Measuring the Benefits of Unilateral Trade Liberalization Part 2: Dynamic Models

Carlos E. J. M. Zarazaga

This series of two articles has examined the potential gains or losses from unilateral trade liberalization predicted by available general equilibrium models of international trade.

Carlos Zarazaga is a senior economist and executive director of the Center for Latin American Economics at the Federal Reserve Bank of Dallas.

The drive for a multilateral trade agreement encompassing the Americas gained momentum about two years ago, with the U.S. Congress poised to grant the president fast-track authority to negotiate Chile's inclusion in the North American Free Trade Agreement (NAFTA). But the series of severe financial crises that rattled the world almost immediately upon NAFTA's inception frustrated the fast-lane approach and slowed progress toward the agreement. Perhaps this delay reflected that policymakers, businesspeople, and even the general public were reconsidering the benefits of trade agreements with crisis-prone partners.

With the prospect of agreement postponed indefinitely, would countries in the area benefit from lowering, even if unilaterally, their trade barriers? This is the issue addressed in this series of two articles begun in the third quarter 1999 *Economic and Financial Review*.

Part 1 concluded that static applied general equilibrium models could make a mild case for unilateral trade liberalization. However, the article raised the possibility that dynamic models, which incorporate the dimension of time, might do substantially better. That conjecture was partially inspired by numerical experiments with models in which the level of capital after the tariff reduction is changed exogenously (from outside the model).

For example, a static applied general equilibrium model by KPMG Peat Marwick (1991) delivers larger welfare gains when the level of capital in Mexico is increased exogenously to make the rate of return to capital the same both before and after NAFTA. The study starts by assuming that the level of capital is the same before and after the inception of NAFTA. Mexico's gains from NAFTA are negligible in this exercise, the equivalent variation of 0.32 percent of GNP.[1] But the assumption of a constant level of capital implies a higher real rate of return to capital after NAFTA. Because this is an unrealistic outcome under free capital mobility, the study lets capital rise to the level needed to ensure that the rate of return is the same before and after NAFTA. Under this assumption, the welfare gain is equal to 4.6 percent in GNP.[2]

Two qualifying comments are in order. First, the capital increase necessary to make the rate of return the same before and after NAFTA is about 8 percent, which is substantial and, for all we know, has not yet materialized, even six years after NAFTA's inception. Second, this expansion in capital is introduced from outside the model. It is impossible to determine, therefore, whether this new capital level is consistent with the optimal consumption (and, thus, saving) decisions of the agents populating the artificial economy. To answer this question, it is necessary to formulate dynamic models that not only lay out the microeconomic foundations of consumers' and firms' behavior by specifying preferences, endowments, and technology but also incorporate the dimension of time in their decisions. The solution to consumers' and firms' maximization problems will dictate the society's desired level of savings and, therefore, of capital, after trade reform. In other words, in dynamic models the level of capital following a trade reform is determined endogenously—that is, within the model—rather than in some ad hoc fashion from outside of it.

Part 2 of this series investigates the extent to which applied dynamic general equilibrium models deliver on their promise of large welfare gains from unilateral trade liberalization.

APPLIED DYNAMIC GENERAL EQUILIBRIUM MODELS OF INTERNATIONAL TRADE

Applied dynamic general equilibrium models, unlike static ones, can address investment issues because they introduce the dimension of time in consumers' and firms' decisions. As a result, consumers can postpone consumption today and save to be able to consume more tomorrow. Recall that in dynamic general equilibrium models, in contrast to static ones, capital accumulation is determined endogenously rather than exogenously.

Operationally, this difference between the models is most apparent in the utility function and budget constraint used to represent consumers' behavior. Intuition suggests that a simple dynamic version of the static utility function presented in Part 1 could be

$$(1) \qquad Welfare = \sum_t^{\infty} \beta^t \sum_i^N \alpha_i \log c_{it},$$

where c_i denotes real consumption of good i, α_i is a parameter that measures the relative importance the representative consumer attaches to each good, t indexes the time of consumption, and β is the factor by which consumers discount the future, with $0 < \beta < 1$.

This formulation of the utility function conveys the idea that consuming a unit of a good in the future is less attractive than consuming this same unit today. Postponing consumption is costly in terms of utility, and that is why a bundle of goods consumed today yields utility $\sum \alpha_i \log c_i$, while that same bundle consumed tomorrow yields the utility $\beta \sum \alpha_i \log c_i$. (Recall that $\beta < 1$ by assumption.)

The intertemporal dimension of the problem also appears in consumers' budget constraint, which typically takes the form

$$\sum_i^N p_{it} c_{it} \leq y_t + (1 + r_{t-1}) b_{t-1} - b_t$$

for each period t, where y_t is the household's endowment in terms of, say, good 1; b_t is the household's net holding of assets measured also in terms of good 1 (positive if the household is a net creditor, negative if it is a net debtor) at the beginning of period t; and r_{t-1} is the real interest rate consumers receive (if net lenders) or pay (if net borrowers) on their previous-period asset holdings.

Rearranging the equation as

$$b_t \leq y_t + (1 + r_{t-1}) b_{t-1} - \sum_i^N p_{it} c_{it}$$

makes it apparent that b_t is the current period savings: the excess of revenues from all sources, the endowment y_t, and interest payments from assets $(1 + r_{t-1})b_{t-1}$, minus current consumption of goods and services, $\sum_i p_{it} c_{it}$.

The presence of savings in the budget constraint makes clear that, in contrast to the static case, a consumer can now borrow or lend (depending on whether a negative or a positive b_t is chosen) to increase or decrease consumption from the level that current income would otherwise permit.

In a dynamic setting, the consumer's problem is no longer to choose the single consumption bundle c_i, but rather, the whole sequence of consumption bundles $\{c_{i,t}\}_{t=0}^{\infty}$ that maximizes utility. Correspondingly, the consumer will choose the sequence of asset holdings $\{b_{i,t}\}_{t=0}^{\infty}$ consistent with the ability to finance that optimal consumption stream.

