IATRC

INTERNATIONAL AGRICULTURAL TRADE RESEARCH CONSORTIUM

Celebrating the First Thirty Years

Edited by

ALEX F. MCCALLA
UNIVERSITY OF CALIFORNIA, DAVIS

ED ROSSMILLER
RETIRED

LAURA BIPES
IATRC ADMINISTRATOR, UNIVERSITY OF MINNESOTA

CONTENTS

INTRODUCTION

At the December 2010 annual general meeting of the International Agricultural Trade Research Consortium (IATRC), the traditional Theme Day was organized as a celebration of the 30 year anniversary of that institution and was titled *Trade in Agriculture: So Much Done, So Much More to Do*. In the aftermath of that meeting a proposal was made that the 30 year history of the IATRC should be written while those who had lived through the full period were still available to provide the necessary institutional memory. The Executive Committee agreed and allocated $2000 to the project as a token of their serious support, while Alex McCalla, Ed Rossmiller and Laura Bipes agreed to see it to fruition. It soon became clear that in the tight fiscal environment of the time, further funding would not be forthcoming. Thus the team decided that if they did most of the work themselves they would be able to publish the results of their efforts as an e-book on the internet, but would not have the resources to produce any paper copies.

They also determined that in addition to the three major papers (unfortunately, the fourth major presentation by Valeria Csukasi, *Future Challenges in Agricultural Trade Negotiations*, is not available to us for inclusion in this manuscript) and the panel presentations at the 30th anniversary theme day, several other documents were available that detailed much of the rationale for the creation of the IATRC, its evolution and its output over the period.

The first of these documents is IATRC Objectives, Organization, Operations and Origins, the so called 'Blue Book', the latest edition of which is Edition VI dated April 2010. The Blue Book is a rolling record of the decisions taken at the meetings of the membership and the Executive Committee and a listing of the various outputs of the Consortium since its beginning. Since the Blue Book is revised and updated periodically and is publically available on the IATRC website

(http://iatrc.org/about/bluebook/BlueBook2010.pdf) it will only be referenced here as needed rather than being reproduced in its entirety.

The second of the documents is *An Analytical History of the IATRC* by Tim Josling, Alex McCalla and T. Kelley White, as requested by the Executive Committee and published in October 1997. It is reproduced here in its entirety.

Another pair of documents that add to the historical picture are the report dated December 2004 to the Executive Committee and the membership as requested by the IATRC Chair, Tim Josling, by the Futures Steering Group consisting of Mike Gifford, Joe Glauber, Stefan Tangermann, Linda Young and Alex McCalla, Chair, and the January 2011 Status Report on IATRC: Progress on Recommendations of the Futures Steering Group by the 2010 Executive Committee. These two documents are also reproduced in their entirety.

PART I

Theme Day

Trade in Agriculture: Much Done, So Much More to Do

CHAPTER 7

The Gains from International Trade Under Monopolistic Competition

ROBERT C. FEENSTRA
UNIVERSITY OF CALIFORNIA, DAVIS

1. INTRODUCTION

Let me thank the International Agricultural Trade Research Consortium, and Alex McCalla in particular, for the opportunity to talk today on the topic of the gains from trade under monopolistic competition. It is no exaggeration to say that this is a topic that has occupied me for my entire career, and will probably continue to do so. As background reading to my lecture today let me mentioned the Zeuthen Lectures that I presented at the University of Copenhagen in 2007, which were published as Feenstra (2010), and from which I draw quite freely. In addition, there is a very recent paper with David Weinstein (Feenstra and Weinstein, 2010). When describing these results I will try to bring out some of the connections to agriculture, though I suspect that my discussants will be able to do that even better than I.

