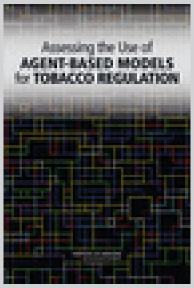

Assessing the Use of Agent-Based Models for Tobacco Regulation

ISBN
978-0-309-31722-1

280 pages
6 x 9
PAPERBACK (2015)

Robert Wallace, Amy Geller, V. Ayano Ogawa, Editors; Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation; Board on Population Health and Public Health Practice; Institute of Medicine

3

Building Effective Models to Guide Policy Decision Making

Although policy makers have long looked to behavioral models to guide their decision making, there is no accepted set of recommendations or best practices for how to manage this process. In accordance with its statement of task, the committee reviewed the uses of agent-based modeling (ABM) in policy decision making and how this method fits into a broader methodological toolkit. The goal of this chapter is to provide guidance on (1) understanding the conditions under which models—specifically individual-level models—are appropriate and useful in aiding policy decisions; (2) elucidating the empirical and theoretical challenges of specifying model inputs and interpreting model outputs appropriately; and (3) providing guidance for navigating key modeling decisions, including determining the appropriate levels of verisimilitude and aggregation, dealing with issues of model specification and evaluation, and quantifying uncertainty. Fortunately for tobacco control policy modelers, many regulatory authorities and academic fields are struggling with related problems in terms of model specification and inference. Their efforts offer a wealth of examples and experiences to draw from.

The organization of this chapter is as follows. The motivation for models in policy decision making is described. The committee articulates specific mechanisms through which human behavior may depend on the behavior of others as well as on features of the local environment. Then the major challenges to getting empirical evidence to adjudicate among these alternative mechanisms are reviewed. Next, a number of key distinctions in modeling are introduced, including micro- versus macro-level models, analytical versus computational models, and models that incorporate varying levels of detail in representing a given process. The appropriateness of each type of model under different levels of uncertainty and data availability is discussed. The committee suggests methodological strategies for specifying individuals' behaviors within micro-level models and for assessing how uncertainty in model inputs translates into uncertainty in model outputs.

THE CHALLENGE OF ANTICIPATING AND UNDERSTANDING POLICY EFFECTS

Policies can backfire when they fail to account for how people change their behavior in response to an intervention. This is known as "policy resistance" in the public health literature (Sterman, 2006) and "blowback" in covert operations. It goes back to old social science literature on the "law of unintended consequences" (Merton, 1936; Smith, 1759). The basic issue is that individuals' behavior often depends on the behavior of other people or features of the social environment, or both. Any policy that aims to change behavior or outcomes can result in a chain reaction of events that can potentially undermine the efficacy of that policy.

This problem arises in many substantive areas. To take an example from tax policy, if workers allocate their time to maximize both earnings and leisure, an overly stringent income tax may lead them to cut back on hours worked, which may in turn reduce total government revenue from taxes (Saez et al., 2012). Within the domain of transportation, antilock brakes can cause people to drive more aggressively, thus partially offsetting their safety benefits (Wilde, 2001). Closer to home for readers of this report, there is evidence that low-tar and low-nicotine cigarettes may actually increase the intake of carcinogens, as people smoke more frequently and hold the smoke in their lungs for longer (HHS, 2010; NCI, 2001).

Although, as the above examples show, a policy may generate negative feedbacks, positive feedbacks may also occur, enhancing the effectiveness of the policy. In the classroom, the provision of tutoring or other special help to some students may indirectly aid the learning of other students as members of the class interact with one another. Persuading one person to stop smoking may influence friends and family to stop smoking as well. Such positive feedbacks are sometimes called *social multipliers* (Manski, 1993).

Whether feedbacks are negative or positive, a central challenge for policy makers is to anticipate how organizations, corporations, and individuals will react to changes in incentive structures and features of the environment. Anticipating this response can be difficult for a number of reasons. One challenge is that knowledge of human behavior is limited and that it is difficult to infer from past behavior how people will respond to novel situations. A related problem is that people's behavior is both influenced by and also influences the behavior of others, through direct interactions (e.g., social influence and peer effects) as well as features of the social environment. This makes it difficult to assess the global effect of a policy or to anticipate its efficacy at different scales of implementation.

For example, a housing policy that encourages a small number of individuals with low income to move to higher-income neighborhoods may appear to successfully accomplish its intended goal of economic integration. However, if that policy were to be expanded to a larger population, the higher-income residents of those neighborhoods might move out (presumably, because the neighborhood has declined), which in the end would leave these lower-income households no better off than before. Conversely, an antismoking policy targeted at a small group of persons may have little positive effect, but one targeted at a larger group may generate a change in social norms that induces persons not within the target group to stop smoking as well. To be maximally effective, policy makers must be able to assess their proposed interventions' total effect, including how affected individuals, organizations, or institutions might adapt to a new environment or change their behavior in reaction to what others are doing.

Anticipating the Effects of Policies

Historically, the "gold standard" for evaluating the effects of a public health intervention has been an analysis of treatment response using data from randomized controlled trials (RCTs). This approach overcomes the fundamental problem of causal inference: For any given treatment unit, the counterfactual outcome is never observed—that is, what would have happened if that unit had or had not received the treatment. By removing the possibility of selection bias, RCTs provide a more rigorous test of treatment effects than do observational studies.

Information gleaned from RCTs alone is often insufficient for guiding policy decision making. Perhaps the most obvious issue is that it may not be feasible or appropriate to carry out the desired RCTs. This is partly due to practical limitations: It is impossible to design RCTs to test all possible policies. There may also be legal or ethical restrictions that make RCTs inappropriate. In some cases quasi-experimental methods (e.g., instrumental variables) or modeling strategies (e.g., propensity score matching) can be used in an attempt to mimic experimental research design, but these approaches may require one to make implausible assumptions in order to produce inferences.

In addition, RCTs are ill-suited for evaluating policy effects when the behavior of different individuals is interdependent. Indeed, the standard analysis of RCTs makes the assumption that one person's treatment outcome is independent of who else received the treatment. When the efficacy of one person's treatment depends on whether others received the treatment, the methodology falls apart. For example, RCTs have limited ability to inform society about the effectiveness of vaccination policies for a population susceptible to infectious disease. An RCT with a small treatment group might provide information about the payoffs to vaccination when a small number of people are vaccinated, but credibly extrapolating from this to a larger treatment population may prove to be impossible, for two reasons. First, any individual's decision whether to get vaccinated may depend on how many others are getting vaccinated. Second, the danger of catching a disease varies with overall rates of vaccination. An RCT examining the effectiveness of a tobacco use cessation treatment program would have similar problems. The treatment of one individual could have beneficial effects on others—for example, on the individual's spouse or peers, who may quit in reaction to the treated individual successfully quitting.