Technical Challenges of International Trade Models with Endogenous Capital Accumulation

The addition of the dynamic dimension could potentially increase welfare because the removal of tariffs can prompt a decline in the cost of the imported goods used in the production of domestic capital goods. The corresponding declines in the unit cost of production of capital goods and, therefore, in their prices provide the necessary incentives for a higher rate of investment. The resulting expansion of the capital stock increases labor productivity and, hence, the output of consumer goods and services. But the introduction of the temporal dimension also raises technical complications worth exploring to understand the limitations of the measures of the benefits from trade liberalization reported below.

One of the challenges facing international economists is the problem posed by a constant discount factor, β. The assumption of a constant discount factor is standard in many intertemporal models but is problematic in international applications, particularly for small open economies. This is because models with a constant β usually generate an explosive (implosive) consumption path, in the sense that consumption as a fraction of income constantly increases (declines) over time. Such paths are highly counterfactual, as consumption–income ratios tend to be stable in actual economies.

The reasons for the odd outcome are outlined in the box entitled "Undesirable Implications of the Constant Discount Factor Assumption." Here, it suffices to say the source of the mischief is the combination of a constant discount factor and the small open economy assumption. Under this latter assumption, a small economy is capable of borrowing and lending unlimited amounts at a constant world interest rate. Of course, this assumption is a good approximation to reality only within certain limits. Eventually, if the economy keeps borrowing without bounds, it will absorb all worldwide available savings, at which point the economy will cease to be small and either the interest rate will rise or the country will be unable to continue borrowing.

International economists dealing with dynamic models—that is, models of endogenous capital accumulation—have tried to solve the problems created by the small open economy assumption in several ways. One popular route has been to abandon the assumption, in Equation 1, that the discount factor, β, is constant over time and assume instead that it is a function of consumption.[3] Mathematically, such an assumption is represented as

$$\beta = \beta(c_t),$$

which says the discount factor, β, at any point in time is a function of consumption.

As the box explains, this alternative assumption may give rise to stationary outcomes—that is, to equilibria with constant consumption–output ratios. Applied dynamic general equilibrium models typically assume the function $\beta(c_t)$ decreases in c_t. In other words, as consumption increases, β decreases. This is not an entirely satisfactory specification because there is little evidence that people discount the future more as they become richer.[4]

Goulder and Eichengreen (1992) offer a different solution to the problem posed by the combination of the small open economy assumption and a constant β. Instead of postulating a variable β, they introduce financial assets (wealth) as a determinant of the utility or welfare function. In particular, they propose evaluating the welfare gains from free trade according to the formula

$$(2) \qquad Welfare = \sum_t^{\infty} \beta^t \frac{\sigma}{\sigma - 1}\left[c_s^{\alpha} b_s^{1-\alpha}\right]^{\frac{\sigma-1}{\sigma}},$$

where α and σ are standard parameters in the literature, assumed to have values that ensure the concavity of the utility function, and b_s is a composite of foreign and domestic assets in real terms.[5] This solves the problem of the lack of stationarity when β is constant because assets are treated as just another good and subject to decreasing marginal utility. This will generally ensure a stationary wealth–income ratio and generate a stationary consumption–income ratio as well.[6]

The idea of including financial assets in the welfare function is not new. In fact, many models studying monetary policy issues assume that money, an asset, is a determinant of the utility function. This practice has met with objections because what utility functions such as that in Equation 2 say, if interpreted literally, is that consumers derive pleasure from the mere fact of holding money or bond issues. This is a highly unattractive proposition, as consumers clearly do not derive utility from the pieces of paper but from what they can buy.[7]

Thus, international economists face the difficulty that the assumption of a constant discount factor standard in closed macroeconomic models is unappealing when applied to small open economies, because it tends to produce the counterfactual outcome that consumption as a fraction of income permanently declines or increases. This prediction has typically been eliminated at the cost of counterintuitive preferences, a factor that must be taken into account in evaluating the quantitative results of the applied dynamic general equilibrium models reported in the next section.

Welfare Gains from Trade Liberalization in Applied Dynamic General Equilibrium Models

Progress in quantifying the benefits of free trade with dynamic models has been slow because of the theoretical difficulties discussed above and other computational issues. The few such models available have a mixed record.

Goulder and Eichengreen (1992) find that a U.S. move to unilateral free trade by removing tariffs from an average rate of 4 percent would cut consumption 0.32 percent, which in turn implies a welfare loss equal to 0.44 percent of GDP.[8] The larger welfare gains dynamic models anticipate do not materialize, therefore, in Goulder and Eichengreen's study.

Undesirable Implications of the Constant Discount Factor Assumption

A constant discount factor in combination with time-separable additive utility functions like the one represented in Equation 1 of the text can lead to counterfactual implications for the consumption path of the model economy.

As explained, in a dynamic setting a consumer will typically maximize the utility function

$$\Sigma_{t=0}^{\infty} \beta^t \log c_t$$

subject to the intertemporal budget constraint

$$c_t = y + (1 + r_{t-1})\, b_{t-1} - b_t\,,$$

where all symbols are as in the text, the endowment y is a constant, the real interest sequence $\{r_{t-1}\}_{t=0}^{\infty}$ is exogenous, and there is only one good (i = 1), with its units redefined so its price is 1. The solution to this problem will be the selection of a consumption sequence $\{c_{i,t}\}_{t=0}^{\infty}$ and an asset-holding sequence $\{b_t\}_{t=0}^{\infty}$ consistent with the ability to finance that optimal consumption stream.

Substituting into the utility function the expression for c_t given by the budget constraint yields the maximization problem

$$\begin{aligned} &Max\ \{\beta^t \ln[y + (1 + r_{t-1})\, b_{t-1} - b_t] \\ &+ \beta^{t+1} \ln[y + (1 + r_t)\, b_t - b_{t+1}] \\ &+ \Sigma_{j=0}^{\infty} \beta^{t+j+2} \ln[y + (1 + r_{t+j+1})\, b_{t+j+1} - b_{t+j+2}]\}, \end{aligned}$$

where the choice variable is b_t in period t, b_{t+1} in period t+1, b_{t+2} in period t+2, and so on, and j and t are time indexes.

