A Political-Economic Analysis of Free-Trade Agreements

By PHILIP I. LEVY *

This paper demonstrates that bilateral free-trade agreements can undermine political support for further multilateral trade liberalization. If a bilateral trade agreement offers disproportionately large gains to key agents in a country, then their reservation utility is raised above the multilateral free-trade level, and a multilateral agreement would be blocked. Bilateral agreements between countries with similar factor endowments are most likely to have this effect. It also follows that bilateral free-trade agreements can never increase political support for multilateral free trade. (JEL F15)

The recent pursuit of bilateral and regional trade agreements, marked most notably by the conclusion of the North American Free Trade Agreement (NAFTA) and the further lowering of trade barriers in Europe, raises questions about the wisdom of this approach to trade liberalization. Governments have asserted that bilateral free-trade negotiations are compatible with the goal of multilateral trade liberalization, but others (e.g., Jagdish Bhagwati, 1992) have questioned whether bilateral arrangements will eventually lead to broader liberalization.

If trade liberalization is to proceed in stages, a formal approach to the process should consider the decision to liberalize at each stage and explore how the decision to liberalize multilaterally is affected by bilateral liberalization. This paper uses a median-voter setting to show a mechanism by which bilateral arrangements may undermine political support for a multilateral arrangement but can never enhance political support for broader free trade.

In the burgeoning literature on preferential trading arrangements and their effects, a number of studies address the impact of bilateral or regional free-trade agreements on trade relations with nonmembers. Unlike this paper, most use the tariff on nonbloc members’ goods as a measure of the agreement’s effects.

Paul Krugman (1991) constructs a model in which the trading world divides symmetrically into blocs. Each bloc—a customs union rather than a free-trade agreement—sets tariffs noncooperatively to take advantage of its market power and move terms of trade in its favor. The larger the individual blocs, the greater is their market power and the higher their tariffs.

Several works address the issues of free-trade agreements (FTAs). Kyle Bagwell and Robert Staiger (1993) focus on the transition period during which an FTA is being formed and posit that the effect of an FTA will be to reduce the volume of trade between the home country and nonparticipants once it is fully implemented. They find that the anticipation of a future drop in multilateral trade volumes interferes with the enforcement of low multilateral tariffs early in the FTA formation process, leading to temporarily higher multilateral tariffs. However, they find that, once an FTA is completed, tariff levels between

---

* Economic Growth Center, Department of Economics, Yale University, P.O. Box 208269, New Haven, CT 06520 (e-mail: levy@econ.yale.edu). I would like to acknowledge the financial support of the Center for Economic Policy Research at Stanford. I also thank the participants in the Stanford Workshop on International Trade, Anne Krueger, Alice Ender, and anonymous referees for their comments. I am especially grateful to Bob Staiger for his guidance. All remaining errors are my own.

1 A recent report of the Council of Economic Advisors (1995 pp. 217–19) considers several arguments for free-trade agreements as “building blocks” or “stumbling blocks” and concludes that they will further multilateral liberalization.

2 Bhagwati (1992) does not formally answer that question but does conclude that many of the arguments in favor of preferential arrangements are of dubious merit. For further general work on regionalism and multilateralism, see the collections edited by Jaime De Melo and Arvind Panagariya (1993) and Kym Anderson and Richard Blackhurst (1993).
the home country and nonparticipants will be no higher.

Martin Richardson (1993) considers the effect of an FTA in a different setting, in which governments maximize a political-support function which gives added weight to export-and import-competing producers' interests. In his model, the small home country has two trading partners supplying a good: the non-FTA partner, with perfectly elastic supply at a low price, and the FTA partner, with perfectly elastic supply at a high price. If an FTA were to make the FTA partner's price less than the tariff-ridden price of the nonpartner, Richardson argues that the home country would lower the tariff against the nonpartner. If it lowered the tariff to the point where the nonpartners's tariff-ridden price was equal to (or just below) the tariff-free price of the partner, the home country would gain tariff revenue on the good and would not harm consumers or import-competing producers (since the domestic post-FTA price would not change).

Panagariya and Ronald Findlay (1994) demonstrate one mechanism by which an FTA could lead to greater protection between blocs. In their model, an exogenously imposed function translates lobbying inputs (labor) into protection. When an FTA is enacted, the labor that was formerly employed lobbying for protection against the FTA partner countries will be released into the labor pool. The wage will be driven down, and thus more labor will be employed lobbying for protection against the rest of the world. Tariffs between blocs should rise.

In Levy (1996), I describe a different mechanism whereby exogenous introduction of a free-trade agreement could induce higher or lower tariffs between trading blocs. It is shown that, in a lobbying framework of the sort originated by Gene Grossman and Elhanan Helpman (1994), FTAs alter pressures for and against trade liberalization by shifting the payoffs to export industries and import-competing industries associated with any given level of protection. The net effect can be to raise or lower barriers between blocs, depending on the characteristics of FTA partners.

The present paper takes a very different approach by portraying national decisions on trade relations as binary choices; countries choose whether to join a free-trade agreement and then choose whether to participate in a broader multilateral agreement.3 In such a setting, there are two readily apparent ways in which bilateral trade agreements could undermine multilateral liberalization: countries could abandon multilateralism in anticipation of future bilateral agreements; or countries could sign bilateral agreements before a multilateral accord is concluded and then lose the desire to pursue multilateralism further. The latter possibility is the topic of this paper.