Finally, traditional analyses of RCTs tell us only what does or does not work; they provide no information on the reasons why an intervention worked or not. Thus, the information gleaned from RCTs and quasi-experimental methods may lack external validity. This makes it difficult to extrapolate the effects of interventions implemented in one context to a different context or to infer the expected effects of novel interventions from knowledge about the effects of prior interventions (Cartwright, 2007; Heckman, 2008; IOM, 2012; Manski, 2013). On the other hand, experimental or quasi-experimental estimates can be used to guide theory and shed light on underlying structural relationships. In a complex world, moving toward a more structural approach—and away from a "black box" analysis of experiments (that is, not having the ability to understand the inner workings of the processes under study)—will help researchers do more than estimate an intervention's causal effect. How a policy is expected to work within the larger social context requires system-level knowledge and a sense of the behavioral mechanisms through which it operates (Sampson et al., 2013).

Structural Models

Structural models use a set of equations or rules—expressed analytically or computationally in programming code—to describe different possible worlds. The specification of the model is dictated by theory, prior knowledge, and other inputs that determine which features of a given process to highlight and which to leave out. These assumptions, combined with data, produce a set of inferences about what will happen under a given set of conditions. This modeling approach includes (but is not limited to) macro-level simulation models, such as system dynamics models, and micro-level simulation models, such as ABMs. The appropriateness of a given modeling strategy depends on the theory brought to bear and on the available empirical evidence.

Structural models typically attempt to capture behavioral relationships or parameters that hold true across a range of social conditions or take those conditions as inputs that affect behavior. This requires a reasonably deep understanding of the incentives that drive behavior. For example, one might observe an association between neighborhood poverty and rates of teenage pregnancy. An example of a superficial model of this process would be one that translates this aggregate correlation into individuals' decisions: As poverty increases, the likelihood of a young woman having a child goes up by some amount. However, this model ignores the underlying motivations for these women's decisions, how other people may influence those decisions (e.g., parents and partners), and how decisions are predicated on these women's beliefs about the benefits and costs of having a child, which may depend on what opportunities are available to them. Without taking into account these underlying motivations for behavior, the hypothetical model is extremely brittle in its ability to make inferences about how women would behave under alternative scenarios. The general point is that the more fundamental—or "deeper"—the relationships captured in a model, the more effective the model is at exploring the implications of a wider set of policy scenarios (Blume, 2015; Heckman, 2008). Thus, an additional criterion for model usefulness is that it has parameters that are sufficiently fundamental to cover all the policies under consideration (Marschak, 1974).

One challenge in using structural models effectively is specifying them in a way that is empirically defensible and that allows for a clear and rigorous quantification of the assumptions embedded in the model (NRC, 2014). To be useful, a model must be able to quantify how uncertainty in the model's inputs translates into uncertainty in the model's outputs (Manski, 2013). Structural models vary widely in how complicated they are. Models that include more parameters and greater verisimilitude do not necessarily make more assumptions than simpler models. This is because in many cases researchers can conceive of models with more parameters than there are available empirical inputs. Regardless of model complexity, models are more credible if parameters are backed up by hard evidence or, at minimum, a well-developed theory. Whatever the level of detail used to represent a process, models that guide policy must be both credible and sufficiently explicit in their assumptions about the process under investigation.

The degree of model verisimilitude may reflect different goals. Some models are designed to run virtual experiments to determine the outcomes that could be expected from implementing different policies. Other models have a simpler goal: to identify the potential pitfalls or unanticipated consequences of a given policy or to get some sense of what RCT design would be needed to accurately assess policy effects. In both cases, if the models are to produce valid inferences, they must be able to capture accurately the distribution of outcomes that might be expected under a given set of conditions and to suggest which outcomes are more or less likely. For some policy makers and model developers, one attraction of ABM is that it allows for almost unlimited detail in representing the process under investigation. More complicated models do not necessarily generate more accurate predictions, especially if data, theory, and other model inputs are insufficient to identify the foundational parameters of the model (Sanstad, 2015).

MECHANISMS THAT GENERATE FEEDBACK BETWEEN BEHAVIOR AND SOCIAL ENVIRONMENTS

There are several different ways that people's actions can be influenced by their environment, which includes both what others are doing ("social interactions") and the institutional, political, and organizational factors that shape people's incentives, such as the regulatory environment. Policies may be more effective when they can directly target the specific mechanism that gives rise to the process under investigation, and thus policy makers need to evaluate an ABM's ability to explicate the behavioral mechanism under investigation. This section reviews the theoretical and empirical literature on mechanisms governing contingent behavior and suggests some ways in which these insights might be fruitfully applied in the domain of tobacco regulation. Note that here the focus is on mechanisms that occur "above the skin" (for example, environmental or societal factors). For a review of structural models that attempt to capture interactions "below the skin" (for example, genetic, metabolic, and neurobiological factors), see Hammond (2015).

Social Interactions

People's behavior is often shaped by what others are doing. This type of phenomenon is often referred to as social interactions, social influence, or spillover effects. Manski (2000) distinguishes between three types of social interactions. First, there are *constraint interactions*, which cause an action to become less desirable and available as more and more individuals engage in it. One example of this is freeway congestion: Freeway driving is most attractive when there are few people on the road and increasingly less desirable as more and more people use the highway. Second, *expectations interactions* occur when people draw inferences about expected outcomes of a given action or about difficult-to-observe attributes of a situation or person based on prior experience or an outside body of knowledge. One case of this is statistical discrimination. For example, an employer may have certain expectations about young workers—that they are more likely to quit their jobs in order to go back to school—and this influences the employer's enthusiasm for hiring from this population. Similarly, teenagers may observe the effects of smoking on older relatives, which shapes their beliefs about the effects of tobacco on health. Finally, *preference interactions* occur when a person's ordering of attractiveness concerning some set of alternatives depends on the choices of others. For example, in the case of "white flight," each time white persons leave a neighborhood because they cannot tolerate the presence of minorities, they leave the neighborhood a bit less white behind them, thereby inducing other whites to exit as well (Schelling, 1978).

These are theoretically distinct processes, each of which suggests different policy interventions, but in practice it is difficult to empirically distinguish them. Moreover, although the cases outlined above represent different instances of *endogenous effects*, people who share the same social context may display similar behavior even in the absence of these social interactions. For example, similar behavior may arise from *contextual effects*, which refer to the way in which people's behavior is shaped by a shared social environment, such as neighborhood composition or school quality. Also, a group of people may share similar behavior or outcomes due to *correlated effects*, which refer to a situation in which people share the same attributes or opportunities. For example, people within the same birth cohort may have similar career trajectories, on average, in large part because they face the same job market conditions. A challenge for researchers who believe they have identified some sort of endogenous behavior is to identify the effects of the social influence apart from shared opportunity structure (Manski, 2007).