The first-order necessary condition with respect to b_t corresponding to this maximization problem is

$$-\frac{\beta^t}{y + (1+r_{t-1})b_{t-1} - b_t} + \frac{\beta^{t+1}}{y + (1+r_t)b_t - b_{t+1}} * (1+r_t) = 0.$$

Dividing both sides by β^t and using the budget constraint again, the following equivalent expression results:

$$-\frac{1}{c_t} + \frac{\beta}{c_{t+1}} * (1+r_t) = 0,$$

which, assuming that $r_t = r$, a constant, takes the form

$$c_{t+1} = \beta * (1 + r) * c_t\,.$$

Assuming, for convenience only, that income is constant over time, the above condition can be represented in terms of consumption–income ratios as

$$\text{(B.1)} \qquad \frac{c_{t+1}}{y} = \beta * (1+r) * \frac{c_t}{y}.$$

In a stationary equilibrium, prices and real consumption–income ratios are constant over time. This implies that the above condition in any stationary equilibrium will take the form

$$\frac{c}{y} = \beta * (1+r) * \frac{c}{y}$$

or, equivalently, that $\beta\,(1 + r) = 1$. This condition will typically be satisfied only by chance. More generally, either $\beta\,(1 + r) < 1$ or $\beta\,(1 + r) > 1$. In the first case, Equation 1 implies that consumption in each period will be lower than in the previous one by the factor $\beta\,(1 + r)$. In other words, the consumption–income ratio will decrease monotonically to 0. In the second case, the opposite is true: the consumption–income ratio increases monotonically over time. The problem is that the implication of the dynamic model in either of these cases is grossly counterfactual, because observed consumption–income ratios are very stable over time.

As the text mentions, one possible way out of this problem is to abandon the assumption of a constant discount factor and assume, instead, that it is a decreasing function of consumption. This can be seen intuitively by replacing the function β with $\beta(c_t)$ in Equation B.1:

$$\frac{c_{t+1}}{y} = \beta(c_t) * (1+r) * \frac{c_t}{y}.$$

Suppose that the function $\beta(c_t)$ is decreasing in c_t and that for certain value c of c_t, $\beta(c_t)\,(1 + r) > 1$. This implies $c_{t+1} > c_t = c$, which in turn implies that $\beta(c_{t+1}) < \beta(c_t)$ and, therefore, that $\beta(c_{t+1})\,(1 + r) < \beta(c_t)\,(1 + r)$. In other words, as consumption increases, $\beta(c_{t+1})$ *decreases*, and so does $\beta(c_{t+1})(1 + r)$ until eventually $\beta(c_{t+n})(1 + r) = 1$ for n large enough. At that point consumption becomes stationary (in the sense that it repeats itself over time) because

$$\frac{c_{t+n+1}}{y} = \frac{c_{t+n}}{y}.$$

However, this way of solving the lack of stationary equilibria in dynamic models with constant discount factor β is somewhat of a mechanical quick fix. Typically, any function $\beta(c_t)$ will be continuous and, therefore, decreasing for some values of c_t (or eventually all of them, as in the example above). At the same time, in any reasonable economic model, consumers want to consume more the wealthier they are. This implies that when $\beta(c_t)$ is decreasing in consumption, it is also decreasing in wealth or, equivalently, that households become more impatient to consume as they get wealthier. Unfortunately, there is no empirical evidence to support this rather ad hoc assumption. The opposite and equally arbitrary assumption that $\beta(c_t)$ is increasing in c_t—that is, that a household's desire to accumulate wealth rises as it becomes richer—cannot be empirically validated either (and introduces the additional technical difficulties mentioned in footnote 4 of the text).

Those gains do seem to materialize—at least for developing countries—in a recent model by Ahearne (1999). It is one of the few dynamic models to quantitatively analyze unilateral trade liberalization's effects for developing countries. Unfortunately, Ahearne focuses on the performance of macroeconomic variables such as aggregate output, consumption, and investment and does not report a measure of welfare, such as that Goulder and Eichengreen report. This omission makes welfare comparison of the two studies difficult. In any case, to the extent the direction of change of consumption and welfare are the same (as they are in Goulder and Eichengreen), Ahearne's outcomes are more favorable to trade reform. He finds that lowering tariffs from an average of 25 percent to 10 percent would result in an increase in consumption of about 3 percent. A reduction to an average rate of 5 percent would raise consumption growth to about 4.5 percent, while the complete removal of tariffs would result in a 6 percent consumption increase.

Six percent consumption growth is by no means negligible and could be seen as an indication that dynamic models can, after all, deliver larger welfare gains from unilateral trade liberalization than their static counterparts. However, it is important to emphasize that the relatively large consumption growth of 6 percent is obtained from removing tariffs originally assumed to be 25 percent. Ahearne's study suggests that consumption growth will be a more moderate 1 percent to 2 percent if the average initial tariff is 4 percent, as in Goulder-Eichengreen.[9] Still, this increase in consumption after trade liberalization seems to reverse Goulder and Eichengreen's negative finding.

Unfortunately, it is difficult to pinpoint the source of the opposite results of these two dynamic models because their features are quite different, from the specification of the utility function to the underlying assumptions about capital mobility.[10] For example, Ahearne's assumption is that the discount factor depends on the level of wealth (or equivalently, consumption), while Goulder and Eichengreen assume that assets enter into the utility function.

Another important difference is that Ahearne assumes the terms of trade are exogenous. Thus, changes in tariffs alter the relative domestic prices but do not change the international terms of trade against the country that liberalizes. This is not the case in Goulder and Eichengreen because they adopt the so-called Armington, or national product differentiation, assumption. As Part 1 explained, this assumption is introduced mainly to account for the puzzling "cross-hauling" in which many countries appear to export the same goods they import.[11] National product differentiation circumvents this problem by assuming each country is the only producer of the good it exports. However, this also means tariffs could help a country exploit its market power. Tariff elimination might be damaging in this case because the optimal tariff typically is not zero under this assumption. This fundamental bias against free trade is absent from Ahearne's model but seems to prevail in Goulder and Eichengreen's.[12]

There are reasons to doubt the welfare losses Goulder and Eichengreen's model delivers because their preferences include assets as a determinant of the welfare function. The resulting welfare measure may reflect consumption changes as much as changes in asset holdings. This is certainly an unappealing way to measure welfare, in light of the general equilibrium theory standard that consumers do not derive utility directly from merely holding assets but from the stream of goods and services those assets can purchase.