The reason I say that I have been working on this topic my entire career is that I actually started when I was an undergraduate at the University of British Columbia, where I wrote my senior essay on the gains from trade with the United States. That was in 1977, more than a decade before the Canada-U.S. Free Trade Agreement of 1989 or the North America Free Trade Agreement or 1994. But back into the 1960s there was discussion by the Economic Council of Canada on the potential benefits of free trade with the United States. Canadian economists saw this as a way for local firms to expand their scale and move down their average cost curves, taking greater advantage of economies of scale. That discussion was limited by not having the modern—by which I mean mathematical—monopolistic competition model to work with. So instead it relied on a limit pricing model developed by two Canadian economists—Eastman and

Skykolt (1967). They demonstrated graphically that increased competition from the United States would lead Canadian firms to reduce markups and expand scale, while at the same time dropping inefficient product lines. These were the ideas that dominated the Canadian policy discussion, and ultimately contributed to the two free trade agreements, with the United States and then with Mexico, too.

I did not make headway on this topic for my senior essay, beyond reading the reporting on the literature, but I did become aware of the huge disconnect between this ideas from the Canadian literature, and the models that I learned about from the first edition of Caves and Jones textbook. The Ricardian and Heckscher-Ohlin models just did not have any room for increasing returns to scale or imperfect competition. I continued to be aware of this deficiency as I did my graduate work at MIT, and it was in my second year there—in 1978—that I saw an early draft of Paul Krugman's (1979) first paper on monopolistic competition and trade, drawing on the mathematical formulation of Dixit and Stiglitz. I immediately thought "this is it!" since Krugman's simple model has all the predictions that were anticipated but not proven in the Canadian literature. Several years ago I was asked to write an article about Krugman's contributions after he won the Nobel prize (Feenstra, 2009), and I recounted this personal experience from my undergraduate days. I sent the article around to my colleagues at Davis, and one of them responded that I honored Krugman by saying that two Canadians thought of it first!

So I was lucky enough to see the development of the monopolistic competition model in trade from the beginning. As you know, the static models of the 1980s gave rise to the dynamic models of the 1990s, and in this new millennium, models that emphasize heterogeneous firms, as I will discuss later. But my own interests have been in the empirical application of these models, and especially in the measurement of the gains from trade. In the monopolistic competition models there are in fact three sources of gains from trade that are not found in the conventional Heckscher-Ohlin model: consumer gains due to the expansion of import varieties; efficiency gains due to increasing returns to scale and/or improved productivity; and welfare gains due to reduced markups.[1]

I will begin by discussing first source of gains from trade, which is the welfare gains from new products in a monopolistic competition model. In the theoretical models those welfare gains would simply be captured by a rise in N, the number of products, but in reality we have to allow for demand to differ across varieties, some of which are demanded in greater quantity than others. So even if we stick with the popular CES framework, what are the welfare gains from new products?

2. CONSUMER BENEFITS FROM IMPORT VARIETY

As I first thought about how to measure the welfare gains from new products, I felt that I had a trick up my sleeve that would allow me to solve this question. That trick came from my other field of study as an undergraduate, besides international trade, which was duality theory. I had benefitted from several graduate courses taught by Erwin Diewert, and I knew that one could capture the consumer gains between two periods with an exact price or cost-of-living index (Diewert, 1976). Erwin always talked about the "new goods" problem, where if a good is not available in one period then we should use its reservation price in the index number formula. The challenge is that the reservation prices generally need to be solved for, which defeats the whole idea of using an exact price index as a short-cut to get welfare gains.

1 These sources are not mutually exclusive. In Krugman (1979), for example, the welfare gains due to reduced markups are identical to the gains from increasing returns to scale: as the scale of firms expands due to trade, average cost is reduced and the gap between price=average cost and marginal cost is also reduced.

But there was a good reason that this price index approach might work for the CES case, namely, that the reservation prices for goods are infinite, so we don't actually need to solve for them. With many goods, the elasticity of demand is approximately equal to the elasticity of substitution, or σ. So a typical demand curve for this utility function is of the form $q = kp^{-\sigma}$, where q denotes quantity, p denotes price, and $k > 0$ is a constant. This demand curve is illustrated in Figure 2.1, and approaches the vertical axis as the price approaches infinity, so the reservation prices of the good is *infinite*. But provided that the elasticity of substitution is greater than unity, then the area under the demand curve is bounded above, and the ratio of areas A/B in Figure 2.1 is easily calculated as $A/B = 1/(\sigma - 1)$. Thus, even with an infinite reservation price, there is a well-defined area of consumer surplus from having the new good available, and measuring this area depends on having an estimate of the elasticity of substitution.

