Journal of Economic Perspectives—Volume 26, Number 2—Spring 2012—Pages 91–118

Gains from Trade when Firms Matter

Marc J. Melitz and Daniel Trefler

The gains from long-distance international trade have been understood and exploited since prehistoric times. Our pre-urban ancestors were benefitting from long-distance trade in obsidian some 10,000 years ago; Plato's Academy was built on the profits of Athenian silver exports; and Rome was not built in a day partly because goods moved too slowly in the vast Roman trade network.

But whereas trade was once dominated by the movement of goods that could only be produced, harvested, or mined regionally, the international trade landscape is now dominated by two striking facts. The first is the rise of intra-industry trade—that is, two-way trade in similar products. Chinese consumers can now buy a midsize car from Toyota (Japan), Kia (Korea), General Motors (United States), and Chery (China). Ditto for consumers in Japan, Korea, and the United States. The second striking fact is that world trade is dominated by huge, innovative, and extraordinarily productive firms. For example, Intel is so large that it is the largest industrial employer in both Oregon and New Mexico and accounts for 20 percent of Costa Rica's exports. China's Foxconn infamously employs 450,000 workers in a *single* one of its many export-oriented electronics factories. These are big companies . . . and if you are reading this document on an Apple computer you know that there are other large companies, too.

The rising prominence of intra-industry trade and huge multinationals has transformed the way economists think about the gains from trade. In the past, we focused on gains that stemmed either from endowment differences (wheat for

■ *Marc J. Melitz is the David A. Wells Professor of Political Economy, Harvard University, Cambridge, Massachusetts. Daniel Trefler is the J. Douglas and Ruth Grant Canada Research Chair in Competitiveness and Prosperity, Rotman School of Management, University of Toronto, Toronto, Ontario, Canada.*

http://dx.doi.org/10.1257/jep.26.2.91. doi=10.1257/jep.26.2.91

iron ore) or inter-industry comparative advantage (David Ricardo's classic example of cloth for port). Today, we focus on three sources of gains from trade: 1) love-of-variety gains associated with intra-industry trade; 2) allocative efficiency gains associated with shifting labor and capital out of small, less-productive firms and into large, more-productive firms; and 3) productive efficiency gains associated with trade-induced innovation.

Back in the 1980s, a "New Trade Theory" was developed that focused on intra-industry trade in differentiated goods produced subject to increasing returns to scale. This theory centered on an elegant tension: consumers love variety and are willing to pay a premium for a desired product, but as the market fragments into niche products, producers struggle to attain the volumes needed to recoup their product development costs. International trade creates a larger marketplace, which means that each firm can operate at a larger scale and hence more firms can survive. The result reads like an advertisement for free trade: lower prices, more varieties. Paul Krugman earned the Nobel Prize in 2008 in large part for his work highlighting how economies of scale and product differentiation lead to intra-industry trade, just as in our example above of midsize cars. See Krugman (1979, 1980), Helpman and Krugman (1985), and Helpman (2011, chap. 4) for a review of love-of-variety gains from trade.

More recently, a second source of gains from trade has emerged from the research of Melitz (2003) and Bernard, Eaton, Jensen, and Kortum (2003). This is the firm-level "reallocation" effect that arises when there is firm heterogeneity. By firm heterogeneity we mean that even within narrowly defined industries some firms are much larger and more profitable than others because, for example, they are much more productive. Globalization generates both winners and losers among firms within an industry and these effects are magnified by heterogeneity. Better-performing firms thrive and expand into foreign markets, while worse-performing firms contract and even shut down in the face of foreign competition. This generates a new source of gains from trade: as production is concentrated towards better-performing firms, the overall efficiency of the industry improves. In this way, globalization raises average efficiency *within an industry*. Why is it that only the better-performing firms grow? Globalization expands markets but also increases competition in those markets. This competition effect dominates for the worse-performing firms while the increased market access dominates for the better-performing firms. Also, a firm's international expansion—whether by exporting, by offshore outsourcing of intermediate components and assembly, or by building plants abroad (multinationals)—entails some up-front fixed costs; and only the best-performing firms have the sales volumes needed to justify these fixed costs.

Our third source of gains from trade comes from the positive impacts of larger markets on *innovation*. New productivity-enhancing products and processes require up-front development costs. Trade integration, by expanding the size of the market, encourages firms to pony up these development dollars, and this in turn raises productivity. Theories of innovation-based gains from trade with homogeneous firms were developed by Grossman and Helpman (1991) and are supported by country-level evidence (Helpman 2004, chap. 5.6). At the firm level, there is a strong relationship between exporting and innovation. For example, Intel and Apple are major patent holders, and Foxconn holds 40 percent of all Chinese patents filed in the United States (Eberhardt, Helmers, and Yu 2011). Of course, this correlation between exporting and innovation does not provide evidence for causality and lacks a framing theory featuring heterogeneous firms. Recently, however, there has been a great deal of theoretical and empirical progress. Lileeva and Trefler (2010) show theoretically and empirically how the market-expanding effects of international integration causally encourage firms to innovate. Verhoogen (2008), Bustos (2011a, b), and Aw, Roberts, and Xu (2011) assess other interesting channels through which trade promotes firm-level innovation. Note that this third source of gains deals with *within-firm* efficiency; in contrast, the second source of gains above deals with *between-firm* or allocative efficiency.

This paper reviews these three sources of gains from trade both theoretically and empirically. Our empirical evidence will be centered on the experience of Canada following its closer economic integration in 1989 with the United States—the largest example of bilateral intra-industry trade in the world—but we will also describe evidence for other countries.

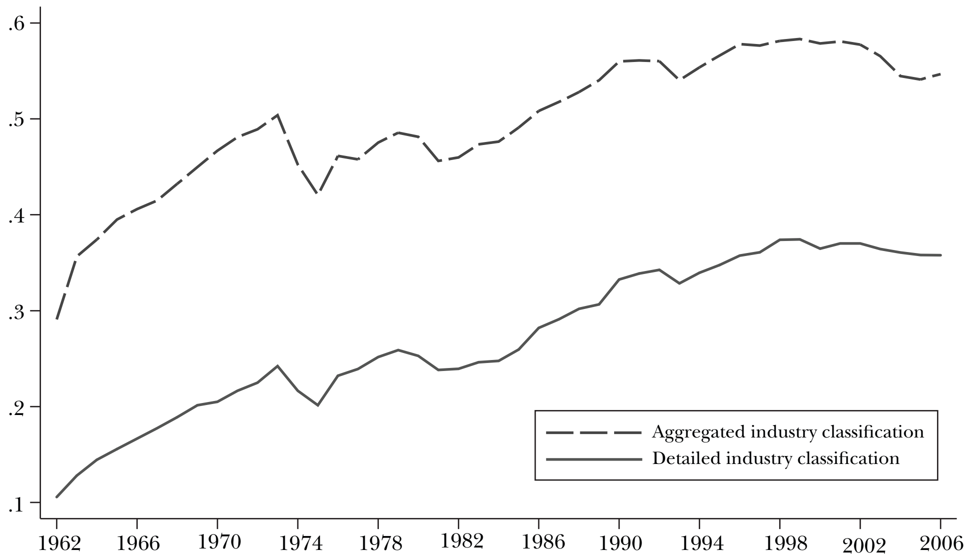
The related literature is huge. Here we focus on firms that expand internationally via exporting as in Melitz (2003) and Bernard, Eaton, Jensen, and Kortum (2003). Another related research topic analyzes how firm boundaries evolve across borders as multinational firms look abroad to "outsource" key parts of their production chain. The interested reader is directed to surveys by Antràs and Rossi-Hansberg (2009) and Helpman (2011, chap. 6).

Gains from Love of Variety (Economies of Scale and Product Variety)

Our first source of gains from trade is intimately related to intra-industry trade. To measure intra-industry trade, one needs to start with a classification system that assigns trade flows to a particular "industry." One can then categorize trade flows as either intra-industry (two-way trade within the same industry classification code) or inter-industry (imports and exports in separate industry codes). The United Nations uses the Standard International Trade Classification, or SITC, to categorize world trade flows. In its most detailed form, the SITC contains 1,161 separate industry codes (that can be consistently traced back over time), but these industries are often aggregated into a smaller subset of industries.

Figure 1 shows the time trend for the share of intra-industry trade according to this most detailed classification, and a more aggregated version with only 59 industry codes. Mechanically, the share of intra-industry trade rises with the level of aggregation for the industrial classification system (after all, with a single aggregate industry code, all trade would be "intra" to this aggregated industry). However, the time trends for the two series are very similar: intra-industry trade grew rapidly from 1962 to the mid 1990s, before stabilizing at a substantially higher level. As countries

Figure 1
World Share of Intra-Industry Trade 1962–2006



Source: Data from Brühlhart (2009). We thank Marius Brülhart for generously sharing his data.
Notes: Figure 1 shows the time trend for the share of intra-industry trade according to the most detailed Standard International Trade Classification (1,161 separate industry codes) and a more aggregated version with only 59 industry codes.

industrialize, they tend to experience a higher share of intra-industry trade because they tend to produce and export differentiated manufactured goods that are similar to other brands of goods that are imported. However, some of the countries with the highest shares of intra-industry trade in 2000 were newly industrializing nations, such as the Czech Republic (77 percent), the Slovak Republic (76 percent), Mexico (73 percent), and Hungary (72 percent); for comparison, the United States had a 69 percent share of intra-industry trade in 2000 (*OECD Economic Outlook* 2002, chap. 6; based on the 59-industry level of aggregation). Most recently, China's share of intra-industry trade has risen above the 50 percent mark.

Why might a country both export and import goods that are similar? As a starting point, consider world trade in automobiles. Consumers in a car-producing country are not limited to buying the car models that are produced domestically: many of those consumers choose to buy models that are produced elsewhere and imported. The extent of this product differentiation is then limited by high fixed start-up costs for a new brand and by the related economies of scale.

We now highlight how the combination of product differentiation and economies of scale generates intra-industry trade using a theoretical example. Notice that this source of gains from trade provides a rationale for trade between two identical countries, which provides a stark contrast with the gains from inter-industry trade that arise from exploiting differences across countries, such as differences in technology (Ricardo) or differences in factor supplies (Heckscher–Ohlin).

