (2003). Policies were, in fact, so successful during the Great Moderation that freshwater economists came to believe that they were not necessary. Lucas (2003) argued that ‘the central problem of depression-prevention has been solved, for all practical purposes, and has in fact been solved for many decades’. Lucas believed that this outcome meant that active countercyclical policies were not necessary. Krugman argues, in his paper below, that Lucas came to believe this precisely because the policies had been so successful.

26 See Woodford (2003, ch. 5, § 3) and Schmitt-Grohé and Uribe (2006). Galí’s text (Galí, 2015) is important, but does not include investment.
At the core of the model is a real analysis of capital accumulation and growth taken from the Solow–Swan–Ramsey growth model. A representative firm decides on investment and so brings about capital accumulation. But unlike in the growth models, output that is saved is not automatically invested. Instead, there is an explicit forward-looking investment function, depending on the expected future need for capital. The extent of investment at any point in time is governed by capital adjustment costs. The equilibrium Ramsey growth path sees investment and capital accumulation exactly keep pace with population growth and technical progress. The representative consumer follows a forward-looking Euler equation. Along the equilibrium growth path consumers hold the equity created by investment. Financial intermediation ensures that this happens, and it does so at a real interest rate that in the long term must be equal to the rate of time preference (since otherwise consumption would not be smoothed). In the short run the growth path is disturbed by shocks to the level and expected rate of change of technology, to the desire to save, and to the financial intermediation process. There is an endogenous ‘neutral’ real rate of interest which can ensure that—despite such shocks—resources remain fully employed. The model can be used to study the effects of technology shocks of the kind studied by growth theorists and real business-cycle (RBC) theorists; we display the effects of such a shock in Appendix II.

Adding nominal rigidities to this model creates the possibility of an output gap in which output, driven by changes in aggregate demand, differs from aggregate supply, so that inflation can emerge. This leads to a role for monetary policy, in the form of a central bank setting the nominal (and real) interest rate, which pins down the rate of inflation. The Taylor rule is one way of representing such a policy. Subject to inflation being controlled, such a policy can also ensure that demand is just sufficient for resources to be fully utilized—what Blanchard and Gali (2005) call the ‘divine coincidence’. Fiscal policy can also stabilize demand, but over time public deficits lead to public debts which, to ensure fiscal solvency, require higher levels of taxes to pay the higher debt interest. Public debt can also crowd out capital, but only if the consumer is treated more subtly, for example in an overlapping generations (OLG) model. The version of the model with nominal rigidities can be used to study the effects of inflation shocks and monetary policy shocks; we display the effects of an inflation shock in Appendix II.

The international version of this benchmark model joins a number of countries together through both trade linkages and asset-market arbitrage. We do not discuss international issues in this paper, but they are covered in this issue by Warwick McKibbin and Andrew Stoeckel (2018).

This is a general equilibrium model of economic growth in which there are also nominal rigidities, and so a need for a nominal anchor. As is well known, this is much more subtle than the partial equilibrium Marshallian model facing Keynes in the 1930s. It is also more subtle than the IS–LM general equilibrium model facing analysts in the 1970s,

---

27 In an OLG the real interest rate can take a higher value, in the long run, than the rate of discount of the representative consumer, since consumption smoothing is possible across a lifetime, but not between generations. This means that higher public debt can raise the real interest rate and so crowd out capital.

28 This version of the model can also be used to study the effects of a demand shock coming—for example—from a change in consumer expenditure or investment expenditure. We do not explicitly display the effects of a shock to consumption in the Appendix, but the results which we show for a negative technology demonstrate the way in which the reduced demand for investment which comes from this shock propagates through the model.
which excluded the microfounded and forward-looking behaviour of consumers and firms, any study of capital and growth, or any of simultaneity in output and inflation.

The simultaneous determination of output and inflation in ‘Taylor-rule macro’ is now widely understood and has routinely been taught to the last generation of students. But the central role of investment, capital, and growth, which is at the centre of the New Keynesian benchmark model described here, is much less well understood. Many popular treatments contain no investment by firms, only consumers who must always consume everything that is produced (see, for example, Clarida, Galí, and Gertler, 1999 (CGG); Galí, 2015). Such a model (deliberately) prevents one from understanding how the process of growth, which creates a demand for capital and investment, interacts with the short-run analysis of aggregate demand, consumption, and savings which is carried out by CGG. But, investment is central in explaining how economies respond to the kinds of shocks analysed by CGG; as well as explaining the process of long-run growth, and the RBC analysis of economic cycles. That is why, in this article, we emphasize the role of capital accumulation and investment in the benchmark model.29

The benchmark model can be used to describe, and understand, a number of recent experiences going well beyond the 1960s and 1970s, including the Asian financial crisis. McKibbin and Stoeckel (2018) discuss some of these experiences in this issue, and in our own paper (Vines and Wills, 2018) we discuss the use of this model in thinking about the question of ‘secular stagnation’.

Nevertheless, this model failed when faced with the global financial crisis. It was not good enough to give any warning of the emergence of crisis in 2008. And it has been of very little help in understanding what to do next. Notwithstanding these failings, there is not yet a new paradigm in sight, not yet a new General Theory for the twenty-first century.

IV. Is the benchmark model fit for purpose?

This is the first of the examination questions that we set our authors. The answer depends on the purpose for which the model is being used.

In his article in this issue, Blanchard (2018) identifies five different purposes and so a need for five different types of model.30 Foundational models should illuminate deep microfoundations. Core models (including our DSGE benchmark model) should provide a generally accepted theoretical framework for the profession, which should be simple enough to teach first-year graduate students.31 Policy models should closely fit data and facilitate policy analysis. Toy models (including IS–LM) are useful to provide insight for students and can provide a quick first pass at problems. Finally, forecasting models should produce the best forecasts possible. An agreement to differ across this wide range of models provides more freedom, less conflict, and makes our task of rebuilding the core model easier. Perhaps one need not worry if the new core model is not foundational, or does not fit the data, or forecast—or even look like a toy.

29 Notice also that, because this is not a model in which capital accumulation is simply determined by savings behaviour, it differs fundamentally from the kind of analysis put forward by Piketty (see Soskice, 2014).
30 For a discussion of the Blanchard taxonomy, see Ghironi (2017).
31 The second half of this sentence comes from Blanchard (2017).
In this issue we find that nearly all of our authors agree that the benchmark New Keynesian DSGE model is flawed. Most also agree with our own view, that it can and should be rebuilt rather than abandoned, though views differ greatly on what exactly will be required. However, some of our authors think that the existing benchmark DSGE model should be discarded, and that we should start again. We think that there is much to learn from their constructive opposition to the rebuilding project, and we discuss those views in this section.

The benchmark model and foundational models

Randall Wright (2018) puts forward a strong defence of foundational models. In his paper he states that RBC models have been good for understanding fluctuations during normal times, endogenous growth models have helped us understand growth and development, search models have been useful for understanding labour market behaviour and unemployment, and microfounded models of exchange have been helpful for understanding monetary issues. His paper is a thoughtful plea for the need to do more foundational work of this kind; the second part of his paper gives an idea of the insight into what can be learned about exchange by modelling the search process.

Randall Wright argues that DSGE models were never suited to the task of studying large crises.32 Joseph Stiglitz (2018) argues that these models have been good at what they are designed to do: explaining the behaviour of the macroeconomy during ‘normal times’ like the Great Moderation, and that the crisis hasn’t disproved this. Wright—no fan of DSGE models—thinks that in good times it is RBC models which have helped us to understand fluctuations. However, they both argue that a benchmark macroeconomic model should be able to explain crises because, as Stiglitz points out, it is crises that have the largest effects on individual well-being. Doctors can’t just treat colds. The answer, they argue, is to blur the boundary between foundational models and DSGE models, dragging the former into the latter.

In the views put forward by Stiglitz and Wright, the reason for the existing benchmark model’s inability to explain the crisis is its theoretical underpinnings—its microfoundations—which are the concern of foundational models. Stiglitz argues that we need microfounded institutions, noting how enforcement costs make selfish individuals and honoured contracts incompatible. He goes on to argue that DSGE makes the wrong modelling choices (as do David Hendry and John Muellbauer (2018, this issue)): complicating simple areas and simplifying complex ones; and identifies eight areas where current microfoundations are flawed. These include: the theories of consumption and expectations; investment and finance; heterogeneous agents and aggregation; and the source of and response to shocks, many of which are related. While these are important topics, it is unclear whether they can all be incorporated into a model that is parsimonious enough to teach to graduate students. There is an important—and valuable—sense in which the paper by Stiglitz declares the whole of our Rebuilding Macroeconomic Theory Project to be an impossible exercise.

Wright goes even further. Like Stiglitz, Wright proposes important suggestions for rebuilding the benchmark model. In particular, he argues that we should allow for frictions in trade from search and matching, incomplete information, and imperfect commitment, which can fundamentally change how the model works. However, the major

32 But see Eggertson and Krugman (2012).
and extraordinary challenge to the profession in his paper is expressed in a key sentence in which he complains that in many economic models, ‘there are gains from trade sitting right there on the table—the outcomes [in the models] are not even in the bilateral core’; Wright thinks that this can lead to ridiculous implications. For Wright, none of the frictions which we discuss in the next section can be allowed into a paper on macroeconomics unless they are properly microfounded as interactions between self-interested and selfish individuals. Institutions such as the use of money (or other assets) in exchange, as well as credit and banking arrangements, should emerge as outputs from, rather than inputs to the model. In the words of Wallace (1988), ‘money should not be a primitive in monetary theory’.

Some correspondence on this issue may be enlightening. David Vines suggested to Randall Wright that he cannot call on the Bursar of Balliol College every morning to renegotiate his salary, expecting Wright’s reply to be ‘Why not? And if you haven’t modelled why not then your model isn’t good enough’. Wright’s much more interesting reply was as follows:

Clearly you could call the Bursar any time you want, but I agree you do not do so very often. I interpret this as meaning it is not too important to call the Bursar all that often. In Keynesian models with Calvo pricing, people really do want to call the bursar or the boss or . . . someone all the time, because it is extremely important in those models. To wit, in such a theory the problem and the only problem with the world is sticky prices (there may be other issues, like monopsonistic competition, e.g., but these are trivial to fix with taxes/subsidies).

Not being able to call the bursar until the Calvo Fairy gives you permission is, in standard Keynesian theory, what causes inefficiencies, too much unemployment, recessions, and—really?—financial crises. [In reality you] . . . do not call the bursar every day because it does not matter; in these models you want to call him but aren’t allowed, and that assumption is truly the root of all evil. To say it just slightly differently, if you were going to lose your job because your wage was about 1 per cent off some notion of equilibrium in the sense of Debreu, then I predict you may well call someone.

In fact, this seems like a surprising response. That is because there is no dollar bill left lying on the table in the New Keynesian model. One of the main points of the paper by Blanchard and Kiyotaki (1987) was to show that not adjusting prices (or wages) is a second-order loss for firms/unions, but a first-order loss for the economy as a whole. Of course, we can agree that the Calvo story does not provide a very deep understanding of the causes of nominal rigidities. But Wright’s response to these inadequacies is to get rid of the Calvo fairy, rather than allowing the Calvo story about the effects of nominal rigidities to operate until we can do better. Wren-Lewis argues in his paper that this kind of approach has caused serious damage to policy modellers’ ability to construct useful models.

Other frictions in current DSGE models do appear to leave arbitrage opportunities on the table; Stiglitz describes this in the latter part of his paper. Wright rules out imposing such frictions unless they can be properly microfounded. Many other authors in this issue, including Stiglitz, think that it is important to include these frictions in models because they describe reality, which is crucial for policy modelling (see Wren-Lewis,
2018). Wright’s ambition of building deep microfoundations into the benchmark model is serious and worthy. But, until that is feasible, there is a risk that removing frictions from the benchmark model will prevent it from saying important things.\(^{33}\)

Andrew Haldane and Arthur Turrell (2018) also advocate deeper microfoundations in the form of agent-based models (ABMs). They would like to abandon the reliance on a single representative agent with rational expectations, and move away from the rigid ‘monoculture’ in macroeconomics in which this has become the norm. Their approach shares similarities with Wright’s, in that every behaviour in the model emerges from the interactions of individual agents facing explicit rigidities. In this approach frictions like nominal rigidities would be emergent, rather than assumed. However, they differ in that Wright believes that ‘while there is naught wrong with numerical work, in general, it is good to have a benchmark that delivers general results by hand’, while the Haldane and Turrell approach would—in addition to using current modelling approaches—aim to explore hypotheses within numerical models, and then seek more parsimonious analytical results \textit{ex post}.

The benchmark model and policy models

Jesper Lindé (2018) acknowledges the Blanchard taxonomy, but wishes to blur the boundary between DSGE models and policy models because—he argues—DSGE models were simple and flexible enough to have successfully informed policy during the crisis. In particular, he claims that such models illustrated the benefits of fiscal stimulus at the zero lower bound, and the risks of fiscal consolidation in a monetary union. In his view the flexibility of DSGE models, coupled with the ability to accommodate various extensions, means there are few contenders to take over from them—even in the policy process. Lindé also adds that while DSGE models might not have forecast the crisis (Blanchard’s fifth purpose), neither did more specialized forecasting models like Bayesian VARs (vector autoregressions).

Hendry and Muellbauer do not agree with Lindé, arguing that DSGE models are ill-suited to policy purposes. They put the poor forecast performance of DSGEs and VARs down to the lack of asset prices, credit, and financial frictions and the use of linear functional forms in both approaches. They take particular aim at the treatment of consumption. In their view, proper modelling of this variable requires two things. First, abandoning the analytical straitjacket that the Euler equation places on consumption. This does not just mean fitting a ‘looser’ form of the same equation, with different coefficients than those imposed by theory. Instead, it must link liquid, illiquid, and housing assets to consumption, even if the theory behind the linkages might not be fully understood. Fundamentally, they argue for evidence-based research: driving theory with data.

Simon Wren-Lewis provides a more extended discussion of the damage done by the DSGE hegemony to the work of building useful policy models. He describes, for the United Kingdom, how this hegemony led to funding for work on policy models being abolished. He also argues that the lack of work on such models explains the inability of

\(^{33}\) Randall Wright’s response to this sentence, when looking at a draft of this article, was ‘perhaps—but for a discussion of the future of macroeconomics we might want to be ambitious and suggest that it is best to explain nominal [and other] rigidities, not to assume them’. But should we let the best be the enemy of the good?
the UK’s policy community to understand the effects of the financial imbalances which were developing in the run-up to the global financial crisis. His approach, like that of Blanchard, is to argue that policy modellers should be free to work in the way described by Hendry and Muellbauer, rather than being constrained by a theoretical straight-jacket. He explains, with great clarity, how such a constraint can operate.

The benchmark model and toy models
Paul Krugman also (implicitly) accepts Blanchard’s taxonomy, but argues that policymakers actually relied on toy models as their default when the financial crisis came (Krugman, 2018). These were versions of the Hicksian sticky-price IS–LM set-up. Such models were, he says, good enough for what they were required to do. He claims that, while many incremental changes have been suggested to the DSGE model, there has been no single ‘big new idea’. This is because the policy responses based on IS–LM were appropriate. In particular, these suggested that large budget deficits would not drive up interest rates while the economy was at the zero lower bound, and that very large increases in the monetary base would not be inflationary, and that the multiplier on government spending was greater than one. Many people were willing to work with DSGE models, and some even considered them superior for many purposes, in agreement with Lindé. But when faced with the question of how to deal with the regime change at the zero lower bound for interest rates, many did not develop new theories, but took Hicksian predictions about policy in a liquidity trap as their starting point. So, Krugman also does not call for the core DSGE model to be rebuilt. One can recognize the MIT method at work here—‘keep it simple, stupid’—which Krugman attributes to Robert Solow (Krugman, 1992).

The benchmark model and forecasting models
Many of our authors—not just Lindé—point out that DSGE models were no good at forecasting the crisis. Some argue that this is a reason why a new framework is needed. This difficulty is not surprising; it comes from the two critical assumptions underpinning DSGE models: the efficient market hypothesis, and rational expectations. The efficient markets hypothesis gives rise to an expectations-augmented yield curve in which there is no endogenous risk premium. Furthermore, a rational expectations model like our benchmark always converges back to the Ramsey equilibrium growth path. Even if there is a very large reduction in private demand which triggers the zero bound, the economy will not collapse because of the forward-lookingness of consumption, investment, and inflation. In such a model, the efficient markets hypothesis means that things can never go seriously wrong because of a risk premium, and the rational expectations assumption of re-convergence to the Ramsey growth path means that there can never be a really serious crisis.

We have some sympathy with the argument that those who built DSGE models really did persuade themselves that the world was like their model. And if the world really is like such a model, then of course the two factors noted in the previous paragraph mean

---

34 The exception to this is his call for further work to be done on pricing behaviour, in order to understand, in particular, why a very large output gap did not lead to disinflation.

35 The list includes Stiglitz, Wright, Haldane and Turrell, and Hendry and Muellbauer.
that you would never expect things to go badly wrong—the Great Moderation will last forever. The paper by Simon Wren-Lewis implies that something like this happened.

V. Can we build a new core DSGE model that is tractable?

The core DSGE model should provide the simplest possible conceptual understanding of macroeconomic processes, rather than being explicitly designed to have policy relevance, or to be foundational, or to forecast. What should such a core DSGE model involve?

(i) Four requirements for a new core model

As already noted, we think that four key points were made by our contributors: the need for financial frictions in the model, a need to relax the rational expectations assumption, the introduction of heterogeneous agents, and underpinning the model with more appropriate microfoundations. We now summarize the contributions of our authors on each of these four issues. In the next section we outline suggestions about how the new core model might be brought together.

Financial frictions

Given that the 2008 crisis originated in the financial sector, which the benchmark DSGE model assumed works frictionlessly, it is natural that almost all authors in this issue mention financial frictions. The assumption of ‘frictionless finance’ had the deep implication that finance had no causal role to play and merely provided the financial intermediation which enabled private-sector expectations about the real economy to be realized. There is general agreement that there is a need to focus on the deep mechanisms underlying these frictions.36

The empirical case for including financial frictions in the core is outlined by Vines and Wills, and the need to integrate finance with the real economy is central to the paper by Hendry and Muellbauer. Since the crisis we have witnessed changes in spreads, changes in the yield curve and deleveraging, and an introduction of new policies like QE and dynamic macro-prudential regulation. Stiglitz further argues that the financial sector is the source of many shocks in a modern economy, either endogenously through the bursting of bubbles, or exogenously through poor policy. Furthermore, diversification does not always dissipate shocks, but can amplify them through contagion.

36 Krugman, by contrast, argues that the lack of financial frictions in the benchmark model is not a major problem. The frictions were well understood, using Diamond–Dybvig model, and the experience of the Asian financial crisis. Wright thinks that this view—that the ‘financial crisis is easy—it’s a bank run like Diamond and Dybvig’—is too simplistic. He notes that in Diamond and Dybvig banks issue simple deposit contracts in an ad hoc way where better contracts are feasible; and that endogenizing contracts might remove runs. This is another example of Wright’s refusal to include frictions which cannot be microfounded as the outcome of an optimal decision.
This has led to a number of suggestions for how core theory should respond. Liquidity constraints are raised by Blanchard, Vines and Wills, and Wright, among others. Balance sheet effects, like a stock of leverage affecting borrowing capacity, are mentioned in the papers by all of Blanchard, Hendry and Muellbauer, Stiglitz, Wren-Lewis, and Vines and Wills; Blanchard (2017) argues that ‘own funds’ affect spending decisions. In summary, ‘stocks should affect flows’: capital for banks and collateral and wealth effects for individuals. Stiglitz argues that risk has first-order effects which are often ignored, seen most clearly in the long time it takes for the collateral of banks to be restored after shocks. Vines and Wills argue that the yield curve should be endogenous—perhaps using a preferred habitat approach—in the hope of reviving the traditions of James Tobin in modern macro. Wendy Carlin and David Soskice (2018, this issue) argue for a need to include a financial accelerator and debt-financed investment in the model, and see a need for including the effects of a leveraged banking system (see also Carlin and Soskice, 2015). Lindé argues for DSGE models with an added financial sector, while Fabio Ghironi (2018, this issue) argues that financial frictions should shed light on the misallocation of resources across heterogeneous firms with market power. Stiglitz warns that it will be difficult to do any of this well.

Wright offers suggestions on how to incorporate these frictions using deep microfoundations. He argues that money, credit, and finance should emerge as outcomes of, rather than inputs to, our theories. By this he means that institutions like monetary exchange and credit should be the results of models, rather than the assumptions underlying them. The types of financial frictions he advocates modelling range across credit, banking, and contracting, set in a dynamic general equilibrium context. The results of doing this can yield new insights. For example, including an explicit role for liquidity can generate self-fulfilling prophecies, like bubbles, booms, crashes, and freezes. This approach shares similarities with Stiglitz and Haldane and Turrell, who all focus on the deep mechanisms underlying the frictions we see in the financial sector.

**Relaxing rational expectations**

The second change to the benchmark model suggested by our authors is relaxation of the requirement that rational expectations hold in all solutions of the model. Some authors, like Lindé, emphasize that the forward-looking behaviour of DSGE models is crucial in the benchmark model, because it allows us to understand how new unconventional policies, like QE and forward guidance, work. In contrast other authors, like Blanchard, Ghironi, Haldane and Turrell, and Stiglitz, all argue that the agents in our models look too far into the future, and that this leads to unrealistic consumption behaviour (the Euler equation) and price-setting behaviour (in Calvo contracting). This can have important implications for policy, for example such forward-lookingness may lead to low estimates of fiscal multipliers as agents overweight the prospect of future tax increases—as noted by Hendry and Muellbauer, and Stiglitz. Blanchard (2017) suggests incorporating finite horizons, not necessarily coming from finite lives and incomplete bequests, but instead from bounded rationality or from myopia. Haldane and Turrell suggest that a less rigid framework, like ABMs, would allow for many different degrees of rationality, and should make it possible to include the effects of heuristics that make sense in an uncertain world with costly information.
**Heterogeneous agents**

The third key addition to the benchmark model suggested by our authors is to incorporate heterogeneous agents: both consumers and producers. To do this a number of authors, including Lindé, Carlin and Soskice, Ghironi, Ricardo Reis (2018, this issue), and Vines and Wills, cite recent work by Kaplan *et al.* (2016) and Ravn and Sterk (2016) that parsimoniously includes both heterogeneous agents and search and matching frictions in a DSGE framework. Haldane and Turrell offer ABM as another way to do this. Stiglitz argues that doing this is crucial because the distribution of income matters, both for demand and for welfare outcomes. He discusses the adjustment to a negative shock; a fall in real wages can reduce demand and increase unemployment if workers have a higher marginal propensity to consume than owners of capital.

Furthermore, Ghironi argues that heterogeneous consumers alone are not sufficient, we also need heterogeneous firms. These should vary in product lines, productivity, size, and trade exposure, and should be allowed to dynamically enter and exit the market. They may also interact, strategically or in networks. He notes that this does not require completely new tools, since endogenous producer entry occurs in endogenous growth models (Romer, 1990), and since heterogeneous agents have been part of trade theory since the work of Melitz. Firm entry and exit over the business cycle affects growth through hysteresis; this approach may help us understand the slowdown since the crisis, with zombie (low-productivity) firms not exiting the market, and new firms not entering. Stiglitz adds to this discussion, suggesting that disaggregating into a number of different sectors would help to explain the structural transformation that may, at present, be contributing to slow productivity growth.

It is important to note that a proper recognition of heterogeneity will put paid to the representative agent method that has been so important in all of what we have discussed so far. Heterogeneity in general eliminates any relation between individual and aggregate behaviour. This means that having a ‘representative agent’ does not, in fact, count as a microfoundation. And attempts to allow for this by ending up, say, with two or three kinds of consumer, as in the work of Ravn and Sterk (2016), do not really get around this aggregation difficulty.

**Better microfoundations**

The final key—and perhaps most serious—addition suggested by our authors is to use more persuasive microfoundations. In the words of Ghironi, macro needs micro. While there is general agreement that we need better microfoundations, there seem to be three different interpretations of what this might mean.

The first approach, as articulated by Blanchard and Krugman, would involve improving the microfoundations in the existing core model. Krugman argues that the main modelling inadequacy identified by the crisis was on the supply side: stagflation had convinced everyone that there was a natural rate of unemployment, but the fact that sustained high unemployment did not lead to deflation during the crisis calls for a rethink. He notes that there have been surprisingly few calls to rethink ideas about inflation and the natural rate. This is essentially because understanding wages and prices is hard; we cannot always start by assuming rational behaviour and that markets reach equilibrium. Blanchard also identifies the *ad hoc* approach to understanding price stickiness as a problem, arguing that the deep reasons behind this, like the costs of
collecting information, probably have implications which reach beyond wage and price setting, and that we ignore these at our peril.

The second approach would bring into the core the approach used in building foundational models. This requires a deeper approach to microfoundations than is currently used, and is advocated by Wright. As already noted, he argues that the use of money, credit, and other assets in facilitating exchange should emerge as outcomes of, rather than inputs to our theories. He argues that there is no canonical financial-macro model with acceptable microfoundations in the way that versions of Mortensen and Pissarides (1994) or Burdett and Mortensen (1998) models are accepted as benchmarks in labour economics. A necessary requirement for any decent microfoundations—he says—should be to be able to price a dollar bill, which will require establishing a price for liquidity. Rather than models that assume the existence of banks, we need models in which banks arise endogenously. He sets a high bar: in general equilibrium there should not only be no gains from trade sitting on the table—as already mentioned—but no gains from changing institutions as well. Inadequately microfounded institutions should be thrown out of models and replaced by institutions which are modelled as the outcomes of structural frictions of an exogenous kind. He suggests that a fruitful avenue for doing this is search theory, since he thinks of the requirement to search as an exogenous friction. We should have—he says—a theory of how people trade with one another, rather than how they trade with their budget constraints. There should be explicit frictions that make the process non-trivial, and institutions should arise endogenously to facilitate this process. He goes on to propose a simple framework for doing this, noting we should aim to do this analytically, at least for core models. It seems there is an analogy in the natural sciences, chemistry has been ‘microfounded’ through the use of explanations from quantum physics.37

The third approach would bring a radically different approach to microfoundations in the core model, relying on simulation methods rather than analytical models. This is agent-based modelling and is advocated by Haldane and Turrell. Like Wright, they suggest modelling the way in which individual agents trade or interact with each other. However, unlike Wright’s proposal, such an approach would be based on many heterogeneous agents, and would need to be solved numerically rather than analytically. Such an approach would—they say—enable one to study complexity, networks, herding, and irrational behaviour. The major advantage of this approach is that it makes it possible to introduce behavioural assumptions for individual agents, and to study systems that are out of equilibrium and in which markets don’t clear. This can lead to phase transitions, like the way that opinions form within populations. Blanchard suggests that this approach may have merit, but criticizes the approach for not having yet produced a core model. Haldane and Turrell respond to this challenge by suggesting that a core model is not required, rather they think that the ABM approach should be seen as offering a flexible, modular toolkit. They argue that such an approach might enable one to study all the following features: nominal rigidities, bounded rationality and limited horizons, and incomplete markets and a role for debt. The difference is that in the ABM framework these will be emergent, rather than primitives.

37 But we would note that this has not led to a demand for all explanations in chemistry to be microfounded in this way.
(ii) **How might we incorporate these suggestions into a new core model?**

A number of authors have made detailed suggestions as to what the new core model might look like, and many valuable proposals can be found in what they say. Here we make two observations about the overall frameworks which are proposed.

In a way, the least radical proposal is that from McKibbin and Stoeckel. They are clearly engaged in building a policy model. But they also see their model as a core model, arguing that it contains a valuable framework within which to carry out simulations of actual events, and of hypothetical possibilities. They see such a framework as valuable because it constrains the outcomes of simulations within the framework of a well-understood theoretical structure—that of the existing benchmark model outlined in section III and Appendix II of this paper—and they think that such a structure causes long-run outcomes to be coherent in well-understood ways. Of course they would allow for short-term deviations from the microfounded rationality embodied in this benchmark model, but they nevertheless see having such a structure as something which is important for a policy model as well as for a core model. Nevertheless their proposal is also a radical one in another sense, since they suggest that the core model should have output disaggregated into a number of different commodities and services, and that it should identify the nature of links between different countries.

In a way, the most radical proposal is that of Carlin and Soskice. They aim to provide a model focusing on the short to medium run; one which is capable of providing a clear explanation of why recovery since the global financial crisis has been so slow. They describe the possibility of a Keynesian unemployment equilibrium, whose existence is underpinned by five fundamental assumptions: a zero bound to interest rates, the absence of disinflation in the presence of high unemployment, strategic complementarities among investors capable of giving rise to multiple equilibria, the assumption that technical progress is embodied in investment so that a low-investment outcome will give rise to a low rate of technical progress, and sufficient myopia among investors and consumers that the possibility of a good outcome in the future does not cast such an optimistic shadow over the present as to rule out the possibility of a Keynesian unemployment equilibrium. Their set-up is one in which there is also a good outcome—what they call a Wicksellian equilibrium—in which resources of the economy are fully employed, the economy grows, and real wages rise, all in the way contemplated by the benchmark New Keynesian model.\(^{38}\)

But their achievement is to show that another different equilibrium is possible, the Keynesian unemployment equilibrium. It is possible to regard their set-up as underpinning the kind of ideas put forward by Krugman in his article. He suggests that, where necessary, we should go on using, with a clear conscience, the kind of IS–LM framework which was the standard model in the 1960s, reassured that this approach proved useful in thinking about the right policy responses at the time of the recent financial crisis.

---

\(^{38}\) Wendy Carlin and David Soskice have already provided, in their macroeconomics textbook, a careful comparison of a number of different modelling approaches with the approach embodied in the benchmark New Keynesian model. See Carlin and Soskice (2015, ch. 16). The model presented in their paper in this issue of the *Oxford Review* could be added to the taxonomy discussed in that chapter in a fruitful manner.
We now make our own suggestion about what needs to be done, drawing on the points identified in the previous section of this article. We argue that four changes are needed. We think that the three behavioural equations of the model: describing consumption, investment, and price-setting, must all be amended or replaced. In addition, we argue that a gap should be introduced between the rate of interest set by monetary policy-makers and the rate affecting consumption and investment decisions.

The treatment of consumption needs to recognize finite horizons, liquidity constraints, overlapping generations (a ‘death distortion’), a distribution of consumers, and also make allowance for the fact that consumers hold housing as major asset class. An overlapping generations structure would make it possible for the equilibrium real interest rate to stay above the marginal product of capital for extended periods (or even permanently), which may help explain the slow recovery. A distribution of consumers with different levels of wealth would make it possible to study inequality, and also the first-order effects of redistribution on aggregate consumption (due to different marginal propensities to consume). Including housing would make it possible to study consumers’ decisions to invest in houses rather than productive capital, which in turn lowers real income in the long term. It appears possible to include this in a simple overlapping generations model (see Barrell and Weale, 2010, and Wang, 2011). However, this will be insufficient if the purpose is to study the house price booms and collapses that Hendry and Muellbauer argue must be examined in a policy model. Such an analysis would require a more complex treatment that also considers the down-payment constraints and varying access to equity withdrawal that characterize home loans.

The Tobin’s Q investment equation needs to be replaced by one which allows for liquidity constraints, and for finite horizons which would dampen the responses to changes in Q. These constraints are important in explaining the downturn in investment immediately after the GFC. But these changes would not help in explaining why, even though equity markets are strong and cash flow is weak, corporate investment in advanced countries is still so low. Here the ideas of Carlin and Soskice seem crucial.

The Calvo-contracts equation for inflation needs to be replaced with one which allows for search and unemployment effects, and inertia (like backward-looking expectations). Relative price adjustment between heterogeneous goods is also important. This is something emphasized by Ghironi, in his model with heterogeneous firms, and McKibbin and Stoeckel, in their examination of the effects on the global structure of relative prices of the rise of emerging-market economies.

Finally, we think that a gap should be introduced between the rate of interest set by monetary policy-makers and the rate affecting the consumption and investment decisions of the private sector. Such a gap may be an important reason why investment in advanced countries has not recovered since 2008. The difficulty here comes from two of the critical assumptions underpinning the benchmark model: the efficient market

39 The importance of the latter can be illustrated by looking at the effects of asymmetric shocks within a monetary union. Kirsanova et al. (2007) and Allsopp and Vines (2008) show that asymmetric negative shocks to a country’s competitiveness in a monetary union might do little damage if the nominal rigidities are of a Calvo kind, because forward-looking wage-and-price-setters will immediately drop prices in the appropriate manner. But, if wages and prices are persistent, this shock can create deflation and a rising real interest rate in that country—because the union-wide nominal interest rate across the monetary union will be unaffected—causing a further reduction of demand and inflation and so perhaps leading to a cumulative deflation.
hypothesis, and rational expectations. The efficient markets hypothesis means that things can never go seriously wrong because of a risk premium, and the rational expectations assumption (and its implication that an economy will eventually re-converge to the Ramsey growth path) means that there can never be a really serious crisis. Inserting a gap between the policy rate and the rate affecting investment decisions may be a way of fixing this.\footnote{A different approach to this issue is put forward by Carlin and Soskice in their paper.}

What we have listed above is already a bit too much like a shopping list, from which it will be necessary to make choices. Furthermore, it is important to add to our shopping list the proposals made in this issue of *OxREP* by Reis and Stiglitz. In a few short pages, Reis gives a very helpful sketch of what could be included in a new core model. And in a much longer contribution Stiglitz produces a framework in which policy choices can be examined; one which can give rise to both Keynesian-type outcomes in the short run, and to growing equilibria over longer periods of time—if shocks are not too large and if the adjustment process in response to these shocks is satisfactory.

We can begin to see—in outline—how many of these changes might be carried out, one by one. But as Keynes learned in the 1930s, markets interact. So one challenge will be to incorporate each of these things in a way that makes sense in general equilibrium. A second challenge will be to do so parsimoniously.

As and when this is done, we think that there will have been a very significant evolution away from the benchmark model that was in place before the global financial crisis. But we do not think of this as a paradigm shift. There will have been many particular changes in content. But perhaps not a sufficient change in method for the outcome to be described as a real change in paradigm.

**(iii) The approach to the data**

The authors in this issue say relatively less about our last examination question, on how models should relate to data. Nevertheless, some important things stand out.

Stiglitz draws on a critique by Korinek (2017) of macro-empirical methods to propose a number of ways in which core and policy models can better relate to data. These include: abandoning the HP filter, which ignores important low-frequency phenomena; allowing for the skewness of time series; finding consensus on measuring the goodness of fit; and ensuring that DSGE models do not impose restrictions that contravene the micro evidence.

If one is trying to estimate a DSGE model, then perhaps the Bayesian method remains the most appropriate one. Roughly speaking, this method, as routinely practised, starts with an analytical structure, imposes that structure on the data through priors, and then invents names for the large errors which may appear in the equations to compensate for the violence that the model does to the data, especially if theory tightly constrains some of the parameters. Such a process is more or less what Romer (2016) ridiculed. An important way to escape the Romer ridicule is to inform these priors by means of detailed work estimating the equations one by one, rather than immediately estimating the system as a whole. Sub-system estimation may also help, to take advantage of connections between related equations while avoiding biases from elsewhere in
the system. But this is difficult to do, partly because the same parameter may appear in more than one equation. Furthermore, as Hendry and Muellbauer argue, if the model excludes highly relevant variables and structural breaks, estimating equations individually is unlikely to help a lot. One must—at the very least—allow for intercept adjustments, and for time-varying parameters.

(iv) A need to change the culture of the macroeconomics profession

A macroeconomics profession which begins to do what we have described in this section of the paper will come to have a rather different culture.

All of the authors in this issue seem to agree with Simon Wren-Lewis that the culture which has emerged in macroeconomics is too hegemonic. Haldane and Turrell use the term ‘monoculture’ to describe what has happened: namely the view that only models with rational expectations and explicit microfoundations of a particular kind could be accepted into the macroeconomic canon. The emergence of a range of different models will certainly serve to undermine this hegemony. A range of core models is likely to emerge. And the recognition of the way in which a policy model has a different purpose from that of a core model is likely to give rise to a situation in which policy models are very different from core models, and—maybe—to be accompanied by a change in the culture to one in which this difference is both tolerated and respected.

Many argue that the discipline has also become too insular. Ghironi argues that macroeconomics needs to overcome the separation between the study of short-run stabilization policy and longer-run growth, something which we have already discussed in this article. Blanchard argues that DSGE modellers might become less insular by looking at other parts of the economics discipline. Ghironi elaborates on this point: he talks about the need to bridge the gap between closed economy macroeconomics and the study of international macroeconomics and international trade. McKibbin and Stoeckel agree with this. In a different vein, Haldane and Turrell note that those who write macroeconomic articles cite other disciplines much less than happens in other fields.

We think that, in a healthier culture, macroeconomists should be allowed to learn more from data, in the way that microeconomists seem to do. As Wren-Lewis argues, there has been an unhelpful movement in two opposite directions. On the one hand, the Lucas critique and the hijacking of the use of the word ‘structural’ to mean ‘microfounded’ has pushed the applied macroeconomists who want to build policy models into a requirement that they deal with data in a very constrained way. On the other hand, there has been a move by those who deal seriously with data into the estimation of VARs, which involve almost no theory. The failure of the twain to meet has—he argues—severely constrained evidence-based progress in macro. We agree with this claim.

How might we move in the required direction? What might pluralist progress involve? A number of authors argue that models should be modular, like Lego, Meccano (Blanchard), or building blocks (Ghironi). The aim might be a core model that can accommodate ‘bolted-on’ extensions. Policy models should focus on issues that are

41 There are a number of other difficulties; some solutions to the problems are offered by Fukac and Pagan (2010).
regularly encountered, while a collection of complementary satellite models might cover less central issues, or ones with less of a general-equilibrium framework. Blanchard argues that we should ‘relegalize shortcuts’, noting that one person’s microfoundations will look like short-cuts to someone else.

Lindé and Stiglitz argue that big and small models should be complementary. Small models might help us to understand new mechanisms, before being incorporated into larger models to test their robustness in general equilibrium. McKibbin and Stoeckel argue that large models—nevertheless ones tightly constrained by theory—can give much real-world insight.

Within each class of model, it may also be that there should be more diversity. Krugman wants a ‘looser-jointed’ benchmark model, one which can be useful without being microfounded. He also argues—in a way very different from what many others are saying—that financial models can be put on the side, only to be used to examine outcomes at times of crisis, rather than being embedded in any core model.

Haldane and Turrell want a ‘rich ecology’ of models, to which selective pressure can be applied by controlled experiments, to see which best fit the facts. They argue that such an approach is used in other disciplines, and that it is also used at the Bank of England, their own institution. Stiglitz ends up arguing that we should teach many kinds of models to graduate students. Nevertheless, he describes a parsimonious three-period model which can be used in a productive way to hold the various ideas together.

Such an approach might be like a collection of maps. A London Tube map is extremely reductive, famously overstating the distance between Paddington and Lancaster Gate which can be walked in 10 minutes. However, for the purpose of navigating the Underground this model is elegantly suited. While a tourist on foot might want a map of famous landmarks, a town planner would need something else. The maps we use to navigate the economy should—we think—be similarly varied.

VI. Conclusion

In his after-dinner talk to the annual NBER workshop in 2017, Olivier Blanchard briefly described what frictions should be included in a new core model. But then, in a disarming aside, he said: ‘[t]his may a hopeless and misguided search. Maybe even the simplest characterization of fluctuations requires many more distortions. Maybe different distortions are important at different times. Maybe there is no simple model . . .’. Nevertheless he then added: ‘I keep faith that there is.’ That is our hope, too.

As support for this journey-in-hope, we think that we have learned two things from this project. The first lesson is that it is time to do away with the microfoundations hegemony. Long ago the University of Cambridge (UK) established DAMTP—the Department of Applied Mathematics and Theoretical Physics. But this did not impede progress in experimental physics at that university. It is time for our subject to allow more room for, and show more respect for, those engaged in building, and using, policy models. These macroeconomists are now doing our equivalent of experimental physics.

The second—and related—lesson is that there needs to be more pluralism. Just like in the sixteenth century, after the Christian Reformation, there may no longer be a true church. It is time to put the religious wars behind us.
Appendix I: Rebuilding macroeconomic theory: the examination questions

We asked contributors to answer all, or some, of the following six examination questions. There are two questions on each of context, content, and method. These were designed to focus on how everything fits together, as distinct from focusing on particular details. All of the questions related to the benchmark New Keynesian DSGE model which was set out for the authors in the manner which we have set it out in section III of this paper.

1. Is the benchmark DSGE model fit for purpose?

What is the purpose of macroeconomics at present? To describe how economies work? To diagnose policy failures? To guide future policy? Does the benchmark model still do these things?

2. What additions to this model are necessary to help us understand growth from now on?

Are slow population growth and a low rate of technical progress endangering the long-run growth process (as envisaged by Summers and Krugman)? If so, how do we best model this? In what way—if at all—is technical progress currently endangered (as imagined by Gordon)? Or is growth mainly driven by the decision to save (as imagined by some interpretations of Piketty)? What is the role of physical infrastructure, public capital, and human capital in the growth process and how is the necessary investment and education to be provided? What are the implications of this growth process for income inequality? How best to think theoretically about the role of financial intermediation in the growth process? What is the relative importance of equity and loan finance? To what extent are leveraged financial institutions necessary for financial intermediation, or might peer-to-peer lending come to supersede these institutions? When will deleveraging stop constraining the growth process, as it has for the last 6 years? To what extent are public deficits and debt necessary in the stimulus of demand and in the provision of infrastructure, public capital, and human capital? Does answering these questions require an approach different from the benchmark model, or simply a modification of that model?

3. What are the important inter-temporal complications?

The benchmark DSGE model assumes perfect foresight, rational expectations, and well-functioning capital markets; policy is assumed to be made either with full credibility, or under discretion; most models have unique solutions across time. Clearly actual history is not like this. In reality information is limited, balance sheet constraints exist, and there are feedbacks from stocks to flows (including from the stock of money to the flow of expenditure). Agents are subject to habits and norms. As a result, there are likely to be multiple equilibria, both real and nominal. What is the role of policy...
in guiding the economy to one particular equilibrium? Most fundamentally, what are
causal connections—both ways—between long-run growth and short-run stabilization
policy? Does answering these questions require an approach different from the bench-
mark model, or simply a modification of that model?

4. **What are the important intra-temporal complications?**

Why does aggregate demand not always adjust to aggregate supply, including after
shocks? How do the resulting fluctuations in capacity utilization affect the growth pro-
cess? The benchmark DSGE model recognizes one reason for this—nominal rigidities.
But should we go beyond this to think about coordination failures? What important
coordination failures result from interactions between agents that are not captured
in such a representative-agent framework (e.g. in the Dixit–Stiglitz model)? In other
words, when do multiple equilibria matter? Does answering these questions require
an approach different from the benchmark model, or simply a modification of that
model?

5. **Should the above questions be discussed with big or small
models?**

Should we, like physicists—or like James Meade and some contemporary DSGE theo-
rists—aim for a general theory of everything? Or should we instead push for a set of
simple, tractable models, each of which just tells part of the story (the MIT method)? In
other words, should we ‘relegalize shortcuts’, as suggested by Blanchard? Has a desire
for completeness meant that we cannot see the wood for the trees? Conversely, has a
desire for elegance caused us to omit important details?

6. **How should our models relate to data?**

Do our models use data appropriately? In particular, should the structures of estimated
models be data-determined (as in VARS) or theory-constrained (as in Bayesian model
building). Is there a mid-way between these extremes?

**Appendix II: The New Keynesian Benchmark Model**

We first set out the components of the real model that underlies the New-Keynesian
benchmark model. For brevity and simplicity, we do not set out the utility-maximization
problem of the representative consumer or the profit-maximization problem of
the representative firm; we simply set out the first-order conditions, and equilibrium
conditions, derived from these two sets of optimization decisions. And, also for sim-
plicity, we assume a constant labour force and an unchanging level of technology; the

---

42 For relevant details see Schmitt-Grohé and Uribe (2006) and Woodford (2003, ch. 5, § 3).
model can be generalized in a straightforward way to incorporate labour force growth and exogenous technical progress.

Notation is as follows: \( C, I, Y, L, K, w, R, \) and \( Q \) represent (respectively) consumption, investment, output, labour supply, the capital stock, the real wage, the (gross) real interest rate, and Tobin’s Q. The model consists of the following eight equations.

\[
\frac{1}{C_t} = \beta R_t E^{t-1}
\]

\[
w_t = \mu C_t L_t^e
\]

\[
Y_t = A_t K_t^\alpha L_t^{1-\alpha}
\]

\[
w_t = (1-\alpha) \frac{Y_t}{L_t}
\]

\[
R_t = E_{t+1}
\]

\[
I_t = K_{t+1} - (1-\delta) K_t + \mu \left( \frac{K_{t+1}}{K_t} - 1 \right)
\]

\[
Q_t = 1 - \xi + \frac{\xi}{2} \frac{K_{t+1}}{K_t}
\]

\[
Y_t = C_t + I_t
\]

Equation (1) is the inter-temporal Euler equation for the representative consumer; equation (2) is the intra-temporal labour-supply equation of the representative consumer, which equates the (real) wage to the marginal disutility of labour.\(^{43}\) Equation (3) shows aggregate supply\(^{44}\) and equation (4) shows that the representative firm employs labour up to the point where the marginal product of labour is equal to the wage. Equation (5) shows that the representative firm carries out investment up to the point at which the marginal product of capital is equal to the real interest rate, plus an allowance for the depreciation of the capital stock, minus any anticipated capital gains on the capital stock, plus an allowance for the marginal costs of capital adjustment.\(^{45}\) Equation (6) shows that capital accumulation is equal to investment, minus depreciation, minus

\(^{43}\) The consumer’s utility function is logarithmic in consumption and decreasing and convex in labour.

\(^{44}\) The production function of the representative firm is Cobb–Douglas.

\(^{45}\) Because of the interaction of Equations (5)–(7), this term has an effect which is increasing in \( \xi \) when the model is simulated, even although \( \xi \) appears in the denominator of Equation (5).
the resources wasted whenever the capital stock is adjusted; adjustment costs are convex and quadratic—the bigger is $\xi$ the greater are these costs. Equation (7) determines Tobin’s $Q$; the equation shows that the larger are adjustment costs of capital (i.e. the larger is $\xi$) the further will $Q$ deviate from unity when the capital stock is away from its equilibrium. Equations (5)–(7), when taken together mean that, when the model is simulated, the larger is $\xi$, the more gradual will be any adjustment of the capital stock to its desired level. Equation (8) shows that aggregate demand is always equal to aggregate supply; the model solves for the real interest rate $R$ which brings this about—effectively by making saving equal to investment. The model is log-linearized around the non-stochastic steady state and solved using the Blanchard–Kahn method; $K$ is a predetermined variable and $C$ and $Q$ are jump variables.

Notice from Equations (5)–(7) that whenever there is a shock to the system, and providing that $\xi > 0$, the real interest rate must move away from the marginal product of capital plus an allowance for depreciation, in order to re-establish an equilibrium in which saving equals investment. This is precisely because there are costs in adjusting the capital stock.

In this model, although all markets clear, there are general-equilibrium interactions between markets. Keynes found that such interactions happened between markets when he postulated sticky wages—implying that the labour market might not clear—and so effectively turned Marshallian analysis into IS–LM. Here the general-equilibrium story is an additional one: the microfoundations of this model mean that interactions between markets happen, even if all markets clear. In particular, the demand for labour depends on consumption and thus on the interest rate and investment. Any shock, such as a fall in investment as analysed below, will therefore shift the demand for labour inwards at any level of the real wage. It will also lead to a lower interest rate, higher consumption and so shift in the supply curve of labour inwards at any level of the real wage. The outcome in the labour market, and in particular the equilibrium real wage, will depend on the shifts in labour demand and supply, which are caused by developments in the product market.

Notice that, in the (implausible) limiting case, in which $\xi \to 0$, the model converges to an RBC model, in which whatever is saved is costlessly invested. In such a set-up, Tobin’s $Q$ always remains equal to unity, and the real interest rate is always just equal to the marginal product of capital plus an allowance for depreciation. There is no need for an endogenous interest rate to make savings equal to investment.46 This chain of reasoning shows why it is essential to include capital adjustment costs in the model.

The parameters used in the simulation below are standard and are shown in Table 1; these parameters correspond to the idea of a quarterly model.

The effects of a sustained negative shock to the aggregate level of productivity, $A$, are illustrated in Figure 1. RBC theorists normally study shocks which are autoregressive, for reasons explained in the text, but it is more revealing to study the effects of a sustained technology shock.47

46 Nevertheless, the shock analysed in the example below would still have general equilibrium effects on the labour market: the reduction in productivity would reduce both labour demand and—by reducing the real interest rate—increase consumption and thereby reduce labour supply.

47 Control engineers teach their students that subjecting a model to sustained shocks reveals things about the structure of the model which are not revealed by transient shocks.
After the shock, profit-maximizing firms will want to hold a lower capital stock, and so the new Ramsey equilibrium must have this feature. That leads to a period of disinvestment; the length and intensity of which depend on the costs of capital adjustment. As a result of this disinvestment, the level of aggregate supply will fall more in the long run than in the short run. The sustainable level of current consumption will fall, and, in the interests of consumption smoothing, consumers will cut their level of consumption immediately. As a result, although aggregate supply is reduced, aggregate demand will also fall even if the real interest rate does not change. If adjustment costs in investment are sufficiently small—as assumed here—then the fall in demand in will
be larger than the fall in supply. This means that the rate of interest will need to fall to ensure that savings remains equal to investment; and that aggregate demand remains exactly equal to the lower level of aggregate supply. In this Ramsey model without financial frictions, the interest rate received by consumers is equal to the cost of capital to investors, after allowing for the costs of adjusting the capital stock.

We can imagine that a perfectly capable central bank sets the interest rate necessary to keep aggregate demand equal to aggregate supply. It does this by manipulating consumption to make savings fall in a way which is consistent with the desired (gradual) decumulation of capital.

We now add nominal rigidities to the model, and a Taylor rule, and so arrive at the full new-Keynesian benchmark model, which we can use to study the behaviour of inflation. We use this model to simulate a cost-push shock.

We proceed by introducing differentiated goods and nominal rigidities in the manner of Calvo (1983). Specifically, we assume a continuum of goods producing firms, \( i \), where \( i \) lies in the range \([0,1]\); these firms operate under monopolistic competition. These goods are aggregated into a single ‘final’ good by means of a Dixit–Stiglitz aggregator in which the elasticity of substitution between the varieties of goods is equal to \( \varepsilon \). Any firm re-optimizes its price in a period with fixed probability \((1 - \kappa)\). (This is the Calvo fairy at work.) With probability \( \kappa \) prices are not re-optimized and are assumed to rise at the average rate of inflation. Finally, we assume—as is standard in the literature—that there is a subsidy to the firms (paid by a lump-sum tax) so that in the steady state firms produce the same quantities as in the flex-price economy, even although each firm faces a downward-sloping demand curve and so can restrict output. This means that the steady state values for \( C, I, Y, L, K, w, R \), and \( Q \) are the same in the full new-Keynesian benchmark model as they were in the real model presented above.

The three new variables in this full model are \( \pi \), the rate of inflation, \( R^n \), the nominal (gross) interest rate, and \( mc \), which represents real marginal cost.

Of course, the model no longer solves for a real interest, \( R \), that would make savings equal to investment in a way which would ensure the absence of inflationary pressure. Instead it allows aggregate demand to move relative to aggregate supply, in a way which creates inflationary pressure. It then allows the central bank to determine the real interest rate by setting the nominal interest rate, in response to developments in the inflation rate, according to a Taylor rule. Equation (9) expresses the real interest rate in terms of the nominal interest rate; equation (10) shows the way in which the central bank sets the nominal interest rate according to a Taylor rule, one in which—for simplicity—there is no output term.

\[
R_t = E_t \left( \frac{R^n_t}{\pi_{t+1}} \right)
\]  

(9)

48 We regard this as a plausible restriction on the parameterization of the model, although opinions on this differ.

49 For brevity and simplicity, we do not set out the profit-maximization problem underlying the price setting behaviour of the representative firm; we simply set out the first-order conditions. For details see Schmitt-Grohé and Uribe (2006) and Woodford (2003).
\[ R_t^n = \beta^{-1} \left( \frac{\pi_t}{\pi} \right)^\delta \]  

(10)

The term \( \beta^{-1} \) at the front of the Taylor rule determines the constant term when the rule is written in linear form; it shows that the real interest rate will converge to its equilibrium value \( \beta^{-1} \) when any disturbances to inflation have disappeared. That is necessary in order for consumption to be in equilibrium—see equation (1) above.

The first eight equations of this new system are identical to those in the real system presented above, except for two crucial alterations. We replace equation (4) with the following equation:

\[ w_t = mc_t (1 - \alpha) \frac{Y_t}{L_t}. \]  

(4a)

This equation solves for levels of real marginal cost, \( mc \), different from unity, thereby enabling aggregate demand to move in a way which is not constrained by aggregate supply. We know, from equation (2), that the economy always lies on the labour supply curve. We also know, from equation (8), that output is demand determined. Consider the effects of an increase in demand, caused by, say, a demand shock. This will cause output to rise, which from the production function—equation (3)—will lead to higher employment. The real wage must rise as the economy moves up the labour supply curve (equation 2), ceteris paribus. But, as more labour is employed, the marginal product of labour must fall (because of diminishing returns). Thus, because the wage is rising (as the economy moves up the labour supply curve), but the marginal product of labour is falling (because of diminishing returns), real marginal cost (which equals the wage divided by the marginal product of labour) must be rising. That is what equation (4a) shows.

We replace equation (5) with the following equation:

\[ R_t = E_t \left[ \frac{\alpha mc_{t+1} Y_{t+1}}{K_{t+1}} + \left( \frac{Q_{t+1} - 1}{2\xi} \right)^2 \right]. \]  

(5a)

The reasoning underlying this equation is similar to that applying to equation (4a), but now the direction of causation is different. Consider again the effects of an increase in demand, caused by a demand shock. This will cause output to rise and so lead to higher employment and a higher wage. But higher output will also raise the marginal product of capital and so encourage investment. But the extent to which investment rises will be constrained by the wedge between the wage and the marginal product of labour, which is what is captured by the real marginal cost variable, \( mc \). Equation (5a) shows the effect of this wedge on \( Q \) and thus on investment.

Finally, equation (11) shows the new-Keynesian Phillips curve, in which inflation in any period depends on the (discounted) expected future rate of inflation plus a term which depends on the logarithm of the real marginal cost of labour to firms50.

---

50 The logarithm of the real marginal cost of labour shows the proportional deviation of real marginal cost from the value of unity which it would have at the intersection of the labour supply and labour demand.
\[ \pi_t = \beta E_t \pi_{t+1} + \frac{(1-\beta)(1-\kappa \beta)}{\kappa} \ln(mc_t) + \mu. \] (11)

This equation effectively shows, for example, that whenever aggregate demand rises above aggregate supply, and so \( mc \) rises above unity, any firm visited by the Calvo fairy will raise its price level. It will do this to an extent which depends on how many other firms are being visited at the same time—i.e. on the size of \( \kappa \)—and also on how high \( \pi \) is expected to be in the next period. The model is solved using the Blanchard–Kahn method; the inflation rate, \( \pi \), is an additional jump variable.\(^{51}\)

We use the following calibration for the two new parameters that we introduce: \( \phi_\pi = 1.5, \kappa = 0.75. \) The first of these corresponds to the parameter used by Taylor when he first introduced his rule; this value ensures that whenever the inflation rate changes the real interest rate moves in the same direction, satisfying the Taylor principle discussed in the text. A value of \( \kappa \) of 0.75 ensures that the Calvo fairy arrives at each firm, on average, once a year. We do not need to calibrate \( \epsilon \). This is because, although the size of \( \varepsilon \) influences the size of mark-up of prices over marginal cost, this mark-up is constant and so disappears when the system is log-linearized.

The term \( \mu \) shows a cost-push shock. For simplicity we specify this as a reduced form shock to the Phillips curve. There are various ways which have been used in the literature to microfound this shock.

We deliberately do not display the results of a technology shock in this model with nominal rigidities, as there are only minor differences to Figure 1. The reason is because the monetary authorities follow the Taylor rule above. When there is a negative productivity shock, inflation will need to fall to induce the central bank to lower the nominal (and thus the real) interest rate. This is needed to prevent demand falling by more than supply. The larger the parameter \( \phi_\pi \) in the Taylor rule, the less inflation needs to fall to achieve this. In the limit, as \( \phi_\pi \) tends to infinity, the response to a technology shock of a model with nominal rigidities becomes identical to that of the real model shown above. Even with the parameters chosen here, the differences between the simulation of the model with nominal rigidities and the results shown in Figure 1 are so small that it is not worth repeating the exercise.

This model with nominal rigidities enables us to display the results of a cost-push shock. Figure 2 shows the responses of the model to a 1 per cent shock to the multiplicative cost-push-shock term, \( \mu \). We assume the shock follows an AR(1) process with persistence 0.8, as is common in the literature; this prevents the results shown in the pictures having a spike, as the result of a spike in inflation.

51 Note that, as \( \kappa \) tends towards unity, the setting of prices tends towards full flexibility. If—in the case of a value of \( \kappa \) close to unity—demand pressures were to make real marginal cost differ from unity, then that would cause a very large disturbance to inflation. In an economy with a nominal anchor provided by a Taylor rule, such as equation (11), the central bank would raise the real interest rate, and that would, of course, moderate the demand pressures. In such an economy, as \( \kappa \) tended towards unity, the economy would tend towards one in which there was perfect price flexibility, output always equalled its natural level, and inflation moved instantaneously to a level at which the real interest was such as to bring about exactly that level of output.
After the shock, all firms are assumed to raise their prices as a result of the shock. The shock to inflation induces an immediate rise in the nominal (and real) interest rate. Consumption falls, and so does investment, because Tobin’s Q falls. Output, employment, and the real wage fall, because workers are always on their labour supply curve. After the shock, all firms are assumed to raise their prices as a result of the shock. The shock to inflation induces an immediate rise in the nominal (and real) interest rate. Consumption falls, and so does investment, because Tobin’s Q falls. Output, employment, and the real wage fall, because workers are always on their labour supply curve. Inflation comes down only gradually, partly because the cost-push shock is autoregressive, but also because each firm moderates their price cuts because they know that some other firms are not yet reducing their own prices.

This model shows that the capital stock is depressed as a result of the disinvestment which happens during the adjustment process, but is gradually rebuilt to its initial level.

Figure 2: Response to a 1% positive cost-push shock

A choice by consumers to lower consumption in response to higher interest rates will be accompanied by a desire for less leisure and so push the labour supply curve to the right, thereby magnifying the fall in the real wage.
This is obviously something which cannot be studied in the Clarida, Gali, and Gertler model, in which all aggregate demand takes the form of consumption. Luk and Vines (2015) show that the method of controlling inflation by monetary policy, which is examined here, rather than partly controlling inflation by fiscal policy, can be costly precisely because it causes disinvestment during the adjustment process, something which needs to be reversed in a way which is costly. But they also find—somewhat surprisingly—that the welfare effects of doing this are small.

References


Gordon, R. (1975), Milton Friedman’s Monetary Framework: A Debate with His Critics, Chicago, IL, University of Chicago Press.
Hayek, F. (1931), Prices and Production, New York, Kelly.
— (1944), The Road to Serfdom, Chicago, IL, University of Chicago Press.


On the future of macroeconomic models

Olivier Blanchard*

Abstract: Macroeconomics has been under scrutiny as a field since the financial crisis, which brought an abrupt end to the optimism of the Great Moderation. There is widespread acknowledgement that the prevailing dynamic stochastic general equilibrium (DSGE) models performed poorly, but little agreement on what alternative future paradigm should be pursued. This article is the elaboration of four blog posts that together present a clear message: current DSGE models are flawed, but they contain the right foundations and must be improved rather than discarded. Further, we need different types of macroeconomic models for different purposes. Specifically, there should be five kinds of general equilibrium models: a common core, plus foundational theory, policy, toy, and forecasting models. The different classes of models have a lot to learn from each other, but the goal of full integration has proven counterproductive. No model can be all things to all people.

Keywords: macroeconomics, methodology, macroeconomic model, DSGE, dynamic stochastic general equilibrium, forecasting, macroeconomic policy making, graduate teaching

JEL classification: B41, E000, E120, E130, E170, E610, A230

I. Introduction

One of the best pieces of advice Rudi Dornbusch gave me was: never talk about methodology; just do it. Yet, I shall make him unhappy, and take the plunge.

This set of remarks on the future of macroeconomic models was triggered by a project, led by David Vines, to assess how dynamic stochastic general equilibrium (DSGE) models had performed during the financial crisis (namely, badly) and to examine how these models might be improved. Needled by David’s opinions, I wrote a PIIE Policy Brief (Blanchard, 2016a). That brief piece went nearly viral (by the standards of blogs on DSGEs). The many comments I received led me to write a second piece, a PIIE Real Time blog (Blanchard, 2016b), which again led to a further round of reactions, prompting me to write a third blog (Blanchard (2017a), which I hoped and fully expected would be my last piece on this topic). I thought I was done, but last February David organized a one-day conference on the topic, from which I learned a lot, which led me to write a fourth blog (Blanchard, 2017b) and now, with this article, my final piece (I promise) on this topic. In doing so I have brought all my previous contributions together into this one longer piece.

*Massachusetts Institute of Technology, and NBER, e-mail: blanchar@mit.edu

I thank Annie Williamson, who has nicely patched the four underlying blog posts together.

doi:10.1093/oxrep/grx045
© The Author 2018. Published by Oxford University Press.
For permissions please e-mail: journals.permissions@oup.com
A common theme runs through all four pieces: we need different types of macro models. One type is not better than the others. They are all needed, and indeed they should all interact. In the final section of my piece I present my attempt at typology, distinguishing between five types.

My remarks about needing many types of models would be trivial and superfluous if that proposition were widely accepted, and there were no wars of religion. But it is not, and there are. As a result, my contribution seems useful. To show how, and why, I have ended up where I am, I have kept the pieces in the order in which I wrote them, rather than reorganize them in a new, more integrated piece.

I limit myself to general equilibrium models. Much of macro must, however, be about building the individual pieces, constructing partial equilibrium models, and examining the corresponding empirical evidence. These are the components on which the general equilibrium models must then build.

II. The challenge to DSGE models

My first piece looked at the case for or against DSGE models. DSGE models have come to play a dominant role in macroeconomic research. Some see them as the sign that macroeconomics has become a mature science, organized around a microfounded common core. Others see them as a dangerous dead end.

I believe the first claim is exaggerated and the second is wrong. I see the current DSGE models as seriously flawed, but they are eminently improvable and central to the future of macroeconomics. To improve, however, they have to become less insular, by drawing on a much broader body of economic research. They also have to become less imperialistic and accept sharing the scene with other approaches to modelization.

For those who are not macroeconomists, or for those macroeconomists who have lived on a desert island for the last 20 years, here is a brief refresher. DSGE stands for ‘dynamic stochastic general equilibrium’. The models are indeed dynamic, stochastic, and characterize the general equilibrium of the economy. They make three strategic modelling choices. First, the behaviour of consumers, firms, and financial intermediaries, when present, is formally derived from microfoundations. Second, the underlying economic environment is that of a competitive economy, but with a number of essential distortions added, from nominal rigidities to monopoly power to information problems. Third, the model is estimated as a system, rather than equation by equation as in the previous generations of macroeconomic models. The earliest DSGE model, representing a world without distortions, was the Real Business Cycle model developed by Edward C. Prescott (1986) and focused on the effects of productivity shocks. In later incarnations, a wider set of distortions, and a wider set of shocks, have come to play a larger role, and current DSGE models are best seen as large-scale versions of the New Keynesian model, which emphasizes nominal rigidities and a role for aggregate demand.¹

¹ While a ‘standard DSGE model’ does not exist, a standard reference remains the model developed by Frank Smets and Rafael Wouters (2007). See Lindé et al. (2016) for a recent assessment with many references.
There are many reasons to dislike current DSGE models. First, they are based on unappealing assumptions. Not just simplifying assumptions, as any model must be, but assumptions profoundly at odds with what we know about consumers and firms.

Go back to the benchmark New Keynesian model, from which DSGEs derive their bone structure. The model is composed of three equations: an equation describing aggregate demand; an equation describing price adjustment; and an equation describing the monetary policy rule. At least the first two are badly flawed descriptions of reality:

Aggregate demand is derived as consumption demand by infinitely lived and foresighted consumers. Its implications, with respect to both the degree of foresight and the role of interest rates in twisting the path of consumption, are strongly at odds with the empirical evidence. Price adjustment is characterized by a forward-looking inflation equation, which does not capture the fundamental inertia of inflation.2

Current DSGE models extend the New Keynesian model in many ways, allowing for investment and capital accumulation, financial intermediation, interactions with other countries, and so on. The aggregate demand and price adjustment equations remain central, however, although they are modified to better fit the data. This occurs in the first case by allowing, for example, a proportion of consumers to be ‘hand to mouth’ consumers, who simply consume their income. In the second case, it occurs by introducing backward-looking price indexation, which, nearly by assumption, generates inflation inertia. Both, however, are repairs rather than convincing characterizations of the behaviour of consumers or of the behaviour of price- and wage-setters.

Second, their standard method of estimation, which is a mix of calibration and Bayesian estimation, is unconvincing.

The models are estimated as a system, rather than equation by equation as in previous macroeconometric models. They come, however, with a very large number of parameters to estimate, so that classical estimation of the full set is unfeasible. Thus, a number of parameters are set a priori, through ‘calibration’. This approach would be reasonable if these parameters were well-established empirically or theoretically. For example, under the assumption that the production function is Cobb–Douglas, using the share of labour as the exponent on labour in the production function may be reasonable. But the list of parameters chosen through calibration is typically much larger, and the evidence often much fuzzier. For example, in the face of substantial differences in the behaviour of inflation across countries, use of the same ‘standard Calvo parameters’ (the parameters determining the effect of unemployment on inflation) in different countries is highly suspicious. In many cases, the choice to rely on a ‘standard set of parameters’ is simply a way of shifting blame for the choice of parameters to previous researchers.

The remaining parameters are estimated through Bayesian estimation of the full model. The problems are twofold. One is standard in any system estimation.

---

2 More specifically, the equation characterizing the behaviour of consumers is the first order condition of the corresponding optimization problem and is known as the ‘Euler equation’. The equation characterizing the behaviour of prices is derived from a formalization offered by Guillermo Calvo and is thus known as ‘Calvo pricing’.
Misspecification of part of the model affects estimation of the parameters in other parts of the model. For example, misspecification of aggregate demand may lead to incorrect estimates of price and wage adjustment, and so on. And it does so in ways that are opaque to the reader. The other problem comes from the complexity of mapping from parameters to data. Classical estimation is *de facto* unfeasible, the likelihood function being too flat among many dimensions. Bayesian estimation would indeed seem to be the way to proceed, if indeed we had justifiably tight priors for the coefficients. But, in many cases, the justification for the tight prior is weak at best, and what is estimated reflects more the prior of the researcher than the likelihood function.\(^3\)

Third, while the models can formally be used for normative purposes, normative implications are not convincing.

A major potential strength of DSGE models is that, to the extent that they are derived from microfoundations, they can be used not only for descriptive but also for normative purposes. Indeed, the single focus on GDP or GDP growth in many policy discussions is misleading: distribution effects, or distortions that affect the composition rather than the size of output, or effects of current policies on future rather than current output, may be as important for welfare as effects on current GDP. Witness the importance of discussions about increasing inequality in the United States, or about the composition of output between investment and consumption in China.

The problem in practice is that the derivation of welfare effects depends on the way distortions are introduced in the model. And often, for reasons of practicality, these distortions are introduced in ways that are analytically convenient but have unconvincing welfare implications. To take a concrete example, the adverse effects of inflation on welfare in these models depend mostly on their effects on the distribution of relative prices, as not all firms adjust nominal prices at the same time. Research on the benefits and costs of inflation suggests, however, a much wider array of effects of inflation on activity and in turn on welfare.

Having looked in a recent paper (Blanchard et al., 2016) at welfare implications of various policies through both an *ad hoc* welfare function reflecting deviations of output from potential and inflation from target and the welfare function implied by the model, I drew two conclusions. First, the exercise of deriving the internally consistent welfare function was useful in showing potential welfare effects I had not thought about, but concluded *ex post* were probably relevant. Second, between the two, I still had more confidence in the conclusions of the *ad hoc* welfare function.

Fourth, DSGE models are bad communication devices.

A typical DSGE paper adds a particular distortion to an existing core. It starts with an algebra-heavy derivation of the model, then goes through estimation, and ends with various dynamic simulations showing the effects of the distortion on the general equilibrium properties of the model.

These would indeed seem to be the characteristics of a mature science: building on a well-understood, agreed-upon body of science and exploring modifications and

\(^3\) In some cases, maximum likelihood estimates of the parameters are well identified but highly implausible on theoretical grounds. In this case, tight Bayesian priors lead to more plausible estimates. It is clear, however, that the problem in this case comes from an incorrect specification of the model and that tight Bayesian priors are again a repair rather than a solution.
extensions. And, indeed, having a common core enriches the discussion among those who actually produce these models and have acquired, through many simulations, some sense of their entrails (leaving aside whether the common core is the right one, the issue raised in the first criticism above). But, for the more casual reader, it is often extremely hard to understand what a particular distortion does on its own and then how it interacts with other distortions in the model.

All these objections are serious. Do they add up to a case for discarding DSGEs and exploring other approaches? I do not think so. I believe the DSGEs make the right basic strategic choices and the current flaws can be addressed.

Let me explain these two claims. The pursuit of a widely accepted analytical macroeconomic core, in which to locate discussions and extensions, may be a pipe dream, but it is a dream surely worth pursuing. If so, the three main modelling choices of DSGEs are the right ones. Starting from explicit microfoundations is clearly essential; where else to start from? Ad hoc equations will not do for that purpose. Thinking in terms of a set of distortions to a competitive economy implies a long slog from the competitive model to a reasonably plausible description of the economy. But, again, it is hard to see where else to start from. Turning to estimation, calibrating/estimating the model as a system rather than equation-by-equation also seems essential. Experience from past equation-by-equation models has shown that their dynamic properties can be very much at odds with the actual dynamics of the system.

III. My initial response to this challenge

I concluded my first post with the following response to the challenge I had set myself. I said, at the time, that DSGE modelling needed to evolve in two ways.

First, I argued, it has to become less insular.

Take the consumption example discussed earlier. Rather than looking for repairs, DSGE models should build on the large amount of work on consumer behaviour going on in the various fields of economics, from behavioural economics, to big data empirical work, to macro partial equilibrium estimation. This work is on-going and should indeed proceed on its own, without worrying about DSGE integration. (Note to journal editors: not every discussion of a new mechanism should be required to come with a complete general equilibrium closure.) But this body of work should then be built on to give us a better model of consumer behaviour, a sense of its partial equilibrium implications, perhaps a sense of the general equilibrium implications with a simplistic general equilibrium closure, and then and only then be integrated into DSGE models. This would lead to more plausible specifications and more reliable Bayesian priors, and this is what I see as mostly missing. I have focused here on consumption, but the same applies to price- and wage-setting, investment, financial intermediation, treatment of expectations, etc. In short, DSGEs should be the architecture in which the relevant findings from the various fields of economics are eventually integrated and discussed. It is not the case today.

Second, it has to become less imperialistic. Or, perhaps more fairly, the profession (and again, this is a note to the editors of the major journals) must realize that different model types are needed for different tasks.
Models can have different degrees of theoretical purity. At one end, maximum theoretical purity is indeed the niche of DSGEs. For those models, fitting the data closely is less important than clarity of structure. Next come models used for policy purposes, for example, models by central banks or international organizations. Those must fit the data more closely, and this is likely to require in particular more flexible, less microfounded, lag structures (an example of such a model is the FRB/US model used by the Federal Reserve, which starts from microfoundations but allows the data to determine the dynamic structure of the various relations). Finally, come the models used for forecasting. It may well be that, for these purposes, reduced form models will continue to beat structural models for some time; theoretical purity may be for the moment more of a hindrance than a strength.

Models can also differ in their degree of simplicity. Not all models have to be explicitly microfounded. While this will sound like a *plaidoyer pro domo*, I strongly believe that *ad hoc* macro models, from various versions of the IS–LM to the Mundell–Fleming model, have an important role to play in relation to DSGE models. They can be useful upstream, before DSGE modelling, as a first cut to think about the effects of a particular distortion or a particular policy. They can be useful downstream, after DSGE modelling, to present the major insight of the model in a lighter and pedagogical fashion. Here again, there is room for a variety of models, depending on the degree of *ad hoc*ery: one can think, for example, of the New Keynesian model as a hybrid, a microfounded but much simplified version of larger DSGEs.

Somebody has said that such *ad hoc* models are more art than science, and I think this is right. In the right hands, they are beautiful art, but not all economists can or should be artists. There is room for both science and art. I have found, for example, that I could often, as a discussant, summarize the findings of a DSGE paper in a simple graph. I had learned something from the formal model, but I was able (and allowed as the discussant) to present the basic insight more simply than the author of the paper. The DSGE and the *ad hoc* models were complements, not substitutes. So, to return to the initial question: I suspect that even DSGE modellers will agree that current DSGE models are flawed. But DSGE models can fulfil an important need in macroeconomics; that of offering a core structure around which to build and organize discussions. To do that, however, they have to build more on the rest of macroeconomics and agree to share the scene with other types of general equilibrium models.

IV. What we agree on and what we do not

A number of economists joined the debate about the pros and cons of dynamic DSGEs, partly in response to my blog post. Among them were Narayana Kocherlakota (2016), Simon Wren-Lewis (2016), Paul Romer (2016), Steve Keen (2016), Anton Korinek (2015), Paul Krugman (2016), Noah Smith (2016), Roger Farmer (2014), and Brad Delong (2016). Here are my reactions to the debate, which I posted in my second blog-post, partly repeating what I had already said.
(i) **Three widely believed propositions**

I believe that there is wide agreement on the following three propositions; let us not discuss them further, and move on.

(i) Macroeconomics is about general equilibrium.

(ii) Different types of general equilibrium models are needed for different purposes. For exploration and pedagogy, the criteria should be transparency and simplicity and, for that, toy models are the right vehicles. For forecasting, the criterion should be forecasting accuracy, and purely statistical models may, for the time being, be best. For conditional forecasting, i.e. to look, for example, at the effects of changes in policy, more structural models are needed, but they must fit the data closely and do not need to be religious about micro foundations.

(iii) Partial equilibrium modelling and estimation are essential to understanding the particular mechanisms of relevance to macroeconomics. Only when they are well understood does it become essential to understand their general equilibrium effects. Not every macroeconomist should be working on general equilibrium models (there is such a thing as division of labour).

(ii) **Two more controversial propositions**

(i) The specific role of DSGEs in the panoply of general equilibrium models is to provide a basic macroeconomic Meccano set. By this I mean a formal, analytical platform for discussion and integration of new elements: for example, as a base from which to explore the role of bargaining in the labour market, the role of price-setting in the goods markets, the role of banks in intermediation, the role of liquidity constraints on consumption, the scope for multiple equilibria, etc. Some seem to think that it is a dangerous and counterproductive dream, a likely waste of time, or at least one with a high opportunity cost. I do not. I believe that aiming for such a structure is desirable and achievable.

(ii) The only way in which DSGEs can play this role is if they are built on explicit micro foundations. This is not because models with micro foundations are ho-lier than others, but because any other approach makes it difficult to integrate new elements and have a formal dialogue. For those who believe that there are few distortions (say, the believers in the original real business cycle view of the world), this is a straightforward and natural approach. For those of us who believe that there are many distortions relevant to macroeconomic fluctuations, this makes for a longer and messier journey, the introduction of many distortions, and the risk of an inelegant machine at the end. One wishes there were a short cut and a different starting point. I do not think either exists.

(iii) **The way forward**

If one does not accept the two propositions immediately above, then DSGEs are clearly not the way to go—end of discussion. There are plenty of other things to do in life.
If, however, one does accept them (even if reluctantly), then wholesale dismissal of DSGEs is not an option. The discussion must be about the nature of the micro foundations and the distortions current models embody, and how we can do better.

Do current DSGEs represent a basic Meccano set, a set that most macroeconomists are willing to use as a starting point? (For Meccano enthusiasts, the Meccano set number is 10. The Meccano company has actually also gone through major crises, having to reinvent itself a few times—a lesson for DSGE modellers.) I believe the answer to this is no. (I have presented my objections in my first piece on the topic.) So, to me, the research priorities are clear:

First, can we write down a basic model most of us would be willing to take as a starting point, the way the IS–LM model was a widely accepted starting point earlier in time? Given technological progress and the easy use of simulation programmes, such a model can and probably must be substantially larger than the IS–LM but still remain transparent. To me, this means starting from the New Keynesian model, but with more realistic consumption- and price-setting equations, and adding capital and thus investment decisions.

Second, can we then explore serious improvements in various dimensions? Can we have a more realistic model of consumer behaviour? How can we deviate from rational expectations, while keeping the notion that people and firms care about the future? What is the best way to introduce financial intermediation? How can we deal with aggregation issues (which have been largely swept under the rug)? How do we want to proceed with estimation (which is seriously flawed at this point)? And, in each case, if we do not like the way it is currently done, what do we propose as an alternative? These are the discussions that must take place, not grand pronouncements on whether we should have DSGEs or not, or on the usefulness of macroeconomics in general.

V. The need for different classes of macroeconomic models

Building on the above discussion, let me now present what I said in my third Peterson blog post (Blanchard, 2017a), published in January 2017. In this piece I set out just one main point: different classes of macro models are needed for different tasks. I focused on just two main classes.

Theory models are aimed at clarifying theoretical issues within a general equilibrium setting. Models in this class should build on a core analytical frame and have a tight theoretical structure. They should be used to think, for example, about the effects of higher required capital ratios for banks, or the effects of public debt management, or the effects of particular forms of unconventional monetary policy. The core frame should be one that is widely accepted as a starting point and that can accommodate additional distortions. In short, it should facilitate the debate among macro theorists.

Policy models are aimed at analysing actual macroeconomic policy issues. Models in this class should fit the main characteristics of the data, including dynamics, and allow for policy analysis and counterfactuals. They should be used to think, for example, about the quantitative effects of a slowdown in China on the United States, or the effects of a US fiscal expansion on emerging markets.
It would be nice if there was a model that did both, namely had a tight, elegant, theoretical structure and fitted the data well. But this is a dangerous illusion. Perhaps one of the main lessons of empirical work (at least in macro, and in my experience) is how messy the evidence typically is, how difficult aggregate dynamics are to rationalize, and how unstable many relations are over time. This may not be too surprising. We know, for example, that aggregation can make aggregate relations bear little resemblance to underlying individual behaviour.

So, models that aim at achieving both tasks are doomed to fail, in both dimensions. The French have an apt expression to describe such contraptions: ‘the marriage of a carp and a rabbit’.

Take theory models. DSGE modellers, confronted with complex dynamics and the desire to fit the data, have extended the original structure to add, for example, external habit persistence (not just regular, old habit persistence), costs of changing investment (not just costs of changing capital), and indexing of prices (which we do not observe in reality), etc. These changes are entirely ad hoc, do not correspond to any micro evidence, and have made the theoretical structure of the models heavier and more opaque. (I have a number of other problems with existing DSGEs, but these were the topics of my previous blogs.)

Now take policy models. Policy modellers, looking to tighten the theoretical structure of their models, have, in some cases, attempted to derive the observed lag structures from some form of optimization. For example, in the main model used by the Federal Reserve, the FRB/US model, the dynamic equations are constrained to be solutions to optimization problems under high order adjustment cost structures. This strikes me as wrongheaded. Actual dynamics probably reflect many factors other than costs of adjustment. And the constraints that are imposed (for example, on the way the past and the expected future enter the equations) have little justification, theoretical or empirical.

So what should be done? My suggestion is that the two classes should go their separate ways.

DSGE modellers should accept the fact that theoretical models cannot, and thus should not, fit reality closely. The models should capture what we believe are the macro-essential characteristics of the behaviour of firms and people, and not try to capture all relevant dynamics. Only then can they serve their purpose, remain simple enough, and provide a platform for theoretical discussions. Ironically, I find myself fairly close to my interpretation of Ed Prescott’s position on this issue, and his dislike of econometrics for those purposes.

Policy modellers should accept the fact that equations that truly fit the data can have only a loose theoretical justification. In that, the early macroeconomic models had it right: the permanent income theory, the life-cycle theory, and the Q theory provided guidance for the specification of consumption and investment behaviour, but the data then determined the final specification. As Ray Fair (2015) has shown, there is no incompatibility between doing this and allowing for forward-looking, rational expectations.

Both classes should clearly interact and benefit from each other. To use an expression suggested by Ricardo Reis, there should be scientific cointegration. But the goal of full integration has, I believe, proven counterproductive. No model can be all things to all people.
VI. Five kinds of models

The participants in the one-day conference which David Vines organized in February 2017 had already heard me expound on the difference between theory models and policy models. But the papers prepared for that meeting in February led me to conclude that it is useful to distinguish between five different types of models. In saying this, I limit myself to general equilibrium models. I do this although—as I have said above—much of macro must, however, be about building the individual pieces, constructing partial equilibrium models, and examining the corresponding empirical evidence, components on which the general equilibrium models must then build.

Foundational models

The purpose of these models is to make a deep theoretical point, likely of relevance to nearly any macro model, but not pretending to capture reality closely. I would put there the consumption-loan model of Paul Samuelson, the overlapping generation model of Peter Diamond, the models of money by Neil Wallace or Randy Wright (Randy deserves to be there, but the reason I list him is that he was one of the participants to the Vines conference. I learned from him that what feels like micro-foundations to one economist feels like total ad-hocery to another), the equity premium model of Edward Prescott and Rajnish Mehra (Mehra and Prescott, 1985), the search models of Peter Diamond, Dale Mortensen, and Chris Pissarides.

DSGE models

The purpose of these models is to explore the macro implications of distortions or sets of distortions. To allow for a productive discussion, they must be built around a largely agreed upon common core, with each model then exploring additional distortions, be it bounded rationality, asymmetric information, different forms of heterogeneity, etc. (At the conference, Ricardo Reis had a nice list of extensions that one would want to see in a DSGE model.)

These were the models David Vines (and many others) was criticizing when he started his project and, in their current incarnation, they raise two issues.

The first is what the core model should be. The current core, roughly an RBC structure with one main distortion, nominal rigidities, seems too much at odds with reality to be the best starting point. Both the Euler equation for consumers, and the pricing equation for price-setters seem to imply, in combination with rational expectations, much too forward-lookingness on the part of economic agents. I, and many others, have discussed these issues elsewhere, and I shall not return to them here. (I learned at the conference that some economists, those working with agent-based models, reject this approach altogether. If their view of the world is correct, and network interactions are of the essence, they may well be right. But they have not provided an alternative core from which to start.)

The second issue is how close these models should be to reality. My view is that they should obviously aim to be close, but not through ad hoc additions and repairs, such as arbitrary and undocumented higher order costs introduced only to deliver more realistic lag structures. Fitting reality closely should be left to policy models.
Policy models
The purpose of these models is to help policy, to study the dynamic effects of specific shocks, to allow for the exploration of alternative policies. If China slows down, what will be the effect on Latin America? If the Trump administration embarks on a fiscal expansion, what will be the effects on other countries?

For these models, capturing actual dynamics is clearly essential. So is having enough theoretical structure that the model can be used to trace the effects of shocks and policies. But the theoretical structure must by necessity be looser than for DSGE: aggregation and heterogeneity lead to much more complex aggregate dynamics than a tight theoretical model can hope to capture. Old-fashioned policy models started from theory as motivation, and then let the data speak, equation by equation. Some new-fashioned models start from a DSGE structure, and then let the data determine the richer dynamics. One of the main models used at the Fed, the FRB/US model, uses theory to restrict long-run relations, and then allows for potentially high order costs of adjustment to fit the dynamics of the data. I am sceptical that this is the best approach, as I do not see what is gained by constraining dynamics in this way.

In any case, for this class of models, the rules of the game here must be different than for DSGEs. Does the model fit well, for example in the sense of being consistent with the dynamics of a VAR characterization? Does it capture well the effects of past policies? Does it allow us to think about alternative policies?

Toy models
Here, I have in mind models such as the many variations of the IS–LM model, the Mundell–Fleming model, the RBC model, and the New Keynesian model. As my list indicates, some may be only loosely based on theory, others more explicitly so. But they have the same purpose. They allow for a quick first pass at some question, and present the essence of the answer from a more complicated model or from a class of models. For the researcher, they may come before writing a more elaborate model, or after, once the elaborate model has been worked out.

How close they remain formally to theory is not a relevant criterion here. In the right hands, and here I think of master craftsmen such as Robert Mundell or Rudi Dornbusch, they can be illuminating. There is a reason why they dominate undergraduate macroeconomics textbooks: they work as pedagogical devices. They are art as much science. But art is of much value.

Forecasting models
The purpose of these models is straightforward: give the best forecasts. And this is the only criterion by which to judge them. If theory is useful in improving the forecasts, then theory should be used. If it is not, it should be ignored. My reading of the evidence is the verdict is out on how much theory helps. The issues are then statistical, from overparameterization to how to deal with the instability of the underlying relations, etc.

VII. Conclusion
The attempts of some of these models to do more than what they were designed to do seem to be overambitious. I am not optimistic that DSGEs will be good policy models unless
they become much looser about constraints from theory. I am willing to see them used for forecasting, but I am again sceptical that they will win that game. This being said, the different classes of models have a lot to learn from each other, and would benefit from more interactions. Old-fashioned policy models would benefit from the work about heterogeneity, liquidity constraints, embodied in some DSGEs. And, to repeat a point made at the beginning, all should be built on solid partial equilibrium foundations and empirical evidence.

References

Kocherlakota, N. (2016), ‘Toy Models’, available at https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxrb2NoZXJsYVtvdGdwMDIzZgMzMTAyZmIzODcxNGZiOGY4Yg
Ending the microfoundations hegemony

Simon Wren-Lewis*

Abstract: The New Classical Counter Revolution changed macromodelling methodology and ushered in the hegemony of microfounded models. Yet, in reality, at any particular point in time there is a trade-off between data coherence and theoretical coherence, and insisting that all academically respectable modelling should be at one extreme of this trade-off is a major mistake. The paper argues that if more traditional structural econometric models (SEMs) had been developed alongside microfounded models, links between finance and the real economy would have been more thoroughly explored before the financial crisis, allowing macroeconomists to respond better to the Great Recession. Just as microfoundations hegemony held back macroeconomics on that occasion, it is likely to do so again. Macroeconomics will develop more rapidly in useful directions if microfounded models and more traditional SEMs work alongside each other, and are accorded equal academic respect.

Keywords: macroeconomic modelling, new classical counter revolution, microfoundations, structural econometric models, empirical coherence

JEL classification: B22, B41, E10

I. Introduction

When Olivier Blanchard moved from being Director of the Research Department of the International Monetary Fund to become a fellow of the Peterson Institute in Washington, he published a piece on the future of DSGE modelling, inspired by preparations for this volume. It was very widely read, and comments led to a second follow-up piece. Reactions to that inspired a third in the series (Blanchard, 2017) which I think it is fair to say startled many academic macroeconomists.

In this third piece, published in January 2017, he argued that there should be two types of model: DSGE models, which he called theory models, and models that were much closer to the data, which he called policy models. He described these two classes of model thus:

DSGE modelers should accept the fact that theoretical models cannot, and thus should not, fit reality closely. The models should capture what we believe are the macro-essential characteristics of the behaviour of firms and people, and not try to capture all relevant dynamics. Only then can they serve their purpose, remain simple enough, and provide a platform for theoretical discussions.