This possibility is addressed using a political-economy approach similar to that of Wolfgang Mayer (1984), in which a simple majority of voters is required to pass a proposal. Agents are presented first with a potential bilateral free-trade agreement and then with a multilateral free-trade agreement. Each potential agreement offers agents new equilibrium prices and product varieties in a trade model of the sort discussed by Helpman and Krugman (1985). A majority of voters must support a trade agreement for passage.

The sequence of votes is important, in that voters have perfect foresight. They will approve a bilateral agreement only if it is preferable to a multilateral arrangement or if the bilateral agreement will not prevent the adoption of a preferred multilateral agreement.

Agents have different holdings of capital and labor and thus react differently to any given proposal. In the rich version of the model, every trade agreement offers agents an increased number of product varieties, which uniformly enhances the welfare of agents. The shifts in goods and factor prices may be beneficial or detrimental depending on an agent's capital-labor ratio. In this approach the voter with the median capital-labor ratio is of primary importance, since that voter will always be in the majority on any vote.

The primacy of the median voter ensures that no proposal that diminishes the median

---

3 An example would be the United States's choice to participate in NAFTA and then its choice of whether to pursue free trade with the fuller membership of the World Trade Organization. In each case, liberalization proceeds through agreements which are adopted or rejected.
voter's utility can ever pass. This means, for instance, that if a multilateral free-trade proposal is not politically feasible under autarky—because the median voter and thus at least half the populace oppose it—then that same multilateral proposal cannot be rendered feasible by any bilateral free-trade agreement. If a bilateral free-trade agreement is politically feasible, it will only raise the reservation utility level of the median voter to which the multilateral proposal will be compared.

For this reason, the paper focuses on cases in which multilateral free trade is politically feasible in autarky. The most interesting cases are those in which the median agent is roughly indifferent between multilateral free trade and the status quo. For the median agent to be indifferent, multilateral free trade must offer a balance of additional product variety (a gain) and adverse price shifts (a loss).

A bilateral free-trade agreement can undermine support for multilateral free trade by offering the median agent disproportionately large gains with relatively small losses. If such a combination raises the utility of the median voter above the level offered by a multilateral free-trade agreement, then the multilateral agreement will no longer be politically viable. This undermining is more likely to occur in bilateral agreements involving countries with similar capital—labor ratios and roughly indifferent median voters. In the extreme case, a bilateral agreement with an identical partner country would bring variety gains without any price shifts. The remaining variety gains offered by a move from the bilateral agreement to multilateral free trade could be insufficient to compensate the median agent for the factor-price losses, in which case the multilateral accord would be blocked.

The kinds of bilateral agreements that would do the least damage to the political feasibility of multilateral free trade would be those that leave the median voter's utility unchanged by combining price shifts with variety gains. This combination could be found in partner countries with capital—labor ratios different from that of the home country. Such agreements would also necessarily engender the most political opposition of any feasible bilateral agreement, since they do the least to enhance the welfare of swing voters.

To develop the point about undermining, this paper begins with a two-good Heckscher-Ohlin model. In this model there are no variety gains to trade, and it is shown that in this setting voters will never forsake multilateral free trade in favor of a bilateral free-trade agreement. If a majority of voters in one country prefers a bilateral trade agreement with a given partner to multilateral free trade, the majority of voters in the partner country will prefer multilateral free trade to the bilateral accord. This result follows from the strict quasi-convexity of indirect utility as a function of the relative price; if a shift in relative prices increases an agent's utility, a further shift in the same direction will increase utility even more. Bilateral free trade can preclude multilateral free trade only in the trivial case when multilateral free trade would result in the same relative price as bilateral free trade, in which case there would be no incentive to trade on a multilateral basis once the bilateral agreement had been struck.

This result is reversed in the second half of the paper with a specific trade model incorporating increasing returns to scale and product varieties. The introduction of product varieties allows agreements that would not be politically feasible in their absence. Specifically, these are agreements in which the median voter in one partner country would suffer from adverse price shifts but is compensated by increased variety gain. The introduction of product varieties into the welfare analysis also allows agreements between identical countries to raise the reservation utility levels in each. Thus, undermining is possible in the latter model, whereas in the former it was not.

An unusual feature of the analysis is the assumption that tariffs are either zero or prohibitive. While this has the virtue of simplicity, the paper's reasoning would carry through with some different fixed level of tariff protection as the alternative to complete liberalization. The lack of tariffs allows a focus on the choice of trading regime and avoids the difficult question of how external tariffs might be determined simultaneously.

Although specific models are used, the lesson of the paper is more general. When agents are roughly indifferent between an initial situation and multilateral free trade and that
indifference results from a balance of gains (increased variety) and losses (adverse price shifts), then any intermediate agreement offering disproportionately large gains will undermine support for multilateral free trade. To the extent that the opinion of the broader public (or a subset thereof) plays an important role in determining trade policy, this paper illustrates a potentially harmful effect of pursuing regional trade agreements.4

The next section will present the Heckscher-Ohlin model along with the voting procedure. The assertion that bilateral agreements cannot preclude multilateral agreements in such a setting will then be proved. In Section II, a specific example of a model with differentiated products and intraindustry trade will be presented, along with a proof that in this setting, bilateral agreements can undermine multilateral agreements. Conclusions and implications will be presented in Section III.