In our theoretical example, two identical countries produce differentiated widget varieties subject to the same constant-returns-to-scale technology. Assume that one worker can produce one widget, but that production of any new variety of widgets requires four workers to cover fixed overhead costs: this implies decreasing average costs of production as the fixed cost is spread over an increasing number of output units (hence the economies of scale). Also, to be specific, suppose that both countries have a fixed supply of 12 workers. If they do not trade, then each country can produce: A) 8 units of one variety, or B) 2 units each of two different varieties.

Allowing countries to trade leads to a new possibility that is better than what either country can achieve on its own. Suppose that each country produces 8 units of one variety and exports 4 of these units to the other country. Consumers are now consuming 4 units of the home variety and 4 units of the foreign variety. This is preferred to either of the no-trade production plans above. Compared to choice B, there is the same number of varieties (2 varieties), but more of each variety (four versus two). Compared to choice A, there is the same number of units (8 units), but more varieties (two versus one). Thus, trade expands the set of consumer choices and eases the tradeoff between consumption units and product variety. Economic integration allows production of each individual variety to be consolidated for the whole integrated market; given increasing returns to scale, this reduces average production costs. At the same time, product variety increases because consumers can buy varieties produced anywhere in the integrated market.

One of the most salient real-world examples of economic integration between similar countries occurred between the United States and Canada. This integration started with the signing of the North American Auto Pact in 1964. Before then, most car models were produced in the United States for U.S. consumers and in Canada for Canadian consumers. High tariffs on auto trade made it uneconomical to export most car models across the border. Because the Canadian auto market was roughly one-tenth the size of the U.S. market, this implied substantial scale disadvantages for production in the Canadian market: labor productivity there was about 30 percent below the U.S. level. The U.S. market was large enough that assembly lines could be dedicated to one particular car model, while Canadian assembly lines had to switch across models, involving costly downtime and reconfiguration costs, and also had to hold substantially higher inventory levels.

The 1964 Pact established a free trade area for autos that allowed manufacturers to consolidate the production of particular car models in one country and export that model to consumers in the other country. For example, General Motors cut in half the number of models assembled in Canada. However, total production of autos in Canada increased as the remaining models produced in Canada supplied the U.S. market as well as the Canadian one. Canadian automotive exports to the United States increased from $16 million in 1962 to $2.4 billion in 1968. That same year, U.S. automotive exports to Canada were valued at $2.9 billion: intra-industry trade in action. Today, $85 billion worth of automotive products cross the

U.S.–Canada border each year—roughly half in each direction. The productivity gains associated with this consolidation were also substantial: by the early 1970s, the Canadian auto industry's 30 percent labor productivity shortfall relative to its U.S. counterpart had disappeared.

Later, this transformation of the automotive industry was extended to include Mexico. In 1989, Volkswagen consolidated its North American operations in Mexico, shutting down its plant in Pennsylvania. This process continued with the implementation of the North American Free Trade Agreement between the United States, Canada, and Mexico. In 1994, Volkswagen started producing the new Beetle for the entire North American market in that same Mexican plant.

This consolidation in response to closer economic integration with the United States was not limited to the auto industry. Following the implementation of the Canada–U.S. Free Trade Agreement in 1989, each Canadian manufacturing industry experienced a dramatic reduction in its product offerings, concentrating on a smaller number of products (Baldwin, Beckstead, and Caves 2002; Baldwin, Caves, and Gu 2005; Baldwin and Gu 2009; Bernard, Redding, and Schott 2011). Baldwin, Caves, and Gu (2005) also report that the decrease in product offerings was accompanied by substantial increases in production runs for individual products. This process is even evident in the Canadian wine industry, an industry that exclusively produced low-end wines that could not possibly compete with Californian giants such as Gallo. In response to the Agreement, Canadian manufactures dramatically reduced the number of varietals produced and focused on the varietals used to produce ice wine. The industry is now healthier than ever (Beamish and Celly 2003).

Another prominent example of economic integration began in 1957, when the major countries of Western Europe established a free trade area in manufactured goods: the European Economic Community, or EEC. Many politicians offered an old-fashioned Ricardian prediction that German manufacturers would eradicate their European competitors. The facts did not treat such predictions kindly: trade within the EEC grew twice as fast as world trade during the 1960s, and intra-industry trade as a share of EEC trade more than doubled from 1960 to 1990. The benefits of the original European Community agreement were about 1 percent of GDP for the largest economies and about 3 percent of GDP for mid-sized economies such as Belgium (Harrison, Rutherford, and Wooton 1989). (These numbers capture more than just pure love-of-variety gains.) Economic integration has continued in Europe as more countries have joined the free trade area that is today called the European Union and as a subset of EU countries adopted the euro as a common currency in 1999. Euro-zone members have experienced strong trade growth, especially intra-industry trade growth, relative to non-EU countries and even relative to EU countries that have not adopted the euro.

A substantial portion of the increased trade that comes with economic integration also delivers increased product variety to consumers. Balistreri, Hilberrry, and Rutherford (2011) show that the worldwide elimination of all trade barriers would raise the number of varieties available by about 3 percent, lower manufacturing prices by a similar amount, and raise world welfare by 2 percent. Most of these gains would accrue to developing countries such as China. Broda and Weinstein (2006) estimate that the number of products available to U.S. consumers through imports tripled between 1972 and 2001, resulting in welfare gains to U.S. consumers equivalent to a 2.6 percent rise in U.S. GDP. Feenstra (2010, table 2.1) examines how worldwide welfare would change if all countries went from autarky to their 1996 levels of trade. He estimates that the welfare gains from increased varieties are comparable to a 12.5 percent rise in world GDP. While the exact magnitudes of the gains from increased variety differ across studies due to differences in what exactly is being modeled, the clear message is that the gains have been very large for developed countries and continue to be large for developing countries.

Trade expands product variety both in final goods (which benefits consumers) as well as in specialized production inputs (which benefits the firms that use those inputs). Ethier (1982) showed that there is a close parallel between these two. Instead of the love-of-variety that accrues to consumers, firms benefit from the increased productivity derived from an increased range of available production inputs. Recent firm-level research has confirmed this product variety benefit for firms that import intermediate inputs. Using Hungarian data, Halpern, Koren, and Szeidl (2005) show that importing many varieties of foreign inputs increases firm productivity by 12 percent. Using Indonesian data, Amiti and Konings (2007) show that a 10 percentage point fall in input tariffs leads to a productivity gain of 12 percent for firms that import their inputs. Kasahara and Rodrigue (2008), Kasahara and Lapham (2012), Topalova and Khandelwal (2011), and Goldberg, Khandelwal, Pavcnik, and Topalova (2010) show similarly large gains for Chile and India. In the context of the Canada–U.S. Free Trade Agreement, Lileeva and Trefler (2010) find that the fall in Canadian tariffs on inputs that Canadian firms purchased from the United States resulted in a 0.5 percent rise in Canadian manufacturing productivity. The Canadian impacts are not nearly as large as effects from developing countries, which suggests that access to a variety of inputs is an important ingredient in the process of economic development.

More variety means more competition, and more competition forces firms to lower their markups and prices. We see evidence of this after the Turkish and Cote d'Ivoire trade liberalizations of 1985 (Levinsohn 1993; Harrison 1994) and in Belgium during the 1994–2004 period of increased integration (Abraham, Konings, Vanormelingen 2009; De Loecker 2011). On the other hand, there was no evidence of falling markups in Mexico after the trade liberalization of the early 1980s (Tybout and Westbrook 1995). We turn next to the gains associated with reallocation of resources across heterogeneous firms within an industry.

Gains from Reallocation at the Firm Level

By the mid-1980s, a large body of theoretical work demonstrated that freer trade could affect productivity by forcing firms to move up or down their average cost curves. Much of the follow-on empirical work assumed that firms were identical, and it made a variety of assumptions that allowed inferences to be drawn from industry-level data (for example, Harris 1984). We now know that the heterogeneity of firms even within narrowly defined industries is a central feature of the data that cannot be ignored (for example, Bernard, Jensen, Redding, and Schott 2007).

Our second source of gains from trade is the result of shifting resources away from less-productive firms and towards more-productive firms. To analyze gains from reallocation of trade between firms, we need a model of trade with heterogeneous firms—that is, in which performance varies across different firms. We can then capture how firms with different characteristics respond differently to trade. Consider the case from the previous example where opening to trade leads to a transition from production plan A in which each country produces two varieties to production plan B in which each country produces one variety. In the real world, those varieties are associated with the firms that produce them. Opening up to trade therefore implies that one of the two firms in each country shuts down, while the remaining firm expands production from 2 units to 8 units. But what are the factors explaining which firms expand and which ones exit?