Ahearne's study may exaggerate the GDP growth from unilateral trade liberalization because he assumes perfect capital mobility. This may not be the case in practice, as evidence suggests that households tend to invest their savings in their home country rather than in foreign ones. Goulder and Eichengreen capture more aptly this reality by assuming that consumers have a bias for domestic assets, and this implicitly limits the capital mobility responsible for the relatively large GDP and consumption gains in Ahearne's model.[13]

PRODUCT VARIETY AND GAINS FROM TRADE LIBERALIZATION

It was argued that dynamic models of international trade have the potential to deliver the large welfare gains from trade liberalization that their static counterparts have failed to produce. The preceding section suggests that dynamic models cannot fulfill those expectations either, except under nonconventional representation of consumers' preferences.

However, one often-heard criticism of all the models discussed so far is that they fail to incorporate the idea that free trade makes possible access to new technologies that enhance the economy's overall productivity. Perhaps this is why dynamic general equilibrium models produce only less-than-striking welfare gains from unilateral trade liberalization.

Consider, for example, the constant-returns-to-scale production function presented in Part 1 of this series:

$$Y = A\,L^{1-\alpha}K^{\alpha}, \tag{3}$$

where Y is aggregate output, L the amount of labor input, K the amount of capital input, and $0 < \alpha < 1$.

In this specification, the total factor productivity, represented by A, is treated as a given parameter, invariant to the trade regime. Therefore, this equation does not capture the idea that trade liberalization will increase an economy's overall productivity. Trade liberalization can raise production only if it leads to the use of more labor or capital inputs.[14] Otherwise, the same amount of labor and capital will produce exactly the same amount of output.

A similar situation arises with the second kind of technology, the increasing-returns-to-scale technology—or, equivalently, decreasing-average-cost technology—considered in Part 1:

$$\textit{Total cost} = F + bQ. \tag{4}$$

Again, notice that tariff policy changes can reduce average costs only if they induce an increase in the quantity of the good. But the basic cost structure, defined by fixed cost F or marginal cost b, is the same regardless of the tariff regime under which countries operate.

This invariance of the overall productivity to the trade regime implicit in conventional production functions has been challenged on several grounds. For instance, an important benefit of international trade is that it gives consumers more choices and offers producers more options in terms of inputs. The advantage of variety is particularly important for economies that can produce only a limited range of goods on their own. This is the case with economies characterized by cost functions such as the one in Equation 4—that is, economies with increasing returns to scale.

Recall from the discussion in Part 1 that in such economies each good is produced by one and only one firm. The number of varieties is determined by the number of firms, which is limited when there are increasing returns to scale. To see this, suppose all firms must pay the fixed cost F in terms of a primary input z (for example, land) and that each economy is endowed with Z units of that good. Each economy on its own will be able to produce, at most, Z/F varieties of goods (for simplicity, we assume Z/F is an integer). The number of goods produced domestically will be limited by that upper bound.[15]

Thus, suppose prohibitively high tariffs or restrictions limiting quantities make trade between two economies disappear. This implies that consumers must make do with domestic goods. Although a consumer would like that 27-inch-screen TV and can afford it, he will have to settle for the smaller domestic model without remote control. Likewise, local producers will have to adjust their technologies to the intermediate inputs domestic firms make available. A construction company may prefer a special kind of foreign-made insulation for a building that will have to withstand extreme temperatures, but the firm will have to use a more expensive and less functional building design to achieve the same results with the less suitable insulators produced domestically.

Next, suppose all barriers to international trade are lifted. Firms in this economy will be able to use both domestic and foreign inputs. The examples above suggest that a larger variety of goods, especially of intermediate inputs, may be associated with aggregate productivity gains not appropriately captured by conventional production functions.

To confront this limitation, economists have started to play with less conventional production functions that incorporate the idea of productivity gains from variety. Such production functions can be constructed by a clever reinterpretation of the conventional constant-returns-to-scale production functions.

For simplicity, assume only one final consumption good is produced with the technology

$$Y = L^{1-\alpha}\left(x_1^{\alpha} + x_2^{\alpha} + \ldots + x_M^{\alpha}\right),$$

where L represents the amount of labor input, x_i represents the quantity of an intermediate input i, $i = 1, 2, \ldots, M$, and $0 < \alpha < 1$. Assuming each intermediate input is used in the same quantity,[16] the technology can be rewritten as

$$Y = M\,L^{1-\alpha}x^{\alpha}. \tag{5}$$

This appears to be the same old constant-returns-to-scale technology of Equation 3, with capital, K, relabeled x and the total factor productivity, A, relabeled M. Indeed, increasing both the amount of labor input L and the typical intermediate input x by h percent would raise production of the consumption good by h percent, which is exactly what is supposed to occur with a constant-returns-to-scale technology.

The trick is that relabeling A as M is not as innocuous as it might appear because now A is not necessarily fixed. In fact, A (or M)—the number of varieties—can be regarded as an input, just as L or each x_i is. In other words, according to this production function, aggregate production of final goods Y requires combining three inputs: the number of varieties of intermediate inputs (M), the quantity used of each of them (x), and labor (L).