Imagine a case in which some correlation is observed between peer group membership and whether and how much a teenager smokes cigarettes. There are four ways in which this result might come about. First, students in the same peer group might influence one another's smoking behavior through the availability of cigarettes or through peer pressure, or both. Second, the students in the same peer group may share similar individual attributes or risk factors (e.g., gender, family resources, parents' education) that affect smoking. Third, the students may affect one another's smoking behavior through behavior other than their own smoking behavior. For example, if students in the peer group are more likely to cut class, and if cutting class leads to higher rates of smoking, a correlation between peer group membership and smoking could be observed. Finally, students may inform one another about the existence and properties of different forms of tobacco (e.g., cigarettes, e-cigarettes, hookah, chewing tobacco). These different pathways would suggest different policy interventions. In the second case, where behavior is a function only of individuals' attributes, there are no social interactions.

One difficulty in identifying social interactions stems from the fact that the average behavior within a group is itself a function of the behavior of group members. Thus, observing a correlation between peer group membership and smoking behavior does not tell us whether peer groups influence the behavior of their members or the groups' behavior is simply aggregating over the behavior of group members. This is known as the "reflection problem" (Manski, 1993, 2000, 2007). If the data contain sufficient variation between and within groups, it may be possible to determine whether individual attributes alone can explain variation in behavior across groups. Even if researchers have reason to believe that group affiliation shapes behavior, they still must distinguish among the different types of social interactions described in the previous paragraph. In many empirical cases, it is difficult or impossible to identify the groups that are actually influencing behavior. The reflection problem is even more difficult to resolve when group affiliation is unknown. See Manski (1993, 2000) and Blume et al. (2010) for more detailed discussion of this issue.

An example of the challenge of identifying the presence of social interactions—let alone determining the nature of those interactions, if they exist—can be found in the debates over the effects of peer influence on smoking. Christakis and Fowler's (2008) study examines whether knowing people who quit smoking makes it more likely that a given individual will also quit smoking. Their results suggested that friends, coworkers, siblings, and spouses had dramatic effects on adults' smoking behavior. For example, they found that a person is two-thirds more likely to quit smoking if his or her spouse also quits. A coworker, sibling, or friend quitting had a smaller, but nontrivial effect—ranging from one-quarter more likely to quit smoking in the case of a sibling to over one-third in the case of a friend. Identifying true "social contagion" effects requires separating out the effects of *homophily* (people's tendency to select others who resemble them on observed or unobserved attributes) and *shared social environment* from the effects of *social influence* (Aral et al., 2009; Shalizi and Thomas, 2011). Indeed, later studies that used more rigorous strategies for controlling for unobserved features of the environment and selection artifacts have found that peers have far more modest effects on smoking behavior (e.g., Fletcher, 2010; Fletcher and Ross, 2012). One lesson here is that policies aimed at encouraging or discouraging the spread of behaviors in networks must be backed by rigorous empirical studies that convincingly separate homophily effects from effects due to social contagion (Aral et al., 2009; Shalizi and Thomas, 2011).

More generally, to be useful for informing regulatory policy, modeling efforts must capture meaningful aspects of the social process under investigation. If the goal is to understand how interdependent human behavior will shape the outcomes experienced under a given policy, a serious empirical effort is required to determine the underlying mechanism at work. It is not enough to hypothesize different mechanisms and use a model to determine whether those mechanisms lead to different outcomes. The model may be misspecified to the point where a "sensitivity analysis"[1] provides no information at all on the true sensitivity of model outputs to inputs (Sanstad, 2015).

Institutional Factors

This chapter has focused on how individuals' behavior influences the behavior of others, but there are also higher-level structures (e.g., tax policy or safety regulations) that shape people's choices. There is evidence across a number of policy domains that if the incentives or risks associated with a given behavior are changed, people will likely behave differently. For example, some evidence suggests that the development of highly effective HIV treatments has been associated with an increase in unprotected sex among people living with HIV in the United States (Katz et al., 2002; Lightfoot et al., 2005).

The behavior of organizations and other coalitions is also influenced by behavioral incentives. Failure to account for those incentives may lead to unexpected and undesirable results. For example, an increase in U.S. corporate taxes may result in some firms decamping for countries with lower tax rates, thereby reducing total U.S. tax revenue (Devereux and Maffini, 2007). Similarly, tobacco companies make strategic decisions that are influenced by the current regulatory environment and will work to counteract the efficacy of policies aimed at reducing smoking. Congress or states may change taxes on cigarettes, for instance, but the tobacco companies may respond by offering coupons or bulk discounts (Arno et al., 1996; Henriksen, 2012; Loomis et al., 2006). A model that aims to predict how people's or firms' behavior might change under a different incentive structure must therefore understand the reasons for their behavior. Later in this chapter, strategies for specifying models of behavior that try to account for these motivational factors are discussed.

Conclusion 3-1: The committee concludes that a deep understanding of human behavior, decision making, and incentive structures is important for agent-based models and other models that are used to understand how interdependent behaviors shape the outcomes of a given policy. Regardless of the model type, if the behavior is not plausible, the model is not likely to be informative.

[1] Sensitivity analysis is "an exploration, often by numerical (rather than analytical) means, of how model outputs (particularly QOIs [quantities of interest]) are affected by changes in the inputs (parameter values, assumptions, etc.)" (NRC, 2012, p. 117).

This conclusion is especially relevant when a model is intended to explore which outcomes might occur (or how people would behave) under alternative scenarios.

There is some debate about how plausible a model's representation of individual behavior must be in order for that model to be informative. Friedman (1953) argued that models do not need to represent the underlying process accurately to be useful—the model only has to predict well. A practical problem with that line of thinking is that it presumes the existence of evidence that the implausible model actually predicts well—but how would this be known ex ante? It can only be discovered ex post. Another problem is that models that do not accurately represent the underlying process under investigation at some degree of fidelity can be brittle, losing their predictive power when conditions or incentives change. This is in contrast to structural models that capture key features of a process (Marschak, 1974).

Sometimes models that inaccurately represent individual behavior yield qualitatively accurate aggregate predictions. For example, Schelling's model (1978) of how individual decisions generate aggregate patterns of segregation includes several implausible assumptions about behaviors (for example, that people make decisions about where to live based on whether their own racial group is the local minority or majority and that there is no cost to moving). The Schelling model did provide the important theoretical and policy-relevant qualitative insight that segregation can emerge even though people have preferences for racial integration. However, it does not give a credible quantitative prediction, and so it does not provide a suitable basis for predicting when segregation will emerge in specified real-world settings. Moreover, it would be a mistake to use the Schelling model to predict how households might respond to pricing incentives, counseling, or other interventions aimed at promoting neighborhood diversity.