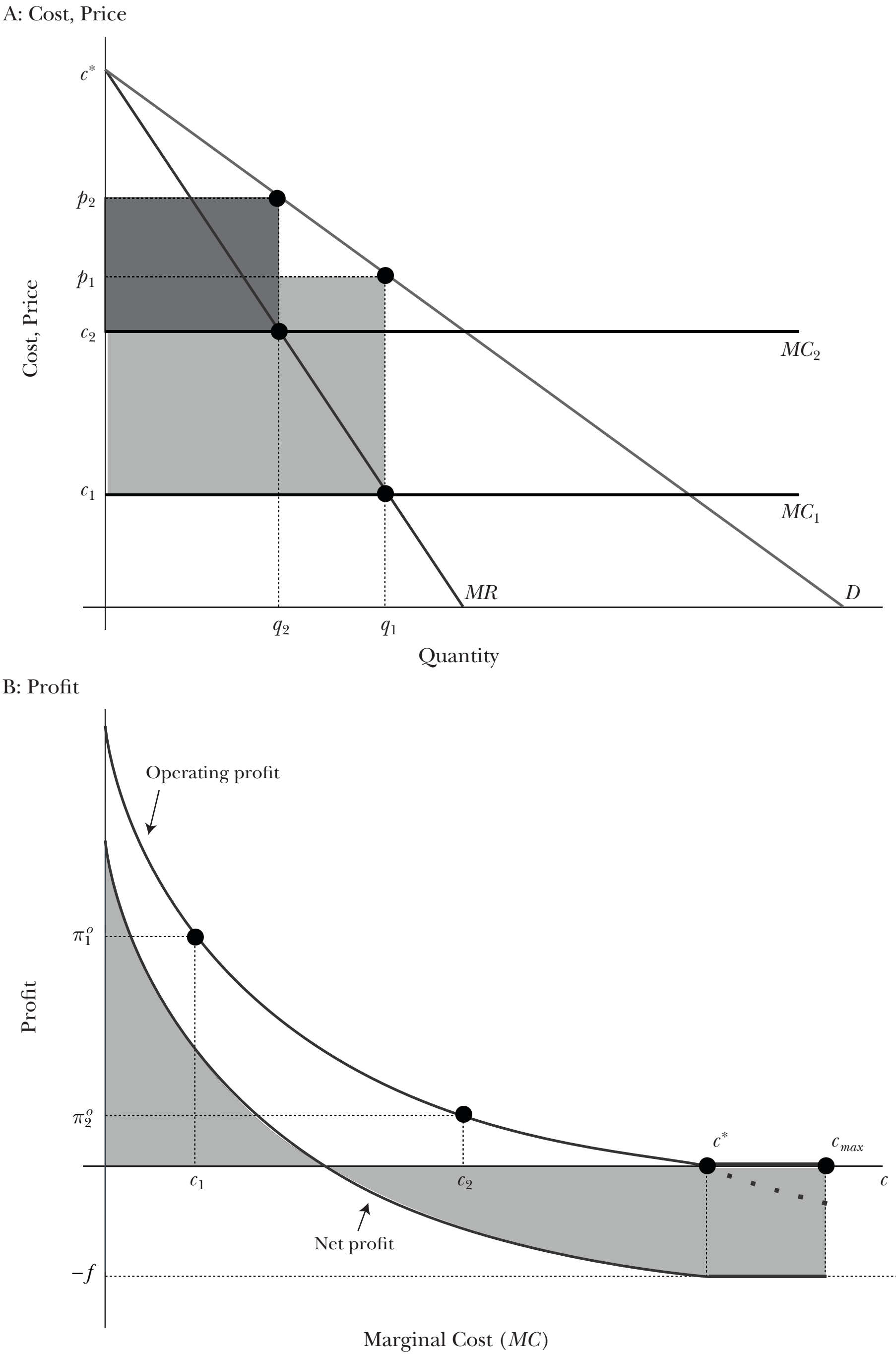
Monopolistic Competitors with Heterogeneous Costs

Melitz and Ottaviano (2008) develop a model of trade that allows for differences across firms; we use a simplified version of that model for the discussion here. Consider a monopolistically competitive industry in which many firms compete by offering different products that are relatively close substitutes for one another—at least as compared to products in other industries. For simplicity, we assume that each firm produces a single product, that demand for all products is symmetric, and that firms differ only with respect to productivity. Specifically, firms differ only with respect to their marginal costs of production c_i, where i indexes firms. (A number of authors have developed related models that allow firms to produce multiple products: for example, Eckel and Neary 2010; Bernard, Redding, and Schott 2011; and Mayer, Melitz, and Ottaviano 2011. Also, demand need not be symmetric: there can be product-quality differences across firms. Such product-quality differences lead to very similar predictions for firm performance as the ones we now discuss for cost differences.)

Panel A of Figure 2 illustrates the price and quantity choices for two monopolistically competitive firms. Both firms face the same downward-sloping *residual* demand curve: residual demand is demand as perceived by the firm, and thus depends on the behavior of other competing firms in the market.[1] On the production side,

[1] The equation for the demand facing a firm that is used in what follows is $Q = S[(1/n) - b(\bar{p} - p)]$, where Q is the quantity of output demanded, S is the total output of the industry, n is the number of firms in the industry, $b > 0$ is a constant term representing the responsiveness of a firm's sales to its price, p is the price charged by the firm itself, and $\bar{p}$ is the average price charged by its competitors. This demand equation may be given the following intuitive justification: If all firms charge the same price, each will have a market share $1/n$. A firm charging more than the average of other firms will have a smaller market share, whereas a firm charging less will have a larger share.

Figure 2
Performance Differences across Firms



Source: Authors.

marginal costs for firm 1 are shown as lower than those for firm 2. In panel A, firm 1 has a lower marginal cost (c_1) than firm 2 (c_2). We also assume that economies of scale exist because of a fixed cost that a firm must incur to develop a product and set up its initial production.

In this setting, each firm maximizes profit by choosing an output level q that equalizes marginal cost and marginal revenue. Firm 1 chooses a higher output level than firm 2 ($q_1 > q_2$), associated with a lower price ($p_1 < p_2$). Firm 1 also sets a higher markup than firm 2: $p_1 - c_1 > p_2 - c_2$; this is a consequence of the marginal revenue curve being steeper than the demand curve. Thus, firm 1 earns a higher operating profit than firm 2: $\pi_1^o > \pi_2^o$, as represented by the shaded areas in panel A of Figure 2. We assume that all firms face the same set-up cost f, so firm 1 also earns higher net profits (subtracting the fixed cost f for all firms). Of course, differences in fixed costs would not affect marginal costs and thus would not affect firm decisions concerning price and output. We can thus summarize the relevant performance differences that result from marginal cost differences across firms in the following way. Compared to a firm with higher marginal cost, a firm with a lower marginal cost will: 1) set a lower price but at a higher markup over marginal cost, 2) produce more output, and 3) earn higher profits.

Panel B in Figure 2 shows how firm profit varies with its marginal cost c_i. Both operating and net profit will be decreasing functions of marginal cost, while the difference between the two is the fixed set-up cost f. Going back to panel A, we see that a firm can earn a positive operating profit so long as its marginal cost is below the intercept of the demand curve on the vertical axis. Let c^* denote this cost cutoff. A firm with a marginal cost c_i above this cutoff is effectively "priced out" of the market and would earn negative operating profits if it were to produce any output (represented by the dotted segment for operating profit in panel A). Such a firm would choose to shut down and not produce (earning zero operating profit but incurring a net profit loss $-f$ due to the fixed cost). Why would such a firm enter in the first place? Clearly, it would not if it knew about its high cost c_i prior both to entry and to paying the fixed cost f.

We assume that entrants face some randomness about their future production cost c_i. This randomness disappears only *after* the set-up cost f is paid and is sunk. Thus some firms will regret their entry decision, as their net profit is negative (they cannot recover the sunk cost f). This is the case for firm 2 in panel B; even though its operating profit is positive, it does not cover the sunk cost f. On the other hand, some firms discover that their production cost c_i is very low and earn a high (and positive) net profit.

Firms consider all these possible outcomes, captured by the net profit curve in panel B when they make their entry decision. Firms anticipate that there is a range of lower costs where net profits are positive (shaded area to the left above the horizontal axis), and another range of higher costs where net profits are negative (shaded area to the right below the horizontal axis). In the long-run equilibrium, firms enter until their *expected* net profit across all potential cost levels c_i is driven to zero. If every cost level c_i from 0 to $c_{\max}$ is equally likely, then this equilibrium is reached when the two shaded areas are equal.[2] Panel B of Figure 2 summarizes the industry equilibrium for a given market size. It shows which range of firms survive and produce (with cost c_i below c^*) and how their profits will vary with their cost levels c_i.

What Changes When Economies Integrate?

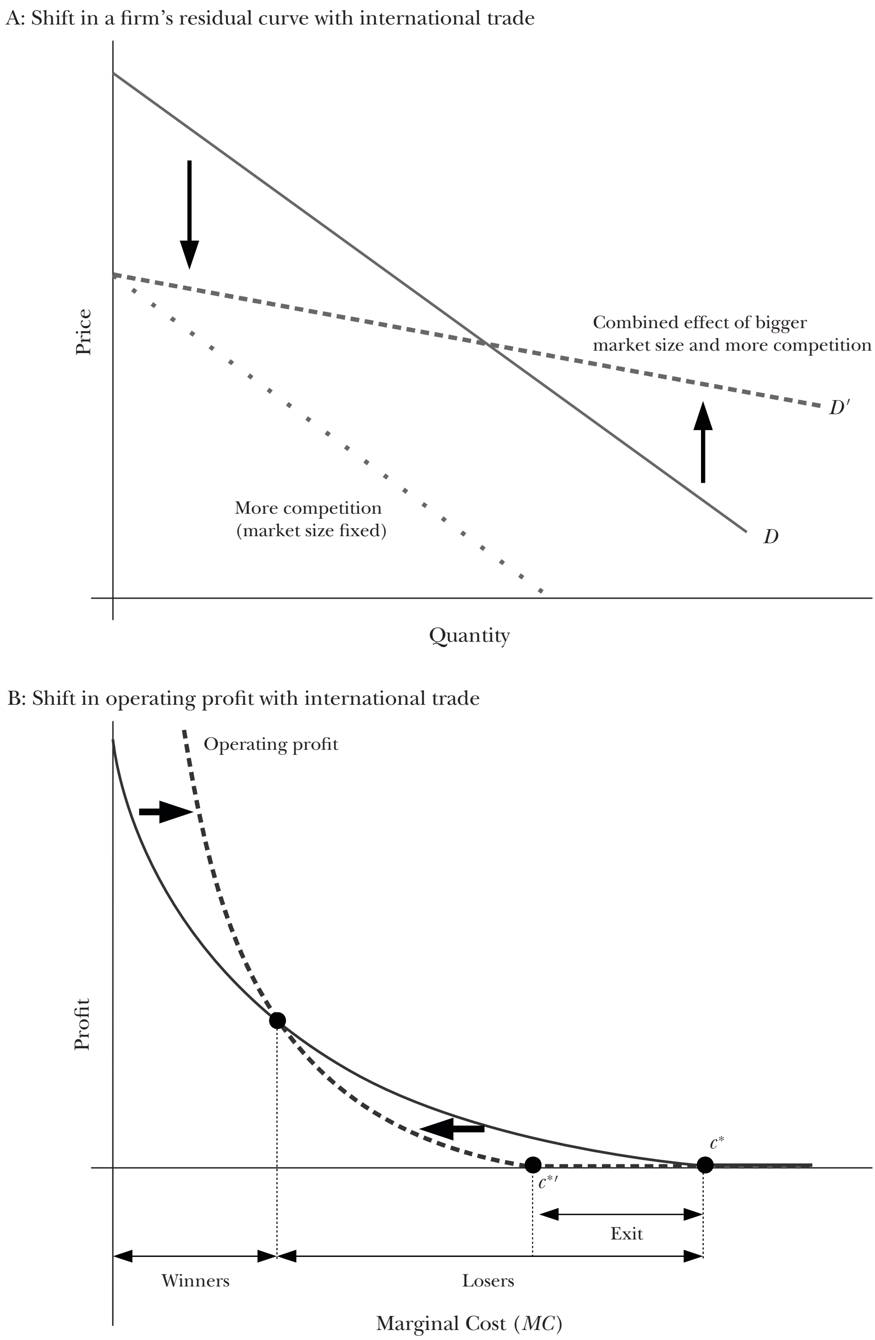
How will the situation faced by these heterogeneous firms alter when economies integrate into a single larger market? A larger market can support a larger number of firms than a smaller market, which implies more competition in the larger market. Increased competition—absent any increase in market size—leads to an inward shift of each firm's residual demand curve. On the other hand, holding competition fixed, a larger market rotates out the residual demand curves for all firms. Putting these two effects of increased competition and greater market size together gives us the combined effect of international trade on the residual demand curve perceived by firms. This change is depicted in panel A of Figure 3, as the shift from demand curve D to D'. The residual demand curve shifts in from the perspective of the smaller firms with lower output levels that operate on the higher part of the demand curve: here, the effect of tougher competition dominates. However, from the perspective of the larger firms that operate on the lower part of the demand curve, the residual demand curve has shifted out: here, the effect of the larger market size dominates.