The reinterpretation of A as the number of intermediate-input varieties represents mathematically the old idea that one-size-fits-all economies will be less productive than highly specialized ones. The intuition is that access to a larger variety of goods will make it more likely that producers will find inputs that better fit the characteristics of their production lines and that consumers will find the products that best fit their tastes and needs.[17]

The gains-from-variety effect can be better understood by comparing the nonconventional production function in Equation 5 with

$$Y = L^{1-\alpha} * M * x.$$

According to this production function, doubling the varieties of intermediate-input goods will have the same effect on output as doubling the amount of each of those inputs, as can be seen from the equalities

$$L^{1-\alpha} * (2M) * x = 2L^{1-\alpha} * M * x = 2Y$$
$$= L^{1-\alpha} * M * (2x).$$

This is not the case with the proposed production function of Equation 5, in which doubling the number of varieties M doubles output, but doubling the quantity of each intermediate input x increases output only by a factor of 2^{α}, which is lower than 2 (recall that $0 < \alpha < 1$).[18]

In other words, in Equation 5 any increase in variety has a larger impact on aggregate production than an identical percentage increase in the quantities of the existing intermediate-input varieties. Loosely speaking, this production function captures the idea that a society cannot easily compensate for the loss of variety with more of the same old stuff. This issue is relevant to measuring the gains from unilateral tariff removal because freer trade policies (even if implemented unilaterally) may give a country access to a larger variety of goods. The welfare gains from such policies may be important if a larger variety of intermediate inputs, as the production function suggests, increases the economy's productivity in the manufacturing of domestic goods.

In the reinterpretation (*Equation 5*) of the conventional constant-returns-to-scale technology (*Equation 3*), the total factor productivity parameter A in the latter would be equal to M before trade liberalization but eventually equal to $(M + \Delta M)$ after trade liberalization, where ΔM represents the additional varieties of intermediate inputs resulting from freer trade.

With a production function like that in Equation 5, the gains from freer trade will come from two sources: the traditional one that tariff reductions will make imported intermediate goods cheaper and thus induce higher output levels of the existing varieties of final goods and services, and the nonconventional one of gains from variety from ΔM. This second effect is a good candidate for boosting the welfare gains from unilateral trade liberalization beyond the negligible to moderate results found by models using more conventional production functions. The remainder of this article reports the results of recent work that has exploited this gains-from-variety approach to build a better case for unilateral free trade.

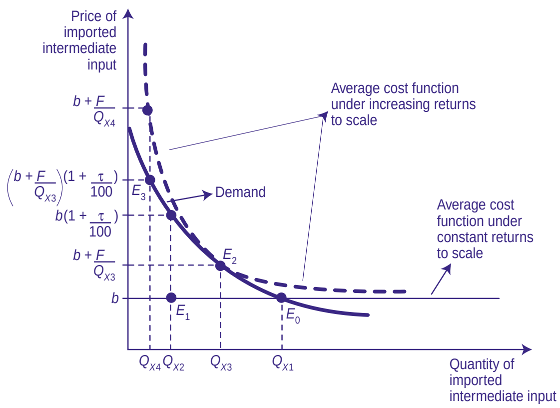
Measuring the Welfare Gains from Variety

In analyzing these studies on welfare gains from product variety, it is important to understand how tariffs reduce the product varieties available to firms and consumers. All the studies discussed below assume that firms face cost functions of the form in Equation 4. Equivalently, they assume all goods are produced with increasing-returns-to-scale technologies.[19] As Part 1 explains, this is the only technology consistent with the national product differentiation assumption, typically introduced to account for the cross-hauling puzzle in trade statistics.

The introduction of increasing-returns-to-scale technologies is not inconsequential for the potential gains from variety with a tariff reduction. The attrition effect tariffs can have on variety starts at much lower tariffs with increasing-returns-to-scale technologies than with constant-returns-to-scale technologies.

Consider a typical final-goods producer's demand for an imported intermediate input produced with a constant-returns-to-scale technology, as represented in Figure 1. Recall that the cost function will look like Equation 4, with F equal to 0, implying a constant marginal and unit average cost of b. Suppose that initially there is no tariff, and the equilibrium demand of the input is at point E_0, with price b equal to the unit and marginal cost and quantity Q_{x1}. Next, suppose an ad valorem tariff of τ percent is imposed on this intermediate input. For the sake of argument, assume the buyers absorb all the burden of the tariff—that is, the foreign producers of those inputs still receive a price b (equal to their unit cost of production) for each unit of the intermediate good they sell to domestic buyers. These buyers will have to pay a price $b(1 + \tau/100)$ for the imported intermediate input. Suppose that at the new equilibrium

Figure 1
The Effect of Tariffs on Product Variety



price the quantity demanded, Q_{x2}, is one-fourth Q_{x1}, the equilibrium quantity before the tariff was introduced. Foreign producers will have to cut the quantity produced by three-fourths. The question is whether they can stay in business doing that. The answer under constant returns to scale is an unambiguous yes because producers will always receive the price b for each unit, regardless of sales level. Since b is also the unit cost of production, they will cover their costs and be able to stay in business whether they sell Q_{x1} units (as in point E_0) or Q_{x2} (as in point E_1). Thus, under constant returns to scale, the available varieties of intermediate inputs will be the same before and after the introduction of the tariff. The only effect of a tariff is that each intermediate input will now be produced in a smaller amount, to match the decline in the quantity demanded as a result of the tariff-induced price increase. Thus, moderate tariffs will tend not to have any visible consequences for product variety under constant-returns-to-scale technologies.[20]

Consider the alternative case in which intermediate inputs are produced with an increasing-returns technology, whose cost function will look like Equation 4. The unit cost of producing Q_{x1} is greater than b by F/Q_{x1}. Therefore, b cannot be an equilibrium price because foreign producers of the imported inputs would suffer a loss. Assume, then, that the equilibrium price for the imported intermediate inputs under increasing-returns-to-scale technology is $b + F/Q_{x3}$. The higher price implies, of course, that the equilibrium quantity, Q_{x3}, will be to the left of the equilibrium quantity under constant returns to scale, Q_{x1}.