I. Bilateral Agreements in a Heckscher-Ohlin Model

This section will consider bilateral agreements in a standard two-good, two-factor5 Heckscher-Ohlin trade model. Let there be many countries, distinguishable only by their fixed endowments of the two factors of production, capital (K) and labor (L). These factors are used in the constant-returns-to-scale production of goods X and Y. The internationally identical technologies will be assumed to be such that K is used relatively intensively in X (and L in Y) with no factor-intensity reversals. Perfect competition will ensure that profits are zero.

Agents in these economies own shares of their country’s capital and labor stocks. If the return to a unit of labor is denoted as w and the return to a unit of capital as r, the income of an agent i is:

\[ I_i = wL_i + rK_i, \]

where \( K_i \) and \( L_i \) are the number of units owned by agent i. Agents are assumed to have identical and homothetic preferences. Income is fully spent on goods \( X \) and \( Y \). Arbitrarily, let \( Y \) be the numeraire good and \( p \) the relative price of \( X \) in terms of \( Y \).

It will be assumed throughout this paper that countries’ relative endowments are sufficiently similar that when they join together in a free-trade area, bilateral or multilateral, there is factor-price equalization.6 Within the trading area, the integrated economy that would result from factor mobility will be achieved instead through trade flows. It is also assumed that tariffs are either zero or prohibitive, so a country only trades with its free-trade partners.7

In this setting, when two countries with different capital–labor ratios form a free-trade area, the resulting relative price will lie between the autarky prices in the two countries and the capital-abundant country will export \( X \) and import \( Y \) (the Heckscher-Ohlin theorem).

When two or more countries join to form free-trade areas, the resulting capital–labor

4 Public opinion is likely to be one important mechanism among several in the actual formation of trade policy. Intuitions about other mechanisms, such as interest-group lobbying, are addressed elsewhere in the literature, as described above. For a general standard reference on public choice, see Dennis C. Mueller (1989). For a recent survey of the political economy of trade, see Dani Rodrik (1994).

5 The assumption that there are only two factors of production is important for the median-voter analysis that follows, in that it permits identification of a median voter (the holder of the median capital–labor ratio). With \( n \) factors of production (\( n > 2 \)), it might be impossible to array voters along a single dimension, which would preclude the identification of a median voter. However, even in a more general case, if one were able to identify a key voter (who determined whether or not an agreement would be adopted) the rest of the paper’s analysis would pertain.

6 A sufficient condition for this to hold is that the endowments of all countries lie in the intersection of the cones of diversification of the country with the highest, and the country with the lowest, capital–labor ratio.

7 This assumption is restrictive; participants in free-trade agreements do trade with nonparticipants. If there were some fixed tariff level that applied against all nonparticipants, then the magnitudes of effects would be altered, but the basic arguments of the present paper would not. If, however, trade barriers against nonparticipants were endogenous, this might or might not affect the basic arguments of the paper, depending on the nature of the endogeneity. Such a case is beyond the scope of this paper, but models of endogenous tariff determination and free-trade agreements can be found in Richardson (1993), Panagariya and Findlay (1994), and Levy (1996).
and imports $Y$, so $e_X > 0$. Similarly, when $p < p^A$ it follows that $e_X < 0$. Thus,

\[ \frac{\partial H}{\partial p} < 0 \quad \text{for } p < p^A \]

\[ \frac{\partial H}{\partial p} > 0 \quad \text{for } p > p^A \]

\[ \frac{\partial H}{\partial p} = 0 \quad \text{for } p = p^A. \]

In a Heckscher-Ohlin model without factor-intensity reversals, $p = p(K/L)$ and $p' < 0$, since $p$ is the price of the capital-intensive good, $X$. Thus one can replace $p$ in $H$ with $K/L$, the capital–labor ratio of the broader integrated economy. It follows that $H$ is strictly quasi-convex in $K/L$ with a unique minimum at $K/L = K_j/L_j$, the autarky capital–labor ratio.

Finally, it can be seen that the parallel between a country in an integrated economy and an individual in a broader economy is exact. In both cases, an economic entity with some fraction of the total capital and labor stock uses those endowments to maximize utility, either through international or interpersonal trade. Therefore, there is an indirect utility function $U(K/L, L_i, K_i)$ for any agent $i$ which is strictly quasi-convex in $K/L$ with a minimum at $K/L = L_i/K_i$.

To illustrate the proposition given above, suppose an agent has a capital–labor ratio slightly higher than that of her country. For purposes of the agent’s welfare, one can index all possible free-trade agreements by the continuum of possible capital–labor ratios that would result. As depicted in Figure 1, any trade agreements with countries less capital-abundant than the agent’s country will increase this agent’s welfare by increasing the

---

8 It should be noted that, as in the proof of Proposition 1, the relative price $p$ decreases monotonically with increases in the capital–labor ratio. I omit descriptions of most price changes in the paper because of this immediate correspondence.
return to capital and raising the price of good $X$, which would then be the export good. Such agreements would lie to the left of the point labeled "Autarky" in Figure 1. Trade agreements that increase the integrated economy’s capital–labor ratio will first hurt the agent. Then, as the capital–labor ratio rises above the agent’s own, the agent’s utility will increase. To the right of point E in Figure 1, this agent’s utility rises above its autarky level.