Panel B of Figure 3 shows the consequences of this demand change for the operating profits of firms with different cost levels c_i. The decrease in demand for the smaller firms translates into a new lower cost cutoff $c^{*\prime}$: Firms with the highest cost levels (above $c^{*\prime}$) cannot survive the decrease in demand and are forced to exit. On the other hand, the flatter demand curve is advantageous to firms with the lowest cost levels: they can adapt to the increased competition by lowering their markup (and hence their price) and gaining some additional market share. (Recall that the high-cost firms are already setting low markups, and cannot lower their prices to induce positive demand, as this would mean pricing below their marginal cost of production.) Thus, the best-performing firms with the lowest cost levels c_i now earn increased operating and net profits. Panel B of Figure 3 illustrates how increased market size generates both winners and losers amongst firms in an industry. Low-cost firms thrive and increase their profits and market shares, high-cost firms contract, and the highest cost firms exit.

In this model, economic integration through market expansion does not directly affect firm productivity. Nevertheless, it generates an overall increase in aggregate productivity as market shares are reallocated from the low-productivity firms with high marginal costs to the high-productivity ones with low marginal costs.

[2] In contrast, when there is no uncertainty about marginal cost because all firms share the same cost c, then entry drives the realized net profit to zero for all firms. With firm heterogeneity, expected net profit is zero, but realized profits will vary as shown in panel B of Figure 2.

Figure 3
Winners and Losers from Market Integration



Source: Authors.

Trade Costs, Export Decisions, and Trade Liberalization

The discussion to this point has implicitly modeled economic integration as a change in market size from a closed economy with no trade all the way to a single combined market with no trade barriers. In reality, initial trade costs are rarely so high as to block all trade prior to liberalization, and liberalization reduces trading costs without fully eliminating them. In a number of ways, this kind of partial trade liberalization has a very similar effect to the simpler case of full integration. With partial trade liberalization, the better-performing firms expand, the worse-performing ones contract, and the worst performing ones exit. This generates the same type of reallocation effect previously described and leads to a rise in aggregate productivity.[3]

However, adding trade costs also allows us to analyze an additional issue: whether firms choose to export. With trade costs, exporting is profitable only for a subset of better-performing firms. Some firms do not export, and instead only serve domestic consumers. We now extend our theoretical model to incorporate trade costs and firms' export decisions. For this purpose, we can no longer analyze a single market: instead, we need to look at firms' decisions in both the domestic and export markets jointly. For simplicity, we consider a special case where both countries are symmetric, so that demand conditions in both the domestic and export markets will be identical.

Assume that a firm must incur an additional trade cost t for each unit of output that it sells to customers across the border. As a result of this trade cost, each firm will set a different price in its export market relative to its domestic market, which will lead to different quantities sold in each market, and ultimately to different profit levels earned in each market. Because we are assuming that each firm's marginal cost is constant and does not vary with production levels, the decisions regarding pricing and quantity sold in each market can be separated: a decision regarding the domestic market will have no effect on the profitability of different decisions for the export market.

Consider the case of firms located in Home. Their situation regarding their domestic (Home) market is exactly as was illustrated in Figure 2, except that all the outcomes such as price, output, and profit relate to the domestic market only. Now consider the export (Foreign) market. The firms face the same demand curve in Foreign as they do in Home (the two countries are identical). The only difference is that each firm's marginal cost in the export market is shifted up by the trade cost t. What are the effects of the trade cost on the firms' decisions regarding the export market? A higher marginal cost induces a firm to raise its price, which leads to a lower output quantity sold and to lower profits (as highlighted in Figure 2).

[3] The more general version of this model analyzed by Melitz and Ottaviano (2008) allows for multiple countries of different sizes and for arbitrary trade costs between any country pair (though the trade costs are proportional to production costs instead of per output unit as in the current version). That paper shows more formally that the effects of multilateral liberalization (all countries proportionally reduce trade costs) are very similar to the case of full economic integration that leads to a single larger market.

We also know that if marginal cost is raised above a threshold level c^*, then a firm cannot profitably operate in that market. Thus, when there are trade costs, some firms will find it profitable to operate in the domestic market but not in the export market because the trade cost pushes their marginal cost for that market above the threshold c^*.

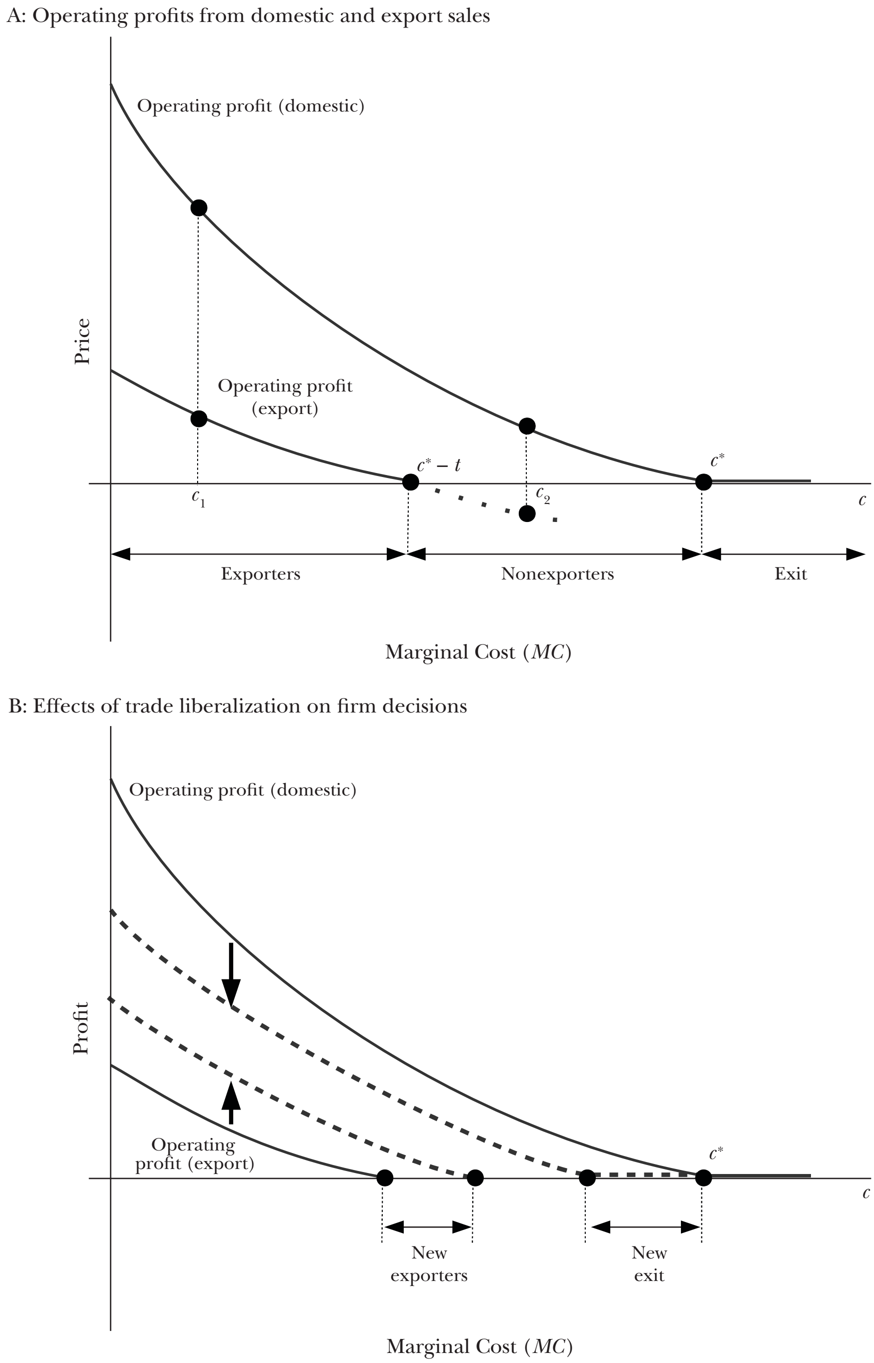
Figure 4 helps to visualize the production and export decisions for all firms based on their marginal cost c_i. Panel A of Figure 4 separates a firm's operating profit into a portion earned from domestic sales, and a portion earned from export sales. (Both portions are functions of a firm's marginal cost c_i as in Figure 2.) Because the only difference between the domestic and export markets is the additional per-unit trade cost t, the horizontal distance between the two curves is equal to the trade cost t. Firm 1 earns positive operating profits from sales in both the domestic and export markets: it will export and reach consumers in both markets. This will be the case for all firms with cost below $c^* - t$. On the other hand, firm 2 only earns positive operating profits from sales in the domestic market—and thus chooses not to export. Any firm with cost above $c^* - t$ will be in this same situation and therefore will not export: those firms only serve their domestic market. As before, the worst performing firms with cost above c^* cannot profitably operate at all (even in their domestic market) and therefore exit.