The challenge is to generalize this one-good example to a case where many new goods are potentially available from trade. To address that case, we do not rely on consumer surplus to measure the welfare gain, as in Figure 2.1, but rather, take the ratio of the CES expenditure functions—dual to the utility function—to derive an exact cost of living index for the consumer. By determining how new goods affect the cost of living index, we will have obtained an expression for the welfare gain from the new products. After solving this problem, we then apply the results to the monopolistic competition model of Krugman (1980).

CES Utility Function

We will work with the non-symmetric CES function,

$$U_t = U(q_t, I_t) = \left[\sum_{i \in I_t} a_{it} q_{it}^{(\sigma-1)/\sigma} \right]^{\sigma/(\sigma-1)}, \quad \sigma > 1, \tag{2.1}$$

where $a_{it} > 0$ are taste parameters that vary with time, and $I_t \subseteq \{1, \ldots, N\}$ denotes the set of goods available in period t at the prices p_{it}. The minimum expenditure to obtain one unit of utility is,

$$e(p_t, I_t) = \left[\sum_{i \in I_t} b_{it} p_{it}^{1-\sigma} \right]^{1/(1-\sigma)}, \quad \sigma > 1, \quad b_{it} \equiv a_{it}^{\sigma}. \tag{2.2}$$

For simplicity, first consider the case where $I_{t-1} = I_t = I$, so there is no change in the set of goods, and also $b_{it-1} = b_{it}$, so there is no change in tastes. We assume that the observed purchases q_{it} are optimal for the prices and utility, that is, $q_{it} = U_t(\partial e/\partial p_{it})$. Then the index number due to Sato (1976) and Vartia (1976) shows us how to measure the ratio of unit-expenditure, or the change in the cost of living for the representative consumer:

Theorem 2.1 (Sato, 1976; Vartia, 1976)

If the set of goods available is fixed at $I_{t-1}=I_t=I$, taste parameters are constant, $b_{it-1}=b_{it}$, and observed quantitites are optimal, then:

$$\frac{e(p_t,I)}{e(p_{t-1},I)} = P_{SV}(p_{t-1},p_t,q_{t-1},q_t,I) \equiv \prod_{i\in I}\left(\frac{p_{it}}{p_{it-1}}\right)^{w_i(I)}, \tag{2.3}$$

where the weights $w_i(I)$ are constructed from the expenditure shares $s_{it}(I) \equiv p_{it}q_{it} \Big/ \sum_{i\in I} p_{it}q_{it}$ as,

$$w_i(I) \equiv \left(\frac{s_{it}(I)-s_{it-1}(I)}{\ln s_{it}(I)-\ln s_{it-1}(I)}\right) \Bigg/ \sum_{i\in I}\left(\frac{s_{it}(I)-s_{it-1}(I)}{\ln s_{it}(I)-\ln s_{it-1}(I)}\right). \tag{2.4}$$

The numerator in (2.4) is the "logarithmic mean" of the shares $s_{it}(I)$ and $s_{it-1}(I)$, and lies in-between these two shares, while the denominator ensures that the weights $w_i(I)$ sum to unity. The special formula for these weights in (2.4) is needed to precisely measure the ratio of unit-expenditures in (2.3), but in practice the Sato-Vartia formula will give very similar results to using other weights, such as $w_i(I) = \frac{1}{2}[s_{it}(I)+s_{it-1}(I)]$, as used for the Törnqvist price index, for example. In both cases, the geometric mean formula in (2.3) applies. The important point from Theorem 2.1 is that goods with high taste parameters a_i will also tend to have high weights, so even without knowing the true values of a_i, the exact ratio of unit-expenditures is obtained.