Next, I use this result to consider whether voters might ever opt for a bilateral trade agreement instead of multilateral free trade. The following voting structure will be assumed. Voters will be asked in a first period whether they would prefer a bilateral trade pact to autarky. Then, in a second period, voters are offered the choice between the existing trade regime (either autarky or bilateral free trade) and multilateral free trade. It is assumed that voters are fully informed about all aspects of endowments, economies, and voting. It is also assumed that the periods are close enough together in time or discount rates are sufficiently low that discounting may be ignored. Thus, agents all vote to maximize their expected utility under the final integrated trading economy. Any proposal that garners the support of a majority of voters will be enacted.

Once a free-trade agreement has been approved, the participating countries retain their rights to veto an extension of the agreement to include a new country or countries. Thus free-trade agreements are distinct from political unions.

While each voter has a single vote and the majority will prevail, one can predict the outcome of any vote by considering the behavior of the voter with the median capital–labor ratio. Suppose, for example, that a trade agreement under consideration would lead to a capital–labor ratio in the resulting integrated economy that was lower than the alternative economywide capital–labor ratio (either under autarky or a bilateral agreement). If this increases the utility of the voter with the median capital–labor ratio, one can deduce from Proposition 1 that all voters with higher capital–labor ratios than this median voter would also gain. Together, these agents must constitute a majority, by the definition of median. The same reasoning applies if an agreement would reduce the median voter’s utility. Therefore, agreements will be approved if, and only if, they enhance the welfare of all participating countries’ median voters. I will call such agreements politically feasible.

Now the major result of this section can be stated:

**PROPOSITION 2:** In a Heckscher-Ohlin setting, there can be no politically feasible bilateral agreements that would supplant a politically feasible multilateral trade agreement.

**PROOF:**

It is useful at this stage to enhance the notation. I will use $k$ as a generic notation for a capital–labor ratio. Next, let $k^c_0$ represent the capital–labor ratio of the median agent in country $c$. Finally, I will distinguish between countries and integrated economies; the latter can consist of a single country (autarky) or two or more countries (free-trade areas). The ratio of all the capital to all the labor in an integrated economy $e$ will be written as $k^e$. Now let the function

$$U_c(k^c_0, k^e)$$

denote the maximum attainable utility of the median agent in country $c$, given the capital–labor ratio in economy $e$, as in Proposition 1.

---

9 This is not the only possible vote ordering, of course. There are permutations in which a multilateral pact could be voted on before a bilateral offer or in which voting sequences are repeated. These permutations can affect some of the results of this paper, and I hope to explore them in later work. For this paper, though, the voting structure is assumed to be the one described above.

10 I assume that agents vote their utility whether or not they believe their vote will decide the election, perhaps as a civic duty. This avoids the issue of agents’ expectations about other agents’ voting behavior. While that issue may be of theoretical interest, empirically one observes elections in which the winner wins by a substantial margin.

11 As an empirical basis for this assumption, it should be noted that the EC has moved the furthest toward political union of any existing regional trade group, yet for admission of new members even the EC has relied upon the unanimous approval of its members.
Although the levels of capital and labor endowments determine utility, the ratios are sufficient to explore the welfare effects of policy changes, and thus the levels are omitted.

Consider two countries, A and B, which might pair to create a free-trade area. In the first stage of voting, voters in both countries decide whether to continue functioning in autarky or whether to join to form a free-trade area (AB). In the second stage, voters will determine whether to maintain the outcome of the first stage or join in a multilateral free-trade area (M) including the other countries. Since the proposition is only concerned with politically feasible multilateral free-trade agreements, assume

\[ U_j(k_0^A, k^M) > U_j(k_0^A, k^A) \]

for each country \( j \) (i.e., all median agents prefer multilateral free trade to autarky).

Returning to the potential free-trade agreement, AB, one can see that if \( k^A = k^B \) then there is no basis for trade between the two countries. Therefore, I will arbitrarily say that \( k^A > k^{AB} > k^B \). For agreement AB to pass, both countries must approve it. This requires that

\[ (2) \quad U_A(k_0^A, k^{AB}) > U_A(k_0^A, k^A) \]

\[ U_B(k_0^B, k^{AB}) > U_B(k_0^B, k^B). \]

Because of the strict quasi-convexity of utility functions, if condition (2) holds, this implies that the median voter in A must gain from decreases in the economy’s capital–labor ratio at \( k^{AB} \), while the median voter in B must gain from increases (see Figure 1).

Next consider the second stage of voting. The aim here is to show that there are no sets of capital–labor ratios (for the two countries A and B and the rest of the world) such that both A and B prefer the bilateral free-trade agreement to multilateral free trade. In relation to free-trade area AB, multilateral free trade can have one of three effects:

(i) Multilateral free trade could leave the integrated economy’s capital–labor ratio unchanged (\( k^M = k^{AB} \)). In this case, there is no basis for trade, so both A and B would be indifferent.

(ii) Multilateral free trade could increase the integrated economy’s capital–labor ratio (\( k^M > k^{AB} \)). In this case,

\[ U_B(k_0^B, k^M) > U_B(k_0^B, k^{AB}) \]

\[ > U_B(k_0^B, k^B) \]

so B would approve of the change. It is possible that

\[ U_A(k_0^A, k^M) > U_A(k_0^A, k^{AB}) \]

only if \( k^M \) is sufficiently greater than \( k^A \), in which case both countries would approve the change. If

\[ U_A(k_0^A, k^M) < U_A(k_0^A, k^{AB}) \]

which will occur if A is capital-abundant relative to the world, then A would want to block multilateral free trade. However, country B would foresee the result and vote against the bilateral free-trade agreement in period 1.