Panel B of Figure 4 summarizes the effects of trade liberalization—a reduction in the trade cost t—for those firm decisions. The figure shows the same two operating profit curves from panel A both before and after (dashed curves) trade liberalization. The operating profit for the domestic market shifts down due to the increase in competition (which shifts the residual demand curve for the domestic market inward as explained earlier). Some of the higher-cost firms that used to produce for domestic consumption no longer earn a positive operating profit after trade liberalization and exit. On the other hand, the operating profit for the export market shifts up due to the lower trade cost. (Increased competition in the export market tends to reduce operating profits there, but this effect is dominated by the direct effect of the trade cost reduction.) A key empirical prediction is that some firms start exporting. Specifically, among the firms that did not export prior to trade liberalization, only the most productive of these start exporting.

Evidence on Gains from Inter-Firm Reallocations

In many ways, the Canada–U.S. Free Trade Agreement is a useful natural experiment for considering the effects of trade integration. The policy experiment is clearly defined: it dealt only with market integration and was not part of a larger package of macroeconomic reforms that often accompany trade liberalization. The enactment of the agreement was largely unanticipated: a Canadian general election was fought on the issue one month before the agreement was to be signed into law and pollsters unanimously predicted that Canada's ruling party—along with the free trade agreement—would be defeated (Brander 1991; Thompson 1993). Thus, evidence about the extent of aggregate productivity changes as a result of reallocations among heterogeneous firms can be sought by looking at the distribution of

Figure 4
Export Decision and Trade Liberalization



Source: Authors.

productivity across Canadian manufacturing plants before and after the agreement, at entrants before and after the agreement, and at the productivity distribution of exporters and nonexporters.

The agreement came into effect on January 1, 1989. Panel A of Figure 5 shows the distribution of labor productivity as measured by value-added per employee across Canadian manufacturing plants both before the agreement in 1988 and in 1996, when there had been time for firm adjustments to occur. For example, the 1996 curve summarizes the productivity distribution of all 35,000 Canadian manufacturing plants in that year. Clearly, the distribution of firms shifted rightward: between 1988 and 1996, the share of low-productivity plants in manufacturing declined and the share of high-productivity plants rose.

The horizontal axis is based on a measure of the log of labor productivity. However, to ensure that dispersion is driven by within-industry rather than between-industry differences in labor productivity, we scale each plant's log productivity by subtracting from it the median log productivity of the plant's four-digit SIC industry. Thus, the median plant in each industry has a score of zero on the horizontal axis. The vertical axis shows the share of plants with the indicated level of productivity. These frequencies are weighted by plant employment; otherwise, tiny plants that account for only a tiny fraction of total employment would dominate the figure.

To get a sense of the degree of productivity dispersion, consider the horizontal axis of Figure 5 and suppose that log productivity at plant A is one unit higher than at plant B. This is equivalent to saying that plant A is three times more productive than plant B. If A is four units higher than B, then A is 50 times more productive than B.[4]

Obviously, labor productivity as shown in Figure 5 is not an identical concept to the horizontal lines showing levels of marginal cost in the theoretical discussion. When marginal costs are low, we typically expect productivity to be high. Therefore, the inverse of marginal costs ($1/c$) is often proxied in empirical work by productivity.

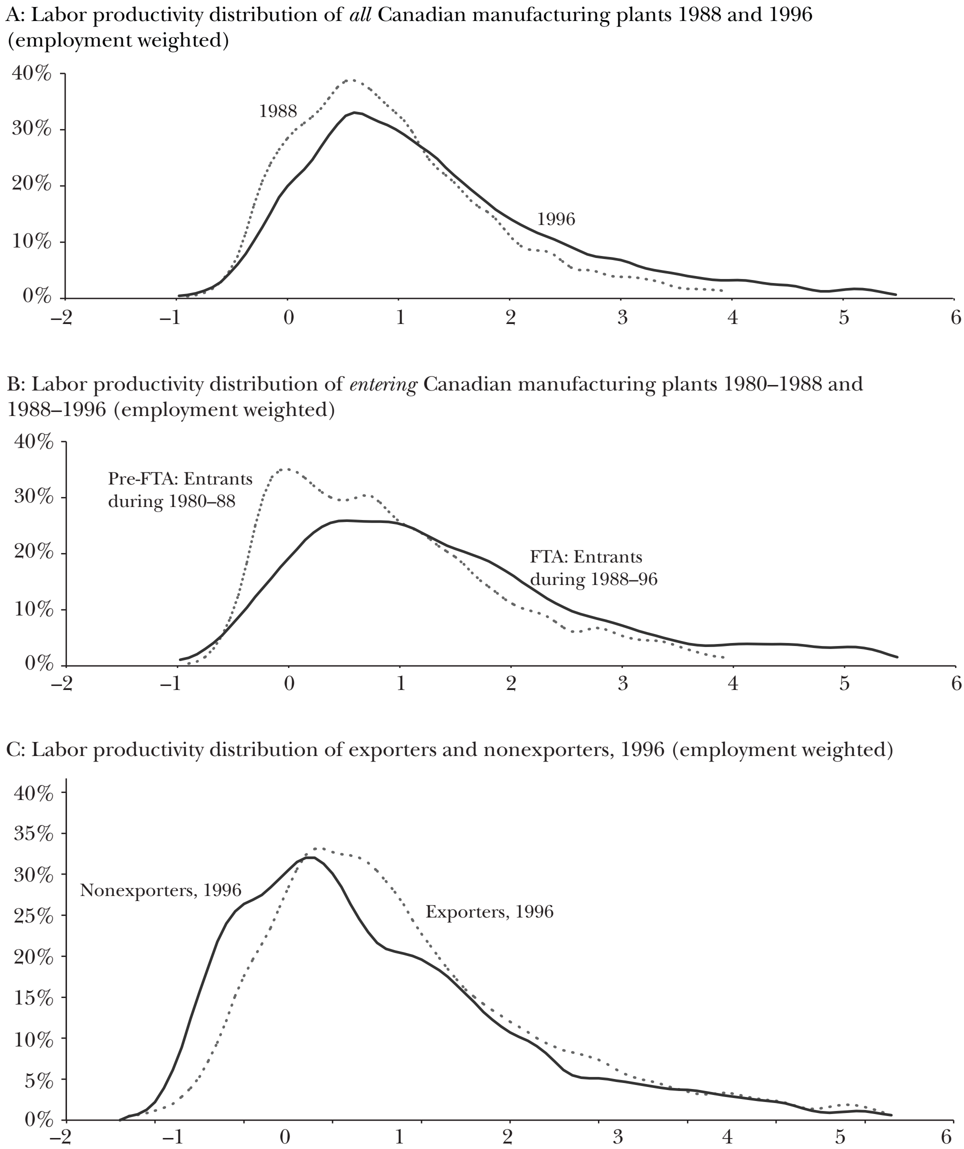
The productivity heterogeneity shown for Canadian manufacturing firms in Figure 5 is a pervasive feature of all economies including, for example, the United States (Bernard, Eaton, Jensen, and Kortum 2003), many European economies (Mayer and Ottaviano 2008; Bartelsman, Hatiwanger, and Scarpetta 2009), as well as China and India (Hsieh and Klenow 2009). Wagner (2007) surveys related studies from countries all around the world, reporting similar patterns regarding widespread firm heterogeneity within industries. Thus, the lessons derived from this example are not specific to the Canadian experience.

What caused the change from 1988 to 1996 in the productivity distribution of Canadian manufacturing firms? It was largely due to the reallocation mechanisms across plants described above. The first of these mechanisms that we examine is

[4] Let φ_A and φ_B be productivities of A and B and suppose that they are 1 unit apart i.e., $\ln(\varphi_A) - \ln(\varphi_B) = 1$. From the property of logs, $\varphi_A/\varphi_B = e^1 = 2.7 \approx 3$. For a difference of 4 units, $e^4 \approx 50$. On a more technical level, the figure is constructed starting with "standardized" log productivities (see the formula in Lileeva 2008, p. 369), which we then multiply by a single scale factor to transform the standardized log productivities into log productivities that are directly comparable with log productivities reported in studies from other countries.

Figure 5

Distribution of Productivity across Canadian Manufacturing Plants before and after the Canada–U.S. Free Trade Agreement



Source: Authors' calculations using data from the Canadian Annual Survey of Manufactures.
Notes: The horizontal axes are based on a measure of the log of labor productivity. However, to ensure that dispersion is driven by within-industry rather than between-industry differences in labor productivity, we scale each plant's log productivity by subtracting from it the median log productivity of the plant's four-digit SIC industry. Thus, the median plant in each industry has a score of zero on the horizontal axis. The vertical axis shows the share of plants with the indicated level of productivity. These frequencies are weighted by plant employment; otherwise, tiny plants that account for only a tiny fraction of total employment would dominate the figure.

the fall in the survival threshold of marginal cost—that is, the leftward shift of c^* in panel B of Figure 4. The empirical counterpart to a fall in c^* is a rise in the break-even level of productivity. One can examine this mechanism by looking either at exit rates or at entry rates. Since a plant may not exit until it has completed the multiyear process of depreciating its fixed capital, it is best to look at entry rates, which adjust more quickly to shocks such as a free trade agreement. Panel B of Figure 5 displays the productivity levels of new entrants to Canadian manufacturing both for the pre-agreement period (1980–1988) and for the period immediately after the agreement came into force (1988–1996). There was a striking decline in the entry rates of plants with productivity near or below the median. To use a sports analogy, in the earlier period many low-productivity plants made the cut and joined the team while in the later period a number of such low-productivity plants no longer made the cut.[5]

The pattern here is confirmed by more formal econometric analysis (specifically, binary-outcome regressions of exit on plant characteristics as well as industry and time controls). For example, Baggs (2005) and Baldwin and Gu (2006) show that the free trade agreement tariff cuts raised exit by a large amount. Lileeva (2008) estimates that the free trade agreement tariff cuts raised exit rates by as much as 16 percent, with all of the increase involving the exit of nonexporters. Bernard, Jensen, and Schott (2006) find similar results for U.S. plants faced with U.S. tariff reductions.