Conclusion 3-1 is most relevant when models will be used to inform policy decisions. Models that do not include an understanding of human behavior, decision making, and incentive structures can be informative for some purposes. However, it is the committee's view that models developed for the purpose of anticipating the effects of policy decisions need to have some anchoring in real-world behavior.

Recommendation 3-1: When developing an agent-based model (or similar modeling approach), the Center for Tobacco Products should consult with subject-matter experts to identify the plausible behaviors and focal processes at work from the beginning of the model development process.

Chapter 4 also discusses the need for input from subject-matter experts throughout the lifespan of a model. An essential feature of ABMs is the representation of agents in the real world, and because agents often have distinct characteristics and behaviors, non-experts can inadvertently misrepresent agent behavior. This makes collaboration with subject-matter experts essential at all stages of model development.

OVERVIEW OF TYPES OF STRUCTURAL MODELS FOR INFORMING POLICY DECISIONS

Thus far, the committee has discussed the role that models can play in guiding policy decisions, and it has reviewed some of the behavioral mechanisms that can lead to feedback between individuals' behavior and the social and regulatory environments. This section provides a high-level overview of the types of models that are used to capture this type of feedback behavior. A number of key features of such models are reviewed, including whether they have an analytical versus computational solution and whether they capture phenomena at the individual or group level. The chapter also addresses the confusion regarding the distinction between microsimulation and ABM, which the committee believes limits researchers' and policy makers' ability to incorporate lessons learned and best practices from the array of studies in different disciplines. This section ends with a discussion on how to define the appropriate level of model specificity and how to anticipate and understand the different forms of equilibrium outputs generated from models.

Analytical and Computational Models

Recall that a structural model is a set of equations or rules for how the individuals or other units in a specified population interact and are influenced by their environment, and it can be implemented both analytically and computationally. The National Research Council defines a computational model as "computer code that (approximately) solves the equations of the mathematical model" (2012, p. 110), whereas analytical models can be solved mathematically in closed form, that is, the solution to the equations used to describe changes in a system can be expressed as a mathematical analytic function (NRC, 2007).

If the behavioral rules or the population structure is sufficiently simple, it may be possible to determine analytically (i.e., mathematically) how the state of the population changes over time and whether the population gravitates toward a steady state (equilibrium). However, if the rules are sufficiently complex or the population is too heterogeneous, it may be impossible to determine the dynamics of the system or to derive the steady state. In some cases, there is no steady state solution (e.g., Salop and Stiglitz, 1982). In this case, the analyst must simulate the process iteratively by applying the rules and updating the population composition to arrive at a final solution or observe the dynamics. Note, however, that some models only determine equilibrium outcomes without specifying the process by which the social system attains equilibrium. For these models, intermediate solutions have no substantive meaning, and thus there are no dynamics.

Even seemingly minor relaxations of assumptions may make it impossible to solve a model analytically. For example, a simple model of disease spread that assumes that interactions among individuals are equally likely ("random mixing") can be solved mathematically. As soon as one relaxes this assumption and allows for different individuals to have varying rates of exposure, the heterogeneity in agents' disease risk makes the resulting model analytically intractable (Blume, 2015; Osgood, 2007). Thus, the results must be simulated.

In general, analytical models have relatively simple specifications of behavioral rules that assume tractable forms of interactions (Grazzini et al., 2013). Models with analytical solutions are often more restrictive than simulation models, as they have fewer parameters and simpler interaction structures. This is not necessarily bad. Analytical models have several advantages over computational models. First, because the equilibrium solutions are derived mathematically, it is possible to identify the whole space of solutions. This is particularly important when there exist *multiple equilibria* for the same set of model inputs, as is often the case with social interactions models. (This point is discussed in greater detail in the next section.) For policy makers it is of great interest to identify the potential for multiple equilibria, as this suggests that, insofar as the model has captured fundamental features of the process, the same starting conditions might, depending on how a process unfolds, end up in very different final states. A possible policy goal may move from a "low level" equilibrium to a "higher level" equilibrium (Moffitt, 2001). Feedback suggests the possibility that, with sufficient understanding of incentive structures, one might "harness" the interactions so that there are bigger payoffs relative to costs. In addition, analytical models can also reveal the path to equilibrium, which may be more important than the equilibrium itself. For example, if the goal of the model is to anticipate outcomes over a finite time horizon, knowing the equilibrium outcome is of little use if it applies to a world that exists decades or centuries into the future.

Second, analytical models allow the analyst to determine the stability of model equilibrium—in other words, how likely it is that the model will return to a given state if it is slightly perturbed away from that state. Whether or not an equilibrium solution is stable may have important policy ramifications. Regardless of the attractiveness of a particular outcome from a social welfare standpoint, if that outcome is highly unstable, it may be impossible to maintain it in real-world situations. Moreover, understanding the "gravitational pull" of different equilibrium solutions provides insight into how the process under investigation may translate from one stable state to the next. Thus, evaluations of policy outcomes must take into account not only the attractiveness of a given result but also the likelihood of maintaining it over time. Analytical models allow researchers and policy makers to take both factors—which equilibria are possible and which are most likely to be sustained—into account. This is much harder to do with simulation models, which may not identify highly unstable equilibrium solutions.

The potential downside of analytical models is that they may only be able to represent a small number of features of a real-world setting and may make simplifying assumptions that reduce the empirical plausibility of the process represented. However, it is a mistake to assume that simply adding more features to a model will provide a better representation of the process under consideration. This is especially true if there is a great deal of uncertainty in how those features should be specified. Researchers need to be clear about what they are giving up for the benefits of added verisimilitude. Users of simple models may be able to understand model behavior very thoroughly, whereas users of more complicated models may lose their grasp on how a given set of results came to be. Therefore, one would only move to a complicated model if the simpler form is understood and there is a reasonably clear and accurate empirical representation of the more complex process.

It is often useful to start with an analytical model and then expand on it slowly, making effort to tie results from computational solutions to their simpler foundations. Examples of this approach include Brown et al.'s (2004) analysis of green belts and Epstein et al.'s (2008) analysis of the coupled spread of disease and fear about the disease. In these cases the researchers took pains to try to understand completely the simple dynamics involved before turning to more complicated and realistic simulations. Furthermore, modelers may use both analytical and computational methods to describe the relationships between individual choice making and aggregate outcomes. For example, a model may be solved analytically to determine the optimal behavior of each individual agent conditional on the behavior of other agents, but then solved using simulation to determine the equilibrium outcome among many agents.