* Blavatnik School of Government, Oxford University, e-mail: simon.wren-lewis@sgs.ox.ac.uk

My thanks to two referees for very helpful comments, but I alone remain responsible for the opinions expressed here.

doi:10.1093/oxrep/grx054

© The Author 2018. Published by Oxford University Press.
For permissions please e-mail: journals.permissions@oup.com
Policy modelers should accept the fact that equations that truly fit the data can have only a loose theoretical justification. In that, the early macroeconomic models had it right: The permanent income theory, the life cycle theory, the Q theory provided guidance for the specification of consumption and investment behavior, but the data then determined the final specification.

This was shocking for most academic macroeconomists. They had long argued that only DSGE models should be used to analyse policy changes, because only these models were internally consistent and satisfied the Lucas critique. Largely as a result of these claims, pretty well all academic papers analysing policy in the top academic journals are DSGE models. Even the Bank of England’s core policy model was a DSGE model, and institutions like the Federal Reserve seemed old fashioned in retaining what appears to be a Blanchard-type policy model as their main model. But here was a highly respected academic macroeconomist arguing that a different type of model should be used for policy analysis.

Blanchard’s position is one that I have been publicly advocating for a number of years, and in this paper I want to give my own reasons why I think Blanchard’s position is correct. I want to go substantially further, and argue that the failure of academic macroeconomics to adopt an eclectic methodological position was the key reason why it was so unprepared for the global financial crisis. If academics ignore Blanchard and insist on continuing the microfoundations hegemony, this will hold back the development of macroeconomics just at it did before the financial crisis.

I start in section II with a brief history of how we got here, focusing on the seminal contribution of Lucas and Sargent. Section III argues that microfoundations modelling is a progressive research programme, and that therefore, unlike most heterodox economists, I do not think it should be abandoned. But I also focus on the central flaw in the claim that only DSGE models can give valid policy advice. The implication is that the revolution inspired by Lucas and Sargent was regressive in so far as it argued for the abandonment of more traditional ways of modelling the economy.

Section IV argues that if these more traditional modelling methods had not been abandoned following the New Classical Counter Revolution, but had been allowed to coexist with microfounded macromodels, then links between the financial sector and consumption behaviour in particular would have been explored before the global financial crisis. I draw on evidence from the UK, where what Blanchard calls policy models continued to be developed until 2000, and the empirical work of Chris Carroll and John Muellbauer. This suggests that the microfoundations hegemony has had severe costs for macroeconomics, and could by implication have significant costs in the future. I would argue that ending the microfoundations hegemony is the most important step macroeconomists can take to improve their policy advice, and advance their discipline. Section V concludes by looking at how the two types of models can not only coexist, but learn from each other.

II. How microfounded models achieved hegemony within academia

Mainstream economists do not generally like talking about methodology. It is what ‘others’ tend to do, where ‘others’ can range from heterodox economists to other social
scientists. As one colleague once asked me, do medics spend a lot of time talking about the methodology of medicine?

This might be a tenable position for a settled and successful science: contract out methodological discussion to the philosophers. But for macroeconomics exactly the opposite is the case. During my lifetime the subject has witnessed what is, without doubt, best described as a methodological revolution.

When I started working on macroeconomic models in the Treasury in the 1970s, consistency with the data (suitably defined) was the key criteria for equations to be admissible as part of a model. If the model didn’t match (again suitably defined) past data, we had no business using it to give policy advice. Consistency with microeconomic theory was nice to have, but it was not an admissibility criteria. Later I built models that were rather less bound by the data and paid more attention to theory, but they were still of the same class.

I call this class of models structural econometric models (SEMs) rather than Blanchard’s term policy models, because that describes the essential features of the model. They are structural in the sense of incorporating a good deal of theory. They are econometric because this is how the equations are typically parameterized.

Among academics in the 1970s, a great deal of work involved econometric investigation of individual equations, or sometimes systems, that could become part of an SEM. However they could also inform theoretical general equilibrium work, which typically started with aggregate equations whose structure would be justified by reference to some eclectic mixture of theory and appeals to this econometric literature. For example Blinder and Solow began a whole literature that examined monetary and fiscal policy in the context of examples of this kind of model that included a government budget constraint (see Blinder and Solow (1973)).

All that began to change as a result of the New Classical Counter Revolution (NCCR). To some, this revolution was a response to the Great Inflation and stagflation, just as Keynesian economics was a response to the Great Depression. But there is a simple flaw in that argument. Paradigm shifts require not only new facts that conventional theory finds it increasingly tortuous to explain, but also an alternative theory to explain those facts. As I have argued in Wren-Lewis (2015), the New Classical revolution did no such thing. Indeed, its initial alternative to SEMs were RBC (real business cycle) models where inflation did not even appear.

We can see this clearly in the seminal paper of the NCCR by Lucas and Sargent (1979), aptly titled ‘After Keynesian Macroeconomics’. They start their article with references to stagflation and the failure of Keynesian theory. A fundamental rethink is required. What follows next is crucial. If the Counter Revolution was all about stagflation, we might expect an account of why conventional theory failed to predict stagflation: the equivalent, perhaps, to the discussion of classical theory in The General Theory. Instead we get something much more general: a discussion of why identification restrictions typically imposed in the SEMs of the time are incredible from a theoretical point of view, and an outline of the Lucas critique. The essential criticism in Lucas and Sargent is methodological: the way empirical macroeconomics has been done since Keynes is flawed.

Using this critique, they argue that traditional SEMs cannot be trusted as a guide for policy. Only in one paragraph do they try to link this general critique to stagflation:
there is no attempt in this paragraph to link this stagflation failure to the identification problems discussed earlier in their text. Indeed, they go on to say that they recognize that particular empirical failures (by inference, like stagflation) might be solved by changes to particular equations within traditional econometric models. Of course that is exactly what mainstream macroeconomics was doing at the time, with the expectations augmented Phillips curve.

If the New Classical Revolution did not provide an alternative theory that would explain stagflation, why was it so successful? The first reason is simply the innovation of rational expectations. The new generation of academics could immediately see that rational expectations were a natural extension of the idea of rationality that was ubiquitous in microeconomic theory. Yet rational expectations was also fatal to the favoured theoretical model of the time, which involved Friedman’s version of the Phillips curve. Put rational expectations into this Phillips curve, and it implied output or unemployment would follow random rather than persistent deviations from their natural levels. At the time, traditional Keynesians did not see the importance of the dating of the expected inflation term (New Keynesian economics came much later), so instead they attacked the idea of rational expectations, which was a fatal error.

A second reason is that a new generation of economists found the idea of microfounding macroeconomics very attractive. As macroeconomists, they would prefer not to be seen by their microeconomic colleagues as pursuing a discipline with seemingly little connection to their own. I suspect a third reason was methodological. The methodology of traditional macro required economists of a theoretical inclination not only to understand and be knowledgeable about the latest empirical results, but also to compromise their theory to be consistent with these results. Whether this was Lucas and Sargent’s original intention or not, as it developed, the new methodology of RBC modelling avoided making difficult decisions in reconciling theory and data, as I discuss at length in the next section.

There was another important development in econometrics that undermined the traditional methodology. Sims (1980) introduced the idea of vector autoregressions (VARs), and at the same time launched a strong attack on identification criteria typically used in more traditional time series econometric techniques. One of his key arguments involved rational expectations. This led to an explosion in VAR building, and therefore a move away from the older econometric techniques that were the foundation of both SEMs and more traditional theoretical analysis of macroeconomic systems.

Both these developments allowed a bifurcation in academic work, and a degree of methodological purity in either case. Macroeconomists could focus on building models without worrying about the empirical fit of each equation they used, and applied time series econometricians could allow the data to speak without the use of possibly dubious theoretical restrictions.

Whatever the precise reasons, microfounded macroeconomic models became the norm in the better academic journals. A paper where some aggregate equations are not clearly and explicitly derived from microeconomic theory will almost certainly not be published in leading journals. In contrast, consistency with the data has become selective: the researcher chooses what ‘puzzles’ they want to explain, and (implicitly) what data they will not attempt to explain. As the arguments of Lucas and Sargent found their way into teaching at the Masters/PhD level, it began to be accepted that microfounded models were the only ‘proper’ way for academics to model the economy.
This hegemony, and the adoption by some key central banks (such as the Bank of England) of DSGE models as their core model, would almost certainly not have been possible without the acceptance of New Keynesian theory, and in particular Calvo contracts, as legitimate microfoundations. It became clear that microfounded models that allowed some form of price rigidity were more successful in reproducing the key features of business cycles than RBC models. As a result, DSGE models were developed which were essentially RBC models with the addition of price rigidities. But the hegemony of DSGE models was not complete, with some important institutions like the Federal Reserve maintaining an SEM as their core model (although they also operate a DSGE model). In addition, many other models used for both forecasting and policy analysis remained essentially SEMs. In the next section I explore why that might be.

III. The microfoundations methodology is progressive but does not dominate

Many macro-modellers today view SEMs as old fashioned, and see microfounded models as simply using better theory. Under this view there has been no methodological change, but simply progression using better theory. But that ignores two things. The first, and less important, is that SEMs are capable of embodying a considerable amount of theory, as we shall see when I describe the UK experience. Second, there is a clear methodological difference between the two classes of models. In microfounded models, internal theoretical consistency is an admissibility criteria: models that are not internally consistent do not get to be called microfounded. That is not the case for SEMs. In some SEMs what we could call ‘external consistency’, consistency with the data, is the selection criterion, but in most that is not a rigid rule, with various restrictions imposed on the basis of theory that might be rejected by the data. As a general class, SEMs are methodologically eclectic in this sense.

The idea of a trade-off in modelling between respecting microeconomic theory and respecting the data will seem puzzling to those more familiar with the hard sciences, although I suspect it is less strange to those involved in medicine. When Adrian Pagan was asked to prepare a report on modelling at the Bank of England (Pagan, 2003), he presented the following diagram (Figure 1). The curve presents an efficient frontier, and unless we know what preferences users have between theoretical and empirical coherence, we cannot rank any models on that curve.

DSGE models, because being microfounded is an admissibility criterion, have a high degree of theory coherence but a much more limited degree of data coherence. They do not completely ignore the data. Even the original RBC models were designed to try and match key correlations in the aggregate data. Calibration of parameters was sometimes based on microeconomic evidence. More recently parameters have been estimated using Bayesian system methods, although these certainly have their weaknesses. Caballero (2010) describes this as ‘Pretence-of-knowledge Syndrome’.

At the other end of the spectrum is a VAR, which simply looks at the historical interaction between a bunch of macroeconomic variables. This only uses theory in selecting the variables and a lag length. An analogy in medicine would be a treatment that was based only on historical evidence (those who did X were less likely to get Y) with no
obvious biological justification, whereas at the other end of the spectrum would be a treatment based on biology but with, as yet, very little evidence that it worked.

The ideal, of course, is to have a model that is both consistent with microeconomic theory and consistent with the data. At any particular point in time that is not achievable either because we need more research or better ideas, or because other problems (such as aggregation) make the ideal unachievable. As with medicine, the option of waiting until this ideal is achieved is not available to policy-makers and those that advise them.

If you accept this framework, then it is clearly quite legitimate to build models that represent some kind of compromise between the extremes of DSGEs and VARs. It is these models that I describe as SEMs. (The various names on Pagan’s curve occupying this middle ground need not concern us here.) SEMs are likely to be attractive to policy-makers because they typically combine enough theory to ‘tell a story’, but give the policy-maker the confidence that the models predictions are consistent with past evidence. To an academic, however, the compromises involved may make model building seem more like an art than a science.

The attraction of microfounded models is that their theoretical content is not just plausible but rigorously derived. An academic paper containing a new DSGE model will derive every equation from individual optimization, even though the derivation for particular equations might have been seen before in countless other papers. Only by doing this can the paper demonstrate internal consistency (with microeconomic theory) for the model as a whole. But how do we know that such modelling is making progress in explaining the real world?
The answer is puzzle resolution. The modeller will select a ‘puzzle’, which is essentially a feature of the data that is inconsistent with the existing microfoundations literature. They then present a new model, or elaboration of the standard model, that can explain this puzzle. This process has weaknesses. It is sometimes not checked whether this method of solving the puzzle is empirically well grounded, or whether it creates some new puzzle of its own.

For example, Korinek (2015) notes some areas where, in an effort to fit aggregate behaviour, DSGE modellers have used elasticities of labour supply that are much greater than most microeconomic studies suggest. This can help create what Pfleiderer (2014) calls ‘chameleon models’: models built on assumptions with dubious connections to the real world but which lead to conclusions that are uncritically applied to understanding our economy. Although Pfleiderer’s main concern is models in finance, I think RBC models qualify as chameleon models in macroeconomics.

These are hazards which microfounded models need to guard against, but if they do so the puzzle resolution procedure can lead to better models. For that reason, I disagree with many heterodox economists who see this methodology as degenerative (to use the terminology due to Lakatos). The puzzle solving strategy does not guarantee that the research strategy is progressive. It could be like trying to fit a square through a triangular shape: every time one puzzle is solved, it creates a new puzzle. However, I do not think there is any clear evidence that this has happened to the microfoundations project as yet.

This is partly because the majority of modellers are what I call pragmatists (Wren-Lewis, 2011). They are prepared to use short-cuts or ‘tricks’ to enable complex theoretical ideas to be condensed into a more tractable form. An example I have already mentioned is Calvo contracts, which are understood to proxy the rational response to menu costs, but another would be money in the utility function to represent transaction services. In Wren-Lewis (2011) I argue that tricks of this kind lead to an important modification in the idea of theoretical consistency, which is partly why those who I have described as microfoundations purists have objected to this practice (Chari et al., 2008).

An example of this tension between purists and pragmatists is explored by Heathcote et al. (2009) in connection with incomplete markets. They describe two modelling strategies. The first involves modelling ‘what you can see’. This is ‘to simply model the markets, institutions, and arrangements that are observed in actual economies’. The second involves modelling ‘what you can microfound’, which is ‘that the scope for risk sharing [in the model] should be derived endogenously, subject to the deep frictions that prevent full insurance’. The contrast between modelling what you can see and what you can microfound is another, rather evocative, way of describing the difference between the goals of pragmatists and purists.

I think this contrast also rather vividly illustrates why a purist strategy cannot ensure that it produces models that are useful to policy-makers. To argue that, because I do not understand a real world phenomenon, I will assume, for the moment, that it does not exist, is an idea any philosopher of science would find difficult to understand. That some macroeconomists seriously argue that you should only model what you can microfound perhaps illustrates the costs of giving little thought to the methodology of your subject.

Pragmatism in microfounded macromodelling can, of course, go too far in an effort to match the data, which is one reason Blanchard gives for having two classes of model. To quote from the piece discussed earlier:
DSGE modelers, confronted with complex dynamics and the desire to fit the data, have extended the original structure to add, for example, external habit persistence (not just regular, old habit persistence), costs of changing investment (not just costs of changing capital), and indexing of prices (which we do not observe in reality), etc. These changes are entirely *ad hoc*, do not correspond to any micro evidence, and have made the theoretical structure of the models heavier and more opaque.

To use the language of Wren-Lewis (2011), these are ‘tricks’ that do not in fact summarize extensive ‘off-model’ microfounded research.

That a research strategy is progressive is not in itself a reason or justification for hegemony. It is quite possible that an alternative modelling strategy might also be progressive, and the two could co-exist. The reason to suspect that might be the case with macroeconomics is that microfounded models tend to be very poor at fitting the data. SEMs, in contrast, can (almost by definition) fit the data much better. Now it could be that the reason for this is entirely spurious correlation, and that the microfounded model is simply the best we can do. But this seems highly unlikely, in part from our own experience with the development of microfounded models. The explanatory ability of these models today is far greater than the simple RBC models of the past. The additional explanatory power of SEMs may simply represent a set of puzzles that has yet to be solved.

The response of most economists who use microfounded models to the idea that SEMs could be an alternative to DSGE models is normally very simple. SEMs are internally consistent only through luck and good judgement, are therefore subject to the Lucas critique, and therefore cannot produce valid policy conclusions. But the counterargument is straightforward. Microfounded models, because they are unable to match so much of the data (or at least much less than an SEM), are misspecified and therefore cannot produce valid policy conclusions. They are implicitly ignoring real world phenomena that we currently do not fully understand.

We can show this using a simple hypothetical example. Suppose we find that the fit of a model with a standard intertemporal consumption function in a DSGE model can be improved by adding a term in unemployment to the equation. The term is significant both in statistical terms and in terms of how it influences model properties. In *ad hoc* terms it is possible to rationalize such a term as reflecting a precautionary savings motive, and as a result the term would be *immediately included under SEM criteria*.

Now it is quite possible that the model including the unemployment term, which would no longer be a DSGE model but which would be an SEM, is no longer internally consistent and fails to satisfy the Lucas critique. A microfounded model of precautionary saving, were it to be developed, might be able to justify the existence of the unemployment term, but with certain restrictions which were absent from the SEM. But the project to build such a model has yet to be undertaken, and will not bear fruit for perhaps years. So the perfectly valid question to ask is: which is the better model today at providing policy advice: the DSGE model which is internally consistent but excludes the unemployment term, or the SEM that includes it? *There is no reason to believe that the advantages of internal consistency outweigh the disadvantage of a likely misspecification relative to an SEM.*

---

1 I am sometimes amazed when I hear DSGE modellers say ‘but all models are misspecified’ as if that was any kind of argument for ignoring the data.
The response that academics should wait until some microfoundations are developed that account for the impact of unemployment on consumption becomes more problematic when you think of the following experiment, which I use again later. Supposing you estimate an SEM by starting with a fairly elaborate DSGE model. You then explore whether you can improve the fit, however defined, of the model by including additional variables in the way described above. Or you might look for dynamics that were significantly different from the DSGE specification, but which nevertheless had an *ad hoc* theoretical justification. My guess is that you would end up with a large number of interventions that significantly improved the fit of the model, and which therefore could improve its ability to do policy analysis. This becomes the new SEM. Should every academic ignore this, and simply wait for the microfoundations that explain all these effects?

Let me take another example where extending the scope of internal consistency within DSGE models actually reduced their ability to analyse policy. (It is also an example which does not involve the Lucas critique, and which therefore also shows that internal consistency is the more general concept.) Before the work of Michael Woodford, DSGE models that included Calvo contracts used an *ad hoc* objective function for policy-makers, typically involving quadratic terms in output and inflation with an empirically based trade-off between the two policy goals. This was an embarrassment for microfoundations modellers, because there was no relationship between the objectives of the assumed benevolent policy-maker and the utility function of the other agents within the model.

What Woodford showed was that it was possible, using Calvo contracts, to derive an objective function with quadratic terms in output and inflation from the utility function of agents in the model, using the property that, under Calvo, general inflation would induce price dispersion and this had microeconomic costs. In other words it was possible to derive an objective function for policy-makers that had the normal form and yet was internally consistent with the rest of the model.

The impact of his analysis was immediate. Within a year or two it became standard for DSGE models to derive their policy-makers' objective function in this way, despite the derivation often being extremely complex and sometimes doubling the length of the paper. But there was a cost. In a typical DSGE model the implied trade-off between inflation and output was not the 1:1 typically used in the past, but one where inflation was far more important than output. Furthermore the data tend to suggest that if output is a proxy for unemployment, then both policy-makers and consumers care more about unemployment than inflation.

This is an example of where a model gained internal consistency, but in a way that made its properties poorer at matching the data and less useful to policy-makers than the inconsistent models that had preceded it.

The typical response to these arguments is that they imply the direction in which DSGE modellers should work, or the puzzles that they need to solve. I absolutely agree: as I have argued before, it is a progressive research strategy, even though the last example is a case where the project took a step back in order to move forward. But that argument avoids the question, which is that at this point in time an SEM may be a more relevant model for policy than a DSGE.

John Muellbauer (Muellbauer, 2016) recently put it like this: ‘Approximate consistency with good theory following the information economics revolution of the 1970s is
better than the exact consistency of the New Keynesian DSGE model with bad theory that makes incredible assumptions about agents’ behaviour and the economy.’ I would phrase the last part of the sentence rather differently, because calling it bad theory and talking of ‘incredible assumptions’ harks back to the traditional Keynesian critiques of the NCCR. It is more that the theory is by now way too simplistic and limited to what is tractable or what ‘tricks’ are regarded as acceptable.

It is also not a valid response to say that the ‘proper role’ for academics is to only work on improving DSGE models, and they should have nothing to do with SEMs. I do not understand the logical basis of that argument. Let me take another example. There was a time when incorporating sticky prices was thought to be completely \emph{ad hoc}. Following the ‘only working with microfounded models’ line, it would mean that academics should not have been involved in any way with models that included price rigidity. Policy-makers, including those in central banks, should have been left to their own devices. This is patently absurd.

Another response to these arguments is to say that the same logic implies that we eventually end up estimating elaborate VARs with \emph{ad hoc} exclusion restrictions. Why not just use a VAR? But an SEM is not a VAR, because it does contain considerable theoretical content. Indeed, if we take the example of an SEM that is an elaborated DSGE model, if we only ever investigate additional dynamics or effects that make some kind of theoretical sense, then we are a long way from a VAR. It is true that the expression ‘some kind of theoretical sense’ is imprecise and may vary from one economist to another, and it is also true that the additions to the DSGE mean that we can no longer be sure of internal theoretical consistency. But on the reasonable assumption that these elaborations that make up the SEM bring us closer to reality, we have good reason to believe that the SEM will give better policy advice than the DSGE it came from.

If there is actually no compelling logical case that can be used to justify the hegemony of microfoundations methodology, why were more traditional SEMs or non-microfounded models eventually excluded from the top journals? This is not central to the subject of this paper so I do not consider this question here (but see Wren-Lewis (2015)). They are more about the sociology of the profession than valid arguments for hegemony.

\section*{IV. Microfoundations hegemony and the financial crisis}

Why is it critical for macroeconomics to end the microfoundations hegemony? I want to argue in this section that this hegemony has held the discipline back, and that a more eclectic environment where SEMs not only coexisted with microfounded models but had equal status with them would allow macroeconomics to advance more rapidly. I want to suggest that this reliance on microfounded models was a key factor in explaining why finance played such a minor role in widely used, state-of-the-art macromodels before the financial crisis, and why these models were only of limited use in assessing the depth and persistence of the recession immediately after the crisis. If, alternatively, SEMs had existed alongside DSGE models, I will argue that the rapid incorporation of a financial sector into DSGE models that we have seen since the crisis would have occurred before the crisis.
This is a strong claim, and as a result I want to proceed in stages, to make three interlinked claims.

(a) The development of SEMs could have continued after the NCCR alongside the development of microfounded models, with SEMs progressing substantially from the models criticized by Lucas and Sargent.

(b) Had this happened, using financial variables to explain trends in consumption would have become an essential part of these SEMs.

(c) If these SEMs had had equivalent academic status to microfounded models, then the latter would have used the former as a means of prioritizing puzzles to investigate. That would in all likelihood have led to the development of microfounded models involving the financial sector before the financial crisis happened.

Claim (a) can in fact be verified empirically. To a large extent it happened in the UK, where academic involvement in SEMs continued until the late 1990s. In particular, the UK’s social science research council (the ESRC) provided a sustained period of funding for a ‘Macromodelling Bureau’, which brought together all the major SEMs in the UK, and compared their properties in a critical way (see Smith, 1990). The funding of the Bureau only came to an end when other UK academics argued that this was not an efficient use of research funding, where efficiency included the ability to publish papers in the top (mainly US) journals.

The models the Bureau analysed in the 1990s were not all like the models that Lucas and Sargent had criticized so effectively. Many embodied rational expectations as routine, and in that and other ways attempted to reduce the impact of the Lucas critique (see Church et al., 2000). They began to look more like the DSGE models that were being developed at the same time. However they remained quite distinct from these models, because most were concerned to include aggregate equations that were able to track the time-series properties of the data on an equation-by-equation basis, and partly as a result they were almost certainly internally inconsistent. (Inconsistency also came from the need to be a ‘horse for all courses’: see Currie (1985).) It is a model of this kind that, for example, is currently used by the US Fed for forecasting and policy analysis (Brayton et al., 2014).

A good example of such an SEM maintained in the UK academic sector was developed by myself with two colleagues at Strathclyde University and called COMPACT. It was built using finance from two substantial 4-year ESRC awards under the Bureau umbrella, and the outputs from each award were rated outstanding by the academics who subsequently reviewed it.

For those who still think of SEMs in terms of embodying pre-NCCR theory, it may be useful to note a few of features of this model. (Darby et al. (1999) provides much more detail.) It included a consumption function based on the Blanchard/Yaari consumption model, but where a proportion of consumers were credit-constrained. The model was New Keynesian, with rational expectations in both price and wage equations. In this sense the basic structure of the model was quite similar to the typical

---

2 Not all academic work on SEMs stopped after the NCCR (see the continuing work of Ray Fair). By ‘equivalent academic status’ I mean that this work was considered publishable in the top academic journals.
reduced form of New Keynesian models, with a forward dynamic IS curve and a New Keynesian Phillips curve. The model also included an investment equation which was based around Tobin’s Q.

Of particular relevance to this paper was the consumption function. Unlike DSGE models at the time, the proportion of credit-constrained consumers was not fixed, but varied with the degree of credit availability, an exogenous variable based on particular financial variables. The inspiration for this specification had come from the work of John Muellbauer and the consumption boom in the UK in the late 1980s.

The Bureau, and work on COMPACT, came to an end at the turn of the century when the funding stream ended. Work on the model did not continue (beyond a brief applied study: Keogh-Brown et al., (2010)), and my own research moved into analysing DSGE models. As a result, claim (b) cannot be verified by example. However, the econometric work of John Muellbauer did continue, and both he and Chris Carroll showed that it was very difficult to model trends in both UK and US consumption behaviour without bringing in some measure of financial liberalization (see for example Aron et al. (2010) and Carroll et al. (2012)).

These trends in the savings ratio, which are quite inconsistent with the standard intertemporal model, were very apparent before the financial crisis, yet as a result of detrending techniques they can be filtered out by empirical work using DSGE models. In contrast, SEMs do not pre-filter the data, and so explaining these trends should be a key requirement for any consumption function in an SEM. This is a good example of how microfoundations can be selective in the empirical puzzles it addresses.

It seems quite reasonable to argue, therefore, that had SEMs continued to be built and developed in the academic sectors of the US as well as the UK, at least some would have incorporated financial effects into consumption behaviour and possibly other areas. If more US (and UK after 2000) academics had been actively involved in SEMs and single-equation econometric modelling, where fitting the unfiltered time series for consumption over a reasonable length of time is critically important, it is likely that this key role for financial variables in determining consumption behaviour would have become both more widely known and explored.

Those operating these SEMs would have understood that to simply treat credit conditions as an exogenous variable, when it was so important in determining the behaviour of such a critical macroeconomic aggregate, was heroic and potentially dangerous. As a result, research would almost certainly have been directed towards understanding this key financial influence, which in turn should have led to a better understanding of financial/real interactions.

Establishing both claims (a) and (b) already means that, once the financial crisis hit, there would have been models available (SEMs) that would have provided a guide to the depth and length of the recession. However, I think we can go further than this, and argue that claim (c) is also likely.

Microfounded models are developed to solve puzzles. The puzzles that are chosen reflect many things, including the likely difficulty in solving the puzzle. I think this is one reason why integrating a financial sector into macroeconomic models was not a high priority before the financial crisis: it is not easy. But I suspect the main reason was that there seemed to be more important puzzles to solve.

If work on SEMs had shown the importance of links from the financial sector for consumption and perhaps other variables before the financial crisis, the need to microfound
those links would have been much clearer. There is an obvious precedent here. RBC models did not include any price rigidities, because these price rigidities at the time appeared to have no basis in microeconomic theory, but central banks were setting interest rates using SEMs that incorporated price rigidities. The Phillips curve was a key macroeconomic relationship. This led macroeconomists to develop New Keynesian theory. Just as price rigidity became a critical puzzle for the microfoundations project to solve, so might financial impacts on consumption if they were embodied in SEMs which in turn were used by policy-makers. At least some of the explosion of work that has been done looking at the role of financial leverage in macro-models since the crisis could have been done before the crisis.

The claim sometimes made by defenders of the microfoundations hegemony that mainstream macroeconomics largely ignored financial to real interactions because there was no empirical reason to do so is clearly false. The importance of these linkages was clearly identified in some of the SEMs that survived in UK academia for a time after the NCCR, and their importance was reaffirmed by empirical work subsequently in both the UK and US. That this was in practice ignored reflects the limited role that empirical evidence plays in the microfoundations project. To put it bluntly, macroeconomists almost everywhere ignored the medium-term behaviour of the largest component of GDP as a result of the methodology they were using. That is a serious problem for the discipline.

V. How DSGE and SEMs can coexist

One of the points that puzzled academics after Blanchard called for two classes of model—DSGE and what he called policy models and I have called SEMs—was that if SEMs provide policy advice what are DSGE models for? One simple answer comes straight from the discussion of section III. We can think of an SEM as incorporating theory in a rough and ready way, but it is clearly better to incorporate it more rigorously. Internal consistency is a goal worth trying to achieve. That alone provides a rationale for the microfoundations project.

A good policy-maker should therefore always want to take note when a DSGE model gives very different answers to an SEM, and SEM modellers should be able to explain the difference, if necessary by using the "method of theoretical deconstruction" outlined in Wren-Lewis et al. (1996). But more importantly, many theoretical macroeconomists will quite understandably want to continue working with internally consistent microfounded models.

Another way of showing how the two types of model can interact is to consider one (but not the only) way to build an SEM, which is to start with a DSGE model and allow the data to add missing dynamics and variable interactions as well as parameter estimates, along the lines of the thought experiment considered in section III. The extent to which you require these additions to be "plausible" in terms of some informal theory is inevitably a matter for judgement, and my own preference is to be quite strict in this regard, but it obviously depends in part on the significance of the change.3

3 The Bank of England came close to this in its BEQM model that preceded its current DSGE model. However it chose to instead create a core and periphery model structure, which created a degree of complexity that proved too much even for the Bank.
In this method of building an SEM, the role of the DSGE model is absolutely clear. But the SEM model also provides something essential to DSGE modellers: a set of puzzles to explain in new work. This shows how the two types of model can not only coexist, but feed off each other in an extremely progressive way.

This example might suggest a division of labour which I think would rather miss the point. A DSGE modeller might suggest that academics could stick with DSGE modelling, and economists in institutions that maintain a large model could build SEMs. This is wrong because it restricts too much what academics can do.

First, it would be foolish to believe that these institutions have the resources to not only maintain an SEM, but also to keep up with the DSGE literature and undertake all the empirical work needed to build good SEMs. One of the lessons from the UK experience with the Modelling Bureau was that the interaction between institutional modellers and those in academia were extremely fruitful.

Second, one of the criticisms that Blanchard makes of DSGE models with which I agree is that these modellers do not take the results of partial equilibrium and micro analysis seriously enough. For example, empirical analysis has for some time emphasized the importance of the precautionary motive in consumption behaviour, but DSGE modellers have only just begun to examine ways (invent tricks) that can get this into their models.

Third, on occasion cases may arise where SEMs or empirical work throw up interactions that are easy to incorporate at the aggregate level, but are difficult to microfound. Historically, price rigidity is a major case in point. In those circumstances, academics might want to quite reasonably explore the properties of such aggregate models.

Fourth, different DSGE models can often lead to the same or similar aggregate specifications. Once again it seems eminently reasonable for academics to start with those aggregates, rather than be forced to choose and reproduce one of the derivations.

These remarks should make it clear that not only can microfounded models and more data based policy models happily coexist, they can learn from each other. I have shown that there is no logical reason why one is superior to the other in providing policy advice. I have argued that the discipline’s response to the financial crisis would probably have been considerably better if this coexistence had happened earlier.

Yet for all this I am not optimistic that this broadening of methodological approaches will happen any time soon. I have had discussions with many academics that show a deep resistance to proposals of this kind. It seems as if the revolutionary spirit embodied in Lucas and Sargent has become embedded into large parts of the discipline. That in turn means that graduates writing their first papers will always run the risk of encountering one of these academics as a referee, and will naturally as a result not take the risk of working outside the microfoundations project. Within academia at least, the microfoundations hegemony seems well entrenched.

References


Where modern macroeconomics went wrong

Joseph E. Stiglitz*

Abstract: This paper provides a critique of the DSGE models that have come to dominate macroeconomics during the past quarter-century. It argues that at the heart of the failure were the wrong microfoundations, which failed to incorporate key aspects of economic behaviour, e.g. incorporating insights from information economics and behavioural economics. Inadequate modelling of the financial sector meant they were ill-suited for predicting or responding to a financial crisis; and a reliance on representative agent models meant they were ill-suited for analysing either the role of distribution in fluctuations and crises or the consequences of fluctuations on inequality. The paper proposes alternative benchmark models that may be more useful both in understanding deep downturns and responding to them.

Keywords: DSGE, representative agent, deep downturns, economic fluctuations

JEL classification: A1, A2, E0, E1.

I. Introduction

Dynamic Stochastic General Equilibrium (DSGE) models, which have played such an important role in modern discussions of macroeconomics, in my judgement fail to serve the functions which a well-designed macroeconomic model should perform. The most important challenge facing any macro-model is to provide insights into the deep downturns that have occurred repeatedly and what should be done in response. It would, of course, be even better if we had models that could predict these crises. From a social perspective, whether the economy grows next year at 3.1 per cent or 3.2 per cent makes little difference. But
crises, when GDP falls and unemployment increases significantly, have large consequences for individual well-being now, as well as for future growth. In particular, it is now well recognized that periods of extended economic weakness such as confronted by the US and Europe after 2008 have significant implications for future potential growth.\(^1\),\(^2\)

While the 2008 crisis, and the inability of the DSGE model to predict that crisis or to provide policy guidance on how to deal with the consequences, precipitated current dissatisfaction with the model, the failings are deeper: the DSGE model fails similarly in the context of other deep downturns.

The DSGE models fail in explaining these major downturns, including the source of the perturbation in the economy which gives rise to them; why shocks, which the system (in these models) should have been able to absorb, get amplified with such serious consequences; and why they persist, i.e. why the economy does not quickly return to full employment, as one would expect to occur in an \textit{equilibrium} model. These are not minor failings, but rather go to the root of the deficiencies in the model.\(^3\)

What we seek from a ‘benchmark model’ in macroeconomics is not always clear. Vines (written correspondence) suggests that we should be looking ‘for a \textit{simple} model which we could teach to graduate students to provide them with a less misleading framework on which to build whatever research project they are engaged with.\(^4\)

Blanchard (2017) suggests:

\begin{quote}
The models should capture what we believe are the macro-essential characteristics of the behavior of firms and people, and not try to capture all relevant dynamics. Only then can they serve their purpose, remain simple enough, and provide a platform for theoretical discussions.
\end{quote}

Thus, a distinction is often made between the core DSGE model as a benchmark model and the variety of expanded models, which introduces a large number of complexities—more shocks than just technology shocks; and more frictions than just nominal wage and price rigidities.\(^4\) To be sure, as many users of DSGE models

\(^1\) This implies that macro-econometrics should use a Bayesian loss function with a high weight in explaining/predicting deep downturns. This is markedly different from assessing models by looking at how well they match particular co-variances or moments. Practitioners of DSGE macro-econometrics claim to use Bayesian estimation approaches, by which they mean they include priors (e.g. concerning the values of certain parameters). I emphasize here another aspect: a loss function which puts high weight on being able to predict the events we care about. In terms of priors, the following discussion make clear that I find some of the priors embedded in the standard model less than persuasive.

\(^2\) DSGE models are, of course, not really a model of medium- to long-term growth: that is determined by factors like the pace of innovation and the accumulation of human capital on which they provide little insight. To understand the former, for instance, one needs a much more detailed analysis of technological progress, including investments in basic research and the transmission of knowledge across and within firms, than any standard macro-model can provide.

\(^3\) There are a myriad of versions of the DSGE model, and some versions may have attempted to address one or the other of the concerns I raise. I cannot in the confines of this short article explain why the purported remedies do not adequately address the problems.

\(^4\) See Smets and Wouters (2003). They introduce a total of ten ‘shocks’—‘two “supply” shocks, a productivity and a labour supply shock, (…) three “demand” shocks (a preference shock, a shock to the investment adjustment cost function, and a government consumption shock), three “cost-push” shocks ((…) to the mark-up in the goods and labour markets and (…) to the required risk premium on capital) and two “monetary policy” shocks’—and multiple frictions, including ‘external habit formation in consumption’, a ‘cost of adjusting the capital stock’, and ‘partial indexation of the prices and wages that cannot be re-optimised’. 
have become aware of one or more of the weaknesses of these models, they have ‘broadened’ the model, typically in an *ad hoc* manner.\(^5\) There has ensued a Ptolemaic attempt to incorporate some feature or another that seems important that had previously been left out of the model. The result is that the models lose whatever elegance they might have had and claims that they are based on solid microfoundations are weakened,\(^6\) as is confidence in the analyses of policies relying on them. The resulting complexity often makes it even more difficult to interpret what is really going on.

And with so many parameters, macro-econometrics becomes little more than an exercise in curve fitting, with an arbitrarily chosen set of moments generated by the model contrasted with reality. Standard statistical standards are shunted aside. Korinek (2017) provides a devastating critique:

> First, the time series employed are typically detrended using methods such as the HP filter to focus the analysis on stationary fluctuations at business cycle frequencies. Although this is useful in some applications, it risks throwing the baby out with the bathwater as many important macroeconomic phenomena are non-stationary or occur at lower frequencies. An example of particular relevance in recent years are the growth effects of financial crises.

> Second, for given detrended time series, the set of moments chosen to evaluate the model and compare it to the data is largely arbitrary—there is no strong scientific basis for one particular set of moments over another. The macro profession has developed certain conventions, focusing largely on second moments, i.e. variances and covariances. However, this is problematic for some of the most important macroeconomic events, such as financial crises, which are not well captured by second moments. Financial crises are rare tail events that introduce a lot of skewness and fat tails into time series. As a result, a good model of financial crises may well distinguish itself by not matching the traditional second moments used to evaluate regular business cycle models, which are driven by a different set of shocks. In such instances, the criterion of matching traditional moments may even be a dangerous guide for how useful a model is for the real world. For example, matching the variance of output during the 2000s does not generally imply that a model is a good description of output dynamics over the decade.

> Third, for a given set of moments, there is no well-defined statistic to measure the goodness of fit of a DSGE model or to establish what constitutes an

\(^5\) Thus, the Smets–Wouters model introduces individual heterogeneity, but everyone has the same preferences, there are no capitalists or workers, and accordingly no differences in marginal propensities to consume. In this context, redistributions have no effect on aggregate demand. Below, we argue that redistributions may matter.

\(^6\) A claim that is already stretched when it comes to the use of an aggregate production function (see below) and the derivation of the demand for money. There are, of course, assumptions that *seem* to provide microfoundations, e.g. the derivation of the demand for money based on the assumption that there is a ‘cash in advance’ requirement. But credit, not cash, or money as it is usually defined, is typically required. See Greenwald and Stiglitz (2003). Many models embrace a ‘cop-out’, putting money into the utility function. Later, I will explain why one can’t rely for policy analyses on results derived from these ‘toy’ models, or models embracing such *ad hoc* assumptions.
improvement in such a framework. Whether the moments generated by the model satisfactorily match the moments observed in the real world is often determined by an eyeball comparison and is largely at the discretion of the reader. The scientific rigor of this method is questionable.

Fourth, [DSGE models] frequently impose a number of restrictions that are in direct conflict with micro evidence. If a model has been rejected along some dimensions, then a statistic that measures the goodness-of-fit along other dimensions is meaningless.

Korinek’s conclusion, that ‘the scientific rigor’ of this methodology is ‘questionable’, must be considered an understatement.7

Sometimes, too, a distinction is made between a ‘policy model’, giving practical advice on what to do in different circumstances, and a model with sound theoretical underpinnings. Thus, the standard Keynesian model might (it could be argued) be good enough for telling us whether and how to stimulate the economy, but the fact that its underlying equations are not microfounded makes it theoretically unacceptable; for good theory, we have to turn to DSGE models. These distinctions are, I think, wrong on two accounts. First, as I explain below, I believe the core DSGE models is not good theory: good theory is based on how firms and households actually behave and markets actually work.8 If credit availability is more important than interest rates, then a model which assumes that there is no credit rationing is bad theory. In the crisis, banks couldn’t get access to funds; they were liquidity constrained. Such constraints are not consistent with the underlying DSGE model. And second, the reason for having a model derived with microfoundations is that a policy change could change certain aspects of previously observed reduced form relationships. One has to have a theory to ascertain whether it would. Good policy requires an understanding of the underlying determinants of behaviour.

So too, short-term policy involves short-term forecasting. As Chairman of the Council of Economic Advisers—established by the US Congress to help ensure that the economy remain at full employment through appropriate macroeconomic interventions—under President Clinton, I had the responsibility for overseeing our forecasts, which were also used in budgetary projections. Though it was before the development of the current generation of DSGE models, many of the considerations upon which we focused are excluded from the standard models. We were, for instance, concerned with changes in expectations. But analyses of expectations were (correctly, in my view) not based on what those might be if it were assumed that individuals had rational

7 I should emphasize: empirical tests should not be limited to standard statistics. On average, black holes don’t exist. They are rare events, but their existence plays a crucial role in confirming the theory. They would not be uncovered through a standard regression. A single experiment (observation) was enough to largely confirm Einstein’s relativity theory.

8 Thus, what is sometimes meant by providing microfoundations is providing foundations based on a particular model of human behaviour, rational individuals with perfect information operating in competitive markets, a model which has been widely discredited. Note that in the extreme ‘microfoundation fundamentalist’ view, an analysis of a giraffe’s behaviour could not include the assumption that it had a long neck, because we cannot explain either why it has a long neck or how it can survive, given that it has a long neck (assuming that standard models of circulatory systems could not explain how blood could be pumped that high).
expectations, or acted as if they did, but on survey data of what expectations are and have been.9

So too, we were concerned with changes in consumption. The determination of consumption is, clearly, a key aspect of any good macro-model. But it is clear that in the short to medium term, the shifts in household savings rates are little related to the considerations on which the intertemporal utility maximization of a representative agent focuses. That model cannot provide accurate predictions of such changes, identifying either shifts in preferences or technology, or even of expectations (especially if those are supposed to be ‘rational’) that give rise to changes in consumption.10

Nor has the DSGE model been useful for policy design, e.g. the best way to ‘deliver’ a tax cut. Behavioural economics has provided, I believe, a persuasive case that savings behaviour is subject to nudges, in ways that are inconsistent with the standard model.11 Importantly, differences in responses to the US tax cuts of 2008 and 2009 have more to do with behavioural economics than the determinants of savings behaviour incorporated into DSGE models.12

But DSGE models seem to take it as a religious tenet that consumption should be explained by a model of a representative agent maximizing his utility over an infinite lifetime without borrowing constraints.13 Doing so is called microfounding the model. But economics is a behavioural science. If Keynes was right that individuals saved a constant fraction of their income, an aggregate model based on that assumption is microfounded. Of course, the economy consists of individuals who are different, but all of whom have a finite life and most of whom are credit constrained, and who do adjust their consumption behaviour, if slowly, in response to changes in their economic environment. Thus, we also know that individuals do not save a constant fraction of their income, come what may. So both stories, the DSGE and the old-fashioned Keynesian, are simplifications. When they are incorporated into a simple macro-model, one is saying the economy acts as if . . . And then the question is, which provides a better description; a better set of prescriptions; and a better basis for future elaboration of the model. The answer is not obvious. The criticism of DSGE is thus not that it involves simplification: all models do. It is that it has made the wrong modelling choices, choosing complexity in areas where the core story of macroeconomic fluctuations could be told

9 Moreover, these surveys show that different groups in the population have distinctly different beliefs (expectations), inconsistent with assumptions of rational expectations and common knowledge. See, for example, Jonung (1981), Jonung and Laidler (1988), and Bruine de Bruin et al. (2010).
10 Of course, ex post, one can sometimes interpret changes in consumption as if there were a change in intertemporal preferences. The theory of consumer behaviour is meaningful, however, only if preferences are stable or change in predictable ways.
11 See Camerer et al. (2011) and the discussion of behavioural economics and savings in the context of macroeconomics. There is ample evidence that individuals’ retirement savings cannot be explained within the standard model. See Hamermesh (1984) and Banks et al. (1998).
12 See Parker et al. (2013).
13 The fact that these constraints markedly change savings behaviour has been long recognized. See, for example, Newbery and Stiglitz (1982), Deaton (1991), Aiyagari (1994), and Carroll (1992, 1997, 2001) provide empirical support. A major criticism of standard DSGE models provided by Hendry and Muellbauer (2018) is that they ignore these constraints and assume ‘cash and other liquid assets, stock market and pension wealth minus household debt, and housing wealth as equally spendable’.
using simpler hypotheses, but simplifying in areas where much of the macroeconomic action takes place.

The complexities of the DSGE model require drastic simplifications: to analyse the model, strong parameterizations are required. We know that the parameterizations used in the DSGE models (e.g. constant elasticity utility functions\(^{14}\)) yield predictions that can easily be rejected (e.g. all individuals have exactly the same portfolio of risky assets and homothetic preferences—unitary income elasticities for all goods and unitary wealth elasticities for all assets\(^{15}\)). To make matters worse, even with all the implausible parameterizations, the large DSGE models that account for some of the more realistic features of the macroeconomy are typically ‘solved’ only for linear approximations and small shocks—precluding the big shocks that take us far away from the domain over which the linear approximation has validity.\(^{16}\)

II. The core of the failing: the wrong microfoundations

The core of the failings of the DSGE model can be traced to the attempt, decades ago, to reconcile macroeconomics with microeconomics. There were two approaches. The first was taken by real business cycle (RBC) theory and its descendant, DSGE, which attempted to reformulate macroeconomics by taking the microfoundations of a simplified version of the competitive equilibrium model—just as that model was being discredited by advances in behavioural economics, game theory, and the economics of information. That strand attempted to explain unemployment and other deviations from predictions of the standard competitive model by looking for the minimal change in that model—assuming (typically nominal) price and wage rigidities.

The second strand (see, for example, Bernanke and Gertler (1989), Kiyotaki and Moore (1997), Greenwald and Stiglitz, 1987\(^{a,b}\), 1988\(^{a,b}\), 1993\(^{a,b}\), 2003) and the references cited there), only entering into the mainstream in the aftermath of the Great Recession, attempted to bring together modern micro with macro, incorporating one or more of the ways in which actual markets are far from perfect (‘market failures’)—besides the possibility of nominal wage and price rigidities, and to which they gave a different interpretation and explanation.\(^{17}\) In doing so, this strand resurrected the thinking of Irving Fisher (1933), who promulgated a quite different version of macrdynamics

---

\(^{14}\) It is never made clear why we should feel better about a model that assumes a constant elasticity of marginal utility than a model that assumes a constant savings rate.

\(^{15}\) This is true if there is more than one good or more than one asset—and clearly, ‘good theory’ has to be consistent with that being the case. See Stiglitz (1969). Our toy models shouldn’t break down when we move from one good to two goods. See the discussion below.

\(^{16}\) More recently, techniques have been developed for solving DSGE models with non-linearities. Though some of the models incorporate some of the features that we argue here should be included in any good macro-model, they still leave out many crucial features, make many unreasonable parameterizations, and test their models using unsatisfactory methodologies, as described elsewhere in this paper. For a survey of the advances in non-linear methodologies, see Fernandez-Villaverde and Levintal (2017), and the references cited there.

\(^{17}\) Indeed, information economics had identified the possibility of real rigidities—in competitive equilibrium, real wages and interest rates could be set at levels at which markets do not clear. Risk aversion and instrument uncertainty provide an explanation for slow adjustments. See Greenwald and Stiglitz (1989). More recently, Fajgelbaum et al. (2017) have provided a simple general model in which information flows more slowly during recessions, and so uncertainty is higher and persists.
than the Hicksian interpretation of Keynes based on wage and price rigidity. Fisher emphasized the consequences of flexibility and debt-deflation. The departures from the standard competitive equilibrium model which play a critical role in these models include incomplete contracts and capital market imperfections, including those associated with imperfect information, incomplete risk markets, and market irrationalities. (Minsky (1986) was particularly influential in the latter.) It is perhaps worth noting that policy-makers in recent downturns—in Japan, Europe, and even the US—have been focused on deflation, on the possibility that prices might fall, not on price rigidities.

In the discussion below, I illustrate the inadequacies of the DSGE framework by focusing on the 2008 crisis. Some advocates of DSGE models say these models were not meant to address ‘once-in-a-hundred-year floods’. There are several responses to this defence. The first is by way of analogy: what would one think of a medical doctor who, when a patient comes with a serious disease, responded by saying, ‘I am sorry, but I only deal with colds’? The second is that not only did the model fail to predict the crisis; it effectively said that it couldn’t happen. Under the core hypotheses (rational expectation, exogenous shocks), a crisis of that form and magnitude simply couldn’t occur.

Crisis bring out into the open deficiencies in the model that are not so apparent in our smaller and more frequent fluctuations. I believe that most of the core constituents of the DSGE model are flawed—sufficiently badly flawed that they do not provide even a good starting point for constructing a good macroeconomic model. These include (a) the theory of consumption; (b) the theory of expectations—rational expectations and common knowledge; (c) the theory of investment; (d) the use of the representative agent model (and the simple extensions to incorporate heterogeneity that so far have found favour in the literature); distribution matters; (e) the theory of financial markets and money; (f) aggregation—excessive aggregation hides much that is of first order macroeconomic significance; (g) shocks—the sources of perturbation to the economy; and (h) the theory of adjustment to shocks—including hypotheses about the speed of and mechanism for adjustment to equilibrium or about out-of-equilibrium behaviour. I cannot review in detail all of these and other failings—such as the failure to include crucial institutional details—in this brief note, and so I am selective, highlighting a few as examples of more general problems. Many of these are related. For instance, the presence of imperfect and asymmetric information leads to credit and equity rationing. Thus, individuals in maximizing their lifetime utility have to take into account credit constraints and, as we have already noted, this gives rise to a markedly different problem than that analysed in the standard DSGE model. One of the reasons that the representative agent model doesn’t work well is that some individuals are credit constrained, others are not. Moreover, numerous studies (see, for example, Kim et al. (2014); Mian and Sufi (2015); Drehmann et al. (2017)) have emphasized the importance of debt for aggregative behaviour; but in a representative agent model, debt (held domestically) nets out, and therefore should have no role. At least at times, short-run to medium-term macroeconomic analysis needs

---

18 The contrast with the Smets–Wouters (2003) model is clear: they assume a complete set of state-contingent securities insuring households against variations in their income.

19 Greenwald and Stiglitz show that there will be real consequences to unexpected disinfation, including through real balance effects. They show that this will be so, whether the resulting redistributive effects are between bank lenders and firms or among firms.

20 The introduction of corporations, with corporate debt owed to households (discussed later), does not address the issues raised here, focusing on household debt.
to pay careful attention to debt and real debt-dynamics. And here, institutional details can matter. The shift from 30-year fixed-rate mortgages to variable rate mortgages with shorter terms played an important role in the crisis, especially when combined with expectations that were not fully rational and credit constraints: when house prices didn’t increase as expected (and it should have been clear that they couldn’t increase forever at those rates) and homeowners faced constraints in refinancing, the bubble broke, and the crisis ensued.21

As I noted earlier, my approach and that of DSGE models begin with the same starting point: the competitive equilibrium model of Arrow and Debreu. It is clear that that model cannot explain many aspects of the economy, including macroeconomic fluctuations. DSGE models begin with the question, what is the minimum deviation from that model required to match macroeconomic behaviour interpreted largely as matching moments? Their first answer was price and wage rigidities, with unanticipated and not fully insured technology shocks. When that failed to do an adequate job, they added multiple shocks and distortions, in a fairly ad hoc way. Standards for what was meant by ‘microfounding’ were similarly ad hoc: putting money into a utility function ‘explains’ money holdings, but tells us nothing about what happens if, for instance, credit availability changes or the probability distribution of monetary emissions changes as a result of a change in monetary policy.22

III. Explaining deep downturns

The approach that I am advocating begins by ascertaining which of the advances in modern microeconomics are most relevant for understanding the fundamental questions of macroeconomic fluctuations: the source of the shocks; amplification—why seemingly small or moderate shocks can have such large effects on macroeconomic variables and individual well-being; and persistence—why the effects of the shocks persist, with say high levels of unemployment long after the initial shock. The interpretation of these deep downturns should translate into policy, explaining, for instance, why government expenditure multipliers may be quite large (consistent with the earlier amplification analysis) and why monetary policy may be relatively ineffective. In this analysis, information imperfections and asymmetries and behavioural economics often play a central role, as do institutions and distributional effects. As I argue below, for instance, the ineffectiveness of monetary policy is not really attributable to the zero lower bound but to the behaviour of banks, the central institution in providing credit to all but the largest firms.23

---

21 This part of the story of the 2008 crisis is now well accepted. See Stiglitz (2010b,c) and Financial Crisis Inquiry Commission (2011).
22 The former might affect the transactions demand for money (credit, not money, is used in most transactions, as noted by Greenwald and Stiglitz (2003)); the latter affects the demand for money as a store of value. See, for example, Tobin (1958), the critique of the portfolio separation theorem (Cass and Stiglitz, 1970), and the implications for monetary and macro-theory (Stiglitz, forthcoming).
23 My argument corresponds closely to that of Hendry and Muellbauer (2018), who note that ‘A major problem with the claim of “theory consistency” is the question of “which theory?” For example, text-book theory, which assumes efficient and close-to-complete markets, well-informed relatively homogeneous agents, little uncertainty, no credit or liquidity constraints, and a stable economy, contrasts with theory that takes account of the asymmetric information revolution of the 1970s and early 1980s associated with Nobel prize winners Stiglitz, Akerlof, and Spence. Relevant theory must incorporate credit and liquidity constraints, incomplete markets with incomplete insurance and high levels of individual and macro uncertainty.’
Because the 2008 crisis was a financial crisis, the standard DSGE models are particularly poorly designed to analyse its origins and evolution: The central problems of finance—bankruptcy, debt, and asymmetric information—simply cannot arise in a representative agent model.24 Who is supposed to lend to whom? And only if the representative agent is suffering from acute schizophrenia can there be issues of information asymmetries, and it is hard to reconcile such schizophrenia with the usual assumptions concerning rationality.

Some DSGE models (e.g. Smets and Wouters, 2003) try to introduce rudimentary finance through having a corporate and a household sector. But the 2008 crisis can’t be explained within that model: it was some households borrowing from others that gave rise to the crisis. Besides, with a representative agent, with or without firms, finance would always be provided in the form of equity—so there still wouldn’t be bankruptcies and debt crises.25

(i) The shocks

The critique of the DSGE models’ relevance for deep downturns in general and the 2008 crisis in particular begins with the source of the crisis itself. For instance, in (most) DSGE models, downturns are caused by an exogenous technology shock. In agriculture, we know what a negative technology shock means—bad weather or a plague of locusts. But what does that mean in a modern industrial economy—an epidemic of some disease that resulted in a loss of collective knowledge of how to produce?26

By contrast the shocks giving rise to economic fluctuations in many, if not most cases, is clearly endogenous.27 The 2008 shock was endogenous, caused by the breaking of the housing bubble—something that markets created, and to which misguided

---

24 The Congressional inquiry into the 2008 crisis called itself the Financial Crisis Inquiry Commission and focused on aspects of the financial sector like credit rating agencies and the role of credit default swaps (CDSs), derivatives, and other complex financial instruments. The standard DSGE models have nothing to say about either of these: these are failings related to its inadequate treatment of the financial sector.

25 There are further criticisms of their particular formulation: with a corporate sector, wealth would have to include the capitalized value of future dividends (they ignore this aspect of wealth). Recent research in macroeconomics has focused on variations in the value of corporations relative to the value of their capital goods, as a result of changes, e.g. in tax laws and market power. See, for instance, Gonzales (2016), Gonzales and Trivin (2017), and Stiglitz (2015, 2016b).

26 Conceptually, there can be a shock to (beliefs about) the total supply of a critical natural resource—with a belief that, for instance, there was a large supply of oil underneath Saudi Arabia being disproved. While the oil crises of the 1970s were a result of an oil shock, it was politics, not a technology shock of the kind incorporated in DSGE models.

27 For example, inventory fluctuations. Later, we note that the economy is best described as adjusting to shocks through a decentralized process of wage and price adjustments. Such adjustment processes themselves may give rise to economic fluctuations. See, for example, Stiglitz (2016a). The one exception to the view just expressed, that shocks are endogenous, relates to open emerging economies, where there is some evidence that many, if not most, come from abroad. But here, too, the DSGE models fail. The economy’s exposure to exogenous risks is endogenous. The ‘rules of engagement’, e.g. the rules governing capital market liberalization, determine the extent to which a country is affected by shocks occurring elsewhere. See Ocampo and Stiglitz (2008) and Stiglitz (2010a,b).
policies may have contributed.\textsuperscript{28} And to the extent that there are exogenous shocks, the extent to which firms and households are exposed to those shocks is endogenous, affected by the structure of the market.

(ii) Finance: preventing excessive risks and designing stable systems

The main problem in crisis prevention today centres around preventing the financial sector from undertaking excessive risks and ensuring the stability of the financial system. Policy-makers recognize that some of the most important shocks to the economy can come from the financial sector.

In standard models, the money demand equation is supposed to summarize all that is relevant for finance; and, indeed, not even that is very relevant—all that matters is that somehow the central bank is able to control the interest rate.\textsuperscript{29} But the interest rate for T-bills is not the interest rate confronting households and firms; the spread between the two is a critical \textit{endogenous} variable.\textsuperscript{30} While large firms may turn to capital markets, small and medium-sized enterprises (SMEs) rely on the banking system. Under current arrangements, the links between aggregate credit creation and the levers controlled by the regulatory authorities, including the central bank, are tenuous and variable. Among the most important levers are regulations that have typically not been included within the ambit of macroeconomic analysis.\textsuperscript{31}

Moreover, finance and the structure of the financial system matter for stability. Understanding the structures that are most conducive to stability, and the central trade-offs (e.g. between the ability to withstand small and large shocks) represents one of the areas of important advances since the crisis.\textsuperscript{32} These were questions not even posed within the DSGE framework—they could not be posed because they do not arise in the absence of a well-specified financial sector, and would not arise within a model with a representative financial institution.

\textsuperscript{28} See, for example, Bernanke (2009), Demyanyk and Van Hemert (2009), Sowell (2009), and Mian and Sufi (2015).

\textsuperscript{29} This hypothesis can be directly tested, and rejected. See Fama (2013).

\textsuperscript{30} This is one of the central points in Greenwald and Stiglitz (2003), who develop a simple model of banking in which the spread is determined. The importance of this spread has been noted in other papers in this issue. See, for example, Vines and Willis (2018). Gilchrist and Zakrajšek (2012) show that the spread has a predictive power on economic activity.

\textsuperscript{31} And again, these levers are typically left out of the standard DSGE model, even though they may be far more effective. The point is simple: if banks are constrained in their lending by capital adequacy or liquidity constraints, changes in those constraints can have large effects, far greater than those generated by the ‘substitution effects’ that arise as returns to T-bills and loans changes. See Greenwald and Stiglitz (2003). Changes in these rules were part of the policy strategy that helped the economy emerge from the 1991–2 recession (Stiglitz, 2003).

\textsuperscript{32} While some work had begun in this area before the crisis (see Allen and Gale (2000), Greenwald and Stiglitz (2003), and Gallegati \textit{et al.} (2008)), the crisis itself provided enormous impetus to research in this area: see Haldane (2009), Allen \textit{et al.} (2010), Roitman \textit{et al.} (2010), Stiglitz (2010\textsuperscript{a,d}), Gai and Kapadia (2010), Haldane and May (2011), Battiston \textit{et al.} (2012\textsuperscript{a}, Acemoğlu \textit{et al.} (2016), and Battiston \textit{et al.} (2016). Complexity in financial structures may make it even impossible to ascertain whether a system is systemically stable. See Roukny \textit{et al.} (2017). So too, additional financial instruments can lead to greater economic instability (Brock \textit{et al.}, 2008; Caccioli \textit{et al.}, 2009), partly because these financial instruments create new betting opportunities, which enhance volatility (Guzman and Stiglitz, 2016\textsuperscript{a,b}).
One of the key reasons that representative agent models fail in enhancing understanding of macro-fluctuations is the pervasiveness of macroeconomic externalities—the actions of each agent (in the aggregate) have macroeconomic consequences which they do not take into account. These externalities help us understand why markets on their own may be excessively fragile and excessively exposed to risks. Such macroeconomic externalities do not arise in RBC models, and only to a limited extent in standard DSGE models. In the presence of incomplete risk markets and imperfect and asymmetric information, pecuniary externalities matter, and the market equilibrium is in general not Pareto efficient (Greenwald and Stiglitz, 1986; Geanakoplos and Polemarchakis, 1986). Corporations may, for instance, undertake excessive debt (in open economies, excessive dollar-denominated debt), implying that in a downturn there may be fire sales, with the resulting decrease in prices having balance sheet effects, amplifying the downturn (see the next section). Banks engage in contracts with each other that may be individually rational, but result in greater systemic risk, particularly in the face of a large shock. RBC models are structured so that these macroeconomic externalities don’t arise, and so markets are always efficient, even in their response to shocks; and in the new Keynesian models with rigid wages and prices that are their successors, they arise only to a limited extent. By contrast, they are at the centre of the alternative models for which we are arguing in this paper, and help explain the significant deviations from efficient outcomes.

In the standard model, issues of systemic risk simply do not arise. The focus was on inflation, as if excessive inflation was the major threat to economic stability. That has not been the case for a third of a century; but the problems posed by financial instability have been recurrent.

One particularly important implication of the kind of models for which I am arguing here, in contrast to the standard DSGE models, is that in the presence of bankruptcy costs, excessive diversification (capital market integration) may result in shocks being amplified, rather than dampened and dissipated—as assumed by the Federal Reserve and predicted by the standard models. Indeed, policy discourse based on assumptions underlying DSGE models had a kind of incoherence: before a crisis, the conventional wisdom called for diversification—as much as possible, e.g. through securitization and financial linkages/risk sharing. After the onset of a crisis, discourse turned to contagion. The word itself, borrowed from epidemiology, suggest the opposite of diversification: were 100 individuals with Ebola to arrive in New York, no one would recommend a policy of diversification, sending two to each state. Contagion arises because of such linkages. Unless one has succeeded in eliminating the prospect of future crises, the

---

34 See, for instance, ch. 7 of Greenwald and Stiglitz (2003).
35 See Farhi and Werning (2016).
36 See, for instance, Bernanke’s remarks concerning the risk posed by the collapse of the sub-prime market (see Bernanke, 2009). Bankruptcy costs introduce a non-convexity in the analysis—suggesting a fundamental way in which the mathematics of DSGE models would have to be altered. Recent models with financial linkages and bankruptcy costs have shown that dense networks, with many linkages, while better able to handle small and uncorrelated shocks, perform more poorly in response to large shocks. For analyses of optimal diversification, see, for example, Stiglitz (2010a,d) and Battiston et al. (2012a,b).
design of an economic system has to take into account the functioning of the system both before and after a crisis, balancing the benefits of linkages before and the costs afterwards. The conventional wisdom never did that. This is not a minor failing, but a major one.

(iii) Amplification and persistence

Beyond explaining the origins of shocks and the extent to which economies get exposed to shocks, an adequate macro-model needs to explain how even a moderate shock has large macroeconomic consequences. One of the key failures in the 2008 crisis was the prediction that even a large sub-prime crisis would not have large economic consequences because the risks had been diversified. Within the DSGE frame that was being used at the time by key policy-makers, this was a natural conclusion. Other models, however, had predicted otherwise, focusing on important amplifiers within the economy—and, indeed, some of these became part of the standard ‘explanations’ of the crisis.

One source of amplification is ‘balance sheet effects’, the contraction in production and investment that arises when firms suffer a shock to their balance sheets. Providing microfoundations for balance sheet effects requires an analysis of why firms can’t replace the lost equity with new equity, i.e. an explanation of equity rationing (see, for example, Greenwald et al. (1984)). Modern information-based finance provides such a theory, and these ideas have already been integrated into simple theoretical and applied macro-models, in models in which firms’ supply and demand decisions are a function of their balance sheets (see Greenwald and Stiglitz (1993b) and Koo (2008)).

Greenwald and Stiglitz show, for instance, how a price shock (resulting from, say, a shock to demand for the product) gets amplified through the firm’s subsequent decisions on how much to produce, how much labour to hire, and how much to invest.

The effects are amplified still further if there is credit rationing. Not only is there a well-developed theory of credit rationing (e.g. Stiglitz and Weiss, 1981), Calomiris and Hubbard (1989) among others have shown that these constraints are binding in important sectors of the economy, and it appears that they are particularly relevant in those sectors subject to large fluctuations in investment. This played out strongly in the evolution of the 2008 crisis where, by 2010, large firms seemed to be sitting on a couple of trillion dollars of cash, while SMEs remained credit constrained.

At the centre of the modern theory of credit rationing, as observed at the macro level, are banks—a critical institution which was missing from DSGE models. This was a particularly peculiar omission because, without banks, there presumably would be no central banks, and it is the central bank’s conduct of monetary policy that is central in those models. The fact that credit is allocated by institutions
(banks), rather than through conventional markets (auctions) is an important distinction lost in the DSGE framework. Greenwald and Stiglitz (2003) model banks as firms, which take others’ capital, in combination with their own, obtaining and processing information, making decisions about which loans to make. They too are by and large equity constrained, but in addition face a large number of regulatory constraints. Shocks to their balance sheets, changes in the available set of loans and their expectations about returns, and alterations in regulations lead to large changes in loan supply and the terms at which loans are made available. Variations in regulations and circumstances of banks across states in the US are helping validate the importance of variation in the supply conditions in banking in the 2008 crisis and its aftermath.38

Given how long it takes balance sheets to be restored when confronted with a shock of the size of that of 2008, it is not surprising that the effects persisted.39 But they seem to have persisted even after the restoration of bank and firm balance sheets. That suggests that this crisis (like the Great Depression) is more than a balance sheet crisis. It is part of a structural transformation, in the advanced countries, the most notable aspects of which are a shift from manufacturing to a service-sector economy and an outsourcing of unskilled production to emerging markets; for developing countries, the structural transformation involves industrialization and globalization. Not surprisingly, such structural transformations have large macroeconomic consequences and are an essential part of growth processes. DSGE models are particularly unsuited to address their implications for several reasons: (a) the assumption of rational expectations, and even more importantly, common knowledge, might be relevant in the context of understanding fluctuations and growth in an agricultural environment with well-defined weather shocks described by a stationary distribution,40 but it cannot describe changes, like these, that happen rarely;41 (b) studying these changes requires at least a two-sector model; and (c) a key market failure is the free mobility of resources, especially labour, across sectors. Again, simple models have been constructed investigating how structural transformation can lead to a persistent high level of unemployment, and how, even then, standard Keynesian policies can restore full employment, but by contrast, increasing wage flexibility can increase unemployment (see Delli Gatti et al., 2012a,b).

38 Hamid Rashid in some recent unpublished analyses has been able to demonstrate this.
39 Reinhart and Rogoff (2009) emphasize that financial crises tend to be long and persistent. But when the economy experiences a deep real shock, it is inevitable that eventually there will be consequences for the financial sector, including banking. The stock market crash of 1929 didn’t turn into a full-scale banking crisis until several years later. Italy’s current banking crisis is the result of its prolonged stagnation. If financial crises are largely the result of deep and prolonged real shocks, then the statement that economic crises associated with financial crises are long lasting says nothing more than that deep and prolonged crises are deep and prolonged.
40 With global warming, even the assumption that variations in weather are described by a stationary distribution is clearly not correct.
41 Knight (1921) distinguished between risk and uncertainty. The standard model with rational expectations models risk. Here, there is fundamental uncertainty.
(iv) Adjustment and equilibrium

One of the reasons that downturns with high levels of unemployment persist relates to the process of adjustment.42 DSGE models don’t address that issue: they simply assume that the economy jumps to the new equilibrium path.43 Though in a model with a single individual, solving for the value of current values of wages and prices which ensures that the transversality condition is satisfied is conceptually clear (the super-smart individual simply thinks through the consequences of choosing any other set of current wages and prices), it is not apparent how that is to be done in the context of a world without common knowledge. If there were a full set of markets extending infinitely far into the future, the problem I described would not occur. But there are not—this is one of the key market failures.44 That the consequences could be ‘resolved’ by the existence of a representative agent provides no insight into how the absence of these markets is addressed in the actual world. Indeed, this problem arises even if the only difference among individuals is their date of birth; with overlapping generations, and at least some individuals not behaving as if there is a dynastic utility function extending over infinity, not only can there exist sunspot equilibrium (Cass and Shell, 1983), but there can be an infinity of paths consistent with rational expectations (see Hirano and Stiglitz, 2017).

Indeed, there is a disparity between an analysis based on the assumption of instantaneous adjustment to a new equilibrium, and what actually happens—and what most policy economists assume. There is a decentralized process of wage and price adjustment, with wages and prices in each market responding to the tightness in that market (in the labour market, that is the simple Phillips curve, asserting that wages rise when labour markets become tight). Obviously, adjustment processes may be more complicated in a macroeconomic environment with inflation, where nominal adjustment would be expected to take into account inflationary expectations.

In the short run, such adjustment processes may be disequilibrating: the fall in wages as a result of unemployment may result in a decrease in aggregate demand, increasing the level of unemployment. This is especially true (an implicit assumption in Keynes; an explicit assumption in Kaldor (1957) and Pasinetti (1962)) if the marginal propensity

42 As we have noted, empirical DSGE models have introduced a large number of factors to smooth out behaviour, e.g. costs of adjustment in investment and habit formation in consumption. Many of these prolong booms, but some should have the effect of shortening the downturn.

43 If, of course, there are costs of adjustment, the size of the jump may be affected by the structure and magnitude of those costs. My critique here parallels that of Hendry and Muellbauer (2018, this issue) who noted ‘the notion that the economy follows a stable long-run trend is highly questionable, despite heroic attempts by policy-makers to stabilize the path’. My critique goes further: DSGE models assume that even without government intervention, the economy is on the (unique) convergent path. Hendry and Muellbauer go on to argue that ‘The world is usually in disequilibrium: economies are wide-sense non-stationary from evolution and sudden, often unanticipated, shifts both affecting key variables directly and many more indirectly. . . . The assumption made in the business-cycle accounting framework that the economy is never very far from its steady-state trend was simply wrong. These models were unable to capture the possibility of the kind of financial accelerator that operated in the US sub-prime crisis and the resulting financial crisis. They ignored the shock amplifying, propagating, and potentially destabilizing processes in the financial accelerator.’

44 Discussed long ago by Hahn (1966). See also Shell and Stiglitz (1967).
consumption (MPC) differs across groups; the lowering of wages shifts income towards
profits, and capitalists’ MPC is lower than that of workers.

What matters is, of course, real wages, and that depends on the adjustment of wages
relative to prices (see Solow and Stiglitz, 1968). Wages and prices may both be falling,
at the same rate, resulting in real wages being constant, a kind of real wage rigidity. The
increase in real balances (real value of money holdings) would normally be expected
to increase spending, but this effect is relatively small, so that the unemployment equi-
librium could still persist for a long time. Moreover, the deflation itself has a depress-
ing effect, since it increases the real interest rate (holding everything else constant). In
addition, if, as assumed in the previous paragraph, different groups in the economy
have different MPCs, the (unexpected) deflation redistributes income from debtors
to creditors, and this depresses aggregate consumption even more (see Eggertsson and
Krugman (2012) and Korinek and Simsek (2016)). (Even more so, in an open economy,
where the creditor is abroad: it is akin to a transfer of income to foreigners, with espe-
cially great effect then on demand for non-traded goods.) Similarly, adjustments of
declines have balance sheet effects of the kind already discussed, with large macroeco-
nomic consequences.

(v) Financial frictions

Not surprisingly, in the aftermath of the 2008 financial crisis there is a growing con-
sensus that at least one critical failing of the standard model is its (non-)treatment of
the financial sector. Financial frictions, as they have come to be called, are impor-
tant. These include credit and equity rationing and the corollary importance of col-
lateral constraints and of banks, to all of which I have already referred. There are, of
course, a variety of information and enforcement problems that can give rise to finan-
cial frictions. Those that might provide the simplest textbook treatment—showing their
potential importance—may not be the most important. This may matter, because dif-
f erent financial frictions may differ in their policy consequences. In particular, theories
based on costly enforcement (e.g. Eaton and Gersovitz, 1981) or costly state verification
(Townsend, 1979) differ markedly from those based on adverse selection and incen-
tives, noted earlier. Similarly, though important macroeconomic externalities may arise
in any of these models (e.g. with incentive compatibility, self-selection, or collateral
constraints), and, typically, the latter are easiest to analyse, that is partly because the
constraint is not adequately endogenized.

Still, for purposes of a simple benchmark model, it may be far better to incorporate
some financial frictions than to ignore them altogether. Indeed, the core teaching model
for macroeconomics that I use entails a three-period model. The centre of attention is

---

45 That is, not incorporated into an adjustment in the interest rate charged. This particular effect would
not arise if debt contracts were fully indexed, but arises whenever there is unexpected disinflation.
46 There can also be balance sheet effects (the financial accelerator), as described earlier.
48 That is, plausibly, the constraint may change with changes in policy and should be endogenously
derived (see Stiglitz and Weiss, 1986). In the short run, that often does not appear to be the case, so that the
standard approach may not be unreasonable for the development of benchmark models.
today, but this is linked to the past (decisions and shocks in the past affect current state variables) and the future. Valuation functions summarize the future beyond tomorrow, and individuals, firms, and banks today may engage in intertemporal trade-offs. Aggregate demand is based on reduced form functions for aggregate consumption and aggregate investment, in which credit constraints and net worth play an important role. Firm production and investment decisions are motivated by value-maximizing firms facing equity constraints (in the short run, they can’t raise equity), and a rising cost of borrowing as they borrow more (reflecting a higher leverage and an increasing expected value of bankruptcy), with a standard production function. The minimal model incorporating distribution entails two classes of households, workers who consume their income and capitalists who maximize their intertemporal utility function within their borrowing constraints. The central bank sets the T-bill rate, and the loan curve is a function of that rate, bank net worth, regulations, and the state of the economy. An adverse shock shifts the loan curve up (i.e. at any level of borrowing, the representative firm has to pay a higher interest rate). The model can then be expanded, depending on the question being posed. For a more extended description of the model, see the Appendix.

IV. Policy

One of the main reasons we want a good benchmark model is for policy. As we have already noted, short-run forecasting models, even when they conceptually begin within a DSGE framework, add in a variety of variables to increase forecasting accuracy. Having a model which matches moments says little about forecasting accuracy. Especially when there is a deep downturn, governments want to do something. Models constructed for analysing small fluctuations are likely to provide little guidance.

Governments make decisions about specific expenditures, and there is no reason to believe expenditure multipliers associated with public investment that is complementary to private investment will be the same, e.g. as for public consumption expenditures. The former crowds in private investment. But DSGE models are unlikely to be able to handle this kind of subtlety, which is at the core of public policy discourse.

Conventional wisdom, partly based on the standard model, is that over time public deficits designed to stimulate the economy lead to public debt, which can crowd out private capital accumulation, harming growth. But that depends on a host of assumptions: (a) if the public expenditure goes to public capital goods or human capital or technology which are complementary to private capital goods, it can crowd in private capital accumulation; (b) in an economy at a zero lower bound, the government can just print money to finance the expenditures. At such times, one is often worried about deflation; any inflationary effects of such money-printing are thus beneficial.

So too, the conventional insight that with rational expectations, multipliers will be low (zero) because of the expectation of future tax increases depends on special assumptions: (a) if the expenditures are for productivity-enhancing public investments,

49 The limiting case of which is credit rationing—households or firms can’t borrow beyond a certain amount.
the conventional multiplier is actually increased with rational expectations; (b) so too, if there had been the expectation of a prolonged economic downturn; some of the ‘leakages’ from spending today are reflected in spending in future demand-constrained periods, increasing incomes in those periods; consumers, taking this into account, spend more today than they otherwise would have spent.50

With financial frictions, monetary policy may be relatively ineffective, not because of the zero lower bound (if that were really the problem, changes in investment tax credits and consumption taxes over time could have altered individuals’ marginal rates of substitution), but because lowering the T-bill rate (or the Fed discount rate) may not alter bank lending much. If that is the case, policies aimed more directly at increasing credit availability to those borrowers for whom it is constrained may prove more effective than conventional monetary policy.51

The central point is that there is a wide range of policies with significant macroeconomic effects that governments consider, and we have to be able to tailor-make models—building off the core model described earlier—to ascertain the effects. Many policies, for instance, may affect a country’s exposure to shocks (full capital market liberalization); others may affect the strength of a country’s automatic stabilizers. Having a simple model that can analyse these effects is crucial. Building such a model from a DSGE framework is unlikely to be as helpful as building one from the framework described above.52

The economy today is going through a structural transformation. The result may be that with current levels and forms of government expenditure and taxation and private expenditures, the economy might fall short of full employment. For all the reasons discussed above and others, the adjustment to a full employment equilibrium may be slow. But even with sticky wages and prices, there exists a set of fiscal policy interventions over time (taxes, expenditures) which could bring the economy back to full employment in the short run, or at least bring it back to full employment faster than would otherwise be the case: not just one, but a multiplicity of such paths, differing, for instance, in their levels of public investment and growth in the short run. Even if one were concerned about the level of debt, there is a balanced budget multiplier—and if the taxes and expenditures are chosen carefully, that multiplier can be quite large. Thus, ‘secular stagnation’ associated with persistent unemployment is not a disease that happens to a country: it is a consequence of policies that can be changed. Again, as we noted

50 See Neary and Stiglitz (1983). So, too, the standard result on the inefficacy of debt-financed expenditures does not hold if there are binding credit constraints and/or if there are life-cycle savers, who are unconcerned about tax payments beyond their horizon.

51 These remarks may provide insight into the relative ineffectiveness of even the so-called non-conventional policies in bringing the economy back to full employment. Without paying due attention to effects on banks’ balance sheets, negative interest rates could even lead to reduce lending activity. We have already noted how these ideas did play a role in the response to the 1991–2 recession by the Clinton administration. So, too, behavioural economics played an important role in the design of the tax cut in the Obama administration.

52 As we have noted, DSGE models begin with the competitive equilibrium model. Variations in that model focused on open economies therefore include an equation assuming uncovered interest parity (UIP). As Hendry and Muellbauer (2018, this issue) point out: ‘There is strong empirical evidence against UIP. Evidence tends to suggest that for small deviations from long-run equilibrium in which purchasing power parity plays an important role, the exchange rate is not far from a random walk, but for large deviations, equilibrium correction is important.’
earlier, building from a DSGE model, with its assumptions of common knowledge and rational expectations, is not likely to be as helpful in designing policies responding to the structural transformation as beginning with a model focusing on financial frictions, as described earlier.\textsuperscript{53}

V. Further critiques

One could go through each of the underlying assumptions of the DSGE model, to explain the role they play—and why they result in a model that fails to predict and explain important aspects of macroeconomic fluctuations, and why ‘reforms’ which are supposed to improve economic efficiency may actually increase macroeconomic volatility.

(i) On the importance of differences in beliefs

I have, for instance, alluded to the assumption of rational expectations. I strongly believe that one cannot fully explain the growth of the housing bubble that played such a large role in the recent crisis within a rational expectations framework.\textsuperscript{54} But clearly, some of the ‘reforms’ in mortgage markets (strongly supported by the Fed Chair at the time) contributed to the creation of the bubble.

Differences in beliefs, too, can play an important role in macroeconomic fluctuations, through what Guzman and Stiglitz call the creation and destruction of pseudo-wealth. When two individuals differ in beliefs, they have an incentive to engage in a bet (or economic transactions which are similar to bets). Both sides, of course, think that they are going to win, so that the sum of their ‘perceived’ wealth is greater than ‘true’ wealth. Until the bet gets resolved, there is an incentive for both to spend more than they otherwise would, if necessary going into debt. The resolution of the bet (the occurrence of the event) means that one side becomes wealthier, the other side less wealthy; but there is more than just a transfer of income: there is a destruction in aggregate wealth leading to a decrease in aggregate consumption. Pseudo-wealth is being created and destroyed all the time, but certain changes—like the creation of new betting markets, e.g. associated with ‘improvements’ in finance, associated with the creation of markets in derivatives and CDSs—can lead to significant increases in aggregate pseudo-wealth; and certain events, like the collapse of the housing bubble, can lead to its net destruction. Fluctuations in pseudo-wealth help explain one of the paradoxes of macroeconomics:

\textsuperscript{53} In such structural transformation, differences in views are likely to be large, giving rise to the possibility of an increase in pseudo-wealth, as described in the next section, and subsequent volatility. Financial structural reform, allowing for more betting, will increase this volatility, and this should have been taken into account in evaluating the benefits. Again, our critique of the use of rational expectations in such situations parallels that of Hendry and Muellbauer (2018), who note: ‘Shifts in the credit market architecture are only one example of structural breaks and evolutions, implying that the notion that the economy follows a stable long-run trend is highly questionable. . . . Uncertainty then becomes radical. Structural breaks also make it hard to sustain the possibility of “rational” or model consistent expectations.’

\textsuperscript{54} See Shiller (2007) and Stiglitz (2010b).
the large fluctuations in the economy in spite of small changes in the physical state variables, the stock of capital, labour, and natural capital.\footnote{There is ample evidence that individuals differ in their beliefs. Note these theories are consistent with each individual believing that he has rational expectations—he is forming his expectations on the basis of all information available to him. But they are not consistent with common knowledge, where everyone has the same beliefs. There can also be ‘negative’ pseudo-wealth, where what individuals believe they are going to pay to creditors is greater than the creditor believes he receives. See Guzman and Stiglitz (2016a,b).}{55,56}

\section*{(ii) Aggregation}

One set of assumptions that is critical, and to which too little attention is given in macroeconomic analyses, concerns aggregation.

Long ago we learned the difficulties of constructing an aggregate production function.\footnote{There are, of course, other possible explanations for this, e.g. sunspot theories, where there may be multiple equilibria. In this short note, I cannot explain the relative strengths of these alternative explanations.}{57} The ‘putty-putty’ model provides great simplification, but one should not claim that any analysis based on it is really ‘microfounded’. While earlier analyses provided a critique of the use of the standard model for equilibrium analysis, e.g. when there is production of commodities by means of commodities or when there are production processes involving capital goods of markedly different durability;\footnote{That is, the relationship between the value of capital (per capita) and output (per capita) may be far different than suggested by the standard production function. For a review, see Stiglitz (1974).}{58} the use is even more questionable for analyses of dynamics: the dynamics of putty-clay models and vintage capital models, for instance, are markedly different from those of putty-putty models.\footnote{For instance, if, during a period of low interest rates, firms install very capital-intensive machines (with a high output per worker), it will be more difficult for the economy to return to full employment: the necessary increase in aggregate demand will have to be greater than it otherwise would have been. See Morin (2014), Cass and Stiglitz (1969). For a more popular discussion, see Aeppel (2012).}{59} It would thus be foolhardy to rely on the putty-clay model for any analysis of dynamics in the short to medium term when such vintage effects can be important.

Even more important is perhaps the aggregation of the whole economy into a single sector, particularly when the underlying stress on the economy is one of structural change, requiring the movement of resources from one sector to another (say agriculture to manufacturing), when there are market imperfections (say in access to credit) impeding the reallocation.\footnote{Emphasized in the work of Delli Gatti \textit{et al.}, noted earlier.}{60}

Policy analyses are also likely to be misguided. Monetary policy is typically presented as an efficient tool. But monetary policy has disproportionate effects on interest-sensitive sectors, thus inducing a distortion in the economy that simply is not evident in a one-sector model (see Kreamer, 2015).

Finally, the use of a representative agent represents an aggregation of the household sector. It is understandable that macroeconomists attempting to microfound macrotheory would want to impose some restrictions: otherwise, any set of demand functions
could be claimed to be microfounded. But assuming a representative agent goes too far, because it eliminates any possibility that distribution matters. There is at least a significant body of thought that argues that the increase in inequality played some, and possibly a critical role, in the build-up to the crisis and to the slow recovery; there are large differences in the marginal propensity to consume between the top 1 per cent and the bottom 80 per cent and, accordingly, anything that affects distribution significantly affects aggregate demand, i.e. has macroeconomic consequences.

VI. Going still further beyond the standard model

The microeconomics of the basic competitive model—as formulated in Arrow and Debreu—has been shown to be flawed by 40 years of economic research. Why should we expect a macroeconomic model based on such microfoundations to work? Most deeply, the standard model is intellectually incoherent and implicitly encourages society to move in a direction which would undermine both efficiency and well-being. It assumes that all individuals are purely selfish and yet that contracts are always fully honoured. Individuals who are fully selfish know that there are enforcement costs, and will not honour their contracts fully, even if the consequence is a loss in reputation. Thus, the Department of Justice and a number of private suits have uncovered the role of pervasive fraud in the securitization process, by many if not most of the credit rating agencies, mortgage originators, and investment banks, consistent with Kindleberger’s (1978) analysis of earlier depressions and panics. While incorporating such behaviour in a standard economic model is difficult, the prevalence of such behaviour is surely out of the spirit of standard DSGE models and more consistent with those models emphasizing institutional arrangements to prevent such behaviour and the exploitation of imperfections of information. Surely, both policies to prevent a recurrence of similar crises and analyses of market dynamics will need to take into account both market and regulatory responses. Most importantly, the inculcation and normalization of a culture of selfishness without moral bounds will lead to an economy that is less efficient with lower individual and societal well-being. Behavioural economics has noted that most individuals systematically behave differently from that model but that embedding

---

61 That is, according to the Mantel–Sonnenschein theorem, in the absence of some restriction, such as the ‘representative agent’ assumption (where all individuals are assumed identical), virtually any aggregate function can be consistent with the standard competitive model. See Mantel (1974), Sonnenschein (1972), and Kirman (1992). There is also a large literature describing the very restrictive conditions under which such household aggregation can be done.

62 Thus, the critique is far more than that the conditions allowing for such aggregation are not satisfied. That is obviously the case, and the fact that it is raises, too, questions about claims that DSGE models are well microfounded.

63 Interestingly, there were provisions of standard contracts in the securitization process designed to mitigate the consequences of moral hazard, but these provisions failed to work as intended, both because of widespread fraud and breach of contract.

64 Though there have been some attempts to incorporate them and their implications into simple microeconomic models, these have not yet been fully brought into macroeconomic analysis. One important variant of the strand of standard macroeconomics does incorporate insights from one particular variant of financial frictions centring on costly state verification.
individuals within a culture of selfishness (where that is taken as the norm) leads to changes in behaviour in that direction.\footnote{For a discussion of some of the recent empirical evidence, see Hoff and Stiglitz (2016).} Macroeconomics is supposed to provide us with models of how the economy actually behaves, rather than how it might behave in a mythical world of infinitely selfish people but among whom contracts are always honoured. Adam Smith, often described as the father of modern ‘selfish’ economics, in his invisible hand conjecture,\footnote{A conjecture, which we noted, turned out to be false whenever there were imperfect risk markets and asymmetric information, except in the very special conditions underlying real business cycle theory and its descendents.} reminds us in his Theory of Moral Sentiments (Smith, 1759):

> How selfish soever man may be supposed, there are evidently some principles in his nature which interest him in the fortunes of others, and render their happiness important to him, though he derives nothing from it except the pleasure of seeing it.