Trade Costs and the Export Decision

A central prediction of the theory is that in the presence of trade costs, only low-cost, high-productivity firms export. Panel C of Figure 5 shows the distribution of Canadian plants separately for exporters and nonexporters. Clearly, the distribution for exporters is to the right of that for nonexporters. On average, Canadian exporters are 40 percent more productive than nonexporters in the same industry (Baldwin and Gu 2003). Since the seminal work of Bernard and Jensen (1995), a huge body of research covering dozens of countries has found this same pattern of higher productivity for exporters relative to nonexporters.[6]

[5] Bernard and Jensen (1999), Trefler (2004), Lileeva (2008), and Lileeva and Trefler (2010) all point out that one must look not just at pre-Free Trade Agreement *levels* (as in Figure 5), but also at pre-FTA *trends*. All of the FTA results reported here hold with pre-FTA controls for both levels and trends. For example, variants of some of the panels in Figure 5 with pre-FTA trend controls appear in Lileeva (2008).

[6] In examining panel C of Figure 5, a critical reader may wonder why there are so many highly productive nonexporters and whether this contradicts the theory. A simple but prominent example may help to explain. Highly productive auto parts plants often cluster around a giant auto assembly plant—Ford, General Motors, and Honda all have major auto assembly plants near Toronto, Canada, that are surrounded by parts suppliers. These parts suppliers are highly productive, but do not directly export. This pattern is clearly not a challenge to the theory: these highly productive plants are supplying parts that are assembled into the autos that are ultimately exported to the United States: highly productive parts suppliers are "indirect" exporters. Of course, there are surely other factors outside the scope of our model that explain why some very productive firms do not export—and conversely—why some low-productivity firms export.

A much more demanding prediction of the theory deals with who will *start exporting in response to falling trade costs.* Panel B of Figure 4 shows that those who start exporting *will* be among the most productive of those who never exported before. To test this prediction, Lileeva and Trefler (2010) examined a sample of over 5,000 Canadian manufacturing plants that had never exported prior to the Canada–U.S. free trade agreement. A very large percentage of these plants (40 percent) started exporting after the agreement came into force on January 1, 1989. Lileeva and Trefler examine whether these plants started exporting because of the U.S. tariff cuts and, more importantly, whether those that started exporting because of the tariff cuts were more productive than nonexporters. To this end, Lileeva and Trefler divide up their sample into quartiles of the 1988 distribution of labor productivity (with the quartiles defined separately for each industry to net out industry characteristics). Only 20 percent of the plants in the bottom quartile of labor productivity started exporting because of the tariff cuts, compared to nearly 60 percent of the plants from the top quartile of labor productivity. (These estimates are from a probit regression in which the dependent variable is 1 if the plant started exporting and 0 if the plant remained a nonexporter. The key independent variable is a plant-specific measure of the change in the U.S. tariff. This measure is described below.) The key conclusion is that, among firms that did not export before trade liberalization, the most productive of these were three times more likely to start exporting in response to the U.S. tariff cuts. This is as predicted in panel B of Figure 4.[7]

Quantifying the Gains from Trade Due to Reallocation across Heterogeneous Plants

In the wake of the Canada–U.S. free trade agreement, Canadian manufacturing productivity rose sharply. We have shown that part of this productivity gain was due to the reallocation mechanisms highlighted by the theory. But how important were these in quantitative terms for productivity growth and overall welfare?

In the conventional approach to estimating the gains from trade, the focus is on measuring welfare, or more specifically, on the income a society would be willing to pay for lower tariffs. These "compensating variation" gains are typically derived from models that 1) make a large number of parametric assumptions (assuming very specific functional forms for preferences that determine the extent of product differentiation, as well as for the utility derived from love-of-variety), and 2) make use of parameter estimates about which we are highly uncertain. In short, a lot of uncertainty surrounds welfare-gain estimates. In the heterogeneous-firms literature, the focus has shifted to estimating productivity gains rather than welfare. The last two decades have seen major improvements in our ability to estimate productivity gains, both because of the creation of

[7] On a related note, profits play a key role in all the mechanisms of our model. Baggs and Brander (2006) confirm that profits move in the expected directions. In particular, they find that falling Canadian tariffs are associated with declining Canadian profits, especially for import-competing firms, while falling U.S. tariffs are associated with increasing Canadian profits, especially for export-oriented firms.

high-quality, firm-level longitudinal data and because of methodological developments aimed at exploiting these data. Thus, although productivity gains are not the same as welfare gains, we have much greater confidence in our estimates of the productivity gains associated with freer trade.

The productivity gains associated with the reallocation of market shares across firms following the Canada–U.S. free trade agreement are usefully broken into two components. First, *the fall in the U.S. tariffs* allowed Canadian plants to export more. This shifted the composition of output towards high-productivity exporters and away from low-productivity nonexporters. Lileeva and Trefler (2010) estimate that the fall in U.S. tariffs causally raised Canadian manufacturing productivity by 4.1 percent via this export-composition channel. Second, *the fall in the Canadian tariffs* led to a shift in domestic market shares—exporters gained market share at the expense of nonexporters. In the extreme, many nonexporters simply went out of business. Trefler (2004) calculates that this selection effect increased overall Canadian manufacturing productivity by 4.3 percent.[8]

Putting these numbers together, we see that the reallocation and selection effects induced by the free trade agreement generated a productivity increase of 8.4 percent (4.1 + 4.3) for Canadian manufacturing. This represents a massive productivity increase in just a short time—especially when one considers that this productivity gain did not come from productivity improvements at the plant level: it only came from the shifting of market shares from less- to more-productive plants.

Canada is not the only country to have experienced such a substantial productivity boost from reallocations driven by trade liberalization. Bernard and Jensen (2004) find that almost half of all U.S. manufacturing productivity growth during 1983–1992 is explained by the reallocation of resources towards exporters. Pavcnik (2002) studies the response of the Chilean manufacturing sector to a massive trade liberalization episode that took place from 1979 to 1986. She finds that two-thirds of the ensuing 19 percent increase in productivity (another example of a massive increase in aggregate productivity) was generated by composition changes within industries due to a reallocation of market shares towards more-efficient producers. Surveys by Greenaway and Kneller (2007) and Wagner (2007) summarize the connections between trade liberalization and aggregate productivity—including this reallocation effect across heterogeneous firms—for a wide range of studies and countries.

[8] Specifically, Trefler (2004) regressed labor productivity growth in the period after the free trade agreement (relative to the pre–agreement period) on U.S. and Canadian tariff cuts mandated by the agreement. He then showed that the *Canadian tariff cuts* raised productivity at the industry level, but not at the plant level. This means that the gains in productivity were coming from selection, rather than from improvements at the plant level. Using this approach, he finds that the free trade agreement raised Canadian manufacturing labor productivity by 5.8 percent of which 4.3 percent was due to the exit associated with the Canadian tariff cuts.

Gains from Rising Within-Plant Productivity

In this section, we move from this "between-plant" effect in which productive plants expand at the expense of less-productive plants to our third source of gains from trade: a "within-plant" effect in which trade raises productivity of individual plants by raising the returns to innovation.

At least as far back as Schmookler (1954), economists have known that the larger the market, the more profitable it is for firms to invest in productivity-enhancing activities. Firms in large markets have the large sales volumes needed to justify incurring the high fixed costs of innovation. The U.S. tariff cuts that were part of the U.S.–Canada free trade agreement greatly increased the size of the market faced by Canadian firms. It should therefore have encouraged Canadian firms to increase their exporting *and* to increase their investments in productivity-enhancing technologies. We start here with a short extension to the theoretical model that captures how larger markets generate incentives for some firms to innovate, and then turn to empirical evidence.

A Theory of Market Size and Firm Innovation

Suppose that an innovation process requires an up-front fixed cost f_I and in return generates a *reduction* in marginal cost Δc_I. That is, innovation reduces marginal cost from c to $c - \Delta c_I$. A firm that produces q units of output and engages in innovation will lower its production costs by $q \times \Delta c_I$. The firm will weigh this cost saving against the fixed innovation cost f_I, and innovate if $q \times \Delta c_I > f_I$ or

$$q > \frac{f_I}{\Delta c_I}.$$

In words, only firms with large volumes q (that is, those with initially lower levels of marginal cost) will find it profitable to innovate. What happens to this firm-level innovation decision when trade is liberalized? Lower trade costs increase an exporter's sales in the export market, and thus increase the exporter's overall output level q. For some exporters, this increase in output will tip the balance in favor of innovating. For some nonexporters, trade liberalization will tip the balance in favor of exporting *and* innovating.

Evidence on Within-Firm Productivity Growth and Trade

For evidence on the link from growth of trade to within-firm productivity, we turn again to Canada's experience with the free trade agreement. Lileeva and Trefler (2010) look at their sample of 5,000 Canadian manufacturing plants that did not export prior to 1988 and divide these plants into those that started exporting after the passage of the free trade agreement and those that did not. In the raw data, the labor productivity of those that started to export rose 29 percent more than for nonexporters; starting to export was highly correlated with within-plant productivity growth. Of course, this 29 percent number does not take into account a serious problem of reverse causality: does exporting lead to increased productivity or does increased productivity lead to exporting?

To answer this question, one needs an instrument for exporting: that is, one needs an event that causes a firm to export but that does not directly affect its productivity growth. As Lileeva and Trefler (2010) show, "plant-specific" tariff cuts fit the bill as an instrument. Consider a single Canadian plant called Lumberjack and the many products it produces. Empirically, products are defined very narrowly, at the six-digit level of the Harmonized System product classification, so that there are thousands of products in manufacturing. For each product produced by Lumberjack, one can calculate the U.S. tariff cut. Averaging these tariff cuts across all of Lumberjack's products yields a plant-specific tariff cut. This plant-specific tariff cut has enormous power in predicting whether a Canadian plant starts exporting and how much it exports. The tariff cut also has no direct impact either theoretically or empirically on a plant's productivity growth. It is thus a novel and valid instrument.