Equilibrium

As mentioned in the previous section, the focus of a model may be to predict the steady states (equilibria) of a system or to predict its dynamics out of equilibrium. When is it useful to focus on out-of-equilibrium versus equilibrium predictions? It depends on how stable an equilibrium is and how long the system being modeled takes to reach the equilibrium. Consider the metaphor of a rocking chair. If the rocking chair is perturbed, it might start out rocking quickly but then settle into a steady state or equilibrium. Because this happens fairly quickly, it may be valuable to develop a model that focuses on equilibrium conditions.

However, it may take a very long time—perhaps decades—for a real social system to reach equilibrium. In this case a model that focuses on equilibrium would be useless, and it becomes important to be able to credibly predict the dynamics that would follow an intervention. Imagine, for instance, that the U.S. Food and Drug Administration initiates an information campaign. People learn something new about tobacco, and their resulting behavior changes in turn affect other people. In this case, the dynamics of social learning would unfold gradually. If the process took only a few months to reach equilibrium, it might suffice to analyze only the equilibrium conditions. On the other hand, if it would take 50 years for the dynamics to play out, then an equilibrium model would be less useful. Thus it is important to consider the speed at which equilibrium is reached, as this has policy implications. For tobacco control policy, not much is known about the time scales over which equilibria may be reached. An example of when it might make sense to examine only equilibrium conditions is a model of the effect of price on smoking prevalence, which falls rapidly following a price increase. Several econometric models have been developed to estimate the final effect of such a price hike (Chaloupka and Pacula, 1999; Chaloupka and Warner, 2000). These models do not try to represent how smoking prevalence changes over time; they focus only on the final value.

Of course, exploring the time to equilibrium requires that the model be initialized in some starting condition that is anchored empirically. Moreover, the model needs to have a meaningful time scale so that "model time" may be mapped onto "real time."

Individual-Level and Aggregated Models

Another decision that must be made in the modeling process is whether the basic units of the model will be at the level of individuals or aggregated groups. The same process may be represented at different levels of aggregation. For example, one might specify a model of teenage smoking behavior that assumes that school attributes influence girls' and boys' smoking decisions differently but that all girls and all boys have the same response to the environment. In this case, one could specify a model that represents the process for girls and boys separately, but not for individual children. Or, by contrast, the analyst could allow for each child to have a unique set of inputs in the decision process and model the process at the individual level.

Both approaches have strengths and weaknesses. Aggregate models are often easier to build and interpret, but they can only handle a limited amount of population heterogeneity. If population heterogeneity is a key feature of the process under consideration, or if the model incorporates individual-specific trajectories or experiences (e.g., work histories), the analyst will likely need to specify the model at the individual level in order to allow each person to have a unique profile.

From an implementation standpoint, aggregate models have certain advantages over individual-level models. They are more straightforward to construct and understand, and they often take less time and computational power to run. Finding empirical data to anchor them may also be easier. In addition, if the analyst wants to simulate the dynamics of a very large population (e.g., the population of the United States), individual-level models can easily become unwieldy. Researchers have to weigh the trade-offs.

Both aggregate and individual models can incorporate feedback effects across levels of analysis. However, it is difficult to incorporate social interaction effects into aggregate models if these effects occur at the local level. (If there are global interactions, for example, where all individuals respond to the total number of people working in the population, individual-specific response functions are not required.) A key challenge in implementing individual-level models is finding the empirical knowledge or data necessary to make them credible. The data demands for an individual model are higher than those for an aggregate level model, especially in terms of the plausible specification of individuals' behavior (see Chapter 6 for a discussion on data needs).

Microsimulation and Agent-Based Models

Within the domain of individual models, some scholars distinguish semantically between two types of models: *microsimulation* and *agent-based models*. Both involve the same basic procedure: Artificial agents are assigned a behavior, and simulation is used to assess the aggregate implications of that behavior. Both modeling approaches are operationalized through computer code. Thus, in theory, anything that is specified as an ABM can be specified as a microsimulation, and vice versa. It is important to note this commonality because, if viewed as two distinct approaches, the two research communities are less likely to benefit from each other's work. However, some argue that there are a number of differences between ABM and microsimulation, both in the research questions they consider and in their common practices.

For example, in a review of the literature on ABM, Macy and Willer (2002) claimed that the difference between ABM and microsimulation is that microsimulations assume no interaction among agents. And, indeed, many microsimulations do attempt to explore how heterogeneous populations respond to some change in policy or incentives, without allowing for interactions among agents. For instance, the Congressional Budget Office's (2007) microsimulation tax model explores how the U.S. population might respond to a change in tax rates, taking into account the fact that different types of people (for example, married and unmarried, men and women) have differential responses. This model does not specify that agents interact; rather, its goal is to compute the net response, taking into account the fact that people's labor force participation is contingent on their expected income after taxes. However, the committee found many examples of microsimulations in which the environment of the agents is generated from agents' previous decisions. As one example, Mare and Bruch (2003) used a microsimulation to determine the equilibrium segregation outcomes implied by agents in a residential mobility model.

Although from a purely technical standpoint microsimulations and ABM are the same modeling enterprise, the committee did find differences in how these techniques tend to be deployed. For instance, microsimulation models typically keep the agents' environments simple and abstract, as these models are anchored in even simpler analytical models for which the dynamics are well understood. ABMs are sometimes grounded in analytical models, but this is not standard practice. Also, ABMs may incorporate highly detailed environments in which the agents interact, drawing on maps and other geographic information. This is technically possible with microsimulations, but in general microsimulations tend to abstract away from spatial features of the agents' environments. In addition to their emphasis on spatial interactions, ABMs tend to emphasize other features of complex systems, including population heterogeneity, adaptation, and learning (Hammond, 2015). In short, there are not fundamental differences between ABMs and microsimulations, but there are historical differences in how these models have been specified and used by their research communities.

Conclusion 3-2: Researchers who use the terms agent-based modeling and microsimulation have different approaches to model specification. However, the committee concludes that from a technical standpoint these are the same enterprise (an individual-level model implemented via computer code). The committee believes that modelers would greatly benefit from best practices and lessons learned from applications that have been performed by the two research communities to address policy questions.

This report is focused on ABM, and it is the committee's sense that agent-based modelers would benefit from drawing on the large literature on microsimulations, especially in the context of policy decision making. For example, microsimulations have been used in tobacco control in recent years, and CTP and other agent-based modelers could look to those examples (Jeon et al., 2012; van Meijgaard et al., 2009). Thus, in the remainder of the report these two methods will be treated as technically the same approach, albeit with different implementation practices.