(iii) Multilateral free trade could decrease the integrated economy’s capital–labor ratio (\( k^M < k^{AB} \)). This case is simply the reverse of case 2. Country A would approve of the change. If B would want to block multilateral free trade, A would foresee the result and vote against the bilateral free-trade agreement in period 1.

Thus, the strict quasi-convexity of utility functions guarantees that no two countries that originally wanted multilateral free trade can establish a bilateral free-trade area that both prefer.

While the proof given above was tailored for a two-country free-trade agreement, the reasoning can be readily extended to regional trade agreements with two or more countries.

**COROLLARY 1:** In a Heckscher-Ohlin setting, there can be no politically feasible regional agreement that would supplant a politically feasible multilateral trade agreement.
Define a regional agreement (R) as a free-trade agreement among any proper subset of the countries in a multilateral agreement. Denote the capital–labor ratio of the combined countries in the regional agreement as $k^R$. Let country A be any country that is capital-abundant relative to the combined countries in R. Let country B be any country that is relatively labor-abundant. Thus, $k^A > k^B > k^R$. For agreement R to pass, all member countries must approve it. One can now repeat the proof of Proposition 2 with $k^B$ in place of $k^R$.

The bilateral argument is demonstrated in Figure 2. A special case is depicted in which the median voters in countries A and B have the same capital–labor ratios as their countries. These capital–labor ratios are the minima of the two utility curves. The capital–labor ratio of a bilateral free-trade area between A and B must lie between these two points. The three cases in the proof given above correspond to situations when multilateral free trade would result in a capital–labor ratio the same as, to the right of, or to the left of the bilateral FTA point. In either of the latter two cases, at least one country would strictly prefer multilateral free trade to the bilateral agreement.

Of course, if the discount rate were sufficiently high or the periods sufficiently far apart, then it would be possible for a bilateral agreement to undermine multilateral free trade. This is a caveat applicable to almost any sequential result. The point remains that, in a Heckscher-Ohlin setting in which voters are asked to consider both a bilateral and a multilateral trade agreement, there cannot be coalitions of countries that would prefer the bilateral agreement to multilateral free trade.\footnote{If countries were able to dissolve a bilateral agreement, then one could see bilateral agreements formed in cases in which one country preferred the multilateral outcome to the bilateral outcome. There would be no incentive to form such an agreement unless sufficient time passed between voting stages. Of course, in such a setting, it would still be the case that a bilateral agreement could not undermine a multilateral agreement.}

It is worth noting that bilateral agreements are feasible in this model if, and only if, the capital–labor ratios of both participating countries lie on the same side of the multilateral integrated economy’s capital–labor ratio. To see how this might happen, return to the example in which $k^A > k^B$. If both ratios were greater than that of the integrated economy under multilateral free trade, it is possible that country B would gain by both an initial increase in the broader economy’s capital–labor ratio and a subsequent decrease sufficiently large that B became relatively capital-abundant. In Figure 2, this situation would involve a multilateral free-trade capital–labor ratio well to the left of country B’s. The same argument could be made if both countries were labor-abundant relative to the multilateral economy. As shown in Proposition 2, though, these politically feasible bilateral agreements could not preclude multilateral free trade.

One can also readily see that a politically feasible bilateral agreement can never increase the political feasibility of multilateral free trade. If multilateral free trade is not politically feasible initially in a country $j$, $U_j(k_0^j, k^M_j) < U_j(k_0^j, k^f_j)$. If a bilateral free-trade agreement $(ij)$ is politically feasible, then $U_j(k_0^j, k^M_j) > U_j(k_0^j, k^f_j)$. By transitivity, then, $U_j(k_0^j, k^M_j) > U_j(k_0^j, k^f_j)$, so there would not be political
support for an expansion from a bilateral trade agreement to a multilateral trade agreement.

Finally, note that the logic of this section would also apply to political unions. A political union between two countries would imply that majority support from the pooled populations would be necessary for expanding the union. Under such an arrangement, there would likely be a new median voter. However, if the voters in the two partner countries are fully informed, they would block the bilateral political union unless they concurred with the predictable ensuing choice on multilateral union.

II. A Differentiated-Product Model

In this section it will be shown that, in a richer model, the result of the previous section can be overturned; there may be bilateral trade coalitions that can supplant multilateral free trade. To demonstrate this, a standard differentiated-product model will be introduced. This model will retain the important features of the Heckscher-Ohlin model but will add another dimension to agents' utility: the number of varieties of the differentiated product that are available. It is this new dimension that allows bilateral agreements to undermine multilateral agreements.

This section will begin by extending the model of the previous section and deriving the effect of trade agreements on agents' welfare. It will be shown that the median voter is still the agent with the median capital—labor ratio. Finally, it will be shown that trade agreements between similar countries can win voter approval and undermine support for multilateral free trade.

The only differences between countries will be their endowments of the two factors of production—capital \((K)\) and labor \((L)\)—and the distribution of factor ownership. I now add the simplifying assumption that each agent \(i\) will be assumed to own one unit of labor and some amount of capital, \(k_i\), where

\[
\sum_{i=1}^{L} k_i = K.
\]

Since labor ownership is assumed not to vary, \(k_i\) also measures an agent’s relative wealth.