Lileeva and Trefler (2010) actually do something fancier than instrumental variables—they estimate the local average treatment effect (LATE). This is the effect on productivity of starting to export *for those plants that started exporting because of the tariff cuts.* Thus, unlike previous studies of the causal impacts of exporting on productivity, their work only uses information drawn from plants that were likely to be affected by the free trade agreement. Using their plant-specific tariff instrument and the local average treatment effect estimator, they establish that the free trade agreement *caused* the productivity of new exporters to rise by 15.3 percent. Since this 15.3 percent rise occurred in plants that accounted for 23 percent of Canadian manufacturing output, the 15.3 percent rise in labor productivity of new exporters raised *overall* Canadian manufacturing productivity by 3.5 percent ($3.5 = 15.3 \times 0.23$; see Table 2 below).

Lileeva and Trefler (2010) then trace the sources of this productivity gain back to investments in productivity. The authors examine advanced manufacturing technologies and rates of innovation at these plants. Table 1 presents the results. Consider the first row, which deals with management techniques essentially associated with lean manufacturing. In the period immediately after implementation of the free trade agreement, 16 percent of new exporters adopted such techniques, 10 percentage points more than nonexporters did. The final column, which reports local average treatment effect (LATE) estimates, shows that 7 of the 10 percentage points was attributable to the effects of increased exporting resulting from the U.S. tariff cuts. As shown in Table 1, similar results hold for the adoption of other technologies and for innovation.

These results break with the discussion of Bernard, Jensen, Redding, and Schott (2007) in this journal, who correctly argue that most careful studies show exporting *does not* raise productivity. Over the years, however, a few careful studies have found otherwise, as in Canada (Baldwin and Gu 2003, 2004; Lileeva 2008), in nine African countries (Van Biesebrock 2005), and in Slovenia (De Loecker 2007). López (2005) provides an overall survey.

Table 1
Innovation Response to Free Trade Agreement by New Exporters

	Raw adoption and innovation rates			*LATE*
	New exporters	*Nonexporters*	*Difference*	*Difference*
Manufacturing information systems	16%	6%	10%	7%
Inspection and communications	18%	10%	8%	8%
Any product or process innovation	30%	20%	10%	8%
Any product innovation	26%	14%	12%	7%

Source: Lileeva and Trefler (2010).
Note: "LATE" is the local average treatment effect.

What has recently buttressed the minority view that a rise in exporting can lead to a rise in productivity is a spate of papers isolating the causal mechanisms through which exporting affects productivity. We have already seen the market-size mechanism of Lileeva and Trefler (2010). Bustos (2011a, b) develops a related model of scale-biased technology choice, which she takes to Argentinean firm-level data for the 1992–1996 period. Bustos (2011b, table 6) shows that firms that began exporting between 1992 and 1996 also increased their technology spending. Bustos (2011a) shows that technology spending increased most in sectors that experienced improved access to Brazilian product markets (through the tariff cuts in the Mercosur regional trade agreement). In a series of studies of Taiwanese electronics exporters, Aw, Roberts, and Winston (2007) and Aw, Roberts, and Xu (2008, 2011) show empirically that there is a complex dynamic interplay between the decisions to export and conduct research and development, with today's decisions about one affecting tomorrow's decisions about the other—and both affecting future productivity. Extending this dynamic methodology to general equilibrium, Shen (2011) also finds strong complementarities between exporting and productivity-enhancing investments among Spanish firms. Bloom, Draca, and Van Reenen (2011) show that import competition from China has led to increases in R&D, patenting, information technology, and total factor productivity among more technologically advanced European firms. Atkeson and Burstein (2010) are the only ones to examine exporting and investing in productivity within a full-blown general equilibrium analysis. They find the general equilibrium feedbacks are important.

A New Dimension of Heterogeneity

In our theoretical model above, firms below a certain productivity threshold should not be exporting. Yet in the empirical work reviewed above, we saw that many low-productivity Canadian plants started exporting in response to U.S. tariff cuts. There is a second puzzle that we have not yet noted: Lileeva and Trefler (2010, table 3) report that the plants that gained most from starting to export (both in terms of productivity gains and increased innovation) were primarily plants that initially had low productivity. That is, among plants that started to export, the benefit was greatest for the least-productive plants.

To see why, consider a firm that is just indifferent between investing and not investing. From the earlier equation, indifference means that $q = f_I/\Delta c_I$, where Δc_I is the reduction in marginal cost or the increase in productivity. Rearranging ($\Delta c_I = f_I/q$) and noting that sales q are increasing in initial productivity, we arrive at a simple conclusion. *Among the set of firms that are just indifferent between innovating and not innovating*, the less-productive, low-q firms must expect larger productivity gains Δc_I from innovation. Lileeva and Trefler's (2010) results strongly confirm this prediction.

Conclusions

Recent research into the welfare gains from intra-industry trade have focused on three sources of gains: 1) gains from increased variety and economies of scale, 2) productivity gains at the industry level from shifting resources away from low-productivity firms and towards high-productivity firms, and 3) productivity gains at the firm level from innovating for a larger market. Each of these mechanisms have proven to be highly important empirically in the context of the exhaustively studied Canada–U.S. free trade agreement, and also appear important in many other less-studied contexts. Indeed, Balistreri, Hillberry, and Rutherford (2011) show that adding firm heterogeneity to standard computable equilibrium models of trade raises the gains from trade liberalization by a multiple of four. Empirical confirmation of the gains from trade predicted by models with heterogeneous firms represents one of the truly significant advances in the field of international economics.

We summarize the causal effects of the free trade agreement on overall Canadian manufacturing productivity in Table 2. As the last row shows, Canadian manufacturing labor productivity rose by 13.8 percent. The idea that a single government policy could raise productivity by such a large amount and in such a short time-span is truly remarkable.

In writing this review, we have focused on the net gains from trade. Yet the model we have developed highlights how intra-industry trade will generate both winners and losers. For example, in the context of the Canada–U.S. free trade agreement, Trefler (2004) shows that 12 percent of the workers in low-productivity firms lost their jobs. Recent research suggests that American workers are similarly struggling in response to the Chinese import surge (Liu and Trefler 2011; Autor, Dorn, and Hanson 2011). Clearly, this suggests an important role for policies that provide an adequate safety net and transitional assistance for those affected workers. The blow to those workers could be cushioned by policies that impede the reallocation process across firms. However, such policies—as opposed to policies that provide some form of direct assistance to the affected workers—would also entail a substantial long-run cost. After all, it is precisely this reallocation process that generates some of the long-run gains that we have described. In addition, policies that seek to impede the

Table 2
Overall Effect of Free Trade Agreement on Canadian Manufacturing Productivity

Selection/reallocation (between plants)	
Growth of exporters (most-productive plants)	4.1%
Contraction and exit of least-productive plants	4.3%
Within-plant growth	
New exporters invest in raising productivity	3.5%
Existing exporters invest in raising productivity	1.4%
Improved access to U.S. intermediate inputs	0.5%
Total	**13.8%**

Source: Trefler (2004) and Lileeva and Trefler (2010).

reallocation process by making firm contractions and expansions costlier would also reduce the potential gains to firm innovation and hence lead to less innovation and further depress the potential long-run gains from trade. Nonetheless, it is important to remember that there are winners and losers from trade liberalization not just among firms, but also among their employees.

■ *Trefler gratefully acknowledges funding from the Social Sciences and Humanities Research Council of Canada (SSHRC) and the Canadian Institute for Advanced Research (CIFAR). His research would not have been possible without the tremendous support of John Baldwin and Alla Lileeva. We thank the editors (David Autor, Chad Jones, John List, and Timothy Taylor) for their invaluable comments.*

References

Abraham, Filip, Jozef Konings, and Stijn Vanormelingen. 2009. "The Effect of Globalization on Union Bargaining and Price-Cost Margins of Firms." *Review of World Economics (Weltwirtschaftliches Archiv)* 145(1): 13–36.

Amiti, Mary, and Jozef Konings. 2007. "Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia." *American Economic Review* 97(5): 1611–38.

Antràs, Pol., and Esteban Rossi-Hansberg. 2009. "Organizations and Trade." *Annual Review of Economics* 1(1): 43–64.

Atkeson, Andrew, and Ariel Tomás Burstein. 2010. "Innovation, Firm Dynamics, and International Trade." *Journal of Political Economy* 118(3): 433–84.

Autor, David H., David Dorn, and Gordon H. Hanson. 2011. "The China Syndrome: Local

Labor Market Effects of Import Competition in the United States." August. http://economics.mit.edu/files/6613.

Aw, Bee Yan, Mark J. Roberts, and Tor Winston. 2007. "Export Market Participation, Investments in R&D and Worker Training, and the Evolution of Firm Productivity." *World Economy* 30(1): 83–104.

Aw, Bee Yan, Mark J. Roberts, and Daniel Yi Xu. 2008. "R&D Investments, Exporting, and the Evolution of Firm Productivity." *American Economic Review* 98(2): 451–56.

Aw, Bee Yan, Mark J. Roberts, and Daniel Yi Xu. 2011. "R&D Investment, Exporting, and Productivity Dynamics." *American Economic Review* 101(4): 1312–44.

Baggs, Jen. 2005. "Firm Survival and Exit in Response to Trade Liberalization." *Canadian Journal of Economics* 38(4): 1364–83.

Baggs, Jen, and James Brander. 2006. "Trade Liberalization, Profitability, and Financial Leverage." *Journal of International Business Studies* 37(2): 196–211.

Baldwin, John R., Desmond Beckstead, and Richard Caves. 2002. "Changes in the Diversification of Canadian Manufacturing Firms (1973–1997): A Move to Specialization." Statistics Canada, Analytical Studies Branch Research Paper 2002179.

Baldwin, John R., Richard E. Caves, and Wulong Gu. 2005. "Responses to Trade Liberalization: Changes in Product Diversification in Foreign- and Domestic-Controlled Plants." Chap. 10 in *Governance, Multinationals and Growth,* edited by Lorraine Eden and Wendy Dobson. Cheltenham, UK: Edward Elgar Publishing.