The earlier Smith was fortunately right, and modern macroeconomics should strive to incorporate behaviour which is consistent with these impulses, just as it does behaviour that is consistent with impulses that may be less noble.\footnote{For an excellent discussion of these two contrasting views of human nature and their implications for economics, see Vines and Morris (2015). For an attempt to incorporate some aspects of these considerations into a formal model, see Greenwald and Stiglitz (1992).} One of the critiques of DSGE modelling is that it and its underlying assumptions have become a dogma, with little incentive to call them into question, especially in a context of peer-reviewed publications.

\section{VII. Concluding comments}

Assumptions matter. All models make simplifications. The question is, as we have said, what simplifications are appropriate for asking what questions. The danger is that the simplifications bias the answers, sometimes in ways that we are not aware of. The DSGE model ignored issues that turned out to be key in the 2008 crisis; not surprisingly, the model neither predicted the most important macroeconomic event in the past three-quarters of a century nor provided good guidance as to the appropriate policy responses. Given the way the models are structured, they could not have predicted such an event. In the run-up to the crisis, monetary authorities focused on inflation rather than on what they should have been focusing on—financial stability; and some of their (especially deregulatory) actions clearly contributed to financial instability. The DSGE models provided them (false) assurance that they were doing the right thing.

Of course, any good macroeconomic model has to be dynamic and stochastic, and present an analysis of the entire economy. But specific assumptions, as we have noted, went into each of these components. We have already discussed several aspects of the assumed dynamics.
Some of the greatest deficiencies, I believe, relate to the treatment of uncertainty, the stochastic element in DSGE models. We have already questioned the underlying presumption in the model of how risks get dissipated through diversification, and that the underlying shocks are exogenous. Also questionable are the typically unstated assumptions concerning risk management. There is ample evidence that risk has first-order effects on firms, households, and banks that are not adequately incorporated into the standard DSGE models. That is why those models had nothing to say about one of the critical questions confronting policy-makers in the 2008 crisis: how best to recapitalize banks. The objective was to enhance lending, especially to SMEs. The way chosen by the US and some other countries, entailing the issuance of preferred shares, can be shown to be far from optimal. There were other ways of bank recapitalization which would have led risk-averse banks to lend more.

In this paper, we have provided many examples of insights that are revealed by simple macroeconomic models with finite periods—insights that are typically obfuscated by the simplifications required by DSGE modelling.

To me, small and big models should be viewed as complementary: one needs to use each to check the results of the other. Perhaps there is some effect that got lost in the three-period simplification. More often than not, it goes the other way. But it is not really a question of small vs big. It is a matter of the careful choice of assumptions. As I have noted, sectoral aggregation is problematic when the underlying macro-disturbance is that of structural change.

That having been said, our models do affect how we think: DSGE models encourage us to think in terms of the economy always moving along a dynamic equilibrium path, and focus our attention on intertemporal substitution. Neither, I suspect, is at the heart of what is really going on in the short to medium term; and as I have suggested, the DSGE has little to say concerning long-term growth. For instance, the belief in the effectiveness of monetary policy has led to the conclusion that its current obvious ineffectiveness is only because of the zero lower bound: if only we could break through that bound the economy could be restored to full employment. Of course, if there were a large enough negative interest rate—if people never had to pay back their loans—there is little doubt that the economy could be stimulated. The question, though, is whether moderate changes, from a real interest rate of say –2 per cent to –4 per cent, would have done the trick, when much larger changes have proven ineffective. The reason for the ineffectiveness lies partly in the fact that lowering the nominal interest rate on T-bills may not lead to a lower lending rate or that lowering the T-bill rate may not lead to an increase in credit availability, as we have already noted. But, as we have also noted, if one really thought that intertemporal prices were the crucial consideration, one could have changed those through tax policies, through changing consumption tax rates and investment tax credits over time.

In the end, all models, no matter how theoretical, are tested in one way or the other, against observations. Their components—like the consumption behaviour—are tested

---

68 This belief encouraged some central banks to move towards negative interest rates, with little success in restoring the economies to robust growth. In some cases, the effects seem negative. Japan’s Central Bank was particularly sensitive to the issues raised here (issues, as we have noted, that were not central in the DSGE models): they worked to mitigate any adverse bank balance sheet effects while maintaining substitution effects.
with a variety of micro- and macro-data. But deep downturns, like the 2008 crisis, occur sufficiently rarely that we cannot use the usual econometric techniques for assessing how well our model does in explaining/predicting these events—the things we really care about. That's why, as I have suggested, simply using a least-squares fit won't do. One needs a Bayesian approach—with heavier weight associated with predictions when we care about the answer. Comparing certain co-variances in calibrated models is even less helpful. There are so many assumptions and so many parameters you can choose for your model, many more than the number of moments you can get from the data; so being able to match all moments in the data does not tell you that your assumptions were correct, and thus does not provide much confidence that forecasts or policies based on that model will be accurate.

Defenders of DSGE models counter that other models did little better than the DSGE models. That is not correct. There were several economists (such as Rob Shiller) who, using less fully articulated models, could see that there was clear evidence of a high probability of a housing bubble. There were models of financial contagion (described earlier in this paper, developed further since the crisis), which predicted that the collapse of the housing bubble would likely have systemic effects. The conviction that this would happen would have been even stronger had the data that the Fed had available to it before the crisis also been available to the public. Policy-makers using alternative models of the kind described here would have done far better both in anticipating the crisis and coping with it than those relying on DSGE models.

Models have consequences even when their predictions and explanatory power are less than stellar. For they affect how households, firms, and, most importantly, policymakers think about the economy. Models which say that the fundamental market failure arises from wage rigidities may be induced to argue for more wage flexibility—to argue that if only we could achieve that, economic performance would be improved.

This essay has argued that the standard DSGE model provides a poor basis for policy, and tweaks to it are unlikely to be helpful. Fortunately, there are alternative frameworks to which modern policy-makers can turn. I have tried here to describe some of the core elements that need to be incorporated into the benchmark models with which we teach macroeconomics to our students. The challenge we should be posing to them is how to develop increasingly sophisticated versions of these into models, small and large, incorporating the various insights provided by a range of ‘partial’ models (such as that of the banking sector) that help us understand the important fluctuations in our economy, and what more we might do to reduce their magnitude and frequency and the human suffering that so often results.

Appendix: Core models for macroeconomic policy

The core models for macro-policy today focus on equilibrium today, and incorporate, in reduced form, insights from recent advances in micro- and macroeconomics, taking on specialized form to focus on one issue or another, one set of circumstances or another. Here, we focus only on closed economy models and on the effect of various policies on output and employment.
(i) **Constructing a barebones model**

The barebones model begins with a neo-Walrasian model, attempting to incorporate the insights of modern microeconomics into the basic aggregate equations, which are not meant to represent the behaviour of a representative individual, but the aggregation of many individuals. Accordingly, a change in, say, a market variable will not only have income and substitution effects, but also distributive effects.

**Basic structure**

The full model has three periods. Most of the analysis focuses on the middle period, labelled period 1, and how policies in that period affect outcomes in that period, identifying the interdependencies between those policies, and the household, firm, and financial sectors. Thus, the model is a *general equilibrium* model, but with a far richer set of interrelationships than characterizes the standard DSGE model. We outline variants in which wages and prices are fully flexible, completely fixed, and intermediate cases. But even when wages and prices are fully flexible, the market equilibrium may deviate substantially from the ‘perfect markets’ equilibrium because of imperfect capital markets, which lead firms to be risk averse and which result in banks playing a central role. Thus, knowing capital and labour supply in this period does not suffice to determine output and employment.

Each period is linked both to the past and the future, and the other two periods (‘zero’ and ‘two’) capture these linkages. What matters is not just firms’ capital stock but also indebtedness (financial structure) and banks’ and households’ balance sheets. Thus, we show how shocks at time zero affect these state variables at time 1.

The objective of bringing in period 2 is to more formally analyse households’, banks’, and corporations’ intertemporal decision-making, e.g. as individuals weigh the marginal utility of consuming today vs the value of a dollar in the future. It enables us to isolate, for instance, the effect of an increase in future productivity on consumption today. This intertemporal problem—in particular, the relevant Euler equations—is really at the centre of attention in DSGE models. The barebones model can be used to study these issues, but as we have suggested in the text, it is almost surely wrong to put them at the centre of macro-analysis.

**Methodological remarks**

The point of the barebones model is to help us understand better how changes in the environment (in an agricultural economy, say, bad weather yesterday or today) or policy affect outcomes, and the channels through which the effects are felt. While one could simultaneously incorporate multiple deviations from the standard competitive model, this is not typically the most insightful way to proceed. Rather, one wants to isolate, say, distribution effects, from balance sheet effects, from the effects of capital constraints, from the regulatory constraints facing banks. We begin by presenting a *general model*, incorporating many, if not most, of the effects discussed in the text (and a few that we touched on only briefly). The real analysis, though, entails specializing this general model.

Underlying the analysis are judgments about the relative importance of different ‘effects’, with some effects being sufficiently small to be ignored—labour supply in the barebones models is assumed inelastic at $L^*$; but it is easy to elaborate on the model...
to incorporate wage effects on labour supply, or even interest rate effects (though that is far-fetched). A key aspect of the analysis is testing for robustness: do results change significantly if, for instance, there is a positive but small labour supply elasticity?

The analysis begins by postulating that the aggregate demand for goods today, \( Y^d \), equals the aggregate supply of goods, \( Y^s \), and the aggregate demand for labour, \( N \), equals the aggregate supply of labour, \( L \). It then goes on to focus on key cases considered in the standard literature: for instance, wage and/or price rigidity; and a no-shirking constraint in the labour market.

It focuses on policy, but not just on the effect of interest rates (which monetary policy is assumed to control), but also on regulatory policy, and on government expenditures. But rather than beginning from the hypothesis that consumption and investment demands are separable from government expenditure, it recognizes that the form of government expenditure can affect either variable.

Policy analysis proceeds by first solving the general equilibrium problem in a reduced-form one-period model—the endogenous variables (in particular output and employment) can be solved for as a function of all the policy variables, \( P \) (given all the other relevant variables describing the economy). It is then a straightforward matter to take the derivative of, say, output or employment with respect to any particular variable, say a particular type of government expenditure. While the resulting expressions are complex and will not be presented here, the analysis can be simplified by assuming that most of the effects are small relative to a few upon which attention is focused. A key difference between the model presented here and the standard DSGE models is that many of the variables that are of first order importance are omitted from that model. The model here, by contrast, pays much less attention to the intertemporal effects upon which the DSGE model focuses. Some of the reasons for this have already been given; others are noted in the final remarks of this Appendix.

The barebones model described below incorporates three effects missing from the standard DSGE models: (i) distribution, (ii) banks, and (iii) credit constraints. One has to pay attention to not just the values of average variables and their changes, but also to dispersion, and changes in dispersion. Knowing that, on average, household wealth was large, with home equity sufficiently great that default was not a significant risk, told one nothing about the state of the economy in 2008: large fractions of households were in fact not able to make their interest payments. Thus, in an important elaboration on the barebones model, dispersions in the relevant state variables (which need to be more formally modelled than here) have to be introduced.

But even in the barebones model, we recognize that the relevant variables are aggregates among heterogeneous agents, and that distribution matters. Thus, a lowering of the interest rate might lower the income of retirees dependent on interest payments, and this distributional income effect may be much larger than any individual substitution effect—including any effect in stimulating investment, except through a collateral effect on credit-constrained firms.

(ii) Period 1

The key ingredients, aggregate demand and aggregate supply, are standard, but their determinants include variables typically left out of DSGE models.
**Aggregate demand**

As in any standard model, we begin by describing aggregate demand:

\[ Y^d = C(N, w, p, r, r_L, t, G, c, E, \xi_b) + I(w, r, p, G, c, E, \xi_f) + G \]

where in the obvious notation \( C = \) consumption, \( w = \) wage rate, \( p = \) price level, \( r = \) T-bill rate, \( r_L = \) (nominal) interest rate at which individuals can borrow (about which we will say more below), \( t = \) tax rate (assumed fixed), \( I = \) Investment, \( c = \) credit availability, \( G = \) government expenditures, and \( E \) represents expectations of the relevant variables (to be described below). \( G \) is a vector of government expenditures, e.g. investment expenditures that are complements (or substitutes) to private investment, consumption expenditures that are substitutes (or complements) to private consumption. Aggregate demand is the sum of consumption, investment, and government expenditures.

Consumption is affected by a variety of market variables, \( m = \{ w, p, N, r_L \} \), wages, prices, employment \( N \), and the borrowing rate \( r_L \); a variety of policy variables \( P = \{ r, t, G \} \) set by the government, the interest rate \( r \), the tax rate \( t \), and (a vector of) government expenditures \( G \); credit constraints \( c \), itself a function of market and policy variables; other variables \( \xi_b \): an individual’s wealth, his liquidity, and his risk perceptions, and like \( c \), these can be a function of both market and policy variables; and his expectations of those variables, \( E \). For instance, if individuals own land, an increase in the interest rate, if it is expected to continue into the future, will lower the value of land. If households hold their wealth in stocks and bonds, consumption smoothing is easy; if, as in many countries, much of household wealth is in the form of housing, households may suffer from illiquidity, so that even if they would like to smooth their consumption over time by dissaving, it may be difficult for them to do so (except possibly by worsening liquidity in future dates). Expectations are complex—especially in idiosyncratic environments (situations like the Great Depression and Great Recession which have not occurred before, at least in the context of an economy that is very similar to the current economy), and it is always difficult to ascertain whether one is in such an environment. Thus, in a deep downturn, like the Great Recession, it is hard even to know what one would mean by ‘rational expectations’. One might, for instance, reasonably assume that if individuals save, some of the savings will translate into consumption in a demand-constrained period, in which case, future wage incomes will increase, and that in turn will lead to higher future consumption today. In short, it is reasonable to think of expectations as being affected by current variables, but the magnitude (and in some cases even sign) of the effects will not always be

---

69 For the moment, we drop time superscripts.

70 It is not assumed that individuals sell all of the labour they wish to. This is important.

71 Because some individuals may be lending (say holding government bonds) and others borrowing, both rates are relevant. The standard DSGE model doesn’t recognize the distinction, and the fact that the rate at which banks lend is an endogenous variable.

72 For instance, if \( T \) is land holdings, \( q \) is the price of land, the value of individuals land wealth is \( qT \). \( q \) itself needs to be solved for as part of the general equilibrium, a function of all the policy and state variables.
clear. To the extent that there is a significant intertemporal substitution effect arising from changes in the real interest rate, one has to model the effect of the market and policy variables on inflationary expectations. In highly reduced form, where expectations themselves are taken to be endogenous, we can write: $C = C(m, P; \xi_h)$.

Similarly, the investment function can be related to the usual variables (with Tobin's $q$ being a function of market and policy variables), with two additions: investment is affected by firm balance sheets (including liquidity) and may be credit constrained. The credit constraint itself can be thought of as being a function of the regulatory variables $R$ as well as the behaviour of banks (or other lenders), itself a function of market and policy variables as well as the net worth, liquidity, and expectations of banks $\{\xi_b, E_b\}$. More formally, in the absence of credit constraints and the expectation of future credit constraints, investment is given by $I^f = (w, r, p, G, E, \xi_f)$, and so long as the firm's liquid assets $L \geq I^d - c$, investment is given by $I^d$. Hence, we can write: $I = \min \{I^d, L + c\}$, where $L$ itself may be a function of market and policy variables.

For simplicity, it is often useful to focus separately on the cases where the credit constraint is and is not binding. But this misses a key point: firm heterogeneity is important, and some firms may be credit constrained, others not.

**Aggregate supply**

Aggregate supply is a function of the same state, market, and policy variables, though it is natural to assume that firms' supply depends on wages relative to prices. The interest rate may be important because it affects the cost of working capital. Again, balance sheets, risk perceptions, liquidity, and expectations of all the relevant variables matter (unlike in the standard neoclassical model). Thus, $Y^s = Y^s\left(\frac{w}{p}, r, E_f, \xi_f\right)$.

**Finance**

$r_L$, the rate at which the firm borrows, depends on the T-bill rate, but there is an endogenously determined spread, affected by the state, environmental, and policy variables affecting banks and their expectations:

---

73 As an example of a complexity which is typically ignored, a lower interest rate today that leads to higher income today with more savings implies that (everything else being constant) wages in the future will be higher. In a standard putty-clay model, this means that firms will choose a more capital-intensive technology, implying that the increased level of investment is higher than we might have expected using a putty-putty model. Standard DSGE models simplify by assuming fully malleable capital. There are marked differences in dynamics (Cass and Stiglitz, 1969). While there is some *ex post* capital labour substitutability, clearly *ex post* substitutability is much less than *ex ante* substitutability. I have not seen a convincing argument that the economy behaves more like putty-putty than putty-clay.

74 As in Greenwald and Stiglitz (1993b).

75 As in Stiglitz and Weiss (1981).

76 Thus, if there were a liquid land market, the value of land (a function of its market price $q$) is an endogenous variable.

77 Thus, in the special cases of this general model focusing on this constraint, we begin with two classes of firms, one liquidity constrained, the other not. But even this misses the more general point: the fraction of firms that are liquidity constrained is an endogenous variable. By assuming a continuum of firms (with different balance sheets) we can endogenously solve for the fraction of firms that are constrained.
In particular, banks are a particular kind of firm which takes its capital, borrows additional funds, and with these resources makes, monitors, and enforces loans, with the difference between the lending rate $r_L$ and the rate at which it can get access to funds (here assumed to be $r$) generating profits; it will charge a lower lending rate if it thinks that there is less risk, if its balance sheets are in better condition, if reserve requirements or capital adequacy requirements are lower, etc.

**Period 1 equilibrium with full flexibility**

Thus, with full flexibility of wages and prices:

$$Y = Y^s \left( \frac{w}{p}, r_L, E_t, \xi_t \right) = Y^d$$  \hspace{1cm} (3)

where $Y$ is actual output. (2) can be substituted into (1) and (3) to obtain aggregate demand and supply as functions of the T-bill rate $r$—but now the financial position of banks, their ability and willingness to make loans, matters too.

With the capital stock fixed, the demand for labour depends on output, $Y$, and in equilibrium, this equals the supply, which is assumed fixed:78

$$N(Y) = L = L^*.$$  \hspace{1cm} (4)

Full employment output is determined by (4), and then, for any given set of policy variables, $P$, expectation (functions) and other (state) variables $\xi$, the level of real wages, $w/p$, must be such as to ensure that $Y^* = Y^s$ (note that $Y^s$ depends just on $w/p$). But given $w/p$, with $Y^* = Y^d$, equation (1) determines nominal wages and prices. With fixed nominal debt, as wages and prices rise together, aggregate demand increases. The real indebtedness of debtors decreases, and the net worth of creditors decreases. If the marginal propensity to consume (from an increase in wealth) of debtors is greater than that of creditors—which we assume—the increase in consumption by debtors more than offsets the decreased consumption of creditors.79

There can be output variability in this model only as a result of shocks to the production function. But a slight extension, allowing the labour supply to be a function of, say, the real wage $L = L(w/p)$ means that (2), (3), and (4) together solve for $w/p$, $r_L$, and $N$; changes in state variables $\xi$ (both of banks and non-bank firms) affect the endogenous variables $\{w/p, r_L, N, Y\}$, and using (1), we can solve for wages and prices separately.

---

78 Again, in the more general case, actual employment is the minimum of the amount of labour demanded and supplied. The barebones model only considers the case where the demand for labour is equal to or less than the supply.

79 In the representative agent model, these two effects cancel. There can be other effects, e.g. on expectations. If investors believe that the increased inflation is not matched by an increase in nominal interest rates going forward, then the expected real interest rate will be decreased, and this will stimulate investment. Similarly, an unexpected increase in the price level could lead to more uncertainty about the future, with ambiguous effects on consumption but probably adverse effects on investment (see Rothschild and Stiglitz (1972), Diamond and Stiglitz (1974)). If differences in the marginal propensities to consume are large, and the increases in prices are modest, the consumption effect is likely to dominate. The barebones model provides a framework within which one can discuss (and model) the significance of these various effects.
By assumption, there is no unemployment, but shifts in the aggregate supply curve (changes in, say, $\xi$ either of firms or banks) lead to changes in the employment level and output. Economic variability is not just about shocks to the production function. And now policy affects the equilibrium not just through standard Keynesian impacts on aggregate demand, but also through shifting the aggregate supply curve (through (2)). Monetary policy matters, but not just through r, but through regulatory variables, and operating both on demand and supply. What matters are not just intertemporal substitution effects (which may be weak), but impacts on net worth of firms and banks, and through credit constraints.

Specialized versions of the model focus on each of these, and on distribution effects. A simple two-class model (workers with life-cycle saving, and capitalists with long-term saving) can easily bring out these first-order distributional effects.

In short, this barebones model can provide insights into the cyclical fluctuations which would occur even in the absence of nominal wage and price rigidities.

Rigid nominal wages and prices

When there is a (nominal or real) rigidity, demand may not equal supply and then it is the short side of the market that prevails:

$$Y = \min\{Y^a, Y^d\}.$$  \hspace{1cm} (5)

With nominal wages and prices fixed at $\hat{w}$, $\hat{p}$, aggregate demand could either be greater or less than supply.

Assume it is less. Then from (5), $Y = Y^d = \hat{Y}$, and employment is then N($\hat{Y}$). Thus, a temporary decrease in the interest rate could, as we noted, lower consumption, because the distributive effect exceeds the substitution effect. If investment is credit constrained, and the value of, say, land used for collateral does not change (much) in response to an announced temporary change in interest rate (as one would expect), then investment would not change. A regulatory change that resulted in banks making more loans would relax the credit constraint, and be more effective in stimulating the economy and increasing employment.

Assume that there were some policies that could lower wages, without instantaneously lowering prices. In a demand-constrained economy, that would lower consumption, even if it would increase aggregate supply. The effects on investment would be ambiguous, since (a) it would make profitable older equipment that might not otherwise have been used; (b) it induces the use of more labour-intensive technology, and thus, for every unit of future increased expected output, it induces less investment; but (c) it makes investment overall more profitable. If the firm(s) are credit constrained, the effect on investment would be zero. In short, there is some presumption that aggregate demand and thus output and employment would be reduced. Even though wages are

---

80 Indeed, much of the recent macro-literature outside of the DSGE/RBC tradition can be seen as doing precisely that.

81 In an open economy, to the extent that the lower wages led to lower prices, demand for exports would increase. But even in an open economy, two-thirds of output is non-traded, so that the lowering of (real) wages of workers can significantly lower aggregate demand.
too high, changing wages alone would have an adverse effect—a standard result in the theory of the second best.

There would be long-run supply benefits once the economy was restored to equilibrium (or to a supply-constrained situation.) The gap between aggregate supply and demand might accelerate price adjustments. Whether one views the change as positive or negative then depends on the social discount rate and the pace of adjustment to the full equilibrium.

But even the price adjustments might not (quickly) restore the economy towards equilibrium; in the short run, matters might get worse. The increased indebtedness in real terms (with nominal contracts) lowers both consumption and investment. Again, demand-side effects may dominate.82

**Efficiency wage unemployment**

Efficiency wages put a constraint on the (real) wages that can be offered (the no-shirking constraint). Similar results obtain in bargaining models.

\[
\frac{w}{p} \geq \phi(N), \text{ with } \phi' > 0 \text{ and } \lim_{N \to 0} \phi = \infty. \tag{6}
\]

Assume the product market clears, and that the constraint binds. Then:

\[
\frac{w}{p} = \phi\left( N\left( Y^{ss}\left(\frac{w}{p}\right)\right) \right) \tag{7}
\]

yielding the equilibrium real wage \((w/p)^*\) and employment levels. Given the policy and state variables, we can solve separately for the market variables \(\{w, p\}\).

In this model (where we continue to assume aggregate demand equals aggregate supply), the effects of policy again arise out of the aggregate supply function—which translates back into a demand for labour: any change that increases the level of output (and thus the demand for labour) at any real wage leads to lower unemployment. Thus, a lowering of interest rates lowers the cost of working capital and thus increases aggregate supply—the willingness of firms to produce and hire labour—at each value of \(w/p\) (given all other variables, such as \(\xi\)), and that leads to a higher equilibrium real wage and a higher level of employment.

(iii) **Extending the model back in time**

The current state variables for households, firms, and banks are affected by what happened in previous periods—including, of course, policies undertaken in previous periods. A model precisely like the one just posited applies to time 0 (the period before the current one). We focus here first not on the decisions in period 0, but on the effects of shocks in period 0 *after* decisions are taken and on government policy responses. We

---

82 This is modelling the Fisher–Greenwald–Stiglitz debt deflation effect.
omit the equations describing period 0, focusing only on those linking events in 0 to state variables in period 1. There are two types of firms, those that produce with a one-period lag (denoted by subscript a), and those that produce with no lag but require capital goods (denoted by subscript b). For one type, what is important is working capital, for the other it is physical capital.\textsuperscript{83} For the first type,

\begin{equation}
W_{1,a} = p_1 Y_{1,a} - (w_{0,a} N_{0,a} - W_{0,a})(1 + r_{0,1}).
\end{equation} \hfill (8a)

The wealth (balance sheet) of a firm going at time 1 is total sales minus what it has to pay back to the bank, which is the difference between its expenditures on inputs (which in this aggregate model is just labour) minus its own wealth. In effect, in this simplified canonical model the firm takes all of its wealth and uses it for its own production.

For the second type, in this simple model, capital lives only one period.

\begin{equation}
W_{1,b} = p_1 Y_{1,b} - w_1 N_{1,b} - (W_{0,b} - K_I)(1 + r_{0,1}).
\end{equation} \hfill (8b)

Thus, current prices, past and current wages, the capital stock, previous period’s production and investment decisions, and last period’s interest rates all affect this period’s wealth.

Equations (8) thus explain how a shock at any date affects wealth and other state variables at later dates. If prices this period are low (e.g. because of a demand shock), wealth will be low. That will affect investment and production, that is, both supply and demand. The interest rate that matters in equations (8) (and in equations (1) and (2) above) will be affected by shocks to the banking system. In real terms, since debt contracts are fixed in nominal terms, banks are better off when the price level is lower than expected; worse off if the default rate is higher, e.g. because of an adverse shock to firms.

Thus, equations (8) give an additional mechanism through which government interventions, like an increase in G, affect the economy: an increase in G increases p, increasing firm wealth and decreasing firm bankruptcies, thus increasing bank wealth. This leads to an increase in firms’ willingness to invest and produce (i.e. a shift in the investment and production functions) and banks’ willingness to lend (the availability of credit, if they ration, and the terms at which they make credit available).

There are other linkages across time: an increase in investment in public capital increases demand for goods at period t but increases productive capacities at t + 1. But if the public capital goods (like roads) are complements to private capital goods (like railroads), then the increased investment in public capital goods today increases the expected return to private capital goods this period, and hence leads to more private investment at time t. The increased private and public capital stocks then reduce the employment needed to generate any level of production next period, with consequences that can easily be worked out.

One of the important insights of this approach is thus that aggregate demand and supply are intertwined. Shocks to the economy and changes in policy are likely to simultaneously affect both.

\textsuperscript{83} This formulation is important because one wants to have the aggregate supply curve depending on the interest rate and balance sheets, which it does in the working capital formulation; but one also wants to have real investment (‘I’) (and aggregate demand) also depending on interest rates and balance sheets.
(iv) Extending the model forward in time

Individuals and firms are always making intertemporal choices, which were in some sense hidden in equations (1) and (3). As we noted in the text, while their decisions have consequences over time (as the previous subsection illustrates), how they behave is an empirical matter, i.e. the extent to which their decisions at any date are based on rules of thumb versus as if they solved a more complex maximization problem, is a matter of some dispute.\(^{84}\) Assume households and firms are forward looking. For instance, in the case of households, we summarize the expected utility of wealth carried forward to the next period (the solution to a complex dynamic programming problem) by \(V(W_{h,2}; \ldots)\). Then households maximize:

\[
U(C_1) + \mathbb{E}\left[V(W_{h,1}(1+r) + (w_iN_i - C_i)(1+r);\ldots)\right]
\]

giving the Euler equation:

\[
U' = \mathbb{E}[V_w(1+r)]
\]

The heart of the analysis is not so much an enquiry into the effect of changes in \(r\), but of changes in uncertainty, with an increase in uncertainty (e.g. about future wages or employment) generating precautionary savings. A fuller analysis would focus more explicitly on (a) aggregating over diverse individuals, with wealth being redistributed as \(r\) changes; and (b) borrowing constraints and imperfections in capital markets, taking into account the large disparity between borrowing and lending rates. The latter means that for large fractions of individuals, consumption is driven by their budget constraint: \(C = wN + r_B B\), where \(r_B\) is the borrowing rate charged these individuals, which is very loosely linked with \(r\) if at all; and \(B\) is the borrowing limit; and for still others, \(C = wN\); because of the kink in the interest rate function, they neither want to be borrowers or lenders.

A similar analysis applies to firms, where again the focus of analysis is on risk and how decision variables affect future risk (including the risk of bankruptcy).

References


\(^{84}\) Indeed, with households in constant flux, with household formation and dissolution having first order effects on well-being, and with household formation and dissolution both being endogenous variables, affected by economic variables, it is clear that the standard model leaves out much that should be important in any real intertemporal maximization problem.


On the future of macroeconomics: a New Monetarist perspective

Randall Wright*

Abstract: This article argues that a pressing goal for macroeconomics is to incorporate financial considerations, but we need models with solid microfoundations. In particular, the use of assets in facilitating exchange, as well as different credit (or other financial) arrangements, should be outcomes of, not inputs to, theories. As a preview, I suggest that mainstream macro does a good job explaining many phenomena, and a financial crisis does not disprove such theory. But understanding crises requires better incorporating factors related to money, credit, banking, and liquidity. The approach called New Monetarist economics can help a lot in this regard.

Keywords: microfoundations, money, credit, banking, liquidity

JEL classification: A00, B00, E00

Bad Guy: ‘You’re interested in money aren’t you? The kind you can spend.’
Bennie: ‘Yes sir. Indirectly.’

Bring Me the Head of Alfredo Garcia, directed by Sam Peckinpah, 1974

I. Introduction

My mandate is to discuss the future of macroeconomics. Before opining on that, the plan is to start by focusing on the past and present. I first try to establish that we have made impressive progress over the last few decades. This progress is manifest in several key areas, including the study of business cycles during normal times (i.e. outside of crises), labour markets and unemployment, development/growth theory, and monetary economics. While that part of the discussion is not crucial for the rest, it is predicated on the idea that it is useful to evaluate our ‘successes’ before embarking on new directions to make improvements in areas some people consider our ‘failures’. As regards

* UW-Madison, FRB Minneapolis, FRB Chicago, and NBER, e-mail: randall.wright@wisc.edu

For input I thank Gadi Barlevy, Todd Keister, Ed Nosal, Victor Rios, Ben Lester, Mei Dong, Jonathan Chiu, Benoit Julien, Xun Gong, Alan Head, Steve Williamson, Chao He, Cyril Monnet, Fabrizio Mattesini, David Andolfatto, Narayana Kocherlakota, and Ricardo Reis. Two anonymous referees also made extremely useful comments, as did several participants in discussions surrounding this ‘Rebuilding of macroeconomic theory’ project. For research support I thank the Ray Zemon Chair in Liquid Assets at the Wisconsin School of Business. More than in a typical disclaimer, all of the above, plus the Federal Reserve Banks of Chicago and Minneapolis, are absolved of responsibility for opinions expressed below.
doi:10.1093/oxrep/grx044
© The Author 2018. Published by Oxford University Press.
For permissions please e-mail: journals.permissions@oup.com
these directions, the most pressing goal in mainstream macro is to incorporate financial considerations with solid microfoundations.

As a way of expressing my theme, I cannot improve on two comments received from, respectively, a participant in a presentation and a referee. First, for many of the issues at hand, a key factor is the notion of liquidity, which has traditionally been neglected (if not outright abused) in both macro and microeconomics. Second, the use of money and other assets in facilitating exchange, as well as different kinds of credit and other financial arrangements (or, more generally, contracts), should be outcomes of, not inputs to, our theories. On both dimensions, I think current models used in policy analyses, including those that try to incorporate financial frictions, are lacking.

I will towards the end sketch a relatively simple version of the kind of model I think may be useful, although that can actually be skipped by those more interested in methodology than equations. By way of preview, let me try to summarize three big-picture points I try to flesh out in what follows.

1. Macro theories that are more or less mainstream do a good job helping us understand real phenomena, such as unemployment, growth, and cycles during normal times, as well as monetary phenomena, like the impact of inflation. The financial crisis does not disprove these theories.
2. Understanding crises at more than a superficial level requires generalizing our theories to better incorporate financial factors related to money, credit, banking, and liquidity, but solid microfoundations are critical.
3. There is a mature body of work called New Monetarist economics, discussed below, that can help in this regard.

II. Where are we?

Whether we need new approaches to macro and monetary economics depends on the issues at hand. Macro was at some point identified by the study of business cycles, defined as recurrent, regular, co-movements in aggregate time series (e.g. output, employment, consumption, and investment) at certain frequencies (e.g. filtering out slowly moving trends). Now if the issue at hand is understanding these fluctuations during normal times then we have made huge progress. Define normal times? Let’s say the period between the Great Depression and the Great Recession, although it may be better to start after the Second World War, or even later, to avoid issues such as price controls, quantity rationing, etc. Indeed, even the shorter period is not homogeneous, containing both the Great Moderation and Stagflation. Still, to be scientific and systematic about the data, as opposed to telling stories about each observation after the fact, I think it is useful to consider some notion of normal times.1

Define success? For RBC (real business cycle) theory, take the bare-bones neoclassical growth model and hit it with impulses that capture in an arguably reasonable way the slow, persistent, and random nature of technical change. The output of the exercise is a set of stochastic processes for the variables of interest. These resemble the

1 As I read Reinhart and Rogoff (2009), crises may well be recurrent at low frequencies, but I still think it is useful to distinguish them from regular or normal business cycles.
data qualitatively and quantitatively—output fluctuates more than consumption but less than investment, the series cohere well, etc. (Cooley (1995) is a standard reference.) This was a surprise to most economists in the 1980s. In retrospect, it should not be controversial. At the quarterly frequency, appropriately filtered aggregate consumption fluctuates about 1 per cent a quarter. How could one expect anything less, given the random nature of technical progress, without even mentioning changes in government spending, taxation, monetary policy, weather, energy prices, and so on? And how could one think that a 1 percent standard deviation in consumption is a big deal? So the business cycle defined in this traditional, albeit narrow, sense is: (a) not a puzzle; and (b) not a problem.

That is different from a view of cycles as market failures virtually by definition. It is now obvious that even if the real world were like Debreu’s (1959) ideal world—and note the use of the subjunctive—fluctuations would emerge from changes in technology, policy, weather, etc. Hence, cycles are not prima facie evidence of inefficiency or a need for intervention, and formalizing, clarifying, and quantifying this was a big contribution. This is relevant because, before moving in new directions, we should know why. It is wrong to say the need arises from a failure to produce a benchmark model of business cycles in normal times, even if such a need may arise from other issues. Yet that’s not the word on the street, or at least on some streets. I am hearing people say we need new theories because those developed by Lucas, Kydland, Prescott, and the rest failed us by neither predicting nor curing the financial crisis. That’s like saying polio research failed us because it didn’t predict or cure AIDS.

Rather than abandon RBC theory, we should teach it like we teach basic calculus to maths students, as received wisdom to be mastered before heading toward the frontier. Yes, there are open issues, including the idea that business cycles affect some people more than others, and the relative impact of different impulses. One does not have to think that only technology shocks are relevant to accept the usefulness of the framework—it can accommodate changes in policy, demographics, household economics, and much more. We may well want to know which shocks account for which aspects of the data, but that constitutes normal science. Moreover, developing the methods of the approach—how we look at data, how we solve models, how we compare them, and the very idea of using computable dynamic GE (general equilibrium) theory in macro—is a big contribution, especially because these methods can also be brought to

---

2 To see how this was controversial, consider the debate between Prescott (1986a,b) and Summers (1986) over the usefulness of the RBC framework. Or consider the remarks that I recall (although it could be my imagination) Rogoff made in discussing the first presentation of Kydland and Prescott (1982) at the NBER. To paraphrase what I think he said: Any major scientific breakthrough has two properties: at first it seems wildly implausible; after careful consideration, it seems completely reasonable. This paper satisfies the first criterion.

3 At the request of a referee, I can clarify what I mean by Debreu’s ideal world. I mean an abstract specification of a set of agents; their preferences; their resources (including property rights or endowments); and the technologies at their disposal. In such an environment one can introduce the notion of competitive equilibrium: taking prices as given, every agent optimizes; and prices are such that when they optimize all markets clear (i.e. the outcome is feasible). With some mathematics, under standard assumptions, it can be shown the concept is not vacuous—equilibrium always exists—and it is efficient in a well-defined sense. This baseline model is typically described as frictionless. Although I understand that it is important for many issues to add frictions, as discussed below, I also believe a good economist ought to first know well the frictionless benchmark.
bear on non-normal times (e.g. Cole and Ohanian (2004) on the Great Depression). To this extent the economics and associated technical developments have been very successful.4

There has been significant progress in other areas, too. Prior to the work of Diamond, Mortensen, Pissarides, and others, unemployment was a thorny issue, and even for those content to say wages can be too high or too low to clear the market, the coexistence of unemployment and vacancies was problematic. It is not problematic for search theorists, and incorporating their ideas into mainstream macro is another success. This is not to say that we know everything—still somewhat of a puzzle is Shimer’s (2005) observation that, for reasonable technology shocks, unemployment in the baseline Mortensen–Pissarides model fluctuates much less than in the data, at least if we dismiss the alternative calibration in Hagedorn and Manovskii (2008).

This is interesting, and reminiscent of getting the labour market in RBC theory to better match the data.5 There are other outstanding problems—e.g. it is still not easy to get capital, risk aversion, and other features of standard macro into the Pissarides (2000) model (but see Krussel et al., 2010). In any case, research on frictional labour markets is a vibrant area, and while we obviously welcome new applications and extensions, I see no need here for a scientific revolution . . . or, maybe it has already happened. In any case, I will not say more about search and unemployment here, since extensive surveys can be found in Rogerson et al. (2005) and Wright et al. (2017); I will have more to say about search generally in what follows.

Another area where we made progress is the study of the determinants of growth and development. Growth theory used to be perhaps good at explaining something, just not growth: standard models either converge to a steady state or a balanced growth path driven by exogenous technical progress. I am not an expert in this area, but I think the endogenous growth theory changed the situation for the better (leading examples include Romer (1986) and Lucas (1988)). There are very important outstanding issues, especially in understanding huge income differences between countries, or maybe understanding better why capital, technology, and production do not flow more easily across borders. Still, there has been much progress, and we now know that growth is potentially far more important than reductions in normal cyclical fluctuations.

To be clear about this, I mean that higher growth is much more important than reductions in the standard deviation of output during normal times; the most recent crisis

4 Several commentators on earlier drafts questioned why I want to ‘celebrate’ RBC theory. As Steve Williamson put it, ‘What’s the conclusion? Apparently that in normal times there isn’t anything important for policymakers to be doing, other than looking after long-run stuff—how big should the government be and how should they collect taxes?’ I actually like that idea—it sounds a lot like parts of Friedman’s (1948) prescription. In any case, let’s not get bogged down, since agreement on the success of RBC theory does not have a lot to do with what follows. In fact, presenting this position was motivated mainly by the way Ricardo Reis is organizing his discussion of similar topics, which involves asking what we should be teaching current graduate students.

5 That can be accomplished in RBC models in various ways, including the addition of non-convexities (Hansen, 1985), home production (Benhabib et al., 1991), or government spending (Christiano and Eichenbaum, 1992). Note that these generalizations all make good economic sense, and are not ad hoc fixes for poor predictions, like those in classical astronomy (as this may come up again, let me say now that I know very little about astronomy, but even apocryphal case studies can be rhetorically useful).
may well have had big effects, but that was not your standard business-cycle contraction, and had more to do with the financial system. I will spend more time on financial issues in what follows, but to wrap up on growth and development, let me say this: We may not yet fully understand why some countries are richer or grow faster than others, but at least we have a set of tools available to describe the situation and evaluate the contributions of various factors (for example, see Parente and Prescott (2000)). This seems like obvious movement in the right direction.

There have also been advances in monetary economics. Not that long ago, the best one could do was impose a CIA (cash-in-advance) constraint saying that agents must trade their endowments for money, then use the proceeds to buy consumption next period. Among many other reasons for finding this unsatisfactory, its first-order prediction is that using money makes us worse off compared to barter, which is really kind of embarrassing. It also forces one to argue it is fine to study monetary policy in models that do not explain what money is or what it does. Worse, some people simply label one argument of utility or production functions money, a practice called ‘reprehensible’ by none other than Tobin (1980).6

Modern theory, an example of which is contained in section V, explains at a level that is deeper (at least deeper than imposing CIA or having currency in utility) what the institution of monetary exchange does and how it is affected by policy. And it matters.

III. What do we need?

At the suggestion of a referee, I start this part of the discussion with the punchline: we need models with microfounded financial frictions. But, based on the situation as sketched above, we do not need a macro revolution, even if it gets the blood flowing to think we might. Instead, we need to focus more attention on certain issues, especially financial issues. However, before elaborating on that, it is useful to first discuss other common modelling devices. Consider nominal rigidities, which feature prominently in virtually all policy-oriented analysis. The Keynesian approach without apology adheres to price stickiness for each and every application, until ‘flexible-price equilibrium’ has become, alas, a retronym, like ‘cloth diapers’ or ‘undecaffeinated coffee’. Imposing sticky nominal prices may help one more quickly move to data and policy conversations, but we should want to know how far we can go explaining the world without such a crutch, and usually we can go pretty far.

6 An alternative that, I think, people attribute to Woodford’s (2003) huge contribution is to ignore monetary considerations entirely. That’s not always bad—it may (or may not) be fine to use Kyland–Prescott for normal fluctuations, Mortensen–Pissarides for labour, and Lucas–Romer for growth, and these are all non-monetary theories. For other issues, I think we need to bring back monetary considerations, broadly defined, as discussed below.
Moreover, even if prices are sticky, it matters why they are sticky. To make this point, let me describe briefly a model of endogenous nominal rigidities with policy implications very different from those that simply impose stickiness as an axiom. In the 2011 Marshall Lecture, published as Head et al. (2012), my co-authors and I propose a model of a monetary economy that delivers nominal price stickiness as an outcome of the search process. That formulation fits very well the standard micro data on price-change behaviour that people in the area claim to be important, yet by design, to make a point, it obeys monetary neutrality in the classical comparative-static sense: the allocation is invariant to a one-time change in the supply of currency. Of course that does not mean monetary policy is irrelevant (e.g. inflation matters).

The result uses a standard micro specification of search behaviour (Burdett and Judd, 1983) that we embed in a dynamic GE monetary framework. In this setting, sellers are literally indifferent between real prices in a certain range, because as the price goes down a seller earns less real profit per unit, but makes it up on the volume. This is not a knife-edge special case, but the endogenous outcome of natural search and information frictions. The model is consistent qualitatively and quantitatively with the observation that some sellers may not adjust their nominal prices after changes in the money supply (or changes in other variables). But monetary policy cannot exploit this, the way it can exploit exogenous nominal rigidities. This simple example makes it transparent that it matters a lot why prices are sticky in the real world. Burdett and Menzio (2016), Liu et al. (2016), and Wang (2016) push this idea further theoretically and empirically.

My bigger goal here is not to stick it to sticky prices, however, it is to contemplate whether recent experience necessitates rejection of standard macro and, if so, rejection in favour of what? I think we need something, just not quite a revolution, and many of the elements that ought to be incorporated are already available, as discussed below. Before getting to that, to add yet more perspective, consider a case where we did need a revolution: Stagflation in the 1970s. The current situation is different, because high-and-rising inflation and unemployment was not a side issue for the then-dominant theory, as it would be a side issue for RBC or search theory. Those theories are not about the Phillips curve, the same way that they are not about marriage or fertility (although extensions of the models can be brought to bear on those issues). In contrast, an exploitable, long-run, negative relationship between inflation and unemployment was a key assumption and the main driver of policy advice in Keynesian economics, and stagflation strikes at its heart. In contrast, the financial crisis does not disprove RBC, search, growth, GE, or game theory.

To further pursue this brief history of thought, when Friedman (1968) pointed out that it should be unexpected inflation that matters, proponents of the Phillips curve approach adapted. Then Lucas (1972), Sargent and Wallace (1976), and others worked out the idea that expectations should not be arbitrary in macro, but should have some connection to actuality. At this point it became difficult to persist in the claim that policy can regularly exploit the Phillips relationship, leading to the so-called Rational Expectations Revolution. It was a revolution in the sense that the crucial building blocks of received wisdom, as articulated by Samuelson and Solow (1960), among others, were overturned. But it is important to recognize that a paradigm shift is only feasible when there is an alternative. At the time the alternative was based on developments in micro theory, mathematics, and (perhaps slightly later) computation.
Lucas (1980), in the spirit of Kuhn (1962), describes this as the evolution of events and ideas. For a new paradigm to emerge there must be: 

(a) core failure of the status quo; and 

(b) a compelling new path.

New Classical Macro was compelling, and attracted many bright, open-minded, young economists who contributed substantially to its development, while others, not surprisingly, defended their old intellectual turf. The new approach was quite technical. These observations make me think we ought to maintain the rigour and discipline of New Classical Macro, or GE theory more broadly, and not shy away from mathematics when it is helpful.

Given that, if the recent crisis was a financial crisis, and there seems a consensus that it was, we should work to understand it better, but that does not render other parts of macro obsolete—we should not be throwing out these babies with the bath water. There are no financial factors in standard RBC, labour, or growth models, although there are attempts to introduce them, with varying degrees of success, but there is nothing like a canonical financial-macro model with accepted microfoundations, say, the way versions of Mortensen and Pissarides (1994) or Burdett and Mortensen (1998) are accepted benchmarks in good macro and micro labour economics.

The closest thing to it may be credit models along the lines of Kiyotaki and Moore (1997), but I see here a need for better foundations, and for making contact with existing research that strives for better foundations. This situation is similar for work following Holmstrom and Tirole (2011) or Diamond and Dybvig (1983), which may contain brilliant ideas, but for macro they not very useful, because they are usually cast in essentially-static (two- or three-period) partial-equilibrium models, while we need dynamic GE models. Why GE? Well, Diamond and Dybvig provide an interesting way to think about some aspects of banking, but what does it imply for labour, housing, or asset markets? Would it be a big deal if you can’t get your funds until after a brief banking holiday? I want to know.

To find out we need to integrate banking into dynamic GE theory. Now Gertler et al. (2015) and references therein try to do this, but those are models with banks as opposed to models of banks. The difference is that banking is imposed as a primitive—by assumption, households cannot lend directly to firms. I want models where banking arises endogenously, e.g. Gu et al. (2013a) and Huang (2016), who, among other things,

---

7 This view of rise of New Classical Macro in the 1970s, due to Stagflation and the advance of modelling techniques, is discussed in Lucas and Sargent (1979), although such history of thought is not uncontroversial (see Wren-Lewis, 2014a,b).

8 This is not surprising: given that it takes a fair bit of mathematics to capture routine physical phenomena, like the orbits of planets (more astronomy), how can it be that shifting a couple of curves allows one to understand much more complex phenomena like a modern economy? I mention this mainly because I am hearing a backlash against mathematical technique and formalism from a few people these days, and not just undergraduates, as in the old days, who seemed to adopt the position mostly out of laziness.

9 Even ignoring the need for dynamic GE, there are issues. I have heard it stated that ‘the financial crisis is easy to understand—it’s a bank run, like Diamond and Dybvig’. But in a typical Diamond–Dybvig model, runs only happen if one assumes banks behave in a particular ad hoc way, issuing simple deposit contracts, when better contracts are feasible. With endogenous contracts, the problems associated with runs can disappear (e.g. Green and Lin, 2003; Andolfatto et al., 2016). Models with ad hoc restrictions on contracts not only fail to help us understand runs or crises, they also provide a weak foundation for analysing government policy, a point made by Wallace (1988). It is important to dig deeper, asking what underlying frictions could potentially generate a run under optimal contracting. See Ennis and Keister (2009, 2010) for an attempt at this based on imperfect commitment.
ask why we need banks in the first place, determine who should act as bankers, and ana-
lyse the trade-off between many small or few large banks, all of which seems relevant
for regulatory reform. They also have bank liabilities—claims on deposits—facilitating
third-party transactions, like notes, cheques, and debit cards throughout history, some-
thing conspicuously absent from most models of banking.

Moving beyond the narrow topic of banking, there are financial macro models, as
evidenced by many of the chapters in the recent *Handbook of Macroeconomics* (Taylor
and Uhlig, 2015). But that research goes only part way, and not very far, to where we
need to be. Moreover, the work is typically so divorced from what I do that it is not clear
how to compare the approaches. I guess the authors feel the same, given there are virtu-
ally no citations in common with the surveys by Lagos *et al.* (2017) and Rocheteau and
Nosal (2017), which are recent, but they have in various forms been around for quite
a while, as have other surveys. Wouldn’t it be better if these literatures communicated?
While I applaud efforts to bring financial considerations into macro, I am disappointed
in the foundations of most policy-related research.

Let me clarify something extremely important: general equilibrium to me does
*not* merely mean multiple optimizing agents, it also means institutions like monetary
exchange and credit arrangements should be results and not assumptions. Relatedly,
contracts should not be constrained by modellers’ whims, even when—or perhaps espe-
cially when—appeals are made to realism. If people in the real world sometimes seem
to trade only non-state-contingent debt, or banks seem to issue only simple deposit
contracts, this should be *explained* and not *imposed*. Laying out explicit features of an
environment that deliver such an outcome as an equilibrium is a way to explain it, but
even that is delicate, because one has latitude in terms of the equilibrium concept. A
purer approach is to show the environment delivers the outcome as an efficient arrange-
ments, i.e. one with no gains from trade sitting on the table.

Good economists know that behaviour and interactions are endogenous, while the
environment, which may include frictions such as spatial or temporal separation, imper-
fect information, or limited commitment, is exogenous. It is also crucial to understand
that a mechanism is logically distinct from an environment. It is fine for Debreu’s agents
to, in some sense, ‘go along with’ competitive markets, since they work well in their
situation—one way to express this involves the connection between equilibrium and
the core. When they work poorly, institutions in model economies should get thrown
out and replaced by something better, subject to the explicit frictions, of course. In too
many models there are gains from trade sitting right there on the table—the outcomes
are not even in the bilateral core, which can lead to ridiculous implications.10

Now here’s a question: if labour economists are responsible for studying unemploy-
ment, business-cycle theorists for studying fluctuations, etc., who is responsible for devel-
oping models of financial considerations? I say monetary economists. To help make the

---

10 I recall (I think) Larry Christiano’s example from an ECB conference. With a fixed nominal wage,
after an increase in nominal prices, the boss calls a worker at 3:00 am yelling, ‘Wake up and get to work!’,
which the worker is obliged to do by assumption in many models. If the boss and worker just chatted about it
for a minute they could both be much better off. Imposing that they can’t sort this out is, to me, and I presume
to Larry, bad economics. To say it slightly differently, a good economist typically will not predict an outcome
will obtain when an individual in the model has an obviously profitable unilateral deviation; I am convinced
that we sometimes need to worry about profitable bilateral (or multilateral) deviation, too.
case, first, let me consider an alternative possibility: financial economists. I try/claim to be a finance professor, as well as a monetary economist, and in my experience, limited as it may be, I see those in the former group interested mainly in an asset's price, and to some extent its liquidity defined loosely and typically poorly. In pursuing this, we take a lot for granted, and often use essentially-static partial-equilibrium models. In contrast, many in the latter group, especially the relatively pure monetary economists, are interested in similar phenomena in the context of dynamic GE models, where institutions, including transactions patterns and financial arrangements, are endogenous.

When I tried to argue in an earlier draft that we need microfounded theories of financial considerations, Narayana Kocherlakota pushed me by asking, which financial considerations? I am happy to consider anything at play in the literature on imperfections in credit, banking, and contracting, although only in a dynamic GE context. To motivate or provide some guidance on how we should go about it, consider this: standard finance (or, for that matter, standard GE) theory predicts the price of fiat currency is 0, because that is the discounted present value of its stream of returns. That is rather flatly rejected by the evidence. Given standard finance cannot correctly price the simplest of all assets—a dollar bill—and given the reason—it does not correctly capture the dollar’s liquidity—it does not seem to me the best approach. Now I am not saying that pricing currency is a sufficient condition for good financial macro; it might be more like a necessary condition.

I therefore think that those interested in financial issues should be informed about developments in monetary economics. By monetary economics I do not only mean the study of how fiat currency can be valued, or how different objects may emerge as media of exchange in equilibrium or efficient arrangements. While those were long-standing issues in need formalization at some point, monetary economics these days is the study of all things related to money, credit, and banking, or even more generally to liquidity. The view that we should be doing ‘liquidity economics’ was pushed a few years ago by John Moore in his plenary talk to the Stockholm meetings of the Society for Economic Dynamics, but I have my own take on it. I present this with some more general discussion in section IV, and with equations in section V.

### IV. Where do we go?

To reiterate my central tenet, money, credit arrangements, and intermediation should be outcomes of models, not inputs into models. There is an approach aspiring to this goal going by the name New Monetarist Economics. Work in the area typically eschews short cuts such as CIA constraints or money in utility/production functions,

11 Maybe data will ultimately decide what friction is most relevant—the key to the crisis?—although from what I remember about the philosophy of science, it is naïve to expect empirics will carry the day by itself.

12 More discussion of methodology and in particular of this label can be found in Williamson and Wright (2010a,b). In the interest of space here I will be brief, but this research lies at the interface between macro and micro—it is meant to be empirically and policy relevant, but also strives for theoretical rigour and consistency, using tools from search, game, bargaining, and mechanism design theory. While most research wants to be rigorous and consistent, New Monetarism is particularly concerned with foundations, although the field has become more policy oriented as the theories have matured and the crisis focused attention on liquidity (again, this is the evolution of events and ideas).
while in other camps these practices are not only considered acceptable, they are flourishing, and branching out, with T-bills, demand deposits, and who knows what else showing up as primitives, justified by appeal to ‘convenience yield’. While we learn something from this work, e.g. the empirical findings of Krishnamurthy and Vissing-Jorgensen (2012), is it not obvious that such reduced-form modelling has flaws?

Explaining this in terms of the Lucas (1976) critique usually falls on deaf ears, for reasons I do not fully understand, but it is incumbent upon me to mention that Lucas has a point. See also Townsend (1987, 1998), and, of course, see Wallace (1998), who proposes this dictum: ‘Money should not be a primitive in monetary theory—in the same way that firm should not be a primitive in industrial organization theory or bond a primitive in finance theory.’ I obviously agree. My preferred endeavour is to explain why assets, including currency, bonds, bank deposits, and so on, show up in payoff functions even though they do not enhance utility or productivity directly. I find it no harder, and more satisfying, but evidently not everyone agrees. Perhaps they do not know the tricks of the trade, and hence think it too challenging. Many papers in this literature use search theory, which seems like the right tool for the job, but could be a foreign language to some macroeconomists, despite its popularity in labour, housing, and other subfields.

Let me say why search theory is useful here. To understand monetary exchange, credit arrangements, and related institutions, they should be endogenous, not primitives. Firms and households, by analogy, are primitives for Debreu, and he makes good progress with that, but his Theory of Value is a model with, and not of, firms or households. The institutions I am interested in have something in common: they exist to facilitate exchange, often intertemporal exchange. Exchange does not need facilitation in Debreu’s world, so the relevant institutions (money, banks, etc.) do not appear in his theory. To achieve a proper understanding of institutions whose raison d’être is mitigating frictions in the exchange process, we obviously need to study environments where exchange is non-trivial. To me, this means environments where agents trade with each other, as opposed to either Debreu’s general- or Marshall’s partial-equilibrium approach, where they trade only with their budget lines. I declare here that search theory is the study of agents trading with each other.¹³

When agents trade with each other, frictions make the process interesting. In addition to the fact that it may take time to find a counterparty, the frictions that hinder intertemporal exchange are limited commitment and imperfect information. The former means we cannot support an ideal credit system simply by committing to honour obligations; the latter means we cannot support it by punishing those who default by taking away future credit, as in Kehoe and Levine (1993), although they actually work in GE and not search theory. As an extreme case, suppose agents are completely

¹³ An early example of search in this sense—equilibrium theory where agents trade with each other—is Diamond (1982); previously it was mainly decision theory where agents trade against (sample from) exogenous distributions of prices, wages, houses, or spouses. Also, I suppose it sounds imperialistic to simply declare search as the study of agents trading with each other, given this is also true in, for example, contract theory. But that integrates well with search, and indeed is enhanced by it, since it captures the idea that you must make contact with someone before you can contract with someone.
anonymous, whence we cannot punish defaulters at all, because we do not know who they are (this is why one does not make loans to complete strangers). Let’s pursue this extreme case for the sake of discussion here, and in a formal model in section V.

Combined with a specification for specialization that means gains from trade cannot be exhausted using direct barter, these frictions make the institution of monetary exchange essential, defined as society being better off by having something that can be used as a medium of exchange (Kocherlakota, 1998; Wallace, 2010). Sometimes the objects that could or should play this role are endogenous (see Kiyotaki and Wright (1989) for a simple example); other times there is a unique candidate for the role and we study how it works (see Kiyotaki and Wright (1991) for a simpler example). If the candidate is fiat currency, meaning intrinsically worthless and unbacked paper, the theory explains how it can be valued for its liquidity—you give me something for it because you think somebody else will give you something for it, and so on. In this way assets can be endogenously valued for their liquidity.

Thus liquidity gives currency a price above its fundamental value of 0 in monetary equilibrium. But wait—if currency can be valued in excess of its fundamental due to liquidity, is that also true of other assets? Yes. Any asset that conveys liquidity can have a ‘moneyness’ about it, which can lead to bubbles, booms, crashes, market freezes, and so on, due to self-fulfilling prophecies (see, for example, Rocheteau and Wright (2013) for details, but a sketch of the idea appears in section V). Thus phenomena like excess volatility, fragility, and hysteresis, based on the self-referential nature of liquidity emerge naturally.14 Perhaps less well known is that similar phenomena arise in pure-credit economies with endogenous debt limits even with no liquid assets (Gu et al., 2013b). This all seems fairly relevant in light of the crisis.

I arrived at these conclusions by studying environments where agents trade with each other. Now one can always claim to get the same insights from reduced-form models, but that smacks of cheap talk. Also, to dispense with the obvious, one can always produce a reduced-form model that is observationally equivalent, with respect to some set of interventions, to a deeper model—just take the latter and derive its reduced-form (the value function). But that’s not so easy without having the original model and, in particular, how would one know the functional form without deriving it?

To me it is self evident that the best way to understand institutions that facilitate trade (money, banking, collateral, etc.) is to analyse exchange in the presence of explicit frictions. Yet I see many papers that, for example, assume A cannot lend to C except through B, then call B a bank, claim this is a theory of banking, and use it to discuss policy or regulation. I need to know why A cannot lend to C. Commitment issues? Private information? Spatial separation? Lay it out! I need to know for two reasons: (a) only then can we check if the assumptions really mean A cannot trade with C; and (b) only then can I see what other implications follow from the assumption. These ideas are

14 The idea that self-fulfilling prophecies, or animal spirits, can lead to complicated dynamics in monetary economies is not new; see Azariadis (1993) for a textbook presentation. But in more modern New Monetarist models, I would argue, the economic intuition is more appealing and the required parameter values are more reasonable. Also, the outcomes are more varied: Gu et al. (2017a,b) provide examples of market freezes and announcement effects that go well beyond what one sees in Azariadis (1993). Also, this has very little to do with fiat currency per se, and more to do with the general nature of liquidity.
not new—see Kareken and Wallace (1980)—but still have force. So why is it rare to see papers that start with explicit frictions and an endogenous role for the relevant institutions? Why is this research a boutique subfield and not bedrock for modern macro? It must be because this approach also has costs. Let’s discuss the costs, and in the process dispel some misinformation.

Those who advocate reduced-form methods often hang their hat on the ability to fit data and thus appeal to policy relevance. That is germane only if New Monetarists were not interested in, or their theories were not conducive to, empirical work. While much their research is indeed theoretical, because many of the issues are conceptual, there are also significant quantitative contributions. I already mentioned work on the cost of inflation (compared to reduced-form models it can be big). Another instance is Berentsen et al. (2011), who integrate microfounded monetary theory with the Pissarides (2000) model of labour to ask how much monetary policy contributes to the behaviour of unemployment (it contributes a sizable fraction). Another is Lagos (2010), who asks how well a liquidity-based model accounts for the equity premium and other puzzles in finance (it does quite well).

Yet another important example is given by Aruoba and Schorfheide (2011), who run a horse race, at the risk of oversimplifying, between New Keynesian and New Monetarist frictions using state-of-the-art econometrics (as I understand it, we win). High-level computational exercises can be found in Molico (2006) or Chiu and Molico (2010, 2011). Even more impressive is recent work by Chiu and Molico (2014) and by Jin and Zhu (2014) that goes beyond stationary equilibrium to study transitions in the distribution of liquidity across agents following monetary shocks. I already mentioned attempts to fit the micro price-change data. Based on these papers, and others in the literature, it is hard to support the allegation that those who stand for microfoundations stand against quantitative economics. Moreover, some of the substantive results in these exercises are rather impressive.

Consider Jin and Zhu (2014). In some of their experiments, a money injection spreads out the distribution of liquidity, effectively making the rich richer and the poor poorer. The rich then work less, but they weren’t working much anyway, while the poor work more, and average output increases. Prices are not sticky—indeed, every trade is executed at terms determined by bilateral bargaining—yet the following is true: while buyers spend more money after the injection, sellers produce more, due to the wealth effect, and hence nominal prices (ratios of dollars to goods

---

15 Here is an example from Williamson and Wright (1994), where there is private information à la Akerlof: producers can bring lemons to the market. So consumers cannot trade directly with producers, right? Not necessarily. There are some informed consumers who never trade for lemons. So, if uninformed consumers do not trade with producers, those with lemons cannot trade at all. Hence, for some parameters, producers bring high quality to market with positive probability, sometimes even probability 1, and the uninformed trade with them. That set-up is crude, but updated incarnations are in Lester et al. (2012) or Li et al. (2012). The point here is that it is not trivial to explicitly describe assumptions that preclude certain trades; also, when one does find the right assumptions, logical consistency dictates that we follow through with all their implications, not just the ones we like.

16 This depends on the details of policy and on parameters, and the effects can generally go either way. In Molico (2006), for example, money injections contract the distribution of liquidity. This should be no surprise: when one takes the market micro structure seriously, of course the effects depend on details of the specification.
changing hands) do not go up on average. In fact, prices change slowly while output rises after a money injection until the distribution settles back down. This is consistent with what looks like neutrality in the long run but not in the short run. However, output does not go up in the short run because prices are sticky; to the contrary, prices appear sticky because output goes up. This certainly should make a few economists rethink their evidence. 17

Another misconception is that New Monetarists restrict attention to models with a 0 wealth effect, or with degenerate distributions of asset holdings, which is related because a degenerate distribution sometimes emerges in certain settings from a 0 wealth effect (see section V). In truth, sometimes we employ preferences with 0 wealth effects, precisely to harness the distribution of asset holdings, as a tractable benchmark, but this is obviously not the case in general as exemplified by all the computational work just mentioned (see also Menzio et al. (2013) and Rocheteau et al. (2015a,b)). There are issues for which it is important to have models with non-degenerate distributions of liquidity, as Wallace (2014) emphasizes, but there are other issues for which this is less important, and good work strives to find the tool for the job.

Another cost of adopting the approach I advocate is that one has to learn a little terminology and technique from search theory. Why pay that cost if many relevant markets seem relatively frictionless? One answer goes like this: even markets where search per se is unimportant are usefully modelled as limiting special cases of frictional markets, with an advantage over the usual Walrasian abstraction, because even in the limit agents trade with each other. Search models have two properties I like: agents trade with each other; and it takes time to find a good counterparty. But one can minimize the latter while maintaining the former in order to concentrate more on frictions like limited commitment or private information.

Having said that, in many applications it is nice to incorporate search per se, because it captures both the intensive and extensive margins of trade. As a recent example, in Rocheteau et al. (2017) entrepreneurs with opportunities to invest must search for funding, as in Wasmer and Weill (2004), and when they find it they can borrow up to a pledgeability constraint, as in Holmstrom and Tirole (2011). This generates extensive and intensive margins of credit rationing, consistent with data. To be clear, the problem for an entrepreneur is not finding a bank—Google Maps solves that—but finding one willing to fund him (by analogy, in job search, the problem for an unemployed worker is not finding a producer, but finding one willing to hire him). Moreover, while search per se may be less important in some financial markets, it is very relevant in others, including over-the-counter asset

17 This bears on another issue, the misconception that New Monetarists ignore nominal rigidity. Jin and Zhu (2014) do not ignore it, they derive it—i.e. it is a result that nominal prices do not move much in the short run. See Rocheteau et al. (2016) for a less impressive but more analytically tractable example of prices looking sticky when they are not. In general, such examples are easy to construct. In Head et al. (2012), as discussed earlier, it is a result that some sellers choose not to adjust nominal prices after changes in the money supply or other variables, and those models fit the micro data very well. But for anyone wed to exogenous stickiness, for whatever reason, that can also be accommodated, as in Aruoba and Schorfheide (2011) (see also Williamson and Wright (2010a,b) for easier examples).
markets (Duffie et al., 2005) and venture capital markets (Silveira and Wright, 2016).\(^{18}\)

In combination with search, New Monetarists often use bargaining, although there are also lots of examples with Walrasian pricing, similar to the job search model of Lucas and Prescott (1974). This may be comforting to those wed to an auctioneer, but I prefer to consider other possibilities.\(^{19}\) In addition to bargaining, options include auctions and price posting with random or directed search, which are neither less tractable nor less realistic than Walrasian pricing. The framework is also amenable to mechanism design, as in Hu et al. (2009). And it matters for policy results, as discussed in, for example, Rocheteau and Wright (2005, 2009) and Aruoba et al. (2011). It also matters for pure theory when, for example, Gu et al. (2013b) show pure credit economies can be unstable with Walrasian pricing and with Nash bargaining, but the economic intuition is very different. In terms of policy, more generally, a standard analysis of optimal fiscal and monetary policy using a microfounded model in Aruoba and Chugh (2008) gives different qualitative and quantitative results than reduced-form models.

For those who want to study issues related to money, banking, and credit, if they were not trained in this tradition, there is a cost to learning about search, matching, arrival rates, bargaining, price posting, etc. This undoubtedly increases the time for the approach to diffuse into mainstream macro. The situation was similar in the economics of labour when Diamond, Mortensen, Pissarides, Burdett, and others were first developing their paradigm, yet it now constitutes the standard benchmark in serious research on unemployment, wage patterns, etc. I think that it is only a matter of time until a similar evolution occurs in the economics of liquidity.

V. An example

To make the paper self contained, following referees’ suggestions, let us now work through an example of a rudimentary New Monetarist model based on Lagos and Wright (2005).\(^{20}\) Recognizing a trade-off between tractability and applicability, we assume that in each period of discrete time a large set of agents interacts sequentially in two distinct markets. First, there is a decentralized market, or DM, where they trade bilaterally, as in search theory. Then there is a centralized market, or CM, as in GE theory. Different objects are traded in the two markets. In the CM agents trade a consumption good \(x\), labour \(l\), and assets, plus, in versions with credit, they settle their accounts. In the DM they trade something else, \(y\), which can be a good, a factor of production, or an asset.

\(^{18}\) Relatedly, search models have an easier time with velocity than CIA models (e.g. Hodrick et al., 1991). In particular, they are not critically dependent on period length: as we go from, say, a quarterly to a monthly calibration, we can scale arrival rates by 1/3 and get the same expected time until a dollar turns over. Lagos and Rocheteau (2005), Liu et al. (2011), and Nosal (2011) model the effect of inflation on velocity using standard elements of search theory, respectively, intensity, entry, and reservation trading decisions.

\(^{19}\) Other advocates of moving away from the auctioneer in macro in include Farmer (2005), but the idea is much older—e.g. Diamond (1984) says, ‘If there is to be a satisfactory micro based theory of money and macro problems, it seems likely that it must dispense with the fictitious Walrasian auctioneer.’

\(^{20}\) Readers less interested in equations can without loss in continuity skip all but the last few paragraphs, which pull back to reconsider the big picture.
Most of the early work on this model has households trading consumption goods in the DM. In some applications this is interpreted and calibrated as a retail sector (Berentsen et al., 2011; Bethune et al., 2016); in others it is interpreted and calibrated as an informal sector (Gomis-Porqueras et al., 2014; Aruoba, 2017). In other applications the agents in the DM are bankers trading reserves (Koeppel et al., 2008; Afonso and Lagos, 2015), innovators trading ideas (Silveira and Wright, 2010), producers trading capital (Aruoba et al., 2011), or investors trading financial assets (Lagos and Zhang, 2015, 2016). The framework is evidently flexible. Different applications also go into more or less detail about the CM; here its role is minimized.

This alternating CM–DM structure is a feature of the example, not the general framework, but for purposes here it has a nice property: we do not need to solve the model on a computer, since rather a lot can be shown analytically. While there’s naught wrong with numerical work, in general, it is good to have a benchmark that delivers general results by hand. Another advantage of the set-up is that at its core is an asynchronicity between expenditures and receipts that is absolutely central to any model of money or credit: as specified formally below, agents may want something in the DM but their income accrues in the CM. Given they are constrained in how much credit they can use, they therefore might choose to bring assets from the current period to use as payment instruments in the next period.

For this presentation let’s say households trade a consumption good \( y \) in the DM, different from \( x \), the numeraire in CM. Both goods are non-storable. Also, suppose \( x \) is produced one-for-one with \( l \), so the CM real wage is 1 (that is easy to relax). In this baseline specification, following older search-based monetary theory, suppose agents in the DM can be buyers or sellers depending on who they meet. One formulation (Aiyagari and Wallace, 1991) models this by saying there are \( K \) goods, and \( K \) types of agents, where type \( k \) consumes good \( k \) and produces good \( k + 1 \) (mod \( K \)). Thus, \( K > 2 \) precludes direct barter in any pair-wise DM meeting. To engender a role for assets in the facilitation of exchange as simply as possible, let’s rule out credit entirely by assuming agents cannot commit and are completely anonymous. Also, for now, the only asset is fiat money, \( m \) (but see below).

If two agents meet in the DM and one likes the output of the other, call the former the buyer and the latter the seller in the meeting. Period utility is \( U(x) - l + u(y) \) for buyers and \( U(x) - l - c(y) \) for sellers, where \( U(x) \), \( u(y) \), and \( c(y) \) have the usual properties. Tractability comes from these pay-offs being linear in \( l \), but again this is a feature of the example, not the framework. All agents have discount factor \( \beta \in (0, 1) \) between one CM and the next. Letting DM and CM value functions at \( t \) be \( V_t(\cdot) \) and \( W_t(\cdot) \), we have

---

21 In the latter case one might identify monetary assets as currency, since the informal sector is said to be very cash intensive; in the former case one might identify them as currency plus demand deposits, since in retail cash, cheques, and debit cards are quite similar, and all are different from credit cards (see, for example, Bethune et al., 2016). Of course to do this justice one might want to add banking and credit explicitly, which is interesting, but goes beyond the scope of this example.

22 As Hicks (1962) put it, ‘We must remember the purposes for which liquid assets are held: to pay existing debts, in so far as they cannot be paid out of new receipts, and to meet necessary new expenditures in so far as they cannot be financed in the same way.’ The formal model simply captures that venerable wisdom with modern dynamic equations.
where $m_t$ and $\hat{m}_t$ are money holdings when the CM at $t$ opens and closes, $\phi_t$ is the price of $m_t$ in terms of $x_t$, and $T$ is a government transfer (or a tax if $T < 0$). For simplicity here $T$ controls the aggregate money supply, $M_{t+1} = (1 + \pi)M_t$, where $\pi > 0$ ($< 0$) comes from injecting (withdrawing) cash. As is standard, assume $\pi > \beta - 1$.

Ignoring non-negativity constraints for now, and eliminating $l$ using the budget equation, we have

$$W_t(m_t) = \max_{x_t, l, \hat{m}_t} \left\{ U(x_t) - l, + \beta V_{t+1}\left(\hat{m}_t\right) \right\} \text{ s.t. } x_t = \phi_t \left( m_t - \hat{m}_t \right) + l + T,$$

Notice $W_t(m_t)$ is linear with slope $\phi_t$. Also, $\hat{m}_t$ is independent of $m_t$, and so the distribution of $\hat{m}_t$ is degenerate across agents entering the DM. This makes (1) about as easy as a two-period problem, since agents start afresh at the end of each CM. Still, the infinite horizon is critical for getting fiat currency valued and, in versions with debt, for getting repayment by threatening to take away agents’ future credit after default. This is not an essentially-static economy.

Bilateral random matching in the DM is summarized by letting $\alpha$ be the probability of meeting someone that produces your good. In equilibrium a buyer always gives up all his cash (at least assuming we are away from the zero lower bound $\iota = 0$, where $\iota$ is the nominal interest rate discussed below). Then, to keep the notation especially easy, let $c(y) = y$ and assume buyers make take-it-or-leave-it offers to sellers (for what it’s worth, in this specification that gives the same outcome as Walrasian pricing). Thus, in exchange for his cash, a buyer gets $y_{t+1} = \phi_{t+1} \hat{m}_t$, so his payment just offsets the seller’s cost. Then we have

$$V_{t+1}(\hat{m}_t) = W_{t+1}(\hat{m}_t) + \alpha \left[ u\left( \phi_{t+1} \hat{m}_t \right) - \phi_{t+1} \hat{m}_t \right].$$

In words, an agent can always get $W_{t+1}(\hat{m}_t)$ by keeping his cash for the CM, but with probability $\alpha$ he spends it, yielding a surplus given by the term in brackets. Symmetrically, an agent may also produce in exchange for someone else’s cash, but that can be ignored when buyers make take-it-or-leave-it offers.

Given (2), the first-order condition for $\hat{m}_t$ from (1) is

$$\phi_t = \beta V_{t+1}'\left(\hat{m}_t\right) = \beta \phi_{t+1} \left[ 1 + \alpha \left[ u\left( \phi_{t+1} \hat{m}_t \right) - 1 \right] \right].$$

Using money-market clearing, $\hat{m}_t = (1 + \pi)M_t$, and letting $z_t = \phi_t M_t$ be real balances, we reduce (3) to

$$z_t = \beta z_{t+1} + \phi_{t+1} \hat{m}_t.$$
where \( \lambda(y) \equiv u'(y) - 1 \) is often called the \textit{liquidity premium}. It is also the Lagrange multiplier on the constraint that a buyer cannot offer more to a seller than he has (more generally, that plus an amount constrained by his debt limit, which is 0 in this example). The difference equation (4) has two steady states, \( z = 0 \) and \( z > 0 \), plus, for some parameters, a variety of bounded dynamic solutions, and these are all equilibrium outcomes.

In some of these dynamic equilibria, \( z \to 0 \) is a self-fulfilling prophecy of hyperinflation; in others, \( z \) follows cyclic or chaotic paths; it can also follow stochastic paths (sunspot equilibria). These and other complicated outcomes are due to the self-referential nature of liquidity, and can be interpreted in terms of bubbles, excess volatility, booms, crashes, and freezes. One can ask what kinds of policies make these outcomes more or less likely, which is of course important, but this is not the place to go into these details. Instead, and importantly, let me emphasize that the same logic applies to real assets, not just fiat currency. Having equity in ‘Lucas tree’, as in standard asset-pricing theory, is like having currency, except that each unit held in the CM generates \( \rho \) units of numeraire, a standard dividend of ‘Lucas fruit’. However, the results are not generally like standard asset-pricing theory unless we shut down the DM by setting \( \alpha = 0 \).

The generalization of (3) for an asset in fixed supply \( A \) and for any \( \rho \) is

\[
\phi_t = \beta(\rho + \phi_{t+1}) \left[ 1 + \alpha \lambda(\rho \hat{m}_t + \phi_{t+1} \hat{m}_t) \right].
\]

(5)

Given \( \rho > 0 \), equilibria with \( \phi = 0 \) and those with \( \phi \to 0 \) vanish, but not necessarily the cyclic, chaotic, and stochastic equilibria, at least as long as \( \rho A \) is not too big. If \( \rho A \) is too big, then a buyer can get \( y^* \), where \( u'(y^*) = 1 \), without giving up all his assets. In this case liquidity is not scare and the unique equilibrium has \( \phi^* = \beta^* \equiv \rho / (1 - \beta) \). Notice \( \phi^* \) is the \textit{fundamental} price of the asset, and \( \phi_t = \phi^* \forall t \) is the unique equilibrium if \( \alpha = 0 \)—i.e. making liquidity abundant is like shutting down the DM in terms of asset prices, if not pay-offs.

Moreover, even with \( \rho < 0 \), as long as \( |\rho| \) is not too big, there are equilibria with \( \phi > 0 \). This shows how an asset with negative yield can be valued for its liquidity. There are also equilibria where these ‘toxic’ assets are valued at \( \phi_t = \phi > 0 \forall t \leq t_t \), then valued at \( \phi_t = \phi_t < \phi_t^* \forall t > t_t \), where \( \phi_t < \phi_t^* \), with \( \phi_t = 0 \) as a special case. This should remind one of a financial crisis. It looks even more like that if, instead of assets being used as payment instruments to finalize DM trade, they are used as collateral in support of credit settled in the next CM. With such credit arrangements, which can also be described in terms of repurchase agreements, the equations are exactly the same—they are simply reinterpretations of the model.\textsuperscript{25}

To consider a few extensions and applications, first let’s go beyond take-it-or-leave-it offers by supposing that in order to get \( y \) the buyer must transfer \( v(y) \) to the seller in

\textsuperscript{25} Many models of collateral assume borrowers can only pledge a fraction of their assets, but that can obviously be assumed here, too. Moreover, it can be microfounded in terms of commitment or information frictions (see the above-mentioned surveys for details concerning this and other claims).
terms of CM numeraire, where \( v(y) \) can be any increasing function with \( v(0) = 0 \). Then (5) still holds, except now \( \lambda(y) = u'(y)/v'(y) - 1 \). Corresponding to different specifications for \( v(y) \) are different ways to determine prices, including Walrasian pricing taking, Nash bargaining, and outcomes from mechanism design. Changing \( v(y) \) affects both the qualitative and quantitative results in interesting ways, and highlighting this is a real strength of the framework.

Second, as long as \( \rho A \) is not too big, so that the liquidity embodied in real assets is scarce, note that real assets and fiat currency can both be used to facilitate DM trade. Indeed, money can be valued along with many other assets, such as neoclassical capital, private or government bonds, housing, etc. With capital and technology shocks, Aruoba (2011) provides a strict generalization of standard RBC theory: when \( \alpha = 0 \) we recover exactly Hansen (1985). This is especially interesting because Hansen’s model does not start with quasi-linearity, but derives behaviour consistent with quasi-linearity by assuming labour is indivisible and agents trade employment lotteries. The same trick works here (Rochetteau et al., 2008).

Going back to a pure currency economy, recall the supply process \( M_{t+1} = (1 + \pi) M_t \). In stationary equilibrium all real variables including \( z_t \) are constant, so inflation is \( \phi_t = 1 + \pi_t \) which is pinned down by the rate of monetary expansion—a version of the quantity equation. Then the Fisher equation \( 1 + \pi_t = (1 + \pi_t) (1 + \rho) \) defines \( \rho_t \) which can be interpreted as the nominal interest rate on a bond that is illiquid (cannot be traded in the DM), just like \( r = 1/\beta - 1 \) is the interest rate on an illiquid real bond. To say more about why such bonds cannot be traded in the DM, one can take different approaches, including asymmetric information as in Lester et al. (2012) or Li et al. (2012), but that would take us too far afield for this presentation. In any case, with this notation, in steady state now (4) reduces to

\[ t = \alpha \sigma \lambda(z). \]  

Monetary policy, by manipulating \( \pi \) or \( t \), influences liquidity, exchange, and welfare. The Friedman rule, \( t = 0 \), happens to be optimal in this example, and yields the first best outcome \( y^* \), but that is not always true. With generalized Nash bargaining, if the buyer has bargaining power \( \theta < 1 \), then \( t = 0 \) is still optimal but yields \( y < y^* \). That might suggest \( t < 0 \) is desirable, but there is no monetary equilibrium with \( t < 0 \)—a New Monetarist version of the zero-lower-bound problem. Further, sometimes \( t > 0 \) can be optimal due to natural externalities in search models or other second-best considerations. Also, notice that aggregate CM output, consumption, and employment are independent of \( t \) here, but that is peculiar to the example, and is not true when \( x \) and \( y \) are not additively separable in utility. Other extensions include adding different types of banking and credit arrangements, labour market frictions, and endogenous growth. It is also worth re-emphasizing that the models are amenable to calibration or estimation. Many policy implications emerge but interested readers should look to the above-mentioned surveys for details and citations to primary sources. The goal here was to whet the appetite.

To conclude this discussion, allow me to address two questions. First, how does the above compare to a New Keynesian model? Well, the approaches are worlds apart. But one thing they have in common is that both can be used to talk about similar substantive issues—the zero lower bound, liquidity traps, open market operations, quantitative
easing, optimal inflation or nominal interest rates, etc. Beyond that, we need more theoretical and empirical work at the intersection of the approaches. Second, how does the model compare to a reduced-form approach? Well, models with CIA constraints are bound to look similar since, after all, they simply assume what we derive, and those who put money in utility functions can start with our $V(m)$ as a primitive as soon as they see our results. But I do not see in the reduced-form literature any analysis of the impact of the arrival rate $\alpha$, bargaining power $\theta$, or more general pricing mechanism $v(y)$. Nor do I see agents trading with each other in those models, or Keynesian models. Having these features in our theories leads to new insights, and ignoring progress along these lines while trying in other ways to force liquidity into mainstream macro is not following best practice.

VI. Conclusion

Let me conclude as I began by summarizing three main points this essay has tried to communicate.

1. Macro theories that are more or less mainstream do a good job helping us understand phenomena such as unemployment, growth, and cycles during normal times, as well as some monetary phenomena, like the impact of inflation. The financial crisis does not disprove these theories.

2. Understanding crises at more than a superficial level requires generalizing macro and microeconomics to better incorporate factors related to money, credit, banking, and liquidity, but solid microfoundations are critical. It is not enough to study models with money, banking, and credit, we need models of money, banking, and credit.

3. There is a mature body of work that speaks to these issues and can help more applied researchers and practitioners.

Jonathan Chiu commented on an earlier draft that the first two points are relatively easy to sell, while the third might take more work. Even people who recognize the need for microfoundations in the abstract could ask, do we need them to understand macro-financial crises? He also suggested arguments to support a positive answer: New Monetarist research highlights explicitly many elements that are commonly regarded as relevant in informal discussions, like information and other problems in ‘illiquid markets’. Additionally, understanding the fundamental roles of financial institutions and credit arrangements is critical for evaluating regulation and reform. Also, understanding the liquidity of different securities is critical for evaluating central bank policy. No one thinks swapping two $10 bills for a $20 bill will do very much, so why is swapping $10 bills for other public or private paper different? The answer depends on details, as described in, for example, Williamson (2012, 2016) or Rocheteau et al. (2016).

Others people suggested I mention that, prior to the crisis, many economists neglected money, liquidity, intermediation, and so on, because they considered this ‘plumbing’ that is best kept behind closed doors. They did not worry about how capital flows from savers to investors, in the same sense that GE theorists do not worry about how households get from their endowments to preferred bundles. But just saying savers must lend
to borrowers via a bank is typically not good enough. I want to know why they cannot lend directly in any model that purports to evaluate policy. A ban on banking shuts down investment in such models, but that is silly, because it ignores the resiliency of incentives. People would surely find ways around a banking ban in reality, and they should be allowed to do the same in our models. This is related to a situation where peso inflation leads to dollarization, something hard to analyze seriously with CIA models. As experience tells us, if the value of a local currency is inflated away very rapidly, dollars can take over. Intermediation is similar: assuming households cannot lend to firms without using banks is akin to assuming dollars cannot be used for transactions in Latin America, which is plainly wrong.

I want also to mention an email conversation with David Vines, who asks:

Why, after spending so much of the paper criticizing *ad hoc*, but simplifying, modelling decisions, does the model [in section V] have different centralized and decentralized markets? You mention that it is a feature of the example, rather than of the whole literature. It seems that more discussion of why this assumption is not *ad hoc* might be helpful.

Let me make two important points. First, having ‘different centralized and decentralized markets’ is a nod to realism. Extreme theories like Arrow–Debreu assume basically no frictions; at the opposite extreme, theories like Diamond–Mortensen–Pissarides or Kiyotaki–Wright assume frictions are rampant. Presumably the real world is somewhere in between. To capture this feature of reality formally, one needs a structure incorporating elements of both kinds of theories. Although there may be, in principle, many ways to encompass the models of Arrow et al. and Diamond et al., I have in practice found one that is quite tractable: alternate them in real time. Now, we can of course shut down either of the markets, in various ways, so this structure nests as special cases formulations with no frictions and those with many frictions, relating to search, bargaining, commitment, etc. To me that is a very attractive feature of the framework laid out in section V.

Second, the assumptions—i.e. the logical building blocks—at the foundation of this kind of model are in a precise sense less *ad hoc* than those in many more mainstream macro models. In case this is not obvious, the idea is this. This specification puts restrictions on how agents meet, what they can credibly promise, and so on. Given these restrictions, the agents do the best they can. More *ad hoc* are assumptions, like sticky prices, which entail agents not doing the best they can, but leave gains from trade sitting on the table, for no explicit reason. If one tries to go deeper into the analysis—say, deeper than starting with a Phillips curve, or even an Euler equation—one is led to problems like the one I discussed in footnote 10 and the surrounding text, where a worker and his employer trade way too much, or too little, labour because the wage is ‘stuck’ at the wrong level. That may or may not be realistic, but in any case, it needs to be explained and not assumed why the pair do not exploit bilaterally profitable deviations that require simply a conversation. That is logically different from, for example, making it difficult for workers and firms to meet in the first place. I am not sure if this is subtle, but to me it is a fundamental difference. Kosher assumptions concerning the environment (e.g. preferences, technology, how agents meet, or what they know) are legitimate; *ad hoc* assumptions concern behaviour of individuals, pairs, or coalitions more generally, as well as institutions, like how payments are made in cash-in-advance models.
In closing, I cannot help but mention how many macroeconomists really seem to enjoy a good crisis—attending the NBER during the Great Recession reminded me of watching The Weather Channel during a hurricane. Perhaps this is understandable. Medical science, I have heard, makes much progress during wars as doctors get chances to practise and experiment with new procedures and treatments. Economists presumably also get much information from extreme events, but too many people are too keen to declare failure and harp on about what’s wrong with macro. There are things wrong with macro, like micro, medicine, and meteorology, but at least some of us are trying to push forward. Since I was asked, my recommendation is to further pursue serious monetary economics, defined as the study of agents trading with each other over time, where explicit frictions make the process non-trivial, and where institutions that facilitate this process arise endogenously. And my prediction for the future of macro? As I said above, I predict it is only a matter of time until the kinds of models discussed in this essay become standard fare for those interested in liquidity.

References


On the future of macroeconomics

— (1986b), ‘Response to a Skeptic’, *FRB Minneapolis Quarterly Review*.


Is something really wrong with macroeconomics?

Ricardo Reis*

Abstract: Many critiques of the state of macroeconomics are off target. Current macroeconomic research is not mindless DSGE modelling filled with ridiculous assumptions and oblivious of data. Rather, young macroeconomists are doing vibrant, varied, and exciting work, getting jobs, and being published. Macroeconomics informs economic policy only moderately, and not more than nor differently from other fields in economics. Monetary policy has benefitted significantly from this advice in keeping inflation under control and preventing a new Great Depression. Macroeconomic forecasts perform poorly in absolute terms and, given the size of the challenge, probably always will. But relative to the level of aggregation, the time horizon, and the amount of funding, macroeconomic forecasts are not so obviously worse than those in other fields. What is most wrong with macroeconomics today is perhaps that there is too little discussion of which models to teach and too little investment in graduate-level textbooks.