Now suppose the same tariff of τ percent is levied on all imported intermediate inputs and the tariff is borne entirely by the domestic buyers of imported intermediate goods. As a result, the price increases from $b + F/Q_{x3}$ to $(1 + \tau/100)(b + F/Q_{x3})$ and the quantity demanded drops to Q_{x4}. It appears, then, that the pair $(b + F/Q_{x3}, Q_{x4})$, represented at point E_3 of the demand curve in Figure 1, is a good candidate for the new equilibrium. But the appearance of two different quantities, Q_{x3} and Q_{x4}, in this pair suggests something is wrong with that conjecture. Indeed, unlike in the constant-returns-to-scale case, producers of the intermediate good experience an increase in the average unit production cost by cutting production of the intermediate input from Q_{x3} to Q_{x4}. In fact, the unit costs will be $b + F/Q_{x4}$, higher than the $b + F/Q_{x3}$ per unit they will receive from the price inclusive of tariff $(1+ \tau/100)(b + F/Q_{x3})$. In other words, at the price $(1+ \tau/100)(b + F/Q_{x3})$ producers of the imported intermediate input will suffer a loss. Therefore, they will have to increase the price (before tariff), say, to $b + F/Q_{x4}$. But this higher producer price will result in a higher user price of $(1+ \tau/100)(b + F/Q_{x4})$, which in turn will reduce demand for the intermediate input even further. This will result in a higher unit cost to produce the imported intermediate input and lead to another round of increases in the domestic price of those imports. Eventually, unit costs will keep rising at a higher rate than the price. This shows up in Figure 1 in the fact that for prices above $b + F/Q_{x3}$, the demand curve stays always to the left of and below the unit average cost curve.[21] This implies producers will always suffer a loss if they cut production below the pretariff level Q_{x3}. Since the tariff reduces demand below that level of production, the producer of the intermediate input will be forced out of business and that input variety will disappear from the market.

Thus, in contrast to the constant-returns-to-scale case, the imposition of even a moderate tariff in the presence of increasing returns to scale may reduce M, the number of intermediate-input varieties available to final-goods producers. This is because the tariff has a "market-size" effect on the intermediate input's unit production cost that was absent in the constant-returns-to-scale technology case.

The next section discusses how these general ideas have been implemented in recent studies that attempt to take into account productivity gains from variety eventually introduced by policies of unilateral trade liberalization.[22]

Welfare Gains from Variety in Static Models of International Trade

I report first a recent study by Klenow and Rodriguez-Clare (1997) because, strictly speaking, their model is static and thus belongs to the class discussed in Part 1 of this series. However, I deferred discussion of this work until now because it is one of the few available studies that explicitly considers the potential gains from variety when measuring the benefits of unilateral trade liberalization.[23]

The Klenow and Rodriguez-Clare model incorporates the product variety effect by assuming that importing firms must pay a fixed cost F to operate and a constant price b for each unit of imported good. For all practical purposes, it is as if the imported goods are produced according to a cost function like that in Equation 4.

Tariffs will make imported goods more expensive and, hence, reduce the demand for them. By the mechanism explained earlier, this smaller market size will eventually leave some importing firms unable to cover their fixed costs, forcing them to shut down. Consumers then suffer because they can no longer find the varieties of goods they had been purchasing. Likewise, local producers will become less productive, as Klenow and Rodriguez-Clare assume a production function of the type in Equation 5 (although a much more complicated one), by which a lower M (number of intermediate-input varieties) results in productivity losses and, therefore, in lower output despite the same capital and labor inputs.

The authors quantify the model using data for Costa Rica and find that removal of a 10 percent tariff can quadruple the gains from unilateral trade liberalization compared with a model in which product variety effects are absent. In particular, they find that imposition of a 10 percent tariff on intermediate goods leads to welfare losses equal to about 2 percent of GDP, as opposed to only 0.5 percent in lost GDP when the variety effect is not taken into account.

Thus, incorporation of gains from variety works in the expected direction of increasing the welfare gains from trade liberalization but keeps them within the moderate ranges of the static models without the gains-from-variety effect, as reported in Part 1.

One possible reason for the moderate gains in the Klenow and Rodriguez-Clare model is that the national product differentiation assumption works against unilateral trade liberalization, as explained above. The model's static nature also could be a factor. Thus, the next step is to see if these limitations are overcome by dynamic models—those that incorporate the dimension of time and, hence, saving and investment decisions—in the context of gains-from-variety effects.

Welfare Gains from Variety in Dynamic Models of International Trade

Quantitative dynamic models measuring the effects of unilateral trade liberalization do not abound. Even fewer of them have tackled the gains-from-variety effect. One such model is a study on Austria by Keuschnigg and Kohler (1996).

As explained earlier, all models of international trade that consider tariffs' effects on product variety must introduce, in one way or another, fixed costs of production. In the case of Keuschnigg and Kohler, it is the local producers of final domestic goods (and not importing firms, as in Klenow and Rodriguez-Clare) that face a fixed production cost. This assumption is the same as in the Klenow and Rodriguez-Clare model, except that tariffs will not change the number of foreign varieties but of domestic intermediate-input varieties supplied to local producers.

Because of this fixed cost, and for the same reasons as in the static models, tariff reductions in Keuschnigg and Kohler increase the market size for every good, eventually making it profitable to import or produce varieties unavailable before. In addition to this static effect, Keuschnigg and Kohler introduce a dynamic one by linking the stock of capital with the number of product varieties.

The intuition behind this additional effect is similar to the one given above when describing why, in the presence of a fixed cost, a fixed factor like land may limit the number of product varieties an economy can produce. The same logic works here, replacing land with capital. Suppose each firm in the economy must pay the fixed cost Keuschnigg and Kohler assume in the form of k units of capital. A given capital stock K could support at most K/k product lines or varieties. Since K is implicitly assumed fixed in static models, any increase in varieties must come through reduction in production costs rather than expansions in the capital stock.

But since Keuschnigg and Kohler allow for investment, the capital stock is not fixed. In fact, reductions of tariffs on intermediate and capital goods can induce a process of capital accumulation for the reasons discussed above.[24] If the capital stock increases by ΔK as a consequence of a unilateral move to trade liberalization, the economy can eventually support the higher number of product varieties $(K + \Delta K)/k$. This capital accumulation effect induces gains from variety in Keuschnigg and Kohler beyond those induced by the market-size effect present in static models described earlier.

Keuschnigg and Kohler calibrate their model to Austrian data and find that removing a 10 percent average tariff would result in welfare gains equal to about a 4 percent increase in GNP if the scale economies are fairly large (or if fixed costs are fairly big).[25] This is more than two times the gains in the Klenow and Rodriguez-Clare static model.