Now consider the case where the set of goods is changing over time, but some of the goods are available in both periods, so that $I_{t-1} \cap I_t \neq \varnothing$. We again let $e(p,I)$ denote the unit-expenditure function defined over the goods within the set I, which is a non-empty subset of those goods available both periods, $I \subseteq I_{t-1} \cap I_t \neq \varnothing$. We sometimes refer to the set I as the "common" set of goods. Then the ratio $e(p_t,I)/e(p_{t-1},I)$ is still measured by the Sato-Vartia index in the above theorem. Our interest is in the ratio $e(p_t,I_t)/e(p_{t-1},I_{t-1})$, which can be measured as follows:

Theorem 2.2 (Feenstra, 1994)

Assume that $b_{it-1}=b_{it}$ for $i \in I \subseteq I_{t-1} \cap I_t \neq \varnothing$, and that the observed quantities are optimal. Then for $\sigma > 1$:

$$\frac{e(p_t,I_t)}{e(p_{t-1},I_{t-1})} = P_{SV}(p_{t-1},p_t,q_{t-1},q_t,I)\left(\frac{\lambda_t(I)}{\lambda_{t-1}(I)}\right)^{1/(\sigma-1)}, \tag{2.5}$$

where the weights $w_i(I)$ are constructed from the expenditure shares $s_{it}(I) \equiv p_{it}q_{it} \Big/ \sum_{i\in I} p_{it}q_{it}$ as in (2.4), and the values $\lambda_t(I)$ and $\lambda_{t-1}(I)$ are constructed as:

$$\lambda_\tau(I) = \left(\frac{\sum_{i\in I} p_{i\tau}q_{i\tau}}{\sum_{i\in I_\tau} p_{i\tau}q_{i\tau}}\right) = 1 - \left(\frac{\sum_{i\in I_\tau, i\notin I} p_{i\tau}q_{i\tau}}{\sum_{i\in I_\tau} p_{i\tau}q_{i\tau}}\right), \tau=t-1,t. \tag{2.6}$$

Each of the terms $\lambda_\tau(I) \leq 1$ can be interpreted as the *period τ expenditure on the good in the common set I, relative to the period τ total expenditure*. Alternatively, this can be interpreted as *one minus the period τ expenditure on "new" goods*

(not in the set I), relative to the period τ total expenditure. When there are more new goods in period t, this will tend to lower the value of $\lambda_t(I)$, which leads to a greater fall in the ratio of unit costs in (2.5), by an amount that depends on the elasticity of substitution.

The importance of the elasticity of substitution can be seen from Figure 2.2, where we suppose that the consumer minimizes the expenditure needed to obtain utility along the indifference curve AD. If initially only good 1 is available, then the consumer chooses point A with the budget line AB. When good 2 becomes available, the same level of utility can be obtained with consumption at point C. Then the drop in the cost of living is measured by the inward movement of the budget line from AB to the line through C, and this shift depends on the convexity of the indifference curve, or the elasticity of substitution.

Measuring the Elasticity of Substitution

We can see that to measure the gains from new goods it is essential to have a good estimate of the elasticity of substitution. Existing estimates at the time of my 1994 article tended to be too low, which would lead to exaggerated estimates of the gains from trade. The reason that these estimates were so low, I believe, was due to the standard simultaneous equations bias: the elasticity of demand cannot be estimated in a demand and supply system without instrumental variables that are orthogonal to the error terms. But in international trade, we are interested in estimating the elasticity of substitution *between source countries for each good*; in other words, we want to measure the Armington (1969) elasticity between source countries. It is difficult if not impossible to find instruments that can be used in every market and country. Feenstra (1994) proposed a method to resolve this problem that makes use of the panel nature of datasets in international trade, i.e. having time-series observations on the amount imported from multiple source countries.

I won't go into the technical details of this new estimator, except to say that it exploits the panel nature of most international trade dataset: many exporting countries supplying a good to say the U.S. over multiple years. I relied on uncorrelated errors from the CES demand and supply curves across countries, and used that moment condition to develop a Generalized Method of Moments (GMM) estimator. In place of conventional instruments one can think of using indicator variables across countries, and identification relies on heteroskedasticity in the data, or what Roberto Rigobon refers to as "identification through heteroskedasticity." The estimator is not hard to implement: it becomes a linear regression across exporting countries, but run on variables that are various second-moments of the data for each country.