Baldwin, John R., and Wulong Gu. 2003. "Export-Market Participation and Productivity Performance in Canadian Manufacturing." *Canadian Journal of Economics* 36(3): 634–57.

Baldwin, John R., and Wulong Gu. 2004. "Trade Liberalization: Export-Market Participation, Productivity Growth and Innovation." *Oxford Review of Economic Policy* 20(3): 372–92.

Baldwin, John R., and Wulong Gu. 2006. "Plant Turnover and Productivity Growth in Canadian Manufacturing." *Industrial and Corporate Change* 15(3): 417–65.

Baldwin, John R., and Wulong Gu. 2009. "The Impact of Trade on Plant Scale, Production-Run Length and Diversification." Chap. 15 in *Producer Dynamics: New Evidence from Micro Data,* edited by Timothy Dunne, J. Bradford Jensen, and Mark J. Roberts. University of Chicago Press.

Balistreri, Edward J., Russell H. Hillberry, and Thomas F. Rutherford. 2011. "Structural Estimation and Solution of International Trade Models with Heterogeneous Firms." *Journal of International Economics* 83(2): 95–108.

Bartelsman Eric J., John C. Haltiwanger, and Stefano Scarpetta. 2009. "Cross-Country Differences in Productivity: The Role of Allocation and Selection." NBER Working Paper 15490.

Beamish, Paul W., and Nikhil Celly. 2003. *Vincor and the New World of Wine.* HBR Case Study, *Harvard Business Review.*

Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum. 2003. "Plants and Productivity in International Trade." *American Economic Review* 93(4): 1268–90.

Bernard, Andrew B., and J. Bradford Jensen. 1995. "Exporters, Jobs, and Wages in U.S. Manufacturing: 1976–1987." *Brookings Papers on Economic Activity* Microeconomics: 67–112.

Bernard, Andrew B., and J. Bradford Jensen. 1999. "Exceptional Exporter Performance: Cause, Effect, or Both?" *Journal of International Economics* 47(1): 1–25.

Bernard, Andrew B., and J. Bradford Jensen. 2004. "Exporting and Productivity in the U.S." *Oxford Review of Economic Policy* 20(3): 343–57.

Bernard, Andrew B., J. Bradford Jensen, Stephen J. Redding, and Peter K. Schott. 2007. "Firms in International Trade." *Journal of Economic Perspectives* 21(3): 105–130.

Bernard, Andrew B., J. Bradford Jensen, and Peter K. Schott. 2006. "Trade Costs, Firms and Productivity." *Journal of Monetary Economics* 53(5): 917–37.

Bernard, Andrew B., Stephen J. Redding, and Peter K. Schott. 2011. "Multi-Product Firms and Trade Liberalization." *Quarterly Journal of Economics* 126(3): 1271–1318.

Bloom, Nicholas, Mirko Draca, and John Van Reenen. 2011. "Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity." NBER Working Paper 16717.

Brander, James A. 1991. "Election Polls, Free Trade, and the Stock Market: Evidence from the 1988 Canadian General Election." *Canadian Journal of Economics* 24(4): 827–43.

Broda, Christian, and David E. Weinstein. 2006. "Globalization and the Gains from Variety." *Quarterly Journal of Economics* 121(2): 541–85.

Brülhart, Marius. 2009. "An Account of Global Intra-industry Trade, 1962–2006." *World Economy* 32(3): 401–459.

Bustos, Paula. 2011a. "Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms." *American Economic Review* 101(1): 304–40.

Bustos, Paula. 2011b. "The Impact of Trade Liberalization on Skill Upgrading: Evidence from Argentina." http://www.crei.cat/people/bustos/Trade_Skill_PaulaBustos.pdf.

De Loecker, Jan. 2007. "Do Exports Generate

Higher Productivity? Evidence from Slovenia." *Journal of International Economics* 73(1): 69–98.

De Loecker, Jan. 2011. "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity." *Econometrica* 79(5): 1407–51.

Eberhardt, Markus, Christian Helmers, and Zhihong Yu. 2011. "Is the Dragon Learning to Fly? An Analysis of the Chinese Patent Explosion." CSAE Working Paper WPS/2011-15.

Eckel, Carsten, and J. Peter Neary. 2010. "Multi-Product Firms and Flexible Manufacturing in the Global Economy." *Review of Economic Studies* 77(1): 188–217.

Ethier, Wilfred. 1982. "National and International Returns to Scale in the Modern Theory of International Trade." *American Economic Review* 72(3): 389–405.

Feenstra, Robert C. 2010. *Product Variety and the Gains from International Trade.* Cambridge MA: MIT Press.

Goldberg, Pinelopi Koujianou, Amit Kumar Khandelwal, Nina Pavcnik, and Petia Topalova. 2010. "Imported Intermediate Inputs and Domestic Product Growth: Evidence from India." *Quarterly Journal of Economics* 125(4): 1727 –67.

Greenaway, David, and Richard Kneller. 2007. "Firm Heterogeneity, Exporting and Foreign Direct Investment." *Economic Journal* 117(517): 134–61.

Grossman, Gene M., and Elhanan Helpman. 1991. *Innovation and Growth in the Global Economy.* Cambridge, MA: MIT Press.

Halpern, László, Miklós Koren, and Adam Szeidl. 2005. "Imports and Productivity." CEPR Discussion Paper 5139.

Harris, Richard. 1984. "Applied General Equilibrium Analysis of Small Open Economies with Scale Economies and Imperfect Competition." *American Economic Review* 74(5): 1016–32.

Harrison, Ann E. 1994. "Productivity, Imperfect Competition and Trade Reform: Theory and Evidence." *Journal of International Economics* 36(1–2): 53–73.

Harrison, Glenn W., Thomas F. Rutherford, and Ian Wooton. 1989. "The Economic Impact of the European Community." *American Economic Review* 79(2): 288–94.

Helpman, Elhanan. 2004. *The Mystery of Economic Growth.* Cambridge, MA: Belknap Press of Harvard University Press.

Helpman, Elhanan. 2011. *Understanding Global Trade.* Cambridge, MA: Harvard University Press.

Helpman, Elhanan, and Paul R. Krugman. 1985. *Market Structure and Foreign Trade.* MIT Press.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics* 124(4): 1403–48.

Kasahara, Hiroyuki, and Beverly Lapham. 2012. "Productivity and the Decision to Import and Export: Theory and Evidence." February 13. http://faculty.arts.ubc.ca/hkasahara/workingpapers/productivity.pdf.

Kasahara, Hiroyuki, and Joel Rodrigue. 2008. "Does the Use of Imported Intermediates Increase Productivity? Plant-level Evidence." *Journal of Development Economics* 87(1): 106–118.

Krugman, Paul R. 1979. "Increasing Returns, Monopolistic Competition, and International Trade." *Journal of International Economics* 9(4): 469–79.

Krugman, Paul R. 1980. "Scale Economies, Product Differentiation, and the Pattern of Trade." *American Economic Review* 70(5): 950–59.

Levinsohn, James. 1993. "Testing the Imports-as-Market-Discipline Hypothesis." *Journal of International Economics* 35(1–2): 1–22.

Lileeva, Alla. 2008. "Trade Evidence and Productivity Dynamics: Evidence from Canada." *Canadian Journal of Economics* 41(2): 360–90.

Lileeva, Alla, and Daniel Trefler. 2010. "Improved Access to Foreign Markets Raises Plant-level Productivity...For Some Plants." *Quarterly Journal of Economics* 125(3): 1051–99.

Liu, Runjuan, and Daniel Trefler. 2011. "A Sorted Tale of Globalization: White Collar Jobs and the Rise of Service Offshoring." NBER Working Paper 17559.

López, Ricardo A. 2005. "Trade and Growth: Reconciling the Macroeconomic and Microeconomic Evidence." *Journal of Economic Surveys* 19(4): 623–48.

Mayer, Thierry, Marc J. Melitz, and Gianmarco I. P. Ottaviano. 2011. "Market Size, Competition, and the Product Mix of Exporters." NBER Working Paper 16959.

Mayer, Thierry, and Gianmarco I. P. Ottaviano. 2008. "The Happy Few: The Internationalisation of European Firms: New Facts Based on Firm-Level Evidence." *Intereconomics: Review of European Economic Policy* 43(3): 135–48.

Melitz, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71(6): 1695–1725.

Melitz, Marc J., and Gianmarco I. P. Ottaviano. 2008. "Market Size, Trade, and Productivity." *Review of Economic Studies* 75(1): 295–316.

OECD. 2002. *OECD Economic Outlook.* June 2002, no. 71.

Pavcnik, Nina. 2002. "Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants." *Review of Economic Studies* 69(1): 245–76.

Schmookler, Jacob. 1954. "The Level of Inventive Activity." *Review of Economics and Statistics* 36(2): 183–190.

Shen, Leilei. 2011. "Product Restructuring, Exports, Investment, and Growth Dynamics." September 24. http://individual.utoronto.ca/lshen/dynamic.pdf.

Thompson, Aileen J. 1993. "The Anticipated Sectoral Adjustment to the Canada–United States Free Trade Agreement: An Event Study Analysis." *Canadian Journal of Economics* 26(2): 253–71.

Topalova, Petia, and Amit Khandelwal. 2011. "Trade Liberalization and Firm Productivity: The Case of India." *Review of Economics and Statistics* 93(3): 995–1009.

Tybout, James R., and M. Daniel Westbrook. 1995. "Trade Liberalization and the Dimensions of Efficiency Change in Mexican Manufacturing Industries." *Journal of International Economics* 39(1–2): 53–78.

Trefler, Daniel. 2004. "The Long and Short of the Canada–U.S. Free Trade Agreement." *American Economic Review* 94(4): 870–95.

Van Biesebroeck, Johannes. 2005. "Exporting Raises Productivity in Sub-Saharan African Manufacturing Firms." *Journal of International Economics* 67(2): 373–91.

Verhoogen, Eric. 2008. "Trade, Quality Upgrading and Wage Inequality in the Mexican Manufacturing Sector." *Quarterly Journal of Economics* 123(2): 489–530.

Wagner, Joachim. 2007. "Exports and Productivity: A Survey of the Evidence from Firm-level Data." *World Economy* 30(1): 60–82.