The two factors receive returns per unit of \(w\) for labor and per unit of \(r\) for capital. Thus, agent \(i\) enjoys an income of

\[
I_i = rk_i + w.
\]

I turn now to the two sectors of production and adopt a specific functional form. \(Y\) will denote the economy’s output of the homogeneous product. The constant-returns production process uses factors in the following way:

\[
Y = \gamma Y K_x^\eta L_x^{\xi(1\sim \eta)},
\]

where all parameters are assumed to be positive and \(\mu \in (0, 1)\). \(Y\) will be assumed to be the numeraire good.

The goods in the \(X\) sector are now differentiated products produced under increasing returns to scale. For an individual variety \(x\), the production function is

\[
x = \gamma_x K_x^{\eta} L_x^{\xi(1\sim \eta)},
\]

where parameters again are positive, \(\eta \in (0, 1)\), and the returns to scale are reflected by \(\xi > 1\). In equilibrium, \(n\) will denote the number of varieties in production, and \(X\) will denote the sum of output over all \(n\) varieties in an economy.13

Agents are assumed to have identical utility functions of the following form:

\[
U = U_x^\alpha Y^{1-\alpha},
\]

\[
U_x = \left(\sum_{i=1}^{n} D_i^\beta\right)^{1/\beta}
\]

\[
\beta = \left(1 - \frac{1}{\sigma}\right)\quad \sigma > 1,
\]

where \(D\) represents the consumption of an individual variety \(x\) and \(i\) indexes the variety. The homotheticity of (7) ensures that expenditure patterns will not vary as income is

13 The results do not depend on the Cobb-Douglas form of the production functions. They do depend, however, on the assumption of homotheticity in production.
redistributed or augmented. The parameter $\alpha$ lies in the range $(0, 1)$. The Spence-Dixit-Stiglitz (SDS) subutility function implies that individuals prefer variety and have a constant cross-price elasticity of substitution ($\sigma$) between varieties.\footnote{An alternative assumption described by Helpman and Krugman (1985) is that subutility preferences are of the Lancaster variety. In this case, consumers have an ideal variety and prefer products which are closer to the ideal. Lancaster preferences would not fundamentally alter the results of this paper but would make the analysis more complicated.}

In this framework, each firm in sector $X$ produces at an identical optimal level of output, denoted by $x$, where

$$ x = \frac{\sigma}{\zeta(\sigma - 1)} $$

and sells at an identical equilibrium price, $p$. For a large number of firms, $\sigma$ approximates the elasticity of demand facing each firm.\footnote{For a fuller discussion of the demand structure, see Helpman and Krugman (1985 Ch. 6). Note also that the fixed level of firm production is a by-product of the Spence-Dixit-Stiglitz approach. Under Lancaster’s approach, $x$ would vary.}

The number of firms, $n$, is thus determined by

$$ X = nx. $$

(9)

Returning to the utility of agents, by substituting equilibrium values into the utility function described by (7) and (8), I derive an indirect utility function for an agent $i$:

$$ U_i = I_i (1 - \alpha)^{1 - \alpha} \alpha^n a^\sigma - 1 (1 - \alpha) $$

(10)

Next, I use this result to explore the effect of any free-trade agreement on agents’ utility.

From equation (10) the change in agent $i$’s utility due to a trade agreement can be described as

$$ \frac{U_i^{FT}}{U_i^{AUT}} = \left( \frac{I_i^{FT}}{I_i^{AUT}} \right) \left( \frac{p^{FT}}{p^{AUT}} \right)^{-\alpha} $$

$$ \times \left( \frac{n^{FT}}{n^{AUT}} \right) \frac{a^{\sigma - 1}}{a^{\sigma - 1}} $$

(11)

where FT denotes values after a free-trade agreement and AUT denotes values in autarky. Define $\lambda_K$ as the percentage increase from an economy’s capital stock to the capital stock of the integrated economy resulting from a free-trade agreement (and $\lambda_L$ as the percentage increase in the labor stock). The final term of (11) can be shown to equal

$$ \left( \frac{n^{FT}}{n^{AUT}} \right)^{\alpha / (\sigma - 1)} $$

(12)

$$ = \left[ (1 + \lambda_K) (1 + \lambda_L)^{1 - \sigma} \right] a^{\sigma - 1} $$

Since $\lambda_K$ and $\lambda_L$ are always greater than zero, this “variety” effect always exceeds one (i.e., it has a positive effect on utility). Note that, if an agreement offered only this variety effect, it would raise the welfare of all agents and win unanimous support.

I call the remaining effect on utility the “comparative advantage” effect.\footnote{Alternatively, one could refer to this as the Stolper-Samuelson effect, since this effect of liberalization on real returns to factors is that identified by W. Stolper and Paul Samuelson (1941).} This is the effect that would remain if $\sigma$ went to $\infty$ and the differentiated product became homogeneous. It is a specific example of the utility arguments underlying Section I. To sign to this effect, first parameterize the shift in the economy’s capital–labor ratio by the variable $\varphi$, where $\varphi$ is implicitly defined by

$$ \frac{K^{FT}}{L^{FT}} = \left( \frac{1 + \lambda_K}{1 + \lambda_L} \right) \frac{K^{AUT}}{L^{AUT}} = \varphi \frac{K^{AUT}}{L^{AUT}}. $$

(13)

Note that the free-trade capital–labor ratio is that obtained by pooling the endowments of the partner countries. I assume, as before, that
endowments are such that the integrated equilibrium can be replicated through trade. Therefore, \( \varphi \) will be greater than 1 if a country's counterpart is relatively capital-abundant and less than 1 if the counterpart is relatively labor-abundant.