High-Dimensional Models and Low-Dimensional Models

Finally, as noted earlier, the model developer must decide on the appropriate level of the model's detail and empirical realism. The appropriate level of model detail depends on the research question, the intended use of the model, and the data that are available to empirically ground the model. It is important to note, however, that no matter what level is chosen, models provide only an imperfect representation of the real world, as computational models in general are not reality mirrors, nor are they intended for this purpose. ABMs can represent anything from low-dimensional, abstract worlds where agents are defined by just one or two attributes and interact in a highly stylized environment based on simple rules, to high-dimensional, highly detailed worlds where agents have many attributes, the environment contains a great deal of information, and agents engage in multiple behaviors (Bruch and Atwell, 2013).

It may be tempting to design an ABM that pulls in all the empirical data and knowledge available in order to create a highly realistic "laboratory" in which to explore policy questions. However, this approach is not usually the most productive, especially at the early stages of modeling, as the available data and knowledge of human behavior are generally not available. While data on demographic, biological, and social characteristics of individuals, families, or other groupings are often collected, data on how those units interact are generally lacking. ABM allows the developers to explore how important various mechanisms are when data are lacking and to assess the potential value of collecting these data; however, this introduces an added layer of uncertainty and raises the possibility of model misspecification. Furthermore, the model can become cumbersome and hard to manage when additional layers of detail are added, and it can be difficult to get clear analytic results (Blume, 2015). The success of a model is not determined by the level of granularity at which it represents a process; rather, its success is based on how successfully it facilitates the understanding of the problem or question under study.

Conclusion 3-3: The committee concludes that low-dimensional and high-dimensional models have complementary virtues and weaknesses. A more complicated model may have greater verisimilitude, but added detail per se does not ensure realism. A low-dimensional model, while abstracting from some features of the real world, may generate forecasts that are easier to understand and interpret.

Recommendation 3-2: The Center for Tobacco Products should develop and employ both low- and high-dimensional models, using both as appropriate to shed light on policy impacts.

SPECIFYING INDIVIDUAL BEHAVIOR IN AGENT-BASED MODELS

This chapter began with a discussion of why policy makers need empirically grounded models to anticipate the effects of their policies. However, those models are only useful insofar as they accurately capture what outcomes would occur under alternative scenarios. A major factor in evaluating the credibility of a micro-level model is whether or not it has captured the core behaviors of individuals or organizations or other units under investigation. This is especially important when the only data available to understand people's response come from a population in which the focal policy has not yet been implemented or has only been implemented on a small scale. Analysts need some way of making empirically defensible claims about how people might change their responses under different conditions. This section discusses different approaches for specifying individual behavior within simulation models. Although a reasonable specification of behavior may not be sufficient for generating a useful model, it is necessary for valid inferences. The point of structural models is to capture fundamental features of the process under investigation. If individual behavior is misspecified, particularly in an individual-level structural model, it is difficult to see the value in the enterprise.

Quantitative Approaches

One approach to specifying individual behavior is to postulate that agent preferences or behaviors are captured by the parameters of a quantitative model. If the behavior under investigation involves discrete changes in agents' attributes—for example, marriage or childbirth—these transitions can be described using coefficients from a discrete-time event history model (Allison, 1982). If the behavior under investigation implies some sort of decision process (for example, the decision to smoke), discrete choice statistical models provide a useful framework for developing an empirically grounded representation of agents' choice behavior (Ben-Akiva and Lerman, 1985; McFadden, 1974). Historically, discrete choice models have been based on a rational actor model of behavior in which individuals have unlimited computational abilities for performing the calculations necessary for evaluating all possible options.

Discrete choice models have become more behaviorally sophisticated in recent years, drawing on largely experimental work in psychology and decision theory that demonstrates that people have limited time for learning about available options, limited working memory, and limited computational capabilities. These choice models allow for "variation in individuals' knowledge of available options; strategies for learning about or evaluating available options; reactions to change in environmental conditions; reactions to past experiences; and susceptibility to social influence" (Bruch and Atwell, 2013, p. 11). For an example of contemporary discrete choice models that incorporate decision makers' cognitive strategies to reduce the demands of evaluating potential options, see Gilbride and Allenby (2004) and Hauser et al. (2010). However, to the best of the committee's knowledge these models have not been applied to problems outside of marketing, so their value for public health applications remains unknown. Regardless of the choice model used, "estimation of relevant coefficients requires information on either revealed preferences (observed choices) or the stated preferences (survey responses to hypothetical choice scenarios) for some population of interest" (Bruch and Atwell, 2013, pp. 11–12). Surveys, observational data, and administrative records are potential sources for this kind of data.

In recent years, a line of work spearheaded by Brock and Durlauf (Blume et al., 2010; Brock and Durlauf, 2001; Durlauf, 2001) has developed discrete choice models that explicitly model social interactions. In other words, the utility or payoff that a person gets from a particular action depends directly on the characteristics or behavior of others. When the characteristics of other reference group members enter the choice function, this reflects contextual effects, as discussed earlier in this chapter. For example, if the availability of female role models influences women's decisions to major in STEM (science, technology, engineering, and mathematics) fields, the number or proportion of available female role models may be incorporated as a background covariate in the model. Alternatively, these variables may capture endogenous effects whereby individuals' choices are contingent on the choices of others. For example, a teenager's decision to engage in some sort of risky behavior may depend on his or her beliefs regarding how many peers are also engaging in that behavior. A complete technical overview of interaction-based models is beyond the scope of this report, but one point worth noting is that the more that individuals' decisions are influenced by the decisions of others (if this influence is positive, it would imply a conformity effect), the greater the likelihood that the social dynamics implied by the process have multiple possible stable outcomes (i.e., equilibria). See pages 66–67 in Durlauf (2001) for a discussion of this issue.

This framework for capturing interdependent decisions has been applied to studies of peer effects on smoking. For example, Card and Giuliano (2013) use information on friendship ties from the National Longitudinal Study of Adolescent Health (Add Health) to estimate discrete choice models of adolescents' choices concerning smoking, sex, and truancy. The researchers find some evidence of social interactions, especially with regard to peer effects on sexual activity. For example, having a best friend who is sexually active increases the likelihood that one is sexually active by 5 percentage points. Weaker evidence also supports peer effects with regard to smoking, marijuana, and truancy.

Qualitative Approaches

Another strategy for modeling decision making is to specify a heuristic rule based on experimental or theoretical knowledge of the process to be modeled and to assume that agents in the model use that rule to make decisions. Heuristics are "rules of thumb" for making decisions under conditions of uncertainty (Kahneman et al., 1982). Heuristics can be invoked both when gathering information to inform decision making and when evaluating information in the actual decision. Heuristics may be combined with a set of weights that specify the relative desirability of various alternative choices. For example, once a set of choice options has been evaluated, one must decide how to go about choosing among them. One option is to use a "satisficing" heuristic—that is, to assume that people are indifferent among various alternative choices as long as they all satisfy some baseline level of acceptability. In the absence of hard evidence about how people go about making decisions, the decision-making mechanism is yet another assumption that goes into model specification. One fruitful area for future research would be pinning down how real people make decisions.