Keywords: methodology, graduate teaching, forecasting, public debate

JEL classification: A11, B22, E00

I. Introduction

I accepted the invitation to write this essay and take part in this debate with great reluctance. The company is distinguished and the purpose is important. I expect the effort and arguments to be intellectually serious. At the same time, I call myself an economist and I have achieved a modest standing in this profession on account of (I hope) my ability to make some progress thinking about and studying the economy. I have no expertise in studying economists. I go to work every day to understand why inflation goes up and down or why some fiscal systems deliver better outcomes than others. Making progress on these questions frequently requires taking detours into narrow technical points on definitions of equilibrium or the properties of statistical estimators. But the focus always remains on understanding the economy, not the profession.

*London School of Economics, e-mail: r.a.reis@lse.ac.uk

This essay was written for the meeting on ‘The Future of Macroeconomic Theory’ organized by David Vines for the Oxford Review of Economic Policy. I am grateful to Chris Adam, John Barrdear, Francesco Caselli, Laura Castillo-Martinez, Wouter Den Haan, Greg Mankiw, Steve Pischke, Jesus Fernandez-Villaverde, Judith Shapiro, Paolo Surico, Silvana Tenreyro, and Randy Wright for comments and conversations.

doi:10.1093/oxrep/grx053

© The Author 2018. Published by Oxford University Press.
For permissions please e-mail: journals.permissions@oup.com
of economics. I personally love reading biographies and delight in thinking about what a young Alfred Marshall would say to a young Kenneth Arrow. Yet, I do not confuse these pleasurable intellectual leisure times with my job as a researcher.

On top of this, asking an active researcher in macroeconomics to consider what is wrong with macroeconomics today is sure to produce a biased answer. The answer is simple: everything is wrong with macroeconomics. Every hour of my workday is spent identifying where our knowledge falls short and how can I improve it. Researchers are experts at identifying the flaws in our current knowledge and in proposing ways to fix them. That is what research is. So, whenever you ask me what is wrong with any part of economics, I am trained by years on the job to tell you many ways in which it is wrong. With some luck, I may even point you to a paper that I wrote proposing a way to fix one of the problems.

While preparing for this article, I read many of the recent essays on macroeconomics and its future. I agree with much of what is in them, and benefit from having other people reflect on economists and the progress in the field. But to join a debate on what is wrong with economics by adding what is wronger with economics is not terribly useful. In turn, it would have been easy to share my thoughts on how macroeconomic research should change, which is, unsurprisingly, in the direction of my own research. I could have insisted that macroeconomics has over-relied on rational expectations even though there are at least a couple of well-developed, tractable, and disciplined alternatives. I could have pleaded for research on fiscal policy to move away from the over-study of what was the spending of the past (purchases) and to focus instead on the spending that actually dominates the government budget today (transfers). Going more methodological, I could have elaborated on my decade-long frustration dealing with editors and journals that insist that one needs a model to look at data, which is only true in a redundant and meaningless way and leads to the dismissal of too many interesting statistics while wasting time on irrelevant theories. However, while easy, this would not lead to a proper debate. A problem that too often plagues these discussions is that each panelist takes turns stating something else that is wrong with economics and pushing in a different direction. By the end, no opposing views are voiced, and the audience feels safe to agree with everything that was said while changing nothing in its day-to-day work, because there seem to be too many alternatives.

With all these caveats in mind, this essay instead provides a critical evaluation of the state of macroeconomics. I discuss four uses of macroeconomics, from those that are, in my view, less wrong, to those that perhaps need more change: research, policy, forecasting, and teaching. To contribute to the debate, I focus on responding to some of the negative verdicts on what is wrong with macroeconomics. The goal is to prevent these criticisms from being read as undisputed facts by the users of knowledge as opposed to the creators of knowledge. In substantive debates about actual economic policies, it is frustrating to have good economic thinking on macro topics being dismissed with a four-letter insult: it is a DSGE. It is worrying to see the practice of rigorously stating logic in precise mathematical terms described as a flaw instead of a virtue. It is perplexing to read arguments being boxed into macroeconomic theory (bad) as opposed

---

1 For my view on these three points, see Mankiw and Reis (2010), Oh and Reis (2012), and Hilscher, Raviv, and Reis (2014), respectively.
to microeconomic empirical work (good), as if there was such a strong distinction. It is dangerous to see public grant awards become strictly tied to some methodological directions to deal with the crisis in macroeconomics. I am not, in any way, claiming that there are no problems in macroeconomics, or that there should be no changes. My goal is not to claim that there is no disease, but rather to evaluate existing diagnoses, so that changes and progress are made in a productive direction.

II. The present of macroeconomic research

Mortality imposes that the future of macroeconomics will be shaped by the youngest members of the profession. There is something wrong with a field when bright young minds no longer find its questions interesting, or just reproduce the thoughts of close-minded older members. There is something right with it when the graduate students don’t miss the weekly seminar for work in progress, but are oblivious of the popular books in economics that newspapers and blogs debate furiously and tout as revolutionizing the field. To evaluate the state of macroeconomic research, as opposed to policy or the history of ideas, one should confront evaluations with evidence on what active researchers in the field are working on. Nobel prizes get most of the attention, and speeches of central bankers about their internal models are part of policy debates. But neither are the right place to look for the direction of the field. More accurate measures of the state of macroeconomics are what the journals have recently published, or what the recent hires of top departments are working on.

A good place to start is to read what some representative young macroeconomists actually work on. Every year, the *Review of Economic Studies* foreign editors select around six economists who have just been on the academic job market to give a tour of a handful of European institutions and present their research. These are not necessarily the best economists, or the ones that had more job offers, but they are typically the candidates that the editors are more excited about and that got more attention in the job market. Because the composition of the jury that picks them is heterogeneous and changes regularly, the choices are arguably not biased in the direction of a particular field, although they are most likely all in the mainstream tradition.\(^2\) Looking at their work gives a sample of what macroeconomic research is today. While they are at the top of the distribution when it comes to quality, these dissertation theses are fairly representative of what modern research in macroeconomics looks like. Here is my short description of what that is for the last 8 macroeconomists (with graduation date, PhD school, and first job in parentheses):

**Martin Beraja (2016, Chicago, MIT)**

Beraja’s job market paper developed a new method to identify the effectiveness of policies within models where the researcher is uncertain about some features of the economy

---

that the data have a hard time distinguishing. His focus is on identification in DSGE models that assume incomplete financial markets and sticky wages and this comes with clear applications to questions of redistribution via fiscal policy across states.

**Arlene Wong (2016, Northwestern, Princeton)**

Wong used micro data to show that it is mostly young people who adjust their consumption when monetary policy changes interest rates. Younger people are more likely to obtain a new mortgage once interest rate changes, either to buy a new home or to refinance an old one, and to spend the new available funds. Her research has painstaking empirical work that focuses on the role of mortgages and their refinancing features, and a model with much heterogeneity across households.

**Adrien Auclert (2015, MIT, Stanford)**

Auclert also focused on how changes in monetary policy affect spending and the macroeconomy, and also emphasized the heterogeneous responses by different households. He argued that when central banks lower interest rates, households whose assets have shorter duration than their liabilities lose out to households whose assets are of longer maturity than their liabilities. He then found that in the data the winners from these cuts in interest rates have higher propensity to spend than the losers, so that cuts in interest rates will boost aggregate spending.

**Gregor Jarosch (2015, Chicago, Stanford)**

Jarosch wrote a model to explain why losing your job leads to a very long-lasting decline in your lifetime wages. His hypothesis was that this is due to people climbing a ladder of jobs that are increasingly secure, so that when one has the misfortune of losing a job, this leads to a fall down the ladder and a higher likelihood of having further spells of unemployment in the future. He used administrative social security data to find some evidence for this hypothesis.

**Luigi Bocola (2014, Penn, Northwestern)**

Bocola tries to explain the depth of the crisis in Italy after 2011. He writes a DSGE model where banks hold sovereign debt, so that bad news about a possible future sovereign default both puts a strain on the funding of banks and also induces them to cut their leverage as a precautionary reaction. This channel for the diabolic loop linking banks and sovereign debt fits reasonably well the behaviour of credit spreads across Italian banks and firms, and predicts that the ECB’s interventions had a small effect.

**Saki Bigio (2012, NYU, Columbia)**

Bigio wanted to understand why banks don’t recapitalize fast enough after suffering large losses during a financial crisis, and this seems to be related to the slump in lending
and real activity that follows these crises. His explanation is that after large losses, banks
are less able to tolerate further losses, which lowers their ability to intermediate, and so
their future profits. Equity holders can then be stuck in a coordination failure, where
no one wants to inject new equity unless others do so as well, banks are stuck in a low
profit equilibrium, and the recovery must come through the slow process of retaining
earnings by banks.

**Matteo Maggiori (2012, Berkeley, NYU)**

Maggiori postulates that countries with more developed financial markets are able to
better deal with lack of funding in a financial crisis. They use this ability to sell insur-
ance to less developed countries, so that in normal times they receive an insurance
premium in the form of capital gains on foreign investments that sustain persistent
trade deficits. During a crisis though, the advanced countries should suffer the heavi-
est of capital losses and a larger fall in consumption, a prediction consistent with what
happened in the United States, but less so with what happened in Germany during the
Euro crisis.

**Joe Vavra (2012, Yale, Chicago)**

Vavra used data on individual prices to find that changes in prices tend to be more
dispersed and more frequent in recessions. He explains this by firms adjusting their
prices more often in recessions, in spite of the costs of doing so, because the volatility
of their firm-specific productivity is higher. But, with this more frequent price adjust-
ment, monetary policy shocks will be less effective at boosting real activity in recessions.

In my reading, this is all exciting work, connected to relevant applied questions, and
that takes data and models seriously. In contrast, in the caricatures of the state of mac-
roeconomics, there are only models with representative agents, perfect foresight, no
role or care for inequality, and a cavalier disregard for financial markets, mortgage con-
tracts, housing, or banks. Supposedly, macroeconomic research ignores identification
and does not take advantage of plentiful microeconomic data to test its models, which
anyway are too divorced from reality to be useful for any real world question. Compare
this caricature with the research that I just described: the contrast is striking. Not a
single one of these bright young minds that are the future of macroeconomics writes
the papers that the critics claim are what all of macroeconomic research is like today.
Instead, what they actually do is to mix theory and evidence, time-series aggregate data
and micro data, methodological innovations and applied policy questions, with no clear
patterns of ideology driven by geography.

Blanchard (2016), Korinek (2015), and Wren-Lewis (2017) worry that the current
standards and editorial criteria in macroeconomics undermine promising ideas, deter
needed diversity in the topics covered, and impose mindless work on DSGEs that brings
little useful knowledge to policy discussions. Smith (2016) emphasizes that we have far
less data than we would need to adequately test our models, and Romer (2016) that
identification is the perennial challenge for social sciences. Smith (2014) and Coyle and
Haldane (2014) characterize the state of economics, not as the perennial glass half full
and half empty, but rather as two glasses, one full and the other empty. In their view, applied empirical economists have been celebrating their successes, while macroeconomists lament their losses.

All of these criticisms contain some truth, but only up to a point. The research that I have just described is diverse, creative, and uses different data to identify causes. Young researchers in macroeconomics today do not seem bound by current standards or afraid to get their hands dirty. They are attacking these big challenges and trying to overcome the criticisms. The data and tools used by applied empirical economists are also used by macroeconomists. This is a sign of a field full of vitality, not of a field in trouble.

One might make the (elitist) criticism that, by focusing on these papers, I have looked only at the disruptive work that may cause scientific revolutions, while the problem is on what goes on in normal macroeconomic science. Table 1 reports the articles published in the latest issue of the top journal in macroeconomics, the *Journal of Monetary Economics*, including their authors, the title of the paper, and the highlights that the authors submitted. These include: theoretical papers on sovereign debt crises and capital controls, applied papers on the interrelation between financial indicators and macroeconomic aggregates, papers looking at extreme events like catastrophes and liquidity traps, and even purely empirical papers on measuring uncertainty in micro data and on forecasting time series in the macro data. There is originality and plurality, and a significant distance from the critics' portrayal of research.

Yet, according to De Grawe (2009), ‘The science of macroeconomics is in deep trouble’, while Skidelsky (2009) thinks that there has already been a ‘discrediting of mainstream macroeconomics’. These opinions express feelings more than facts, so it is hard to debate them. But if the collapse in the reputation of macroeconomists was as large as they claim, there should be hints of it at least in some rough measures of academic output and prestige. Space in the top journals in the economics profession is scarce. If macroeconomics was in a crisis, journals would, at least slowly, publish fewer and fewer articles on macroeconomics. From the demand side, general interest journals would not be interested in publishing articles that non-macroeconomists have no interest in reading. From the supply side, enough articles in a field must be written for a select few to be of sufficient quality to pass the difficult standards of these top journals.

Card and Della Vigna (2013) split the papers published in the top general-interest journals in the profession according to their field. They find no discernible change in the share of articles on macroeconomics over the last four decades. Figure 1 uses their approach, with some slight changes, in plotting the share of articles on macroeconomics, identified by a JEL code of E, that were published in the official journals of the two largest regional associations in economics, the American Economic Association and the European Economic Association. The sample goes from the start of 2000 to the end of 2016, so there are roughly as many years after the start of the Great Recession as there are before. Publication in the two journals follows the same trend: if anything, the share of papers in macroeconomics has been increasing over time. Figure 1 plots also the share of working papers published by the National Bureau of Economic Research (NBER) on macroeconomic topics to account for possible lags in the decline in macroeconomics due to publication delays. While there was a temporary decline in the share of macroeconomic papers right after 2008, for the past 5 years it has been steadily rising, and it is now at the highest level of the past 12 years.
## Table 1: Articles in the *Journal of Monetary Economics*, vol. 84, December 2016

<table>
<thead>
<tr>
<th>Authors</th>
<th>Title</th>
<th>Highlights</th>
</tr>
</thead>
</table>
| Gilles Chemla, Christopher A. Hennessy | Government as borrower of first resort | - A privately informed firm issues debt to a speculator and investors in safe assets.  
- With high uninformed safe asset demand, the private sector may pool at risky debt.  
- The government can increase welfare by issuing safe bonds, crowding out risky debt.  
- Government may eliminate risky debt and portfolio distortions, reducing investment.  
- Government debt can accommodate risky debt and distortions, encouraging investment. |
| David S. Miller                 | Commitment versus discretion in a political economy model of fiscal and monetary policy interaction | - Microfounding fiscal policy affects monetary policy decisions.  
- Time inconsistency is alleviated by the politically distorted fiscal authority.  
- Monetary responses mitigate the political distortion's effect.  
- Price commitment results in lower welfare as it eliminates monetary responses. |
| Vasco Cúrdia, Michael Woodford | Credit frictions and optimal monetary policy                         | - A positive average spread has little quantitative effect in the transmission of shocks.  
- Time variation in credit spread affects the relation between spending and policy rate.  
- Time variation in credit spread affects the relation between inflation and real activity.  
- Basic NK optimal target criterion is approximately optimal with credit spread.  
- The target criterion can be implemented by an augmented forward-looking Taylor rule. |
- It makes the term structure of interest rates decreasing, because of prudence.   
- It makes the term structure of risk premia increasing, because of risk aversion.   
- The uncertain trend or volatility of growth has a strong impact on asset prices.   
- The uncertain frequency of catastrophes plays a similar role. |
| Daniel Shoag, Stan Veuger       | Uncertainty and the geography of the great recession                  | - Local policy uncertainty during the Great Recession matches unemployment outcomes.   
- This relationship is robust to numerous controls.   
- Increased uncertainty contributed to the severity of the Great Recession. |
| Zhu Wang, Alexander L. Wolman   | Payment choice and currency use: insights from two billion retail transactions | - Rich transactions data covering payment patterns for 3 years, thousands of stores.  
- Consistent with theory of consumers' threshold transaction size for cash use.  
- Across transaction size, cash share falls and dispersion across locations rises.  
- Cash share displays weekly and monthly cycles, correlated with transaction volume.  
- Over the longer term, cash share has declined, largely replaced by debit. |
| Andrea L. Eisfeldt, Tyler Muir  | Aggregate external financing and savings waves                        | - Provide external finance cost time series using firm financing and savings decisions.  
- Estimated average cost of external finance is 2.3 per cent.  
- Provide evidence of external finance cost shocks.  
- Formally reject nested model without external finance cost shocks.  
- Document external finance and savings waves. |
### Table 1: Continued

<table>
<thead>
<tr>
<th>Authors</th>
<th>Title</th>
<th>Highlights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adrien Auclert, Matthew Rognlie</td>
<td>Unique equilibrium in the Eaton–Gersovitz model of sovereign debt</td>
<td>The Eaton–Gersovitz model is widely used for empirical analyses of sovereign debt markets. We show that the model with exogenous default value and short-term debt admits a unique equilibrium. - This counters the common view that sovereign debt markets are prone to multiple equilibria. - Multiplicity requires altering the timing of the model, or considering long-term debt.</td>
</tr>
<tr>
<td>Gianluca Benigno, Huigang Chen, Christopher Otrok, Alessandro Rebucci, Eric R. Young</td>
<td>Optimal capital controls and real exchange rate policies: a pecuniary externality perspective</td>
<td>- A new literature studies the use of capital controls to prevent financial crises. - If the exchange rate policy is costly, capital controls become part of the optimal policy mix. - This mix combines capital controls in tranquil times with exchange rate policy in crisis times. - It yields more borrowing, fewer and less severe crises, and higher welfare than capital controls alone.</td>
</tr>
<tr>
<td>Marco Cozzi, Giulio Fella</td>
<td>Job displacement risk and severance pay</td>
<td>We study the insurance role of severance pay in the presence of displacement risk. - Post-displacement earnings losses are sizeable and persistent due to loss of tenure. - Asset markets are incomplete. - We find that severance pay entails substantial welfare gains. - These welfare gains are negligible if earnings losses are not persistent.</td>
</tr>
<tr>
<td>Michael Abrahams, Tobias Adrian, Richard K. Crump, Emanuel Moench, Rui Yu</td>
<td>Decomposing real and nominal yield curves</td>
<td>- A term structure model for nominal and inflation-indexed government bonds. - Model is used to decompose yields into expectations and risk premia. - Variations in nominal term premia are primarily due to movements in real term premia. - LSAP announcements lowered yields mainly through a reduction of real term premia. - Monetary policy surprises primarily affect real forwards through real term premia.</td>
</tr>
<tr>
<td>Domenico Giannone, Francesca Monti, Lucrezia Reichlin</td>
<td>Exploiting the monthly data flow in structural forecasting</td>
<td>- A framework for combining structural models and now-casting is proposed. - Conditions for deriving the monthly dynamics of the model are discussed. - Linking the model with auxiliary variables improves now-casting performance. - The proposed model traces in real time the shocks driving the business cycle.</td>
</tr>
<tr>
<td>Lena Mareen Boneva, R. Anton Braun, Yuichiro Waki</td>
<td>Some unpleasant properties of loglinearized solutions when the nominal rate is zero</td>
<td>- We show that it matters how one solves the New Keynesian model at the zero lower bound (ZLB). - The nonlinear solution exhibits new types of ZLB equilibria that cannot occur using a loglinearized solution. - Fiscal multipliers are small and orthodox at the ZLB for a large and plausible set of parameterizations of the model. - The New Keynesian model can be used to make a case for supply-side fiscal stimulus at the ZLB. - In situations where a labour tax rate cut increases employment, the government purchase multiplier is about one or less.</td>
</tr>
<tr>
<td>Yang K. Lu, Robert G. King, Ernesto Pasten</td>
<td>Optimal reputation building in the New Keynesian model</td>
<td>We study how reputation building affects the optimal committed policy. - The reputation building effect can overturn the conventional policy prescriptions. - The reputation building effect is quantitatively important. - The reputation building effect is relevant over a large parameter space.</td>
</tr>
</tbody>
</table>
A related criticism of macroeconomics is that it ignores financial factors. Macroeconomists supposedly failed to anticipate the crisis because they were enamoured of models where financial markets and institutions were absent, as all financing was assumed to be efficient (De Grawe, 2009; Skidelsky, 2009). The field would be in denial if it continued to ignore these macro-financial links. Figure 2 checks this hypothesis in the article database, measuring the share of papers in the journals that have both the E and the G JEL fields, so they contain research at the intersection of both macroeconomics and finance. The figure shows that research in macro-finance has increased continuously over the sample. The share of macro-finance papers more than doubled for both the American Economic Review (AER) and the NBER from pre- to post-crisis, but was already on the rise since 2000. Of the increase in the macro share on average between 2000–7 and 2009–16, which was 3.7, 2.0, and 5.1 percentage points for the AER, Journal of the European Economic Association (JEEA), and NBER respectively, a very large part of it is accounted by macro-finance papers, which increased by 4.3, 1.3, and 3.9 per cent, respectively. Almost half of all macroeconomic papers in the AER in 2012 were also listed as finance papers. A more anecdotal piece of evidence comes from the 2012 survey by Brunnermeier et al. (2013) on macroeconomics with financial frictions. It runs for 93 pages, it cites 177 references, most written before the crisis, and it references six other books and surveys that the authors state that one must read to get a full picture of the research on the intersection between macroeconomics and financial factors. One can safely argue that there is a hole in our knowledge of macro-financial interactions; one might also argue more controversially that economists have filled this hole with rocks as opposed to diamonds; but it is harder to argue that the hole is empty.
Finally, on the demand side, macroeconomics can only have a future if there are still academic jobs for the young macroeconomists. Figure 3 shows the share of job postings in ‘Job Openings for Economists’, the main board for job advertisements for freshly minted PhDs, that again list macroeconomics as identified by its JEL code as the desired hire. The share is remarkably constant over the past 15 years. At least for now, the marketplace seems to continue to appreciate what macroeconomists do.

Surely, when looking back in the future, some current directions of research will have turned out to have been unproductive or even misguided. Journals have many flaws, and editors and referees are naturally biased towards propagating old paradigms, and to stick up for their turfs. But my reading of the evidence is that macroeconomic research is not on the path to self-destruction implied by its critics. Looking at the current research frontier led to a different description from the one that one gets from the critics, and one that is at odds with the pessimistic tone of their criticisms.

III. The performance of macroeconomic policy

Among all fields of economics, macroeconomics seems to be one of the ones that attracts the most attention from the popular media. At the same time, macroeconomists are very far from running the world. In deciding the size of the budget deficit, or whether a fiscal stimulus or austerity package is adopted, macroeconomists will often be heard by the press or policy-makers, but almost never play a decisive role in any of
the decisions that are made. Most macroeconomists support countercyclical fiscal policy, where public deficits rise in recessions, both in order to smooth tax rates over time and to provide some stimulus to aggregate demand. Looking at fiscal policy across the OECD countries over the last 30 years, it is hard to see too much of this advice being taken. Rather, policy is best described as deficits almost all the time, which does not match normative macroeconomics. Moreover, in popular decisions, like the vote in the United Kingdom to leave the European Union, macroeconomic considerations seemed to play a very small role in the choices of voters. Critics that blame the underperformance of the economy on economists vastly overstate the influence that economists actually have on economic policy.

One area where macroeconomists have perhaps more of an influence is in monetary policy. Central banks hire more PhD economists than any other policy institution, and in the United States, the current and past chair of the Federal Reserve are distinguished academic macroeconomists, as have been several members of the Federal Open Market Committee (FOMC) over the years. In any given week, there are at least one conference and dozens of seminars hosted at central banks all over the world where the latest academic research is discussed. The speeches of central bank governors refer to academic papers in macroeconomics more than those of any other policy-maker.

Looking at the major changes in the monetary policy landscape of the last few decades—central bank independence, inflation targeting, financial stability—they all

---

3 Not even economists think they had much of an impact on the Brexit vote; see den Haan et al. (2016).
followed long academic literatures. Even individual policies, like increasing transparency, the saturation of the market for reserves, forward guidance, and balance-sheet policy, were adopted following academic arguments and debates. In the small sub-field of monetary economics, one can at least partially assess its successes and failures in the real world by judging how central banks have done over the past few decades.

Every central bank that I know of in the developed world is in charge of keeping inflation low and stable. Some central banks have this as their only goal, others as one of several, but there is strong agreement across societies as reflected in central bank mandates that central banks can control inflation in the long run and keeping it stable is their main task. Figure 4, reproduced and updated from Reis (2016), compares the performance of four major central banks with regards to the measure of the price level that is stated in their legal mandates. In solid black is the actual outcome, in dashed grey is the target moving forward since a 2 per cent target was officially adopted, and in dotted grey is a hypothetical target from extrapolating the 2 per cent backwards in time. The hypothetical is important for the United States, since it had long been noted that the Federal Reserve behaved as if it had a target of 2 per cent even before this was decided. Comparing actual and expected, the conclusion for the United States, the Eurozone, and Canada is clear: monetary policy has been remarkably successful. For the United Kingdom, the price level drifted upwards after the crisis, although in its defence, the Bank of England interpreted its mandate as stating that bygones are

**Figure 4:** Actual price level and targets in four major central banks

Notes: The target price level is in the dashed grey line from the date of the announcement of the target forward, the hypothetical target is the extension of the target backwards in time (dotted grey line), and the actual price level is in the solid black line. All are normalized to equal zero at the date of adoption of the target, except for Canada that is normalized to zero in 1998. For the United States, the inflation target was adopted in January of 2012 using the personal consumption expenditures deflator as the reference measure. For the Euro area, the target was adopted in January 1999 for the harmonized consumer price index. For Canada, the target for the total consumer price index was adopted in 1991. For the Bank of England, the current target for the consumer price index target was adopted in December 2003. The target for all four is a 2 per cent annual growth in the price level. The vertical axis is in a log scale.
bygones when it comes to past deviations, so that since 2011, the slope of the price level has been approximately on target.

Another way to judge the performance of macroeconomics as applied to central banking is through the response to the crises of the last decade. Macroeconomists did not prevent the crises, but following the collapse of Lehman or the Greek default, news reports were dominated by non-economists claiming that capitalism was about to end and all that we knew was no longer valid, while economists used their analytical tools to make sense of events and suggest policies. In the United States in 2007–8, the Federal Reserve, led by the certified academic macroeconomist Ben Bernanke, acted swiftly and decisively. In terms of its conventional instruments, the Federal Reserve cut interest rates as far as it could and announced it would keep them low for a very long time. Moreover, it saturated the market for reserves by paying interest on reserves, and it expanded its balance sheet in order to affect interest rates at many horizons. Finally, it adopted a series of unconventional policies, intervening in financial markets to prevent shortages of liquidity. Some of these decisions are more controversial than others, and some were more grounded in macroeconomic research than others. But overall, facing an adverse shock that seems to have been as serious as the one behind the Great Depression, monetary policy responded, and the economy recovered. While the recession was deep, it was nowhere as devastating as a depression. The economic profession had spent decades studying the Great Depression, and documenting the policy mistakes that contributed to its severity; these mistakes were all avoided in 2008–10.4

Turning to the Eurozone crisis, many agree that the intervention of the European Central Bank (ECB) in defending the euro ‘whatever it takes’, in Mario Draghi’s famous words, was decisive in preventing a collapse of European sovereign debt markets. In turn, while other European and national authorities had difficulty agreeing on a response to the crisis, the ECB intervened quickly and decisively, and the supply of credit stayed up, even in the periphery countries with banking problems. Again, most of the interventions, both in stopping the sovereign debt crisis, and in using longer-term liquidity interventions, were justified and based on academic papers in macroeconomics. Without taking credit away from the policy-makers who had the courage to implement these policies, like the practical men in Keynes’s famous quotation, they were following the principles of macroeconomists.5

A separate criticism of macroeconomic policy advice accuses it of being politically biased. Since the early days of the field, with Keynes and the Great Depression, macroeconomics was associated with aggressive and controversial policies and with researchers who wore other hats as public intellectuals. More recently, during the rational-expectations microfoundations revolution of the 1970s, early papers had radical policy recommendations, like the result that all systematic aggregate-demand policy is ineffective, and some leading researchers had strong political views. Romer (2016) criticizes modern macroeconomics for raising questions about what should be obvious truths, like the effect of monetary policy on output. He lays blame on the influence that Edward Prescott, Robert Lucas, and Thomas Sargent had on the field. Krugman (2009), in turn, claims the problem of macroeconomics is ideology, and in particular

---

4 See Reis (2009) and Blinder (2013).

5 See Baldwin et al. (2015) and Brunnermeier and Reis (2017).
points to the fierce battles between different types of macroeconomists in the 1970s and 1980s, described by Hall (1976) in terms of saltwater versus freshwater camps. These features of the history of thought in macroeconomics are worth pointing out and discussing. But if they were crucial for diagnosing the state of the field, then they should stand out as very different from what happens in other fields in economics. Yet, labour economics also has a history of heated debates and strong ideological priors, as well as continuous re-examination of truths previously held as obvious, such as the effects of the minimum wage on employment or of immigration on wages.6 The father figures of modern public economics, such as Anthony Atkinson, Joseph Stiglitz, or Martin Feldstein, have also actively participated in popular debates with strong views in their role as public intellectuals. Researchers in both fields frequently make policy prescriptions, and their work is picked up by the media. These fields have been publicly promoted by the profession more than that of macroeconomists: of the last ten winners of the John Bates Clark medal, a prize given by the American Economic Association to honour economists under the age of 40, five have been researchers who list labour or public economics as one of their main fields of research.7 Macroeconomics does not stand out from labour and public economics in the features that the critics point out when they single it out for criticism.

The point is not to claim there are weaknesses in different fields of economics. The point is rather to note that macroeconomics is not all that special relative to the other fields. Economists across all fields were in part surprised by the crisis, but also eager to study it and analyse it. Economic theorists understood that we needed to invest more time in characterizing the role of speculation and sudden shifts in equilibrium; industrial organization economists turned their attention to auctions run by central banks and to the operation of payment systems; and financial economists realized how little attention we had paid to understanding rare events or to the measurement of systemic risk. There have been important debates on methods in development economics and in labour economics.8 Researchers in these fields, as in macroeconomics, perpetually feel dissatisfied with the state of their knowledge and work every day to improve it. Data have expanded and progress was made, but this is true both in microeconomics and macroeconomics.

To conclude, some of the diagnoses of the crisis in macroeconomics presuppose that macroeconomics is very different from the rest of economics, in having an outsized influence on policy, having more ideological researchers, or being especially hit in its credibility and methods by the crisis. This section noted that this specialness of macroeconomics is more apparent than real. As such, explanations for the problems of macroeconomics today that are too field specific may miss the target.

IV. Poor forecasting yes, but relative to what?

One way that macroeconomics stands out from other fields in economics is in how often it produces forecasts. The vast majority of empirical models in economics can be very

---

7 They are, in reverse chronological order, Roland Fryer, Raj Chetty, Amy Finkelstein, Emmanuel Saez, and Daron Acemoglu. The full list is here: https://www.aeaweb.org/about-aea/honors-awards/bates-clark
8 Angrist and Pischke (2009), Wolpin (2013), and Deaton and Cartwright (2016), among others.
successful at identifying causal relations or at fitting behaviour, but they are never used to provide unconditional forecasts, nor do people expect them to. Macroeconomists, instead, are asked to routinely produce forecasts to guide fiscal and monetary policy, and are perhaps too eager to comply. As I wrote in Reis (2010): ‘by setting themselves the goal of unconditional forecasting of aggregate variables, macroeconomists are setting such a high bar that they are almost sure to fail.’

Forecasting is hard. Forecasting what people will do when their behaviour is affected by many interrelated personal, local, and national variables is even harder. Forecasting when the forecasts cause changes in policy, which make people change their choices, which in turn make it required to revise the forecasts, is iteratively hard. Forecasting when economic agents themselves are forecasting your forecast to anticipate the policies that will be adopted, involves strategic thinking and game theory that goes well beyond the standard statistical toolbox. Very few economists that I know of would defend themselves too vigorously against the frequent criticisms of forecasting failures by economists. As is regularly shown, macroeconomic forecasts come with large and often serially correlated errors.9

At the same time, the way that forecasts are mis-read and mis-interpreted is part of the problem. As much as economists state that their forecasts are probabilities, and come with confidence bands, they are reported in the media always as point estimates. The Bank of England struggled to introduce fan charts as a way to display the uncertainty in its policy forecasts. Moreover, the supposedly most embarrassing forecast errors come with regards to large crises. Yet, these crises are rare events that happen once every many decades. Since typical economic time series only extend over a little more than one hundred years, statistically forecasting the eruption of a crisis will always come with large imprecision.10

Compare how economics does relative to the medical sciences. Analogies across sciences are always very tricky, and must be taken with a large grain of salt. Moreover, surely economists are still far from being as useful as dentists, as Keynes dreamed of, let alone to have made a contribution to human welfare that is close to the one made by doctors or biologists. The comparison to make is much narrower and more limited, restricted only to how economic forecasts compare to medical forecasts.

Imagine going to your doctor and asking her to forecast whether you will be alive 2 years from now. That would sound like a preposterous request to the physician, but perhaps having some actuarial mortality tables in her head, she would tell you the probability of death for someone of your age. For all but the older readers of this article, this will be well below 50 per cent. Yet, 1 year later, you have a heart attack and die. Should there be outrage at the state of medicine for missing the forecast, with such deadly consequences?

One defence by the medical profession would be to say that their job is not to predict time of death. They are driven to understand what causes diseases, how to prevent them, how to treat them, and altogether how to lower the chances of mortality while trading this off against life quality and satisfaction. Shocks are by definition unexpected, they cannot be predicted. In fact, in practice, most doctors would refuse to

\[9\] See, for instance, Edge and Gurkaynak (2010) and Wieland and Wolters (2012).
\[10\] For assessments of the state of forecasting see Clemens and Hendry (2011) or Elliott and Timmermann (2013).
answer the question in the first place, or they would shield any forecast with a blank statement that anything can happen. This argument applies, word for word, to economics once the word ‘disease’ is replaced by the words ‘financial crisis’.

A more sophisticated defence would note that medical sciences are about making conditional forecasts: if you make some lifestyle choices, then your odds of dying change by this or that much. These forecasts are at best probabilistic. Medical science can quantify in terms of conditional probabilities how certain behaviours affect mortality. Moreover, once the disease sets in, health researchers have given us the tools to understand what just happened to your body, rationalize it, and predict which treatments have some chances of helping, with what side effects. These lead to better choices and to better treatments, and they are a major contribution of the biomedical sciences to knowledge and human welfare.

Economics is not so different, even in 2007–8. Within days or weeks of the failure of Bear Sterns or Lehman Brothers, economists provided diagnoses of the crisis, and central banks and finance ministries implemented aggressive measures to minimize the damage, all of which were heavily influenced by economic theory. Economic concepts such as asymmetric information, bank runs, the role of liquidity, saturating the market for reserves, and forward guidance at the zero lower bound, all provided concrete interpretations of the crisis, suggestions for policies, and discussion of trade-offs. The economy did not die, and a Great Depression was avoided, in no small part due to the advances in economics over many decades.

Too many people all over the world are today being unexpectedly diagnosed with cancer, undergo enormously painful treatment, and recover to live for many more years. This is rightly hailed as a triumph of modern oncology, even if so much more remains to be done. After suffering the worst shock in many decades, the global economy’s problems were diagnosed by economists, who designed policies to respond to them, and in the end we had a painful recession but no melt-down. Some, somehow, conclude that economics is at fault.

At the same time, a doctor examining you in an emergency room can predict quite accurately how quickly the virus in your body will spread, and what the state of your health will be in 24 hours. Biologists and chemists can make remarkably sharp predictions of what will happen to your body after you take a certain medicine. Economists surely do not come even close to this. Perhaps, but the equivalent to these successes would be for me to crunch through the data on sales, customer characteristics, and others at the coffee shop downstairs, run many experiments varying the prices in the menu, and then use the economic model of a demand curve to predict what happens to coffee sales over the next week if we double the price. I conjecture that the economic forecast would be quite good. Macroeconomists are instead asked to predict what will happen to the changes in the CPI or GDP over the next 1–5 years. The comparison of forecast quality must be made for the same time horizon and for a similar level of aggregation. The fairer comparison would be to ask doctors to predict what will happen to the changes in the annual number of patients that eventually die after being admitted to an emergency room due to a stroke. For these similar units, my guess is that medical forecasts will look almost as bad as macroeconomic forecasts.

Currently, the major and almost single public funder for economic research in the United States is the National Science Foundation. Its 2015 budget for the whole of social, behavioural, and economic sciences was $276m. The part attributed to its social
and economic sciences group was $98m. The main public funder of health studies in the United States is the National Institute of Health (NIH), but there are many more, including several substantial private funders. The NIH’s budget for 2015 was $29 billion. Its National Institute of Allergy and Infectious Diseases alone received $4.2 billion in funding. A very conservative estimate is that society invests at least 40 times more trying to study infectious diseases, including forecasting the next flu season or the next viral outbreak, than it does in economics. More likely, the ratio of public investment to science devoted to predicting and preventing the next disease is two or even three orders of magnitude larger than the budget of science dedicated to predicting and preventing economics crises. There is no simple way to compare the output per unit of funding across different fields, but relative to its meagre funding, the performance of economics forecasting is perhaps not so bad.

A detour for another comparison may drive the point of this section in. There has been much progress in weather forecasting, such that predicting the weather over the next few days is done with less uncertainty than it was a decade ago. Forecasting the weather is an activity that takes as many or more resources as forecasting the economy, and that also affects a series of policy choices and economic decisions. Comparing macroeconomic forecasts to forecasts of average temperature or precipitation over the next 1–5 years, as opposed to over the next few days, it is far from clear that economics forecasting is doing so poorly.

To conclude with the most important message, yes, economic models do a poor job forecasting macroeconomic variables. This deserves to be exposed, discussed, and even sometimes ridiculed. Critics like Haldane (2016) are surely right, and the alternatives that they propose for improvement are definitely worth exploring. If nothing else, this may help the media and the public to start reporting and reading forecasts as probabilistic statements where the confidence bands or fan charts are as or more important than the point forecasts. But, before jumping to the conclusion that this is a damning critique of the state of macroeconomics, this section asked for an evaluation of forecasting performance in relative terms: relative to other conditional predictions on the effectiveness of policies, relative to other forecasts for large diverse populations also made many years out, and relative to their accuracy per dollar of funding. From these perspectives, I am less convinced that economics forecasting is all that far behind other scientific fields.

V. Redirecting the criticisms to teaching macroeconomics

If I replace ‘macroeconomic research’ with ‘macroeconomics as taught in entryway classes’ in the critiques of macroeconomics, they seem more on point. The doubts raised in this essay applied to the descriptions of the state of our knowledge, or to what is current macroeconomic research. Like Rodrik (2015) in his overall defence of economics, the validity of the criticisms and the scope for reform seem much clearer to me in regards to how macroeconomics is taught and how it is used by policy-makers. The popularity of criticisms of macroeconomics with the press and audiences in interdisciplinary debates likely has less to do with research, which most people know and care little about, but rather with their exposure to macroeconomics in the way it is taught and used in policy discussions.
At the undergraduate level, I see a productive debate taking place. The leading textbook in intermediate macroeconomics, *Macroeconomics* by N. G. Mankiw (see Mankiw, 2015), is regularly revised, and many chapters have been changed significantly in the last decade to address the issues raised by the crisis. In the fringes, there are new entrants to this market and healthy competition of ideas and approaches, including more radical changes, such as the one in the core-econ.org project. Macroeconomics is not alone here, as similar debates take place for instance in econometrics.\(^\text{11}\)

At the graduate level, there is more room for improvement. To start, empirical work in macroeconomics today includes a rich set of tools and approaches. Macroeconomists need to be trained in time series, and also to understand the fundamental identification problems, and the rich datasets that can be used to test behaviour. There are classic empirical questions around which one could structure an entire class in core macroeconomics, and taking the model to the data is today not an after-thought but an integral part of almost all research projects. Macroeconomics could be taught in a much more data-driven way than is done today.

Moreover, teaching is still tied to a benchmark frictionless neoclassical framework in the core graduate classes in macroeconomics, and this deserves to be questioned. Researchers in modern macroeconomics have made much progress in the last three decades to provide alternatives to the assumptions of full risk-sharing, full information, flexible prices, or lump-sum taxes, to name a few. For each of these assumptions, there are separate, tractable, simple, analytical models that could be taught in an introductory class. The challenge is to bring these together in a bare-bones model that can provide a new benchmark. I put forward that spending more effort debating what should be in such a model and trying to write it down would lead to the highest marginal return produced by debates on the state of macroeconomics.

This is a debate worth having, especially as I am sure that many would disagree with what weight empirical work should have in the core sequence, or with what ingredients should be part of the core model. Criticisms and discussions of macroeconomics focused on this discussion would be more constructive and get the wider community of macroeconomists involved. With more people pursuing graduate studies and higher demand for workers trained in advanced economic tools, graduate-level macroeconomics, especially at the Masters level, cannot be taught as if its only role was to train future academic researchers. As Mankiw (2006) and Blanchard (2017) emphasize, there is an important role for macroeconomists as engineers, as opposed to scientists, and this requires small usable models.

**VI. A small contribution to the discussion of a new core model**

A core model makes stark assumptions that make the model incredible, but also tractable and insightful on important economic mechanisms that are broadly applicable to many features of the world. The model should be simple, but it can still be richer than it is today. It is especially important that it takes what are today seen as imperfections or frictions as

\(^\text{11}\) Angrist and Pischke (2017).
the benchmark, while it makes the flexible, full-information limits become the special cases. As a starting point for the discussion, a teacher could structure a core model as follows.\(^\text{12}\)

(i) **Capitalists**

A fraction of the population lives for many periods and has time-separable preferences between consumption and leisure with exponential discounting. They use funds to consume and to invest in a series of financial instruments, including equity, bonds, deposits, government bonds, as well as a full set of Arrow–Debreu securities. They collect income from owning these assets, taxed at a capital-income tax rate, and income from working, taxed at a labour-income tax rate.

Setting up and solving this intertemporal problem allows students to learn about dynamic optimization. The Euler equation with respect to bond holdings provides the vehicle to teach intertemporal consumption smoothing. The Euler equation with respect to equity holdings provides the pretext to discuss the equity premium, and portfolio choice more broadly. Combining these (and maybe others) teaches students about no-arbitrage opportunities and about the role of the stochastic discount factor. Finally, working through the set of Arrow–Debreu securities, one can introduce the concept of full risk sharing, what complete markets actually means, and basic aggregation theorems, culminating in the definition of a representative agent. Throughout, taxes can be dropped and reintroduced along the way, so that their distortional effects are seen. Also throughout, there is empirical evidence on the equity premium, tests of risk sharing, the effects of distortional taxes, and debates on what is behind the Frisch elasticity for a representative agent.

One ingredient that I have left out, and which deserves further discussion, is whether to have finite or infinite lives. In particular, one could discuss the infinitely-lived agent as a succession of generations that are altruistically linked, so the discount factor measures preferences for future children. This would allow a discussion the Blanchard–Yali impure-altruism model, which gives some of the insights from overlapping-generations models.

(ii) **Workers**

Another fraction of the population refers to workers. Ideally, a household could transition between the two types in the population, so the separation would not be so stark. Allowing for these transitions is too much for the core model, but would provide a good problem set. Workers have the same preferences as capitalists, over the same goods and paying similar taxes. However, they differ in two important regards.

First, workers have idiosyncratic labour income shocks, perhaps because of shocks to their employment status and their health. These shocks lead to an income distribution and a discussion of inequality connected to the data. Moreover, they allow for a

\(^{12}\) I will leave out the many references to the literature in the description below. They would be too many, as they should be, since the ingredients below reflect the accumulation of knowledge over many articles and many years.
treatment of government transfers programmes, tied to health and employment status, which dominate the government budget today. Ideally, one would have endogenous unemployment, providing a good contrast with exogenous health. This can be modelled with a version of a search and matching framework that makes stark assumptions on how vacancies are posted and on separation, in order to keep the dynamics of unemployment simple.

Second, workers can only trade a bond between themselves, which is in net zero supply. Moreover, they are subject to a ‘maximally tight’ borrowing constraint, whereby they cannot borrow. Solving their problem introduces students to the standard incomplete-markets models and to precautionary savings. These can be taught by inspecting the Euler equation of this problem. Moreover, the equation has an easy solution: the joint assumptions of a bond in zero net supply and no-borrowing constraints, imply that in equilibrium every worker has zero savings. This matches the empirical evidence that a large fraction of the population has zero wealth, and can be used to discuss identification problems between this model and the Keynesian hand-to-mouth model, which assumes, rather than derives, zero savings. Clearing the market for the bonds gives an equation for the interest rate that moves with precautionary savings.

Left out is housing. Ideally, for both capitalists and workers, one would include a third good, beyond consumption and leisure, that is durable. This would allow students to understand the concept of user costs, and to understand how a good can also be an illiquid asset. Moreover, this would allow for a discussion of the empirically important connection between housing and borrowing.

(iii) Goods-producing firms

A continuum of firms exists, each producing a variety of consumer goods that are assembled according to a constant-elasticity aggregator. Each firm hires labour and capital goods to produce under a neoclassical production function. They are monopolistically competitive, and students can learn from their problem about the pricing choices of a monopolist.

Importantly, each firm sets its price only infrequently because of nominal rigidities. The simplest way to model them, which is not easily rejected by the data, and so can be discussed together with empirical evidence, is to assume sticky information. Each period a randomly drawn fraction of firms update their information, while the remainder keep to old price plans. Class discussions and evidence on rational expectations can happen at this stage. The model allows for a simple derivation of the Phillips curve relating marginal costs to inflation, and the importance of this relation in the history of macroeconomic thought.

Here, one can also discuss entry of firms subject to fixed costs and entry lags, together with exogenous firm exit. Under stark assumptions, this model of endogenous variety can be made simple with limited dynamics over time. It allows for a discussion of mark-ups and of basic firm dynamics. It can also be extended to give a brief discussion of innovation and endogenous growth.
(iv) Capital-producing firms and banks

A competitive sector produces capital goods but must borrow funds from banks. There are many banks, and each has an individual net worth that it can accumulate over time, subject to an exogenous end of life, and the payment of dividends to the capitalist owners. The relations between banks and other agents are subject to incentive constraints, which reflect themselves in different forms. Banks cannot take on outside equity, but only deposits, from capitalists. They attract these deposits and make loans subject to a leverage constraint derived from their ability to abscond with part of the funds. One can also discuss here, and introduce, adjustment costs to investment, together with restrictions on borrowing that are linked to the value of collateral.

Financial institutions would enter the core model using the tools of limited commitment. They can be complemented with a discussion of liquidity in financial markets. The basics of new monetarist models can be discussed, as banks may be subject to idiosyncratic withdrawal shocks, and at the same time borrow from each other in interbank markets that are modelled as being over the counter and requiring search for counterparties.

(v) The government

One branch of the government is the fiscal authority. It chooses taxes for labour and capital, collects dividends from the central bank, and finances social transfers as well as government purchases. Students can be taught about (naive) hyperbolic discounting by considering policy-makers that are tempted by the present because of elections and rotation in power. This allows for a potential conflict between the government and the private sector, and allows for the government not to be treated as benevolently as benchmark models today do.

The other branch of the government is the central bank, which is independent, and so has its own budget constraint and set of tools. Discussing the fiscal backing of the central bank, or in other words its effective independence, one can introduce students to hyperinflation and the informative data that result from these episodes. This section of the core can then focus on how central banks control inflation, by discussing the determination of the price level in general equilibrium. It also allows one to teach the history of inflation across the world.

(vi) DSGE

Finally, all of these parts can be brought together in a dynamic stochastic general equilibrium (DSGE) model. Note that, in each of the segments that I discussed, after going over the microfoundations and exploring several topics, one is left with only one or two key equations. Therefore, by the end, there are probably only about 10 or so equations left. Teaching students how to solve these as a system, through the tools of linearization, likelihood, and filtering, and culminating with estimation, would complete the class and connect its different ingredients to the data on business-cycle fluctuations and the effect of fiscal and monetary policy.
Note that in each these parts, there are one or two parameters that can be set to some limit (say zero) and in doing so recover the frictionless rational-expectations growth model. If the fraction of workers goes to zero, tax rates are zero, the degree of information rigidity is zero, firm entry costs are zero, banks can abscend with zero of the depositors’ money, the fiscal authority has no present bias, and so on, then one recovers the core model that currently dominates. But, and this is the main point, this is now a special case, rather than the benchmark. Moreover, stated as this limit, it becomes clear how much this old core leaves out, and how rich is the modern macroeconomics that builds on the new core.

I am sure that almost every reader of this article will disagree with at least one of the ingredients I listed above, or have a favourite one that should be included. Myself, I doubt that if I were to teach such a class for a few years, I would not conclude that some ingredients are best left out, and others should be brought in. My point here is not to provide the answer to what a new graduate core should be, but to show that there can be one. In the description above, I used core models that are taught in the first few weeks of field classes in macroeconomics with financial frictions, incomplete markets, or monetary economics. What remains is to bring them together as a core, as opposed to a disparate collection of benchmarks for different subfields of macroeconomics. It will take some effort to move beyond the brief description above, but this seems feasible. Organizing conferences where these ingredients are debated, and providing incentives for researchers to pedagogically defend their versions of core models, would be a fruitful future path for debates on the state of macro.

VII. Conclusion

I have argued that while there is much that is wrong with macroeconomics today, most critiques of the state of macroeconomics are off target. Current macroeconomic research is not mindless DSGE modelling filled with ridiculous assumptions and oblivious of data. Rather, young macroeconomists are doing vibrant, varied, and exciting work, getting jobs, and being published. Macroeconomics informs economic policy only moderately, and not more than, nor all that differently from, other fields in economics. Monetary policy has benefitted significantly from this advice in keeping inflation under control and preventing a new Great Depression. Macroeconomic forecasts perform poorly in absolute terms and given the size of the challenge probably always will. But relative to the level of aggregation, the time horizon, and the amount of funding, macroeconomic forecasts are not so obviously worse than those in other fields. What is most wrong with macroeconomics today is perhaps that there is too little discussion of which models to teach and too little investment in graduate-level textbooks.

References


Skidelsky, R. (2009), ‘How to Rebuild a Shamed Subject’, *Financial Times*, 5 August, available at https://www.ft.com/content/dfc9294a-81ef-11de-9c5e-00144feabdc0


Good enough for government work?
Macroeconomics since the crisis

Paul Krugman*

Abstract: This paper argues that when the financial crisis came policy-makers relied on some version of the Hicksian sticky-price IS-LM as their default model; these models were ‘good enough for government work’. While there have been many incremental changes suggested to the DSGE model, there has been no single ‘big new idea’ because the even simpler IS-LM type models were what worked well. In particular, the policy responses based on IS-LM were appropriate. Specifically, these models generated the insights that large budget deficits would not drive up interest rates and, while the economy remained at the zero lower bound, that very large increases in monetary base wouldn’t be inflationary, and that the multiplier on government spending was greater than 1. The one big exception to this satisfactory understanding was in price behaviour. A large output gap was expected to lead to a large fall in inflation, but did not. If new research is necessary, it is on pricing behaviour. While there was a failure to forecast the crisis, it did not come down to a lack of understanding of possible mechanisms, or of a lack of data, but rather through a lack of attention to the right data.

Keywords: macroeconomic models, pricing behaviour, measurement and data

JEL classification: E10, E30

I. Introduction

It’s somewhat startling, at least for those of us who bloviate about economics for a living, to realize just how much time has passed since the 2008 financial crisis. Indeed, the crisis and aftermath are starting to take on the status of an iconic historical episode, like the stagflation of the 1970s or the Great Depression itself, rather than that of freshly remembered experience. Younger colleagues sometimes ask me what it was like during the golden age of economics blogging, mainly concerned with macroeconomic debates, which they think of as an era that ended years ago.

Yet there is an odd, interesting difference, both among economists and with a wider audience, between the intellectual legacies of those previous episodes and what seems to be the state of macroeconomics now.

Each of those previous episodes of crisis was followed both by a major rethink- ing of macroeconomics and, eventually, by a clear victor in some of the fundamental debates. Thus, the Great Depression brought on Keynesian economics, which became...
the subject of fierce dispute—and everyone knew how those disputes turned out: Keynes, or Keynes as interpreted by and filtered through Hicks and Samuelson, won the argument.

In somewhat the same way, stagflation brought on the Friedman–Phelps natural rate hypothesis—yes, both men wrote their seminal papers before the 1970s, but the bad news brought their work to the top of the agenda. And everyone knew, up to a point anyway, how the debate over that hypothesis ended up: basically everyone accepted the natural rate idea, abandoning the notion of a long-run trade-off between inflation and unemployment. True, the profession then split into freshwater and saltwater camps over the effectiveness or lack thereof of short-run stabilization policies, a development that I think presaged some of what has happened since 2008. But I’ll get back to that.

For now, let me instead just focus on how different the economics-profession response to the post-2008 crisis has been from the responses to depression and stagflation. For this time there hasn’t been a big new idea, let alone one that has taken the profession by storm. Yes, there are lots of proclamations about things researchers should or must do differently, many of them represented in this issue of the Oxford Review. We need to put finance into the heart of the models! We need to incorporate heterogeneous agents! We need to incorporate more behavioural economics! And so on.

But while many of these ideas are very interesting, none of them seems to have emerged as the idea we need to grapple with. The intellectual impact of the crisis just seems far more muted than the scale of crisis might have led one to expect. Why?

Well, I’m going to offer what I suspect will be a controversial answer: namely, macroeconomics hasn’t changed that much because it was, in two senses, what my father’s generation used to call ‘good enough for government work’. On one side, the basic models used by macroeconomists who either practise or comment frequently on policy have actually worked quite well, indeed remarkably well. On the other, the policy response to the crisis, while severely lacking in many ways, was sufficient to avert utter disaster, which in turn allowed the more inflexible members of our profession to ignore events in a way they couldn’t in past episodes.

In what follows I start with the lessons of the financial crisis and Great Recession, which economists obviously failed to predict. I then move on to the aftermath, the era of fiscal austerity and unorthodox monetary policy, in which I’ll argue that basic macroeconomics, at least in one version, performed extremely well. I follow up with some puzzles that remain. Finally, I turn to the policy response—and its implications for the economics profession.

II. The Queen’s question

When all hell broke loose in financial markets, Queen Elizabeth II famously asked why nobody saw it coming. This was a good question—but maybe not as devastating as many still seem to think.

Obviously, very few economists predicted the crisis of 2008–9; those who did, with few exceptions I can think of, also predicted multiple other crises that didn’t happen. And this failure to see what was coming can’t be brushed aside as inconsequential.
There are, however, two different ways a forecasting failure of this magnitude can happen, which have very different intellectual implications. Consider an example from a different field, meteorology. In 1987 the Met Office dismissed warnings that a severe hurricane might strike Britain; shortly afterwards, the Great Storm of 1987 arrived, wreaking widespread destruction. Meteorologists could have drawn the lesson that their fundamental understanding of weather was fatally flawed—which they would presumably have done if their models had insisted that no such storm was even possible. Instead, they concluded that while the models needed refinement, the problem mainly involved data collection—that the network of weather stations, buoys, etc. had been inadequate, leaving them unaware of just how bad things were looking.

How does the global financial crisis compare in this respect? To be fair, the DSGE models that occupied a lot of shelf space in journals really had no room for anything like this crisis. But macroeconomists focused on international experience—one of the hats I personally wear—were very aware that crises triggered by loss of financial confidence do happen, and can be very severe. The Asian financial crisis of 1997–9, in particular, inspired not just a realization that severe 1930s-type downturns remain possible in the modern world, but a substantial amount of modelling of how such things can happen.

So the coming of the crisis didn’t reveal a fundamental conceptual gap. Did it reveal serious gaps in data collection? My answer would be, sort of, in the following sense: crucial data weren’t so much lacking as overlooked.

This was most obvious on the financial side. The panic and disruption of financial markets that began in 2007 and peaked after the fall of Lehman came as a huge surprise, but one can hardly accuse economists of having been unaware of the possibility of bank runs. If most of us considered such runs unlikely or impossible in modern advanced economies, the problem was not conceptual but empirical: failure to take on board the extent to which institutional changes had made conventional monetary data inadequate.

This is clearly true for the United States, where data on shadow banking—on the repo market, asset-backed commercial paper, etc.—were available but mostly ignored. In a less obvious way, European economists failed to pay sufficient intention to the growth of interbank lending as a source of finance. In both cases the institutional changes undermined the existing financial safety net, especially deposit insurance. But this wasn’t a deep conceptual issue: when the crisis struck, I’m sure I wasn’t the only economist whose reaction was not ‘How can this be happening?’ but rather to yell at oneself, ‘Diamond–Dybvig, you idiot!’

In a more subtle way, economists were also under-informed about the surge in housing prices that we now know represented a huge bubble, whose bursting was at the heart of the Great Recession. In this case, rising home prices were an unmistakable story. But most economists who looked at these prices focused on broad aggregates—say, national average home prices in the United States. And these aggregates, while up substantially, were still in a range that could seemingly be rationalized by appealing to factors like low interest rates. The trouble, it turned out, was that these aggregates masked the reality, because they averaged home prices in locations with elastic housing supply (say, Houston or Atlanta) with those in which supply was inelastic (Florida—or Spain); looking at the latter clearly showed increases that could not be easily rationalized.
Let me add a third form of data that were available but largely ignored: it’s fairly remarkable that more wasn’t made of the sharp rise in household debt, which should have suggested something unsustainable about the growth of the 2001–7 era. And in the aftermath of the crisis macroeconomists, myself included (Eggertsson and Krugman, 2012) began taking private-sector leverage seriously in a way they should arguably have been doing before.

So did economists ignore warning signs they should have heeded? Yes. One way to summarize their (our) failure is that they ignored evidence that the private sector was engaged in financial overreach on multiple fronts, with financial institutions too vulnerable, housing prices in a bubble, and household debt unsustainable. But did this failure of observation indicate the need for a fundamental revision of how we do macroeconomics? That’s much less clear.

First, was the failure of prediction a consequence of failures in the economic framework that can be fixed by adopting a radically different framework? It’s true that a significant wing of both macroeconomists and financial economists were in the thrall of the efficient markets hypothesis, believing that financial overreach simply cannot happen—or at any rate that it can only be discovered after the fact, because markets know what they are doing better than any observer. But many macroeconomists, especially in policy institutions, knew better than to trust markets to always get it right—especially those who had studied or been involved with the Asian crisis of the 1990s. Yet they (we) also missed some or all of the signs of overreach. Why?

My answer may seem unsatisfying, but I believe it to be true: for the most part what happened was a demonstration of the old line that predictions are hard, especially about the future. It’s a complicated world out there, and one’s ability to track potential threats is limited. Almost nobody saw the Asian crisis coming, either. For that matter, how many people worried about political disruption of oil supplies before 1973? And so on. At any given time there tends to be a set of conventional indicators everyone looks at, determined less by fundamental theory than by recent events, and big, surprise crises almost by definition happen due to factors not on that list. If you like, it’s as if meteorologists with limited resources concentrated those resources in places that had helped track previous storms, leading to the occasional surprise when a storm comes from an unusual direction.

A different question is whether, now that we know whence the 2008 crisis came, it points to a need for deep changes in macroeconomic thinking. As I’ve already noted, bank runs have been fairly well understood for a long time; we just failed to note the changing definition of banks. The bursting of the housing bubble, with its effects on residential investment and wealth, was conceptually just a negative shock to aggregate demand.

The role of household leverage and forced deleveraging is a bigger break from conventional macroeconomics, even as done by saltwater economists who never bought into efficient markets and were aware of the risk of financial crises. That said, despite the impressive empirical work of Mian and Sufi (2011) and my own intellectual investment in the subject, I don’t think we can consider incorporating debt and leverage a fundamental new idea, as opposed to a refinement at the margin.
It's true that introducing a role for household debt in spending behaviour makes the short-run equilibrium of the economy dependent on a stock variable, the level of debt. But this implicit role of stock variables in short-run outcomes isn't new: after all, nobody has ever questioned the notion that investment flows depend in part on the existing capital stock, and I'm not aware that many macroeconomists consider this a difficult conceptual issue.

And I'm not even fully convinced that household debt played that large a role in the crisis. Did household spending fall that much more than one would have expected from the simple wealth effects of the housing bust?

My bottom line is that the failure of nearly all macroeconomists, even of the saltwater camp, to predict the 2008 crisis was similar in type to the Met Office failure in 1987, a failure of observation rather than a fundamental failure of concept. Neither the financial crisis nor the Great Recession that followed required a rethinking of basic ideas.

III. Not believing in (confidence) fairies

Once the Great Recession had happened, the advanced world found itself in a situation not seen since the 1930s, except in Japan, with policy interest rates close to zero everywhere. This raised the practical question of how governments and central banks should and would respond, of which more later. For economists, it raised the question of what to expect as a result of those policy responses. And the predictions they made were, in a sense, out-of-sample tests of their theoretical framework: economists weren't trying to reproduce the historical time-series behaviour of aggregates given historical policy regimes, they were trying to predict the effects of policies that hadn't been applied in modern times in a situation that hadn't occurred in modern times.

In making these predictions, the deep divide in macroeconomics came into play, making a mockery of those who imagined that time had narrowed the gap between saltwater and freshwater schools. But let me put the freshwater school on one side, again pending later discussion, and talk about the performance of the macroeconomists—many of them trained at MIT or Harvard in the 1970s—who had never abandoned their belief that activist policy can be effective in dealing with short-run fluctuations. I would include in this group Ben Bernanke, Olivier Blanchard, Christina Romer, Mario Draghi, and Larry Summers, among those close to actual policy, and a variety of academics and commentators, such as Simon Wren-Lewis, Martin Wolf, and, of course, yours truly, in supporting roles.

I think it's fair to say that everyone in this group came into the crisis with some version of Hicksian sticky-price IS-LM as their default, back-of-the-envelope macroeconomic model. Many were at least somewhat willing to work with DSGE models, maybe even considering such models superior for many purposes. But when faced with what amounted to a regime change from normal conditions to an economy where policy interest rates couldn't fall, they took as their starting point what the Hicksian approach predicted about policy in a liquidity trap. That is, they did not rush to develop new theories, they pretty much stuck with their existing models.
These existing models made at least three strong predictions that were very much at odds with what many influential figures in the political and business worlds (backed by a few economists) were saying.

- First, Hicksian macroeconomics said that very large budget deficits, which one might normally have expected to drive interest rates sharply higher, would not have that effect near the zero lower bound.
- Second, the same approach predicted that even very large increases in the monetary base would not lead to high inflation, or even to corresponding increases in broader monetary aggregates.
- Third, this approach predicted a positive multiplier, almost surely greater than 1, on changes in government spending and taxation.

These were not common-sense propositions. Non-economists were quite sure that the huge budget deficits the US ran in 2009–10 would bring on an attack by the ‘bond vigilantes’. Many financial commentators and political figures warned that the Fed’s expansion of its balance sheet would ‘debase the dollar’ and cause high inflation. And many political and policy figures rejected the Keynesian proposition that spending more would expand the economy, spending less lead to contraction.

In fact, if you’re looking for a post-2008 equivalent to the kinds of debate that raged in the 1930s and again in the 1970s—a conflict between old ideas based on pre-crisis thinking, and new ideas inspired by the crisis—your best candidate would be fiscal policy. The old guard clung to the traditional Keynesian notion of a government spending multiplier somewhat limited by automatic stabilizers, but still greater than 1. The new economic thinking that achieved actual real-world influence during the crisis and aftermath—as opposed, let’s be honest, to the kind of thinking found in this issue—mostly involved rejecting the Keynesian multiplier in favour of the doctrine of expansionary austerity, the argument that cutting public spending would crowd in large amounts of private spending by increasing confidence (Alesina and Ardagna, 2010). (The claim that bad things happen when public debt crosses a critical threshold also played an important real-world role, but was less a doctrine than a claimed empirical observation.)

So here, at least, there was something like a classic crisis-inspired confrontation between tired old ideas and a radical new doctrine. Sad to say, however, as an empirical matter the old ideas were proved right, at least insofar as anything in economics can be settled by experience, while the new ideas crashed and burned. Interest rates stayed low despite huge deficits. Massive expansion in the monetary base did not lead to inflation. And the experience of austerity in the euro area, coupled with the natural experiments created by some of the interregional aspects of the Obama stimulus, ended up strongly supporting a conventional, Keynesian view of fiscal policy. Even the magnitude of the multiplier now looks to be around 1.5, which was the number conventional wisdom suggested in advance of the crisis.

So the crisis and aftermath did indeed produce a confrontation between innovative new ideas and traditional views largely rooted in the 1930s. But the movie failed to follow the Hollywood script: the stodgy old ideas led to broadly accurate predictions, were indeed validated to a remarkable degree, while the new ideas proved embarrassingly wrong. Macroeconomics didn’t change radically in response to crisis because old-fashioned models, confronted with a new situation, did just fine.
IV. The case of the missing deflation

I’ve just argued that the lack of a major rethinking of macroeconomics in the aftermath of crisis was reasonable, given that conventional, off-the-shelf macroeconomics performed very well. But this optimistic assessment needs to be qualified in one important respect: while the demand side of economy did just about what economists trained at MIT in the 1970s thought it would, the supply side didn’t.

As I said, the experience of stagflation effectively convinced the whole profession of the validity of the natural-rate hypothesis. Almost everyone agreed that there was no long-run inflation–unemployment trade-off. The great saltwater–freshwater divide was, instead, about whether there were usable short-run trade-offs.

But if the natural-rate hypothesis was correct, sustained high unemployment should have led not just to low inflation but to continually declining inflation, and eventually deflation. You can see a bit of this in some of the most severely depressed economies, notably Greece. But deflation fears generally failed to materialize.

Put slightly differently, even saltwater, activist-minded macroeconomists came into the crisis as ‘accelerationists’: they expected to see a downward-sloping relationship between unemployment and the rate of change of inflation. What we’ve seen instead is, at best, something like the 1960s version of the Phillips curve, a downward-sloping relationship between unemployment and the level of inflation—and even that relationship appears weak.

Obviously this empirical failure has not gone unnoticed. Broadly, those attempting to explain price behaviour since 2008 have gone in two directions. One side, e.g. Blanchard (2016), invokes ‘anchored’ inflation expectations: the claim that after a long period of low, stable inflation, price-setters throughout the economy became insensitive to recent inflation history, and continued to build 2 per cent or so inflation into their decisions even after a number of years of falling below that target. The other side, e.g. Daly and Hobijn (2014), harking back to Tobin (1972) and Akerlof et al. (1996), invokes downward nominal wage rigidity to argue that the natural rate hypothesis loses validity at low inflation rates.

In a deep sense, I’d argue that these two explanations have more in common than they may seem to at first sight. The anchored-expectations story may preserve the outward form of an accelerationist Phillips curve, but it assumes that the process of expectations formation changes, for reasons not fully explained, at low inflation rates. The nominal rigidity story assumes that there is a form of money illusion, opposition to outright nominal wage cuts, that is also not fully explained but becomes significant at low overall inflation rates.

Both stories also seem to suggest the need for aggressive expansionary policy when inflation is below target: otherwise there’s the risk that expectations may become unanchored on the downward side, or simply that the economy will suffer persistent, unnecessary slack because the downward rigidity of wages is binding for too many workers.

Finally, I would argue that it is important to admit that both stories are ex post explanations of macroeconomic behaviour that was not widely predicted in advance of the post-2008 era. Pre-2008, the general view even on the saltwater side was that stable inflation was a sufficient indicator of an economy operating at potential output, that any persistent negative output gap would lead to steadily declining inflation and eventually outright deflation. This view was, in fact, a key part of the intellectual case for
inflation targeting as the basis of monetary policy. If inflation will remain stable at, say, 1 per cent even in a persistently depressed economy, it’s all too easy to see how policy-makers might give themselves high marks even while in reality failing at their job.

But while this is a subjective impression—I haven’t done a statistical analysis of recent literature—it does seem that surprisingly few calls for a major reconstruction of macroeconomics focus on the area in which old-fashioned macroeconomics did, in fact, perform badly post-crisis.

There have, for example, been many calls for making the financial sector and financial frictions much more integral to our models than they are, which is a reasonable thing to argue. But their absence from DSGE models wasn’t the source of any major predictive failures. Has there been any comparable chorus of demands that we rethink the inflation process, and reconsider the natural rate hypothesis? Of course there have been some papers along those lines, but none that have really resonated with the profession.

Why not? As someone who came of academic age just as the saltwater–freshwater divide was opening up, I think I can offer a still-relevant insight: understanding wage- and price-setting is hard—basically just not amenable to the tools we as economists have in our kit. We start with rational behaviour and market equilibrium as a baseline, and try to get economic dysfunction by tweaking that baseline at the edges; this approach has generated big insights in many areas, but wages and prices isn’t one of them.

Consider the paths followed by the two schools of macroeconomics.

Freshwater theory began with the assumption that wage- and price-setters were rational maximizers, but with imperfect information, and that this lack of information explained the apparent real effects of nominal shocks. But this approach became obviously untenable by the early 1980s, when inflation declined only gradually despite mass unemployment. Now what?

One possible route would have been to drop the assumption of fully rational behaviour, which was basically the New Keynesian response. For the most part, however, those who had bought into Lucas-type models chose to cling to the maximizing model, which was economics as they knew how to do it, despite attempts by the data to tell them it was wrong. Let me be blunt: real business cycle theory was always a faintly (or more than faintly) absurd enterprise, a desperate attempt to protect intellectual capital in the teeth of reality.

But the New Keynesian alternative, while far better, wasn’t especially satisfactory either. Clever modellers pointed out that in the face of imperfect competition the aggregate costs of departures from perfectly rational price-setting could be much larger than the individual costs. As a result, small menu costs or a bit of bounded rationality could be consistent with widespread price and wage stickiness.

To be blunt again, however, in practice this insight served as an excuse rather than a basis for deep understanding. Sticky prices could be made respectable—just—allowing modellers to assume something like one-period-ahead price-setting, in turn letting models that were otherwise grounded in rationality and equilibrium produce something not too inconsistent with real-world observation. New Keynesian modelling thus acted as a kind of escape clause rather than a foundational building block.

But is that escape clause good enough to explain the failure of deflation to emerge despite multiple years of very high unemployment? Probably not. And yet we still lack a compelling alternative explanation, indeed any kind of big idea. At some level, wage
and price behaviour in a depressed economy seems to be a subject for which our intellectual tools are badly fitted.

The good news is that if one simply assumed that prices and wages are sticky, appealing to the experience of the 1930s and Japan in the 1990s (which never experienced a true deflationary spiral), one did reasonably well on other fronts.

So my claim that basic macroeconomics worked very well after the crisis needs to be qualified by what looks like a big failure in our understanding of price dynamics—but this failure didn’t do too much damage in giving rise to bad advice, and hasn’t led to big new ideas because nobody seems to have good ideas to offer.

V. The system sort of worked

In 2009 Barry Eichengreen and Kevin O’Rourke made a splash with a data comparison between the global slump to date and the early stages of the Great Depression; they showed that at the time of writing the world economy was in fact tracking quite close to the implosion that motivated Keynes’s famous essay ‘The Great Slump of 1930’ (Eichengreen and O’Rourke, 2009).

Subsequent updates, however, told a different story. Instead of continuing to plunge as it did in 1930, by the summer of 2009 the world economy first stabilized, then began to recover. Meanwhile, financial markets also began to normalize; by late 2009 many measures of financial stress were more or less back to pre-crisis levels.

So the world financial system and the world economy failed to implode. Why?

We shouldn’t give policy-makers all of the credit here. Much of what went right, or at least failed to go wrong, reflected institutional changes since the 1930s. Shadow banking and wholesale funding markets were deeply stressed, but deposit insurance still protected a good part of the banking system from runs. There never was much discretionary fiscal stimulus, but the automatic stabilizers associated with large welfare states kicked in, well, automatically: spending was sustained by government transfers, while disposable income was buffered by falling tax receipts.

That said, policy responses were clearly much better than they were in the 1930s. Central bankers and fiscal authorities officials rushed to shore up the financial system through a combination of emergency lending and outright bailouts; international cooperation assured that there were no sudden failures brought on by shortages of key currencies. As a result, disruption of credit markets was limited in both scope and duration. Measures of financial stress were back to pre-Lehman levels by June 2009.

Meanwhile, although fiscal stimulus was modest, peaking at about 2 per cent of GDP in the United States, during 2008–9 governments at least refrained from drastic tightening of fiscal policy, allowing automatic stabilizers—which, as I said, were far stronger than they had been in the 1930s—to work.

Overall, then, policy did a relatively adequate job of containing the crisis during its most acute phase. As Daniel Drezner argues (2012), ‘the system worked’—well enough, anyway, to avert collapse.

So far, so good. Unfortunately, once the risk of catastrophic collapse was averted, the story of policy becomes much less happy. After practising more or less Keynesian policies in the acute phase of the crisis, governments reverted to type: in much of the
advanced world, fiscal policy became Hellenized, that is, every nation was warned that it could become Greece any day now unless it turned to fiscal austerity. Given the validation of Keynesian multiplier analysis, we can confidently assert that this turn to austerity contributed to the sluggishness of the recovery in the United States and the even more disappointing, stuttering pace of recovery in Europe.

Figure 1 sums up the story by comparing real GDP *per capita* during two episodes: Western Europe after 1929 and the EU as a whole since 2007. In the modern episode, Europe avoided the catastrophic declines of the early 1930s, but its recovery has been so slow and uneven that at this point it is tracking below its performance in the Great Depression.

Now, even as major economies turned to fiscal austerity, they turned to unconventional monetary expansion. How much did this help? The literature is confusing enough to let one believe pretty much whatever one wants to. Clearly Mario Draghi’s ‘whatever it takes’ intervention (Draghi, 2012) had a dramatic effect on markets, heading off what might have been another acute crisis, but we never did get a clear test of how well outright monetary transactions would have worked in practice, and the evidence on the effectiveness of Fed policies is even less clear.

The purpose of this paper is not, however, to evaluate the decisions of policy-makers, but rather to ask what lessons macroeconomists should and did take from events. And the main lesson from 2010 onwards was that policy-makers don’t listen to us very much, except at moments of extreme stress.

This is clearest in the case of the turn to austerity, which was not at all grounded in conventional macroeconomic models. True, policy-makers were able to find some economists telling them what they wanted to hear, but the basic Hicksian approach that did pretty well over the whole period clearly said that depressed economies near the zero lower bound should not be engaging in fiscal contraction. Never mind, they did it anyway.

**Figure 1:** Slumps and recoveries in two crises, real GDP *per capita*, pre-crisis = 100

*Note:* The x-axis shows the number of years after each crisis.
Even on monetary policy, where economists ended up running central banks to a degree I believe was unprecedented, the influence of macroeconomic models was limited at best. A basic Hicksian approach suggests that monetary policy is more or less irrelevant in a liquidity trap. Refinements (Krugman, 1998; Eggertsson and Woodford, 2003) suggested that central banks might be able to gain traction by raising their inflation targets, but that never happened.

The point, then, is that policy failures after 2010 tell us relatively little about the state of macroeconomics or the ways it needs to change, other than that it would be nice if people with actual power paid more attention. Macroeconomists aren’t, however, the only researchers with that problem; ask climate scientists how it’s going in their world.

Meanwhile, however, what happened in 2008–9—or more precisely, what didn’t happen, namely utter disaster—did have an important impact on macroeconomics. For by taking enough good advice from economists to avoid catastrophe, policy-makers in turn took off what might have been severe pressure on economists to change their own views.

VI. That 80s show

Why hasn’t macroeconomics been transformed by (relatively) recent events in the way it was by events in the 1930s or the 1970s? Maybe the key point to remember is that such transformations are rare in economics, or indeed in any field. ‘Science advances one funeral at a time,’ quipped Max Planck: researchers rarely change their views much in the light of experience or evidence. The 1930s and the 1970s, in which senior economists changed their minds—e.g. Lionel Robbins converting to Keynesianism, were therefore exceptional.

What made them exceptional? Each case was marked by developments that were both clearly inconsistent with widely held views and sustained enough that they couldn’t be written off as aberrations. Lionel Robbins published *The Great Depression*, a very classical/Austrian interpretation that prescribed a return to the gold standard, in 1934. Would he have become a Keynesian if the Depression had ended by the mid-1930s? The widespread acceptance of the natural-rate hypothesis came more easily, because it played into the neoclassical mindset, but still might not have happened as thoroughly if stagflation had been restricted to a few years in the early 1970s.

From an intellectual point of view, I’d argue, the Great Recession and aftermath bear much more resemblance to the 1979–82 Volcker double-dip recession and subsequent recovery in the United States than to either the 1930s or the 1970s. And here I can speak in part from personal recollection.

By the late 1970s the great division of macroeconomics into rival saltwater and freshwater schools had already happened, so the impact of the Volcker recession depended on which school you belonged to. But in both cases it changed remarkably few minds.

For saltwater macroeconomists, the recession and recovery came mainly as validation of their pre-existing beliefs. They believed that monetary policy has real effects, even if announced and anticipated; sure enough, monetary contraction was followed by a large real downturn. They believed that prices are sticky and inflation has a great deal of inertia, so that monetary tightening would produce a ‘clockwise spiral’ in unemployment
and inflation: unemployment would eventually return to the NAIRU (non-accelerating inflation rate of unemployment) at a lower rate of inflation, but only after a transition period of high unemployment. And that’s exactly what we saw.

Freshwater economists had a harder time: Lucas-type models said that monetary contraction could cause a recession only if unanticipated, and as long as economic agents couldn’t distinguish between individual shocks and an aggregate fall in demand. None of this was a tenable description of 1979–82. But recovery came soon enough and fast enough that their worldview could, in effect, ride out the storm. (I was at one conference where a freshwater economist, questioned about current events, snapped ‘I’m not interested in the latest residual.’)

What I see in the response to 2008 and after is much the same dynamic. Half the macroeconomics profession feels mainly validated by events—correctly, I’d say, although as part of that faction I would say that, wouldn’t I? The other half should be reconsidering its views—but they should have done that 30 years ago, and this crisis, like that one, was sufficiently well-handled by policy-makers that there was no irresistible pressure for change. (Just to be clear, I’m not saying that it was well-handled in an objective sense: in my view we suffered huge, unnecessary losses of output and employment because of the premature turn to austerity. But the world avoided descending into a full 1930s-style depression, which in effect left doctrinaire economists free to continue believing what they wanted to believe.)

If all this sounds highly cynical, well, I guess it is. There’s a lot of very good research being done in macroeconomics now, much of it taking advantage of the wealth of new data provided by bad events. Our understanding of both fiscal policy and price dynamics are, I believe, greatly improved. And funerals will continue to feed intellectual progress: younger macroeconomists seem to me to be much more flexible and willing to listen to the data than their counterparts were, say, 20 years ago.

But the quick transformation of macroeconomics many hoped for almost surely isn’t about to happen, because events haven’t forced that kind of transformation. Many economists—myself included—are actually feeling pretty good about our basic understanding of macro. Many others, absent real-world catastrophe, feel free to take the blue pill and keep believing what they want to believe.

References

Eichengreen, B., and O’Rourke, K. (2009), ‘A Tale of Two Depressions’, Voxeu.com, 1 September.


Stagnant productivity and low unemployment: stuck in a Keynesian equilibrium

Wendy Carlin* and David Soskice**

Abstract: A major challenge is to build simple intuitive macroeconomic models for policy-makers and professional economists as well as students. A specific contemporary challenge is to account for the prolonged slow growth and stagnant productivity that has followed the post-financial crisis recession, along with low inflation despite low unemployment (notably in the UK). We set out a simple three-equation model, which extends the core model in our two recent books (Carlin and Soskice, 2006, 2015) to one with two equilibria and two associated macroeconomic policy regimes. One is the standard inflation-targeting policy regime with equilibrium associated with central bank inflation targeting through monetary policy. It is joined by a second, Keynesian policy regime and equilibrium, with a zero lower bound (ZLB) in the nominal interest rate and a ZLB in inflation in which only fiscal policy is effective (Ragot, 2015). Our approach is related to the Benigno and Fornaro (2016) Keynesian–Wicksellian model of growth with business cycles. It diverges from New Keynesian models because although we attribute model-consistent expectations to the policy-maker, we do not assume that these are the basis for inflation and growth expectations of workers and firms. We compare our approach to Ravn and Sterk’s related multiple equilibrium New Keynesian model (Ravn and Sterk, 2016).

Keywords: stagnation, liquidity trap, zero lower bound, strategic complementarity, multiple equilibria, inflation-targeting

JEL classification: E32, E43, E52, O42

I. Introduction

In this paper we set out a three-equation model with two alternate parts—one summarized by a vertical and the other by a horizontal long-run Phillips curve. The former is appropriate to ‘normal’ macroeconomic conditions and the latter to a world following a
financial crisis that occurred in a low-inflation environment. In doing so we borrow from some of the most interesting recent developments in macroeconomics that seek to analyse the long drawn-out post-financial crisis period of low growth and very low inflation by introducing heterogeneous agents, incomplete markets, and precautionary savings.

However, in contrast to many New Keynesian models, we pay attention to investment. In particular, we model investment as a multiple equilibrium strategic complementarities game in which uncertainty is the key to the choice of equilibrium. The simple two-part three-equation model with mechanisms that help account for recent experience builds on our earlier work on small tractable macroeconomic models targeted at policy-makers and students, and motivated by the need to understand contemporary macroeconomic problems.\(^1\)

The aim is to highlight mechanisms that could explain the very slow recovery from the financial crisis, which has specific and unusual features. Since the crisis, the UK, along with other high-income countries, has been characterized by very low growth of per capita GDP, low core inflation, and a real interest rate close to zero. The slow recovery has lasted longer than the overhang of the financial effects of the crisis, which suggests other mechanisms are at play in addition to the deflationary deleveraging implied by private- and public-sector debt overhang effects and the impairment of the financial system emphasized by, among others, Reinhart and Rogoff (2009, 2010). A further puzzling aspect of current macroeconomic performance is the persistence of low growth and low and stable inflation while unemployment has fallen to low levels, and in the UK and some other countries, the employment rate has reached historically high levels.

We set out the ‘Wicksellian–Keynesian’ model in section III. The model we develop has two equilibria and two associated macroeconomic policy regimes. The Wicksellian equilibrium is the familiar one in a model of an inflation-targeting central bank: inflation is at target, unemployment is at the equilibrium of the labour market, and the real interest rate is positive and at its ‘natural’ or ‘Wicksellian’ level. Of particular interest for current developments is the other—Keynesian—equilibrium. In this equilibrium, inflation is constant, but at zero. The nominal interest rate is at the zero lower bound (ZLB), also zero. The implication is that the economy is at a labour market equilibrium with the real interest rate equal to zero. The economy becomes trapped in a Keynesian equilibrium because a sufficiently large fall in aggregate demand in a low inflation environment takes it to the ZLB and renders stabilizing monetary policy ineffective. A floor on wage inflation at zero prevents a downward deflationary spiral and explains why the economy remains in the Keynesian equilibrium.

We introduce investment as a multiple equilibrium game of strategic complementarity in which investment is ‘trapped’ at a low level in the Keynesian equilibrium, and

---

\(^1\) In the first book (Carlin and Soskice, 1990), we introduced wage- and price-setting agents with the result that involuntary unemployment characterized the equilibrium of the labour market. The model was well-suited to analysing the trends and cross-country comparisons in unemployment across high-income countries that first emerged in the 1970s. In the second (Carlin and Soskice, 2006), the model was extended to introduce an inflation-targeting central bank and used to analyse policy-making by contemporary central banks. In the third (Carlin and Soskice, 2015), inflation-targeting was extended to the open economy. We also showed how the policy-maker’s intended equilibrium at target inflation and equilibrium unemployment could be subject to a financial cycle fuelled by a financial accelerator process based on the use of housing as collateral and on a value at risk model of investment bank behaviour.
together with a model of precautionary saving accounts for persistent low aggregate
demand. Consistent with models of endogenous growth, we model productivity growth
as a function of investment (productivity increases are embodied in the capital stock),
so that productivity growth is also at its ZLB in a Keynesian equilibrium with very low
net investment.

When the economy is in the Wicksellian regime, fluctuations in aggregate demand
and productivity shocks are stabilized by conventional inflation-targeting monetary
policy. In the Keynesian regime, monetary policy is ineffective. We discuss circum-
cstances under which fiscal policy can release the economy from the stagnation trap.

The paper is organized as follows. In section II, we summarize stylized facts of the
period since the global financial crisis, focusing on the UK but presenting data for the
US, Germany, and Japan as well. The basic two-part Keynesian–Wicksellian three-
equation model is set out in section III. In section IV, we explain the persistence of low
investment and stagnant productivity in the Keynesian equilibrium. In section V we
investigate the puzzle of a Keynesian equilibrium at high employment. Section VI pulls
together the microeconomic assumptions underlying the model so as to provide a com-
parison with two other small models—the standard New Keynesian model of Clarida
et al. (1999) and a multiple equilibrium New Keynesian model proposed by Ravn and
Sterk (2016).

II. Economic performance since the global financial crisis

The motivation for developing the simple model in this paper is the unusual combin-
ation of characteristics of economic performance experienced by major economies fol-
lowing the global financial crisis. The features are particularly pronounced in the UK,
but as the following charts illustrate, are present elsewhere. The charts show data for
the UK, US, Germany, and Japan. We date the post-crisis period from 2010 and it is
shaded in each figure.

Figure 1 shows long-run hourly productivity trends for the economy as a whole from
1973. In all four countries, productivity growth has fallen below 1 per cent per annum,
with that in the UK close to zero. Figure 2 charts the downward trend in the real inter-
est rate (on 10-year sovereign bonds) for the G-7 countries (weighted by GDP). The real
interest rate appears to have shifted down to a rate close to zero, after being between 2
and 4 per cent for the previous two decades.

Figure 3 presents data for the UK, US, Germany, and Japan on the growth in capital
services per hour worked over the period since 1985. There is a common, dramatic fall
in the rate of capital deepening across countries in the 2010–15 period. This applies to
Germany and Japan, which did not experience domestic leverage-based financial crises
in their economies in 2008–9, as well as to the US and UK.

From Figure 4 we see that, with the exception of Germany, real wage growth (in terms
of the consumer price index, CPI) has more or less ceased in the post-crisis period.

2 In order to concentrate in this paper on the Keynesian regime, we do not integrate into the two-equilib-
rium model the destabilizing financial cycle that can take the economy away from the Wicksellian equilibrium
as developed in ch. 6 of Carlin and Soskice (2015).
Figure 1: Trend hourly productivity growth (1973–2015)

Notes: Labour productivity, total economy, is defined as GDP per hour worked and its growth rate is calculated as first natural-log difference. Trends are estimated using the Hodrick-Prescott filter.

Figure 2: G7 real sovereign bond yield (1985–2016)

Notes: Real 10-year government bond yield, G7 weighted by GDP. Dashed lines show the minimum and the maximum.
Source: OECD.
Figure 3: Capital deepening: growth in capital services per hour worked (1985–2015)

Notes: Total economy, percentage change at annual rate and percentage of GDP.

Figure 4: Real compensation per hour worked, deflated by CPI (1995–2016)

Dataset: LFS.
Figure 5 shows that in the post-crisis period, developments in the labour markets of these economies are also notable: unemployment rates have fallen to levels below 6 per cent and employment rates are, with the exception of the US, historically high.

Finally, in spite of labour market tightness, nominal wage growth as shown in Figure 6 has been between −0.5 and 3 per cent per annum in all four countries in the post-crisis period. Core CPI inflation remains low (Figure 7), but with no deflationary spiral following the financial crisis.