The larger welfare gains in Keuschnigg and Kohler again demonstrate that omitting the time dimension and capital accumulation may lead to a fairly sizable underestimation of the benefits of unilateral trade liberalization.

However, like many models reported above, Keuschnigg and Kohler's does not produce striking welfare gains. The potential for positive terms-of-trade effects from the national product differentiation assumption may be responsible for this. In fact, Keuschnigg and Kohler report that with a milder terms-of-trade effect, the gains from removing a 10 percent average tariff could be as large as 7 percent of GNP.

It is important to remember that the gains-from-variety effect may be a dangerous concept to play with. However beneficial to the case for unilateral trade liberalization, it may paradoxically undermine the very case for free market policies that it is meant to buttress. This literature typically appeals to increasing returns to scale, and, in the presence of such technology, markets cannot achieve the social optimum without government intervention.[26]

Dynamic models that incorporate gains-from-variety effects seem to have more potential for delivering nonnegligible welfare gains. At the same time, these models include significant increasing-returns-to-scale technologies, a somewhat problematic feature because it opens the door to government intervention and may undermine the case for free market policies that the gains-from-variety effect is meant to boost.

A clear message from the quantitative experiments these two articles report is that neither the introduction of time nor of product variety effects can completely overpower the strong force against unilateral removal of tariffs introduced in almost all models by the national product differentiation assumption.

The strength of such a force is suspect, especially in models that assume monopolistic competition. As explained in Part 1, that assumption puts the market power at the firm—rather than at the country—level, which, in principle, should weaken the case for trade barriers introduced by the national product differentiation assumption. Perhaps more weight should be given to the models discussed in this article that mitigate the country market power effect of national product differentiation. These models deliver moderate to sizable welfare gains from unilateral trade liberalization.

Thus, the advantages of unilateral trade liberalization are cause for optimism. But to the extent that country market power is perceived as important in evaluating alternative trade policies, countries may balk at the prospect of a unilateral free trade policy. Hence, the suspension or slow progress of negotiations for multilateral trade agreements is cause for concern.

The measures of welfare gains (or costs) from unilateral trade liberalization policies presented in this article should be considered with caution for several reasons. First, dynamic models generally pose formidable technical challenges that—in the particular applications discussed in this article—have been circumvented with not completely appealing shortcuts. Second, the theoretical foundations of the national product differentiation assumption and the bias it introduces against unilateral trade liberalization remain controversial, especially in models that assume a monopolistically competitive market structure. Furthermore, the empirical measures of the market power implicit in such an assumption are imprecise. Certainly, international trade researchers still have a lot of work to do before the benefits of unilateral trade liberalization policies can be confidently assessed.

CONCLUSION

Free trade advocates consider the denial of fast-track authority to the U.S. president a worrisome development. The concern is that lack of interest in multilateral trade agreements will create a backlash against the free trade policies Latin American countries adopted in the 1990s. The fear is warranted if each country in the region perceives it will experience welfare losses from adopting free trade policies when some if its major trading partners do not.

This series of two articles has examined the potential gains or losses from unilateral trade liberalization predicted by general equilibrium models of international trade. Negligible to moderate gains are found in static as well as dynamic models that do not incorporate gains from product variety. The results confirm that the omission of the dimension of time and, hence, of capital accumulation can undermine the case for unilateral trade liberalization.

NOTES

The author wishes to thank David Gould, Gregory Huffman, Evan Koenig, and, especially, Steve Brown for comments on earlier drafts that substantially improved the contents and organization of the ideas in both articles of this series. Any remaining errors are, of course, mine.

1 As Part 1 explains, the equivalent variation in income is the change in income that consumers should experience without a trade liberalization to replicate the level of utility they would attain with it. A negative equivalent variation in income implies that consumers are worse off after trade liberalization.

2 Although this model is frequently put in the applied general equilibrium category, it does not strictly belong there because it implicitly assumes an excess supply of labor (that is, the labor market is in disequilibrium) in Mexico, at least before NAFTA.

3 One of the first authors to implement this approach in the context of a small open economy was Mendoza (1991).

4 The alternative assumption that β is an increasing function of c_t is not less problematic, as it creates the possibility of multiple equilibria, an issue beyond the scope of this article.

5 As explained in Part 1, concavity is a mathematical property that captures the idea that consumers have a taste for variety.

6 The counterpart for the lack of stationarity of the consumption–income ratio with a constant β is the lack of stationarity of the wealth–income ratio.

7 The introduction of money in the utility function has been justified as a shortcut to capture the notion that money facilitates trade. In fact, Feenstra (1986) shows that under certain conditions, transaction costs in trade will operate as if money were an argument of the utility function. However, the same reasoning applies if transaction costs are associated with buying and selling bonds or equities. It is on these grounds that Poterba and Rotemberg (1987), for example, include short-term government debt (and not just fiat money) in the utility function. This shortcut to modeling transaction costs explicitly may be useful for addressing certain monetary policy questions, but its application to the issue of the welfare consequences of alternative trade policies is more controversial.

8 For the reasons given in Part 1, consumption growth may not be a good measure of well-being, particularly in models in which labor supply is endogenous. For example, households may consume more after trade liberalization but also work harder, so the welfare gains may not be nearly as large as the increase in consumption would otherwise suggest. That's why most applied studies of international trade, like Goulder and Eichengreen's, report the *equivalent* variation in income rather than the *actual* variation in income (or GDP).

9 I arrived at this figure by assuming that the change in consumption from a removal of tariffs will be a linear function of the original average tariff rate for tariffs in the 0 percent to 10 percent range.

10 Both models assume several sectors, but the details of the disaggregation and technologies in each of them differ. Both models introduce frictions in the investment process but differ in the details. Goulder and Eichengreen assume that changing the capital stock from its current level is costly in terms of resources, while Ahearne assumes it is costly in terms of time—that is, that it takes several periods to bring the capital stock to the desired level.