Data may be used to specify agents' behavior by using ethnographic or participant observations that can provide information on the motivations, strategies, or "rules of thumb" that drive decision making. For example, Hoffer et al. (2009) use ethnographic data to calibrate their ABM of heroin markets. In contrast to statistical specifications of behavior, a qualitative model of behavior is typically formulated as a set of rules governing human action or, alternatively, as a set of rules for interpreting information. One can also combine quantitative and qualitative data on behavior. For example, if experiments reveal a systematic bias in how people perceive their environment, an adjustment could be made to the inputs of a statistical model of behavior.

MODEL UNCERTAINTY AND POLICY DECISION MAKING

As should be obvious from the discussion thus far, models cannot predict the future with certainty. Models can mislead policy makers if modelers present their findings with greater certitude than is warranted. A good model will quantify how uncertainty in the model's inputs translates into uncertainty in what outcomes are most likely under a given policy and will generate a range of predictions that reflect that uncertainty (Manski, 2013;

Wagner et al., 2010). The key issue is separating what is known from what is unknown. Note that this is a very different enterprise from conducting a "parameter sweep" type of sensitivity analysis, which merely provides insight into the workings of the model itself and not into the relationship between the model and the actual world. Uncertainty is only meaningful if the model is anchored in key features of the process under investigation. At a minimum, this might be a simple model that includes an empirically defensible representation of individuals' behavior and interaction.

Once analysts have generated a set of credible model outputs, they must use that information to draw some kind of conclusion about the best course of action. The challenge for the policy maker is to evaluate candidate policy outcomes and weigh the risks and benefits. Thus, to use models effectively to guide policy decisions, the model user needs a rule for translating uncertain predictions into a policy decision.

Conclusion 3-4: The committee concludes that the common exercise of sensitivity analysis does not suffice to measure the uncertainty in model-based forecasts. Sensitivity analysis may provide some insight into the workings of the model itself, but it does not per se assess the potential relationship between model findings and the real world.

Recommendation 3-3: When the U.S. Food and Drug Administration uses the findings of any model, the agency should take into account the uncertainty of findings in order to evaluate policy outcomes and weigh the risks and benefits appropriately.

CONCLUSION

In this chapter, the committee provides an overview of the use of ABMs in policy decision making and explicates how ABM fits into a larger set of modeling approaches. The committee found that ABMs could play an important role in policy decision making and offer useful insights that are not possible with a more aggregated approach. However, to provide meaningful inferences, ABMs must at a minimum include a plausible representation of individual behavior. This may be a fruitful avenue for future research. Moreover, models must provide some account of how uncertainty in model inputs translates into uncertainty in model outputs. To use these models effectively, policy makers will likely need to develop a rule for translating these uncertain predictions into a policy decision.

REFERENCES

Allison, P. D. 1982. Discrete-time methods for the analysis of event histories. *Sociological Methodology* 13(1):61–98.

Aral, S., L. Muchnik, and A. Sundararajan. 2009. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences of the United States of America* 106(51):21544–21549.

Arno, P. S., A. M. Brandt, L. O. Gostin, and J. Morgan. 1996. Tobacco industry strategies to oppose federal regulation. *JAMA* 275(16):1258–1262.

Ben-Akiva, M. E., and S. R. Lerman. 1985. *Discrete choice analysis: Theory and application to travel demand*. Vol. 9. Cambridge, MA: MIT Press.

Blume, L. 2015. Agent-based models for policy analysis. Paper commissioned by the Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation (see Appendix B).

Blume, L. E., W. A. Brock, S. N. Durlauf, and Y. M. Ioannides. 2010. Identification of social interactions. In *Handbook of social economics*, Vol. 1b, edited by J. Benhabib, A. Bisin, and M. Jackson. Amsterdam, Netherlands: North Holland, Elsevier.

Brock, W. A., and S. N. Durlauf. 2001. Discrete choice with social interactions. *The Review of Economic Studies* 68(2):235–260.

Brown, D. G., S. E. Page, R. Riolo, and W. Rand. 2004. Agent-based and analytical modeling to evaluate the effectiveness of greenbelts. *Environmental Modelling & Software* 19(12):1097–1109.

Bruch, E., and J. Atwell. 2013. Agent-based models in empirical social research. *Sociological Methods & Research* 1–36.

Card, D., and L. Giuliano. 2013. Peer effects and multiple equilibria in the risky behavior of friends. *Review of Economics and Statistics* 95(4):1130–1149.

Cartwright, N. 2007. Are RCTs the gold standard? *BioSocieties* 2(1):11–20.

CBO (Congressional Budget Office). 2007. *The effect of tax changes on labor supply in CBO's microsimulation tax model.* Washington, DC: Congressional Budget Office.

Chaloupka, F., and R. Pacula. 1999. The impact of price on youth tobacco use. In *Changing adolescent smoking prevalence: Where it is and why. Smoking and Tobacco Control Monograph 14*. Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health, National Cancer Institute.

Chaloupka, F. J., and K. E. Warner. 2000. The economics of smoking. In *Handbook of health economics edition 1*. Vol. 1, edited by A. J. Culyer and J. P. Newhouse. Amsterdam, Netherlands: North Holland, Elsevier. Pp. 1539–1627.

Christakis N. A., and J. H. Fowler. 2008. The collective dynamics of smoking in a large social network. *The New England Journal of Medicine* 358(21):2249–2258.

Devereux, M. P., and G. Maffini. 2007. *The impact of taxation on the location of capital, firms and profit: A survey of empirical evidence.* Oxford University Centre for Business Taxation. https://ideas.repec.org/p/btx/wpaper/0702.html#biblio (accessed March 18, 2015).

Durlauf, S. N. 2001. A framework for the study of individual behavior and social interactions. *Sociological Methodology* 31(1):47–87.

Epstein, J. M., J. Parker, D. Cummings, and R. A. Hammond. 2008. Coupled contagion dynamics of fear and disease: Mathematical and computational explorations. *PLoS ONE* 3(12):e3955.

Fletcher, J. M. 2010. Social interactions and smoking: Evidence using multiple student cohorts, instrumental variables, and school fixed effects. *Health Economics* 19(4):466–484.

Fletcher, J. M., and S. L. Ross. 2012. *Estimating the effects of friendship networks on health behaviors of adolescents.*Cambridge, MA: National Bureau of Economic Research.

Friedman, M. 1953. *Essays in positive economics*. Chicago, IL: University of Chicago Press.

Gilbride, T. J., and G. M. Allenby. 2004. A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Science* 23(3):391–406.

Grazzini, J., M. G. Richiardi, and L. Sella. 2013. *Analysis of agent-based models*. Laboratorio R. Revelli, Centre for Employment Studies. http://www.laboratoriorevelli.it/_pdf/wp135.pdf (accessed March 18, 2015).