To summarize, we need a model that will help to explain how an economy with an inflation-targeting central bank can exhibit the following features:
- virtual stagnation of productivity, capital services per hour and real wages;
- a real interest rate close to zero;
- low unemployment and high employment rates;
- low growth of nominal wages.

III. The two-part two equilibrium model

To provide a framework for interpreting the stylized facts of the period since the financial crisis presented in section II, we set out a small macro model with two equilibria—a
Keynesian one and a Wicksellian one. By including a second type of equilibrium, this model extends the closed economy three-equation model (Carlin and Soskice, 2005, 2006, 2015). We begin with an intuitive explanation of the two equilibria and then show how the model is represented graphically.

The standard components of the model are the IS curve, where aggregate demand is a function of the real interest rate, the Phillips curve derived from the interaction of aggregate demand and wage and price-setting (PC), and the central bank’s monetary policy reaction function (MR). Underlying the Phillips curve are the price-setting real wage curve, which is the real wage implied by profit-maximizing price setting, and the wage-setting real wage curve (resulting from wage-setting in a labour market with incomplete contracts). Equilibrium employment requires that the price-setting and wage-setting real wages are equal. At equilibrium, inflation is constant and there is involuntary unemployment. The central bank has an inflation target and implements its policy by affecting the real interest rate and hence aggregate demand. Finally, the equilibrium is characterized by a real interest rate determined by the IS relation at equilibrium employment.

A minimal amount of information about the microeconomic foundations of the model is required in order to understand the way it works. This is developed further in section VI. Firms produce differentiated products and set prices. In the standard case (Dixit and Stiglitz, 1977), consumers maximize CES utility functions implying a constant elasticity demand function, \( \varepsilon \), and constant returns with productivity \( \lambda \), hence the price-setting real wage is \( w_{PS} = \frac{\varepsilon - 1}{\varepsilon} \). This holds in both the Keynesian and Wicksellian parts of the model.

The wage-setting real wage differs in the two parts. In the Wicksellian case the wage-setting real wage \( w_{WS} \) can be seen as an explicit or implicit bargain or agreement in which \( w_{WS} \) is higher the higher the rate of employment, \( n \), and the higher is productivity \( \lambda \), scaled by \( \alpha \) as a measure of the bargaining effectiveness of the labour force. Thus \( w_{WS} = \alpha \lambda n \).

Hence, equilibrium employment in the Wicksellian case is given by

\[
\begin{align*}
    w_{W,c} &= w_{PS} = \frac{\varepsilon - 1}{\varepsilon} = w_{WS} = \alpha \lambda n_{W,c} \\
    \rightarrow n_{W,c} &= \frac{1}{\alpha \left( \frac{\varepsilon}{\varepsilon - 1} \right)}
\end{align*}
\]

(where the interpretation is that equilibrium employment is reduced by low product market competition and stronger labour market bargaining power \( \alpha \)).

In the Keynesian case the key assumption is that nominal wage inflation is bounded by a floor around zero, which captures the ability of employees to coordinate on opposition to nominal wage cuts. For simplicity, we assume the floor is at zero. The desire of wage-setters to avoid the negotiating costs involved in a rational expectations-based Phillips curve explains the widespread use of compensation for previous inflation in wage-setting, and we make that assumption, referring to the ‘compensation’ Phillips curve.

Equilibrium employment in the Keynesian state, absent nominal wage cuts, is that real wages remain at \( w_{WS} = w_{PS} = \frac{\varepsilon - 1}{\varepsilon} \) whenever \( n < n_{K,c} \). Hence \( n_{K,c} = n_e \), where \( n_e \) is the employment rate determined by the level of aggregate demand.
Thus there are two states of the economy with employment at equilibrium and two associated macroeconomic policy regimes. In one of the equilibria, the inflation rate is equal to the central bank’s inflation target. This is the policy-maker’s intended equilibrium. The real interest rate at the intended equilibrium is variously referred to as the stabilizing, natural, or Wicksellian interest rate. For this reason, we call this the Wicksellian equilibrium of the model.

In the second equilibrium, unemployment is higher than at the Wicksellian equilibrium. The real interest rate and inflation are both zero. Whereas employment in the Wicksellian equilibrium is determined in the labour market by the unique intersection of the wage- and price-setting curves, in the second equilibrium, it is determined by the level of aggregate demand. For this reason, it is called the Keynesian equilibrium of the model. Although determined by aggregate demand, inflation is constant at the equilibrium so it must also be a labour market equilibrium where wage- and price-setting real wages coincide.

Before setting out the model graphically, it is useful to contrast the role of aggregate demand in defining the Keynesian equilibrium with the way fluctuations in aggregate demand play out in the neighbourhood around the Wicksellian equilibrium. Beginning at a Wicksellian equilibrium with inflation at target, a negative shock to aggregate demand leads to a fall in employment and output. Given lags in policy-making, inflation falls and pulls forecast inflation below the target. The central bank takes the new compensation-based expectations-augmented Phillips curve as its constraint and optimizes by setting the interest rate (lower) using its monetary rule to boost aggregate demand and output. Inflation rises and the economy returns to equilibrium. If the shock to demand is permanent, the new equilibrium will have a lower stabilizing real interest rate, but all other characteristics of the equilibrium, including the unemployment rate, are unchanged.3

Now consider a large negative demand shock. Output, employment, and inflation fall, and the central bank seeks to stabilize as above. However, it is prevented by the ZLB on the nominal interest rate from creating the positive output gap required to get the economy on a path back to its intended equilibrium. Given the persistence of a negative output gap, inflation would be expected to fall, taking the economy further away from equilibrium. Falling inflation pushes up the real interest rate, which depresses aggregate demand further. The dynamics of the model suggest an unstable path with growing deflation and rising unemployment.4 But this does not match the data well. Indeed, there appears to be a floor to inflation close to zero. In Japan, inflation has fluctuated around zero for almost 20 years.

If we assume a ZLB not only to the nominal interest rate but also to inflation, then instead of a deflationary spiral, the economy will settle at equilibrium with the real and nominal interest rates, and inflation all equal to zero. Maximum employment will be determined by aggregate demand when the real interest rate is equal to zero. Once the economy is at the ZLB for wage inflation, then there is no negative bargaining gap (where the wage on the wage-setting curve is below that on the price-setting curve) because inflation cannot fall. This explains why there is a labour market equilibrium

---

3 This is set out in detail in Carlin and Soskice (2005) and in Carlin and Soskice (2015, ch. 3).
associated with low aggregate demand when the economy is at the ZLB for the nominal interest rate and inflation. As in the Wicksellian equilibrium, there is involuntary unemployment. Unlike the standard business cycle fluctuation away from the Wicksellian equilibrium discussed above that produces a rise in cyclical unemployment to which the central bank can successfully respond, the additional unemployment in the Keynesian equilibrium is ‘structural’.

We now set out the model in graphical form. There is more detail on the assumptions about the microeconomic mechanisms in section VI. In the lower panel of Figure 8 is the labour market diagram, with the wage- and price-setting curves. The intersection of the two solid curves determines the equilibrium employment level in the Wicksellian equilibrium at $n^{W,e}$, labelled $I$. This implies a vertical ‘long-run’ Phillips curve in the middle panel, indicating that there is no long-run trade-off between inflation and unemployment. The downward-sloping $MR_i$ curve maps the tangencies between ellipsoid central bank indifference curves centred on the equilibrium labelled $I$ and positively sloped ‘short-run’ expectations Phillips curves (not shown). It indicates the central bank’s best response to any deviation in inflation from target: if inflation rises above target, the central bank uses the $MR_i$ at that rate of inflation to identify the negative output gap with employment below $n^{W,e}$ that will put it on the path back to target

Figure 8: Wicksellian and Keynesian equilibria
inflation. The central bank uses the IS curve to calculate the change in the policy interest rate that will produce the required output gap. In equilibrium, the economy is at the intersection of the MR and PC curves. The real interest rate in equilibrium depends on the position of the IS curve: with IS₁, the stabilizing real interest rate is \( r^{W} \).

The regime in which there is a Keynesian rather than a Wicksellian equilibrium is defined by the ZLB in the nominal interest rate and in inflation. In Figure 8, the Keynesian equilibrium is labelled II with employment at \( n_{K} \) below the Wicksellian level. Note that IS₁ intersects the horizontal axis to the left of IS and hence at employment below the minimum Wicksellian equilibrium employment level, which is defined by a real interest rate of zero.

The contrast between the two regimes is shown in Figure 8:

- In the top panel, the IS₂ curve is to the left of the grey IS. The IS curve labelled IS marks the boundary between the Wicksellian and Keynesian regimes.
- In the middle panel, the ('long-run') PC is horizontal at the ZLB of inflation, \( ZLB_{\pi=0} \) and vertical in the Wicksellian regime.
- In the bottom panel, the wage-setting curve, \( w^{WS} = w_{t} \), is horizontal when \( ZLB_{\pi}=0 \). In the Wicksellian regime, the labour market equilibrium is at the intersection of the upward-sloping WS curve and the horizontal PS curve.

To show graphically how an economy gets stuck in a Keynesian equilibrium we look at the way the central bank’s ability to respond to a large negative output gap is constrained by the ZLB on the nominal interest rate.

In Figure 9, we show how the economy adjusts from an initial Wicksellian equilibrium, I, to the Keynesian equilibrium II. Initially, the negative demand shock leads to a fall in output and employment (A to B). From the lower panel, we can see that the ('short-run') Phillips curve for expected inflation at the target of 2 per cent implies deflation at the lower employment level (point C). However, by assumption, inflation cannot fall below zero, shown by the dashed part of the Phillips curve beneath the horizontal axis. Given expected inflation of zero, the central bank’s constraint is the Phillips curve indexed by \( \pi^{E}=0 \): the central bank wishes to shift the economy next period to point D on the MR curve. But this is not feasible because of the ZLB on the real interest rate: the economy will settle at point E with inflation, nominal interest rate, and real interest rate equal to zero. The economy is in a Keynesian equilibrium.

To conclude this section, we draw out the implications for macroeconomic policy in the two regimes. In the neighbourhood of the Wicksellian equilibrium, when the economy is disturbed by demand, supply, or inflation shocks, the central bank uses changes in the interest rate to alter aggregate demand so as to keep the economy close to the inflation target at equilibrium unemployment. Given the policy-maker’s objectives, monetary policy is both necessary and sufficient as stabilizer. The equilibrium fiscal multiplier is zero because equilibrium output and employment are determined by the supply side. Higher government spending will ceteris paribus result in a higher stabilizing real interest rate; otherwise, the equilibrium is unaffected.

5 This is set out in detail in Carlin and Soskice (2005) and in Carlin and Soskice (2015, chs 3 and 12).
Monetary policy is ineffective in the Keynesian equilibrium. An expansionary fiscal policy sufficient to raise inflation above the target (point C in Figure 10) will take the economy out of the Keynesian equilibrium and restore the role of monetary policy. With inflation above target, the central bank will raise the nominal interest rate to get the economy on to the $MR_1$ curve to the left of the Wicksellian equilibrium (at point D). The usual adjustment process of a falling interest rate (from point E) and rising employment will take the economy back to the Wicksellian equilibrium along the $MR_1$ curve at $n_{W,e}$ (Figure 10).

But as we will see in the next two sections, a revival of aggregate demand—either via fiscal policy or via consumption—may not be sufficient to release the economy from a Keynesian equilibrium, even if it is able to deliver low unemployment. To explain this we need to introduce productivity growth and its relationship to investment (in section IV) and in section V, to show how a Keynesian equilibrium can persist at low unemployment.

**IV. Low investment and stagnant productivity in the Keynesian equilibrium**

In this section we look at the role of aggregate demand, and in particular at investment and productivity growth. In our model of two equilibria, investment plays a major role. We come back to consumption at the end of the section.

As the data in section II make clear, a specific characteristic of the post-financial crisis period is that productivity growth has virtually ceased in the UK and is unusually
low in all four countries. It is also the case that the growth of capital services per worker is exceptionally low. Our hypothesis is that low productivity growth is a consequence of low investment (which entails a model in which productivity growth is endogenous), and that persistent low investment is the outcome of coordinated beliefs of firms around pessimistic expectations of market growth.

Caballero documents the lumpiness of investment in firm-level data, which is consistent with its sunk-cost character (Caballero, 1999; Bloom et al., 2007; Bloom, 2014). Because of the limited second-hand markets in capital goods, net investment decisions are largely irreversible and are made when management has strong and positive views about the future. In a Keynesian equilibrium, in the absence of strongly held optimistic expectations, firms invest at a low level because the potential loss if expectations are mistaken is low. The $i$th firm’s investment will depend on its degree of confidence in the expected growth of its market. Bloom (2014) argues that uncertainty is strongly positively related to recession.

The return to investment to each firm depends on aggregate output. Thus, firms are influenced by the actions of other firms and a natural way to model the problem is as a game of strategic complementarities among firms where there are two stable equilibria and an unstable one (Vives, 2005). The stable equilibria are a low investment level characterized by high uncertainty and pessimistic expectations, and a high investment level associated with high certainty and optimism.

In modelling expectations, we refer to Tuckett’s psychological behaviouralist approach where formally independent decision-makers are either in a (radically) uncertain state or coordinate on an optimistic state, which is believed in with high probability (Tuckett, 2012; Tuckett and Nikolic, 2017; King, 2016). We refer to this as the
Tuckett hypothesis. Without specifying exactly how optimistic expectations come about, individual decision-makers will adopt optimistic expectations simply based on shared optimism (perhaps airing a shared argument or narrative—‘this time is different’, for example). According to Tuckett, it is not that individual decision-makers look to see the weakness or lack of rationality of the shared optimism, so much as their concern to ensure the optimism is indeed shared. This then has the effect ex post that decision-makers do take the ‘right’ decisions jointly that enable the high investment equilibrium.

Bringing together the assumption that firm $i$’s expectations of its market growth depend on aggregate demand in the economy (consumption plus the investment of all other firms) with the Tuckett hypothesis about the role of uncertainty, provides a model of two stable investment equilibria in which the switch from one to the other occurs because of changes in beliefs in a setting of radical uncertainty. By radical uncertainty, we simply mean to signal the difference between a situation where probabilistic decisions are possible and one in which they are not (King, 2016).

This model is represented graphically in Figure 11. On the horizontal axis is aggregate investment per worker for all but the $i$th firm; on the vertical is investment by the $i$th firm. A symmetric equilibrium exists along the 45-degree line. The S-shaped function indicates that there are two stable equilibria (at low and high investment). At the low one, high uncertainty and pessimism keep investment in the rest of the economy—and therefore in the $i$th firm—low. At the high one, certainty and optimism generate high investment in all the $j$ firms—to which the $i$th firm’s best response is to invest at a high level.

From the geometry in Figure 11, when the S-shaped function cuts the 45-degree line from above, the equilibrium is stable: for example, the equilibria $(I, T)$ are both stable. When the function cuts the 45-degree line from below, the equilibrium is unstable: following a disturbance, investment will converge to either a low or a high equilibrium.

Vives’s strategic complementarities game captures the slope and existence conditions for the two stable and one unstable equilibria noted above, and nicely characterizes the
difference between stable Tuckett ‘pessimistic’ low expectations at low levels of investment, $I_{\text{Low}}$, and stable ‘optimistic’ high expectations with high investment, $\bar{I}$, with rapidly changing expectations in between as the $i$th and the other firms make major upwards expectational adjustments.

Using the diagram we can provide an account of how the economy initially trapped in the lower stable equilibrium, jumps to a high investment equilibrium. As uncertainty falls, the S-shaped function shifts up. If uncertainty continues to fall, the S-shaped function will eventually be tangential to the 45-degree line—a further fall will see the lower equilibrium disappear and investment will rise until the optimistic stable equilibrium is reached with investment at $\bar{I}$.

This picture is obviously greatly simplified. But it enables us to see how a move from a high-level coordinating investment-equilibrium can occur in ‘one jump’ to a low level equilibrium. This might plausibly be as a result of a financial crisis, either directly or indirectly. An historically significant event such as the financial crisis is a commonly observed signal to which many can respond with changing beliefs about macroeconomic uncertainty. And given the initial conditions of low inflation and the small gap between the prevailing interest rate and the ZLB, this in turn moves the economy from a Wicksellian to a Keynesian equilibrium.

Since it is typically the case that much new technology and hence productivity growth is embedded in new investment, a move to a low investment equilibrium is a likely explanation for a period of very low productivity growth. Productivity remains stagnant in the Keynesian equilibrium, with the lower bound of productivity growth at zero, because investment is at a very low level.

The market growth expectations of firms are affected not only by the investment decisions of other firms but also by consumption spending. Because of incomplete insurance and credit markets, workers in this model self-insure through precautionary savings to cover periods of unemployment rather than buying insurance or being able to borrow when unemployed. This reflects asymmetric information by banks (and insurance companies) in the face of idiosyncratic risk, and moral hazard in the face of aggregate unemployment, which is characteristic of the Wicksellian as well as of the Keynesian equilibrium.

Instead of saving primarily for ‘smoothing’ motivations (as in the standard Euler equation approach based on complete credit markets and insurable unemployment risk), many households are ‘hand to mouth’ consumers who spend what they earn. Rather than spending on the basis of future income expectations, spending is based on past income. In a recessionary environment household expenditure is based on (low) past income.

Moreover, if households are nervous about the possibility of future unemployment, financial institutions are unlikely to be prepared to lend to those who have become unemployed. Instead as Challe and Ragot argue, households have to self-insure rather than borrow; thus households save in response to recessionary fears (Challe and Ragot, 2016).

What our model shows is that there may be many Keynesian employment equilibria, below the unique Wicksellian equilibrium. As an example, take a simple aggregate demand equation, with ‘hand to mouth’ consumption, and profits saved, and with a low investment equilibrium given by $I = I_{\text{Low}}$, and no other component of aggregate demand. Consumption is equal to $\lambda \frac{\varepsilon - 1}{\varepsilon} n_d$, the price-setting real wage multiplied by $n_d$. Hence $\lambda n_d = \lambda \frac{\varepsilon - 1}{\varepsilon} n_d + I_{\text{Low}} \rightarrow n^K = n_d = \frac{\varepsilon}{\lambda} I_{\text{Low}}$ is a Keynesian equilibrium. If $I_{\text{Low}}$
changes because of coordinated changes in expectations, then a new Keynesian equilibrium emerges.

In a much more sophisticated model, Benigno and Fornaro’s paper on stagnation traps (2016) also sets out a dual Keynesian–Wicksellian framework with broadly similar assumptions. Their concern is to analyse growth and cycles together within an explicit Aghion–Howitt type endogenous growth framework. The major difference in the results is that in Benigno and Fornaro’s model both the Wicksellian and the Keynesian employment equilibria are unique, whereas we emphasize the range of possible Keynesian equilibria once the economy is in the stagnation trap.

In terms of our simplified approach, in the Keynesian part of the model, a Keynesian equilibrium is given by

$$n_{Ke}^K = \frac{\varepsilon}{\lambda} L_{Low}$$

and there may be many such equilibria reflecting different values of $L_{Low}$.

By contrast, in Benigno and Fornaro’s model, the low investment level is uniquely pinned down with the level of employment in the Keynesian equilibrium:

$$L_{Low} = L_{Low}(n^K).$$

In this case the expected value of the employment rate, $n^K$, will be equal—at least eventually—to the actual value of the employment rate. Hence

$$n_{Ke}^K = \frac{\varepsilon}{\lambda} L_{Low}(n^K)$$

so that there is a unique Keynesian equilibrium.

In both models, the Wicksellian equilibrium is unique, $n_{We}^W$, so that

$$n_{We}^W = \frac{\varepsilon}{\lambda} T_{High}(n_{We}^W, r^W).$$

Thus the major difference in the characterization of the stagnation trap between the Benigno and Fornaro model and the approach set out here is the uniqueness or non-uniqueness of the Keynesian equilibrium.

In both cases if productivity growth is embedded in investment, it is plausible to assume that (very) slow productivity growth is associated with the Keynesian (stagnant) equilibrium. And a jump up to a Wicksellian equilibrium, again with productivity growth embedded in investment, now implies a jump to a higher level of productivity growth and in consequence growth in real wages.

While our approach is less sophisticated than that of Benigno and Fornaro, and while the latter usefully integrates stagnation traps with endogenous growth theory, the broadly similar result is a contrast between perhaps long-lasting periods of unemployment with very slow productivity growth and moves (which may be quite rapid) to Wicksellian equilibria with more normal productivity growth and relatively high employment rates. Wicksellian equilibria may also be long-lasting (although subject to destabilizing financial cycles).

V. A Keynesian equilibrium at high employment

The two-equilibrium model can account for a lengthy period during which monetary policy is paralysed and unemployment is high. Yet the data in section II suggest that although several features of the Keynesian equilibrium continue to persist in the UK and elsewhere (very low real interest rates, low inflation, and ineffective monetary policy), unemployment rates are now low and employment rates are at historically high levels.

One hypothesis for the UK’s performance consistent with the model is that the weakness of investment persists, which is reflected in Figures 1 (productivity) and 3 (capital
deepening), but that aggregate demand has revived through consumption spending. Although real wages have not been rising, households have had access to cheap credit, especially for cars and consumer durables, in the low interest rate environment. This has shifted the IS curve to the right.

Increased aggregate demand in the context of zero productivity growth (a consequence of the continued weakness of investment) implies higher employment. According to Figure 5, the employment rate is higher than it was prior to the financial crisis (in three of the four countries). We need to explain how employment rose above the pre-crisis Wicksellian equilibrium level without triggering inflation—and, therefore, to explain how a Keynesian equilibrium with the real interest rate and inflation close to zero can co-exist with low unemployment.

In graphical terms, this requires a downward shift in the wage-setting curve, which in the context of the nominal interest rate at the ZLB and wage inflation at zero, implies the extension of the flat part of the wage-setting curve to the level of employment consistent with a new lower unemployment rate at a Wicksellian equilibrium. Figure 12 illustrates. The new Wicksellian equilibrium is marked by I and the new Keynesian equilibrium at low unemployment where the economy is trapped is labelled II.

The graphical illustration of the high employment Keynesian equilibrium begs the question of why the wage-setting curve should have shifted downwards. Anything that weakens the reservation position of employees shifts the curve down: lower unemployment benefits, less legal protection for workers, weaker unions, and increased labour supply (for example, as a result of migration) can all increase the expected cost of job loss. New forms of contracts, including the so-called zero hours contracts in the UK, have increased job insecurity. All of these developments, some of which are associated with the continued shift of employment out of industry and with new forms of work organization facilitated by new technology, reduce the wage the employer needs to pay in order to secure adequate worker effort at any given unemployment rate.

Figure 12: A Keynesian equilibrium at low unemployment
VI. Microeconomic mechanisms underlying the two-equilibrium model: comparison with standard New Keynesian and heterogeneous agent (HANK) models

In this section we contrast the mechanisms sustaining the Keynesian and Wicksellian equilibria in the small model we have set out with those in two other models. The first is the widely cited three-equation representative agent New Keynesian model of Clarida et al. (1999) and the second is a new small model motivated by the empirical issues of the post-crisis stagnation (section II), which is called a heterogeneous agent New Keynesian model, HANK (Ravn and Sterk, 2016). All three models can be presented graphically using IS, PC, and MR-type curves, although as we shall see, the underpinnings of the curves reflect quite different assumptions about the behaviour of the economic actors. This comparison will help to clarify that although ours is a heterogeneous agent model with multiple equilibria like Ravn and Sterk’s model, it is not ‘New Keynesian’. We call it HAWK for short—heterogeneous agent, Wicksellian-Keynesian model. Neither the Clarida et al. nor the Ravn and Sterk model includes investment.

Unlike real business cycle and New Keynesian models, our model does not impose model-consistent expectations on all agents. Model-consistency of expectations implies that it is common knowledge that all relevant agents have full knowledge of the whole model and know that everyone else does. In particular, we part company with the New Keynesian modelling of the Phillips curve. The requirement of model-consistency of expectations combined with sticky prices implies that with a New Keynesian Phillips curve new information leads to jumps in inflation, which are counter-factual.6 Some agents—such as those in the foreign exchange market and Central Banks—do form model-consistent expectations given the expectations formed by others.7

Since the actors in our model set wages, prices, output, and interest rates, we require that a relation in the model describing how x takes decisions such as setting wages, or prices, or output, or the interest rate should be comprehensible to the actor in question. Few price-setters would understand Calvo price-setting, let alone see this as approximating how prices are in fact set. The requirement in NK DSGE (New Keynesian dynamic stochastic general equilibrium) models for wage and price variables to jump on to a perfect foresight path runs counter to our requirement for comprehensibility of actions to actors. Whereas in a price-taking model, comprehensibility is not relevant, it is important in a model where the strategies of actors are central.

As in our previous work, the microfoundations of the modelling in our two-equilibrium HAWK model are post-Walrasian in the sense of Stiglitz (1993), Bowles and Gintis (2000), and Howitt (2006). Principal–agent relationships arising from asymmetric information problems characterize interactions in the labour market between workers and employers, and in the credit market between borrowers and lenders. A macroeconomic model defined by principal–agent problems is by definition a heterogeneous

---

6 As an example of the concern of policy-makers, the brief for the Bank of England’s 2016 review of its in-house New Keynesian model says: ‘More generally on expectations, the rational expectations assumption within COMPASS produces responses that too often appear implausible (large and front-loaded).’ They call for ‘an alternative expectations scheme that was well tested, well understood and gave rise to plausible impulse responses’.

7 See ch. 16 of Carlin and Soskice (2015) for a detailed comparison of the CS model and the New Keynesian dynamic stochastic general equilibrium model.
agent model and, because information problems make it impossible to write complete labour or credit market contracts, there is no frictionless benchmark. Unemployment and credit rationing are present in any equilibrium.

Before drawing the contrasts with the two New Keynesian models, we summarize the assumptions of our model.

(i) Key assumptions

The model is of a closed economy. The ten key features of our HAWK model (some shared by Clarida et al. and Ravn and Sterk as explained below) are that:

i. in the Blanchard–Kiyotaki (1987) tradition, there are Keynesian microfoundations with firms setting prices and producing what is demanded at that price based on Dixit–Stiglitz (1977) monopolistic competition; this applies to the Wicksellian as well as the Keynesian part of the model (as in Clarida et al. and Ravn and Sterk);

ii. it is a heterogeneous agent model, with employed and unemployed workers; and there is a distinct class of equity owners of firms, so that the effect of income inequality can be understood; we assume all profits are saved along Kalecki–Pasinetti lines; (as in Ravn and Sterk);

iii. with the exception of the Central Bank, we assume agents do not form model-consistent expectations about the model as a whole; this is on empirical grounds and distinguishes the model from New Keynesian ones;

iv. asymmetric information and moral hazard imply (a) incomplete loan and insurance markets—generating credit constraints and precautionary savings (as in Ravn and Sterk and Challe and Ragot (2016)); and (b) involuntary unemployment at any labour market equilibrium;

v. there are no constraints on the timing of price-setting, but incomplete contracts and 'influence costs' (Milgrom and Roberts, 1990) concerning employees' productivity contributions make periodic wage-setting efficient; this is moreover overwhelmingly the case in practice;

vi. because model-consistent expectations are ruled out, the desire to avoid the negotiating costs involved in an expectations-based Phillips curve explains the widespread use of compensation for previous inflation in wage-setting, and we make that assumption, referring to the compensation Phillips curve;

vii. in line with much empirical evidence there is de facto (something like) a ZLB in nominal wage inflation in wage-setting; a simple analytic explanation may be that nominal wage-cutting is often highly visible across groups of employees: it thus constitutes a focal point in a coordination game in which employees can choose whether or not they feel well-disposed towards the employer in their effort decision; in any case we assume that wage-inflation has a ZLB;

viii. we assume that investment decisions by firms are actions in a multiple equilibrium dynamic game of strategic complementarities with stable optimistic and pessimistic choices (Vives, 2005; Caballero, 1999);

ix. while oversimplified, we assume productivity is largely embodied in the capital stock so that low investment implies static productivity and, given the profit mark-up, static real wages; there is no exogenous technical progress;
x. Given the presence of an inflation-targeting central bank (which imposes a ceiling on inflation), we assume a ZLB on the nominal rate of interest; with a ZLB on nominal wage and price inflation, there is a non-negative lower bound on the real rate of interest.

The two parts of the model can be summarized in the following equations where the superscript $ps$ refers to precautionary savings:

1. The Keynesian part:

\[
W = \frac{\lambda (\varepsilon - 1)}{\varepsilon} = w^{R,e} \\
\pi = 0 \\
r = 0 \\
c = nh(1 - s^{nw}(n)) \\
I = I_{Low} \\
y = c + I + g \\
n = y / \lambda \\
\dot{\lambda} = 0 \\
n^{K,e} = n < \frac{\varepsilon - 1}{\alpha \varepsilon}
\]

Note that, ignoring precautionary savings:

\[
y = wn_d + I_{Low} + g \\
\rightarrow \lambda n_d = \frac{\lambda}{\varepsilon} n_d + I_{Low} + g \\
\rightarrow n^{K,e} = n_d = \left( \frac{\varepsilon}{\lambda} \right) (I_{Low} + g) \\
\rightarrow y = \frac{\lambda}{\lambda} n_d = \varepsilon (I_{Low} + g) \\
\rightarrow \Delta y = \varepsilon \Delta (I_{Low} + g)
\]

which takes us from one Keynesian equilibrium to another.

2. The Wicksellian part:

\[
W = \frac{\lambda (\varepsilon - 1)}{\varepsilon} = w^{W,e} \\
\pi = \pi_1 + \alpha (n - n^{W,e}) \geq 0 \\
r = r^{W} + \tau (\pi - \pi') \geq 0 \\
c = wn \\
y = c + T_{High}(r) + g \\
n = y / \lambda \\
n^{W,e} = n < \frac{\varepsilon - 1}{\alpha \varepsilon}
\]
The equilibrium multiplier is zero since \( y^{w,e} = \lambda n^{w,e} \).

**(ii) Comparison with the New Keynesian models**

The representative agent New Keynesian model (Clarida *et al.*, 1999) has a transparent modelling of central bank behaviour that relates to how central banks describe their deliberations and actions.\(^8\) However, when subjected to shocks, the model incorporates the unappealing ‘jumping’ behaviour of inflation characteristic of the New Keynesian Phillips curve.\(^9\) In the Clarida *et al.* model, there is a single equilibrium (type I) at the policy-maker’s inflation target so it is not a useful model for addressing the problems of a low growth, low inflation equilibrium. By assuming a single representative agent, it omits key features of the world that we are seeking to understand.

In response to these challenges, in the last few years models with heterogeneous agents have been developed (e.g. Kaplan *et al.* (2016) and the many papers cited there, and by Ravn and Sterk, 2016, 2017). In many such models consumption behaviour varies according to the employment status of the agent. This is often tied in with an extension to the search and matching labour market model (SAM) in which the assumption that there are complete insurance markets for unemployment risk is dropped, with the result that unemployment becomes a major concern.

In these models, new mechanisms can lead to large fluctuations in aggregate demand and employment. They emerge from the joint presence of a precautionary savings motive, unemployment risk, and nominal rigidities:

intuitively, a wave of pessimism among households about their employment prospects could be self-fulfilling as the increased desire to build precautionary savings reduces aggregate demand, causing firms to hire fewer workers when prices are sticky and stabilization policy is insufficiently responsive. (Ravn and Sterk, 2016, p. 2).

When these features are brought into a typical quantitative NK DSGE model, it can only be solved numerically. From the perspective of a policy-maker, this is problematic because the modeller is not able to make the causal mechanisms and policy transmission channels transparent, and the presence and nature of multiple equilibria cannot be rigorously analysed (Ravn and Sterk, 2016).

To deal with these difficulties, Ravn and Sterk (2016) set up a small-scale analytical version of a HANK–SAM model with incomplete markets, search and matching frictions, and nominal rigidities. It is a model with three agents/classes in equilibrium: unemployed, employed, and owners/entrepreneurs. Their paper delivers a model that can be represented graphically and the qualitative predictions compared with the model of Clarida *et al.* (1999). Neither model includes investment.

In Ravn and Sterk’s model there are three equilibria—the central bank’s intended equilibrium, a second, liquidity trap equilibrium at the ZLB, and a third ‘unemployment trap’ equilibrium with low aggregate demand, high unemployment, and low

\(^8\) For an example, see Mark Carney’s speech January 2017 at the LSE entitled ‘Lambda’ (Carney, 2017).

\(^9\) The characteristics of the New Keynesian Phillips curve are explained in ch. 16 of Carlin and Soskice (2015).
inflation, but at an interest rate above the ZLB. Positive feedbacks between unemploy-
ment and precautionary savings reinforce the unemployment trap.

We simplify the Ravn–Sterk model by concentrating on the HANK element, abstract-
ing from the details of the search and matching model of the labour market. We explain
how a simplified version of the Ravn–Sterk model works and then compare it with the
two-equilibrium HAWK model presented earlier in this paper. This helps to bring out
important features of the two models as they relate to the modelling of consumption,
the Phillips curve, and the monetary policy rule.

The Ravn–Sterk model is a major departure from the benchmark representative
agent New Keynesian model with its unique equilibrium at target inflation and labour
market clearing. Ravn and Sterk’s model has an inflation-targeting central bank, but
there is involuntary unemployment at the intended equilibrium, and there are ‘belief-
driven’ fluctuations so the economy can become trapped in equilibria characterized by
very low inflation and high unemployment.

The motivation for Ravn–Sterk’s model is very similar to that for the two-equilibrium
model presented above: to match the post-crisis reality where low inflation and high
unemployment persisted for a number of years. And, as we shall see, the interaction
between the risk of unemployment and the motive for precautionary saving plays a role
in both models (there is no investment in their model). In their words:

The unemployment trap that can arise . . . offers an alternative perspective of
secular stagnation which ties together low real interest rates, high unemploy-
ment and low activity. Moreover, because of the low job-finding rate, there is a
strong incentive for precautionary savings which drives down the real interest
rate. Intriguingly, the unemployment trap can occur in our model purely because
of expectations and thus does not rely on sudden changes in population growth,
technological progress or financial tightening. (Ravn and Sterk, 2016, p.17)

The Ravn and Sterk model is shown in the figure below, where a measure of labour
market tightness is on the horizontal axis and the inflation rate is on the vertical. The
diagram captures the three equations in the model (the ‘IS’, the ‘PC’, and the ‘MR’).
Ravn and Sterk combine the IS and monetary rule into a single equation called EE.

We compare Figure 13 with Figure 8, which shows our two-equilibrium
Wicksellian–Keynesian model.

The first observation is that there are multiple equilibria in both models where the
PC and EE (or MR) curves intersect. In each model, the equilibrium intended by the
policy-maker is labelled I. The Clarida et al. model has a single equilibrium, which
is the one intended by the policy-maker. But in both Ravn and Sterk and our model,
there is a second equilibrium labelled II at which monetary policy is ineffective when
the interest rate is at the ZLB. The Ravn and Sterk model has a third equilibrium, III,
which is called the unemployment trap, where inflation is low but the economy is not
constrained by the ZLB on the nominal interest rate.

Looking at the geometry of Figures 8 and 13, it is evident that the structural rela-
tionships summarized in the Phillips curve and the EE (Monetary policy / IS) curve
that produce the two equilibria (I and II) are very different. When comparing the two
figures, the striking differences are in the slopes of the PC and EE curves. Notice first
that the Phillips Curve in the Ravn–Sterk model is very flat—it must be flatter than the
EE curve to produce the third unemployment trap equilibrium (III), and second, that
the EE curve is positively sloped in Ravn–Sterk and negatively sloped in our HAWK model. In Ravn and Sterk's model, the EE curve is negatively sloped only in the special case where the economy is at the ZLB.

The models have similar objectives (tractability, transparent mechanisms, relevance to contemporary conditions) and superficially similar results: an equilibrium at the policy-maker’s intended outcome and a ZLB equilibrium. Both incorporate an inflation-targeting policy-maker and heterogeneous agents so as to include precautionary savings, uninsured unemployment risk, and feedback at high unemployment between precautionary savings and unemployment. Yet the nature of the key relationships tying together the supply side (PC), the demand side (IS), and the policy-maker (MR/EE) are quite different as signalled by the contrasting geometry in the figures.

We next provide some intuition behind the differently sloped Phillips and EE/MR curves in the two models. In the online Appendix, we set out the equations of a simplified version of the Ravn and Sterk model in order to explain in more detail the sources of the differences.

Three key modelling choices lie behind the differences between the Phillips curve and EE/MR relations in the Ravn and Sterk and our HAWK model illustrated in Figures 8 and 13.

Ravn and Sterk assume:
- high price adjustment costs in equilibrium, which is necessary to flatten the ‘long run’ Phillips Curve and permit the third equilibrium to arise;
- there is no instrument for precautionary saving (in spite of the presence of the precautionary motive for saving), which makes the EE curve positively sloped; and
- a model of central bank behaviour without a specific inflation target. If there is an inflation target, then the third ‘unemployment trap’ equilibrium disappears.
The relatively flat equilibrium Phillips curve

The reason for the upward-sloping PC in the Ravn and Sterk model is their assumption that more rapid adjustment of prices is costly. As we show in the Appendix, dropping the terms in the Rotemberg price adjustment cost from the Ravn–Sterk first order condition produces a wage-setting real wage (as a function of $n$) equal to the price-setting real wage at a unique degree of labour market tightness, just as in our HAWK model. This results in a vertical PC relation (adding inflation adjustment costs to the HAWK model, similarly produces an upward-sloping PC).

In the steady-state equilibrium, the rise in inflation permits the real wage to increase and the exactly corresponding fall in the real profit mark-up enables the competing claims on output of wages and profits to be met and thus preserves constant inflation equilibrium in the conditions of a tighter labour market. It is not clear what justifies the assumption of price adjustment costs following wage changes, and certainly why they should be so high as to flatten the Phillips curve to the degree necessary.

The positively sloped EE curve

This curve incorporates the IS relation along with the central bank’s monetary rule. Recall that in our model, the equivalent to this curve is downward sloping: when inflation is judged too high by the central bank, it will tighten monetary policy by raising the interest rate according to the IS relation in order to create a negative output gap so as to reduce inflation. In the Ravn and Sterk model, the curve is upward-sloping. This means that although higher inflation makes the policy-maker want to reduce the output gap in order to dampen inflation (by raising the interest rate in the normal way), the expectation of this policy response will only be consistent with the consumption behaviour of employed workers if employment and the wage are in fact higher.

To explain this, we need to look more closely at an assumption that appears to be made for technical reasons. We find it has strong implications (the argument is set out using equations in the Appendix). The assumption is that although employed workers have a precautionary savings motive, they cannot implement it because there is no savings instrument. Ruling out saving implies that those in employment must be so-called ‘hand to mouth’, which means that their consumption each period is equal to their wage. And yet, employed workers are assumed also to satisfy the Euler equation, which means they are making their optimal consumption choice. To bring their desired consumption dictated by the Euler condition into line with the wage (hand to mouth consumption) requires that when inflation is higher and the central bank is expected to raise the interest rate, employment and the wage must be higher. This produces combinations of higher inflation and higher employment that satisfy both the IS relation and the monetary rule. This is not an intuitively appealing mechanism.

The third ‘unemployment trap’ equilibrium

In the simplified version of the Ravn and Sterk model set out in the Appendix, we find that when central bank behaviour is adjusted to bring it into line with the more standard Clarida et al. approach with an inflation target chosen by the policy-maker, the specific equilibrium that their paper prioritizes (the unemployment trap, III) disappears because the EE curve becomes downward-sloping as it is in our HAWK model.
VII. Conclusion

We set out a small analytical model that has recognizable central-bank and private-sector behaviour and includes heterogeneous agents and models of consumption and investment that are consistent with two equilibria: the intended one and the one at the ZLB. The ZLB equilibrium is characterized by a lower bound of zero on wage inflation and productivity growth as well as on the nominal interest rate.

This small model provides a way of accounting for the persistence of the long post-crisis stagnation—without recovery to the neighbourhood of the normal Wicksellian equilibrium nor a destabilizing deflationary process. Getting stuck in the stable Keynesian regime is due to the combination of low net investment (keeping productivity growth close to the ZLB), low aggregate demand (rendering monetary policy impotent because of the ZLB on the nominal interest rate), and the functioning of the labour market that sets a lower bound at zero on wage inflation, and prevents a deflationary spiral.

We attach particular importance to investment as a key driver, with investment modelled as a game of strategic complementarities and low and high coordinating equilibria. Once in a low equilibrium, the economy can get caught in a prolonged period of stagnation. In addition, productivity growth is embedded at least in part in investment: hence investment-induced stagnation can tie down productivity growth to very low levels.

In the context of this model, the more recent phenomenon, most starkly present in the UK, of the combination of low and stable inflation, a very low real interest rate, and productivity growth close to zero with low unemployment can be accounted for by developments on both the supply and demand side. In particular, it is consistent with, on the one hand, a weakening of the bargaining power of labour (due to policy, organizational and technological changes), and on the other, to a revival of consumer demand based on the extension of consumer credit in the conditions of low interest rates. Weaker worker bargaining power implies lower unemployment at a new Wicksellian equilibrium and extends the zone of the Keynesian ZLB trap to higher levels of employment, without upward pressure on inflation. More buoyant consumer demand under the ZLB conditions expands employment. However, without a shift to the optimistic scenario about future growth, investment and productivity growth do not revive.

In the Wicksellian regime, monetary policy is effective in stabilizing supply, demand, and inflation shocks—additional policies are required to prevent a financial cycle-induced crisis. In the Keynesian regime, monetary policy is ineffective in restoring employment to the Wicksellian equilibrium. Fiscal policy may be effective in enabling the economy to escape the stagnation trap, but for a return to normal conditions of positive productivity growth, a reduction of uncertainty about expected market growth is necessary to revive investment and, with it, productivity and real wage growth.

References


Macro needs micro

Fabio Ghironi*

Abstract: An emerging consensus on the future of macroeconomics views the incorporation of a role for financial intermediation, labour market frictions, and household heterogeneity in the presence of uninsurable unemployment risk as key needed extensions to the benchmark macro framework. I argue that this is welcome, but not sufficient for macro—and international macro—to tackle the menu of issues that have been facing policy-makers since the recent global crisis. For this purpose, macro needs more micro than the benchmark set-up has been incorporating so far. Specifically, artificial separations between business cycle analysis, the study of stabilization policies, and growth macro, as well as between international macroeconomics and international trade, must be overcome. I review selected literature contributions that took steps in this direction; outline a number of important, promising directions for future research; and discuss methodological issues in the development of this agenda.

Keywords: DSGE, heterogeneous firms, macroeconomic policy, market entry and exit, monopoly power, producer dynamics, reallocation, structural reforms, trade

JEL classification: E10, E32, E52, F12, F23, F40

I. Introduction

The global financial crisis (GFC) of 2008 and the Great Recession (GR) that followed have prompted a re-examination of the toolkit for macroeconomic analysis that should be used as benchmark for teaching, research, and policy advice.

The emerging consensus is that the new benchmark should make it possible to explain why the crisis happened, why the recovery was extremely slow, and the connection between cyclical dynamics and longer-run growth. The consensus is also that benchmark macroeconomics can no longer afford to abstract from such features of reality as heterogeneity across agents, uninsurable risk, and unemployment. The consensus points toward a framework that includes endogenous capital formation, a role for the financial sector, and...
a merger of the HANK and SAM frameworks.\(^1\) While I agree with the consensus and the promise of the HANK–SAM marriage, I think focusing exclusively on introducing financial intermediation and heterogeneous-agent unemployment in the benchmark macro framework is not sufficient.\(^2\) I believe that even those major changes to the standard New Keynesian toolbox would leave it unable to capture other mechanisms of first-order importance for macroeconomics. The field would continue being perceived as incapable of confronting reality (regardless of whether this perception is entirely fair or not).

Specifically, it seems to me high time for macroeconomics to move beyond the representation of firm behaviour in terms of production by a constant number of symmetric firms that produce either the same good under perfect competition or a fixed range of goods under monopolistic competition between a continuum of firms. Unemployment in the aftermath of the GFC happened also because a large number of firms failed and firms that did not fail reduced the number of active production lines. Credit market freezing was central to firm failures and decisions to cut production lines. Heterogeneous efficiency across firms implied that only the most efficient producers were able to survive, but their activity was slowed down by stagnant demand. Exposure to trade became the culprit for job losses that were most often caused by technological advances and/or by labour market rigidities that prevented effective reallocation of labour across firms, sectors, and geographical areas.\(^3\) Understanding the very slow speed of recovery since the crisis and the connection between a cyclical phenomenon (such as the GR) and longer-term dynamics requires us to understand also the slow-down of US business dynamism that happened in the last several years. Less firms are being created, and firms grow more slowly than they used to.\(^4\) Reduced firm entry and creation of new product lines along the recovery path generate hysteresis effects with long-run output consequences. In this light, it is no surprise that the US and other economies have been struggling with low productivity growth.

I strongly believe that, if we want to provide students and young researchers with a set of macro tools that can address the most important questions of the last decade—and those likely to arise in the future—macro needs micro (MNM—probably the sweetest acronym you will ever see in macroeconomics): The standard toolkit for macroeconomic analysis of fluctuations and policy must be extended to include producer-level dynamics of entry and exit, heterogeneity across firms, and the implications of these dynamics and heterogeneity for the macroeconomy.

This does not require the creation of a completely new set of tools: endogenous producer entry is a standard feature of endogenous growth models (for instance, Romer, 1990), firm heterogeneity and its implications for domestic versus foreign market entry and exit are standard features of trade theory since Melitz (2003). What needs to be recognized is that market entry and exit do not matter only for the long-run growth in the

\(^1\) HANK stands for heterogeneous-agent New Keynesian, as in Kaplan et al. (2016), and SAM stands for search-and-matching, as in the models of unemployment that build on Diamond (1982a,b) and Mortensen and Pissarides (1994). Ravn and Sterk (2016) provide an example of tractable ‘marriage’ of HANK and SAM.

\(^2\) This is independent of whether HANK–SAM and financial intermediation are included in closed- or open-economy macro models.

\(^3\) The fact that only large firms tended to survive made it easier to move toward a situation of monopsony in the labour market, with firms implementing contractual arrangements (such as non-compete clauses) that contributed to increasing labour market rigidity.

\(^4\) See Economic Innovation Group (2017) and Sparshott (2016a,b,c).
absence of uncertainty that growth economists usually focus on; domestic and foreign market entry and exit by heterogeneous producers do not matter only in the steady-state, balanced-trade environment most trade economists restrict their attention to. Growth is the result of entry and exit decisions that are taken under uncertainty during the business cycle. These decisions will contribute to shaping the cycle, and longer-run growth will be affected by cyclical dynamics through hysteresis effects. Heterogeneity will crucially affect the allocation of resources across producers and aggregate productivity. It will be among the determinants of what the economy trades and how it responds to foreign competition.\(^5\)

If macroeconomics aims to address the dynamics of the last decade and the economic issues that have been central to recent political outcomes, artificial separations between modelling of business cycles and longer-term dynamics must be abandoned, and the same must happen to similarly artificial separations between macroeconomic and trade modelling.\(^6\)

A growing literature has made significant inroads into the development of the type of framework I am suggesting and has built a strong case for its empirical relevance. Some results of this literature have begun informing policy advice in important ways. I briefly review the state of the art in this area below, focusing on selected contributions to model development. This literature provides the foundation for extensions of the framework in a number of directions.

The rest of the paper is organized as follows. Section II summarizes what the literature has already accomplished by briefly describing the main ingredients of some representative, existing models and their key results. Section III describes what I consider the most important directions for future research. Section IV discusses methodological issues. Section V concludes.

### II. The state of the art

This section reviews selected contributions to the state of the art in macro and international macro theory in which producer-level dynamics contribute to fluctuations. I focus on models that assume monopolistic competition or other forms of monopoly power, as they lend themselves most directly to providing the foundation for sticky-price extensions. The set of papers I mention is by no means intended to be a complete survey of the existing literature, and it includes much of my own work. It is the set of papers that allows me most transparently to describe how producer-level dynamics can

---

\(^5\) Absence of entry and exit dynamics from the foundation of the basic New Keynesian framework is the result of the fact that monopoly power, but not the free entry condition, was necessary as a stepping stone to introduce price stickiness in the model. This contributed crucially to the separation between (New Keynesian) business cycle macro, growth macro, and trade theory. According to Feenstra (2003), a constant number of firms ‘violates the spirit of monopolistic competition’.

\(^6\) I would argue that this should be true also from the perspective of trade research—i.e. that trade economists should move beyond the fiction of balanced trade and the steady-state focus of their models, and they should recognize that cyclical dynamics can have effects that seal the fate of the most major trade policy decisions.
be integrated in models of fluctuations, and to connect this to present-day questions of interest.\(^7\)

(i) Closed economy

Bilbiie, Ghironi, and Melitz (2012) provide a benchmark model of fluctuations with monopolistic competition and endogenous producer entry subject to sunk costs. The model—referred to as the BGM model below—assumes that consumers derive utility from having access to a larger set of products, but the existence of entry costs implies that only a subset of the products consumers would like to have access to is actually available at each point in time. The consumption aggregator is not restricted to the familiar Dixit–Stiglitz (1977) specification, but it takes a general homothetic form. Different from earlier business cycle models with monopolistic competition and endogenous entry that assume fixed entry costs and a free entry condition that implies zero profits on a period-by-period basis, the BGM model assumes sunk costs and a time-to-build lag. Entrants spend the first period setting up their production lines, and they begin producing and generating profits only in the following period. Free entry then equates today’s sunk entry cost (which requires use of labour) to the expected present discounted value (EPDV) of profits from tomorrow to the infinite future, with discounting adjusted for an exogenous probability of firm destruction. Formally, the entry condition in the symmetric equilibrium of the BGM model is 

\[ v_t = \left(\frac{w_t}{Z_t}\right)f_{E,t}, \]

where \( v_t \) is the EPDV of profits from \( t + 1 \) on, \( f_{E,t} \) is the sunk entry cost (in units of effective labour), \( w_t \) is the real wage (in units of consumption), and \( Z_t \) is exogenous aggregate labour productivity.\(^8\)

Firms finance their entry costs by issuing shares in the stock market, and this provides the general equilibrium link between entry decisions and the optimizing behaviour of the representative household. In this model economy, investment takes the form of creation of new production lines, financed by households with their savings. The price of investment is determined by the Euler equation for share holdings. With separable, log-utility from consumption:

\[ v_t = \beta(1-\delta)E_t \left[ \left( \frac{C_t}{C_{t+1}} \right) \left( v_{t+1} + d_{t+1} \right) \right] \]

\(^7\) By focusing on models in which firms have monopoly power, I completely omit the vast literature that builds on Hopenhayn (1992). For more references and discussion than I can cover here, see the papers I mention and references therein. For the open economy case, an extensive reading list is available at http://faculty.washington.edu/ghiro/ITMSyllabus.pdf.

\(^8\) The presentation of the model assumes a one-to-one identification between a producer, a product line, and a firm, and I use these terms interchangeably below. This was to facilitate relating our model to the New Keynesian literature, where individual producers in the usually assumed Dixit–Stiglitz continuum are referred to as firms. However, our preferred interpretation—consistent with relative empirical importance—is that every profit-maximizing unit should be interpreted as a product line at a possibly multi-product firm whose boundaries we are leaving unspecified by exploiting continuity and the assumption that firms remain of negligible size relative to the size of the market.
Where $\beta \in (0, 1)$ is the familiar discount factor parameter, $\delta \in (0, 1)$ is the exogenous probability of firm destruction that applies to all firms (including new entrants) at the end of each period, $C_t$ is consumption, and $d_t$ denotes firm profits, distributed to households as dividends. Forward iteration of this equation in the absence of bubbles returns the expression for the EPDV of profits in the free entry condition. Aggregate accounting implies the standard equality between aggregate demand—the sum of consumption and investment (the price of shares times the number of new entrants, $N_{E,t}$)—and income (the sum of labour income and profits generated by the number $N_t$ of producing firms): $C_t + v_t N_{E,t} = w_t L_t + N_t d_t$, where $L_t$ is the amount of labour employed by the economy. The price of shares is the key, endogenously determined relative price that determines the allocation of resources between consumption of existing products and creation of new ones.

Even if the benchmark version of the model assumes that each good is produced using only labour (in linear fashion, as in the most basic New Keynesian model), the number of active producers in any given period behaves very much like the capital stock in the simplest real business cycle (RBC) model: $N_t = (1 - \delta)(N_{t-1} + N_{E,t-1})$. As this law of motion shows, the number of producing firms is predetermined and does not respond to shocks on impact, but it then adjusts gradually in response to stochastic disturbances to aggregate productivity.9

The benchmark version of the model is simple enough to be literally solved with pencil and paper in log-linearized form, even though the details of the solution are not included in the published paper. There, Bilbiie, Melitz, and I use calibration to illustrate the properties of the model numerically. We show that it does at least as well (or as poorly—beauty here is in the eye of the beholder) as the basic RBC set-up with respect to the familiar set of business cycle moments these models are usually evaluated against, but, in addition, it replicates successfully data properties such as the cyclicality of profits and producer entry. With translog preferences (which imply that products become more closely substitutable as their number increases), the model does a remarkable job of matching the cyclicality of the labour-share-based measure of mark-ups in the US economy used by Rotemberg and Woodford (1999).

Bilbiie, Melitz, and I set up the model intentionally to keep it as simple and as clean as possible, thus abstracting from many features of reality. For instance, we do not introduce heterogeneity across producers, and we assume that exit happens only as a result of exogenous firm destruction.10 Absence of heterogeneity and endogenous exit implies that the model does not feature hysteresis. It also makes it possible to solve it reliably using log-linear approximation, thus obtaining results that are transparent to most macroeconomists—and no, this does not mean that we are married to log-linearization! In a nutshell, it is not unfair to characterize our model as Romer (1990)

---

9 The paper also presents a version of the model in which production combines labour and physical capital.

10 There is no fixed cost in the model in addition to the initial sunk cost of entry. This implies that, once firms have entered, they would never exit, unless hit by the exogenous ‘death’ shock. Endogenous exit would require heterogeneity to avoid situations where all firms would want to exit. We discuss in the paper the reasons why properly calibrated exogenous exit is a reasonable approximation of reality for the purposes of our exercise.
minus long-run growth and plus uncertainty, with preferences that are not restricted to Dixit–Stiglitz.

Since our paper was circulated and published, a large number of extensions and applications have been written and published by many scholars, including explorations of the role of monetary in the presence of endogenous producer dynamics in sticky-price versions of the model.11 Once one introduces heterogeneity and endogenous exit, and assumes the appropriate externality in entry costs, the framework can generate both hysteresis and endogenous growth, making it possible to study the questions that are of so much interest nowadays.12,13

We address normative issues in Bilbiie et al. (2016), and I studied optimal fiscal policy in the BGM model in Chugh and Ghironi (2015), but I focus next on work by Dhingra and Morrow (2014) that—although not developed in a dynamic, stochastic environment—addresses a very important question: what is the optimal amount of product variety in the presence of monopolistic competition and heterogeneous productivity across firms? Does the market equilibrium coincide with the solution to a social planning problem? Dhingra and Morrow study the conditions under which this happens. Their analysis complements what Bilbiie, Melitz, and I did in our 2016 CEPR DP, which focused on the DSGE case without firm heterogeneity.

Now consider the following. Since the results of Hsieh and Klenow (2009) and work by others on the consequences of resource misallocation across firms (for instance, Restuccia and Rogerson (2013) and Fattal Jaef (2016)), one of the mantras we have been hearing from the policy community is that countries should implement structural reforms designed to facilitate reallocation of resources to high-productivity firms and exit by the low-productivity ones. A problem that I see in that discussion is that it is often completely disconnected from discussion of why those low-productivity firms exist. If it is because they are kept alive (or in an undead, zombie state) by distorting an otherwise efficient outcome, yes, reforms should be implemented that ‘kill’ those firms and reallocate resources to the more efficient ones. But heterogeneous productivity—with low-productivity firms existing in equilibrium—may also be the efficient outcome of consumer demand of differentiated products and endogenous entry of producers that satisfy that demand. Dhingra and Morrow’s paper helps us understand when this might be the case, and the discussion of reallocation in the policy debates should become very aware of their results. Then, if one combines BGM and Dhingra–Morrow, it becomes possible to study the consequences of product market reforms that facilitate entry and reallocate resources across heterogeneous firms in a dynamic model.

11 An incomplete list of references includes Bergin and Corsetti (2008), Faia (2012), Lewis (2013), and Bilbiie et al. (2014). On fiscal policy, see Chugh and Ghironi (2015) and Colciago (2016). Some of these contributions explore the consequences of strategic interactions among firms of non-negligible size in models of oligopolistic competition. I return to this topic below.
12 A technical challenge that should be tackled in the case of translog preferences would be how to ensure that long-run growth would not imply a downward trend in the mark-up, which would be inconsistent with the evidence. But the problem would not arise with standard Dixit–Stiglitz preferences.
13 Anzoategui et al. (2015) show how endogenous technology adoption and R&D extensive margin dynamics—which, like BGM, share key features with Romer (1990)—can result in persistent business cycle fluctuations. Comin and Gertler (2006) introduced the concept of medium-term business cycles and showed that a model with endogenous R&D and entry can replicate these lower-frequency fluctuations.
environment that makes it possible to trace the effects of reforms from their short-run impact all the way to their long-term outcomes. I return to the topic of structural reforms below, but the events of the last decade—and the prominent role that market reforms have taken in recommended policy menus—underscore how important it has become to go beyond the static, long-run analysis of Blanchard and Giavazzi (2003) and understand the effects of reforms in fully dynamic, stochastic settings. Most of the work referenced above assumes monopolistic competition among a continuum of firms. It thus lends itself naturally to incorporation of sticky prices, as in some work I mentioned. But this means it also lends itself naturally to exploration of the role of monetary policy (not to mention fiscal policy) in affecting the dynamics triggered by exogenous shocks and other policies (such as changes in market regulation) and potentially contributing to longer-term effects once hysteresis is accounted for. Once firm heterogeneity is included, one can study the implications of macroeconomic policy for changes in characteristics of the distribution of productivity across firms—and this in average productivity and, if the model includes long-run growth, the long-run growth rate of the economy. These analyses, if performed, would complement the focus on the distributional effects of monetary policy implied by household heterogeneity in the HANK framework by focusing on the production side of the economy. The ongoing debate on ‘secular stagnation’, low productivity growth, and hysteresis effects suggests that these are exercises it would be important to perform.

Finally, it is important to note that the focus of BGM and the afore-mentioned literature on monopolistic competition does not imply that attention should be restricted to this form of interaction between producers. The New Keynesian macro literature and much trade literature settled on monopolistic competition because, under assumption of continuity (or of a sufficiently large number of producers), it makes it possible to accomplish basic goals (introducing sticky prices or having welfare benefits from product variety) while avoiding the issue of strategic interactions between firms of non-negligible size. Once firms that are not of negligible size relative to the size of the market are included in the model, one needs to take a stand on their mode of competition, on why they do or do not collude, etc. Peter Neary has been advocating for a long time that trade theory should move beyond monopolistic competition and study other forms of market power. In macro, Federico Etro and co-authors have been developing very interesting versions of the BGM model and of its extension to the sticky-price environment that explore the consequences of Bertrand or Cournot competition. See, for instance, Colciago and Etro (2010), Etro and Colciago (2010), and Etro and Rossi (2015a, b). The results of these papers and others in this area provide a starting point for further exploration of the implications of strategic behaviour by large firms for the questions facing policy-makers today and, possibly, in the future.

14 Cacciatore and Fiori (2016) make an important contribution in this direction by using a version of BGM extended to incorporate SAM frictions in the labour market.
15 See, for instance, Neary (2016) and Neary and Tharakan (2012). De Blas and Russ (2015) explore the consequences of Bertrand competition in an extension of the widely used Ricardian model by Bernard et al. (2003).
(ii) Open economy

In Ghironi and Melitz (2005—GM below), we made a start at bridging the gap between modern international macroeconomics and trade theory by incorporating the Melitz (2003) trade model in a DSGE model of international business cycles. In a nutshell, we developed a true dynamic, general equilibrium Melitz model with uncertainty.\(^\text{16}\) The model shares several features with the BGM model described above, with two major differences. As in the original Melitz model, we assume that entrants face uncertainty about their firm-specific productivity at the time when they commit to sunk entry decisions into their domestic economies. Upon entry, producer-specific productivity is drawn from a continuous distribution (assumed to be Pareto when we solve the model). Firm-specific productivity remains fixed thereafter, but production (which, as in the benchmark version of BGM, uses only labour) is subject to aggregate, country-specific productivity shocks. In terms of the BGM model details I presented above, the value of the firm and firm profits in the symmetric equilibrium, \(v_t\) and \(d_t\), are replaced by average firm value and profit, \(\bar{v}_t\) and \(\bar{d}_t\), i.e. the firm value and profit evaluated at an appropriately defined, market-share-weighted average of firm-specific productivity.\(^\text{17}\)

The second key difference relative to BGM is that GM develops a two-country model in which producers decide endogenously whether to export output to the foreign market. Trade entails two types of costs: standard iceberg costs and fixed costs. Because of these fixed costs, only sufficiently productive firms—those whose firm-specific productivity is above an endogenously determined cut-off—export to the foreign country. Aggregate shocks cause the cut-off productivity for exporting to fluctuate, and thus cause changes in the composition of the consumption baskets across countries. (Average profits, \(\bar{d}_t\), thus combine average profits from domestic sales and average export profits.) As we show, the micro-level features of the model cause deviations from purchasing power parity that would be absent without trade costs.\(^\text{18}\)

We show that the model sheds new light on a classic issue in international macroeconomics: the Harrod–Balassa–Samuelson (HBS) effect, or the evidence that richer countries are characterized by higher prices and an appreciated real exchange rate. Textbook theory (for instance, Obstfeld and Rogoff, 1996) assumes that the effect is caused by differences in productivity growth between traded and non-traded sectors. In our model, a completely aggregate increase in home productivity causes real appreciation because of entry and endogenous non-tradedness.\(^\text{19}\) Thus, we provide a new perspective on the

---

\(^\text{16}\) While Melitz (2003) refers to the model as dynamic, general equilibrium, and characterized by behaviour under uncertainty, the extent to which it indeed has those characteristics is not what macroeconomists would have in mind: Melitz (2003) focuses on a steady-state environment; the financing of sunk costs incurred by firms upon entry is not really modelled; and the only uncertainty is that on firm-specific productivity that firms face before entry in the domestic economy. We address those limitations in GM by developing a fully dynamic model in which entry costs are financed by households (as in BGM) and firms are subject to stochastic, country-specific shocks to aggregate productivity.

\(^\text{17}\) Once these changes are made, aggregate accounting implies the same equality between total demand and total income as in BGM under assumption of financial autarky. When countries are allowed to trade bonds, aggregate accounting implies a standard law of motion for net foreign bond holdings.

\(^\text{18}\) As in BGM, domestic entry is financed by households through purchases of shares in firm equity. We assume that firms are fully owned domestically (i.e. there is no international trade in equities) for simplicity. Hamano (2015) studies the implications of international trade in equities in the GM model.

\(^\text{19}\) All goods are tradable in GM, but some of them are non-traded in equilibrium.
HBS effect that helps explain evidence and complements the traditional theory. Because we intentionally set up the model to allow reliable solution by log-linearization, we can delve deep into it with pencil and paper (even if, different from the basic BGM, we cannot solve it fully), and we obtain analytical results that make intuitions very transparent. Numerical examples then serve the purpose of illustrating those intuitions.

In the second part of the paper, we show that the calibrated model (with a calibration that, if anything, is chosen to match micro-level data) does at least as well (or as poorly) as the standard international RBC (IRBC) model at replicating standard business-cycle moments, and it does better on some dimensions. In a follow-up paper (Ghironi and Melitz, 2007), we show that the calibrated model sheds new light on the cyclicality of net and gross trade flows while being less subject to some problems of the IRBC set-up concerning the cyclicality of the terms of trade.

While Melitz and I were pleased to see the model perform at least as well as the IRBC framework (and better in some dimensions), from my perspective, the real contribution of our paper (and of BGM) was to show how a mechanism that we thought important (and that evidence discussed in the papers increasingly suggested important not just for long-run phenomena) could be embedded in a macro set-up without huge costs in terms of tractability and intuition, and that the exercise would shed valuable light on important questions (the HBS effect). In my opinion, BGM and GM were much more about the mechanisms—endogenous entry in domestic and export markets—and their implications than about the numbers generated by the specific calibrations per se. This emphasis on mechanisms is something I return to below.

Since the publication of GM, a fast-growing literature has developed at the intersection of international trade and international macroeconomics, with contributions covering a wide range of theoretical and empirical issues. It is fair to say that this literature has done a lot to remove the artificial separation of these two fields that I mentioned in the Introduction, but more needs to be done, especially in recognizing that producer-level dynamics and firm heterogeneity should become part of our benchmark thinking and toolkit.

With respect to issues that have become central in present-day policy discussions, very interesting work has developed versions of the model suitable for studying the determinants and consequences of offshoring in a DSGE environment. For instance, Zlate (2016) shows that a model with endogenous offshoring can successfully replicate a number of empirical features of US–Mexico interdependence, including dynamics of firm offshoring decisions that our standard international macro models are silent about.20 Cacciatore (2014) develops a version of GM that incorporates SAM frictions in the labour market and studies the consequences of trade integration in this framework, tracing the dynamics triggered by trade integration from the impact effect of the policy change to the long-run consequences, and studying also how trade integration affects the characteristics of the international business cycle. Trade economists mostly focus on steady-state models and the long-run gains from trade when debating the effects of trade integration. But the devil is in the dynamics! The world is never in

20 Zlate (2016) develops a model of vertical foreign direct investment (FDI), in which US firms decide to offshore production to Mexico in order to produce output that is then imported back to the US and sold to US consumers. See Contessi (2015) for a model in which firms decide to offshore in order to serve the foreign market (horizontal FDI).
steady state. As recent and ongoing events are making painfully clear, it is the dynamics of adjustment (or the rigidities that interfere with adjustment) that are determining electoral outcomes and the fate of proposed or existing policies (trade and others). Trade economics should move past the fiction of steady state and balanced trade when studying the effects of trade integration. Cacciatore’s work is an important step in that direction.\textsuperscript{21}

Two important developments in the trade literature that I consider especially promising for their spillovers for macro research are the study of granularity and that of global value chains (GVCs). Di Giovanni and Levchenko (2012) build on Gabaix (2011) to make an important contribution. As Gabaix showed, under the appropriate assumptions about the distribution of firm size, idiosyncratic shocks across firms do not wash out in the aggregate. If the economy is ‘granular’, i.e. it features a fat tail of disproportionately large firms, shocks to these firms become a driver of the aggregate business cycle. Di Giovanni and Levchenko begin by showing that smaller, more open economies tend to be more granular than large ones. They then develop a multi-country Melitz-type model with granularity, and they show that trade integration tends to increase granularity. This is an intuitive consequence of a key property of the Melitz model: the model implies that trade reallocates market share toward the relatively more efficient firms, which become bigger (increased market share of more efficient firms also results in an endogenous increase in average firm productivity). In a granular environment, large (more efficient) firms becoming larger implies more granularity. This poses obvious questions for the debate on the consequences of trade integration, and also for the ongoing discussions on structural reforms. It is a research area that macroeconomists should pay much attention to.

The same is true of GVCs. Fragmentation of production across borders has changed the nature of trade, resulting in increasing importance of trade in value added rather than traditional trade. Bems and Johnson (forthcoming) and Johnson and Noguera (2012) have made important contributions to our understanding of the phenomenon, and Duval \textit{et al.} (2016) have explored the implications for international business cycles. GVCs have key implications for how we think about competitiveness, because exchange rate changes no longer have only the standard effect on trade of making purely domestically produced goods cheaper (or more expensive) for foreigners. The international macroeconomics of GVCs is only at its beginning, and it is an especially important research area—also for a better understanding of what we would stand to lose with trade wars.

Coming to interdependence between trade and macroeconomic policies, inroads have been made by Bergin and Corsetti (2016) and Cacciatore and Ghironi (2012; CG below).\textsuperscript{22} Bergin and Corsetti (2016) show how accounting for extensive margin dynamics can reconcile the traditional preference of policy-makers for boosting manufacturing competitiveness with the incentive to appreciate the terms of trade embedded in New

\textsuperscript{21} Alessandria and Choi (2014) and Alessandria \textit{et al.} (2014) also study the consequences of trade integration in DSGE models with endogenous entry and exit decisions. Their models incorporate additional features of producer dynamics, but they abstract from labour market frictions.

\textsuperscript{22} See also Cooke (2014, 2016). Hamano and Pappadà (2017) focus on the interaction between monetary policy, producer dynamics, and external imbalances.
Keynesian open economy models since Corsetti and Pesenti (2001). In CG, we develop a version of the GM–Cacciatore model that incorporates sticky wages and prices, and we study the consequences of changes in trade integration for the optimal conduct of monetary policy, as well as the role of monetary policy in the dynamics triggered by a possible return to past levels of tariffs. Interdependence between trade and monetary policy is all over the map in policy discussions and documents—just think of the role of the creation of the Single Market in the run-up to the euro. The New Keynesian open economy literature has studied the consequences of openness for optimal monetary policy and alternative exchange rate regimes in models in which openness is characterized by changes in the degree of home bias in consumer preferences or parameters of technology. But to the extent that openness is the outcome of trade policy actions, proxying policy by varying a parameter of preferences may be very misleading: after all, those are the famous structural parameters we would like to keep invariant to policy. Embedding trade microfoundations in the international macro framework makes it possible to perform a deeper analysis of the consequences of changes in trade policy for monetary policy. This is what we make a start at doing in CG.23

III. What next?

The discussion in the previous section hinted at a number of research directions that I consider promising for the future. In addition to those, there are four directions that I view as especially relevant for future macro theory research. Two—financial intermediation and household heterogeneity—have already become part of the emerging ‘consensus future’ of macroeconomics. I briefly discuss below how research in these areas would connect naturally to issues related to producer dynamics and heterogeneity. Existing research already yielded results that could be used to introduce these areas of ongoing work at the end of a first-year, PhD macro sequence—at least, that is what I would do if I taught the second semester (or the third quarter) of such sequence. The other two research directions I focus on build and expand on themes I mentioned in section II.

(i) Financial intermediation

The work I reviewed in the previous section makes strong simplifying assumptions with respect to the role of financial markets. Entrants finance their sunk entry costs by issuing equity in frictionless stock markets (except for an assumption of extreme home equity bias that prevents international equity trading in most open economy models). There is no role for financial intermediation and associated frictions. Reality reminded us brutally in the last decade of the possible consequences of abstracting from these

23 In Barattieri et al. (2017), we use a small open economy version of CG to study the macroeconomic consequences of tariff shocks. We show that the model replicates empirical evidence on the responses to such shocks, and we evaluate whether protectionism can be beneficial by raising inflation when the economy is mired in a liquidity trap. Our conclusion is negative.
features. Moreover, empirical work had already documented the importance of bank finance—and the consequences of changes in the characteristics of banking markets—for the dynamics of producer entry and exit before the GFC (for instance, Cetorelli and Strahan, 2006).

Stebunovs (2008) made a start at modelling the results of this empirical literature in a version of BGM in which firms must borrow from intermediaries with monopoly power that compete in Cournot fashion over the number of loans they issue. He showed that balancing the portfolio expansion effect of extending more loans with the profit destruction externality that producer entry imposes on all producers results in intermediaries erecting a financial market barrier to entry in the form of a mark-down in the bank’s valuation of an extra loan (i.e. an extra productive unit) relative to the perfectly competitive benchmark.

Formally, the Euler equation for financing of entry in the BGM model is replaced by an Euler equation that determines the value of an additional producing firm at time $t+1$ in the portfolio of loans extended by a financial intermediary:

$$q_t = \beta E_t \left[ \left( \frac{C_t}{C_{t+1}} \right) \left( 1 - \frac{1}{H} \right) d_{t+1} + (1 - \delta) q_{t+1} \right]$$

where $H$ is the number of financial intermediaries that compete in the market. The number $H$ plays a similar role to that of the elasticity of substitution across products ($\theta$) in the familiar continuous model of monopolistic competition: With endogenous output, the assumption $\theta > 1$ is necessary to ensure strictly positive output in equilibrium. The case $\theta \to \infty$ corresponds to perfect competition. Here, $H > 1$ is necessary to ensure that intermediaries finance a positive number of entrants: If $H = 1$ (absolute monopoly in the banking market), the Euler equation above implies $q_t = 0$, and the economy is starved of entry and, eventually, production. If $H \to \infty$, the banking market becomes perfectly competitive, and the Euler equation that determines $q_t$ becomes that of perfectly competitive finance, as in BGM (except for the difference that $q_t$ is the value of a firm producing with certainty at $t+1$, and thus $(1 - \delta)$ multiplies only $q_{t+1}$). As long as $H$ is finite, the Euler equation implies a mark-down of $q_t$ relative to the perfectly competitive scenario.24

The entry condition $v_t = (\hat{w}_t / Z_t) f_{E,t}$ of BGM is replaced by $q_t = \left\{ \hat{w}_t / \left( (1 - \delta) Z_t \right) \right\} f_{E,t}$. A reform that increases competition in local banking markets (such as the scenario explored empirically by Cetorelli and Strahan (2006)) causes $H$ to rise and boosts entry of non-financial establishments by narrowing the gap between $q_t$ and its value under perfect competition.

Notz (2012) shows that Stebunovs’s results hold also in a model in which the financial contract is a more standard debt contract and does not assume that intermediaries extract all the profits of the firms they finance in repayment of their loans. More recently, Bergin, Feng, and Lin (2014; BFL) develop a version of BGM that builds on Jermann and Quadrini’s (2012) to incorporate financing constraints and a mix of

24 This is akin to Hayashi’s (1982) result that monopoly power results in a mark-down of the marginal valuation of capital relative to its average valuation in capital accumulation decisions by firms. See Cacciatore et al. (2015) for an open economy extension of Stebunovs’s analysis.
equity versus bond finance. BFL show that their model replicates several features of data in response to financial shocks.

These examples illustrate how finance can be embedded in models with producer dynamics, but much more work in this area is needed, especially to address the implications of borrower heterogeneity (here, heterogeneity across producers), asymmetric information, and the open economy dimension. We still lack a consensus model of producer heterogeneity and financial frictions that would allow us to address the role of financial intermediation for misallocation of resources across producers with market power. Manova (2013) built on Melitz (2003) to develop a benchmark model of financial frictions and trade with heterogeneous firms. She highlighted how financing requirements associated with trade can result in trade participation by a smaller set of firms in the presence of frictions, and her work provided the theoretical foundation for the explanation of the ‘great trade collapse’ of 2008–9 that highlighted the drying up of trade finance as the key source of the collapse. However, we still do not have a consensus framework that embeds financial frictions and a role for intermediation in a dynamic international macro model with uncertainty. More research in this area is needed in order to address a number of interesting positive and normative questions.

(ii) Heterogeneous households

Kaplan et al. (2016) introduce uninsurable employment risk in the New Keynesian model to address the inability of the framework to study the distributional consequences of monetary policy. Ravn and Sterk (2016) show how the framework can be tractably combined with search-and-matching frictions in the labour market. The SAM model has been incorporated in a number of international macro and macro papers with producer dynamics, some of which I reviewed above. However, to the best of my knowledge, these models abstract from uninsurable employment risk, with implications for their results and the range of questions they can address. Introducing meaningful household heterogeneity in models with producer dynamics will make it possible to study the connection between distributional issues and firm dynamics.

For instance, evidence shows that, on average, exporting firms are larger and more productive than firms that serve only the domestic market. The evidence also shows that, on average, exporters pay higher wages than non-exporters. How do changes in trade policy that affect job creation and destruction by exporters and non-exporters affect the distribution of income on impact and along the dynamics toward the new long-run steady state? How does this interact with macro policy and the exchange rate? How do uninsurable employment risk and household asset accumulation shape the effects of large policy shocks, such as the possible dismantling of the North American Free Trade Agreement (NAFTA) or the withdrawal of the Federal Reserve from international financial regulation arrangements? And what is the effect (if any) of limited household participation in financial markets (a possible interpretation of market incompleteness) on the dynamics of firm entry and exit? In BGM and GM, firms finance their entry costs by issuing equity purchased by the representative household on the stock market. The extent to which households participate in the stock market may have implications for the extent to which different firms rely on alternative sources of finance, and it may importantly affect the consequences of shocks and macro policy actions for both the
distribution of household income and the distribution of activity across firms. In the light of ongoing events and policy discussions, these are interesting, important questions that our macroeconomic framework should address.25

In addition to these two research areas, I view two directions of study as especially important to build an overall framework of analysis suited to tackle present-day (and future) positive and normative issues: one is the consequences of granularity, networks, and strategic interactions between firms; the other is interdependence across policies, within and across countries, and policy-regime change. However, I would reserve covering work in these areas for second-year, field PhD courses.

(iii) Granularity, strategic interactions, and networks

The benchmark macro model with monopolistic competition assumes a continuum of measure-zero producers that interact with each other in non-strategic fashion. Producers respond to aggregates but not to individual competitors. This is true even in models with heterogeneous producers, such as frameworks that allow for heterogeneous productivity across firms. But the research by Gabaix (2011) in a closed-economy environment and its extension by di Giovanni and Levchenko (2012) to the consequences of international trade highlight the importance of allowing differences in firm size to have meaningful implications for the consequences of idiosyncratic shocks across firms. Moreover, once we begin entertaining the idea that firms in our macro models should no longer be measure-zero entities, the assumption of non-strategic monopolistic competition becomes less and less tenable. In small open economies, policy-makers pay attention to the decisions of individual large firms that represent a disproportionate portion of the economy in taking their decisions. Even in large economies, expansions or contractions of industry giants at the centre of large networks of transactions, and interactions between such key firms, have ripple effects that can propagate to aggregate consequences and non-negligible spillovers abroad.

Integrating these mechanisms in macro models will be important to answer a number of questions. For instance, as I discuss below, structural reforms designed to increase product market flexibility have become part of the policy menu invoked by policy-makers to improve economic performance in a number of countries. But how do reforms impact economies (domestic and foreign) in the presence of granularity, networks, and strategic interdependence between large firms (and, sometimes, between these firms and the policy-makers themselves)? Although there is a growing literature on granularity, networks, and strategic interactions (some of which I briefly mentioned above), we simply do not know enough in this area, and we need to know more.26

Similarly, we do not know enough about the implications of GVCs for macroeconomic policy, structural reforms, or even the dynamic consequences of changes in trade

25 Research should also address the extent to which SAM is a satisfactory model of unemployment for the purpose of addressing these and other questions, or whether Michaillat’s (2012) version with rationing unemployment should be preferred.
26 On granularity, see also Carvalho and Grassi (2017) and di Giovanni, Levchenko, and Mejean (2014, 2017). The literature on networks includes Carvalho (2010), Acemoglu et al. (2012), Baqae (2016), and Grassi (2016).
The establishment of GVCs has resulted in fragmentation of production into networks that cross multiple country borders, with product components or unfinished products often crossing a given border repeatedly before the finished product is available to consumers in its final destination. As hinted above, this implies that the standard notion of the competitiveness effect of exchange rate changes is no longer valid. We need a dynamic model of roundabout production across country borders to begin understanding the consequences of different macro policies and exchange rate arrangements in this environment, and we need dynamic models of GVC formation—say, a dynamic, stochastic version of Antras and Chor (2013)—to study these questions more deeply and to understand the consequences of market reforms and changes in trade policy. The threat that established GVCs might unravel if protectionist pressures led to trade wars makes the need for this research all the more urgent.

(iv) Policy interdependence and the dynamics of policy-regime change

The difficulties facing policy-making since 2008 have highlighted the importance of multi-pronged approaches to tackling crises and persistent recessions. Calls for policy packages have become a mantra for policy-makers at the highest level. In many instances, calls for multi-pronged policy-making are combined with exhortations (or promises) to engage in stronger international coordination of economic policies (for instance, see G20 (2016) and Lagarde (2016a,b)). The menu of this multi-pronged approach usually includes monetary policy, fiscal policy, and structural reforms, where, depending on the situations, the latter include reforms of financial, labour, and product markets. Macro-prudential policy is also often added to the menu.

This rich menu of policies, and their interdependence within and across countries, raises questions about how these policies interact with each other—again, within and across countries. Evaluating the possible benefits of coordinating policies across countries—or across policy-makers within a given country—requires attention to specifying policy-maker objectives, strategy spaces, and asymmetries across countries (or policy-makers) that can impinge on the possible gains from coordinating policies. I discuss below some recent work on the interaction of monetary policy and structural reforms to which I contributed. Although this work yielded valuable insights, the analysis of product and labour market reforms that it performs is based on assuming that characteristics of market regulation in reforming economies (think of euro area countries) are exogenously (and fully credibly) adjusted to US levels, taking the United States as a benchmark for flexibility. No stand is taken on whether this is optimal for the countries involved, and not even on whether existing levels of regulation are optimal for the United States. The topic of optimal structural reforms (like that of optimal trade policy in a dynamic general equilibrium setting under uncertainty) is a major open area for research. When one recognizes that reforms interact with macroeconomic

27 For recent examples of calls for appropriately designed policy packages, see Draghi (2016a), G20 (2016), Lagarde (2016a,b), and Praet (2016).
28 Draghi (2016b) contains an explicit call for research on ‘interdependence in interdependence’.
policy (as Draghi (2015) and other policy-makers clarified in unequivocal terms), one needs to address the implications of strategic interactions between governments (in charge of regulation) and central banks. And when one recognizes that policies have spillover effects across country borders, it becomes clear that we need to understand international interdependence within and across policy areas, and within and across borders. This is an area where we have barely begun scratching the surface. Past theoretical literature explored interdependence between monetary and fiscal policies within and across countries (Eichengreen and Ghironi, 2002; Dixit and Lambertini, 2003; or Beetsma and Jensen, 2005, in a New Keynesian framework) and between monetary and trade policy (Basevi et al., 1990), but much more work needs to be done to yield results that can provide a more reliable road map for the understanding of positive and normative issues in this area.

The same is true of policy-regime transition. Theoretical macro analysis (or international macro analysis) most often compares different policy regimes ‘in a vacuum’, without addressing the actual issue of how the economy transitions from one regime to the other, and what policy-makers should do to ensure an orderly transition. In international macro, we have a vast literature on exchange rate crises and a more recent literature on sudden stops as the result of occasionally binding constraints, but we do not really have a consensus way of modelling orderly versus disorderly regime transition. While I could rattle off a bunch of consensus references on crises, I cannot come up with a consensus reference for modelling no-crisis endogenous regime transition, and what ensures that the transition would indeed be a no-crisis one. I think this should be an important area for future research, and I think this is the area where we may need the most significant departures from existing methodologies for macro modelling.

I am optimistic that developments in theoretical and computational tools have put us in a position to make significant progress in the near future along all four research directions I described, and I think we should encourage our PhD students to pursue them. A better understanding of the economy and formulation of better policy advice will follow from a deeper understanding of producer dynamics, of the interaction of these dynamics with finance and the behaviour of heterogeneous households, and of how strategic decisions of large firms shape policy-maker responses and/or are shaped by them.

29 The four research areas I focused on are by no means the only ones I consider important for the future development of macroeconomic theory. An example of another new area of research that I consider very important and that I think we should mention in second-year field teaching is immigration. Given recent events and evidence, I believe it is important that international macroeconomists start moving beyond the assumption of immobile labour we usually make and start exploring the implications of labour mobility. Dmitriev and Hoddenbagh (2012) and Farhi and Werning (2014) made a start at this in their models of monetary unions, connecting the literature that employs the basic New Keynesian setup to Mundell’s (1961) seminal work on optimum currency areas. Mandelman and Zlate (2012) went one step further in their modelling of immigration, treating it as an entry decision subject to sunk costs. They embedded this mechanism in a model of US–Mexico interdependence and showed that the model does a remarkable job of replicating cyclical patterns of immigration from Mexico into the United States and of remittances from the latter back to Mexico. Mandelman and Zlate (2016) extend the model to incorporate task trade and skill upgrading to study the role of offshoring and immigration dynamics in shaping observed US labour market polarization. Much more work in this area is needed.
IV. Methodological issues

From a methodological standpoint, I believe that the path to progress lies in not being dogmatic and in recognizing that different types of models can be useful for different purposes, as Blanchard (2017) argued. Within such a flexible, non-dogmatic approach, DSGE models can serve very important purposes for theoretical analysis and the application of theory to questions of positive and normative nature, including policy evaluation exercises.

To return to a theme I mentioned above, analysis of policy packages (be it positive or normative) requires models to include all the features that are key to disentangling and understanding the effects of different policies, and how they interact with each other. \(^{30}\) DSGE models have the potential to fulfil this task successfully. By building on the appropriate level of microfoundation, they stand the best chance of disentangling the various channels through which the policies that are called for are transmitted and interact with each other. By being dynamic, the models can help us understand the differences between short- and long-run effects of different policy actions—and how different parts of policy packages can complement or substitute for each other over time. By being stochastic, the models recognize that policy operates in an uncertain environment, where consumers, firms, and policy-makers take their decisions without perfect knowledge of the future; that the effects of reforms can depend on business cycle conditions, and reforms themselves can alter the characteristics of the business cycle. Finally, general equilibrium implies that prices and quantities are jointly determined by the constraints and optimality conditions of the model, with no imposition of a priori assumptions on how policy should affect any price or quantity.

Importantly, the defining characteristics of DSGE modelling that I just mentioned (microfoundation—even if, strictly speaking, there is no M in DSGE, dynamics, uncertainty, and joint determination of prices and quantities by the model’s constraints and optimality conditions) do not necessarily include rational expectations and reliance on exogenous productivity shocks as the sole source—or even as a source—of cyclical fluctuations. DSGE analysis does not require the most standard Euler equation that ties expected growth in the marginal utility of consumption to the ex ante real interest rate, nor does it require all those ingredients (or solution techniques) for which DSGE research has become the object of a barrage of criticism from academics, bloggers, and journalists. We may want to use some or all of those ingredients and techniques because, after all, models are never meant to be photographs of reality, and it is useful to establish benchmark, transparent results in simplified frameworks that can then guide our understanding of the implications of working with more realistic assumptions. But nothing in the DSGE approach constrains us to using any of those ingredients. Even the level of microfoundation we want to embed in our models is ultimately a decision that must be taken based on the balance between complication, clarity, and empirical plausibility of assumptions and results. Related to this point, I am not advocating in this paper that all models in macro and international macro from now on should include all the building blocks I discussed above (producer entry, firm heterogeneity, openness of the economy, financial intermediation, HANK, and SAM). The

\(^{30}\) Some of the material in this and the next paragraph repeats points I made in Ghironi (2017b).
choice of what to include in any given model should still be guided by the questions we want to address: Our framework and approach—and our teaching—should be flexible enough that building blocks can be added or subtracted depending on the question of interest and the goal of our analysis—whether it is purely theoretical or more applied.\(^{31}\)

As I pointed out above, my view on the balance of micro and macro that we should incorporate in DSGE analysis is that macro needs (more) micro than the established benchmark has been incorporating, with a focus on producer dynamics and interactions to supplement the increasing attention to financial and labour markets. Note that this is important not just for the sake of microfoundations and elegance: it is important for the models to address key features of real world dynamics, to fit the narrative of policy-makers, and to avoid potentially misleading results.