One can also parameterize agent \( i \)'s capital holdings by \( \rho_i \), where \( \rho_i \) is implicitly defined by

\[
(K/L)_i = k_i = \rho_i \left( \frac{K^{AUT}}{L^{AUT}} \right).
\]

Therefore, \( \rho_i > 1 \) if agent \( i \) is relatively capital-abundant, \( \rho_i < 1 \) if agent \( i \) is relatively labor-abundant, and \( \rho_i = 1 \) if agent \( i \) has the same capital–labor ratio as the country.

With these parameterizations, the comparative-advantage effect is

\[
\left( \frac{I_i^{FT}}{I_i^{AUT}} \right) \left( \frac{p_i^{FT}}{p_i^{AUT}} \right)^{-\alpha} = \varphi^{1/1+\theta} \left( \frac{\rho_i + \theta}{\varphi + \theta} \right)
\]

where\(^{17}\)

\[
\theta = \frac{1}{\mu(1-\alpha) + \eta \alpha} - 1.
\]

Thus, the comparative-advantage effect of an agent's utility change can be seen to depend on the capital abundance of the agent and the capital–labor ratio of the partner country. It will prove useful to work with the natural logarithm of equation (15). Define

\[
f(\varphi, \rho, \theta) = \frac{1}{1 + \theta} \ln \varphi + \ln \left( \frac{\rho}{\varphi + \theta} \right) - \ln(\rho + \theta)
\]

so that \( f > 0 \) implies an increase in utility and \( f < 0 \) implies a decrease in utility due to a free-trade agreement. Then,

\[
\frac{\partial f}{\partial \varphi} = \frac{\theta(\varphi - \rho)}{\varphi^2(1 + \theta)(\varphi + \theta)},
\]

which describes the change in desirability of an agreement as the factor abundance of a partner country changes.

It is now possible to establish the identity of the median voter in this new framework.

**PROPOSITION 3:** The median voter for a country will be the agent with the median capital–labor ratio in that country.

**PROOF:**

Consider a given free-trade agreement characterized by \( \varphi \). Consider the agent with the median capital–labor ratio, whom we can identify as \( \rho_{\text{median}} \). From the log of the utility ratio, \( f(\varphi, \rho_{\text{median}}, \theta) \), one can derive

\[
\frac{\partial f}{\partial \rho} = \frac{\theta(1 - \varphi)}{(\rho + \varphi \theta)(\rho + \theta)}.
\]

The sign of this expression depends only on whether the partner country is relatively capital-abundant (\( \varphi > 1 \)) or labor-abundant (\( \varphi < 1 \)). One can thus say, in terms of the comparative-advantage effect, that, if the partner country is relatively capital-abundant, then all agents with \( \rho > \rho_{\text{median}} \) are worse off than the median agent and all agents with \( \rho < \rho_{\text{median}} \) are strictly better off. Since variety gains affect all voters equally, this demonstrates that the agent with the median capital–labor ratio is the median voter.

If a median voter gains from both the comparative-advantage effect and the variety effect, the arguments from Section I against a bilateral agreement undermining a multilateral agreement will still apply. The interesting counterexample will occur when the median voter is abundant in the same factor as the partner country. Equation (17) indicates that as \( \varphi \) moves from 1 toward \( \rho_{\text{median}} \), the corresponding comparative-advantage effect on the median voter's utility will be negative and declining.
Balanced against this comparative-advantage loss is a gain from the variety effect. In terms of Figure 1, it is now possible for trade agreements to occur in the region between Autarky and E if, and only if, variety gains compensate for the adverse shifts in goods and factor prices.

I now return to the sequential consideration of a free-trade agreement among a subset of countries and a multilateral free-trade agreement involving the entire set and state the following results:

**PROPOSITION 4:** No agreement involving a proper subset of countries can render politically feasible an otherwise infeasible agreement involving the full set of countries.

**PROOF:**
Return to the notation of Proposition 2, in which \( k_0^c \) represents the capital–labor ratio of the median agent in country \( c \) and the ratio of all the capital to all the labor in an integrated economy \( e \) will be written as \( k^e \). Let the possible values of \( e \) be: MFT for multilateral free trade, \( c \) for autarky in country \( c \), and FTA for a free-trade agreement involving a subset of the countries. The utility notation must now also include \( n^e \) to represent the number of varieties in economy \( e \). Now the utility function for an agent in country \( c \) can be written as

\[
U_c(k_0^c, k^c, n^c).
\]

Suppose that

\[
U_c(k_0^c, k^\text{MFT}, n^\text{MFT}) < U_c(k_0^c, k^c, n^c).
\]

This implies that multilateral free trade is politically infeasible. An agreement involving a proper subset of countries is politically feasible if, and only if,

\[
U_c(k_0^c, k^\text{FTA}, n^\text{FTA}) \geq U_c(k_0^c, k^c, n^c).
\]

But that implies

\[
U_c(k_0^c, k^\text{FTA}, n^\text{FTA}) > U_c(k_0^c, k^\text{MFT}, n^\text{MFT}).
\]

Therefore multilateral free trade must remain infeasible.