Hammond, R. A. 2015. Considerations and best practices in agent-based modeling to inform policy. Paper commissioned by the Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation (see Appendix A).

Hauser, J. R., O. Toubia, T. Evgeniou, R. Befurt, and D. Dzyabura. 2010. Disjunctions of conjunctions, cognitive simplicity, and consideration sets. *Journal of Marketing Research* 47(3):485–496.

Heckman, J. J. 2008. Econometric causality. *International Statistical Review* 76(1):1–27.

Henriksen, L. 2012. Comprehensive tobacco marketing restrictions: Promotion, packaging, price and place. *Tobacco Control* 21(2):147–153.

HHS (U.S. Department of Health and Human Services). 2010. *How tobacco smoke causes disease: The biology and behavioral basis for smoking-attributable disease: A report of the Surgeon General.* Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention, Office on Smoking and Health.

Hoffer, L. D., G. Bobashev, and R. J. Morris. 2009. Researching a local heroin market as a complex adaptive system. *American Journal of Community Psychology* 44(3-4):273–286.

IOM (Institute of Medicine). 2012. *Scientific standards for studies on modified risk tobacco products*. Washington, DC: The National Academies Press.

Jeon, J., R. Meza, M. Krapcho, L. D. Clarke, J. Byrne, and D. T. Levy. 2012. Chapter 5: Actual and counterfactual smoking prevalence rates in the U.S. population via microsimulation. *Risk Analysis* 32(Suppl 1):S51–S68.

Kahneman, D., P. Slovic, and A. Tversky. 1982. *Judgment under uncertainty: Heuristics and biases*. Cambridge, MA: Cambridge University Press.

Katz, M. H., S. K. Schwarcz, T. A. Kellogg, J. D. Klausner, J. W. Dilley, S. Gibson, and W. McFarland. 2002. Impact of highly active antiretroviral treatment on HIV seroincidence among men who have sex with men: San Francisco. *American Journal of Public Health* 92(3):388–394.

Lightfoot, M., D. Swendeman, M. J. Rotheram-Borus, W. S. Comulada, and R. Weiss. 2005. Risk behaviors of youth living with HIV: Pre-and post-HAART. *American Journal of Health Behavior* 29(2):162–171.

Loomis, B. R., M. C. Farrelly, and N. H. Mann. 2006. The association of retail promotions for cigarettes with the Master Settlement Agreement, tobacco control programmes and cigarette excise taxes. *Tobacco Control* 15(6):458–463.

Macy, M. W., and R. Willer. 2002. From factors to actors: Computational sociology and agent-based modeling. *Annual Review of Sociology* 28(1):143–166.

Manski, C. F. 1993. Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies* 60(3):531–542.

———. 2000. Economic analysis of social interactions. *Journal of Economic Perspectives* 14(3):115–136.

———. 2007. *Identification for prediction and decision*. Cambridge, MA: Harvard University Press.

———. 2013. *Public policy in an uncertain world:Analysis and decisions*. Cambridge, MA: Harvard University Press.

Mare, R. D., and E. E. Bruch. 2003. Spatial inequality, neighborhood mobility, and residential segregation. On-Line Working Paper Series, California Center for Population Research, University of California, Los Angeles. http://escholarship.org/uc/item/8xm4q2q7 (accessed March 18, 2015).

Marschak, J. 1974. Economic measurements for policy and prediction. In *Economic information, decision, and prediction*, Vol. 7–3. Netherlands: Springer. Pp. 293–322.

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. Chapter 4. In *Frontiers in econometrics*, edited by P. Zarembka. New York: New York Academic Press. Pp. 105–142.

Merton, R. K. 1936. The unanticipated consequences of purposive social action. *American Sociological Review* 1(6):894–904.

Moffitt, R. A. 2001. Policy interventions, low-level equilibria, and social interactions. Chapter 3. In *Social dynamics*, edited by S. N. Durlauf and H. P. Young. Washington, DC: Brookings Institution Press. Pp. 45–82.

NCI (National Cancer Institute). 2001. *Risks associated with smoking cigarettes with low machine-measured yields of tar and nicotine.* NIH publication No. 02-5074. Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health, National Cancer Institute.

NRC (National Research Council). 2007. *Models in environmental regulatory decision making*. Washington, DC: The National Academies Press.

———. 2012. *Assessing the reliability of complex models: Mathematical and statistical foundations of verification, validation, and uncertainty quantification*. Washington, DC: The National Academies Press.

———. 2014. *Advancing land change modeling: Opportunities and research requirements*. Washington, DC: The National Academies Press.

Osgood, N. 2007. Using traditional and agent based toolsets for system dynamics: Present tradeoffs and future evolution. Presented at International Conference of the System Dynamics Society, Boston, MA. http://www.systemdynamics.org/conferences/2007/proceed/papers/OSGOO488.pdf (accessed March 18, 2015).

Saez, E., J. Slemrod, and S. H. Giertz. 2012. The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature* 50(1):3–50.

Salop, S., and J. E. Stiglitz. 1982. The theory of sales: A simple model of equilibrium price dispersion with identical agents. *The American Economic Review* 72(5):1121–1130.

Sampson, R. J., C. Winship, and C. Knight. 2013. Overview of: "Translating causal claims: Principles and strategies for policy-relevant criminology." *Criminology & Public Policy* 12(4):585–586.

Sanstad, A. H. 2015. Assessing agent-based models for regulatory applications: Lessons from energy analysis. Paper commissioned by the Committee on the Assessment of Agent-Based Models to Inform Tobacco Product Regulation (see Appendix C).

Schelling, T. 1978. *Micromotives and macrobehavior*. New York: Norton.

Shalizi, C. R., and A. C. Thomas. 2011. Homophily and contagion are generically confounded in observational social network studies. *Sociological Methods & Research* 40(2):211–239.

Smith, A. 1759. *The theory of moral sentiments*. Edited by D. D. Raphael and A. L. Macfied. Oxford, UK: Oxford University Press.

Sterman, J. D. 2006. Learning from evidence in a complex world. *American Journal of Public Health* 96(3):505–514.

van Meijgaard, J., J. E. Fielding, and G. F. Kominski. 2009. Assessing and forecasting population health: Integrating knowledge and beliefs in a comprehensive framework. *Public Health Reports* 124(6):778–789.

Wagner, W., E. Fisher, and P. Pascual. 2010. Misunderstanding models in environmental and public health regulation. *NYU Law Environmental Law Journal* 18:293–356.

Wilde, G. J. S. 2001. *Target risk 2: A new psychology of safety and health: What works? What doesn't? And why…* Toronto, Canada: PDE Publications.