Interdependence across policies is an excellent example of what I have in mind. As ECB President Mario Draghi began his campaign in favour of structural reforms designed to increase the flexibility of product and labour market in the euro area, macroeconomists became naturally interested in the question of whether implementing such reforms during a recession and when the central bank is constrained by the zero lower bound (ZLB) on interest rates can be especially costly. The question was initially addressed by Eggertsson, Ferrero, and Raffo (2014; EFR below) in a paper where they use the off-the-shelf basic New Keynesian framework and model reforms essentially as exogenous cuts to price and wage mark-ups. EFR concluded that reforms boost external competitiveness and improve the external balance, but they can be very costly at the ZLB because of their deflationary effect. The EFR article was an important starting point for discussion of the effects of reforms during recessions and at the ZLB, and it received a lot of attention in the media, in the policy community, and among academics. Economists at policy institutions started using their much richer DSGE models to simulate the effects of reforms modelled in the same way. But this modelling approach completely abstracts from any of the product and labour market dynamics that policymakers have in mind when talking about structural reforms. It is sufficient to read the opening statements to Draghi’s press conferences or his speeches to pick up countless references to entry barriers, product market competition, job creation and destruction, and features of micro-level behaviour that the EFR modelling approach abstracts from.

Is it important to include such features in the model? The answer from the work I did with Cacciatore and Fiori (Cacciatore et al., 2016a—CFG below) and with these authors and Duval (Cacciatore et al., 2016b,c) is a strong yes. In this work, which provided the model foundation for the discussion of structural reforms in the April 2016 issue of the IMF’s *World Economic Outlook* (IMF, 2016) and for the advice the IMF has been giving since, we show that incorporating micro-level dynamics has important consequences for results, and that the ZLB should not in itself be a reason to delay reforms, at least of some types. In CFG, we show that implementing reforms in an environment of exceptional macro policy expansion is a way not only to smooth short-run costs

\(^{31}\) I most strongly reject the criticism that DSGE models are designed based on ‘cherry picking’ the facts to be matched, and the implication that other modelling approaches would be superior by not being subject to this problem. Every model will only explain the behaviour of the variables it incorporates. Every model-builder, within each modelling approach, is engaging in cherry picking, and every model will fail on some empirical dimensions that some of us may find very relevant. The only model that avoids this problem is called reality.
of reforms, but also to bring long-run benefits closer to the present, as Draghi (2015) argued. Of course, our models abstract from many relevant micro-level features, and they make assumptions that the anti-DSGE crowd views as mortal sins. But I view the type of nuanced policy advice our models helped the IMF give since last April as a clear success of this DSGE work, and many other successes are out there that the critics fail (or simply refuse) to acknowledge.\footnote{For an interesting contribution to the debate on structural reforms that accounts for firm heterogeneity and endogenous producer exit, see Hamano and Zanetti (2017).}

So, yes to DSGE, and yes to micro.

Within this approach, I believe the most productive way to proceed, especially for teaching and academic research, should be to focus on mechanisms rather than \textit{ad hoc} tweaks to the mathematical specifications of preferences and technologies or adjustment costs. Failures of models to address important features of what we aim to explain should be studied by asking ourselves what mechanisms are missing from the framework for it to succeed—and by mechanisms I mean deeper features than most tweaks to preferences, technology, and/or the introduction of adjustment costs. The same argument applies to shocks: When a model fails to explain a set of data I am interested in, given a set of shocks on which I (more or less) believe the empirical literature has told me what I should expect, my preferred approach is to ask what deep mechanism(s) is the model missing rather than adding shocks all over the model in order to ‘force’ it to fit the data. It is this deep-mechanism-driven approach that led me to develop the agenda with Melitz and those with Cacciatore and others based on the incorporation of producer dynamics. By focusing on mechanisms and keeping the framework as simple and ‘clean’ as we could, we were able to develop models such that, even if analytical solution is not feasible, one can go deep enough into the model with pencil and paper that the intuitions for results become quite transparent.\footnote{Put differently, one can incorporate micro-level dynamics into macro models without necessarily turning them into black boxes or ‘kitchen sinks’. The same was accomplished by much DSGE literature by many scholars in other areas. This literature therefore is not subject to Blanchard’s (2017) criticism of existing DSGE research, and it has already been doing what Blanchard recommends DSGE modellers should be doing. See Ghironi (2017a) for more discussion of this point.}

Of course, this is not to deny that there is a role for tweaks and adjustment costs. We used adjustment costs ourselves in Cacciatore \textit{et al.} (2016b,c) because of the IMF’s interest in a quantitative model (and we used an \textit{ad hoc} shock to the Euler equation for bond holdings in Cacciatore \textit{et al.} (2016c) to push the economy against the ZLB), but, in academic research and for teaching purposes, the cleaner framework should be preferred, with numerical results (which are qualitatively the same as in the more quantitative model) intended more for illustration than for close empirical relevance.

Thus, even taking as given this preference for cleaner models that may miss several features of the data, there should be no dogmatic preclusion against the use of model tweaks in both academic and (especially) policy research, and there should be no preclusion against the use of what Wren-Lewis (2017) refers to as SEMs (structural econometric models), or of ‘toy’ models (in the language of Kocherlakota (2016)), be they...
microfounded or not.\textsuperscript{34} If a SEM or a toy model make it possible to address the question of interest, my view is that we only stand to learn from comparing their results to those of different types of DSGE models (the ‘clean’ variety I prefer and the more quantitative set-ups that have become predominantly associated with the DSGE label since Christiano et al. (2005) and Smets and Wouters (2007)). When results are similar, we will perhaps feel more comfortable about them, and when they differ, we will have new research questions to ponder.\textsuperscript{35}

So, macro needs micro to talk about phenomena that are very relevant to explain reality and address policy questions, but, just as important, we should teach our students to be flexible and non-dogmatic, and we should equip them with the tools to tailor the modelling approach and the specific models they use as researchers to the purposes of their research.\textsuperscript{36}

V. Conclusion

Over 20 years ago, Paul Krugman wrote:

\begin{quote}
I would like to know how the macroeconomic model that I more or less believe can be reconciled with the trade models that I also more or less believe. . . . What we need to know is how to evaluate the microeconomics of international monetary systems. Until we can do that, we are making policy advice by the seat of our pants. (Krugman, 1995)
\end{quote}

Answering that call for research at the intersection of international trade and international macroeconomics is as important now as it was then. In fact, this paper essentially argued that the scope of Krugman’s call and the answer to it must include also research that does not focus on the open economy dimension of macroeconomics. Macro—whether international or not—needs micro: MNM!

This paper has summarized several existing contributions to answering this call, it has outlined key next steps in this programme, and it has discussed methodological issues for this agenda. Contrary to the doom-and-gloom view of macroeconomics that dominates newspaper articles and popular blogs, and that has been put forth also by some very notable scholars, I believe that macroeconomics did not regress in the last 30 years; that it did commit mistakes, but it also delivered a number of important, valuable results; and that there is a bright future for the field if we can avoid being dogmatic—and if the sociology of the journal publication business does not stymie many promising efforts.

\textsuperscript{34} Baldwin and Krugman (1989) is a great example of how to use non-explicitly-microfounded toy modelling to incorporate producer dynamics in the analysis of classic international macro questions. Mehra and Prescott (1985) is an excellent example of a DSGE, microfounded, toy model that yielded very important insights.

\textsuperscript{35} Agent-based modelling is another research area that should receive more attention. Hommes and Iori (2015) presents a collection of recent applications of this approach. See also Assenza and Delli Gatti (2013).

\textsuperscript{36} This also requires avoiding the ‘Don't read, just write’ approach to dissertation research that some pursue. As a field, we will be less likely to lock ourselves into any box if our students have been exposed to and encouraged to read on a ‘variety of boxes’ (and how to manipulate them) on the way to their job market papers.
References


Ghironi, F. (2017a), ‘Not All DSGEs Are Created Equal’, blog post, 26 January.


An interdisciplinary model for macroeconomics

A. G. Haldane* and A. E. Turrell**

Abstract: Macroeconomic modelling has been under intense scrutiny since the Great Financial Crisis, when serious shortcomings were exposed in the methodology used to understand the economy as a whole. Criticism has been levelled at the assumptions employed in the dominant models, particularly that economic agents are homogeneous and optimizing and that the economy is equilibrating. This paper seeks to explore an interdisciplinary approach to macroeconomic modelling, with techniques drawn from other (natural and social) sciences. Specifically, it discusses agent-based modelling, which is used across a wide range of disciplines, as an example of such a technique. Agent-based models are complementary to existing approaches and are suited to answering macroeconomic questions where complexity, heterogeneity, networks, and heuristics play an important role.

Keywords: macroeconomics, modelling, agent-based model, consumption

JEL classification: A12, B22, B40, C63, E21, E03, E27

I. Introduction

The economic and financial crisis has arguably spawned a crisis in the economics and finance profession. Much the same occurred after the Great Depression of the 1930s, when economics was rethought under Keynes’s intellectual leadership (Keynes, 1936). The challenge today, to academics and to policy-makers, may be every bit as great. A spotlight has been thrown on the models used prior to the crisis. Critics argue that they were too restrictive and were not supported by empirical evidence. Academic work is beginning to plug these holes and new perspectives are being sought. Rising to the challenge of modelling a broader set of economic circumstances requires macroeconomists to think afresh, perhaps seeking inspiration from other disciplines with different methodological approaches. This paper outlines the difficulties facing macroeconomic
modelling, examines how a more pluralistic methodology could be beneficial, and offers one constructive path for rising to these challenges.

The structure of the paper is as follows. Section II presents evidence of insularity in economics and considers why that insularity may not be optimal. Section III discusses this ‘macroeconomic mono-culture’, perhaps best exemplified by the dominance of the representative economic agent framework operating with rational expectations; it calls for a more diverse set of macroeconomic models, including ones which experiment with different agent behaviours. In section IV, the interdisciplinary origins of the agent-based approach are described. Section V contrasts the philosophy of agent-based modelling with other macroeconomic modelling approaches. Section VI asks what agent-based models could do for macroeconomics, and section VII concludes.

II. An insular discipline?

Economics, particularly macroeconomics, has historically been rather insular as a discipline, at least in comparison with other subjects. This is not a new observation. It has been puzzled over previously (for instance in Hausman (1992)). And, despite progress more recently, there remains evidence to suggest that the Great Financial Crisis has not dispelled entirely this insularity.

Figure 1 shows that over the period 1950–2010 economics papers in academic journals have cited papers in other disciplines less frequently than the average. In turn, those other disciplines have consistently failed to cite economics research. This suggests that new ideas have not flown freely whether into, or out of, economics. Insularity is partly a natural consequence of stratification because academic disciplines exist as partially ring-fenced areas of enquiry where specialists develop and focus their attentions (Jacobs, 2014). But it appears to be peculiarly strong in economics, as reported in Fourcade et al. (2015).

This situation has improved over recent years, as shown in Figure 2. Nonetheless, economics remains more self-referential than many other, arguably, less interdisciplinary subjects. Even mathematics, considered by many a model of ‘purity’, sits higher on Figure 2 than does economics. Inter-disciplinary work involves extra costs (Yegros-Yegros et al., 2015), particularly those arising from coordination across subject areas, from the most respected journals having a natural bias to that field and from the difficulties of accurate reviewing across subjects. However, there is no reason to assume that these factors are more acute in economics than in other disciplines.

Alongside this evidence of inter-disciplinary insularity in economics, there is evidence of greater intra-disciplinary insularity, too. The number of authors per paper in economics is lower than in other major disciplines, as shown in Figure 3. The ‘big science’ phenomenon has seen increasing returns to scale from academic collaboration over recent years, with scientists clubbing together in ever-larger teams, sometimes collaborating over several generations, in order to deliver the most dramatic breakthroughs (Aad et al., 2012; Abbott et al., 2016). Wuchty et al. (2007) show that research by larger teams is better cited, on average, than research authored by a solo author across all disciplines. Interestingly, this is also true for the top five journals in economics, where a paper with four authors picks up 61 per cent more citations than a paper with one author (Card and DellaVigna, 2013).
Despite these benefits, most natural sciences rank higher than economics on measures of intra-disciplinary insularity. As the size of the datasets available to macroeconomists grows, working in small teams risks missing out on significant economies of scale in research. And experiments, for some people the gold standard of the scientific method, are associated with an increase in the number of authors per paper (Hamermesh, 2013).

Lack of diversity is another signal of insularity. A recent study found that around 43 per cent of the articles published in the top four economics journals were authored by scholars connected to one of the editors at the time of publication (Colussi, 2017). Female authors accounted for fewer than 15 per cent of papers in three of the top economics journals in 2011. Does this matter for the quality of economic research? Diversity is generally deemed to be good for innovation (Page, 2008; Carney, 2017). And the risks of the opposite, a research mono-culture, are likely to be serious (Bronk, 2011; Bronk and Jacoby, 2016).

The strongest evidence on this comes, interestingly enough, from economists themselves. When asked in 2006 to agree or disagree with the statement ‘In general, interdisciplinary knowledge is better than knowledge obtained by a single discipline’, close to 60 per cent of academic economists said that they strongly disagreed versus an average of 21 per cent for practitioners of sociology, political science, psychology, finance, and history (Fourcade et al., 2015).
Figure 2: The number of citations from a shown discipline to other disciplines

![Figure 2](image2)


Figure 3: The mean number of authors per paper for the top 20,000 papers by number of citations in each Scopus subject category for each year for selected subjects and years

![Figure 3](image3)

Source: Scopus.
Macroeconomics is the sub-discipline of economics that is probably most visible to the general public. Yet in the UK, scientists are trusted (see Figure 4(a)) in a way that economists are not (see Figure 4(b)) by almost every group in society. This has many likely causes. It cannot be explained by the greater technical content of economics. It might be explained, at least in part, by the way this technical content is conveyed to a lay audience. That, in turn, may generate a lack of trust in economists’ models and methods, in a way which does not apply to scientists.

Yet precisely because economics combines elements of both the natural and social sciences, its points of disciplinary tangency are likely to be greater than for many other subject areas. Economics has a long reach. As Keynes (1924) said:

**Figure 4:** Net trust by different groupings

Notes: Data shown are the results from a poll of 2,040 British adults over 14–15 February 2017. ‘Leave’ and ‘Remain’ refer to respondents’ voting choice in the UK’s referendum on membership of the European Union.

Source: YouGov.
the master-economist must possess a rare combination of gifts. . . . He must be mathematician, historian, statesman, philosopher—in some degree. He must understand symbols and speak in words. He must contemplate the particular in terms of the general and touch abstract and concrete in the same flight of thought. He must study the present in the light of the past for the purposes of the future. No part of man's nature or his institutions must lie entirely outside his regard.

Keynes’s description of the economist is as relevant today as it was in 1924, though these days we could easily add to the list statistician, computer scientist, psychologist, and even evolutionary biologist. Macroeconomics has much to gain from taking inspiration from other disciplines; and other disciplines could in turn benefit from a better understanding of economics (Stern, 2016), including macroeconomic modelling techniques (Tasoff et al., 2015).

Some specific examples can illustrate some of the potentially fertile areas recently inhabited by cross-disciplinary research in economics. In recent epidemics, the largest economic costs have arisen not from the direct effects of deaths, but from changes in people’s behaviour in response (Avian Flu Working Group, 2006; Keogh-Brown et al., 2010; Sands et al., 2016). The same is true of the collateral damage to the economy caused by recessions and crises. For example, epidemiological models have shown promise in explaining financial contagion at times of financial crisis (Haldane and May, 2011; Arinaminpathy et al., 2012).

A second example is technology and innovation. There are many examples today of transformative technologies which could have macro effects. Digital markets are improving matching processes and reducing information asymmetry. ‘Big data’ promises to identify new risk factors, while artificial intelligence (AI) and robotics offer substantial gains in productivity. There are risks, too—algorithms could collude on price, or operate in a way that is prejudiced (Ezrachi and Stucke, 2016). Text analysis methods, aided considerably by new algorithms such as Word2Vec (Mikolov et al., 2013), are improving macroeconomic forecasting (Baker et al., 2016; Nyman et al., 2015). An understanding of, and an ability to use, computer and data science techniques is likely to be increasingly useful for economists.

A third example is climate change. The Bank of England has recently published research on some of the implications of climate change (Batten et al., 2016). The effect of global temperatures on productivity, growth, and financial stability are highly non-linear and strongly negative beyond a threshold temperature (Burke et al., 2015). Understanding those macroeconomic effects requires a fusion of expertise from the natural and social sciences.

A fourth example, and perhaps the most successful of all, is behavioural economics—the fusion of psychology and the economics of choice (Tversky and Kahneman, 1975). This is reshaping the view of how economic agents make decisions and could in time help to configure models of the macroeconomy and macroeconomic policy. It is already likely to have helped move economics up the table in Figure 2. Plainly, though, there is a distance left to travel.

III. The ‘macroeconomic mono-culture’

Why did macroeconomics become insular? The antecedent subject of political economy took a much broader view, as did many of the economists of the nineteenth and early
twentieth centuries. The dominance of the field by a single methodology may offer a clue. This had its origins in the ‘New Classical Counter Revolution’ of the 1970s in which Lucas, Sargent, Kydland, and Prescott (Kydland and Prescott, 1982; Lucas and Sargent, 1979) overturned the use of ‘structural’ or ‘policy’ modelling of aggregate macroeconomic variables. These structural econometric models, said Lucas and others, had made several mistakes, including saying little about the stagflation of the 1970s and making implausible theoretical assumptions in order to match the data.

Lucas’s most famous critique of these models was that they were not robust to changes in policy; that they did not allow for agents’ behaviours to change as the incentives of those agents changed (Lucas, 1976). Lucas’s critique is reasonable in that robust models should seek to explain how changes in policy might affect aggregate outcomes. In practice, no model is fully Lucas critique-proof; it is a matter of degree.

Lucas and others developed a methodology which they believed would not fall foul of the critique. The paradigm which Lucas believed best achieved these grounded macroeconomic fluctuations is so-called ‘microfoundations’. Simply put, this said that macroeconomic behaviour should be built up from the aggregation of the individual actions of self-interested, typically optimizing, agents. Being grounded in optimizing behaviour, these self-interested behaviours were less susceptible to change when aggregate macroeconomic relationships changed.

In practice, such microfoundations often became closely associated with a particular type of self-interested behaviour, namely optimization with rational expectations (Lucas, 1972, 1987; Muth, 1961). The weak form of rational expectations is as follows: let $I_{t-1,i}$ be the information set available to agent $i$ at the beginning of time period $t$. This agent has an individual expectation $E_{t-1,i} \{ p_{t+s} \}$ for the value of $p$ in period $t+s$ with $s \geq 0$. Define $E_{t-1,i} \{ p_{t+s} | I_{t-1,i} \}$ as the true expectation of $p_{t+s}$ given the available information. The weak form of rational expectations is then defined by

$$E_{t,i} \{ p_{t+s} \} = E_{t,i} \{ p_{t+s} | I_{t-1,i} \} + \epsilon_{t,i} ; E_{t,i} \{ \epsilon_{t,i} | I_{t-1,i} \} = 0$$

where $\epsilon_{t,i}$ is an error term. For this weak form to be true, the objective probability distribution $f$, where $p \sim f$, must exist. In a model, it requires that an agent with full access to the information set $I$ be able to accurately predict the price up to the error. The stronger form of rational expectations, as specified in Muth (1961), adds to the weak form the assumptions that each agent knows the behaviours and decisions that other agents will take, the true values of any deterministic exogenous parameters governing the evolution of the economy, the properties of any probability distributions governing stochastic exogenous variables (such as $f$), and the realized values of endogenous variables. Coupled with these expectations are assumptions about agent optimality. Typically, this takes the form of an assumption that agents maximize their discounted sum of expected future utilities, subject to a budget constraint.

These microfoundations have a number of desirable properties. They can serve as a useful approximation of real-world behaviour, in at least some situations. They condense the world into a small number of readily observable factors, which can help in determining what is driving what. And they are sometimes more analytically tractable, and are certainly more analytically elegant, than most of the alternatives.

They do, however, have a number of limitations. For one, they invoke strong assumptions (such as rational expectations and optimization) which are often not borne out
in the data (Estrella and Fuhrer, 2002). One defence of rational expectations is that, even though individuals may not be rational, their irrationalities cancel at the aggregate level. Shaikh (2016) has shown how several different microeconomic behaviours can lead to the same macroeconomic outcomes. But it is well-known from other disciplines that, in general, heterogeneous micro-level behaviour combines to generate complex, non-linear responses and emergent behaviour at the macro-level. There are also simple scenarios where the strong version of rational expectations breaks down, for instance in minority games (Arthur, 2006). In general, there are relatively few cases in which aggregation to the macro-level undertaken in models which use rational expectations are completely sound (Kirman, 1992).

One plausible alternative set of ‘microfoundations’ would draw instead on empirically observed behaviours among consumers, firms, and governments. The observed behaviours are often termed heuristics or ‘rules of thumb’. These often arise from models which draw on insights from psychology to understand human behaviour (Tversky and Kahneman, 1975). What constitutes rationality is itself not clear or well defined (Simon, 1959). Indeed, in a world of Knightian uncertainty (Knight, 1921), imperfect information, altruism, and costly computation, there is an emerging body of evidence suggesting that heuristics are more ‘rational’, at least as measured by performance, than the narrowly-defined rationality in which agents compute their optimal course of action without any limitations to their information gathering or processing ability (Gode and Sunder, 1993; Hommes, 2006; Haldane and Madouros, 2012; Aikman et al., 2014; Assenza et al., 2017). This is sometimes called the ‘ecological rationality’ of heuristics (Gigerenzer and Brighton, 2009).

The models which arguably dominate the rational-expectations-cum-optimization methodology in macroeconomics are so-called dynamic stochastic general equilibrium (DSGE) models (Smets and Wouters, 2003). In their most stripped-down form, these DSGE models have a unique equilibrium with deviations that are small and smooth, no role for stock variables and micro behaviours of agents which can be simply and linearly aggregated into the behaviour of a representative agent with rational expectations. The majority of central banks take the DSGE framework as their starting point, including the Bank of England (Burgess et al., 2013).

There has been much debate over the pros and cons of DSGE models which we do not seek to repeat here (Colander et al., 2009; Fair, 2012; Smith, 2014). Particularly strong criticism has included that their additions resemble the epicycles of the Ptolemaic system of astronomy (Fagiolo and Roventini, 2012, 2017), their representative agents are ‘stochastic Robinson Crusoes’ (Summers, 1986), and that the models as a whole are ‘post-real’ (Romer, 2016). Post-crisis, researchers are adding heterogeneity in agent types, a role for the financial sector, and bounded rationality for some agents. Some models use exogenous shocks with fat-tailed, rather than Gaussian, distributions (Ascari et al., 2015). These are all useful additions. There are earlier rational expectations models with heterogeneity added via probability mass functions; see, for example, Heathcote (2005). The originators of the pre-crisis workhorse DSGE model (Smets and Wouters, 2003) recently published a paper containing many of these modifications, including a zero lower bound, non-Gaussian shocks, and a financial accelerator. They find that the extensions go: ‘some way in accounting for features of the Great Recession and its aftermath, but they do not suffice to address some of the major policy challenges associated with the use of nonstandard monetary policy and macroprudential policies’ (Lindé et al., 2016).
There are also new research directions, such as the heterogeneous agent New Keynesian DSGE models (Kaplan et al., 2016; Ravn and Sterk, 2016), which feature more than one type of equilibrium.

Despite these important modifications, it seems likely some features of economic systems will remain very difficult to reproduce in a DSGE setting—for example, crisis dynamics. DSGE models have struggled to simultaneously explain the stylized facts observed in real economic systems (Fukac et al., 2006). Table 1 shows a selection of these stylized facts. An important example is the distribution of GDP growth seen historically, which is shown in Figure 5. Around 18 per cent of the data across the UK, US, Germany, and Japan fall outside the best-fit normal distribution. DSGE models do not tend to reproduce, except by construction, the excessively large variation in GDP growth seen in the historical data (Ascari et al., 2015).

The global financial crisis provides a good example of these limitations. DSGE models struggled to explain either how it started or how it propagated. Figure 6 shows the range of

<table>
<thead>
<tr>
<th>Stylized fact</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endogenous self-sustained growth</td>
<td>Burns and Mitchell (1946); Kuznets and Murphy (1966); Zarnowitz (1984); Stock and Watson (1999)</td>
</tr>
<tr>
<td>Fat-tailed GDP growth-rate distribution</td>
<td>Fagiolo et al. (2008); Castaldi and Dosi (2009)</td>
</tr>
<tr>
<td>Recession duration exponentially distributed</td>
<td>Ausloos et al. (2004); Wright (2005)</td>
</tr>
<tr>
<td>Relative volatility of GDP, consumption, and investment</td>
<td>Stock and Watson (1999); Napoletano et al. (2006)</td>
</tr>
<tr>
<td>Cross-correlations of macro variables</td>
<td>Stock and Watson (1999); Napoletano et al. (2006)</td>
</tr>
<tr>
<td>Pro-cyclical aggregate research and development investment</td>
<td>Walde and Woitek (2004)</td>
</tr>
<tr>
<td>Cross-correlations of credit-related variables</td>
<td>Lown and Morgan (2006); Leary (2009)</td>
</tr>
<tr>
<td>Cross-correlation between firm debt and loan losses</td>
<td>Foos et al. (2010); Mendoza and Terrones (2012)</td>
</tr>
<tr>
<td>Distribution of duration of banking crises is right-skewed</td>
<td>Reinhart and Rogoff (2009)</td>
</tr>
<tr>
<td>Distribution of the fiscal costs of banking crises to GDP ratio is fat-tailed</td>
<td>Laeven and Valencia (2013)</td>
</tr>
<tr>
<td>Firm (log) size distribution is right-skewed</td>
<td>Dosi et al. (2007)</td>
</tr>
<tr>
<td>Fat-tailed firm growth-rate distribution</td>
<td>Bottazzi and Secchi (2003, 2006)</td>
</tr>
<tr>
<td>Productivity heterogeneity across firms</td>
<td>Bartelsman and Doms (2000); Dosi et al. (2007)</td>
</tr>
<tr>
<td>Persistent productivity differential across firms</td>
<td>Bartelsman and Doms (2000); Dosi et al. (2007)</td>
</tr>
<tr>
<td>‘Lumpy’ investment rates at firm-level</td>
<td>Doms and Dunne (1998)</td>
</tr>
<tr>
<td>Counter-cyclicality of firm bankruptcies</td>
<td>Jaimovich and Floetotto (2008)</td>
</tr>
<tr>
<td>Firms’ bad-debt distribution fits a power law</td>
<td>Di Guilmi et al. (2004)</td>
</tr>
<tr>
<td>Firm sizes fit a Taylor power law</td>
<td>Gaffeo et al. (2012)</td>
</tr>
</tbody>
</table>
forecasts for UK GDP growth produced by 27 economic forecasters (including the Bank) in 2007. Pre-crisis forecasts were very tightly bunched in a range of 1 percentage point. The methodological mono-culture produced, unsurprisingly, the same crop. These forecasts

**Figure 5:** The distribution of year-on-year growth in GDP, 1871–2015

![Figure 5](image)

Source: Hills et al. (2016).

**Figure 6:** Range of GDP forecasts in 2007Q4

![Figure 6](image)

Source: Haldane (2016).
foresaw a continuation of the gentle undulations in the economy seen in the decade prior to the crisis—the so-called Great Moderation (Bernanke, 2004). At the time, the damped oscillations of the Great Moderation seemed to match well the smooth motion of DSGE models.

The most important point here is not that this set of models did not forecast the precise timing of the crisis. Almost by definition, costly financial crises cannot be forecast because, if they could, central banks and governments would take actions to prevent them. The real problem was that these models said nothing about the probability of a serious crisis arising endogenously at any time, or about the downstream consequences for the economy of a crisis once it had struck. The absence of non-rational expectations, heuristics, and non-linear amplification channels was probably key in explaining these problems.

The more general point is that a single model framework is unlikely to best serve the needs of macroeconomists in every state of nature. At its best, the scientific method calls for carefully controlled experiments applied to a rich ecology of models, enabling gradual selection of those models which best fit the known facts. But macroeconomists rarely, if ever, have the luxury of running experiments. And even when they can, experimental model validation is hard because the macroeconomy is a complex system of interacting parts in which cause and effect are difficult to separate.

Facing these constraints, it is likely that a patchwork of models will be more resilient than a single methodology. A group of genuinely distinct models, in competition to match the moments of the real world, are likely to produce a far richer set of insights than a single class of models, however aesthetically beautiful. In other words, what may be needed is a ‘Cambrian Explosion’ in macroeconomic modelling.

From forecasting, there is evidence that combining two or more models leads to greater predictive power than using one model alone (Stock and Watson, 2006; Timmermann, 2006; Silver, 2012). This has been specifically demonstrated for inflation forecasting by Norges Bank in Bjørnland et al. (2012). It is likely to be true qualitatively as well as quantitatively; what one model does not pick up or explain well may be explained by another. A set of models which are distinct, but plausible, will be more informative jointly, perhaps especially so when they disagree. This ‘zoo of models’ approach has also been adopted at the Bank of England (Burgess et al., 2013).

What types of animal are likely to be most useful in the zoo? At a high level, two types: single equilibrium ‘type I’ economic models for dealing with close to equilibrium fluctuations; and more complex, multiple equilibrium ‘type II’ economic models for dealing with far from equilibrium fluctuations. Type I models are stationary and approximately linear. behaviours are well anchored and close to optimizing. Agents’ interactions are predictable and aggregate to something close to a single representative agent (Kirman, 1992; Solow, 2008).

Type II models capture behaviours which are ‘irrational’ or ‘heuristically rational’ and heterogeneous. Uncertainty, as distinct from risk, is acute. Aggregate behaviour in these models is likely to be fat-tailed and often emergent. Type II models are also likely to help educate us by playing out scenarios we did not expect when constructing the model, combining known micro-level features to produce unexpected aggregate outcomes. This is the ‘if you didn’t grow it, you didn’t explain it’ philosophy described in Epstein (1999).

Type II macroeconomic models should be able to explain how economic features, including crises, can arise endogenously. To give a concrete example, a type I analysis of the crisis would invoke an exogenous crisis shock—in which a large number of consumers default on a significant fraction of their loans—to understand how the great financial crisis then
evolved (Kumhof et al., 2015). A type II model would show how crises unfolded, not due to any exogenous shock, but as a natural consequence of the rules and behaviours of the agents within the model of the economic system over time. This provides a mechanism for exploring policies which reduce the frequency or severity of crises.

As Krugman and others have pointed out (Colander et al., 2009; Krugman, 2011), some partial (dis)equilibrium ways of thinking about the economic system could have helped in understanding the crisis before it unfurled, including the ideas of Bagehot (1873), Leijonhufvud (2000), Kindleberger (2001), and Minsky (2008). Before the New Classical Counter Revolution, structural econometric models were also popular (Blanchard, 2017)—the US Federal Reserve’s FRB/US macro model fits into this category (Brayton and Tinsley, 1996) (see Welfe (2013) for a review of these types of model). However, all of these models tend to operate at the aggregate level and in general equilibrium, rather than aggregating from the agent-level. Because of that, they are less well-suited to tackling problems with a high degree of agent heterogeneity or with shifting equilibria.

The hegemony of the ‘representative agent with rational expectations’ approach runs deep in macroeconomics. It has similarities to Newtonian physics. But this quasi-mechanistic view, while sometimes a useful approximation, is not a good representation even of modern-day physics. Modern physics research deals with complex systems, emergent behaviours, vast simulations, and outcomes which are probabilistic and stochastic beyond what is implied by the Gaussian distribution. As an example with parallels in economics, aggregate thermodynamic variables like temperature are not sufficient to describe the rich dynamics of systems which are far from equilibrium; granular descriptions of individual particles and their distributions are required (Turrell et al., 2015a). There are well-developed tools for tackling problems that are not analytically tractable, that include complex behaviours, and that feature a high degree of heterogeneity. In the next section we describe one of them.

IV. The model from Monte Carlo

In the 1930s, physicist Enrico Fermi was trying to solve a particularly difficult problem: the movement of one of the types of particle which make up the atom, the neutron, through background material. This was a tricky calculation as the neutrons had a distribution over several variables including position $r$, energy $E$, and direction $\Omega$ according to a distribution $f(r, E, \Omega, t) drdEd\Omega$.

The full equation tracks the position, direction, and energy evolution of $N$ particles over time with terms representing numerous discrete, continuous, source, and sink interactions of the neutrons. Initially, Fermi and other scientists tried to solve the entire problem analytically in single equation form, but it proved difficult to solve for all but the simplest cases.

Fermi developed a new method to solve these problems in which he treated the neutrons individually, using a mechanical adding machine to perform the computations for each neutron in turn. The technique involved generating random numbers and comparing them to the probabilities derived from theory. If the probability of a neutron colliding were 0.8, and he generated a random number smaller than 0.8, he allowed a ‘simulated’ neutron to collide. Similar techniques were used to find the outgoing direction of the neutron after the collision. By doing this repeatedly, and for a large number
of simulated neutrons, Fermi could build up a picture of the way neutrons would pass through matter. Fermi took great delight in astonishing his colleagues with the accuracy of his predictions without, initially, revealing his trick of treating the neutrons individually (Metropolis, 1987).

The more general method of using random numbers to solve problems soon got a name that reflected its probabilistic nature: Monte Carlo. It was further developed by Fermi, Stanislaw Ulam, John von Neumann, and others (Metropolis and Ulam, 1949; Metropolis et al., 1953; Metropolis, 1987). It found wide usage because of the way it naturally weights the scenarios that are explored by the probability of them occurring. It is efficient in problems with a high number of dimensions, and effective in reproducing all of the moments of a distribution function. Monte Carlo is a standard technique in finance, where it is used to calculate the expected value of assets.

Monte Carlo simulation remains widely used in a range of disciplines known under different names, including individual-based models in biology and ecology, agent-based models (ABMs) in economics, and multi-agent systems in computer science and logistics, as described in Turrell (2016). Recent applications in physics include calculating how beams of particles could destroy cancerous cells (Bulanov and Khoroshkov, 2002; Arber et al., 2015), and how to produce energy from nuclear fusion reactions (Lindl et al., 2004; Spears et al., 2015). They have made a mark in ecology (Carter et al., 2015), where they have been used to model endangered species; in epidemiology (Degli Atti et al., 2008), where they have been used to make detailed predictions of how influenza could spread given demographics and transport links; and for the decentralized behaviour of autonomous vehicles (Ernest et al., 2016). As in physics, their use in epidemiology sees a set of difficult-to-solve differential equations being replaced with simulations of individual agents. Monte Carlo simulation passes the market test, too. It has been used for project finance modelling under uncertainty, forecasting mortgage repayment rates (Geanakoplos et al., 2012), redesigning the rules of the NASDAQ stock exchange (Bonabeau, 2002), and for simulating the transport of people (Heppenstall et al., 2011).

A number of firms, such as Sandtable and Concentric, offer bespoke agent-based models for commercial applications.

Agents in these models might include the consumers in an economy, fish within a shoal, and even galaxies within the Universe, as in Davis et al. (1985). As well as interacting directly with each other, agents might also have a connection to their environment—for instance, banks subject to regulation or whales migrating across the ocean. The behaviours or rules that agents follow depend on the question of interest. Some models have many different types of agent, perhaps firms, workers, and governments. These may themselves differ, so that while all workers have a chance to be employed by a firm and receive a wage, the human capital and marginal propensity to consume of each worker could be different and determined according to an empirical distribution. A schematic of an agent-based model in economics is shown in Figure 7. Heterogeneous agents interact both with each other within a network structure, and with the wider environment.

The important feature of ABMs is that they explain the overall evolution of a system by simulating the behaviour of each individual agent within it and then explicitly combining their micro-level behaviours to give a macro-level picture. Each agent is a

---

1 Monte Carlo techniques are also used in econometrics, in the estimation of Bayesian models.
self-contained unit which follows a given set of behavioural rules. This ‘bottom-up’ approach is very much in the spirit of a ‘microfoundations’ approach, though it differs fundamentally in how it then aggregates to the macro level.

V. How are agent-based models in economics different?

There are important differences between ABMs in the sciences and in economics. In economics, agent-level behaviours are not known to the same level of accuracy as the laws of nature which govern the interactions between, for example, particles. In economics, behaviours can change over time in response to the environment. Agent-level assumptions in economics thus need to be rigorously tested and varied.

Partly as a consequence of the inherent uncertainty in agent behaviour, ABMs in economics can match the data only probabilistically. They tend to match moments and reproduce stylized facts. Some argue that this is naïve compared to more exactly matching the historical evolution of variables over time. But the latter approach, often using forcing processes, seems to suggest an implausible level of precision. An ABM is a way to generate many possible, plausible realizations of variables in exactly the same manner as different possible price paths are generated in Monte Carlo option pricing. Another way to think of this is in terms of the trade-off between bias and variance.
Very broadly, ABMs attain lower bias at the cost of higher variance. Model errors are squared in the bias but only linear in the variance (Friedman et al., 2001). Models with forcing processes aim at low variance at the cost of higher bias.

Where do ABMs fit into the wider modelling landscape? Figure 8 shows this schematically. Models lie on a spectrum. Statistical models rarely say anything about heterogeneous agents. DSGE models say more, and ABMs yet more. But ABMs are not useful for every problem. While in principle it would be possible to use them in forecasting, there are already models potentially better equipped for this, such as dynamic factor models (Stock and Watson, 2011) and machine learning (Chakraborty and Joseph, forthcoming). ABMs are better placed to produce conditional forecasts, where a particular policy is being explored. This is how epidemiologists use agent-based models, too (Degli Atti et al., 2008): rather than attempt to forecast the specific time that a virus outbreak will happen, they identify the risk factors for a virus to break out and subsequently spread. In Figure 8, ABMs nest DSGE models as a special case with little heterogeneity, no stock variables, and a particular set of assumptions about agent behaviour. Broadly defined, ABMs will add most value when problems revolve around heterogeneity, complexity, non-linearity, emergence, heuristics, and detailed rules.

The models in Figure 8 require quite different approaches and ABMs have a modelling philosophy which is distinct from their closest neighbours. The archetypal DSGE model

**Figure 8: A spectrum of macroeconomic modelling**

Notes: Macroeconomic ABMs may be thought of as lying within a wider modelling space, here shown as having two axes. Internal consistency is best represented by strongly microfounded behaviour, while external consistency is demonstrated by agreement with the data. On the other axis is the degree of agent heterogeneity which the model can include, with representative agents on one end and heterogeneity along many dimensions at the other. The within variation of each model type is likely to be larger than the variation between them, but the figure illustrates their approximate location within the wider modelling space.
comes bundled with a number of assumptions, including rational expectations. Rather than offering a ‘core’ model, ABMs are a flexible toolkit for solving complex problems involving heterogeneous agents. One could use rational expectations, but there is no requirement to do so. This flexibility is one reason why ABMs populate every field from military war games to ecology, and why it is impossible to write down a representative ABM.

As a demonstration of this, Table 2 shows a non-exhaustive list of consumption functions in different macroeconomic ABMs. In DSGE models, a core model has emerged partly because the need for analytical tractability forces the modeller to choose from a restricted set of behaviours. ABMs are a generalization to more behaviours and more agents. This means ABMs are often bespoke, adapted to the particular question they are answering. For that reason, they have been criticized as ‘black box’. Their bespoke nature means that there is also a bigger cost to understanding them.

Yet there is no reason why an ABM could not have what Olivier Blanchard (Blanchard, 2017) has outlined as being required for a core macroeconomic model: nominal rigidities, bounded rationality and limited horizons, incomplete markets, and a role for debt. Indeed, these features are found in many macroeconomic ABMs. One of the major differences is the degree to which these models are solved computationally. The process in a typical DSGE model is to specify agents’ behaviours and analytically aggregate them, assuming in the process that markets clear. Usually they would then be linearized and, in the last step, solved numerically. In a typical ABM, the process is quite different. They are solved numerically at the agent level, one behaviour at a time.

This has a number of implications for the strengths and weaknesses of ABMs. Being free of the need to specify equations which can fit together and be solved analytically can be liberating for some problems. This does not mean that ABMs cannot be represented mathematically—there is a theorem which says that all ABMs which can be computed numerically have an explicit mathematical representation2 (Leombruni and Richiardi, 2005; Epstein, 2006). But usually the direct mathematical translation would be too unwieldy to transcribe and would only be interesting in special cases. ABMs are better thought of as algorithms for aggregating the behaviours of individual actors than as systems of equations.

If there is a downside to this flexibility it is that analytical certainty must be replaced with numerical convergence. Sinitskaya and Tesfatsion (2015) highlight this by comparing numerical solutions to a lifecycle optimization problem within an ABM to a known analytical solution. The numerical solution methods do not reach precisely the same outcome as the analytical solution. But for most problems in macroeconomics, the accuracy in how a problem is posed is likely to be a far larger source of error than the lack of precision in the numerical solution of that problem. Indeed, this is another example of the bias-variance trade-off.

The modelling philosophy also differs with respect to interpretation of results. It is useful to think of ABMs as a machine for generating many alternate realizations of the world. Just as in real experiments, it is useful to make many repeats with both control and treatment groups. If the future distribution of one particular variable were truly uncertain, say income \( y \) such that \( y \sim f \), an ABM could be run not just with millions of possible draws from the same distribution \( f \) but also from an entirely different parametric or empirical distribution \( g \). In an ABM, it is possible to do this by changing a single

\[ f \]

2 Every agent-based model is computable by a Turing machine, and every algorithm computable by a Turing machine may be expressed via sets of partial recursive functions.
Table 2: Examples of consumption used in different macroeconomic agent-based models

<table>
<thead>
<tr>
<th>Consumption model description</th>
<th>Consumption model references</th>
<th>ABM references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inflation dependent fraction of permanent income; consumption given by ( c_t = \alpha \hat{y}<em>t ), where ( \alpha = k_t - f(l_t - E</em>{\pi_{t+1}} - r_t) )</td>
<td>( \hat{y}_t ) is permanent income from Friedman (1957).</td>
<td>Salle et al. (2013)</td>
</tr>
<tr>
<td>Aggregate consumption as sum of incomes of all employed and unemployed; ( C_t = \sum_i \sum_j y_{ij} )</td>
<td>Hand-to-mouth consumers, as described in Campbell and Mankiw (1989).</td>
<td>Dosi et al. (2010)</td>
</tr>
<tr>
<td>Fixed propensities ( \alpha_1, \alpha_2 ) out of expected real disposable income and expected real wealth; ( c_t = \frac{1}{E_{\pi_{t+1}}} (\alpha_1 y_t + \alpha_2 W_t) )</td>
<td>Godley and Lavoie (2007)</td>
<td>Caiani et al. (2016)</td>
</tr>
<tr>
<td>Adaptive (with memory parameter ( \xi )) expectation of income and fixed fraction of wealth (based on buffer-stock); ( c_t = \hat{y}_{t+1} + (1 - \xi) y_t + 0.05W_t )</td>
<td>Carroll (1997, 2009)</td>
<td>Assenza et al. (2015)</td>
</tr>
<tr>
<td>Concave, monotonically increasing bounded above fraction of real income; ( c_t = \min \alpha \left{ \frac{y_t}{p_t}, \frac{y_{t+1}}{p_{t+1}} \right} )</td>
<td>Carroll and Kimball (1996); Souleles (1999)</td>
<td>Lengnick (2013)</td>
</tr>
<tr>
<td>Wealth growth ( \Delta W ), average historical consumption of other households ( \bar{c}_t ) and a reference consumption level ( \bar{c}_t ); ( c_t = \bar{c}<em>t + \alpha \frac{\Delta W_t}{\bar{c}<em>t} + \beta (\bar{c}</em>{t+1} - c</em>{t+1}) ) where ( \alpha &gt; 0, \beta &gt; 0 )</td>
<td>Abel (1990); Jawadi and Sousa (2014)</td>
<td>Guerini et al. (2016)</td>
</tr>
<tr>
<td>Numerically solved utility maximization subject to intertemporal budget constraint with leisure ((1 - l)) and wealth ( W_t ); ( \max_{c_t} \left{ \sum_{t=1}^{\infty} \beta^{-t} u(c_t, 1 - l, W_t) \right} ) such that ( W_{t+1} - \rho c_t &gt; 0, c_t &gt; 0 )</td>
<td>Friedman (1957)</td>
<td>Sinitskaya and Tesfatsion (2015)</td>
</tr>
<tr>
<td>Consumption as a fixed fraction of wealth and income; ( c_t = \alpha (W_t + y_t) )</td>
<td>Godley and Lavoie (2007)</td>
<td>Gatti and Desiderio (2015); Gualdi et al. (2015)</td>
</tr>
<tr>
<td>Buffer-stock with mean backward-looking income ( \hat{y}_t ) and target wealth-to-income ratio ( \phi ); ( c_t = \hat{y}_t + \alpha (W_t - \phi \hat{y}_t) )</td>
<td>Carroll (1997)</td>
<td>Chan and Steiglitz (2008); Cincotti et al. (2010); Dawid et al. (2012, 2014)</td>
</tr>
</tbody>
</table>

Notes: For full details see the references.

In more analytical models, it might not be possible at all to do so. Due to the complexity of the interactions within an ABM, it may also be necessary to treat the outcomes as one would real experimental data, with the machinery of hypothesis testing and confidence intervals.
As with all models, it is nonsensical to have more free parameters than imposed or calibrated parameters. ABMs have been accused of being full of free parameters, but this is bad modelling rather than an intrinsic feature of ABMs. One measure of all models is how parsimoniously they describe the empirical data. Macroeconomic ABMs can match both micro and macro stylized facts for most, if not all, the entries in Table 1. For instance, both Gualdi et al. (2015) and Caiani et al. (2016) match the frequency and the extent of the cyclicality in productivity, nominal wages, firms’ debt, bank profits, inflation, unemployment, prices, and loan losses, while Dosi et al. (2015) reproduce, among other features, the distributions of output growth and the duration of banking crises. By simulating every individual actor within an economic system, all moments of distribution functions are accessible to ABMs.

If, as described in Wren-Lewis (2016b), DSGE models maintain internal consistency by sacrificing some external consistency, then ABMs are more of a bridge between internal and external consistency: they sacrifice some internal consistency by allowing agents to have behaviours which are not hyper-rational. Nonetheless, the combination of bespoke models, flexibility, and a tendency to focus on complex systems can mean that communication is a challenge for agent-based modellers. ABMs are rarely as easy to write down in equation form as, say, the simple three-equation New Keynesian DSGE model. There may not always be an easy fix: complex real systems must sometimes be described with complex simulations, at least initially.

Science has had to deal with the trade-off between bigger, more feature-packed models and more abstract but easier to digest models for a long time. Compromises have emerged. In physics, it is common to use a complex simulation of a system as a way of initially exploring hypotheses or discovering new phenomena. Once a specific effect is identified within the complex simulation, a purely theoretical model (or a much simpler numerical model) is built to explain its salient features. An example is in Turrell et al. (2015b), in which a very fast way for lasers to heat matter was first identified in a model with over $10^7$ agents and then explained with a handful of parameters in an analytical model using differential equations. Another example of a theoretical model being overturned by simulation is found in Sherlock et al. (2014). The poor performance of the original theory, accepted for decades, would have been difficult to understand or describe analytically without the use of an ABM.

Gualdi et al. (2015) are the exemplars of this approach in a macroeconomic ABM. A rich and complex model is boiled down to a much simpler ABM which retains, and explains, the same phenomenon. The difference between this approach and simply beginning with a smaller, perhaps analytical, model is that the bigger and more realistic model genuinely surprises the researcher with a relationship or phenomenon that they did not expect. It is then the researcher’s job to unpick this effect and interrogate it further.

The bottom line is that whether an ABM is good or bad will depend on its specific assumptions, how it is used, and how the results are interpreted. The lack of restrictions on modelling assumptions can be a risk. As it is the flexibility of ABMs that gives rise to this risk, the benefits of this flexibility need to be significant to justify this cost. It is argued in the next section that they are.
VI. What could agent-based models do for economics?

The strength of ABMs for economics lies in their flexibility relative to the other models shown in Figure 8. Fabio Ghironi has identified topics in this special issue (Ghironi, 2018) which are likely to be important for macroeconomic models in the future. They are topics which can be difficult to capture in established models. But, as shown in Table 3, ABMs are already delivering new perspectives on each of them.

In general, ABMs are well-suited to situations where interactions between agents really matter, where heuristics dominate, where the heterogeneity of agents is important, where policies have agent-level implications, where granular data are plentiful, and where analytical methods fail. Non-Gaussian distributions, non-linear equations, time-inconsistent choices, boundedly rational behaviours—it is possible to solve all of these numerically at the agent-level. Given that many real world problems involve these features, ABMs have many potential uses in economics (Tesfatsion, 2002).

A great deal of research in networks, particularly as applied to financial systems, has shown that there are emergent properties at the system-level which arise out of interactions at the agent-level (Battiston et al., 2007; Gai et al., 2011). As an example, Bardoscia et al. (2017) show that as banks integrate and diversify at the agent-level, they can increase the system-wide risk of instability because they create cyclical dependencies which amplify stress. ABMs can get at this behaviour where agent-level interactions lead to counter-intuitive behaviour at the macro-level. Herding effects (Alfarano et al., 2005) are another example.

One of the most important use cases of ABMs is to explore microfoundations other than rational expectations. Table 2 shows how ABMs run the whole gamut of behavioural assumptions for consumption. There is an emerging body of experimental evidence suggesting that heuristics can outperform ‘rational’ behaviours (Gigerenzer and Brighton, 2009). To have confidence in the conclusions of ABMs, there is a need for a wider body of work on how best to include realistic agent-level behaviours; see Gabaix (2016) for one example of this. And in understanding agent behaviours, advanced machine-learning techniques are likely to be useful in approximating decisions made by real people. Initial steps down this path have already been taken by DeepMind (Leibo et al., 2017).

ABMs are perhaps most easily distinguished from other models by their ability to model heterogeneity (Hommes, 2006). There is mounting evidence of the importance of

<table>
<thead>
<tr>
<th>Topic</th>
<th>ABM reference</th>
<th>ABM description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial intermediation</td>
<td>Ashraf et al. (2017)</td>
<td>Analysis of the role that banks play in firm entry and exit.</td>
</tr>
<tr>
<td>Heterogeneity of firms and</td>
<td>Assenza et al. (2016)</td>
<td>Exploration of how exporting and domestic firms become heterogeneous in productivity.</td>
</tr>
<tr>
<td>households</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Granularity and networks</td>
<td>Bardoscia et al. (2017)</td>
<td>Shows how diversification can lead to increased systemic risk in networks of banks.</td>
</tr>
<tr>
<td>Policy interdependence</td>
<td>Popoyan et al. (2016)</td>
<td>A macroeconomic model exploring the interdependence between macroprudential regulation and monetary policy.</td>
</tr>
</tbody>
</table>
household, firm and consumer heterogeneity in both status (such as wealth) and behaviour (Guvenen, 2011; Alfì et al., 2009; Gabàix, 2011). A special example is when agents’ stock variables matter. Under the assumptions of the archetypal DSGE model it does not matter that variables such as wealth or debt are not tracked because flows implicitly describe the information about stocks which are relevant for the model. However, there are situations in which richer information on stocks are relevant, such as when heterogeneous agents target a particular level of wealth or debt (Elmendorf, 1996; Muellbauer and Murata, 2009) or when wealth feeds back into behaviour (Cooper and Dynan, 2014).

Models which include accounting identities at the aggregate level have existed for some time (Wren-Lewis, 2016). But the recent ABM of Caiani et al. (2016) has applied them at the agent level to capture the interaction of agents’ decisions through the balance sheet channel. Gatti and Desiderio (2015) show the role of firms’ balance sheets in monetary policy. Policies which act heterogeneous also need to be modelled—for instance, the Bank of England Financial Policy Committee’s policy in 2015 to ensure ‘that mortgage lenders do not extend more than 15 per cent of their total number of new residential mortgages at loan to income ratios at or greater than 4.5’. For monetary policy, one of the important practical channels for influencing consumption relies for its effectiveness on agent heterogeneity: those who gain from policy easing have higher marginal propensities to consume than those who lose (Auclert, 2015).

The chair of the US Federal Reserve, Janet Yellen, has recently commented (Yellen et al., 2016) that:

Economists’ understanding of how changes in fiscal and monetary policy affect the economy might also benefit from the recognition that households and firms are heterogeneous. For example, in simple textbook models of the monetary transmission mechanism, central banks operate largely through the effect of real interest rates on consumption and investment. Once heterogeneity is taken into account, other important channels emerge. For example, spending by many households and firms appears to be quite sensitive to changes in labor income, business sales, or the value of collateral that in turn affects their access to credit-conditions that monetary policy affects only indirectly. Studying monetary models with heterogeneous agents more closely could help us shed new light on these aspects of the monetary transmission mechanism.

An ABM could be used to look at how heterogeneity along several inter-related dimensions affects policy transmission—for instance, marginal propensity to consume as a function of net assets and demographics. Inequality is one type of heterogeneity which is already being explored; see Gibson (2007) and Caiani et al. (2016) for examples where initially egalitarian distributions, such as income, become unequal endogenously. ABMs are not the only way to include heterogeneity, but they do offer an easier route to a higher degree of heterogeneity. Just as Fermi and his colleagues found with neutrons, there is a point where the trade-off between increasingly elaborate mathematical models and numerically solved models favours the latter. New types of DSGE model are closer to that cross-over point, but the other side of the trade-off remains worthy of attention.

The ability to model systems without analytical constraints makes some problems easier to study. Lord Stern, author of a significant review of the economics of climate change (Stern, 2007), has called for agent-based modelling as a way to more realistically incorporate the macroeconomic trade-offs of climate change (Stern, 2016).
Lamperti et al. (2017) for an example of a macroeconomic ABM which takes up this challenge. It facilitates the analysis of systems which are out of equilibrium—systems where markets do not necessarily clear, which are in dynamic disequilibrium, or which transition between different equilibria. One of the unique perspectives from this type of model is showing how a disequilibrium can emerge as the result of the self-interested (but not necessarily hyper-rational) choices of individual agents.

A dramatic example is the non-linear, dynamic macroeconomic ABM of Gualdi et al. (2015), in which output and employment discontinuously collapse to new values depending on the propensity of firms to hire and fire new staff, in addition to firms’ levels of indebtedness. These discontinuous shifts from one equilibrium to another are well known from physics as ‘phase transitions’. They have also been identified in contagion in financial networks (Watts, 2002; Gai and Kapadia, 2010) and in the way that opinions and narratives can shift within a population (Sornette, 2014; Shiller, 2017). De Grauwe (2010) develops an ABM that includes this opinion shift effect in an otherwise New Keynesian set-up.

There are undoubtedly problems, too. Calibration, for instance, is currently a relatively weak area for macroeconomic ABMs. Although the best have been successful in reproducing an impressive range of stylized facts, calibration techniques and standards tend to vary substantially across models. The generalized method of moments is the most commonly used approach (Franke and Westerhoff, 2012), but techniques based on vector auto-regressions are also being developed (Guerini and Moneta, 2017).

ABMs have already delivered impressive results on some partial (as opposed to general) economic and financial systems. ABMs have provided plausible explanations for phenomena in financial markets, including fat tails, clustered volatility, and bubbles (Cutler et al., 1989; Hong and Stein, 1999; Lux and Marchesi, 1999; Alfi et al., 2009). Lux and Marchesi (1999) showed that in order to reproduce the fat tails seen in the distribution of absolute returns in markets, a number of market participants who trade based not on fundamentals, but based on optimism and pessimism (known as noise traders) are required. In their model, agents are able to switch groups based on the performance of strategies. The market fundamentals follow a Gaussian distribution so that the fat-tailed distribution of returns is solely due to the interactions between the different types of traders. In the model, if agents see an opportunity to profit by becoming noise traders, they switch strategies. In the short term, they can ride a wave of optimism or pessimism, and so enjoy larger absolute returns than the variation in fundamentals would imply. But the deviation from underlying value has limits because of the remaining fundamentalist agents and, eventually, prices must partially revert. This simple agent-based model provides a compelling explanation for one of the puzzles of financial markets.

At the Bank of England, ABMs have been used to aid understanding of both the corporate bond market (Braun-Munzinger et al., 2016) and the UK housing market (Baptista et al., 2016). The former model, shown schematically in Figure 9(a), was used to study how investors redeeming the corporate bonds held for them by open-ended mutual funds can cause feedback loops in which bond prices fall. The non-linear loop induces price overshoots in the bond index (shown schematically in Figure 9(b)) and the model looks at possible ways to reduce the extent of the overshooting. The model is calibrated to granular empirical data, and produces a reasonable match to the distribution of daily log-price returns. In one scenario, the fraction of funds using passive trading strategies is increased. The increased presence of passive funds is price and yield stabilizing in the median case, but it introduces a tail-risk of very large price falls (yield
Figure 9: Design of an ABM of the corporate bond market. (a) A schematic of the ABM of trading in the corporate bond market. (b) A schematic of how initial price and yield changes are amplified and propagated in the model.

Source: Braun-Munzinger et al. (2016).
rises), as shown in Figure 10. The reason is that when there are few actively trading funds, the market-maker is more likely to observe a sudden glut of positive or negative net demand and so create major price movements in response. This was not a result which was expected a priori.

A UK housing market model developed at the Bank of England explores the links between macroeconomic stability and house price cycles (Geanakoplos et al., 2012; Erlingsson et al., 2014). Capturing these cyclical dynamics is not straightforward. One potential reason is that the housing market comprises many types of agent: renters, first-time buyers, owner-occupiers, sellers, and buy-to-let landlords. These agents are all represented in the ABM, and are heterogeneous by age, bank balances, income, gearing, and location. Additionally there is a banking sector (a mortgage lender) and a central bank; all are shown schematically in Figure 11(a). The combination of the actions of these agents gives rise to the cyclical dynamics seen in Figure 11(b). These arise endogenously in the model.

The inclusion of an explicit banking sector is important as banks provide mortgage credit to households and set the terms and conditions available to borrowers. The lending decisions of the banking sector are subject to regulation by the central bank, which sets loan to income, loan to value, and interest cover ratio policies. The model reproduces the probability mass function of the share of loans by loan-to-income band given by the Product Sales Database of UK mortgages, as shown in Figure 12. The model has been used to look at several scenarios, including policy scenarios. For example, a policy which applies at the agent-level—such as no more than 15 per cent of new mortgages being at loan-to-income ratios at or greater than a given multiple—serves to dampen boom-and-bust cycles in house prices at the aggregate level.

**Figure 10:** The distribution of outcomes for the median yield of a corporate bond index over the 100 trading days after a sudden increase in the loss rate on the index

---

**Notes:** Percentiles refer to repeated model runs.

**Source:** Braun-Munzinger et al. (2016).
Figure 11: Design and calibration of an ABM of the UK housing market: (a) a schematic of the agents and interactions in the housing market model, (b) a benchmark run of the housing market model showing house price index cycles

Source: Baptista et al. (2016).
VII. Conclusion

Economics has been unusually insular and trust in economists is low. At the centre of this insularity has been a particular type of microfounded behaviour. But the type of microfoundation embedded within mainstream macroeconomic models is far from the only, or in some cases the most plausible, choice. And microfounded models are not the only kind of models that are useful for making sense of aggregate economic fluctuations. A more diverse approach to macroeconomic modelling may be beneficial when making sense of the economy and when setting policy to shape the economy.

Drawing on the last seven decades of simulation, this paper presents one complementary modelling approach. In agent-based models (ABMs), aggregate behaviour is built around the behaviours of individual agents. In that sense, it is very much in the spirit of the ‘microfoundations revolution’, even if the modelling philosophy is different. These approaches have already proven their worth in partial (dis)equilibrium settings, both in finance and the natural sciences.

But there is work to be done before ABMs enjoy the same standing in the mainstream economic literature as their cousins. Among the more fundamental issues which need addressing are:

- what are the most realistic behaviours to incorporate, at the agent-level, once the assumptions of representative agents with rational expectations are relaxed?
- how should information flow between agents within the macroeconomy best be captured? and
- in which situation are ABMs and DSGE models likely to give similar/dissimilar results?

Figure 12: Once calibrated, the housing market ABM reproduces the loan-to-income probability mass function for the United Kingdom

Source: Baptista et al. (2016).
Alongside these technical questions, modellers must also ask themselves how the results of complex simulations of complex systems can best be communicated to researchers and policy-makers alike. Challenges aside, ABMs are a promising complement to the current crop of macroeconomic models, especially when making sense of the types of extreme macroeconomic movements the world has witnessed for the past decade.

References


Wicksell, K. (1918), Wicksell’s metaphor appears in a footnote to his review of Karl Petander’s Goda och darliga tider, Ekonomisk Tidskrift.


The financial system and the natural real interest rate: towards a ‘new benchmark theory model’

David Vines* and Samuel Wills**

Abstract: The 2008 financial crisis revealed serious flaws in the models that macroeconomists use to research, inform policy, and teach graduate students. In this paper we seek to find simple additions to the existing benchmark model that might let us answer three questions. What caused the boom and crisis? Why has the recovery been slow? And, how should policy respond to that slow recovery? We argue that it is necessary to add financial frictions to the benchmark model. This allows us to study the effects of leveraged financial institutions, and of a yield curve based on preferred habitats. Such features will cause endogenous changes in the natural real interest rate and the spread between that interest rate and the rate which influences expenditure decisions. They are likely to radically change the way in which the model responds to shocks. We point to some promising models that incorporate these features.

Keywords: credit frictions, dynamic stochastic general equilibrium, theory, natural real interest rate, risk premia

JEL classification: A23, B22, B40, E30, E40, E44, E50, E51

I. Introduction

Seismologists say that their profession proceeds in two ways: with better models, and more earthquakes. In 2008 there was an economic earthquake, and our models are still catching up. During an earlier earthquake in the 1930s Keynes found that the classical model did not address his policy concerns. In the 2000s we have found that the New Keynesian dynamic stochastic general equilibrium (DSGE) model does not address our concerns. Like the other contributions to this issue, our paper offers a perspective on how to fix the benchmark model.

*Balliol College, Political Economy of Financial Markets Programme at St Antony’s College, and Institute for New Economic Thinking (INET) in the Oxford Martin School, Oxford University; and Centre for Economic Policy Research, e-mail: david.vines@economics.ox.ac.uk

**School of Economics, University of Sydney; Oxford Centre for the Analysis of Resource-Rich Economies, University of Oxford; Centre for Applied Macroeconomic Analysis, Australian National University, e-mail: samuel.wills@sydney.edu.au

We thank Christopher Bliss for very useful comments, and Dylan Smith for helpful conversations about the material in section IV. We are grateful to the authors of this double issue of the Oxford Review of Economic Policy for many useful discussions. Wills acknowledges support of ESRC grant number ES/K009303/1.

doi:10.1093/oxrep/grx061

© The Authors 2018. Published by Oxford University Press.
For permissions please e-mail: journals.permissions@oup.com
This paper is written from the vantage point of the graduate-school classroom. One of us was teaching the existing benchmark model during the crisis, and the other was a graduate student at that time. Both were frustrated at the gap between the models being taught in class and the issues being discussed in practice. In this paper we aim to address that gap.

We seek to find an addition, or additions, to the New Keynesian DSGE model that will help us provide answers to three questions. First, what caused the mid-2000s boom and subsequent global financial crisis? Second, why has the recovery since then been so slow? Third, how should policy respond to this slow recovery? We want to find answers of a simple, teachable kind.

We argue that the most pressing change to the New Keynesian DSGE model needed to answer our questions, although there are many possible answers, is to incorporate financial frictions.

The existing benchmark New Keynesian DSGE model relegates credit to the background (see section II of the paper). The two particular examples of such a benchmark model which we focus on in this paper are those created by Clarida, Gali, and Gertler (CGG, 1999) and the more detailed Smets and Wouters model (SW, 2007), which includes capital accumulation.\(^1\) Both of these models assume that credit markets work without frictions. Households and firms have unlimited access to credit markets which instantly clear, and insurance is fully available against idiosyncratic risks such as losing a job. Monetary policy influences only the price of credit, rather than its availability, and the resulting change in this price is costlessly transmitted to the rest of the economy.

In reality the nature of financial frictions appears to have undergone significant changes since the crisis of 2008. There has been considerable deleveraging by banks, coming from an increase in bank capital ratios (see section III). This deleveraging appears to have caused large and persistent increases in the spread between retail and policy rates, and policies have been adopted to respond to this spread. Furthermore, there have been significant changes in the yield curve, apparently caused by changes in the willingness of banks to undertake liquidity transformation. The policies now being used to manage the recovery from the crisis, like macro-prudential regulation and quantitative easing (QE), appear designed to manipulate the yield curve, a possibility which is not captured in the benchmark model.

We believe that it is necessary to add a microfounded model of credit, and of financial frictions, to the New Keynesian DSGE model, in order to properly respond to these real-world developments.\(^2\) Three methods of incorporating credit frictions into the model appear useful.

Liquidity constraints, as in Ravn and Sterk (2016), lead to counter-cyclical risk premia (that is, the premium is high when aggregate demand and output are low). This leads unemployment-fearing households to engage in precautionary savings. That, in turn, lowers aggregate demand and the natural rate of interest. However, such a framework does not allow households to borrow at all and so is not suitable for studying macroprudential policy levers that have been important since the crisis (see section IV(i)).

---

\(^1\) For additional presentations of the CGG model, see also Woodford (2003) and Galí (2015). It is the SW model which we take as the benchmark DSGE model in our Assessment paper at the beginning of this double issue of the Oxford Review of Economic Policy.

\(^2\) See Reinhart and Rogoff (2011) for an extended discussion of the effects of changes in credit conditions.
Leverage, as modelled by Curdia and Woodford (2016), can endogenously increase the probability of financial crisis. This can also lead to counter-cyclical risk premia—i.e. to risk premia rising when there is a downturn—to compensate banks for the cost of default. Such a framework also allows credit markets to affect aggregate demand and the natural interest rate, and is well-suited to studying unconventional policies like macro-prudential regulation and QE. However, it does not allow for effects of credit conditions on the shape of the yield curve (see section IV(ii)).

A yield curve, driven by a preferred habitat model along the lines of Vayanos and Villa (2009), would address some shortcomings of the other frameworks. Incorporating these features would help us to shed light on the ability of QE to twist the yield curve (see section IV(iii)). It would also remedy a historic omission: the absence of James Tobin’s work on portfolio balance from the two other models which are considered in this paper.

As we noted at the beginning of this paper, and in our Assessment article (Vines and Wills, 2018) at the beginning of this issue of the *Oxford Review of Economic Policy*, there is an analogy here with the developments which occurred in the 1930s. Keynes inherited Alfred Marshall’s classical model. He found, in his discussions at the Macmillan Committee in 1930, that this model could not help him to understand the unemployment which had afflicted Britain ever since the country had returned to the Gold Standard in 1925 at an overvalued exchange rate. He needed to amend the classical model. By the time he had written the *General Theory*, he had added sticky wages in the labour market. This new feature completely changed how the whole model worked. It meant that a fall in animal spirits that lowered investment might mean that unemployment could emerge as an equilibrium outcome. Keynes had to invent the consumption function, the multiplier, the IS curve, and liquidity preference in order to understand the full implications of what he was saying. But now, many years later, we understand these details and can explain Keynes’s ‘new’ system to our students in a straightforward way.

By incorporating credit frictions into our model, we seek to add three new features to the existing benchmark DSGE models. The first is a natural rate of interest that is affected by credit market conditions. The second is a spread between the natural interest rate and the interest rate faced by consumers and investors that is also caused by credit conditions. The third is a market for imperfectly substitutable risky assets that can produce a yield curve which could be twisted by policy. It is our ambition to show that these three additions can change the way in which the whole system works. And we want to be able to explain what happens in a simple way. Indeed, we want to be able to do this in a way which is as simple to explain as the Keynesian multiplier.

### II. The existing benchmark New Keynesian DSGE model

#### (i) A description of the model

We draw a distinction between theory models and policy models, following Blanchard (2018, this issue); Wren-Lewis (2018, this issue) calls the latter ‘structural economic

---

3 We discuss the chain of reasoning in our Assessment article in this issue (Vines and Wills, 2018).
models’. As Blanchard and Wren-Lewis argue, one of the purposes of policy models is to actually fit data, so that they can be used to provide relevant policy advice. In the present paper we are only concerned with theory models, rather than the analysis of policy.

The focus of this paper is the question of how to create a theory model that is suitable for the core graduate-school course in macroeconomics. We use the term ‘benchmark model’ to describe the New Keynesian theory model(s) that is, or are, taught in a graduate-school course. The main ones are the three-equation CGG model, and the more detailed SW model which includes endogenous capital accumulation. Our aim is to ‘fix’ these benchmark models by adding a treatment of credit market frictions.

Of course, when one is trying to fix something, it is first a good idea to get clear what one is trying to fix. The letter inviting authors to contribute to this issue of the Oxford Review of Economic Policy explicitly put forward the SW model, rather than the CGG model, as the benchmark model. It is true that some graduate courses in New Keynesian macroeconomics do not present the SW model, concentrating entirely on the CGG model instead. But this can create confusion for students because, by omitting capital, the CGG model is harder to reconcile with the Ramsey model and the real business cycle (RBC) model, both of which are taught to graduate students in other modules. By contrast, the SW model has the advantage of incorporating all of the Ramsey/RBC insights in a model that also includes the effects of changes in aggregate demand and so is Keynesian in form. As a result, we argue in our Assessment that the SW model is the benchmark New Keynesian DSGE model (Vines and Wills, 2018). Nevertheless, to simplify our exposition in the present paper we use the smaller three-equation CGG model that abstracts from capital accumulation, and so from an endogenous natural level of output. But throughout the paper we discuss what the implications would be of allowing for capital accumulation and for the resulting endogenous natural level of output.

The SW model can be described as follows. There is an IS curve, which has two components. First, there is a forward-looking Euler equation for the consumption of a representative consumer. Second, there is a forward-looking equation for investment by the representative firm which shows that investment is driven by Tobin’s q, the movements of which are determined by the size of capital adjustment costs. The natural level of output is determined by a production function employing the capital stock, labour, and the level of technology. Aggregate demand can differ from the natural level of output because of nominal rigidities, creating an output gap. That gap causes inflation, in a way described by the forward-looking Phillips curve which depends on Calvo price-setting. Monetary policy is represented by a Taylor rule, which may or may not be the product of optimization by a forward-looking central bank. This policy determines the nominal interest rate, and thus the real interest rate, which feeds into both the Euler equation for consumption and the arbitrage equation that determines Tobin’s q.

Collapsing the SW model into a three-equation CGG model means abstracting from investment—consumption is the only form of private-sector demand. The CGG model therefore also abstracts from the effect of capital accumulation on the natural level of output, which will be important in some of what follows. For this reason, when we are discussing the CGG model we will keep track of the effects of any changes to the natural rate of interest rate, and to the natural level of output, in order to bear in mind the way in which this endogeneity works in the SW model.
It is useful for what follows to set out the key equations of the CGG model. These are an IS curve (showing aggregate demand relative to aggregate supply, or the output gap), a Phillips Curve (showing the inflation which results from the output gap), and a Taylor rule (showing the interest rate set by monetary-policy-makers). A fourth equation can be added if the natural level of output is allowed to vary, which changes the natural rate of interest.

\[
\dot{y}_t = E_t[\dot{y}_{t+1}] - (i_t - E_t[\pi_{t+1}] - r^n_t)
\]

\[
\pi_t = \beta E_t[\pi_{t+1}] + \theta \dot{y}_t,
\]

\[
i_t = \phi_y \pi_t + \phi_y \dot{y}_t + r^n_t
\]

\[
r^n_t = \rho + aE_t[\Delta y^n_{t+1}]
\]

In this model the actual level of output is demand determined; the output gap, \(\dot{y}_t = y_t - y^n_t\), is the difference between the actual level of output and the underlying natural level of output, \(y^n_t\), that would be attained in the absence of any nominal frictions. The IS curve comes from a forward-looking Euler equation for consumption: the difference between consumption today and what it is expected to be in the future is driven by the gap between the real interest rate and its natural level. The Calvo–Phillips curve shows an equation for \(\pi_t\)—the rate of inflation relative to the inflation target set by policy-makers—which is driven by expected future inflation and by the output gap. (The parameter \(\beta\) shows the effect of the time rate of discount.) The nominal interest rate, \(i_t\), is controlled by the central bank and is the only interest rate in the economy. It is modelled as following a Taylor rule which depends on the deviation of inflation from its target, on the output gap, and on the natural real rate of interest, \(r^n_t\), which is the real rate that would prevail in the absence of any nominal rigidities (as in Wicksell, 1898, p. 102). The natural level of output is exogenous in this CGG model, unlike in the SW model. It can be assumed to grow along a long-run balanced path due to population growth and technical progress, but that underlying growth process is ignored (again, unlike what happens in the SW model). The natural real interest rate, \(r^n_t\), is exogenous in this model and always equal to the time rate of discount, \(\rho\), except—as shown in the fourth equation—when the natural level of output \(y^n_t\) is expected to change. This system has the property that in the steady state the inflation rate will equal its target, output will be at its natural level, and the real interest rate will equal its natural level.

The only friction in the model is that prices are adjusted infrequently and in an overlapping manner. We assume that this happens with a fixed probability for any particular firm (Calvo, 1983); this probability affects the size of the coefficient \(\Theta\). To allow prices to be set like this, there are many firms who compete monopolistically. The distortion

---

4 The equations of the SW model are set out in Appendix II of our Assessment article (Vines and Wills, 2018).

5 For a full derivation see Clarida et al. (1999), Woodford (2003), or Galí (2015).

6 The coefficient \(\alpha\) depends on underlying microfounded parameters of the model, including the intertemporal elasticity of substitution. The model can be extended to make this natural level of output a function of capital, as in the SW model.
to the natural level of output caused by monopolistic power is corrected by a simple subsidy to firms (financed by lump-sum taxes), but the distortion caused by infrequent price-adjustment and the resulting dispersion of relative prices remains present in the model; this is what makes deviations of inflation from its target costly.

In such a set-up the job of the policy-maker is relatively easy. In the face of technology or demand shocks, a policy-maker following the above Taylor rule will vary the nominal interest rate, and in turn the real interest rate, in such a way that perfectly stabilizes inflation (assuming some conditions for determinacy are observed—Sargent and Wallace (1975)). As nominal rigidities are the only friction in the model, stabilizing inflation also closes the output gap, in a result that has been described as a ‘divine coincidence’ (Blanchard and Galí, 2005). The result of following this Taylor rule is that, in the absence of inflation, the nominal interest rate will perfectly track the natural real interest rate, $r^*$. But cost-push shocks to inflation introduce a trade-off between stabilizing inflation and the output gap.

Both this CGG framework, and the more general SW model which it simplifies, have many advantages. However, the framework which they present has important omissions. It relies on a single interest rate and so does not allow for frictions in the transmission of monetary policy. Indeed, the only role for credit in this model is through the effects of the interest rate on consumption; and since there is just one representative consumer, and there are no investors, everything that is produced is consumed and there are no credit flows at all. The stock of debt does not matter (in the simple model presented here, there is none), nor does it matter that banks create money, with the associated asymmetries in information, constraints on liquidity and capital, and varying risk exposures over time. Financial markets are complete, so that any idiosyncratic unemployment risk can be insured away, and individuals don’t suffer from losing their jobs; they are always on their labour supply curve. Price-setting occurs in a very stylized way. Actual variables don’t influence natural variables. By relying on this model for teaching we are implicitly telling our students that once inflation has been controlled, and the output gap has been closed, then our job as macroeconomists is done. But there are important omissions in this treatment. In particular, they clearly limit the ability of the model to deal with our three key questions: why was there a boom before the Great Financial Crisis, why has the subsequent recovery taken so long, and what can policy do about it?

(ii) Attempting to explain secular stagnation using the existing benchmark model

One school of thought argues that our questions can be answered without radically changing the benchmark SW model, but simply by changing the exogenous variables that are fed into it. The ‘secular stagnation’ hypothesis argues that the slow recovery since the crisis may have been caused by slower population growth, a reduced rate of technological progress, and an increase in savings (Summers, 2014, 2015; Gordon, 2012; Eichengreen, 2015).

In this argument the first two changes—slower population growth and a reduced rate of technological progress—will reduce the demand for investment in a way which can be captured within the SW framework without appealing to credit market conditions.
They lower the natural rate of interest, but do not affect the gap between that interest rate and the interest rate affecting households. Investment demand is lower because less population and technology in the future will also require less capital along the Ramsey growth path. This may have a very large effect on the short-run flow of investment, since there is a need for a lower long-run stock of capital per effective worker. The short-run effect on investment might be particularly large if the adjustment costs are small. All else equal, a large fall in investment demand will cause a large fall in the natural rate of interest.

In addition, global savings rates have risen, particularly in China. Higher Chinese savings have been attributed to marriage market competition (Wei and Zhang, 2011), and the demographic shock of the one-child policy (using an overlapping generations framework by Coeurdacier et al. (2015)), see Figure 1. While in the SW model and CGG such increases in savings cannot be studied, an extension of the SW model in an overlapping generations manner—to make possible this study—is relatively straightforward.

*Figure 1: China population profile, 2016*

Note: China’s one child policy created large deviations in its population profile. A large generation will soon be retiring, to be replaced by a much smaller one entering the labour force.


See McKibbin and Vines (2000) for an analysis of the East Asian financial crisis of 1997–8 along these lines. This approach is less effective in explaining the rapid boom in the run-up to the global financial crisis in 2008, since there was no obvious increase in the rate of population growth or in the rate of technical change. Perhaps it sees the increase in investment in US housebuilding as resulting from an increase in the ability of sub-prime borrowers to purchase housing (i.e. an increase in sub-prime entitlements) leading to a shortage of housing, or perhaps it requires a reduction in the risk premium which takes it into the territory discussed below.

See Summers (2014, 2015) and Coeurdacier et al. (2015) and our discussion in section III. Such an increase in savings is not something that can be explained within the SW framework as it stands. But one can add something like an overlapping generations framework to this set-up, so that one might provide this explanation without appealing to credit-market conditions.
There is also evidence that the natural rate of interest has fallen considerably since 2008. Coeurdacier et al. (2015) suggest that high Chinese savings led to global imbalances and to a sustained fall in the global real interest rate. Barsky et al. (2014) find that the natural rate fell from approximately 4 per cent in 2007 to –6 per cent in 2009, where it has remained since, using an estimated variant of the SW model. Carvalho et al. (2016) also find that higher longevity and lower population growth has caused a long-run fall in the equilibrium real interest rate.

Nevertheless, slow population and technology growth and intergeneration savings effects do not appear to be enough to explain the low natural real interest rate since 2008. Furthermore Barsky et al. (2014) find that the natural rate is both volatile and pro-cyclical, which the SW model cannot help us to explain.

Instead, we must turn to other explanations, which introduce countercyclical fluctuations in the natural rate of interest, and a gap between this interest rate and the interest rate which influences spending.

III. Real world changes in financial frictions

Since the 2008 crisis, credit markets have changed considerably. Figure 2 shows that during the 2008 crisis the spread between banks’ borrowing and lending rates increased sharply, as did the spread of both over the bank rate. The increased spread between borrowing and lending persisted for over 2 years, and the spread over the bank rate for much longer.

**Figure 2**: UK interest rates, 2003–12, %

Note: Since the financial crisis the spread between borrowing and lending rates, and between both and the bank rate, has increased persistently. Secured loans are 2-year, 75% LTV, fixed-rate mortgages. Deposits are fixed-rate bond deposits.


9 If such high Chinese saving is responsible for the low natural interest rate globally, then the problem might be solved in the coming 5 years as an unusually small generation (the children of the first one-child generation) enters the workforce.

10 Summers (2014) argues that only 10 per cent of the reduction in potential GDP can be attributed to slower technological growth.
Since the crisis the capital ratios of banks also increased considerably (see Figure 3). Prior to the crisis, an average of 6 per cent of the risk-weighted assets of major UK banks was held in Tier 1 capital. By the time the Basel III accord was introduced in 2012 this had risen to 10 per cent. Since then the capital ratio has risen a further 6 per cent as a result of the new requirements imposed by the Basel III accord.

Furthermore, during the 2008 crisis the shape of the UK yield curve changed considerably, becoming dramatically steeper as the Bank of England lowered the policy rate while longer-term yields remained largely unchanged (see Figure 4).

It is now thought that these changes in the credit markets are important in understanding both the crisis and the recovery. Summers (2015) suggests that the changes have contributed to the current climate of low investment and high savings, in addition to the demographic factors outlined in section II(ii). Investment may be low because the relative price of capital goods has declined; the process of innovation requires less

**Figure 3:** Major UK banks’ capital ratios

![Chart showing major UK banks’ capital ratios](chart3.png)

*Notes:* After the 2008 crisis the Tier 1 capital ratios of UK banks increased significantly, and this has continued since the introduction of the Basel III accord in 2012.


**Figure 4:** UK government liability spot yield curve

![Chart showing UK government liability spot yield curve](chart4.png)

*Note:* During the financial crisis the UK gilt yield curve became dramatically steeper.

capital; financial intermediation has become more costly, due to larger spreads of the
type shown in Figure 2; and collateral requirements are higher. Similarly, savings may
be higher because loans require more collateral ex ante, and costly financial interme-
diation is causing firms to finance investment from retained earnings. Eggertsson and
Mehrotra (2014) and Hamilton et al. (2016) argue that a deleveraging shock has created
an oversupply of savings.

If credit markets were completely independent of monetary policy, then such effects
could be studied separately. However, changes in credit markets are influenced by mone-
tary policy, which is why study of them must be embedded within a full macroeconomic
model. Giavazzi and Giovannini (2010), for example, argue that inflation targeting can
‘increase the likelihood of a financial crisis’, and Summers (2015) notes that there are ‘financial stability consequences of protracted periods of zero interest rates’.

A more nuanced, microfounded treatment of credit markets is thus needed for the
benchmark model to shed light on policies employed since the crisis. Macroprudential
policy has been introduced in a number of places to stabilize the risks of financial crisis,
distinct from monetary policy’s aim of stabilizing inflation, in the interests of matching
the number of instruments to the number of targets (Tinbergen, 1952). To appropriately
comment on macro-prudential policy, the core model must have a role for its most
likely levers: reserve requirements and lending standards. As macro-prudential and
monetary policies will not be conducted in isolation, credit frictions must be nested
inside a model that describes both. Quantitative easing has also been used extensively
in Japan, the UK, the US, and Europe, and understanding this again requires financial
intermediation and a term structure of interest rates.

IV. Endogenizing financial frictions in the benchmark model

(i) Incorporating liquidity constraints

One way of incorporating credit frictions into the benchmark model is to restrict
households’ access to borrowing. Recently work has begun on a heterogeneous agent
new Keynesian framework with search and matching frictions in the labour market
(HANK–SAM). Kaplan et al. (2016) impose constraints both on household liquidity
and the ability of households to own firm equity; they need to solve their model numer-
ically. Ravn and Sterk (2016) simplify this set-up by assuming that households cannot
borrow at all, and cannot insure against idiosyncratic unemployment risk. The result
is a model with three types of households: asset-poor households who are employed,

11 The 2012 Basel III accord allows for countercyclical bank capital buffers to be applied by national
regulators. The UK Prudential Regulation Authority (PRA) was established in the same year with that purpose
in mind. The US Fed implicitly applies countercyclical bank capital ratios in the way it defines its crisis
scenarios for stress-testing. A crisis is defined as unemployment rising by 4 percentage points, or reaching 10
per cent, whichever is greater. When unemployment is low these scenarios become more conservative, and by
extension so, too, do capital ratios.

12 Such as loan-to-valuation ratios, used widely, or restrictions on mortgage lending to investors, as in
Australia.
asset-poor households who are unemployed, and asset-rich households. The model turns out to be simple enough that closed-form solutions are possible. The set-up is able to provide important insights into the distributional effects of shocks, and the role of precautionary savings which emerge as a consequence of labour-market uncertainty.

The model has two important frictions: sticky prices, which appear in the typical way in the New Keynesian Phillips Curve, and incomplete financial markets, which are important because they affect aggregate demand through precautionary savings, the need for which arises as a result of uncertain labour-market conditions. A sketch of the model is provided by the following set of equations.

\[ \dot{c}_{e,t} = E_t [\bar{c}_{e,t+1}] - (i_t - E_t [\pi_{t+1}] - r^*) \]

\[ \pi_t = \beta E_t [\pi_{t+1}] + \delta \bar{m}c_t(\theta_t, \theta_{t+1}, A_t) \]

\[ \theta_t = \gamma \dot{c}_{e,t} \]

\[ i_t = \phi_x \pi_t + \phi_y \dot{\theta}_t \]

\[ r^* = g(E_t[\theta_{t+1}]) \]

The set-up modifies the equations presented previously in the following way. The first equation is a modified Euler equation for \( \dot{c}_{e,t} \), the consumption of those asset-poor households who are employed. (The behaviour of this variable drives the dynamics of the model, since unemployed asset-poor households are unable to borrow, and the behaviour of asset-rich households is ignored.) The last term in this equation captures the fact that the natural rate of interest is now endogenous, as shown in the final equation; this captures the wedge in the natural interest rate introduced by the wish for precautionary savings due to the risk of unemployment. A less tight labour market increases the risk of unemployment, and in turn increases precautionary savings, so households must be offered a lower interest rate to be prevented from saving. This is the first key new idea in the model. The second equation shows that the incomplete-markets wedge modifies the Calvo–Phillips curve so that the level of marginal costs, \( \bar{m}c_t \), now depends on technology, \( A_t \), and on the current and future values of labour-market tightness, \( \theta_t \). The second important new idea in the model is that at low levels of labour-market tightness, an increase in aggregate demand has a much smaller marginal effect on inflation. This non-linearity in the Phillips curve arises because there is little hiring taking place at low levels of demand, so that any change in demand hardly increases the inflationary pressure caused by firms bidding for workers. The third equation captures the fact that everything that is consumed must be produced using labour, which in turn affects labour-market tightness, \( \theta_t \). The next equation shows the Taylor rule for the interest rate, which in this set-up depends on the degree of labour-market tightness, rather than on the output gap. The final equation shows the way in which the natural rate of interest \( r^* \) depends on labour-market tightness; a less tight labour market will lower the natural rate of interest.
This framework has several appealing features, not least its tractability. It creates the possibility of three different steady states. The first is the ‘typical’ steady state: in normal times when the labour market is well-utilized, the behaviour of the model around the steady state has features typical of the CGG model. The second is a liquidity trap at the zero lower bound (as in Eggertsson and Woodford, 2003; Eggertsson and Krugman, 2012). The third is a new ‘unemployment trap’ characterized by low hiring. In this equilibrium there is low demand, due to precautionary savings by the employed caused by the incomplete-financial-markets friction; and low inflation, due to the labour market friction of sticky prices. However, inflation is not very low (or negative) because of the non-linearity in the Phillips curve identified above. This means that the Taylor rule does not give rise to a monetary policy which is sufficiently loose to prevent the low-level equilibrium from emerging.

This last feature offers a useful perspective on secular stagnation, because it places precautionary savings by unemployment-fearing workers at centre stage. The model also offers a channel through which monetary policy can affect risk premia, and in turn business cycles. In the benchmark model a weak labour market will cause households to smooth consumption and save less. In this framework, a weak labour market increases precautionary savings, as workers prepare for the possibility of unemployment, which further weakens labour conditions. The additional savings lower the natural real interest rate, as seen in the equation for \( r^n \). This is consistent with empirical evidence for the co-movement between labour-market tightness and the real interest rate which is observed in reality (Barsky et al., 2014; Ravn and Sterk, 2016). Monetary policy can reduce the movements in risk premia described by the model by stimulating aggregate demand when it is low, and vice versa.

But the shortcoming of this approach, for the purposes of the discussion in this paper, is that it gives no role to credit and leverage. In this framework, households are restricted from borrowing completely. There is thus no scope for analysing macroprudential policies, like capital adequacy ratios and loan-to-valuation ratios, which are currently under consideration by policy-makers globally.

(ii) Incorporating leverage and risk premiums

An alternative avenue for incorporating credit frictions into the benchmark model is to introduce a stock of leverage, which affects both an interest rate spread, and the natural rate of interest.