**PROPOSITION 5:** An agreement involving a proper subset of countries can render politically infeasible an otherwise feasible agreement involving the full set of countries.

**PROOF:**
Using the same notation, suppose that

\[
U_c(k_0^c, k^\text{MFT}, n^\text{MFT}) > U_c(k_0^c, k^c, n^c)
\]

so that multilateral free trade is initially politically feasible. If

\[
U_c(k_0^c, k^\text{FTA}, n^\text{FTA}) > U_c(k_0^c, k^\text{MFT}, n^\text{MFT})
\]

then the subset agreement will have rendered an otherwise politically feasible multilateral agreement infeasible. For an agreement that does this, consider an FTA between country \( c \) and a country that is identical to country \( c \) in every respect. Such an agreement would leave goods and factor prices unchanged.\(^\text{18}\) Thus, the income of any agent \( i \), with capital \( k_i \), remains unchanged by the opening of trade. By equation (10), utility increases with \( n \). All agents in each economy benefit from the liberalization, so under any distribution of capital, agents would choose the free-trade agreement over autarky.\(^\text{19}\) If \( k_0^c \) is such that

\[
U_c(k_0^c, k^\text{MFT}, n^\text{MFT}) - U_c(k_0^c, k^c, n^c)
\]

is sufficiently small (i.e., the variety gain just outweighs the comparative-advantage loss for the median voter), then the variety gains offered by the FTA would render the move to multilateral free trade undesirable (since the comparative-advantage loss would be

\(^{18}\) The result that prices remain unchanged requires homotheticity in production. The fixed-optimal-output result, as stated above, is particular to the Spence-Dixit-Stiglitz subutility assumption. With increasing returns, if the optimal output level increased, it is possible that the relative price of the differentiated product would fall while factor returns remained constant. This could further enhance welfare.

\(^{19}\) This proposition readily extends to an agreement involving \( n \) identical countries, where \( n > 2 \), since a larger group simply implies a larger number of varieties. Thus, the results apply to regional agreements as well as to bilateral agreements.
unaffected). This would necessarily hold for the identical partner country as well.

This result can be generalized beyond the case of countries with equal capital–labor ratios. For any pair of countries, the closer their capital–labor ratios are, the more variety gains they offer in proportion to comparative-advantage effects, the more popular the agreements are likely to be, and the more potentially damaging they are to multilateral free trade.

It is of some interest to explore how a country could avoid undermining support for multilateral liberalization in this setting. One answer, of course, is to pursue only a policy of multilateral liberalization from the start. If bilateral free-trade agreements are to be sought, though, the answer would be to pursue agreements with countries or groups of countries with different factor endowments. A capital-abundant country needs partner countries with less capital and more labor. A free-trade agreement will preserve the feasibility of multilateral free trade so long as it balances variety gains and comparative-advantage effects in such a way as to leave the median voter in each country preferring multilateral free trade to the subset free-trade agreement.

III. Conclusion

This paper has shown that in a Heckscher-Ohlin setting it is politically impossible for a bilateral trade agreement to supplant multilateral free trade. In contrast, it was shown that, in a model with differentiated products and variety gains, bilateral free trade can undermine support for multilateral free trade. Conditions were described under which this could happen. To explain the difference, it is important to understand that the differentiated-product setting permits trade agreements that would have been politically impossible in the setting of Section I by allowing gains through variety gains as well as price shifts. In the Heckscher-Ohlin setting, a voter’s utility depended solely on the capital–labor ratio. For undermining to occur in the differentiated-product setting, the median voter would have to experience lower utility under free trade in the absence of variety gains and higher or equal utility once variety gains are taken into account.

Throughout, the distribution of factors was shown to be crucial. The task of maintaining political support for multilateral free trade when countries negotiate side agreements in a differentiated-product setting was shown to require a clear understanding of the political situation (i.e., factor distribution) in all participating countries as well as a careful selection of those participating countries. The general principle emerged that the more politically popular a bilateral agreement is, the more likely it is to undermine political support for further multilateral liberalization.

The applicability of these results to more intricate models merits further research. Still, the general lesson should remain: intermediate accords can upset the balance of gains and losses offered by multilateralism and can therefore undermine political support. Only when such balances are impossible (as in the Heckscher-Ohlin setting) do these concerns subside.

As stated above, all such difficulties are readily avoidable if countries are restricted to pursuing multilateral liberalization. This is not necessarily an argument against more lenient World Trade Organization rules of the sort endorsed by Alan Deardorff and Robert Stern (1991). The World Trade Organization must accommodate the strong desires of its most powerful members if it is to survive as an institution. Were it to veto major policy initiatives put forward by the United States or the European Community countries, it would be more likely to come apart at the seams than to prevail. Instead, this paper suggests that those powerful member countries might wish to return their attentions to the task of multilateral liberalization.

REFERENCES

Bhagwati, Jagdish. “Regionalism and Multilateralism: An Overview.” Columbia Univer-


Levy, Philip I. “Free Trade Agreements and Inter-Bloc Tariffs.” Mimeo, Yale University, 1996.