11 As Part 1 discusses, Japan may appear to import and export cars simply because of the way trade statistics are reported. For example, Japan could be importing convertibles and exporting vans. Although these are different products, they might appear simply as "cars" in the broad categories used in trade statistics, giving rise to an apparent cross-hauling puzzle.

12 Unfortunately, none of these authors report an optimal tariff for their models. One conjecture worth exploring is that the 4 percent initial tariff Goulder and Eichengreen assume in their benchmark case is much closer to the optimal tariff than the 25 percent rate Ahearne assumes. Obviously, removing an optimal tariff will cause welfare losses while removing a nonoptimal one might enhance welfare.

13 In fairness, Ahearne himself reports that GDP gains in his model are more moderate under limited international capital mobility.

14 As explained in the previous section, trade liberalization can result in a higher capital stock, which implies a higher use of capital input in the production process.

15 It is important to note that models of monopolistic competition (as opposed to perfect competition) have established rigorously how the number of goods and the amount produced of each will be determined in a decentralized economy, using two conditions: that each firm will maximize profits by producing the output level at which the marginal revenue equals the marginal cost, and that free entry ensures that in equilibrium no firm will capture monopolistic rents. For a didactic presentation of this material, see Krugman and Obstfeld (1991), chapter 6.

16 This assumption is only for expositional convenience. Strictly speaking, it is a result, not an assumption, that can be obtained as the equilibrium outcome of a monopolistic competition model in which final goods are produced from inputs that enter symmetrically (that is, have the same elasticity of substitution) in a constant elasticity substitution production function. For a more formal discussion, see Romer (1987).

17 Ethier (1982) was among the first to propose production functions of this type. For a nontechnical but persuasive presentation of the benefits of variety in production and consumption, see Cox and Alm (1999).

[18] Thus, if $\alpha = 1/2$, output will increase by $2^{1/2} = \sqrt{2} \cong 1.41$.

[19] Recall that under increasing-returns-to-scale technologies, in contrast to constant-returns-to-scale technologies, the average unit cost declines with the quantity produced.

[20] The tariff will affect M only if it gets so high that the price $b(1+ \tau/100)$ is at or above the demand curve's intersection with the vertical axes, where the equilibrium quantity demanded will be zero. As anticipated, the tariff need not be that high to affect product variety in the presence of increasing-returns-to-scale technologies.

[21] For a situation like this to emerge, the demand curve must intersect the vertical axis. Not all utility functions will induce that property. For example, the demand functions induced by the logarithmic utility functions in Part 1 never intersect the vertical axis. Of course, the studies in Part 2 specify utility functions that do induce that property on the demand for the relevant goods.

[22] Recall that the parameter M, the number of varieties in the nonconventional production function (*Equation 5*), can be interpreted as a measure of the overall efficiency of technology because it plays the same role as A, the total factor productivity parameter, in the more standard production function (*Equation 3*).

[23] The Klenow and Rodriguez-Clare model contains many interesting details that cannot be discussed here without sacrificing the focus of the article. Therefore, I sketch only those features of the model whose understanding is essential to trace the fundamental forces behind its welfare results.

[24] Because Keuschnigg and Kohler use an overlapping generation model, they do not have to confront Ahearne's difficulty of how to introduce the time preference parameter β in agents that never die.

[25] The welfare gains were computed taking into account that the capital stock will gradually adjust to its new long-run equilibrium level after the trade reform is implemented.

[26] In more technical terms, the Second Welfare theorem does not hold under increasing returns to scale; therefore, a Pareto optimum cannot typically be implemented by a free market economy. Dixit and Stiglitz (1977) show, for example, that corrective measures could involve taxes on some goods and subsidies on others. By analogy, it is not difficult to envision environments in which the remedies would involve tariffs on some imports and subsidies on some exports.

REFERENCES

Ahearne, Alan (1999), "Trade Liberalization and Capital Accumulation in Developing Economies: A Quantitative Analysis," manuscript, International Finance Division, Board of Governors of the Federal Reserve System.

Cox, W. Michael, and Richard Alm (1999), "The Right Stuff: America's Move to Mass Customization," Federal Reserve Bank of Dallas *1998 Annual Report*, 3–26.

Dixit, A., and J. Stiglitz (1977), "Monopolistic Competition and Optimum Product Diversity," *American Economic Review* 67 (June): 297–308.

Ethier, Wilfred J. (1982), "National and International Returns to Scale in the Modern Theory of International Trade," *American Economic Review* 72 (June): 389–405.

Feenstra, Robert C. (1986), "Functional Equivalence Between Liquidity Costs and the Utility of Money," *Journal of Monetary Economics* 17 (March): 271–91.

Goulder, Lawrence H., and Barry Eichengreen (1992), "Trade Liberalization in General Equilibrium: Intertemporal and Interindustry Effects," *Canadian Journal of Economics* 25 (May): 253–80.

Keuschnigg, Christian, and Wilhelm Kohler (1996), "Commercial Policy and Dynamic Adjustment Under Monopolistic Competition," *Journal of International Economics* 40 (May): 373–409.

Klenow, Peter J., and Andres Rodriguez-Clare (1997), "Quantifying Variety Gains from Trade Liberalization," manuscript, Graduate School of Business, University of Chicago.

KPMG Peat Marwick/Policy Economics Group (1991), "The Effects of a Free-Trade Agreement Between the U.S. and Mexico" (Washington, D.C.: U.S. Council of the Mexico–U.S. Business Committee, May).

Krugman, Paul R., and Maurice Obstfeld (1991), *International Economics: Theory and Policy*, 2nd ed. (New York: Harper Collins).

Mendoza, Enrique (1991), "Real Business Cycles in a Small Open Economy," *American Economic Review* 81 (September): 797–817.

Poterba, James M., and J. J. Rotemberg (1987), "Money in the Utility Function: An Empirical Implementation," in *New Approaches to Monetary Economics*, ed. William A. Barnett and Kenneth J. Singleton (New York: Cambridge University Press), 219–40.

Romer, Paul (1987), "New Theories of Economic Growth," *American Economic Review* 77 (May, Papers and Proceedings, 1987): 56–62.