The classic approach for incorporating credit is the ‘financial accelerator’ framework of Bernanke, Gertler, and Gilchrist (BGG, 1999). The financial accelerator describes how exogenous macroeconomic shocks are amplified and propagated by credit markets. In the original paper, unleveraged banks lend to leveraged firms. The lending contract is nonlinear: in good times firms keep their excess profits over the interest rate, while in bad times they go bankrupt and banks seize the residual value of their capital. Banks must charge a risk premium, or spread, over their cost of finance to cover the real cost of ‘inspecting’ bankrupt loans. As firms’ net worth is procyclical (due to profits and asset prices), the risk premium is countercyclical, which enhances the swings in borrowing, investment, spending, and production. This type of analysis requires the full equipment of the benchmark SW model. In the absence of the bankruptcy costs, it collapses.
back to that model. But by including bankruptcy costs the model introduces a wedge between the interest rate received by savers and that paid by investors, in ways which are important for our purposes. Extensions to this model allow for leveraged banks to borrow from households, who also receive a bankruptcy risk premium over the risk-free rate (Luk and Vines, 2011; Ueda, 2012). Nevertheless, this kind of model is disappointing in that, with reasonable parameter values, the quantitative importance of the resulting changes in the risk premium does not seem to be large (see Luk and Vines, 2011).

More recent, and simpler, examples are provided by Curdia and Woodford (2010, 2011, 2016) and Woodford (2012), who relax the benchmark model’s representative agent assumption and allow for two types of households: borrowers and lenders. Each type of agent differs by their marginal utility of consumption, which introduces a motive for financial intermediation. Borrowers face an interest rate spread over the policy rate driven by two financial frictions: a real cost of originating and monitoring loans (or a risk premium), and a mark-up in the financial sector from monopolistic competition. The model can be summarized in the following sketch,

\[
\dot{y}_t = E_t[\dot{y}_{t+1}] - \left( ii(i_t, L_t) - E_t[\pi_{t+1}] - r^a_t \right)
\]

\[
\pi_t = \beta E_t[\pi_{t+1}] + \theta \dot{y}_t + \Omega(L_t)
\]

\[
i_t = \phi_x \pi_t + \phi_y \dot{y}_t + r^a_t
\]

\[
r^a_t = \rho + a E_t[\Delta y^a_{t+1}(L_{t+1})]
\]

\[
L_{t+1} = \rho L_t + \gamma \dot{y}_t + \epsilon
\]

where \(ii(i_t, L_t)\) is the average rate of interest for borrowers and savers, which is a function of the policy rate, \(i_t\), and a credit spread which depends on the stock of leverage, \(L_t\). The Phillips curve now includes a real distortion from credit frictions that increases with leverage, \(\Omega(L_t)\), and raises the marginal cost of supplying goods to borrowers. This in turn raises inflationary pressure for any given level of real activity. An increase in leverage allows the natural level of output to expand, but it also introduces default risk into the economy; these effects both influence the rate of interest \(r^a_t\). Leverage evolves dynamically as a function of stock of existing leverage and the output gap in the previous period.

Incorporating credit into the model in this way has some advantages. Unlike the baseline model, it attempts to model the spread shown in Figure 2, allowing it to vary with the stock of leverage in a way which influences the natural rate of interest. This is in addition to the influence of capital accumulation on the natural rate of interest which is seen in the SW model. Doing this goes towards modelling a core aspect of past financial crises, as noted by Reinhart and Rogoff (2011) and by financial market participants such as Soros (1987).

Second, such a set-up breaks the ‘divine coincidence’ which was discussed earlier in relation to the benchmark—allowing feedback from actual variables to natural variables.
(see Woodford, 2012). In models with credit frictions like this, monetary policy has real effects on the flexible price equilibrium (see also De Fiore and Tristani, 2011). Curdia and Woodford (2016) show that in such a setting, the central bank’s objective becomes one of stabilizing not just inflation, and thus closing the output gap, but also one of moderating the risk of future financial crisis. Thus, if the interest rate is the only policy instrument available—i.e. in the absence of macro-prudential policy—the central bank should be willing to lean against credit booms, even at the expense of disturbing its targeting of inflation and the output gap.

Third, it opens the possibility of studying other unconventional policies. Quantitative easing has already been studied in this framework by introducing a central bank that holds commercial bank reserves (Curdia and Woodford, 2011). However, the framework does not yet include a yield curve, which, as discussed below, seems important for a study of quantitative easing.

It would be possible to introduce a macro-prudential policy-maker into the Curdia and Woodford (2016) model. Giving such a policy-maker a distinct objective function and a separate set of tools—like the loan-to-valuation ratio or the kind of bank capital requirements depicted in Figure 3—might make it possible for such a model to provide useful and tractable insights into the interaction between macro-prudential and monetary policy. A goal for that direction of inquiry may be to produce a macro-prudential rule akin to the Taylor rule for monetary policy.

Woodford (2012) offers a simple treatment of a model of this kind in which the strength of financial frictions switches endogenously between high and low states. Curdia and Woodford (2016) extend this set-up to that of fully microfounded setting where financial frictions evolve as a forward-looking average of all future credit spreads.

(iii) Incorporating portfolio balance, multiple maturities, and the yield curve

A third approach to credit frictions is to incorporate a yield curve. The benchmark models, and each of the models discussed above, do not include a yield curve. All the interest rates in BGG, Curdia and Woodford, and Ravn and Sterk, are rates on short-term loans. The models assume that the expectations-augmented theory of the yield curve holds; so that a long-term loan is equivalent to rolling over, period-by-period, a sequence of short-term loans charging the short-term rate. But interest rates at different maturities might well diverge from the expectations-augmented yield curve implicit in the benchmark model, because demand for, and supply of, financial assets might not be substitutable across different time periods. The resulting shape of the yield curve might come to influence investment, the natural level of output, and the natural rate of interest.

If the shape of the yield curve matters, then studying modern policies like quantitative easing in models which include only short rates (like Curdia and Woodford, 2011) will be inadequate. In fact, quantitative easing since the crisis has aimed to influence the shape of the yield curve by purchasing long-term assets in exchange for short-term ones. If the expectations-augmented yield curve theory held, then such purchases would not have affected asset prices as they did.
To properly study the yield curve, the benchmark model needs to better incorporate risk aversion. This research agenda would essentially bring James Tobin’s insights on portfolio balance into modern benchmark models. In the second half of the twentieth century Tobin introduced to macroeconomics the idea that volatility—a second-order phenomenon—can have first-order effects. This gave rise to his theory on portfolio balance, which acknowledges that there exist a wide variety of assets that vary in risk and return, and established that the mixture demanded by people will depend on their risk aversion. By contrast, in the Curdia and Woodford model discussed above, the spread between the policy rate and the interest rate in the IS curve does not rely on risk-averse asset holders, but instead on leverage and the costs of financial intermediation. And in BGG the spread again depends on leverage, which increases the probability of bankruptcy losses. But banks are risk-neutral, and only charge an interest rate spread which covers only the expectation of those losses; when productivity falls this spread rises, in line with the expected costs of default.

To understand why the risk premium varies at different maturities, we must find an alternative to the expectations-augmented yield curve. One promising avenue is micro-founded models that incorporate the preferred habitats hypothesis in which borrowers and lenders have different characteristics at different maturities.

The preferred habitats approach uses modern portfolio theory to analyse how investment portfolios are allocated at various maturities, incorporating varying risks and degrees of risk aversion. Such analysis is in its infancy. A relatively recent and useful contribution is that by Vayanos and Villa (2009). The model consists of households who supply bonds and possess a preference for maturity that varies over their lifetime, giving rise to habitats. Arbitrageurs bridge the gaps between habitats by buying and selling bonds of different maturities. These arbitrageurs process limited liquidity and are risk averse. For example, a household may need to borrow to buy a house, and so will issue a bond of a particular maturity, whose date of final payment depends—say—on the expected date of receipt of income in the future. An arbitrageur will need to purchase this bond, in exchange for assets at a different maturity issued by other households. Such a purchase will expose the arbitrageur to liquidity risk, for which he or she will charge a risk premium, the size of which depends on his or her degree of risk aversion. An increase in risk aversion by these arbitrageurs will lead to an increase in the natural rate of interest. In turn this will reduce investment and so the natural level of output.

The model put forward by Vayanos and Vila is so far only a partial-equilibrium model; it has not yet been nested within a full DSGE framework. Doing this would create a theory of the yield curve that is useful for macroeconomic analysis. For example, the use of such a model might shed light on how an increase in savings might increase the spread between the long-term rate of interest and the long-term rate of interest influencing investment, since those holding the savings might be unwilling to hold extra-long-term assets unless they were offered extra yield to compensate them for a loss in liquidity. A framework of this kind could also allow us to study quantitative easing, by examining its effect on the yield curve, rather than by examining its effect on central bank reserves, as is done by Curdia and Woodford (2011). One might also study how macroprudential policies affect different parts of the yield curve.
V. Conclusion

The Great Depression of the 1930s revealed serious flaws in the classical model of Alfred Marshall, which Keynes set out to remedy. The 2008 financial crisis similarly revealed flaws in the current benchmark New Keynesian DSGE models. This paper has attempted to describe simple ways to remove one of the apparent flaws—the lack of convincing financial frictions—in a way which might help us answer our three questions: what caused the boom and crisis, why has the recovery been so slow, and how should policy respond? We have described some existing models that might do this in a sufficiently parsimonious way as to be taught to first-year graduate students. While all those described hold promise, the framework in Curdia and Woodford (2016) is of particular interest because of its ability to capture the dynamics of leverage that appear to have been important during the crisis and the recovery. Properly incorporating financial frictions into the benchmark DSGE model will radically change the way it responds to shocks: by allowing for endogenous movements in the natural real interest rate and the spread between that interest rate and the rate which influences expenditure decisions. This will give an appropriately central role to leveraged financial institutions, and the policies that affect them, in the models we offer to the next generation of economists.

References


Wicksell K. (1898), Interest and Prices.
DSGE models: still useful in policy analysis?

Jesper Lindé*

Abstract: This paper discusses the usefulness of DSGE models in monetary and fiscal policy analysis. While the recent crisis has exposed some weaknesses in these models, I argue that DSGE models currently have few contenders to replace them as core models in the policy process. The prominent role for forward-looking behaviour and their simplicity make DSGE models very suitable for policy analysis. In addition, DSGE models are flexible enough to be used for many purposes, while other models are often more limited in terms of the questions they can address. As a result, I argue that improved DSGE models—modified to take the lessons of the recent crisis into account—will remain as a workhorse tool in many policy institutions for a long time to come.

Keywords: macroeconomic models, New Keynesian economics, financial crisis, core and satellite models, monetary and fiscal policy, criteria for useful policy models

JEL classification: E52, E58

I. Introduction

Designing quantitative macroeconomic models for monetary and fiscal policy analysis has proved to be a daunting task. By extending the basic stochastic growth model taught in macroeconomics PhD courses with some real rigidities in the form of habit persistence in consumption, variable capacity utilization, investment adjustment costs, and allowing for gradual nominal wage and price adjustment, Christiano, Eichenbaum, and Evans (2005; CEE henceforth) managed to develop a general equilibrium model which featured an empirically realistic monetary policy transmission mechanism for a key set of aggregate quantities and prices. Building on this structure, Smets and Wouters (2003, 2007) showed that a variant of the CEE model amended with a suitable set of shocks could match all the variation in seven key time series for the United States. In

* Sveriges Riksbank and CEPR, e-mail: jesper.l.linde@gmail.com

I am grateful for very useful comments from Stefan Laséen, Doug Laxton, Ulf Söderström, David Vines, Anders Vredin, and Samuel Willis. Discussions with and support from Olivier Blanchard and David Vines have also been instrumental in the completion of this article. This paper was written while the author was a resident scholar at the International Monetary Fund, and the author is grateful to the fund for its hospitality and stimulating environment. Any views expressed in this article are solely the responsibility of the author and should not be interpreted as reflecting the views of the Executive Board of Sveriges Riksbank or those of any other person associated with the Riksbank.

doi:10.1093/oxrep/gxr058
© The Author 2018. Published by Oxford University Press.
For permissions please e-mail: journals.permissions@oup.com
addition, Smets and Wouters (2007) explicitly showed that the forecast performance of their model was not evidently subpar relative to a Bayesian vector autoregressive model (VAR), arguably a reasonable multivariate benchmark.

Hence, the interest in using these relatively simple models in monetary policy analysis exploded and, as documented in Coenen et al. (2012), many central banks and leading policy institutions (e.g. the International Monetary Fund (IMF) and the European Commission) developed and started to use dynamic stochastic general equilibrium (DSGE) models extensively prior to the global financial crisis. In some institutions, such as the US Federal Reserve, DSGE models were mostly used to generate alternative scenarios around the baseline forecasts. At these larger institutions, other models were often used to construct and interpret the baseline forecast. But in some small institutions with more limited resources (and hence lacking the resources to develop and maintain several structural macro models), DSGE models were also sometimes used to inform the staff and policy-makers about the baseline projection.1

Following the unexpected sharp decline in economic activity in the US and around the world in the autumn of 2008 and spring of 2009, a hotly debated issue has been the role and type of models used by central banks and other key policy institutions; see, for example, Krugman (2009), Blanchard (2016), and the other contributions in this issue.2

Critics of DSGE models have put forward the view that modern macroeconomic tools developed during the ‘Great Moderation’—particularly DSGE models—proved to be of little use when the United States and the rest of the world entered the Great Recession (see, for example, Hendry and Mizon, 2014; Romer, 2016; Hendry and Muehlbauer, 2018; and Stiglitz, 2018). Others—see, for example, Blanchard (2016) and the contributions in Gürkaynak and Tille (2017)—seem to articulate the view that there is nothing fundamentally flawed with DSGE models per se and that they can be useful for many purposes, although the particular models actually used in policy analysis as we entered the recession did not include all relevant features (e.g. more elaborate financial sector and intermediation).

The different views on the usefulness of DSGEs for addressing pressing policy issues leads to a divergence in thinking as to how to go forward. Some seem to think that DSGE models used by policy institutions should be revised to better account for financial frictions, non-linearities, and heterogeneity. Others seem to think that they should be abandoned altogether for alternative approaches; these observers seem to think that many assumptions in these models are fundamentally flawed: rational expectations, homogeneous agents, efficient markets, one stable equilibrium, etc. For instance, Romer (2016) writes in his paper: ‘For more than three decades, macroeconomics has gone backwards. . . models attribute fluctuations in aggregate variables to imaginary causal forces that are not influenced by the action that any person takes.’

While I have invested considerable amounts of my time and human capital in DSGE models, my prior certainly covers the possibility that DSGEs may be too misspecified

1 For example, the Sveriges Riksbank (Swedish Central Bank) started to use its DSGE model Ramses as a core model in 2006. Sims (2007) recognizes that the Riksbank was the first institution to use an estimated DSGE model as a core model; see Adolffson et al. (2007) for details of this model and how it was used in the policy process. Norges Bank (see, for example, Brubak and Sveen, 2009) and Czech Central Bank (see, for example, Clinton et al., 2017) were also early in adopting DSGE models as a core tool in the policy process.

2 In addition, Blanchard (2017a) provides a useful list of references.
and simplistic to be useful for normative (welfare) analysis. However, given that we have few or no alternative approaches today that are more helpful than DSGEs in daily quantitative policy analysis at central banks and treasuries, my punch line in this article is to argue that DSGE models will continue to be a core tool in positive policy analyses, at least in the foreseeable future. In a nutshell: it takes a new (better) model to beat an old model. Even so, I should be careful to say that I believe that other models—both with forward- and backward-looking behaviour—can be important complements and sometimes even substitutes to DSGEs, depending on the question addressed and resources available for model building and maintenance.

Throughout the paper I use the perspective of a policy economist at a central bank or treasury serving his or her principals. I think this perspective is fitting, and my only comparative advantage relative to the other prominent economists contributing to this special issue. In a sense, my paper is less ambitious than some of the others: I simply argue that DSGEs will play an important role going forward; I leave to others to discuss how DSGEs should be improved. Even so, I outline some criteria that I believe future models must pass to serve as a core tool in the policy process.

The remainder of this article is organized as follows. In section II, I put the critique against DSGEs, and the need for a complete rethinking in macroeconomics in some historical perspective. I also briefly compare the forecasting properties of DSGE models to a sophisticated reduced form model during the most acute phase of the crisis; a period during which DSGEs are said to have performed very poorly. Moreover, I argue that DSGEs, notwithstanding their severe limitations in forecasting performance, were quite helpful to policy-makers during the crisis. In section III, I discuss where we go from here, the role DSGE models will play in future policy analysis. I set up a number of requirements for a core model in the policy process and argue that DSGEs, despite their evident drawbacks, are among the very few models out there which satisfy these criteria. In this section, I also present some suggestions on how we can develop and organize the work on so-called ‘core’ and ‘satellite’ models. Finally, in section IV, I offer some brief concluding remarks.

II. The critique against DSGEs and their usefulness during the crisis

When Keynes transformed macroeconomics in the 1930s, GDP in advanced economies including the United States was way off its long-term growth trend path, as can be seen in Figure 1 below. The economic decline and elevated persistent unemployment rates inspired Keynes to work on *The General Theory*, in which he challenged the reigning framework where prices and wages could move and equilibrate all markets and outlined

---

3 In Adolfson *et al.* (2016) we argue that model misspecification rather than problems with identification are central to understanding the widely held belief that Maximum Likelihood estimation of DSGE models is not feasible. In an environment with model uncertainty and misspecification, Bayesian techniques offer a natural way to estimate models by allowing the econometrician to examine the performance of the model in a region of the parameter space that can be deemed a priori plausible.
an alternative framework with slow adjustment in wages and prices with an important role for economic policy in macroeconomic stabilization.\textsuperscript{4}

Moving forward to the 2008 crisis and the current outlook, an interesting aspect of Figure 1 is that it demonstrates that the current deviation from long-term trend is modest, and is expected to remain so (at least for the United States, as projected by the Congressional Budget Office). Hence, it is not evident that a similar transformation of macroeconomics is called for. The current crisis does not stand out in any way like the crisis in the 1930s. From this long-term perspective, it merely implied a return to the assumed linear trend path.\textsuperscript{5}

Even so, I fully recognize that the crisis exposed severe weaknesses in standard DSGE models in use at key policy institutions. For instance, DSGE models did not see the crisis coming. But neither did professional forecasters (who were not running DSGEs), so criticizing DSGE models for not predicting the crisis is simply unfair. More troublesome for DSGE models is that outcomes during the crisis were completely outside the posterior densities as shown in Figure 2 below.

In Figure 2, the left panel shows yearly GDP growth projections (4-quarter change) with the Smets and Wouters (2007) model estimated on data up to the crisis (1965Q1–2007Q4) and the right panel shows the projection for a Bayesian VAR (BVAR) estimated on the same seven time series with a prior on the long-run means of the series (i.e. the steady state) set to be consistent with the priors used in the DSGE model. Evidently, the crisis is well outside the feasibility set spanned by the DSGE model. However, the right panel in the figure shows that the same is in fact true for the BVAR model. Hence, the Great Recession was also a highly unlikely tail event according to

\textsuperscript{4} See Temin and Vines (2016) for an exposition of Keynes’s thinking at the time and a comparison with the economic challenges facing Europe today.

\textsuperscript{5} I would like to thank Marc Giannoni for the inspiration for this figure.
the BVAR model. Given that both the BVAR and the DSGE are linearized models, the relatively high degree of similarity of the two model forecasts is not completely surprising. We also see that the uncertainty bands are roughly equal in size in both the DSGE and in the BVAR model (if anything slightly smaller in the BVAR). This finding is neither obvious nor trivial as the DSGE model does not have a short-lag BVAR representation.

Clearly linearized DSGEs underestimated the severity of the crisis as late as in 2008Q3, and thereby demonstrated some key limitations going forward.6 Even so, in its defence it should be said that US GDP growth outcome during 2008Q4 was also well below uncertainty bands for other professional forecasters (e.g. the Federal Reserve).7 Furthermore, standard reduced form models of inflation determination also had problems in coping with inflation outcomes during the crisis. For instance, take the Rudebusch–Svensson (1999) accelerationist Phillips curve:

$$\pi_{t+1} = \alpha_{\pi_1} \pi_t + \alpha_{\pi_2} \pi_{t-1} + \alpha_{\pi_3} \pi_{t-2} + \alpha_{\pi_4} \pi_{t-3} + \alpha_y y_t + \varepsilon_{t+1},$$

where $\pi_t$ is annualized inflation measured with the GDP deflator, and $y_t$ is the Congressional Budget Office (CBO) measured output gap (the results are little affected if unemployment minus a constant non-accelerating inflation rate of unemployment (NAIRU) is used instead). This equation was originally estimated by Rudebusch and

---

6 Lindé et al. (2016) show that the uncertainty bands of a variant of the DSGE model with financial frictions embedded contains the actual outcome provided that the forecast conditions on the actual interest rate spread in 2008Q4 (difference between Baa-AAA yields). However, even conditional on the interest rate spread, the median projection severely underestimates the actual fall in GDP. Lindé et al. argue that allowing for the time-varying role of financial frictions and shocks mitigates the tension between projected and actual outcomes, but this is not a structural explanation, so further work is needed as suggested by Rajan (2005).

7 For instance, in its October 2008 Greenbook, issued on 22 October, the Federal Reserve staff projected that the GDP growth rate in 2008Q4 would not undershoot –3 per cent even if more financial stress materialized, whereas the actual outcome was below –4 per cent.
Svensson for the sample period 1961Q1–1996Q2 (marked by the vertical solid line); they discovered its remarkable empirical in-sample fit. As shown in Figure 3, which depicts the actual (solid) and fitted (dotted) values of their regression when the sample is extended to include data after the second quarter of 1996, the model as originally estimated holds up extraordinarily well out-of-sample until the beginning of the crisis in 2008. But during the crisis (the dotted vertical line marks the third quarter of 2008), there appears to be a change in the relationship and the model predicts that inflation should have fallen much more than it actually did. So reduced form models also have trouble in accounting for the evolution of key variables during the crisis.

Despite their shortcomings as a forecasting tool during the crisis, DSGE models—or more generally structural macro models with forward-looking expectations—nevertheless proved extremely useful to policy-makers during the crisis. Stylized DSGE models were useful to point out the benefits of fiscal stimulus (Woodford, 2011; Eggertsson, 2011; Christiano et al., 2011) in pro-longed liquidity traps, and Coenen et al. (2012) corroborated their findings in empirical policy models. These predictions were subsequently supported in empirical work by Blanchard and Leigh (2014).

During the European debt crisis, rising debt levels and financial market stress forced many European countries to pursue aggressive austerity to put their fiscal house in order. In this context, DSGE policy models could be used to point out the potential difficulties in pursuing aggressive fiscal consolidation under currency union membership and zero lower bound (ZLB) constraints (see, for example, Corsetti et al. (2013),
and Erceg and Lindé (2010, 2013)), and that a more gradual approach to consolidation could be much less painful in terms of forgone output under some circumstances. My narrative is that these insights eventually turned into actual policy in the euro area.

When it comes to monetary policy, the partly forward-looking determination of consumption and investment decisions in DSGE models—which implies that both current and expected future real interest rates matter for economic outcomes today—was key to understanding the merits of the so-called unconventional monetary policies pursued by the Federal Reserve and other central banks (e.g. Bank of England and the European Central Bank (ECB)) during the global recession and the European debt crisis. The unconventional monetary policy tools the Fed employed to stem the financial crisis in 2008–9 and to strengthen the recovery during 2010–14 mainly consisted of forward guidance about the future path of the federal funds rate and large-scale asset purchases of private and public longer-term securities (see, for example, Bernanke, 2013). Forward guidance, or expanded guidance to keep future policy rates lower for longer, boosts near-term economic activity and supports the inflation outlook by putting downward pressure on long-term real yields. Because short-term interest rates in the United States following the crisis were expected to remain close to their effective lower bound for quite some time, even without guidance about future rates, the Fed decided to supplement its interest rate policies and forward guidance with large-scale asset purchases, often referred to as LSAPs (or QE—quantitative easing). LSAPs were open market purchases of longer-term US Treasury notes and mortgage-backed securities, with the objective to reduce term premiums and long-term yields and thereby provide further stimulus to investment and consumption. Woodford (2012) discusses in detail the theoretical transmission channels (including balance sheet effects) of forward guidance and LSAPs in the context of DSGE style models. There is also a growing body of empirical literature suggesting that these policies had intended effects on financial markets and the real economy, see, for example, Swanson (2017) and Rogers et al. (2016) for effects on financial markets, and Weale and Wieladek (2016) and Wu and Xia (2016) for effects on aggregate quantities and prices.

III. Where do we go from here?

A key question is: what is required from macro-models to be useful going forward? In this section, I discuss it from a policy perspective, i.e. what is required of a core model to be useful in a policy environment to address most positive (i.e. non-normative) questions. In the first subsection, I present the criteria for a core policy model. In the next subsection, I apply these criteria to some well-known model classes, including DSGEs. Last, I discuss challenges in the development of core and satellite models which meet these criteria.

---

8 See also De Graeve and Lindé (2015) for a formal exposition in a simple New Keynesian DSGE model.
(i) **Criteria for a useful core policy model**

Below, I list my criteria for a useful ‘core’ policy model. By ‘core’ model, I mean a structural model that the institution (e.g. central bank or treasury) uses as its main policy model in the forecasting process. This use normally involves computing the baseline forecast (subject to adding judgements; see, for example, Svensson (2005) for a formal treatment) and alternative scenarios around that baseline (e.g. stronger productivity growth or higher oil prices than in the baseline projection).

**Five criteria for a core policy model:**

1. alternative scenarios and counterfactuals in line with the institutional view;
2. credible communication about future policies matter for outcomes today in a way that is in line with the institutional view;
3. historical decompositions in line with the institutional view;
4. forecasting performance not too subpar relative to other advanced time series tools (important to build credibility with senior staff and policy-makers);
5. simple and transparent enough so that knowledge about the essentials of model properties can be well understood by principals (i.e. senior staff and policy-makers) and retention is feasible when model builders leave or move on to other assignments.

With this in mind, let me now briefly discuss what I mean by each of these criteria in some more detail. The first criterion means that the effects of various shocks need to be in line with the views endorsed by leading policy-makers in the institution. As an example, when it comes to tighter (looser) monetary policy compared to the baseline projection, it is natural that output and inflation should be below (above) their baseline paths in such an alternative scenario. But not only the signs matter, the magnitudes and timing of the deviations from the baseline path matter. Important here is that the core policy model must behave in line with the institutional view (if it deviates from the prior view, then deviations should be convincing enough to change that view) for a range of shocks that are often considered key in understanding the economic outlook (for example, fiscal stimulus and consolidation through spending and/or various tax rates, transient or permanent productivity changes, changes in foreign demand, or higher (lower) oil prices). When the Federal Reserve Board developed and considered replacing its existing open economy model, FRB Global (see Brayton et al., 1997), with an open economy New Keynesian DSGE model, a key step was to relate and explain any deviations of the effects of a broad set of shocks in the DSGE from the existing model (see Erceg et al., 2006). If the effects of important shocks in the proposed model do not comply with the priors of the institution, and the deviations cannot be convincingly explained so that senior staff and policy-makers change their positions, the usefulness of the model will be seriously impaired. The second criterion is also crucial—in many policy institutions (e.g. central banks), there is a firmly held belief that credible announcements of future policies matter for economic outcomes today.⁹ Importantly, this criterion rules out backward-looking models by definition, as in those models only

---

⁹ In the context of monetary policy, Woodford (2012) argues forward guidance has important effects on the economy. In the context of fiscal policy, ‘fiscal foresight’ (see, for example, Leeper et al., 2009) implies that announced credible changes to fiscal policy have important real effects via forward-looking expectations.
actual changes in policy instruments impact the economy. The ability to use models for counterfactual simulations of alternative scenarios—for fundamental shocks or alternative policies for a given state—is singled out by Coenen et al. (2017) as the key reason why DSGEs have retained their status as a leading policy tool among central banks.

While satisfying the first two criteria is necessary and allows the model to be used for alternative scenarios, it is not sufficient for full integration into the forecasting and policy analysis process. For a model to establish credibility as a forecasting tool and have an impact on the baseline projection, my experience is that it also has to satisfy two additional criteria. The first of these—listed as the third above—is that the model must imply a historical decomposition, i.e. the role various shocks have played to explain fluctuations in key variables historically and in the present, which adheres well with the beliefs of the institution. If the economic forces at work filtered out by the model do not sing to the staff involved in the forecasting process, they will doubt the role of the model as a forecasting tool. Moreover, if the historical forecasting track record of the model is in line with other competing tools and the forecasts look sensible, it will add to the credibility of the model. Adolfson et al. (2007) showed that the DSGE model developed and introduced as the core policy model at the Riksbank in 2006—named Ramses—had sensible historical decompositions and forecasting performance in line with official Riksbank projections.10 This was necessary to build credibility for the model and fully integrate it into the forecasting process.11 Important here is that forecasting performance is not only in terms of Root Mean Square Errors—an integral part is the story behind the forecast contour. However, as this model lacked financial frictions and an explicit bank sector, it was subsequently replaced (during the global financial crisis) with a model augmented with a more elaborate financial sector which allowed for both endogenous and exogenous movements in credit spreads (see Adolfson et al., 2013). Since national accounts and other data fed to the policy models are only available with a lag (and sometimes measured with significant errors in real-time), the forecasting performance of core models is often improved by conditioning on extraneous information from high-frequency models (i.e. nowcast tools) and the recent international outlook (often crucial for small open economies, and often taken from, for example, the IMF or a suitable set of recent credible global forecasts); see Benes et al. (2010) for a discussion. Finally, note that these last two criteria imply that the model has to be estimated and its empirical performance thoroughly evaluated in several dimensions (preferably both validated as a system, as discussed above, and also equation by equation as suggested by Blanchard (2017a)).

10 Juillard et al. (2007) and Edge et al. (2010) corroborate the findings in Adolfson et al. (2007) on US data during the pre-crisis period by finding that a medium-scale DSGE performs favourably relative to other state-of-the-art forecasting models during the Great Moderation period. Edge and Gürkaynak (2010), however, argue that the model’s absolute forecasting performance is still unsatisfactory.

11 Obviously, the Riksbank never took the model projections as its official projections. All models rely on simplifications and are inevitably incomplete as descriptions of the real world. It follows that the forecasts were not mechanically taken from a single model, but affected by other models and judgements. Nevertheless, Ramses influenced the thinking of the staff and had some impact on the official projections, especially for the revisions of the forecasts. See Iversen et al. (2016) and Lindé and Reslow (2017) for further details on the role of model projections and judgements in the Riksbank policy process.
Fifth and last, for the core model to be successful in a policy institution, modellers have to be able to explain the key mechanisms in the model clearly for senior non-modeller staff and policy-makers. For smaller institutions with limited resources to spend on model development and maintenance, it is also crucial that the model is sufficiently simple and transparent so that transfer of knowledge of the model essentials from model developers to users is feasible. For this reason, it may sometimes be desirable to make less theoretically elegant short-cuts in model design, provided that it allows the model to explain certain aspects of the data almost equally well in a more simple and transparent way.

Notice that these criteria mean that the model often reflects the priors of the policy-making institution on how the economy functions. The views of the functioning of the economy at policy institutions are formed over time and to change them is often a long process. Thus, one important reason why many policy models did not feature an elaborate financial sector and the possibility of large adverse shocks stemming from deteriorating financial market conditions prior to the recent financial crisis was the widely held belief that such scenarios were highly unlikely. Many policy-makers, including the subsequent Chairman of the Fed, Ben Bernanke, believed in the Great Moderation (see, for example, Bernanke, 2004) and macroeconomic performance from the mid-1980s to 2006 also suggested that an abrupt adverse economic scenario was highly unlikely. This view was little informed by DSGE models. Thus, I think Romer’s (2016) and Stiglitz’s (2018) argument that DSGE models misled policy-makers into believing that financial frictions were not essential to understand business cycles (i.e. trigger a large abrupt contraction in output) should be nuanced somewhat. Prior to the financial crisis of 2008, many models at policy institutions did not feature a prominent role for financial frictions (see, for example, Coenen et al., 2012) because there was a belief both among model builders and many policy-makers that such frictions were not critical in the various simulations policy models were used for. The specification of policy models is often the outcome of an active interplay between demand (policy-makers) and supply (staff model builders). Still, in institutions where policy-makers have less well-informed insights into the underlying model assumptions and their implications, model builders carry the brunt of the responsibility for model design.

(ii) Which macro models satisfy the core criteria?

When I think about the current set of macro models which satisfy all these criteria, I think medium-scale DSGEs like the Smets and Wouters (2007) model augmented with a financial sector and a complete GDP identity dominate. At central banks, a flexible price–wage real business-cycle (RBC) model does not pass the test for a core model because it does not feature any material effects of monetary policy on aggregate quantities. To qualify to serve as a core policy model at central banks, the model must have a conventional monetary transmission mechanism, i.e. hump-shaped and gradual effects on output and price inflation following a monetary policy shock (cf. the first criterion in the previous subsection). The strength of the Christiano et al. (2005) and

---

12 For open economies, it thus has to be an open economy model with potentially significant foreign influences.
Smets and Wouters (2003, 2007) papers was to show that essentially two perturbations of the basic RBC model—gradual adjustment in nominal wages and prices, as well as real rigidities in the form of habit formation and investment adjustment costs—were sufficient in order to develop a model which could possibly qualify for policy analysis.

Other models satisfy some of these criteria, but not all of them jointly. For instance, structural backward-looking models and VARs easily satisfy the fourth criterion, and with some ingenious identifying assumptions also the first and third criteria. However, due to their backward-looking nature, they do not satisfy the second criterion, which will severely impair their ability to function as a workhorse model. State-of-the-art backward-looking time series models are great tools in the model arsenal, but cannot be considered core material. Likewise, traditional large-scale macroeconometric models in the Klein (1947) tradition cannot be considered core material. At least not at central banks and treasuries, where expectations about future policy actions should be able to have material effects on outcomes today.

Another interesting class of model is semi-structural forward-looking models like the Federal Reserve Board’s FRB/US (see, for example, Brayton and Tinsely (1996) and Brayton et al. (2014)) and the IMF’s Global Projection Model (GPM; see, for example, Carabenciov et al., 2013). These models essentially have neoclassical properties in the long run, while the behavioural equations are formulated to allow considerable flexibility in accounting for the short-run properties of the data, including allowing expectations to be formed by either adaptive or rational expectations. They may, for instance, impose backward-looking expectations or rational inattention among households, but rational expectations in financial markets. With a possible question mark about the forecasting performance of these models without an extensive use of well-informed judgement (add factors in the behavioural equations), these models clearly also fulfil the five criteria I set out above. However, building and maintaining these large-scale models often require lots of resources and human capital, and for smaller policy institutions with limited resources to spend on model development, DSGEs are competitive given their stronger microeconomic foundations and the ability to get economists trained in academia to work on them (including help from prominent academics on various issues). Therefore, it is not surprising that core policy models used by smaller central banks are often DSGEs. Larger institutions like the Federal Reserve Board, the ECB and IMF, can afford the luxury of having several structural models in their arsenal, but at least one is often a DSGE (see the survey of policy models by Coenen et al. (2012)).

Blanchard (2017a, b) points out that we need different classes of models for different purposes, and I agree. Specifically, he argues that we need five classes of macroeconomic model: foundational models, DSGE models, policy models, toy models, and forecasting models. However, Blanchard’s reasoning seems to suggest that New

---

13 Agent-based models discussed by Haldane and Turrell (2018) and ‘New Monetarist Models’ with deep microfoundations advocated by Wright (2018) are other possible frameworks, but I am not aware of any papers documenting that these models fulfil the criteria listed. For instance, for monetary policy purposes, I am not aware of any papers documenting that these models feature a realistic monetary transmission mechanism. But subject to further progress (see the discussion in section III(iii)), such models may become contenders as core material.

14 See also Sims (2007) for a discussion of the pros and cons of DSGE and semi-structural macroeconomic models.
Keynesian DSGE models fit the data too poorly to be used as core policy models. So we should mainly have DSGEs for thinking about various theoretical mechanisms and not impose the straightjackets of general equilibrium and microfoundations on core policy models. Here I think one has to be careful with terminology. As I see it, DSGE models like that of Smets and Wouters have to a certain extent relaxed these straightjackets by allowing for a flexible modelling of real and nominal rigidities as well as a rich set of shocks (of which some might sometimes be less straightforward to interpret and relate to the real world, as noted by Romer (2016)). So in a nutshell, I think that empirically validated and estimated DSGE models that are used for monetary policy analysis by, for example, the Riksbank, the ECB (see Coenen et al., 2010), the Federal Reserve Banks of Chicago and New York (see Brave et al. (2012) and Del Negro et al. (2013)) qualify as core policy models in Blanchard’s (2017b) terminology. They can be used for routine scenario analysis purposes and to do counterfactual exercises, and their empirical performance is often good enough to be useful benchmarks in the forecasting process once they are conditioned on the staff’s nowcast (the current quarter for which various indicators often have superior information plus occasionally also one or two quarters ahead, and possibly also the international outlook). See Iversen et al. (2016) for a comparison of stored real-time forecasts of the Riksbank DSGE model and a Bayesian VAR model with official Riksbank forecasts. Even so, given that the statistical performance of DSGEs is still likely subpar relative to the most sophisticated time-series tools, as noted by Pagan (2003) in his trade-off frontier between theoretical and empirical coherence and more recently argued by Hendry and Muellbauer (2018) and Hendry and Mizon (2014), it is useful to complement and compare the forecasts in the DSGE with projections generated by an advanced reduced form model in the forecasting process; for instance, a BVAR model with a comparable set of observables used to estimate the DSGE. This ensures that the forecasts of the core policy model are cross-checked with state-of-art-forecasting tools, the last class of models discussed by Blanchard (2017b).

(iii) Development of new models and a role for a suite of models

Although my view is that DSGEs are a leading contender to serve as a core model, we of course need to improve on these models. The recent crisis has shown that the pre-crisis generation of core macro models did not integrate financial markets sufficiently well (often intentionally as discussed previously), see Lindé et al. (2016) for in-depth analysis and discussion. We also had hints/knew about other key shortcomings prior to the crisis. In closed economy models, some key issues were the consumption Euler equation with its infinite discounting and the failure of the pure expectations hypothesis of interest rates without time-varying risk premia (see, for example, Sarno et al. (2007) and the references therein) with potentially implausible implications for the potency of forward guidance). For open economy models, we had the empirical shortcomings of the uncovered interest rate parity (UIP) condition and difficulties in accounting for the strong influence of foreign shocks through trade channels. In addition, the important implications of heterogeneity in wealth and income for the effectiveness of the policy tools employed to deal with the recent crisis have increased the appetite among policymakers for models with heterogeneous agents (see, for example, Kaplan et al., 2016).
With this in mind, I discuss two issues pertaining to the development of new core and satellite DSGE models: (i) how to do model development, and (ii), how to deal with the fact that not all issues can be addressed with one single model.

When it comes to model development, I fully agree with Blanchard’s (2017a) suggestion that smaller models should be developed first to understand new mechanisms. These mechanisms can then, once properly understood, be integrated into the larger core model as deemed necessary. To facilitate this process we should strive for making core models modular, so that they can be stripped down in a way that highlights key transmission channels and policy trade-offs.

In this process, testing the robustness of findings in small satellite models in more large-scale models similar to the core model remains key. As Blanchard (2016) points out, macroeconomics is about general equilibrium, and it is not clear that a particular result which holds up in a simple model survives in a more realistic framework. Let me give two examples. My first example is based on work by Lemoine and Lindé (2017). In this paper, we study the merits of credible and gradual hikes in the sales tax rate in a protracted liquidity trap (i.e. a situation in which the policy rate is expected to be constrained by its effective lower bound for a considerable period of time, say 2 years or longer). Now, since the consumption Euler equation without habit is given by

$$\frac{C_t}{\sigma} = \beta E_t \left( \frac{1 + i_t}{1 + \pi_{t+1}} \right) \left( 1 + \tau_{c,t} \right) \frac{C_{t+1}}{\sigma},$$

where $\sigma$ is the inverse of intertemporal substitution elasticity and $\beta$ is the discount factor, $i_t$ is the nominal policy rate, $\pi_t$ is inflation, $C_t$ private consumption, and $\tau_{c,t}$ the sales tax. Under lump-sum taxation and Ricardian equivalence, this equation implies that a gradual and credible rise in the sales tax rate is in some sense isomorphic to a cut in the nominal interest rate. In a protracted liquidity trap, the positive effects on the output gap will outweigh the drag on potential output, and overall output may expand as expected inflation rises and the real interest rate gap (i.e. the difference between the actual and potential flex-price real interest rate) falls. So the simple model with fixed aggregate capital suggests that a gradual hike in the sales tax is a potent policy to stimulate output and at the same time increase tax revenues. However, in a more empirically realistic model with some fraction of hand-to-mouth consumers and endogenous capital and financial frictions, the impact of the same path for the sales tax is, in fact, negative unless some other more distortionary tax (e.g. the labour- or capital-income tax) is cut aggressively to balance the budget. Hence, the benign effects of such a policy action in the simple stylized model is not necessarily robust in a more empirically realistic large-scale model. So when building stylized models to study the effectiveness of different policy actions, we are advised to keep an eye on whether the results generalize in the core model environment (for example, with endogenous capital and both price and wage stickiness as well as other real frictions).

My second example is taken from a recent paper with Olivier Blanchard and Christopher Erceg in which the net effect of a policy action depends crucially on the relative strength of two opposing forces in the economy. Specifically, Blanchard et al. (2017) study the extent to which higher government spending in core euro area economies (e.g. Germany and France) boosts economic activity in peripheral euro area economies (e.g. Italy and Spain). Within a currency union context, they identify two
opposing forces. From an aggregate perspective, the effects of a core spending hike are likely to depend on how strongly monetary policy reacts to the induced rise in euro area output and inflation. Outside of a liquidity trap, the ECB would raise interest rates in real terms, which would dampen private demand in both the core and periphery; and unless periphery net exports rose enough to compensate, periphery GDP would likely fall. To gauge the strength of the two channels (positive effects of higher net exports and negative effects of higher real interest rates) and make an assessment of the net effect of a core fiscal expansion on periphery GDP, one would need an estimated empirically relevant model with plausible estimates of relative price adjustment and other real rigidities. Estimation is therefore key, a simple qualitative model is not enough.15

Turning to the second issue, it is of course not possible to address every issue with the core model. But as alluded to earlier, this is not necessary as some issues are not routinely studied, or not easily understood within the context of a larger core model. The core model should only contain mechanisms and shocks that needed to be discussed routinely and are fairly well understood. Other issues which only become relevant occasionally are better analysed with smaller ‘satellite’ models. Satellite models—or specific purpose models—only have to satisfy the first two criteria in section III(i) As an example: suppose that the housing market, which is hard to model in DSGE models without significant alterations in model structures, is omitted in the core model, but that the possibility of a severe correction in the housing market makes it necessary for the staff to brief their principals on the aggregate effects of such a correction and what monetary policy can do to mitigate them. How could the staff, subject to pressing time constraints, do this? They could construct or take an off-the-shelf satellite model with housing and calibrate/estimate it so that it features (a) plausible effects of housing price shocks (in line with VAR evidence, say) and (b) a monetary policy transmission mechanism in line with the core model. This model could then be used to assess what monetary policy can do to deal with the severe adverse correction in the housing market, including a monetary policy response which deviates somewhat from the historical behaviour (estimated monetary policy rule) through ‘Odyssean’ forward guidance and other unconventional policies if deemed necessary. Satellite models are also useful to study other issues that does not come up regularly in the policy process, for instance the implementation of new macroprudential regulation.

To sum up, a formidable challenge for the part of the macroeconomic profession involved in developing general equilibrium models based on microeconomic foundations is how to improve on the current framework while at the same time keeping it tractable (see, for example, the discussion in Lindé et al. (2016)). Therefore—given our limited understanding of how the economy functions—I am also in favour of alternative approaches (for instance, the one advocated by Hendry and Muellbauer (2018) and work on agent-based models under bounded rationality assumptions, as suggested by Haldane and Turrell (2018)) and hope they thrive in academia and at larger policy institutions which can muster the necessary resources.16 Alternative

15 This was an example where aggregate data could be used to shed light on relevant trade-offs; I also think there are many occasions where micro data can be very fruitful to macroeconomic model building, as discussed by Ghironi (2018) in his contribution to this issue.

16 A recent paper by Fagiolo and Roventini (2017) also discusses the merits of agent-based models relative to DSGE models.
model frameworks are important to challenge mainstream thinking, and ultimately the selection between alternative frameworks is a horse race in terms of usefulness and empirical relevance. Furthermore, it is important to keep in mind that different models often complement each other and can serve different roles; they are not always substitutes.

IV. Concluding remarks

I have brought forward a list of criteria which I believe macro models need to satisfy to be useful and influence monetary and fiscal policy decisions. I have argued that forward-looking models based on micro foundations—DSGEs—satisfy these criteria and will therefore likely remain as a key policy tool in the foreseeable future, at least at smaller institutions with limited resources to spend on model development and maintenance. DSGE models are simple and flexible enough to be used for many purposes, while other models are often more limited in terms of the questions they can address. However, other models may be more useful for specific purposes, and it is therefore desirable—resources permitting—for a policy institution to have additional models in its arsenal.

Much effort is currently under way to remedy important flaws in pre-crisis variants of DSGE models in several dimensions. Better integration of financial markets, assessing the impact of the representative agent assumption, relaxing the infinite horizon assumption in consumer behaviour, for instance. The ‘to do list’ for DSGE modellers is long. But importantly, alternative frameworks are also being developed. This is good, in my opinion, as it challenges the mainstream thinking, and it is conceivable that alternative frameworks may one day prove equally—or even more—fruitful in economic policy-making.

References

Lindé, J., and Reslow, A. (2017), ‘It’s a Myth that the Riksbank’s Forecasts have been Governed by Models’, *Sveriges Riksbank Economic Review*, 1, 28–50.


The future of macroeconomics: macro theory and models at the Bank of England

David F. Hendry* and John N. J. Muellbauer**

Abstract: The adoption as policy models by central banks of representative agent New Keynesian dynamic stochastic general equilibrium models has been widely criticised, including for their simplistic micro-foundations. At the Bank of England, the previous generation of policy models is seen in its 1999 medium-term macro model (MTMM). Instead of improving that model to correct its considerable flaws, many shared by other non-DSGE policy models such as the Federal Reserve's FRB/US, it was replaced in 2004 by the DSGE-based BEQM. Though this clearly failed during and after the global financial crisis, it was replaced in 2011 by the DSGE COMPASS, complemented by a ‘suite of models’. We provide a general critique of DSGE models for explaining, forecasting and policy analyses at central banks, and suggest new directions for improving current empirical macroeconomic models based on empirical modelling broadly consistent with better theory, rather than seeking to impose simplistic and unrealistic theory.

Keywords: DSGE, central banks, macroeconomic policy models, finance and the real economy, financial crisis, consumption, credit constraints, household portfolios, asset prices

JEL classification: E17, E21, E44, E51, E52, E58, G01.

I. Introduction

After the global financial crisis, quantitative models used by central banks, as well as prevailing fashions in macroeconomic theory and the training of graduate students in economics, came in for heavy criticisms, some mentioned in Blanchard (2018, this issue). This paper is about policy modelling at central banks rather than the syllabus of first-year graduate courses in macroeconomics, though there are implications for the training of bank economists. The paper begins by reviewing some of the above criticisms, particularly of the New Keynesian dynamic stochastic general equilibrium (DSGE) models used at
central banks. To help appreciate the contrasts between this approach and the policy models that preceded it, section II critically analyses the Bank of England's 1999 policy model, an example of what Wren-Lewis (2018, this issue) calls a ‘structural econometric model’ (SEM). Section III provides a general critique of New Keynesian DSGE models. Sections IV and V explain, and drawing on section III, critically analyse the Bank of England's 2004 and 2011 applications of the New Keynesian DSGE approach. Section VI suggests directions for improvement in policy modelling and section VII concludes. Thus, sections II, IV, and V, analysing specific practice at the Bank of England, illustrate the more general critiques of section III and the proposals for improved modelling in section VI.

In the most fundamental and scathing of post-crisis critiques, Buiter (2009) said:

The Monetary Policy Committee of the Bank of England I was privileged to be a ‘founder’ external member of during the years 1997–2000 contained, like its successor vintages of external and executive members, quite a strong representation of academic economists and other professional economists with serious technical training and backgrounds. This turned out to be a severe handicap when the central bank had to switch gears and change from being an inflation-targeting central bank under conditions of orderly financial markets to a financial stability-oriented central bank under conditions of widespread market illiquidity and funding illiquidity. Indeed, the typical graduate macroeconomics and monetary economics training received at Anglo-American universities during the past 30 years or so, may have set back by decades serious investigations of aggregate economic behaviour and economic policy-relevant understanding. It was a privately and socially costly waste of time and other resources.

Most mainstream macroeconomic theoretical innovations since the 1970s (the New Classical rational expectations revolution1 associated with such names as Robert E. Lucas Jr., Edward Prescott, Thomas Sargent, Robert Barro etc., and the New Keynesian theorizing of Michael Woodford and many others) have turned out to be self-referential, inward-looking distractions at best. Research tended to be motivated by the internal logic, intellectual sunk capital and aesthetic puzzles of established research programmes rather than by a powerful desire to understand how the economy works—let alone how the economy works during times of stress and financial instability. So the economics profession was caught unprepared when the crisis struck.

He went on to criticize the complete markets paradigm of the established research programme: ‘Both the New Classical and New Keynesian complete markets macroeconomic theories not only did not allow questions about insolvency and illiquidity to be answered. They did not allow such questions to be asked.’

Muellbauer (2010) noted that: ‘the recent generation of DSGE models failed to incorporate many of the liquidity and financial accelerator mechanisms revealed in the global financial crisis’. He argued:

Underlying conceptual reasons for the failure of central bank models of the DSGE type include their typical assumptions about representative agents, perfect

---

1 See Wren-Lewis (2018) and Hoover (1994) for an illuminating intellectual history of this revolution.
information, zero transactions costs, and of efficient markets. For most of these models, with the notable exception of Bernanke et al. (1999), and others who also incorporate a financial accelerator for firms, e.g. Christiano et al. (2003), Christiano et al. (2009), it is as if the information economics revolution, for which George Akerlof, Michael Spence and Joe Stiglitz shared the Nobel Prize in 2001, had not occurred. The combination of assumptions, when coupled with the trivialisation of risk and uncertainty in these supposedly stochastic models, and the linearisation techniques used in their solution, render money, credit and asset prices largely irrelevant. The calibration/estimation methods currently used to apply these models to the data typically ignore inconvenient truths.

Caballero (2010) criticized the ‘pretence of knowledge syndrome’:

the current core of macroeconomics—by which I mainly mean the so-called dynamic stochastic general equilibrium approach—has become so mesmerized with its own internal logic that it has begun to confuse the precision it has achieved about its own world with the precision that it has about the real one. This is dangerous for both methodological and policy reasons.

Romer (2016) shares Cabellero’s perspective and is critical of the incredible identifying assumptions and ‘pretence of knowledge’ in both Bayesian estimation, e.g. Smets and Wouters (2007), and the calibration of parameters in DSGE and real business cycle (RBC) models.

Stiglitz (2018, this issue) uses insights from the information economics revolution, to which he made such major contributions, to criticize many aspects of recently fashionable macroeconomics.

As Reis (2018, this issue) notes, academic research since 2010 has moved on; for example, there are DSGE models incorporating financial frictions and banking, and some models that do not linearize around a steady state, so permitting more serious fluctuations and the potential of multiple solutions. The heterogeneous agent research agenda, incorporating liquidity constraints and uninsurable idiosyncratic income risk, of Kaplan et al. (2016) and others, has abandoned the complete markets paradigm. Moreover, microeconomic evidence with macroeconomic implications is being studied more intensively and such research is being published. Lindé (2018, this issue) argues that ‘improved DSGE models—modified to take the lessons of the recent crisis into account—will remain as a workhorse in many policy institutions for a long time to come’.

Blanchard (2018) points to a number of failings of DSGE models and recommends greater openness to more eclectic approaches. He acknowledges the usefulness of theoretical models with stylized simplifications to explore important macroeconomic issues, but questions whether even the newer generation of DSGE models are suitable as central bank policy models. He recommends the co-existence of this kind of research with work on policy models such as the FRB/US model, in which economic theory is used more loosely and the data are allowed to speak more freely than in typical DSGEs. This is an example of what Wren-Lewis terms a ‘structural econometric model’ (SEM). Up to 2003, the Bank of England maintained such a model and the

2 He usefully distinguishes foundational models making a deep theoretical point, DSGE models designed to explore macro implications of a set of distortions, and ‘toy models’ such as RBC or simple IS–LM models.

3 We follow his terminology for now but return to the proper meaning of ‘structural’ in section III.
UK Treasury continues to do so. The Bank of Canada and the Netherlands central bank have implemented new SEMs since 2010 and the European Central Bank (ECB) is actively developing SEMs for five major Eurozone economies in its ECB-MC project. Blanchard also usefully distinguishes policy models or SEMs from forecasting models, judged purely on how well they forecast.

As the above outline of the paper explained, our discussion of practice at the Bank of England illustrates more general themes of criticism and proposals for repair work. We share Wren-Lewis’s regret that the Bank of England and other central banks, with the exception of the US Federal Reserve, abandoned their SEMs before the financial crisis, instead of improving such models by addressing weaknesses such as those discussed in section II below. There are many parallels with developments at other central banks, see the general critique of New Keynesian DSGEs in section III, and connections with the wider macroeconomic debates addressed in the collection of papers in this issue. In our view, the Bank of England quarterly model (BEQM) introduced in 2004, based on the DSGE approach, was a lurch in the wrong direction—see section IV. Despite the failure, not unique to the Bank of England, of BEQM during and after the financial crisis, in 2011 a new DSGE model, the ‘central organising model for projection analysis and scenario simulation’ (COMPASS), was introduced by the Bank. Section V explains how the ‘CAST’ plan (‘COMPASS and suite transition’) has been used for modelling, forecasting, and policy analysis, and provides a critique.

II. The medium-term macro model of 1999 (MTMM)

We begin with the Bank’s 1999 MTMM, updated in 2000, and describe its broad structure, namely IS–LM theory plus a wage-price sector, broadly consistent with New Keynesian theory. As already noted, MTMM was a ‘structural econometric model’, incorporating a good deal of economic logic. The core macro-econometric model consisted of about 20 equations determining key endogenous variables. There were a further 90 or so identities defining relationships between variables, and about 30 exogenous variables whose paths had to be set. GDP was determined in the short term by the components of aggregate demand—private consumption, investment (including inventory investment), government consumption, and net exports. In the longer term, GDP was determined by supply-side factors, which determined potential output. Domestic firms were modelled as producing a single composite good using an aggregate production function of the Cobb–Douglas form. So output was determined in the long run by the evolution of the capital stock, labour supply, and total factor productivity (TFP). These variables were assumed to be unaffected by the price level or the inflation rate. Credit stocks and flows played no role in the model and broad money (M4) was represented by a demand equation that essentially played no further role. The financial sector was not explicitly represented, though there were equations for short and long interest rates.

Price level dynamics and the adjustment of actual output towards potential were broadly determined by the interaction between aggregate demand and supply, augmented by explicit relationships for aspects of wage- and price-setting. Firms set domestic output prices as a cyclically varying mark-up over unit labour costs. Firms were also assumed to set the level of employment, with unemployment defined by subtracting
employment from the labour force. Real wages were determined by bargaining in an
imperfectly competitive labour market. Inflation expectations had an explicit role
in wage determination. Price responses were sluggish, so there was slow adjustment
towards both real and nominal equilibria.

However, there were some serious problems, some of which also arise in FRB/US
and in other SEMs mentioned above, thus raising generic issues of contemporary rel-
ence. Our discussion is therefore of more than historic relevance. The consumption
function had an equilibrium-correction form with the long-run solution given by net
worth, labour income, and a real interest rate. Net worth treated cash and other liquid
assets, stock market and pension wealth minus household debt, and housing wealth as
equally spendable. These are highly implausible restrictions, contradicted by empirical
evidence: see sections III and VI below. Short-run dynamics plausibly allowed for nega-
tive cash flow effects of higher base interest rates on spending, negative effects from a
rise in the unemployment rate—interpreted as a kind of income uncertainty effect—
and positive effects from the current growth rate of real gross housing wealth, though
this is likely to be fairly endogenous. There was no explicit consideration of income
growth expectations, in marked contrast to the Federal Reserve’s 1996 FRB/US model:
see Brayton et al. (1997). Even worse, there was no allowance for the radical liberaliza-
tion of credit conditions after 1979, although, since house prices are correlated with
credit conditions, the growth rate of housing wealth should have picked up some of the
short-term consequences.

Moreover, the house price equation itself was grossly mis-specified, taking no account
of shifts in mortgage market credit conditions, mortgage rates, and housing supply, as
reflected in the stock of housing. The long-run solution depended only on earnings and
the 10-year index-linked bond yield, though the latter was not significant (t=1.2) and
the speed of adjustment was an incredibly slow 3.4 per cent per quarter (t=2.0), so the
long-run solution was poorly determined.

With these defects, MTMM would not have captured well the credit-fuelled house price
boom of the 2000s, nor the related rise in the consumption/income ratio. It would have
largely missed the rising vulnerability of consumption to higher consumer debt relative to
income. Nor would it have handled well the dramatic credit crunch beginning in 2008 and
its large joint impact on consumption together with the fall in house prices. With the omis-
sion of nominal mortgage rate effects from the house price equation and its slow speed
of adjustment, it would not have captured adequately the relatively rapid feed-through of
monetary policy to UK aggregate demand. Net worth did rise in the period before 2008
and fall afterwards, permitting some linkages in the right direction between asset prices
and the real economy, but understating the size and dynamics of the relationship.

The other private-sector demand component, business investment, was also poorly
determined, with t-ratios all below 2, and a weak long-run solution. Housing invest-
ment was assumed to be proportional to business investment, a heroic assumption.
Aggregate business investment is notoriously difficult to model, with expectations of
future profits, liquidity, credit constraints for smaller companies, and risk premia, all
potentially playing a role. Moreover, with lumpy adjustment costs at the micro level,
there may be non-linearities in the response of aggregate investment to drivers: see

One place where expectations were formally embedded is the wage equation.
Staggered wage-setting with some forward-looking elements and considerable nomi-
nal inertia make sense, as does the long-run solution, bringing in indirect employment
taxes, the price level, productivity, unemployment, and slow structural changes (e.g. the fall in unionization) picked up by the Hodrick–Prescott filter. The treatment of expectations allowed both model-consistent and more backward-looking alternatives, and the coefficients were freely estimated and well determined. The diagnostics for the wage equation look reasonable. The import and export price equations were well determined and had backward-looking behaviour, but with plausible exchange rate transmission. However, the equation for the GDP deflator had a weak long-run solution, just tied to unit labour costs. It is remarkable that international prices and the real exchange rate played no role, nor did property prices, which should feed through slowly into rental costs: see Aron and Muellbauer (2013a) for US evidence. One suspects that the rise in UK inflation after the exchange rate depreciation of 2008–9 would not have been well captured by MTMM.

The other weakness in the price sector was the exchange rate equation itself, driven by the arbitrage condition defined by uncovered interest parity (UIP). There is strong empirical evidence against UIP. Evidence tends to suggest that for small deviations from long-run equilibrium in which purchasing power parity plays an important role, the exchange rate is not far from a random walk, but for large deviations, equilibrium correction is important: see Taylor et al. (2001).

As this discussion has illustrated, there was plenty of scope for model improvement within the same broad structure. However, the fashion among macroeconomists and many other central banks had shifted away from this kind of model towards New Keynesian DSGE models.

III. Misplaced origins of DSGEs at central banks

(i) What is ‘structural’?

Concerns about the non-constancy of parameters in models, under changed states of nature, have been a contentious issue in empirical econometrics since Robbins (1932), and discussed (inter alia) also by Frisch (1938) and Keynes (1939). Haavelmo’s (1944) classic article on the probability approach in econometrics is the first systematic definition of ‘structural’. Haavelmo contrasts the potential lack of autonomy of empirical regularities in an econometric model with the laws of thermodynamics, friction, and so forth, which are autonomous or ‘structural’ because they ‘describe the functioning of some parts of the mechanism irrespective of what happens in some other parts’, see Hoover (1994). Simon (1952, 1953) explicitly links the structural (or invariant)/non-structural distinction in terms of the representation of causality. It is the omission of any causal relation related to those remaining, not only those deriving from the expectations of economic agents about the functioning of the system that produces non-invariance.

Wren-Lewis (2018) and Hoover (1994) trace how the New Classical Revolution in macroeconomics gained dominance, and why SEMs were displaced as policy models by DSGEs. The Lucas (1976) critique of then current large econometric policy models was the key. Because the parameters of those models were not ‘structural’, i.e. not policy-invariant, they would necessarily change whenever policy (the rules of the game) was changed. However, there was a subtle shift in the meaning of ‘structural’:
Given that the structure of an econometric model consists of optimal decision rules for economic agents, and that optimal decision rules vary systematically with changes in the structure of series relevant to the decision maker, it follows that any change in policy will systematically alter the structure of econometric models. Lucas (1976, p.41).

He called into question the prevailing large-scale econometric models that lacked microfoundations of optimal decision rules of economic agents.

Lucas’s critique can be challenged on five fronts. First, one can question the use of the word ‘optimal’ in defining ‘structure’. Many economists and other social scientists question the information-processing ability of all individuals to make optimal decisions, see Simon (1972) and Kahneman (2003). There is now greater recognition of widespread cognitive biases and the use of heuristic decision rules, see Tversky and Kahneman (1974), and of sheer ignorance, see the evidence on financial illiteracy, Lusardi (2016).

Second, one can challenge Lucas on the nature of econometric models, seen as the aggregation of individualistic decision rules, denying collective decision-making, market externalities, and game-theoretic interactions between agents or groups of agents.

Third, one can challenge the assumption by Lucas of rational or model-consistent expectations, implying the ability of agents to rapidly learn about changes in the economic environment. As we discuss further below, the rational expectations hypothesis runs into fundamental problems when structural breaks shift the underlying probability distributions of the variables so that the ‘law of iterated expectations’ fails. Indeed, the evidence of the intermittent failures by the most proficient professional forecasters, which are not necessarily due to the Lucas critique, suggests that agents are often not able to extract the ‘structure of series relevant to them’ (see, for example, Ericsson and Irons, 1995). For example, it is doubtful that UK households in 2017 have much idea about the Bank of England reaction function on quantitative easing, while forecasters have, for years, predicted some normalization of UK interest rates and been systematically wrong. Moreover, changes to the parameters of zero-mean variables are difficult to detect, so that changes in policy rules do not necessarily induce the forecast failure claimed by Lucas (see Hendry and Doornik, 1997).

Fourth, one can challenge the claim that any change in policy will systematically alter models. Most obviously, if the relevant exogenous policy variable is included in the model, changes in the variable need not change the structure of the model. Sims (1980, p. 12) agrees and gives an example of a shift in an exogenous tax rate, previously also subject to shifts. By ‘policy change’ Lucas means a shift in a policy reaction function which would need to be included in the model. Under rational or model consistent expectations, this is liable to result in complex model changes. Even more important than shifts in policy reaction functions, are structural shifts in the economy (globalization, financial innovation, financial regulation and deregulation, the collapse of Bretton Woods, the formation of OPEC cartel, the decline in trade union power, new technologies). These are not mostly about changes in reaction functions by forward-looking policy-makers. Most of these changes are effectively exogenous—at least not the product of purposive policy decisions. Paradoxically, this makes it easier to incorporate them in models by conditioning on them, while Lucas’s rational expectations story of policy rule changes is somewhat harder to deal with.
Finally, one can challenge the claim that the Lucas critique has ubiquitous relevance by checking the empirical evidence of its relevance. There are many tests for the critique as discussed by Favero and Hendry (1992). The SEMs of the day were attacked not only because they were potentially not ‘structural’, but because of the accusation that they contained ‘incredible restrictions’, partly because, if agents had rational expectations, individual behavioural equations would reflect complex, system-wide properties of the data. As Wren-Lewis notes, the proposition by Sims (1980) that vector autoregressive regressions (VARs) should be used as a way of modelling economic relationships without imposing theoretical restrictions took hold:

Both these developments allowed a bifurcation in academic work, and a degree of methodological purity in either case. Macroeconomists could focus on building models without worrying about the empirical fit of each equation they used, and applied time series econometricians could allow the data to speak without the use of possibly dubious theoretical restrictions.

Unfortunately, because of the need to counter ‘the curse of dimensionality’—the explosive increase in the numbers of parameters in a VAR as the number of variables and lags increases—in practice VARs imposed new restrictions. These are on the number of variables and lags it is feasible to include, and/or via the application of Bayesian priors and/or exclusion restrictions common in ‘structural’ VARs. Moreover, a-theoretical VARs made it harder than in SEMs to formulate a constructive response to the Lucas critique by including relevant modifications of strategic equations. In this respect, the explicit treatment of income expectations in FRB/US, and that model's flexibility in comparing different formulations of the expectations process was surely preferable to pretending that VARs were somehow immune to the critique. It is also a good example of the fourth of the challenges to the Lucas critique listed above.

Macroeconomists turned to DSGE models claiming to be microfounded in individual optimization and incorporating rational expectations. However, the use of the word ‘structural’ to describe an equation or a subset of a model has been hijacked by followers of Lucas (1976) away from its original meaning of invariant to shifts elsewhere in the system to mean ‘microfounded’. But rationalization in terms of optimizing behaviour does not guarantee invariance, nor is it true that invariance always fails in models where such microfoundations are lacking. Central to the claim to be microfounded was a set of further simplifying assumptions, in particular ‘representative agents’ and linear budget constraints, needed to obtain tractable models with what were seen to be desirable steady-state properties. If economies did not deviate too far from such steady-state paths, one could justify linear or log-linear approximations around such paths with a lack of concern for uncertainty even with locally non-linear relationships.

---

4 Sims (1980, pp. 13–14) appears to sympathize with this view, arguing that if policy shifts were not rare, agents would marginalize with respect to them, so that they would be part of the dynamics in a vector autoregressive system.

5 In the context of US inflation models, Aron and Muellbauer (2013a) find that standard methods of choosing lag lengths systematically discard important longer-run information, and that this is true for all information sets considered.
(ii) Microfoundations built on sand

The first element of our view that New Keynesian DSGE models had the wrong microfoundations concerns the representative agent assumption. As is well known, the conditions for exact aggregation of demand functions so that aggregate behaviour corresponds to that of an average household, are very restrictive: demands are linear functions of income and wealth, with households sharing the same marginal propensities. With optimizing behaviour under linear budget constraints, preferences need to have a very specific form, see Deaton and Muellbauer (1980, ch. 6), though, as Muellbauer (1976) showed, there is a more general notion of a representative household than ‘average’, consistent with slightly more general but still very restrictive preferences. With heterogeneous credit constraints across households, even such restrictions on preferences would be of no help in obtaining exact aggregation. Kirman (1992), Carroll (2000), Hoover (2001), and Stiglitz (2018) are among many criticizing the representative agent assumptions of RBC and New Keynesian macroeconomics.

Instead of representative agent economics, stochastic aggregation theory suggests we can often still make good progress with aggregate data even if behaviour at the micro-level looks different from an aggregate model. An excellent example is Houthakker (1956), who showed that a Leontief production function, with no substitution, and Pareto distribution of the parameters at the micro-level, implied substitution at the macro-level as if it arose from an aggregate Cobb–Douglas technology. The functional form at the micro-level could hardly differ more radically from that at the macro-level. In the literature on lumpy adjustment costs, micro behaviour switches discretely from no adjustment to adjustment when some micro-thresholds are reached. In the aggregate, however, behaviour is smooth, as explained by Bertola and Caballero (1990). A recent applied example of stochastic aggregation comes from models of aggregate mortgage delinquency and foreclosure rates (Aron and Muellbauer, 2016). A key driver is the proportion of mortgages with negative equity: if the distribution of mortgage debt to equity is fairly stable, a shift in the ratio of average debt to average equity shifts non-linearly the fraction of borrowers with negative equity. Since bad loans restrict the ability of banks to extend new credit, negative equity is an important non-linear element in the business cycle feedback loop.

The second element of our critique of the microfoundations of New Keynesian DSGE models is of their adoption of the complete markets paradigm, implicitly denying the asymmetric information revolution of the 1970s, see the quotations from Hoover says of representative agent models: ‘While they appear to use the mathematics of microeconomics, the subjects to which they apply that microeconomics are aggregates that do not belong to any agent. There is no agent who maximizes a utility function that represents the whole economy subject to a budget constraint that takes GDP as its limiting quantity. This is the simulacrum of microeconomics, not the genuine article.’

In contrast with exact linear aggregation, in stochastic aggregation, assumptions on the joint distributions of the data allow aggregate behaviour to be represented by parameters of the distributions, such as means, variances, and covariances.

The distribution of debt/equity is approximated by a logistic function defined on a cubic in the debt/equity ratio. The mean debt to mean equity ratio, with slight trend adjustment, is then used to generate estimates of the UK proportion of borrowers with negative equity, consistent with cross-section snapshots. Negative equity is highly significant in explaining aggregate defaults.
Buiter (2009) and Muellbauer (2010) in section I. In these DSGEs, households discount temporary fluctuations in income to maintain spending in the face of shocks, thus providing a stabilizing anchor to the economy, in turn justifying the rational expectation that shocks will prove temporary.

This old-fashioned textbook view of consumption behaviour was challenged by Deaton (1991), Carroll (1992, 1997, 2001), and Aiyagari (1994). Given uninsurable individual income risk and liquidity constraints, the result of asymmetric information, they show that households engage in buffer-stock behaviour to ameliorate income risk and discount expected future income at higher rates than assumed by the textbook model. Moreover, given heterogeneous income processes, heterogeneous liquidity constraints, and heterogeneous asset ownership, there will be considerable heterogeneity in the discount rates used by different households. On average, discount rates applied to expected incomes will be far higher than those of the textbook model.

This has profound implications, as the important paper by Kaplan et al. (2016) demonstrates. They contrast two general equilibrium models: a representative agent New Keynesian (RANK) model and a heterogeneous agents New Keynesian (HANK) model, and show that the monetary policy channel works quite differently in the latter. An important feature of their model, which is shared with their earlier papers on fiscal policy in the context of wealthy ‘hand-to-mouth’ consumers (Kaplan et al., 2014; Kaplan and Violante, 2014), is that consumers own not only buffer stocks in the form of liquid assets but also illiquid assets, typically earning higher long-run returns. However, there are lumpy transactions costs in trading in and out of such assets and households face borrowing limits.

To keep the HANK model tractable, Kaplan et al. (2016) adopt a highly simplified view of housing. A heterogeneous agent model which incorporates somewhat more realistic features of housing and credit markets with important consumption and monetary transmission implications has been developed by Hedlund et al. (2016). Both papers imply that since heterogeneous households, facing idiosyncratic micro uncertainty and radical macro uncertainty, discount income expectations with much higher weights on near-term expectations, aggregate behaviour cannot be adequately approximated by RANK-style models.

(iii) The omission of shifting credit constraints, household balance sheets, and asset prices

The asymmetric information revolution of the 1970s provided microfoundations for the application of credit constraints by the banking system; see Stiglitz (2018) for further discussion. In many countries, shifts in these constraints were among the most important structural changes in the economy; see the example of the US discussed below. Thus, a third criticism of New Keynesian DSGE models, linking closely with the previous, is the omission of debt and household balance sheets, including housing, crucial for understanding, together with shifts in credit availability, consumption, and macroeconomic fluctuations. The US Federal Reserve did not abandon its large non-DSGE econometric policy model FRB/US, but it too was defective in that it also relied on the representative agent permanent income hypothesis, which ignored shifts in credit constraints and mistakenly lumped all elements of household balance sheets,
debt, liquid assets, illiquid financial assets (including pension assets), and housing wealth into a single net worth measure of wealth. Because housing is a consumption good as well as an asset, consumption responds differently to a rise in housing wealth than to an increase in financial wealth; see Aron et al. (2012). Second, different assets have different degrees of ‘spendability’. It is indisputable that cash is more spendable than pension or stock market wealth, the latter subject to asset price uncertainty and access restrictions or trading costs. This suggests estimating separate marginal propensities to spend out of liquid and illiquid financial assets. Third, the marginal effect of debt on spending is unlikely just to be minus that of either illiquid financial or housing wealth. The reason is that debt is not subject to price uncertainty and it has long-term servicing and default risk implications, with typically highly adverse consequences.

There is now strong micro evidence that the effect of housing wealth on consumption, where it exists, is much more of a collateral effect than a wealth effect; see Browning et al. (2013), Mian et al. (2013), Windsor et al. (2015), Mian and Sufi (2016), and Burrows (2017). The economics of this are further discussed in section VI below. As mortgage credit constraints vary over time, this contradicts the time-invariant housing wealth effect embodied in FRB/US.

The importance of debt was highlighted in the debt-deflation theory of the Great Depression of Fisher (1933). Briefly summarized, his story is that when credit availability expands, it raises spending, debt, and asset prices; irrational exuberance raises prices to vulnerable levels, given leverage; negative shocks can then cause falls in asset prices, increased bad debt, a credit crunch, and a rise in unemployment.

Of structural changes, the evolution and revolution of credit market architecture is often the single most important. In the US, credit card ownership and instalment credit spread between the 1960s and the 2000s. The government-sponsored enterprises—Fannie Mae and Freddie Mac—were recast after 1968 to underwrite mortgages. Interest rate ceilings were lifted in the early 1980s. Falling IT costs transformed payment and credit screening systems in the 1980s and 1990s. More revolutionary was the expansion of sub-prime mortgages in the 2000s—driven by the rise of private label securitization backed by credit default obligations (CDOs) and swaps. The 2000 Commodity Futures Modernization Act (CFMA) made derivatives enforceable throughout the US with priority ahead of claims by others, e.g. workers, in bankruptcy. This permitted derivative enhancements for private label mortgage-backed securities (PMBS) so that they could be sold on as highly rated investment grade securities. A second regulatory change was the deregulation of banks and investment banks. In particular, the 2004 Securities and Exchange Commission (SEC) decision to ease capital requirements on investment banks increased leverage to what turned out to be dangerous levels and further boosted PMBS; see Duca et al. (2016). Similar measures to lower required capital on investment

10 In recent years, several empirical contributions have recognized the importance of the mechanisms described by Fisher (1933). Mian and Sufi (2014) have provided extensive microeconomic evidence for the role of credit shifts in the US sub-prime crisis and the constraining effect of high household debt levels. Focusing on macro-data, Turner (2015) analyses the role of debt internationally with more general mechanisms, as well as in explaining the poor recovery from the global financial crisis. Jordà et al. (2016) have drawn attention to the increasing role of real estate collateral in bank lending in most advanced countries and in financial crises. An early exploration by King (1994) of micro-foundations of Fisher’s model seems to have had little subsequent policy impact at the Bank.
grade PMBS increased leverage at commercial banks also. These changes occurred in the political context of pressure to extend credit to poor.

(iv) The missing financial accelerator

In the 1980s and early 1990s, major credit market liberalization had occurred in Norway, Finland, Sweden, and the UK, causing credit, house price, and consumption booms which were followed by busts—precursors of the US sub-prime crisis. In the financial accelerator feedback loops that operated in the US sub-prime crisis, falls in house prices increased bad loans and impaired the ability of banks to extend credit. As a result, household spending and residential investment fell, increasing unemployment and reducing incomes, feeding back further on to lower asset prices and credit supply. Figure 1, due to John Duca (see Duca and Muellbauer, 2013), illustrates the feedback loops in the US sub-prime crisis.

These feedback loops involve non-linearities and amplification. For example, falls in house prices, driving up the incidence of negative equity, cause, via bad loans, a sharper contraction in credit availability than rising house prices cause an expansion of credit availability. Moreover, a contraction in credit availability itself feeds back on to lower house prices. The combination of lower credit availability, which lowers the spendability of housing collateral, even at given house prices, and lower house prices, had a multiplicative effect in lowering consumption in the US sub-prime crisis; see Duca and Muellbauer (2013).

Figure 1: The financial accelerator in the US sub-prime crisis

Such mechanisms were entirely missing in New Keynesian DSGE models, and hardly represented in those DSGE models, such as Bernanke et al. (1999) and Christiano et al. (2003), which incorporated a financial accelerator only for firms. Iacoviello (2005)
and the estimated DSGE model of Iacoviello and Neri (2010) did introduce housing into DSGE. They assume two representative households, patient and impatient, present in a fixed proportion. Patient households apply a loan-to-value constraint when offering mortgage loans to the impatient households, a kind of financial friction. But because of the assumption of infinitely lived or dynastic households, saving for a down-payment, one of the most important saving motives in industrial countries, is omitted. In their closed economy model, without banks and foreclosures, and assuming a frictionless and efficient housing market, transmission and amplification of monetary or other shocks via housing is extremely limited. For example, their model implies that aggregate home equity withdrawal, the excess of households’ mortgage borrowing over acquisitions of housing, is always negative. In practice, US home equity withdrawal was strongly positive for much of the period from 2001 to 2006, and in the peak quarters was of the order of 10 per cent of that quarter’s household income. However, this fact, and the realized foreclosures, were not in the set of salient data chosen by Iacoviello and Neri for their model calibration. Indeed, for their calibrated model, they compare the correlation between consumption growth and house price growth with and without the financial friction. Without the friction, the correlation is 0.099, the result of the common influence of the shocks\textsuperscript{11} on house prices and consumption. With the friction, the correlation rises to 0.123. One would be tempted from this to conclude, but quite wrongly, that financial frictions have little impact on the macroeconomy. This is the opposite of what Figure 1 implies.

(v) Rational expectations

The world is usually in disequilibrium: economies are wide-sense non-stationary from evolution and sudden, often unanticipated, shifts both affecting key variables directly and many more indirectly. Technology, globalization, both in trade and in finance, trade union power, credit conditions, monetary and fiscal policy rules and other legislation, social mores, skills, wars, resource and financial crises, climate, demography, health and longevity, and income and asset distributions all change over time. These, and other innovations keep perturbing the economic system in ways that not even rational individuals can foresee (Uber and Airbnb are two recent examples). DSGEs therefore suffer a double flaw: they are making incorrect assumptions about the behaviour of the agents in their model, and are also deriving false implications therefrom by using mathematics that is invalid when applied to real economies, as we now discuss. For formal proofs, see Hendry and Mizon (2014), which we now summarize.

Structural changes are a key source of forecasting error, as noted above. However, the mathematical basis of DSGEs fails when events suddenly shift the underlying distributions of relevant variables. The ‘law of iterated expectations’ becomes invalid because an essential, but usually unstated, assumption in its derivation is that the distributions involved stay the same over time. Economic analyses with conditional expectations (‘rational expectations’) and inter-temporal derivations then also fail, so DSGEs become unreliable when they are most needed.

\textsuperscript{11} The major shock driving real house prices is a ‘preference’ shock, which Romer (2016) ironically terms a ‘caloric’ shock in contrast to the ‘phlogiston’ of productivity shocks, the major driver of real residential investment in their model.
To explain the law of iterated expectations, consider a very simple example—flipping a coin. The conditional probability of getting a head tomorrow is 50 per cent. The law of iterated expectations says that one’s current expectation of tomorrow’s probability is just tomorrow’s expectation, i.e. 50 per cent. In short, nothing unusual happens when forming expectations of future expectations. The key step in proving the law is forming the joint distribution from the product of the conditional and marginal distributions, and then integrating to deliver the expectation.

The critical point is that none of these distributions is indexed by time. This implicitly requires them to be constant. The law of iterated expectations need not hold when the distributions shift. To return to the simple example, the expectation today of tomorrow’s probability of a head will not be 50 per cent if the coin is changed from a fair coin to a trick coin that has, say, a 60 per cent probability of a head.

To explain more formally, failures of the intertemporal law of iterated expectations arise because, by definition:

$$E_{t}[E_{t} [y_{t+1} | y_t]] = \int \int y_{t+1} f_{t}(y_{t+1} | y_t) dy_{t+1} f_{t} (y_t) dy_t$$

which is not equal to $$E_{t+1} [y_{t+1}]$$ when $$f_{t+1} (\cdot)$$ is not equal to $$f_{t} (\cdot)$$. Here $$f_{t} (\cdot)$$ is the density of the distribution at t and $$E_{t}$$ is the expectation using that distribution.

Conversely, for the law of iterated expectations to hold over time, all of the distributions involved must remain the same, yet shifts in endowments, asset allocations, policy variables, longevity, etc., are occurring intermittently. Worse, the largest shifts, as in a financial crisis, cause the greatest failures of models based on falsely assuming that the law of iterated expectations holds. When sensible economic agents realize that intertemporal planning inevitably forces them to revise after changes, then DSGE models now additionally impose the wrong behaviour for the agents. The key implication is that models based on the law of iterated expectations are intrinsically non-structural, and hence are not a viable basis for economic policy: they are susceptible to the Lucas critique for all policy changes that shift the distributions of either policy instruments or targets, yet most policy interventions change means or variances of relevant variables.

Rational expectations (RE) are asserted to correspond to the conditional expectation of the future, given all available information formed from the actual current distribution of variables, so are in fact irrational when distributional shifts occur. The failure to realize this major problem for RE derives in part from the sloppy mathematical derivations that are all too common, as if forgetting that expectations are integrals over the associated distributions, as explained above.

Building on the well-known theorem that in constant distributions, the correct conditional expectation would be the minimum mean-square error (MMSE) predictor, by not indexing the expectations operator for the distribution over which the integral is being calculated, the MMSE result is supposedly ‘proved’ in many settings where the calculus is completely inapplicable. This includes most DSGE derivations, but especially applies to so-called New-Keynesian Phillips curves (NKPCs) where the expression:

$$y_{t+1} = E[y_{t+1} | I_t] + \nu_{t+1}$$

is written, such that taking conditional expectations of both sides apparently proves that:
and hence that agents can form an unbiased predictor of $y_{t+1}$ at time $t$. Using an undated and unindexed conditional expectations operator makes it seem that the error at $t+1$ is unbiasedly predicted. At time $t$, agents cannot predict the random component $v_{t+1}$ in advance, but that does not establish that the expectation is unbiased for $y_{t+1}$ because the underlying distribution can shift. The apparent proof relies on not writing $E$ with a subscript for the distribution integrated over $f_t(\cdot)$, to clarify that today we expect a zero-mean error. Written correctly:

$$y_{t+1} = E_f[y_{t+1} | I_t] + v_{t+1},$$

and hence that:

$$E_f[v_{t+1} | I_t] = 0.$$

Unfortunately, an $E$ with the subscript $E_f$ is needed to prove that expectations are actually unbiased, which would be a crystal-ball predictor, already knowing the entire distribution of $y_{t+1}$ at time $t$. Castle et al. (2014) show that the NKPC is an artefact of not handling shifts, which are then proxied by substituting the actual future value $y_{t+1}$ for $E[y_{t+1} | I_t]$. The jumps in future values of variables induced by RE are implausible outside of ‘fully efficient’ speculative markets, which probably do not exist, noting that an efficient market implies that future price changes are unpredictable, but the converse does not follow: that future price changes are unpredictable does not entail that a market is efficient—it could be completely erratic.

The NKPC effectively takes the forward second difference of the log price level. High degrees of differencing make data that are non-stationary or subject to shifts look stationary. Omitted variables or shifts in an inflation model push the estimated dynamics towards finding such a differenced specification. Inflation, as seen in the dynamics of an aggregate price index, is almost certainly influenced by relative price adjustment, missing in simplifying assumptions of one good or symmetric goods in New Keynesian models. Evidence of such relative price adjustment and of shifts in inflation dynamics in the US comes from Aron and Muellbauer (2013a).

Another example of the consequences of specification errors comes from Leeper et al. (2017) who use Bayesian methods to estimate a DSGE model broadly of the Smets and Wouters (2007) type for US quarterly data from 1955 to 2014. The utility function incorporates external habits and, with relatively diffuse priors, the data suggest values of the habit parameter close to 1, implying that forward-looking households derive utility from changes rather than levels of consumption and leisure. This is close to second differencing the data for forward-looking households. As the data set excludes not only credit, asset prices, and shifts in credit constraints and other regime shifts, but also oil prices, the estimates are pushed towards second differencing to cope with these omissions. The estimated degree of habit formation is a symptom of mis-specification.

(vi) Misunderstanding forecast failure

A failure to understand forecasting theory also seems partly responsible for the false belief that ‘theory-based models’ are essential. A typical assertion comes from Coletti et al. (1996) at the Bank of Canada:
The inability of relatively unstructured, estimated models to predict well for any length of time outside their estimation period seemed to indicate that small-sample econometric problems were perhaps more fundamental than had been appreciated and that too much attention had been paid to capturing the idiosyncrasies of particular samples.

The econometric theory in Clements and Hendry (1995) allows for the forecasting model to be mis-specified for the data generation process, with parameters estimated from inaccurate observations, on an integrated-cointegrated system, intermittently altering unexpectedly from structural breaks. That theory helps explain the prevalence of forecast failure, can account for the results of forecasting competitions, and explains some of the good performance of so-called ‘consensus’ forecasts. However, it also demonstrates that in a stationary world, forecast failure cannot be shown to be necessarily due to ‘poor econometric methods’, ‘incorrect estimation’, or ‘data-based model selection’, although such mistakes will increase interval forecasts: shifts relative to the prevailing situation are needed for forecast failure. Claiming that forecasts are based on an ‘economic theory based model’, howsoever ‘structured’, will not by itself counter any of the causes of forecast failure and associated policy failure unless it can anticipate and model shifts.

Intercept corrections (ICs) in equilibrium correction models discussed by Clements and Hendry (1996) are a way of getting any forecast ‘back on track’ after a location shift, irrespective of its source, and assuming it is a location shift (so relatively permanent). ICs are not a magic bullet—‘forecast’ failure will still occur—unless they represent crystal-ball information about the shift that is going to occur (e.g. knowing a dock strike is imminent). But they can help avoid systematic failure after shifts, if appropriately implemented—as can other devices. After forecast failure, there are still non-trivial issues about how to do intercept correction, e.g. how to separate temporary from permanent components of shifts. There are parallels here with the difficulty forecasters have had in deciding whether and to what degree the UK’s post-crisis capacity output had permanently declined.

In DSGEs an approximate equivalent to intercept correction would be to add the previous period’s residual to every equation, which would reduce systematic mis-forecasting resulting from a structural break or from the omission of non-stationary variables. Even a poor model that, in Charles Goodhart’s words, ‘excludes everything I am interested in’ (cited in Buiter, 2009) can thus be protected from the worst consequences of forecast failure, supporting Blanchard’s distinction between policy models and forecasting models.

(vii) The lack of flexibility of DSGEs

A number of authors in this issue, including Lindé, Reis, McKibbin and Stoeckel, and Wren-Lewis, suggest that DSGEs are a flexible tool. Indeed, one can adapt a closed economy DSGE model for an open economy or add a terms-of-trade shock to a model (as does the Bank of Canada’s TOTEM). One can add a fixed fraction of myopic households or a risk premium. One can add a housing sector to a model which previously did not include one. But as our discussion above of Iacoviello and Neri (2010) indicates, that particular effort probably subtracted from understanding because of the efficient market, rational expectations, representative agent framework where the consumption Euler equation is the crucial link between the present and the future.
The consumption Euler equation is thus the key mechanism for the operation of model consistent-expectations. This makes it the main straitjacket of the representative agent DSGE approach. In section VI we advocate its replacement by a solved-out ‘credit-augmented’ consumption function, incorporating the discounted present value of future incomes, using an average discount rate far higher than in standard textbook permanent income models. Such a replacement has fundamental implications as explicit expectations mechanisms are then needed for the other behavioural equations, also of the solved-out form. This allows a much more modular approach, as for example in FRB/US, allowing heterogeneity in expectations between households and firms. Within DSGE models resting on an aggregate or sub-aggregate Euler equation, this kind of modularity is hard to achieve. The Bank of England’s BEQM and its successor COMPASS struggled with these issues, as we now explain.

IV. The Bank of England Quarterly Model of 2004 (BEQM)

In his review of modelling for the Bank of England, Pagan (2003) claimed a trade-off between theory-consistency and empirical relevance; see Wren-Lewis (2018) for a picture and discussion. He notes defects in the ability of MTMM to track the falling saving ratio of the 1980s and in the wage-price system to match the inflation data. But he notes that perhaps the biggest problem with MTMM was not being ‘state of the art’, presumably in theory-consistency. Bluntly, MTMM was no longer fashionable as it was not sufficiently in line with the New Keynesian perspective on the ‘science of monetary policy’; see Clarida et al. (1999) for a summary of the contemporary view. The ‘optimal’ point on the trade-off between theory and evidence posited by Pagan is chosen by a tangent to the highest envelope, determined by the preferences of the policy-makers. There are some qualifications and difficulties with Pagan’s concept.

(i) A unique, universally agreed theory (an ‘agreed-upon conception of the way in which the economy is thought to function’ in Pagan’s own words) is essential if the ‘theory consistency’ of models is to be meaningful enough to even draw the axes on the diagram, as otherwise a different trade-off measure is needed for every theory class. But theories conflict. For example, textbook theory, which assumes efficient and close-to-complete markets, well-informed homogeneous agents, little uncertainty, no credit or liquidity constraints, and a stable economy, contrasts with theory that takes account of the asymmetric information revolution of the 1970s. Many economists, including Stiglitz (2018) as well as ourselves, argue that relevant theory should incorporate credit and liquidity constraints, incomplete markets with incomplete insurance and high levels of individual and macro uncertainty, as discussed in sections V and VI. If the data fit better an empirical approximation to the second theory than an empirical approximation to textbook theory, what meaning can be given to ‘theory consistency’?

(ii) Theories evolve, delivering new insights, thereby requiring a redefinition of models being ‘consistent’ each time. Pagan, indeed, acknowledges that the frontier evolves.12

12 He also says: ‘At any point in time, there will be a frontier of “best-practice” models that shows the combinations of empirical and theoretical coherence that are attainable. There is no precise way of determining this frontier but sometimes opinions form about what is on and off the frontier.’
(iii) Even a single fixed theory does not imply a unique model: it matters how theory models are implemented empirically, how unobservables in the theory are measured, and how expectations are treated.

(iv) Empirical coherence is not unique either: alternative data measures exist, revisions occur, estimation uncertainty is rarely accurately calculated, and the selection criteria for empirical models differ.

(v) Most fundamentally, the diagram suggests theory consistency is bound to reduce, rather than improve, empirical coherence and suggests non-evidential preferences could legitimize falsified theory, making macroeconomics in danger of being unscientific, when the ‘microfoundations’ aim was precisely the opposite. In other words, the trade-off precludes achieving both theory consistency and empirical relevance.13

BEQM, see Bank of England (2004) and Harrison et al. (2005), was claimed to be an advance on MTMM. Its core is a DSGE model in which rational expectations play a major role. This generates long-run steady-state solutions for the key endogenous variables, obtained by solving the core model forward. In the long run, all variables in the model settle on paths that grow consistently with each other in a sustainable equilibrium. Short-run dynamics are handled by embedding these long-run solutions in equilibrium-correction equations in which short-run frictions can explain temporary deviations from the long-run solutions. It is essential to the logic of the approach that only stationary variables can represent such frictions, which do not affect the long-run paths.

The household sector is central to the whole model. A Blanchard–Yaari model of inter-temporal choice aggregates over different households to a representative agent form of a life-cycle consumption function driven by permanent labour and transfer income and net financial wealth. The equilibrium-correction form of the consumption function is fairly similar to that in MTMM, with four key differences in the long-run solution. First, the measure of consumption excludes housing. Second, income is replaced by permanent income. Third, net financial wealth replaces net worth including housing. Finally, theory implies a long-term interest rate response of consumption which is calibrated, not estimated as in MTMM.

Given the slightly different definition of consumption, the equation standard errors are not exactly comparable, but 0.01 in BEQM compared to 0.006 in MTMM suggests a considerable worsening of the match of the consumption function to the data. The slower speed of adjustment, 0.12 in BEQM vs 0.17 in MTMM, also suggests an even less well-functioning long-run solution. This is confirmed by the t-ratio of 1.5 on net financial assets in BEQM vs 3.6 on net worth in MTMM. Simulations suggest that the response of consumption to the short-term interest rate, mainly imposed by assumption rather than tested, was a good deal stronger than in BEQM; see Bank of England (2004). Though there is a treatment of expectations consistent with simple textbook theory, the indications are that compared to MTMM, BEQM is even less data-coherent. Also, as in MTMM, BEQM’s assumption that liquid assets minus debt, and pension and stock market wealth are equally spendable and

---

13 The notion of a trade-off between theory and empirical coherence is often hard to sustain in physics, taking Newton’s Laws as an example.
that shifts in credit conditions are irrelevant is empirically untenable. And it seems that BEQM’s assumption of a zero marginal propensity to spend out of housing wealth does even more violence to the data over this period than MTMM’s assumption that the marginal propensity to spend out of housing wealth is the same as for other assets.

Poor though the fit of MTMM’s investment equation was, BEQM’s is even worse. Further, according to the logic of the model, the gaps between key variables and their long-run solutions are supposed to be stationary. However, augmented Dickey–Fuller (ADF) tests show evidence of non-stationarity for consumption, investment, and exports; see Harrison et al. (2005, p.116). These results for the three major components of aggregate demand do not instil confidence in the approach. The worse fit for these equations compared to MTMM is repeated in many other components of BEQM, for example in the wage equation. Simulations in Bank of England (2004) suggest that the response of CPI inflation to a rise in interest rates is much more rapid in BEQM than in MTMM, suggesting a larger role for forward-looking expectations and less nominal inertia in BEQM. The circumstantial evidence from the estimated equations is that the data do not support such a rapid response.

V. The Bank’s COMPASS model and CAST: its structure and a critique

(i) The structure of COMPASS

If BEQM was a lurch further away from data coherence than MTMM, the introduction in 2011 of the Central Organising Model for Projection Analysis and Scenario Simulation, COMPASS, could be regarded as another milestone on this road. However, there was an explicit recognition that the core model was incomplete and that more weight would be put on the ‘suite of models’ for complementary input into the policy-making process. BEQM was felt to be too complicated and cumbersome for practical advice to the Monetary Policy Committee (MPC), so a simpler DSGE model was sought with more informal methods of introducing practical frictions than the core/non-core form of BEQM. It is interesting to note that the worse fit of BEQM compared to MTMM, and the fundamental tests rejecting the stationarity of the deviations between key endogenous variables and their long-run solutions, do not seem to have been an important factor in the search for a new model.14 However, BEQM’s ‘breakdown’ during the financial crisis may have been a consideration.

The basic idea was to supplement the new DSGE model by stationary ‘wedges’ representing frictional deviations from the fundamental model using, in a fairly informal way, some information from the Bank’s ‘suite of models’. It could be said that for true

14 Apart from coherence with theory and data, three other criteria were listed (Burgess et al., 2013, p. 5): ‘Tractability. The central organizing model should be easy to use, easy to understand, reliable and robust. Flexibility. It should be possible to examine easily the implications of alternative economic assumptions (e.g. the implications of different parameter values) on the behaviour of the central organizing model. Comprehensiveness. The central organizing model should provide adequate coverage of the key economic mechanisms.’
believers, this had the advantage of protecting the core model from potentially contradictory empirical evidence. It seems that, unlike in BEQM, no formal tests of the stationarity of the ‘wedges’ have been carried out, perhaps because various de-trending procedures were thought to have removed the issue, or because of restrictions imposed in the ‘suite’. However, the Bank has published model details and forecast evaluation data and encouraged external evaluation and critique, e.g. in the Stockton Review of 2013, and by a further external review in 2016 of the Bank’s forecast infrastructure, the review of CAST (COMPASS and Suite Transition) to which the authors of the present paper contributed. The Bank also established in 2014 its internal Independent Evaluation Office, see Ashley and Paterson (2016), which examined forecast performance in some detail; IEO (2015).

We now describe in a little more detail how COMPASS, in conjunction with the Bank’s suite of models, has been used for modelling, forecasting, and policy analysis. In 2011, Bank staff introduced this new forecasting platform to assist the MPC in producing its quarterly economic projections. The set-up, which is still in use, comprises a DSGE model supported by a suite of other models. The latter serve three purposes: they provide forecast profiles for sectors of the economy that are absent from COMPASS (‘missing channels’); they provide complementary sectoral analysis as a cross-check on the plausibility of COMPASS profiles; and they inform the calibration of COMPASS responses on occasions when the unadulterated profiles seem implausible.

Burgess et al. (2013) explain how the Bank’s forecasting platform operates, drawing on the Chari et al. (2007) business cycle accounting (BCA) framework as a primary justification for integrating ‘off-model’ information into the model. They argue that it is possible to represent a wide variety of models in terms of a very simple ‘prototype’ RBC model that contains stochastic disturbances to the equilibrium conditions (known as ‘wedges’). However, the ‘wide variety of models’ referred to are all within the DSGE class with various added frictions, such as adjustment costs, labour hoarding, search and match labour markets, and even simple forms of financial frictions. In this way of thinking, externally computed ‘shocks’ derived from the suite of models could be interpreted as determining temporary stochastic disturbances to the underlying equilibrium conditions implied by COMPASS.

COMPASS contains 13 key domestic economic variables, including output, consumption, investment, government spending, exports, and imports. Employment is measured by total hours of work, and prices include export and import prices, CPI, nominal wages, the short-term interest rate and the sterling exchange rate. World prices and world output are also modelled, making 15 endogenous variables in all. These are subject to 18 ‘structural’ shocks. An important early aim was to use shock decompositions to help the MPC better understand the nature of the shocks hitting the economy, though this has proved more difficult and less useful than anticipated. One innovation relative to BEQM was to introduce a time-varying risk premium shock. Note that money, credit, house prices, and other asset prices or balance sheets play no role. So the risk premium shock is one of the few ways in which any connection with the global financial crisis could be made.

Any macroeconomic policy model for an open economy needs to embody a production function, and relationships which allocate output between domestic and foreign purchasers, and demand between imports and domestic production. The simple assumptions built into COMPASS do not seem particularly controversial on this score,
though the lack of variable labour and capital utilization, and hence of short-run labour hoarding, make it harder to interpret productivity data within COMPASS. The random walk assumption for total factor productivity, one might have thought, should increase the degree of uncertainty faced by economic agents, including policy-makers. This makes it particularly odd that precautionary, buffer-stock behaviour and the shorter-time horizons induced by uncertainty play no role, as further discussed below.

Apart from the production function, budget constraints, and accounting identities, the behavioural relationships in COMPASS are Euler equations for consumption, and first-order conditions for optimal investment and employment, and pricing equations embodying some stickiness combined with an important role for expectations. The data are detrended and Bayesian estimation is used to obtain structural parameter estimates. Iterative forward solution methods are used to obtain forecasts.

(ii) Forecasting failures

The failures during the global financial crisis of the DSGE-based model BEQM, adopted by the Bank in 2004, have been well documented, as the Daily Telegraph reported in August 2010: http://www.telegraph.co.uk/finance/economics/7935732/Bank-of-England-overhauls-forecast-model-after-errors.html. However, many other central banks, the International Monetary Fund (IMF), and the Organization for Economic Cooperation and Development (OECD) reported similar failures during the global financial crisis.

That article quotes Michael Saunders, now on the MPC: ‘(BEQM) misses the role of credit and financial conditions’; ‘Liquidity strains and money market dysfunctions have no role in BEQM’; ‘The need for a major overhaul is quite large.’ The article reports on the large sums the Bank was spending on the new model, which came into operation in 2011. But, as noted above, COMPASS has a similar DSGE structure to BEQM, though it is simpler and easier to operate. Its 15 endogenous variables also exclude credit and financial conditions.

Fawcett et al. (2015a) examined what the forecast performance of COMPASS would have been, working with real-time data, had it been in existence in 2007: it would also have failed very badly over the financial crisis. Common sense about the economy, using a wider set of information, as embodied in private-sector expectations and by the MPC, forecasted better than COMPASS. In their Bank Underground blog, Fawcett et al. (2015b) write:

forecasts of UK GDP growth from standard DSGE models fare particularly badly during and after financial crises, and possibly in the presence of structural breaks more broadly. An obvious conclusion to draw is that these models might perform better if they contained meaningful financial frictions.

15 They also compare GDP and inflation forecasts for 2001–13 produced by COMPASS with MPC forecasts incorporating judgement as published in the Inflation Report and forecasts from a statistical suite of forecasting models including VARs. There is no outright winner for both GDP and inflation over the entire period and all horizons. COMPASS forecasts are better for inflation beyond a year ahead, but worse for GDP at all horizons and for short-term inflation.
In defence, it is argued that COMPASS was meant to be complemented by ancillary models and by MPC judgement. However, it is questionable whether the way in which information from the suite of models has been combined with COMPASS has adequately compensated for the failings of COMPASS both as a forecasting platform and for understanding the economy. As ex-MPC member David Miles says (private communication):

In my time (early 2009 to autumn 2015) the MPC was pretty well aware of the enormous limitations and weaknesses of COMPASS. I think the ‘suite of models’ was a way of trying to do an emergency repair job on a machine that was not on its own of great use—particularly in 2009 and 2010 when credit conditions were at times dire. I think most of us learned much more from talking to companies up and down the country than from what the models were producing.

Fawcett et al. (2015b) argue:

It should also be noted that forecast accuracy in and of itself is not the only relevant criterion for assessing the utility of a structural model like COMPASS, which also has a role to play, for example, in providing scenario analysis.

However, if the scenario analysis is actively misleading, a model could actually be damaging understanding of the economy and policy capability, as investigated by Castle, Hendry and Martinez (2017).

(iii) Problems with business cycle accounting interpretations

In the business-cycle accounting (BCA) framework underlying COMPASS and the suite, identifying shocks and their dynamic reactions in the model is key to interpretability. We are critical of evaluating models by impulse response functions (IRFs) showing reactions to shocks: these show the dynamic properties of a model, but not the data generating process unless the model is complete and correct. There is also an identification issue as changing the intercept or the error by the same amount in the equation being shocked cannot be distinguished: but unless super exogeneity holds, the effect will not be the same. Pagan (2017) shows the further problem that integrated, in other words, non-stationary, measurement errors are generated by current methods of calculating IRFs for changes in variables.

More fundamentally, our criticisms of the New Keynesian DSGE model in section III invalidate the BCA framework. The information economics revolution implies that the omission of shifts in credit constraints, balance sheets, and asset prices from DSGE models was a capital mistake as these are non-stationary variables and not ‘stationary wedges’. To put it another way, the assumption made in the BCA framework that the economy is never very far from its steady-state trend was simply wrong. These models were unable to capture the possibility of the kind of financial accelerator that operated in the US sub-prime crisis and the resulting financial crisis. They ignored the shock amplifying, propagating, and potentially destabilizing processes in the financial accelerator. Shifts in the credit market architecture are only one example of structural breaks and evolutions which make it hard to sustain the possibility of ‘rational’ or model consistent expectations.
Deep in the justification for why the economy is never far from its steady-state trend is the representative agent, rational (model consistent) expectations, permanent income view of household behaviour shared by COMPASS and BEQM. In that view, as discussed in section III, households discount temporary fluctuations in income to maintain spending in the face of such shocks, thus providing a stabilizing anchor to the economy, in turn justifying the rational expectation that shocks will prove temporary. To be fair, COMPASS assumes that only 70 per cent of households behave in this way and that 30 per cent are rule-of-thumb households who just spend current income. However, the fact that the mass of households follow the permanent income view still plays a major stabilizing role in the model.

While Deaton and Carroll emphasize idiosyncratic income uncertainty at the micro level, there is also often huge uncertainty at the macro level. Nothing makes this more obvious than the large forecasting errors made by the Bank, despite access to the best information in the economy and the most powerful information-processing capability in the UK. This macro uncertainty reinforces the Deaton–Carroll argument that, on average, households discount the future much more heavily (have shorter horizons) than the textbook permanent income hypothesis which underlies COMPASS. This undercuts the assumption that the economy is never far from its steady-state trend and the shock-decompositions of the BCA approach.

As David Miles says:

While I was on the MPC I always found the ‘shock decompositions’ derived from the Bank’s DSGE model rather hard to interpret—being both somewhat arbitrary and difficult to rationalize. The failure to model the credit side of the economy and its time-varying interaction with spending was a weakness, albeit one shared with many macroeconomic models. The Bank’s models also proved of limited value in understanding the biggest issue post financial crisis—namely whether the awful path of UK labour productivity reflects permanent changes or is significantly transitory.

(iv) UK-specific failures to capture the financial crisis

Section III explained the feedback loops operating in the US sub-prime crisis. In the UK, similar feedback loops operated but with important differences. In the absence of a pre-crisis construction boom, the halving of residential investment had a less severe effect on GDP. Importantly, the prevalence of floating rate mortgages meant that the cuts in interest rates fed through quickly to improve the cash flows of many mortgage borrowers and stabilized house prices, limiting losses on the lenders’ mortgage books. As Aron and Muellbauer (2016) show, rates of mortgage repossessions in the UK were of the order of 10 per cent of the US rates, and government policy also helped by improving income support for those in payment difficulties and encouraging forbearance. This meant that part of the feedback loop operating in the UK via the household sector was substantially less severe than in the US. The UK’s severe credit crunch owed more to losses on commercial property and developer loans, fines for financial misdeemeanour, and losses on overseas activities, such as Royal Bank of Scotland’s disastrous takeover of ABN-Amro in 2007, which outweighed losses on the mortgage books of most domestic UK commercial banks. The Bank of England and the government
eventually realized that the credit crunch was severely handicapping a UK recovery, despite quantitative easing (QE), and low rates, so the Funding for Lending Scheme and other measures were introduced.¹⁶

(v) Problems with the Bank’s ‘suite of models’

Unfortunately, the Bank has for some time had somewhat of a blind spot concerning the role of credit and the housing market in the UK economy, despite its historical interest in home equity withdrawal (HEW), e.g. Davey (2001) and Reinold (2011), and the maintenance of quarterly historical data on HEW back to 1970. HEW is about the housing collateral channel by which house prices, together with credit conditions, influence consumption. This channel is entirely absent from COMPASS and only plays a small role in the suite, to be discussed further below. The theory and UK evidence for such a channel was presented in Muellbauer (2008); see further discussion in section VI, and a comparison with the US and Japan provided in Aron et al. (2012).

Balance sheet effects on consumption via the stock market and housing in the UK are clearly important; see US and international evidence surveyed in Cooper and Dynan (2016). The Bank’s suite does now include a three-equation estimated model for consumption, unsecured debt, and household broad money; see Cloyne et al. (2015). This imposes the restriction that broad money, stock market and pension wealth have the same marginal propensity to spend, and minus that on debt, though the short-term dynamics includes separate housing and financial wealth terms.¹⁷ Mortgage debt appears in another part of the system and is driven by the Fernandez-Corugedo and Muellbauer (2006) index of credit conditions, gross housing wealth, and the mortgage rate. In turn, changes in mortgage debt have a short-term effect on housing wealth, determined in the long run by income and the mortgage interest rate. There are therefore some transitory linkages from credit conditions and interest rates via housing wealth to consumption, and there is a direct effect from interest rates in the consumption equation but no time variation in the size of the effect.¹⁸

Balance sheets had previously been modelled in the Financial Stability Review; see Benito et al. (2001). Balance sheets are forecast by cumulating sectoral deficits/surpluses that are generated by the main macro forecast. Butt and Pugh (2014) provide commendable research on credit spreads for households and companies and these spreads apparently now play a role in modelling financial flows. A much more comprehensive non-DSGE model integrating the real side and financial flows and balance sheets with

¹⁶ Judging from speeches by MPC members at the time, Adam Posen deserves particularly ‘credit’ for understanding the sensitivity of the economy to the credit crunch despite its omission from models, and proposing policies to ease it.

¹⁷ Footnote 26 in Cloyne et al. (2015) explains that the housing wealth effect is excluded from the long-run solution for consumption (as in BEQM).

¹⁸ Cloyne et al. (2016) find evidence for a negative effect of changes in mortgage rates on spending using micro-data, first explained in a theory model by Jackman and Sutton (1982). With micro-data, it can be hard to establish the time-varying character of effects, but with a coherent model and aggregate time-series, Muellbauer (2008) and Aron et al. (2012) find that with higher debt relative to income, the cash-flow effect is larger, but partly offset by easier credit access. The implication, missing in Cloyne et al. (2015), is that in the post-2007 credit crunch, with high inherited debt levels, the cash-flow effect would have been particularly powerful.
stock-flow consistency has recently been produced by Burgess et al. (2016). With more emphasis on the financial side and a short sample for 1998–2014, the real economy relationships such as consumption and investment functions take a fairly rudimentary form. This should nevertheless be a useful prelude to greater future integration between the real and the financial sides of a policy model with richer behavioural relationships.

Recently, the Bank has included Geoff Meen’s house price model for the Office for Budget Responsibility (OBR) in the suite, which does embody equilibrium correction and has a mortgage rationing effect for the post-2007 period.\(^{19}\) However, while the main economic forecasts feed into balance sheet forecasts, the latter apparently play relatively little role in the main economic forecasts, though there have been periods when extreme discrepancies between the two have led to shifts in the forecasts of the former.\(^{20}\)

There is a disconnect between COMPASS and the suite on one side and the financial stability responsibilities of the Bank on the other. The term ‘financial stability’ does not occur in Burgess et al. (2013), nor is there any reference to the important insights in Aikman et al. (2009) and the RAMSI model of Burrows et al. (2012). It seems as though the linkage is almost entirely one way, from the ‘real economy’ to finance. The contradictory lesson from the global financial crisis apparently remains to be learned, though the greater openness to fresh thinking under Governor Mark Carney and Chief Economist Andy Haldane augur well.

While it is on the demand side and in the weak links between the real economy and finance where arguably the most severe problems arise (see the views of David Miles quoted above), there have also been issues in modelling pricing behaviour as the failure to forecast the inflationary implications of sterling’s depreciation in the financial crisis makes clear. The prevalent strategy of linearizing a DSGE around a supposed steady state, entails that individual equation specifications are not rigorously investigated, thereby missing potentially important shifts and non-linearities, as with the role of a non-linear response of wages to price inflation leading to wage-price spirals in Castle and Hendry (2014).

\section*{VI. Towards better macroeconomic policy models}

To improve policy models, central banks need research that merges theory-driven and data-driven approaches, rather than treating them as adversaries. In micro-econometrics, great strides have been made in evidence-based research. In macro, the methodological blinkers discussed above have hampered evidence-based progress in developing policy models. Blanchard (2018) writes:

\begin{quote}
For these models, capturing actual dynamics is clearly essential. So is having enough theoretical structure that the model can be used to trace the effects of shocks and policies. But the theoretical structure must by necessity be looser than for DSGE: aggregation and heterogeneity lead to much more complex aggregate
\end{quote}

\(^{19}\) The quarterly speed of adjustment of under 6 per cent suggests either very slow adaption of house prices or some omitted variables.

\(^{20}\) However, according to Burgess et al. (2013), p. 47: ‘The profiles from the balance sheet model are used regularly as an input to MPC discussions, and judgements about financial conditions and the nominal environment are often fed back into the projections in COMPASS.’
dynamics than a tight theoretical model can hope to capture. Old-fashioned policy models started from theory as motivation, and then let the data speak, equation by equation. Some new-fashioned models start from a DSGE structure, and then let the data determine the richer dynamics.

It is possible that, in future, the generation of vast amounts of micro-data from administrative sources rather than surveys subject to selection bias and large measurement errors, may allow quantitative models for the whole economy to be constructed. Ideally, such macro-models would be based on statistically tested models of micro-behaviour, aggregated up from micro-data on millions of households and many thousands of firms. Testing should establish whether such models best assume full information optimizing behaviour at the micro-level or heuristic behaviour rules adopted in agent-based modelling approaches (Haldane and Turrell, 2018, this issue). In the absence of such data, there is an important place for policy-relevant models using aggregate data, general enough to be consistent with plausible micro-behaviour and with plausible assumptions about information and market structure. Such models should be able to encompass insights from multiple stylized models, and use aggregate time series data to learn about the relevance of these insights.

Given the fundamental role of structural breaks in forecast failure, more attention should be paid to disequilibrium states, especially when forecasting, as after a location shift, all equilibrium-correction models naturally converge back to their previous equilibrium means, leading to systematic forecast failure. However, there are methods for robustifying forecasts after location shifts are detected using intercept shifts that help avoid systematic forecast failure (see Clements and Hendry, 1996, and Castle et al., 2015).

These methods also suggest a possible way forward for replacing rational or model-consistent expectations assumed in DSGE models by hypothesizing that agents use such robust methods, at least in the absence of good survey data on expectations. It is sometimes argued in favour of rational expectations that there is something wrong with an account of economic policy in which the authorities can consistently exploit an informational advantage over the public without the public ever catching on, and where what the model predicts is consistently at variance with what it assumes that the public expects. Robust forecasting methods avoid such objections as well as mitigating the failure of the rational expectations hypothesis in the face of structural breaks and radical uncertainty.

The remainder of this section begins by discussing the encompassing approach for evidence-based model selection. It continues by discussing the pros and cons of system estimation, proposing sub-system estimation as the most sensible option. Examples of sub-systems are discussed, with applications to the household sector. A UK illustration is given of how the specification of the consumption function can be improved. It is followed by highlighting some of the insights for economic policy derived from this kind of evidence-based research.

---

21 He is therefore critical of the FRB/US model’s derivation of the lag structure from optimization by representative agents facing polynomial costs of adjustment.
(i) **Encompassing**

Encompassing aims to reconcile the plethora of empirical models that often exist to ‘explain’ some economic phenomenon. If there are several different empirical models, even if each appears viable when evaluated against its own information set, all but one must be either incomplete or incorrect, and all may be.

A natural way to reduce the set of ‘locally admissible’ models is to test whether any one model can account for the results found by all the others, thereby making them inferentially redundant. In a Monte Carlo setting, knowing the data generating process (DGP) would enable a researcher to derive the properties of all models thereof (called population encompassing), and hence account for their simulation outcomes. Encompassing asks whether any given empirical model can mimic this property to account for the results of other empirical models of the same dependent variables. Many of those models will be mutually non-nested, in that they use different sets of explanatory variables, and hence could capture different features of the DGP. Knowing the DGP would explain such outcomes, but it is manifestly more demanding to ask of another empirical model to statistically encompass rival contenders; see Bontemps and Mizon (2008) for an excellent overview.

Consider models M1 and M2 of a variable denoted $y$, with its (unknown) DGP denoted M0. To implement a progressive research strategy, encompassing must satisfy the three properties that define a partial ordering across models:

(i) reflexive: a model explains itself;
(ii) anti-symmetric: if M1 explains M2, then M2 cannot both be different and explain M1;
(iii) transitive: if M1 explains M2 and M2 explains M3, then M1 explains M3.

Creating a ‘super model’ M4, say, that nests all possible contenders would achieve (i)–(iii), but need not provide a useful or interpretable explanation. Thus, the concept of parsimonious encompassing is used in practice, which requires a small model to account for the results of a larger model within which it is nested. This still requires (i)–(iii), but also that M1 (say) explains M4.

Encompassing tests are based on two implications of the encompassing principle. First, if one model encompasses a rival, then it ought to be able to explain the predictions of that rival model. Alternatively, as the encompassed model does not capture all the features of the DGP which are captured by the encompassing model, the first model ought to be able to predict some mis-specifications of the rival model. The latter is a powerful criterion when structural breaks occur, demanding of M1 that it can account for the shifts in M2.

In the context of model selection, encompassing tests of any simplified model against the nesting model provide a key statistical control, revealing if reductions have lost relevant information. This enables tight significance levels to be used for individual selection decisions to retain or omit variables, thereby avoiding ‘overfitting’ without losing valuable information; see Hendry and Doornik (2014).

(ii) **System specification and sub-system estimation**

Developing better models entails using theory more loosely but broadly consistent with microfoundations of the new generation of heterogeneous agent models with...
idiosyncratic uninsurable household income uncertainty and liquidity constraints. (Dis)-equilibrium correction models with interpretable long-run conditional solutions are invaluable. Although system specification is needed, system estimation is not required. Sub-system estimation is better, as system estimation overstates the benefits of imposing restrictions assuming the model is correct versus the costs of using a misspecified model. Instead, we stress the importance of testing, especially for structural change. Indeed, a ‘correct’ theory can be rejected by data evidence simply because of an overly restrictive empirical implementation; see Hendry and Mizon (2011).

System estimation might be adopted (a) because of efficiency gains from ‘full-information’ methods, or (b) to impose cross-equation restrictions on parameters believed to be the same, or (c) for deriving an estimate of a missing or ‘latent’ variable, see below. On the first, Hendry (1976) provides a comprehensive analysis, and finds such gains are small and dependent on the correct specification of the system. On the second, efficiency gains can result when the parameters involved are indeed identical, and when empirical identification is difficult. However, these gains can largely be achieved in estimating a sub-system of closely related equations, for example for the household sector or for the non-financial corporate sector, rather than the system as a whole where the imposition of erroneous restrictions can result in inconsistent estimates.

What is important is system specification, and one of the major flaws of models like BEQM or COMPASS was to omit key sectors, possibly in a misguided attempt to keep the system tractable for joint estimation. Rather, there are major advantages to adopting the approach proposed in Hendry and Doornik (2014) and Hendry and Johansen (2015) for each equation, then checking the resulting system for coherence and completeness. In their method, the theory information is retained unaffected by only selecting over additional orthogonalized candidate variables, longer lags, and nonlinear functions thereof, as well as tackling outliers and location shifts. When the theory model is correct, identical distributions of parameter estimates are obtained as from direct fitting of that model to the data, but when the theory specification is incomplete, incorrect, or changes over time, selecting from the additional features can deliver an improved model; see http://voxeu.org/article/data-mining-more-variables-observations for an explanation and example.

An example of sub-systems estimation applied to the household sector is the LIVES structure—a ‘latent interactive variable equation system’; see Duca and Muellbauer (2013). Where data on hard-to-measure phenomena such as credit conditions are not available, latent variables can be used to link with observable data, e.g. on credit spreads, other measures of loan conditions, and bank lending survey data. State-space methods or spline functions can be used to update the latent variable or variables, responding quickly to evolving data and incorporating new information, so adapting to structural changes. Coherent multi-equation models for major sectors of the economy have the advantage that sensible economic interpretations can be made at the sectoral level. The coherence of an integrated approach means that one is not relying just on one piece of data or one equation for inference. An example of such a model is the six-equation system for the German household sector in Geiger et al. (2016) and a similar model for France in Chauvin and Muellbauer (2017). As noted above, sub-system estimation is preferable to the attempt to estimate all the equations in a large system simultaneously. To learn from data, one needs reasonably general formulations allowing
hypothesis tests for reductions to more parsimonious equations—far more feasible at the sub-system level.

The sub-system LIVES approach focuses on financial innovations/regulatory shifts as key sources of breaks for the household sector equations. The latent variables pick up everything, subject to smoothing, not explained by the variables in the system. The flexible use of dummies or the state-space alternative to represent the latent variables should allow the system to respond quickly to structural breaks as new data arrive. The latent variables resemble intercept corrections in individual equations but with cross-equation restrictions and economic interpretations. This is co-breaking, see Hendry and Massmann (2007), not just in the intercepts but including interactions with selected regressors.

The task of developing LIVES sub-systems is not trivial. Given the cross-equation restrictions in this set-up, automatic model selection methods to reduce general specifications to parsimonious ones, such as those available in the Doornik and Hendry (2013) Autometrics software, cannot (yet) be applied. The German and French six-equation models consist of equations for consumption, liquid assets, unsecured debt, mortgage debt, house prices, and permanent income. There are two main latent variables, credit conditions for unsecured and for mortgage debt. One of the biggest problems concerns the role of demographic variables. These are typically I(2) or I(1) variables so the risk of spurious regression is large and empirical identification relative to the latent variables is not trivial. Long samples, cross-country consistency, judgements on what parameter magnitudes are sensible given ranges of variation in the data, and qualitative conclusions from economic theory and country-specific institutional knowledge are all helpful in arriving at parsimonious models with strong economic interpretations.

For France, the latent variable measuring mortgage credit conditions tracks closely, since 1990, minus the ratio of non-performing bank loans to total bank loans to the private sector. This implies an important and necessarily non-linear linkage between the household sector and the banking system. It also supports the economic interpretation given to this latent variable.

(iii) Improving the UK consumption function: economic and policy insights

In the UK, the down-payment constraint and access to home equity loans has varied over time, with a relaxation in the 1980s, tightening in the early 1990s, followed by a relaxation which ended when the credit crunch began in 2007–8. In turn, the funding for lending scheme, help-to-buy and the recovery in commercial bank balance sheets have led to a more recent relaxation. These shifts have altered the relationship between house prices and consumption. They have also altered the short-term impacts of changes in nominal mortgage rates on aggregate spending. Joint modelling of consumption and

22 I(2) indicates that twice-differencing is needed to convert such a variable to stationarity, while I(1) indicates that once-differencing is enough.

23 In Duca et al. (2016), one of the main drivers of US house prices is a measure of the median loan-to-value ratio for first-time home buyers, a good proxy for credit conditions. The model for the latter includes a non-linear measure of lagged falls in house prices in the previous 2 years.
household balance sheets would help understand these phenomena, debt dynamics, and the implications of debt for consumption, and should be central to the information base for policy formation, within a larger model where the feedback loops can be more fully explored.

A graphical illustration below, based on a 2012 update24 of the UK consumption function explained in Aron et al. (2012), gives an indication of how these phenomena played out for consumption. The long-run ‘credit-augmented’ solution for the log consumption to income ratio generalizes in several respects the textbook permanent income form, which depends on the log ratio of permanent to current income and on the net worth to income ratio. First, it splits net worth into three categories, net liquid assets defined as liquid assets minus debt, illiquid financial assets (stock market and pension wealth), and housing wealth. Second, shifting mortgage credit conditions, measured by a credit-conditions index, has an intercept effect capturing the shifting implications for saving behaviour of variations in the down-payment constraint, and interacts with housing wealth, capturing how access to home equity finance varies with access to credit. Third, the discount rate used to discount expected income growth is higher than the real interest rate to account for income uncertainty and liquidity constraints. Moreover, the coefficient on the log ratio of permanent to current income is freely estimated,25 instead of imposing a coefficient close to one implied by the textbook permanent income model.

To incorporate shifts in credit constraints such as the down-payment constraint for a mortgage, the disaggregation of balance sheets, a role for house prices, income uncertainty, interest rates, and demography, the long-run version26 of the credit-augmented generalized aggregate consumption function is:

\[
\ln \left( \frac{c}{y} \right) = \alpha_0 + \alpha_1 \frac{r}{y} + \alpha_2 \ln \left( \frac{y^p}{y^c} \right) + \gamma_1 \frac{NLA_{t-1}}{y} + \gamma_2 \frac{IFA_{t-1}}{y} + \gamma_3 \ln \left( \frac{hp_{t-1}}{y_{t-1}} \right) + \gamma_4 \frac{HA_{t-1}}{y} + \gamma_5 \text{demog},
\]

Here \( c \) is consumption, \( y \) is income, \( r \) is a real interest rate, \( \theta \) is an indicator of income uncertainty, \( y^p/y^c \) is the ratio of permanent to current income, \( NLA \) is liquid assets minus debt,27 \( IFA \) is illiquid financial assets, \( hp \) is an index of house prices, \( HA \) is gross housing wealth, and \( \text{demog} \) captures the effect of demography on consumption. Key coefficients can be time varying because of shifts in credit conditions.

The intercept \( \alpha_0 \) increases with greater availability of non-housing loans and of mortgages, as the need to save for a down-payment is reduced. The coefficient measuring

24 The update was prepared for a presentation by Muellbauer for a December 2012 Bank of England Monetary Policy Roundtable. Up to 2001, the credit conditions index comes mainly from Fernandez-Corugedo and Muellbauer (2006) with three smooth transition dummies capturing later shifts, estimated just from the consumption equation, rather than from a full LIVES system for consumption and household balance sheets.

25 It also introduces an interaction effect with the credit conditions index, as easier credit access should allow households to be more forward looking.

26 The dynamic version includes partial adjustment, and changes in the unemployment rate, income, and interest rates.

27 It is possible to disaggregate net worth into four main elements, with a separate coefficient on debt. However, relative to a common alternative restriction, the assumption that mortgage debt can just be netted off gross housing wealth, assuming that the coefficient on debt is minus that on liquid assets is better supported by the data.
the sensitivity of down-payment requirements to house prices relative to income, \( \gamma_{3t} \), should become less negative as access to mortgages rises. If access to home equity loans increases, the coefficient, \( \gamma_{4t} \), measuring the marginal propensity to spend out of housing wealth, should increase. One might also anticipate that expectations of future income growth, captured in \( \alpha_{3t} \), would have a larger effect on consumption when credit constraints ease. It is also possible that \( \alpha_{1t} \), the sensitivity of consumption to the real interest rate, might be affected by credit conditions. This is the functional form underlying the following graphical decompositions.

Figure 2 shows a rise from 1980 to 2007 in the log ratio of consumption to income, measured by non-property disposable household income. Some of this is attributable to the rise in the credit conditions index (CCI)—note the jump after 1979 induced by credit market liberalization under Mrs Thatcher. But the interaction effect between CCI and housing wealth/income has even sharper effects and captures much of the time variation. Note that in the 1970s, when credit was heavily rationed, this effect was essentially negligible. In the early 1990s, much of the fall in consumption relative to income is explained by the credit crunch and the decline in house prices, and the combination was even more pronounced after 2008 in the global financial crisis, when the credit crunch was even more severe. However, income growth expectations measured by estimates of log permanent to current income from a forecasting model, also explain some of the variation in log consumption/income. For example, in the first half of the 2000s, high levels of current income relative to permanent income (low log permanent/current income) offset some of the rise in consumption relative to income that would have been induced by the house price boom. And in the global financial crisis, when current income fell, permanent income does provide some stabilization for the consumption/income ratio. The remainder of the long-run effects are shown in Figure 3.

Figure 2: Long-run decomposition of log consumption/income into effects of credit conditions and their interaction with housing wealth/income and into log permanent/current income
The rise in illiquid financial wealth relative to income makes a substantial contribution to the rise in consumption relative to income from 1980 to 2000, and explains some of the decline in the aftermath of the collapse of the dotcom stock market boom in the early 2000s. Crucially important, the other major story told by Figure 3 is the effect of the long-term build-up in debt implied by the decline in liquid assets minus debt, relative to income. This is the pay-back for the credit liberalization and boom in asset prices which boosted consumption but built up debt burdens, and illustrates the vulnerability of the household sector to high debt levels when asset prices fall and access to new credit contracts. The estimated long-run effect of real interest rates is relatively small.

The two figures show only long-run effects. The estimated short-run dynamics also reveal two further important effects. One is a highly significant negative effect of the change in the unemployment rate on consumption, also a feature of the 1999 MTMM consumption function. The second is a negative effect of the change in the nominal borrowing rate, a mix of the mortgage rate and base rate, capturing the asymmetric short-run effect on borrowers compared with savers. This is weighted by the debt/income ratio and also includes an interaction with CCI: for given CCI, a higher debt/income ratio implies a more negative effect of higher interest rates on spending. But

---

28 The estimated model imposes the restriction, empirically supported, that the coefficient on debt equals minus the coefficient on liquid assets. For a range of countries, we find aggregate housing ‘wealth’ effects of zero or negative before mortgage credit liberalization, the marginal propensity to consume for net liquid assets between 0.08 and 0.16, for illiquid financial assets between 0.015 and 0.03, and speeds of adjustment typically 0.4–0.7 for aggregate consumption, somewhat lower for non-durables.
with easier access to credit, refinancing is easier when interest rates rise, so softening the impact. The implication is that in 2008–10, when debt/income reached record levels, but access to new credit was very constrained, the impact of lower interest rates on aggregate UK consumption was particularly strong, evidence for monetary policy effectiveness. The 1999 MTMM consumption function also included a negative effect from the rise in the nominal borrowing rate on consumption, but with no account taken of the time variations in the effect.

Of course, such partial equilibrium decompositions tell only part of the story of policy transmission. For example, higher consumption feeds into higher output, higher asset prices, and lower unemployment, adding to the direct channel of transmission. Integration into a larger model of the economy is necessary to capture the full feedback loops, many of which are missing in COMPASS.

(iv) Wider economic and policy insights

There are close parallels between consumption behaviour in the US and the UK, though with mainly fixed-rate mortgages, the response of consumption and house prices to lower interest rates in the US is necessarily slower than in the UK.\(^{29}\) The above generalization of the textbook permanent income model is an example of the encompassing approach as it encompasses the textbook model as a special case. It is also an example of the looser, and in our view more relevant, application of theory. In contrast to the FRB/US consumption function which incorporates no shifts in credit constraints and aggregates the household balance sheet into a single net worth concept, contradicted by micro evidence, see section III (iii), it no longer corresponds to a representative agent optimizing model. The claimed microfoundations of the FRB/US consumption function do not save it from parameter instability: the estimated speed of adjustment for data up to 2009 of 0.19 falls to 0.10 for recent data.\(^{30}\) This is clear evidence against treating the FRB/US consumption function as a ‘structural’ equation in the classical sense of invariant to shifts in the economic environment.

Because of its omissions, the FRB/US model failed to give proper warning of risks faced by the US economy after 2007. At the Jackson Hole conference in 2007, Mishkin (2008) reported the results of FRB/US simulations of a 20 per cent decline in real house prices spread over 2007–8. The standard version of the model simulated GDP lower than the baseline by 0.25 per cent in early 2009 and consumption lower by only 0.6 per cent in late 2009 and 2010. The simulations suggested a rapid recovery of residential investment given the lowering of the policy rate in response to the slowing economy. FRB/US failed to include a plausible model of house prices and so also missed the feedback from the credit crunch back on to house prices modelled in Duca et al. (2011,

\(^{29}\) Duca and Muellbauer’s (2013) version of the US consumption function uses a latent variable to model the shifting marginal propensity to spend out of housing assets as access to home equity finance has varied, capturing the multiplicative effects of the credit crunch and falling house prices discussed after Figure 1 above.

\(^{30}\) Even at 0.19, Mishkin (2008, pp. 391–2) expresses unease about the slow transmission of wealth effects on consumption.
Consistent with this time series evidence, Favara and Imbs (2015) provide strong micro-evidence for the link between credit supply and house prices in the US. Among the findings of the LIVES models for the household sectors of Germany and France discussed above are that major shifts took place in French credit conditions compared to Germany’s, and these help explain the radically different patterns of house price developments in the two countries. Ignoring post-1980 shifts has catastrophic effects on the French consumption and house price equations and rather less serious ones for Germany. In contrast to the UK, higher house prices relative to income in Germany and France tend to reduce aggregate consumption, other things being equal. The interpretation is that conservative lending practices and high down-payment constraints force many younger households to save for a housing down-payment. Moreover, many renters may be more cautious anticipating that higher rents will follow higher house prices. However, as the French mortgage market has liberalized, so these negative effects of higher house prices relative to income have softened. These findings suggest that monetary transmission in Germany is very different from the US and UK, and somewhat different from France.

The three-equation LIVES model for the household sector in Canada of Muellbauer et al. (2015) reveals striking differences between Canada and its neighbour in linkages between house prices and aggregate consumption. The ATM-like role of home equity in the US is far smaller in Canada, highlighting the importance of institutional differences. In Canada, unlike in many US states, mortgage contracts are ‘full recourse’, meaning that defaulters’ other assets and income flows can be legally drawn upon by mortgage lenders, and pursued for years. There is no tax relief for mortgage interest payments. The banking sector is concentrated and with compulsory mortgage insurance and federal oversight, lending standards are high, with almost no sub-prime lending.

In Japan, too, high down-payment constraints and conservative mortgage lending practices are factors implying a small negative effect from higher house prices on consumption, given income and other asset prices; see Aron et al. (2012). The radically different structure of Japanese household balance sheets compared to those of the US implies that one-size-fits-all ideas derived from the US about monetary policy transmission are wrong, as explained in Muellbauer and Murata (2011). No G7 economy has such a high ratio of liquid assets (mainly bank and post office saving deposits) to income, or such a low rate of stock-market participation (though Germany is not far behind). There is strong empirical evidence that lower real interest rates on deposits reduce aggregate consumption in Japan, given income and other asset prices. This explains why aggregate consumption in Japan has failed to respond to low and now negative policy interest rates or, indeed, to forward guidance on higher future inflation: many pre-and post-retirement households can hardly be enthused by the promise that the real value of their liquid assets will be further eroded in the future with the real income stream remaining negative for longer. Lower policy rates transmit to the real economy in other ways, such as the real exchange rate, the stock market, and higher investment including in residential housing, but overall monetary transmission in Japan is almost certainly weaker than in the US. This suggests that the ‘zero lower bound’ on nominal

31 For South Africa, the omission of credit conditions has similarly catastrophic effects for models of consumption and household debt, Aron and Muellbauer (2013b).
interest rates is not the issue for monetary policy in Japan, paralleling scepticism by Stiglitz (2018) on discussions of the zero lower bound in standard models.

These insights have applications to the Chinese economy. While mortgage markets have developed a great deal—household debt/GDP rose from 34.5 per cent in January 2014 to 44.4 per cent in January 2017—down-payment constraints are still far more stringent than in the US. The easy assumption that higher house prices in themselves will fuel higher aggregate consumption in China is certainly wrong. The easier credit flows that drive up house prices stimulate residential construction which creates higher employment and incomes. And though easier credit to households can temporarily stimulate consumption, given house prices, this is offset by the increased saving for down-payments by many younger households induced by higher house prices relative to incomes. When the credit expansion comes to an end, the hit to residential construction will affect employment and income, while still high house prices and high levels of debt constrain consumption. These are major problems for the hoped-for transition of the Chinese economy away from investment and export-led growth to consumption-led growth.

With countries going through major demographic transitions, empirical evidence on the implications is of great relevance. Empirical evidence from the LIVES models for Germany and France suggest that demographic effects on aggregate consumption conditional on household portfolios disaggregated into the main asset and debt classes are small. However, these portfolios, including the important debt components, themselves appear to be quite sensitive to demographic change, implying a slow but important feed-through of demography to household saving rates. Further research on these lines should illuminate the role of demography in the secular stagnation feared by some economists.

Empirical insights not available from micro-cross-sections or short panels include that, contrary to simple textbook models, there is a major role in Germany and France for nominal mortgage interest rates in driving house prices and the mortgage stock. The role of real rates as embodied in ‘user cost of housing’ increases with leverage. Another is that increased access to unsecured credit reduces demand for liquid assets—ignored by previous research on household demand for money. Finally, evidence for the buffer stock role of both liquid assets and unsecured debt comes from the negative reaction of the former and the positive reaction of the latter to a rise in the unemployment rate.

The careful distinction between the demand for and the supply of credit in LIVES models helps understand the paradox of the frequently found positive correlation between economic growth and credit growth and the negative one with debt levels. The evidence uncovered that, in aggregate, debt has far more negative effects on consumption than stock market or housing wealth have positive effects, has sobering implications for the extended use of monetary policy, including large-scale purchases of government bonds. In the short run, monetary policy, and, of course, provision of liquidity to banks and other institutions under liquidity stress, are important policy levers. But if extended periods of low interest rates and low returns on safe assets drive up debt/income and prices of risky assets to levels beyond what would be sustainable under moderately higher interest rates, the boost to spending will be reversed later.
VII. Conclusions

Blanchard (2018) argues that DSGEs are unlikely to serve as useful macro-econometric policy models and recommends their co-existence with policy models where economic theory is applied more loosely and empirical evidence is used more intensively and more creatively. Wren-Lewis (2018) calls these ‘structural econometric models’ (SEMs), though see our discussion of the term ‘structural’ in section III. He regrets the abandonment, in preference to further testing and model development, by central banks such the Bank of England of their SEMs, exemplified by the Bank of England’s 1999 MTMM, discussed in section II. In section III we provided a general critique of the use at central banks of New Keynesian DSGE models as policy models.

We traced much of the problem back to the Lucas critique. Followers of Lucas effectively hijacked the term ‘structural’ from its earlier usage by Haavelmo (1944), the Cowles Commission, and Simon (1952, 1953) as referring to the invariance of a model or sub-model to shifts in the economic environment to now mean ‘microfounded’. The pincer movement of the Lucas critique and Sims’s (1980) advocacy of VARs supposedly free of the ‘incredible restrictions’ imposed in SEMs led to the intellectual fashion turning against SEMs. However, ‘microfounded’ was based on the illusion that macro data can be treated as the result of optimizing behaviour of representative agents. Moreover, the simplified textbook assumptions of linear budget constraints and close to complete markets contradict the fundamental insights of the asymmetric information revolution of the 1970s.

These imply buffer-stock behaviour of households faced with liquidity constraints and undiversifiable individual uncertainty, contradicting the textbook permanent income hypothesis and the adequacy of an aggregate Euler equation for consumption underlying RBC and New Keynesian DSGE models; see Muellbauer (2010) on the extensive empirical evidence against aggregate consumption Euler equations. The information revolution implies that the omission of shifts in credit constraints, balance sheets, and asset prices from DSGE models was a capital error. As a result, these models were unable to capture the possibility of the kind of financial accelerator that operated in the US sub-prime crisis and the resulting financial crisis. Shifts in the credit market architecture are only one example of structural breaks and evolutions, implying that the notion that the economy follows a stable long-run trend is highly questionable, despite heroic attempts by policymakers to stabilize the path. Uncertainty then becomes radical. Structural breaks also make it hard to sustain the possibility of ‘rational’ or model consistent expectations. The law of iterated expectations necessary to derive DSGE models then breaks down. To illustrate, the uncertainty caused by Brexit emphasises the fundamental role of macro-economic uncertainty, and should the UK exit the EU and the Single Market without a viable trade deal, many distributions will shift radically, so the mathematics of DSGE models will again fail when good policy models are most needed.

The experience at the Bank of England illustrates these more general themes. Although the Bank of England’s 1999 MTMM was mis-specified as explained in section II, BEQM discussed in section IV was even worse, and COMPASS whose structure is the subject of section V, worse still, though the hope was that by supplementing it with a suite of models, it could still be a useful policy guide.

The poor forecasting performance of both BEQM and COMPASS during the financial crisis was discussed. Unfortunately, key omissions in the suite of models contributed to the inability of the suite to properly correct for the omissions in COMPASS.
Such omissions, shared with the Bank's old MTMM, the current Treasury model, FRB/US and models at the Bank of Canada, and (to a degree) at the Netherlands central bank, include mis-specifications of the consumption function. The two most important are the omission of shifts in credit availability and the assumption that all assets and liabilities can be aggregated into a single net worth measure to drive consumption. The latter is patently absurd. It seems incredible that pension wealth is as spendable as cash-in-hand (though recent changes in UK pension legislation may have made pension pots more spendable for some than before). Moreover, inter-temporal theory, even in the absence of credit constraints, implies that housing, a consumption good as well as an asset, has a different relationship with consumption than cash or stock market wealth. When there are credit constraints in the form of down-payment constraints for a mortgage, saving for a down-payment becomes an important component of aggregate saving. And access to home equity loans, far from universal or constant over time, has implications for the relationship between house prices and aggregate consumption, as confirmed by micro evidence discussed in section III (iii).

Section VI discussed improvements in research methodology in the form of the encompassing approach to take into account the implications of different theories to better use evidence in the design of macro policy models. We recommended sub-system estimation in preference to complete systems estimation. The latent interactive variable equation system, LIVES, for a multi-equation approach to jointly model consumption together with main elements of balance sheet and house prices for the household sector is an example of such a sub-system. It enables the extraction of hard-to-observe directly shifting credit conditions as latent variables. An illustration was provided of a better UK consumption function, able, for example, to capture the interaction of the credit crunch and falling house prices in the financial crisis. Some important empirical insights from this kind of evidence-based research to policy debates were discussed. The relevance of this kind of research raises questions about whether the current training of graduate students in economics is adequate for those wishing to work in central banks.

To conclude, a generic issue in debates about policy models concerns their theoretical foundations. A major problem with the claim of ‘theory consistency’ is the question of ‘which theory?’ For example, textbook theory, which assumes efficient and close-to-complete markets, well-informed relatively homogeneous agents, little uncertainty, no credit or liquidity constraints, and a stable economy, contrasts with theory that takes account of the asymmetric information revolution of the 1970s and early 1980s associated with Nobel prize winners Stiglitz, Akerlof, and Spence. Relevant theory must incorporate credit and liquidity constraints, incomplete markets with incomplete insurance and high levels of individual and macro uncertainty. In our view, approximate consistency with relevant theory, as illustrated above, is preferable to closer consistency with highly stylized ‘textbook’ theory.

References


Modelling a complex world: improving macro-models

Warwick J. McKibbin* and Andrew Stoeckel**

Abstract: Macro models have come under criticism for their ability to understand or predict major economic events such as the global financial crisis and its aftermath. Some of that criticism is warranted; but, in our view, much is not. This paper contributes to the debate over the adequacy of benchmark DSGE models by showing how three extensions, which are features that have characterized the global economy since the early 2000s, are necessary to improve our understanding of global shocks and policy insights. The three extensions are to acknowledge and model the entire global economy and the linkage through trade and capital flows; to allow for a wider range of relative price variability by moving to multiple-sector models rather than a single good model; and to allow for changes in risk perceptions which propagate through financial markets and adjustments in the real economy. These extensions add some complexity to large-scale macro-models, but without them policy models can oversimplify things, allowing misinterpretations of shocks and therefore costly policy mistakes to occur. Using over-simplified models to explain a complex world makes it more likely there will be ‘puzzles’. The usefulness of these extensions is demonstrated in two ways: first, by briefly revisiting some historical shocks to show how outcomes can be interpreted that make sense within a more complex DSGE framework; then, by making a contemporary assessment of the implications from the proposed large fiscal stimulus and the bans on immigration by the Trump administration which have both sectoral and macroeconomic implications that interact.

Keywords: macroeconomics, models, risk, relative prices, DSGE

JEL classification: C02, C5, C68, D58, D9, E17, E62, F4

‘Everything should be made as simple as possible, but no simpler.’
(attributed to Albert Einstein)

I. Introduction

The debate over how to make macro-models useful is not new. One contribution to that debate was made by this journal 17 years ago (*Oxford Review of Economic Policy*, 2001).

* Centre for Applied Macroeconomic Analysis, Australian National University, and The Brookings Institution, e-mail: warwick.mckibbin@anu.edu.au

** Centre for Applied Macroeconomic Analysis, Australian National University, e-mail: abstoeckel@gmail.com

We would like to thank David Vines for very helpful guidance and Adrian Pagan for excellent comments. The views expressed in the paper are those of the authors and should not be interpreted as reflecting the views of any of the above collaborators or of the institutions with which the authors are affiliated, including the trustees, officers, or other staff of the ANU or The Brookings Institution.

A useful summary of the long history of model development and new directions for research in macroeconometric models can be found in Hall et al. (2013).

doi:10.1093/oxrep/grx056
© The Authors 2018. Published by Oxford University Press.
For permissions please e-mail: journals.permissions@oup.com
vol. 16, no. 4, 2000). Then, as now, the debate had many dimensions; one paper by McKibbin and Vines (2000) argued that, for real-world policy analysis, large structural models that incorporated both inter-temporal optimization and stickiness were required. Inter-temporal budget constraints and their role in determining asset prices were needed, along with short-term stickiness in wages, adjustment costs in investment, and rule-of-thumb behaviour by consumers and producers. The authors went on to show how several macroeconomic ‘puzzles’ raised at the time by Obstfeld and Rogoff (2000) could be explained once these features were included in a large structural macro-model. The model they used, incorporating these two features that they focused on, was the G-Cubed model.

Inter-temporal optimization and stickiness are now routinely included in the class of DSGE models used today that seek to analyse actual policy issues. But three other features included in the G-Cubed model, namely, global linkages in capital and trade, the role of relative prices in transmitting shocks across countries, and changes in financial risk, tend to be omitted from benchmark DSGE models that deal with real-world policy. That omission turns out to be important. In this paper we show why these three aspects are important to explain some of the macroeconomic puzzles that have arisen during the past 18 years, but particularly with the emergence of large developing economies as key parts of the global economy and with the onset of the global financial crisis and its aftermath.

Our view is that there need not be such a clear distinction between large policy models and the ideas at the foundation of the DSGE models. There is much that the discipline of DSGE models can contribute to improving larger-scale policy models. The G-Cubed model of McKibbin and Wilcoxen (2009, 2013) and the MSG2 model of McKibbin and Sachs (1991) show how to build policy models with DSGE microfoundations but enough nominal and real frictions to better understand recent experience. We are closer to the arguments of Lindé (2018) that the approach of the DSGE framework is useful in many respects. We also agree with Wren-Lewis (2018) that large-scale models are useful, although we find the DSGE framework is a useful guide to improving large-scale models. We disagree with the Hendry and Muellbauer (2018) view that there is little to be gained from the DSGE approach. The G-Cubed model is an example of a model that is consistent with the benchmark DSGE model as set out in section III of the Vines and Wills paper (2018), but also incorporates short-run rigidities, less than fully rational behaviour, and heterogeneous firms in many sectors.

(i) **What are we trying to understand?**

To make better policy decisions implies better understanding of the economic system and consequences of any disruptions that occur. Understanding any economy in the world, including the US, requires an acknowledgement that the world is highly integrated. The fact that a number of US macroeconomic models still assume a closed US economy without exchange rates is a major puzzle and a major problem for most questions of macroeconomic policy. Understanding large complex systems, like the global economy, requires large-scale models. But what features to include, what to leave out, and for what purpose? Blanchard (2017) makes a useful contribution here and distinguishes five types of macro model—each with a different aim in mind. One of those classes of models, which G-Cubed falls into, is policy models. Their purpose, as
Blanchard describes it, is ‘to help design policy, to study the dynamic effects of specific shocks, to allow for the exploration of alternative policies’. These models should fit the data and capture actual dynamics and have ‘enough theoretical structure that the model can be used to trace the effects of shocks and policies’.

Two examples Blanchard gives of such use are: ‘If China slows down, what will be the effect on Latin America?’ and ‘If the Trump administration embarks on a fiscal expansion, what will be the effects on other countries?’ There are plenty more examples one could give. What does Brexit do for the UK, Europe, and the rest of the world? What could happen to economies if rising protection in the US leads to a trade war? How does the emergence of a billion people into the global economy as producers and consumers change the propagation of global shocks? The potential list is a long one. Notice something in all of these examples—the interest is in how a shock in one economy might affect other world economies and what the implications are for policy. That is, not only are we interested in how a shock in one economy affects that economy, but in a highly globalized world, understanding how these shocks spill over into other economies is also crucial. In addition, it is critical to understand how shocks then reverberate back into the originating country from the rest of the world. Shocks are transmitted by trade and financial linkages. We have to be able to understand the basis of trade and capital flows and how various events affect these.

The level, direction, and composition of trade depend on transaction costs (tariffs, subsidies, transport cost, other taxes) and comparative advantage. So relative differences between countries—whether from natural endowments (think energy, agriculture, and mining), from physical and human capital (and how specific and fixed that is), or from relative productivity differences—all matter for trade. It follows that we need a minimum set of countries and sectors to reflect these relative differences. This is particularly important in a world economy that has experienced extremely large changes in relative prices of food versus energy versus manufactured goods versus services over the past 15 years since the large emerging economies of China, India, and Brazil have been transforming the world economy. Nearly every shock of interest that can be imagined will affect sectors differentially, generate changes in relative prices, and cause differences in trade flows that contribute to the overall macroeconomic outcome for an economy. The same is true for financial assets and capital flows—and, of course, these financial aspects are linked to the real sectors.

Capital flows are affected by asset market arbitrage conditions. Important considerations here are: the substitutability between financial assets; satisfying inter-temporal constraints on international debt (besides all the domestic constraints); and the arbitrage conditions linking the rate of exchange-rate change to the difference between home and foreign interest rates and risk premia. It is this role of risk, how it is modelled in G-Cubed, and what insights we gain from including this aspect, along with the role of relative price changes, that is the focus of this paper. We now expand on the role of risk and relative prices.

(ii) What can be important and is largely missing from the usual macro story

Three key features of the world economy since 2000 have been: the emergence of large developing countries such as China and India as well as ASEAN economies and other
emerging regions into global production and consumption chains; large shifts in risk premia in different markets and large swings in the relative prices of manufactured goods, energy, mining, agriculture, and services. These must have had macroeconomic implications, yet they are largely missing in most macroeconomic models of advanced economies.

Figure 1 shows one measure of risk from 2001 to 2009. Clearly large spikes in risk occur during the Dot-Com crisis starting in 2000 and leading up to the financial crisis in 2008.

Figure 2 shows the price of energy, mining, agriculture, durable and non-durable manufacturing, and services from 2000 to 2008, just before the global financial crisis.

This paper summarizes applications of the G-Cubed multi-country model that incorporate these relative prices and risk shocks that are relevant for understanding macroeconomic adjustment and considering contemporary macroeconomic policy.

II. The approach of G-Cubed

(i) Brief summary of the general approach and key features

The features of the benchmark DSGE model described by Vines and Wills (2018) in the introduction to this volume are all incorporated in the G-Cubed model used here.

Figure 1: The Lehman Brothers’ bankruptcy and risk premia

Notes: From McKibbin and Stoeckel (2009). Weekly data. Risk premium on inter-bank borrowing approximated by the rate on 1 month Euro-dollar deposits less the Federal funds rate. Risk premium on corporate bonds measured as the yield on BAA-rated corporate bonds less the 10-year Treasury bond yield.

Data source: Federal Reserve Board.
At the core of the model is a real model of the process of capital accumulation and growth, one that has been taken over from a multiple-sector version of the Solow–Swan growth model, and from the work of real business-cycle (RBC) theorists. Capital accumulation is driven by the decision to invest by the firms that exist in each sector of the model. But unlike in the RBC model, what is saved is not automatically invested. Instead, there is an explicit investment function; the investment decision is forward looking, depending on the expected future need for capital. The extent of investment is governed by the costs of adjustment of capital. The long-run equilibrium Ramsey growth path for this model is one in which the capital accumulation caused by investment exactly keeps pace with population growth and technical progress. In the short run there is a distinction between the optimizing firms and rule-of-thumb firms in each sector of each economy. In the long run both firms follow the same decision rules.

The representative consumer is based on a forward-looking Euler equation but with a probability of death included, as in Yaari (1965), Blanchard (1985), and Weil (1989). This means that the model approximates an overlapping generations (OLG) model and deviates from the standard infinitely lived agent models. Along the equilibrium growth path consumers hold the equity created by investment. Financial intermediation ensures that this happens, although the existence of non-Ricardian equivalence caused by the positive probability of death means that the real interest rate can deviate from the sum of the pure rate of time preference plus the long-run growth of productivity. In the short run, shocks to the level of technology, and to its expected rate of...

**Figure 2**: Commodity, manufacturing, and services prices, US$, 2000=100

change, disturb this growth path; so do shocks to the desire to save, and to the financial
intermediation process. There is an endogenous ‘neutral’ real rate of interest which can
ensure that—despite such shocks—resources eventually become fully employed. The
model can be used to study the effects of technology shocks, of the kind studied by
RBC theorists, but also demand shocks.

The addition of nominal rigidities to this model creates the possibility of an output
gap in which output, driven by changes in aggregate demand, is different from aggre-
gate supply, so that inflation can emerge. This leads to a role for monetary policy, in
the form of a central bank setting the nominal (and short-term real) interest rate; such
policy can pin down the rate of inflation. The Henderson–McKibbin–Taylor rule is one
way of representing policy. Subject to inflation being controlled, such a policy can also
ensure that demand is just sufficient for resources to be fully utilized.

The G-Cubed model importantly joins all countries in the world together through
both trade and asset-market arbitrage.

Fiscal policy can stabilize demand, but over time government deficits lead to public
debts which, to ensure fiscal solvency, require higher levels of taxes to pay the higher
debt interest. Public debt can also crowd out capital because of the Approximate OLG
feature of the model. This was explored at great length in McKibbin (2006) in a version
of the G-Cubed model that incorporated annual cohorts of agents to demonstrate the
impact of global demographic change on the real rate of interest and global growth.

Extensive detail of that model’s features is in McKibbin and Vines (2000) and is not
necessary to repeat here, except for the key features of the model (amended for the ver-
sion of G-Cubed used here) so there is some context in which to interpret the dynamics
and the insights described later. The key features are:

- specification of the demand and supply sides of economies;
- the real side of the model is disaggregated to allow for the production of six types
  of goods and services;
- international trade in goods and services produced by the six sectors, and in finan-
  cial assets;
- integration of real and financial markets of economies with explicit arbitrage link-
  ing real and financial rates of return adjusted for equity risk premia for domes-
  tic assets and country risk premia for international arbitrage through uncovered
  interest parity conditions;
- inter-temporal accounting of stocks and flows of real resources and financial
  assets;
- imposition of inter-temporal budget constraints so that agents and countries can-
  not forever borrow or lend without undertaking the required resource transfers
  necessary to service outstanding liabilities;
- short-run behaviour is a weighted average of neoclassical optimizing behaviour
  based on future income streams and Keynesian current income;
- full short-run and long-run macroeconomic closure with macro dynamics at an annual
  frequency around a long-run Solow/Swan/Ramsey neoclassical growth model;
- the model is solved for a full rational-expectations equilibrium at an annual fre-
  quency over 100 years;
- fiscal and monetary policy adjustments are made in response to shocks to meet
domestic objectives for activity, employment, and target inflation, with rules vary-
ing across countries.
The version of the model used here has six sectors: energy, mining, agriculture, manufacturing durables, manufacturing non-durables, and services. The first three sectors are important for their links to resource endowments that vary widely across the globe. Separating manufacturing into durables and non-durables is important because interest rate changes have differential effects. Also, there is a capital goods producing sector in each economy but it is not traded.

In the current version there are 17 countries/regions as set out in Table 1. More aggregation to make the model smaller is possible, but the grouping below has proved to be optimal to appreciate most shocks of interest from a macro perspective.

Important for policy DSGE models noted by Blanchard earlier is the need to fit the data and capture the actual dynamics. The G-Cubed model is calibrated to the data using a mix of time series econometrics and calibration of parameters following the approach common in computable general equilibrium (CGE) models.

III. Some lessons from the model for understanding history

(i) Asian and global financial crisis

In McKibbin (1998) the crisis in Asia was modelled as a rise in country risk in the crisis economies. That is, foreign investors suddenly needed a high rate of return to hold assets in the crisis economies. The rise in country risk immediately caused a capital outflow and sharp real and nominal exchange rate depreciation and a collapse in private consumption and investment. A sharp recession in the crisis economies resulted.

The analysis was undertaken at the start of the crisis so two possible scenarios were modelled. The first looked at a permanent rise in risk; the second a temporary rise in risk. In the permanent case, the recovery was weak, investment flat, and the capital stock eventually fell to new lower levels. The temporary risk shock led to a deep recession but a ‘V-shaped’ recovery. The actual outcome was a flatter V-shaped recovery with a permanently lower capital stock and an investment rate at half of the pre-crisis levels. This outcome was in between the temporary and the permanent risk shock scenarios, but the model simulations did a good job in stylizing the adjustment to the

<table>
<thead>
<tr>
<th>Table 1: G-Cubed model countries/regions (version 140V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>United States</td>
</tr>
<tr>
<td>Japan</td>
</tr>
<tr>
<td>United Kingdom</td>
</tr>
<tr>
<td>Germany</td>
</tr>
<tr>
<td>Rest of euro area</td>
</tr>
<tr>
<td>Canada</td>
</tr>
<tr>
<td>Australia</td>
</tr>
<tr>
<td>Korea</td>
</tr>
<tr>
<td>Rest of advanced economies</td>
</tr>
</tbody>
</table>

2 See McKibbin and Wilcoxen (2013) for more details.
shocks during and after the crisis. As argued by Corbett and Vines (1999), the results for changes in expected productivity had a similar effect as the risk shocks. The existing capital stocks across the economy were too large and had to be reduced, but this was expensive given real rigidities in changing capital stocks and highly flexible financial flows.

The same methodology was used by McKibbin and Stoeckel (2009) to model the global financial crisis (GFC), although a larger range of shocks was needed because the actual shocks were different to those experienced by crisis economies in the Asian financial crisis. At the time of analysis, which was in the midst of the crisis, it was clear what the shocks were (although the persistence was unknown) and the announced policy responses. The shocks were a bursting of the housing bubble in the US market; a sharp rise in global risk premia in equity markets and in risk assessments by households and firms; and the rise in risk was larger in the crisis economies of the United States and United Kingdom. The policy responses were large changes in interest rates and bailouts of banks by central banks in the US and UK and a global but asymmetric response of fiscal expansions. The actual crisis was a combination of shocks in the US and UK and global changes in risk as well as direct spillovers through trade and capital flows adjustment.

The lessons from that modelling exercise were that the recessions were large, particularly in the US and UK, but the global fiscal stimulus and monetary reactions offset the severity of the risk shocks and the bursting of the housing bubble. The recovery, particularly in housing, would be slow because there was an excess of housing capital that needed to be run down over time.

Another interesting insight was that the sectoral effects of the crisis were very different. The distinction between durable and non-durable manufacturing in each economy was important. The demand for non-durable goods is driven by income, wealth, and relative prices of these goods. The demand for durable goods was largely for building the capital stock in the production sectors and the household's capital stock in the household sector. A large rise in risk in the model meant that the expected value of buying a durable good yielding a return over many years collapsed in present value terms. The rise in risk meant a rise in the rate at which future income would be discounted and therefore there was a collapse in the demand for durable goods. The durable goods sector in the model experienced the largest fall in output. This was true in all countries. It meant that countries that exported durable goods (such as Japan and Germany) experienced a large trade shock as part of the financial crisis, relative to countries with a different composition of trade and a different source of income generation.

One of the side effects was that the model predicted that global trade would contract by more than global GDP and the impact across countries would reflect both the financial risk shocks but also the composition of the production of each economy. When decomposing global production versus global trade it is clear that the share of durable goods in global trade is higher than the share of durable goods in global production. There are some countries which have a comparative advantage in durable goods production for investment purposes. Because of the disaggregation of manufacturing into durable and non-durable goods, the shift in the ratio of global trade to GDP was well captured by the model.

While in both cases the authors did not claim they could predict the movement of risk premia, the modelling showed the important consequences of changes in risk in various markets and the usefulness of the model in evaluating outcomes and the implications of alternative policy responses.
(ii) Growth and relative price shocks from emerging economies

In McKibbin and Cagliarini (2009), the authors explore the impact of growth in the emerging world, particularly China and India, on relative prices in the global economy from 2000 and on the macroeconomic adjustments in the global economy. They considered a sharp rise in productivity growth in China but focusing on the relatively larger rise in productivity in China in the manufacturing sector. The productivity shocks were particularly interesting because they captured the rise in Chinese productivity growth relative to other countries, as well as the rise in manufacturing productivity that reduced the relative price of manufactured goods to other sectors. The second shock was a fall in the risk of investing in emerging economies, particularly in China. The third shock was a relaxation of global monetary policy with a focus on the relative shifts across countries.

One interesting result was that the shocks in overall productivity in China led to changes in the global allocation of capital and trade flows. The shifts in relative productivity growth within China led to a large change in the relative price of manufactured goods relative to energy, agriculture, and services in China. This relative price change also spilled over into a similar relative price change in all economies. The response of central banks was critical for how this propagated in different economies. While a flexible exchange rate could offset the change in the overall price level it could do little to offset any changes in relative prices occurring through trade flows.

A further interesting result was that a change in US monetary policy (of the scale of monetary relaxation that was observed from 2001 to 2004) not only changed the US price level as expected, but it also led to a persistent change in relative prices within the US economy. The capital intensity of production and the elasticity of demand in response to income and price changes in general equilibrium, meant that a fall in interest rates would impact demand and supply across sectors differently. The more capital-intensive sectors would get more investment for a fall in interest rates and the supply response, although eventually returning to baseline, persisted for up to a decade. This monetary policy was shown in the model to change relative prices for a number of years during the adjustment back to a monetary neutral long run. It was interesting that a country like Australia experienced an appreciation of the currency which partly offset the transmission of the US monetary policy to the Australian economy, but the nominal exchange rate change did nothing to offset the relative price changes emanating from changes in US monetary policy. This is not surprising because there were price changes across the six sectors but only one exchange rate which offset the average price change. Much of what is understood about the neutralizing role of flexible exchange rate in offsetting monetary shocks is from one-good models, but the story is much richer and realistic in a multiple-good world.

IV. An application: implications for understanding the impact of Trump’s economic policies

In this section we summarize some results from McKibbin and Stoeckel (2017) which focused on President Trump’s proposed economic policies. We explore the impacts of deporting illegal immigrants and the impact of a plausible fiscal package implemented
by the Trump Administration. This shows both the macroeconomic and sectoral consequences and how they interact. It also answers a question posed by Blanchard (2018) in this issue on the global impact of the Trump Administration’s fiscal policy.

(i) Immigration

A major area of policy reform with significant potential economic impacts is the Trump Administration’s policies on immigration. The simulation explored in McKibbin and Stoeckel (2017) is that by the deportation of illegal workers, the US labour force would be reduced by 1.5 per cent on average, with the largest falls in agriculture and durable manufacturing. The shock was designed to be consistent with the Pew Research study (Krogstad et al., 2017) on the distribution of illegal workers across the US economy.

After the deportation of immigrants, the economy will need to settle down on a new Ramsey growth path, with a lower level of population. Because the labour force is lower, the stock of capital will need to be lower, and so there is a fall in investment. This leads to a lower interest rate, currency depreciation, and current account surplus, as a way of preserving aggregate demand as the economy adjusts to the new Ramsey equilibrium path.

Since there are more illegal immigrants in some sectors than others, the effects on sectoral output differ significantly. These are shown in Figure 3. Durable manufacturing and agriculture are initially the hardest hit because that is where the reductions in workforce are highest and the durables sector is also adversely affected by the loss of investment caused by the loss of output. The decline in output means there is less need for the existing stock of capital, which means less investment than otherwise. Economy-wide

![Figure 3: Effects on USA output from deporting illegal immigrants, % deviation from baseline](image-url)

investment could be 1.4 per cent lower by 2019 and 2020 than baseline (Figure 4, top right panel). The composition of investment across sectors is very different. The largest falls occur in the sectors losing more workers because the return to capital in those sectors is hit hardest. In addition, the durable goods sector provides most of the investment goods for the rest of the economy, so that sector takes a second hit from the aggregate slowdown in investment. Output in durable manufacturing could be 3.2 per cent below baseline by 2020, and for agriculture, 2.5 per cent lower.

With less output in each sector, aggregate real GDP must be lower. This is shown in Figure 4. Real GDP could be 1.6 per cent below base by 2020, before settling at 1 per cent below what it might otherwise have been from 2023 onwards because the workforce is permanently smaller. There is an interesting aggregate dynamic due to the sectoral shocks. Less activity and less investment means less demand for borrowed funds so real interest rates are less than baseline by some 45 basis points in 2017 and 2018 (Figure 4, bottom left panel). Less borrowing than otherwise, some of which would have come from overseas, means the current account deficit (as a % of GDP) is not as large and the temporary gain is shown in the bottom right panel of Figure 4. That implies the trade balance is better than otherwise (not shown), the equilibrating mechanism being a small initial real depreciation of the real exchange rate.

It is interesting that the stated goal of the policy to improve job prospects in the manufacturing sector is not achieved. In aggregate, GDP is lowered by the policy of deporting migrant workers but particularly output falls in the manufacturing sector.

**Figure 4:** USA macroeconomic effects from deporting illegal immigrants, deviation from baseline
(ii) Fiscal policy—tax and spending (important sectoral and macroeconomic effects)

McKibbin and Stoeckel (2017) also modelled a change in US fiscal policy. Although the actual policies were unclear at the time of writing, and still are, the modelled policy was:

- cut corporate tax rate by half;
- cut household tax rate by a quarter;
- increase spending on goods and services by 2 per cent of GDP to capture the infrastructure spending, defence spending, and the cost of building a wall on the Mexican border.

The combination of company tax cuts, personal tax cuts, and the spending of an extra 2 per cent of GDP, all unfunded, expand the fiscal deficit. The net effect on the deficit is shown in the top left panel of Figure 5, which is to expand the fiscal deficit by an extra 5 per cent of GDP (with little variation) relative to the baseline or ‘business-as-usual’ case. There is an immediate boost to consumption (top right panel) of 3 per cent above baseline in 2017, this falls to 1 per cent in 2018, stays that way until 2024, and thereafter declines to be below baseline by 2 per cent beyond 2040. The reason for that pattern is the extra debt to fund the consumption has to be paid for, in effect bringing future consumption forward.

The company tax cuts boost after-tax profits, lifting the return on capital and encouraging new investment (next panel in sequence). Investment rises to nearly 4 per cent above baseline in 2018 and thereafter declines quickly to 1 per cent above base before tapering back to baseline thereafter. Commensurate with the incentive to invest, there is an incentive to hire workers, so real wages rise, peaking at 2.4 per cent above baseline in 2022 in the pattern shown in the right-hand panel, second row, of Figure 5.

Since consumption and investment are large components of GDP (a common measure of economic activity), the next panel shows the spike that can be expected. Real GDP rises by 1.7 per cent above base in 2017 and falls close to baseline by 2022 and remains there thereafter. This gain is in terms of extra GDP over the level of real GDP that would have prevailed in the absence of the stimulus. In familiar growth terms (the year-on-year change), the next right-hand panel shows the extra 1.5 per cent growth in 2017 but slightly lower growth than otherwise in the years to 2024 because the extra boost to activity is declining after 2017. It makes it obvious that the short-term benefits of Trump’s fiscal policies are at the expense of the future.

Higher real wages and more activity initially sound impressive, which they are, but the trouble is that GDP is a measure of economic activity in the country regardless of who owns the productive assets. The substantial borrowing by the US as a consequence of the fiscal policies means we should consider net payments abroad of interest, profits, dividends, and rents—something the G-Cubed model does. In other words, we should consider the total income earned by residents of a country. That is, gross national product (GNP) is a better measure of the welfare of US residents. Once we consider payments abroad, the extra borrowings, capital inflow, and what happens to the trade balance, the picture is not as rosy.

3 GNP = GDP + net income from (or payments to) abroad.
Real GNP (bottom left panel of Figure 5) is higher than baseline in 2017, in line with the extra GDP, but is back to baseline by 2020, and steadily declines below baseline thereafter with the on-going borrowings needed to fund the deficit and therefore rising.
interest payments abroad. The net gain to US residents (the integral under the real GNP curve) is now negative.

The adjustment story is again familiar from the standard Ramsey growth model with OLG consumers. The US economy returns to a Ramsey growth path with a larger government and permanently higher government debt. The fiscal closure in the model is what is called an ‘incremental interest payments rule’ in which, within each country, a lump sum tax is levied on households to cover changes in interest payments on the stock of government debt. Thus, the long-run change in US government debt will be the long-run change in the fiscal deficit (in this case roughly 5 per cent of GDP) divided by the long-run growth rate in the model (1.4 per cent productivity growth in all sectors plus zero population growth). This works out to be an additional 357 per cent of GDP, which is clearly extremely large but not explosive because the tax rate adjusts to prevent an explosion of interest payments. It also takes more than a century to converge to this point.

The trade balance also initially moves into deficit as part of the early financing of this debt expansion, but eventually moves into substantial surplus as exports must exceed imports in the process of transferring resources to the foreign owners of US assets. The real exchange rate, although initially appreciating as capital flows into the US economy, must eventually depreciate to enable exports to rise above imports and meet this inter-temporal budget constraint facing the entire US economy. Because of the existence of non-Ricardian consumers, interest rates have to rise permanently in the US to crowd out investment and achieve a balanced growth equilibrium. Lower investment results in a lower capital stock. With less capital, output will be lower. This makes it even harder to generate resources to service foreign holders of US assets. Thus, the real exchange rate must depreciate even more in the long run, to generate the transfer of resources to foreign owners of capital and debt, than would be the case in a model with Ricardian consumers. It is hard to see how the approach proposed by Hendry and Muellbauer (2018) could capture this critical, externally driven, structural adjustment process.

As mentioned above, a string of extra deficits and borrowings expands the government debt ratio, which could be double what it might otherwise be by 2040 (right-hand panel). This outcome is problematic: Congressional Budget Office (CBO) projections of future government debt in the US under existing policies show a large and rising debt burden to around 150 per cent of GDP by 2047 (CBO, 2017). Recall that this model incorporates any extra tax revenue from the extra domestic activity, but these results show that the tax cuts dominate. The expansion is not ‘self-financing’.

It might be thought the extra early stimulus would boost inflation in the US. It tends to, but there is more going on; the capital inflow necessitated by the extra corporate and government borrowing causes the exchange rate to appreciate, lowering import prices, the fall being enough initially to offset any tendency for inflation to rise. Inflation could fall to 0.7 percentage points below baseline in 2018 (Figure 6, top left panel). Note short-term rates rise (right-hand panel) because, under the monetary policy rule, the Fed is targeting both activity and inflation and there is some policy tightening as a result.

Extra borrowing to fund the deficit causes long-term interest rates to rise (left-hand panel); rates settling at nearly 1 percentage point higher than otherwise over the long term because of the existence of non-Ricardian consumers. Higher rates encourage capital inflow which must be matched by deterioration in the current account deficit.
This deficit (as a percentage of GDP) is 5.7 per cent worse in 2018 than it would otherwise be. The trade balance shows a similar pattern early on since it is a large component of the current account, but later there is a difference (right-hand panel). The deviation from baseline in the trade deficit gradually narrows, and, although not shown on the chart, would turn positive at some distant point in the future to fund...
the increasing interest payment component of the current account on the balance of payments. This initial deterioration of the trade balance will cause a problem for the Trump Administration, given its rhetoric on this measure outlined during the election campaign and subsequently.

The mechanism by which the trade balance deteriorates to enable the capital inflow is an appreciation of the real exchange rate. The US dollar appreciates by 12 per cent in 2017 (expressed in real effective terms) relative to what it otherwise would have been. The appreciation encourages imports (and seen earlier keeps inflation low) and discourages exports. Agricultural exports are hit the hardest (Figure 6, bottom left panel) and could be 26 per cent below baseline in 2020, the year of the next presidential election. Note that total agricultural output will still be higher in 2020 (bottom right panel) because domestic consumption (around 80 per cent of production by volume) is higher as a result of the extra total domestic consumption shown earlier in Figure 5. But, had an even more disaggregated model been used, it would show that some mid-western states, heavily dependent on agricultural exports, would be worse off under the higher dollar, creating political problems come election time.

The decline in exports is across the board, so a similar set of arguments will apply to manufacturing, even though the value of total domestic manufacturing output should be slightly higher in 2020 because total consumption is higher. Under current policy proposals, it will not be possible for President Trump to stimulate the domestic economy and promote exports. That is why it is important to understand the economics behind the policies and avoid a resort to trade protection being the resolution of this contradiction.

There are important repercussions for other economies. Two other countries with large current account surpluses and bilateral trade surpluses with the United States that have already drawn the ire of President Trump are China and Germany. The effects of President Trump’s fiscal programme on Germany and China are shown in Figure 7. The extra borrowing by America to fund the expansion has to come from domestic and overseas sources. Germany and China are two such sources. So the capital outflow from these two countries must cause their current account surpluses to rise and with them the trade surpluses (Figure 7, top four panels). In 2017 Germany’s trade surplus could rise by 1.9 per cent (as a share of GDP) over baseline and for China the increase could be 1.6 per cent. Notice the different trends between the trade balance and current accounts. As income earnings continue to rise on the extra loans made to America, the current account (as a share of GDP) continues to increase over baseline. But eventually this extra income will be spent and the trade balance would fall below baseline if the graphic was extended far enough out.

The end result is that for Germany and China, net exports rise, especially in the early years, and mirror the deterioration of the trade balance in the United States. The mix of export changes is different across the sectors, though—shown in the bottom panels of Figure 7. The differences are due to changing exchange rates across countries due to differences in comparative advantage and in capital flows.

More net exports could be thought to encourage more investment in Germany and China. But higher interest rates in the US also lead to higher interest rates in Germany.

---

4 In May this year President Trump described Germany’s trade surplus with America as ‘very bad’ and it needed to change (Jacoby, 2017).
Figure 7: Effects on Germany and China from US fiscal programme, deviation from baseline

and China due to linkages between capital markets. This effect dominates and so investment in both Germany and China falls initially below baseline (Figure 7, second bottom panel). This fall in investment is more important in China than for Germany since investment in China as a share of GDP is much higher (over 40 per cent of GDP). So although net exports in China are higher, this positive effect on activity is outweighed by the negative effect of investment and GDP falls below baseline by 1.4 per cent in 2017 (bottom panel). For Germany, investment as a share of GDP is smaller than for China and net exports larger so on balance there is a small initial boost to GDP which rises by 0.6 per cent in real terms in 2017. It falls below baseline after that because the boost from net exports declines while the fall in investment continues to grow for another 2 years before returning to baseline. Longer term, China and Germany are worse off in terms of real GDP by around 0.4 per cent below baseline.

V. Conclusion

This paper argues that important aspects of the world economy that tend to be missing in most macroeconomic models are the importance of global linkages in trade and financial markets, the role of relative prices (sectoral disaggregation), and changes in risk premium. These have been important sources of shocks for most countries since 2000. Despite this, it is still the case that there are closed economy models being used for policy analysis in major economies, with no role for the exchange rate or for the major sources of economic growth in the global economy. Many models are missing the sources and consequences of the major shifts in risk and relative prices since 2000 and therefore the nature of the most important shocks is not identified because there is no detailed structural model.

While aggregate models will remain useful for exploring some issues, there is also a case for using models with a consistent macroeconomic closure, such as found in DSGE models, but allowing for greater complexity in the modelling of adjustment at the sectoral level to aggregate and sectoral shocks. With the world economy continuing to undergo major structural change due to the growth of large emerging countries, and differential degrees of technological innovation in different sectors, future shocks are more likely to involve significant changes in relative prices and risk. If these features are not incorporated into our macro models for policy analysis, then it is likely that many structural problems will be interpreted as cyclical issues at the aggregate level. This is problematic for understanding the consequences of the shocks and the nature of policy design.

As Blanchard (2017) wisely argues, there is a role for many different types of models for addressing different types of questions. There is a role for more complex but internally consistent models that capture observed changes in the global economy that must be relevant for the macroeconomic policy design in all open countries. The examples in the paper drawn from studies using the G-Cubed multi-country model, demonstrate the insights that can be gained from allowing for complexity when it is relevant.
References


The sea of learning has no end

Encourage your library to subscribe to this journal by completing a library recommendation form at:
academic.oup.com/journals/pages/librarians
Articles published online ahead of print are available to read and cite with our Advance Article service.
Stay alert to the latest content

Sign up to any of the following free alerts to receive the latest content straight to your inbox:

- **New issue alerts**
  Receive alerts via email each time an issue is published online

- **Article Activity alerts**
  Receive alerts when a specific article is cited or corrected
  Track topics / authors / articles

- **RSS feeds**
  Receive alerts with the ‘headlines’ from the latest journal content

- **Advance Articles**
  View papers online weeks ahead of print

Visit [academic.oup.com/my-account/register](http://academic.oup.com/my-account/register) ‘Email Alerts’