In practice, the elasticities of substitution obtained from this method are very plausible. Feenstra (1994) considered products like mean's leather athletic shoes, or cotton knit shirts, or various types of steel, and obtained estimates of the elasticity between 3 and 8. These were much higher values than obtained previously, and more in line with what trade economists would expect. Feenstra even added gold bullion and silver bullion as additional test case, and obtained estimates of the elasticity of substitution for each of these products of 25 and 40, respectively. Those high estimates are essentially infinite, indicating that there is perfect substitution between country sources of gold or silver. To conclude, for the six products analyzed by Feenstra (1994), the new method for estimating the elasticity of substitution worked very well indeed.

At about the same time as publishing this article, I worked on another project that dealt with agriculture and made use of the GMM estimator. This project was joint with Andrew Rose (Feenstra and Rose, 1993), and I relied heavily on advice from my colleagues at UC Davis. Andy and I were asked to write a paper on the prospective impact of NAFTA on

California agriculture, and we decided to focus on the distorted market for water in California. We reasoned that with NAFTA allowing the import of fresh fruits and vegetables from Mexico, that would water use in California. Our calculations suggested that thee gains from reduced water use would be at least as large as the conventional gains from trade. But to estimate the gains, I needed to have Armington elasticities for the various crops, so I applied my GMM estimator to asparagus and other vegetables. I recall that one of the discussants at the conference said that throwing a GMM estimator at asparagus was a bit like using a sledgehammer to swat a flee, but still, it worked well enough!

Gain from Import Variety for the United States

Broda and Weinstein (2006) apply the above methods to measure the gains from trade for the U.S. They define a *good* as either a 10-digit Harmonized System (HS) category, over the period 1990–2001, or as a 7-digit Tariff Schedule of the United States (TSUSA) category, for the earlier period 1972–1988.[2] The imports from various source countries are the *varieties* available for each good. The ratio $(\lambda_t/\lambda_{t-1})$ is constructed for each good, using the expenditure on new and disappearing source countries. In addition, they estimate σ for each good. For the TSUSA categories of goods, they estimate roughly 12,000 values for σ, with a median value of 3.6. For the HS system, they estimate about 14,500 value for σ, with a median value of 2.9. In both cases the distribution of elasticities is highly skewed towards the right (so the mean values are much larger than the medians).

Putting together the ratios $(\lambda_t/\lambda_{t-1})$ for each good with the elasticity of substitution, they obtain $(\lambda_t/\lambda_{t-1})^{-1/(\sigma-1)}$, which is aggregated over all goods. For the TSUSA data they used 1972 as the base year and measured the gains from new supplying countries up to 1988, and then for the HS data they used 1990 as the base year and measured the gains from new supplying countries up to 2001. Summing these, they obtain an estimate of the gains from trade for the U.S. due to the expansion of import varieties, which amount to 2.6% of GDP in 2001.

Two features of Broda and Weinstein's methods deserve special mention. First, by measuring the expenditure on new supplying countries relative to a base year, they are following the hypothesis of Theorem 2.2 that the "common" set of countries should be those with *constant* taste parameters. In contrast, when countries first start exporting goods, it is reasonable to expect that the demand curve in the importing country shifts out over some number of years as consumers become informed about the product. Broda and Weinstein are allowing for such shifts for new and disappearing countries after the base year, and all such changes in demand for these countries are incorporated into the λ_τ terms in Theorem 2.2. That is the correct way to measure the gains from new import varieties.[3]

Second, Broda and Weinstein (2006) did not incorporate any changes in the number of U.S. varieties into their estimation, nor did they include the U.S. as a source country in the estimation of the elasticity of substitution for each good. That is the correct approach only under the limited case where *the number of U.S. varieties is constant*. While turns out to be true under our assumptions in the model of Krugman (1980), it is certainly not the case in more general models: we could expect that increases in import variety would result in some reduction in domestic varieties. In that case, the gains from import varieties would be offset by the welfare loss from reduced domestic varieties. That potential loss was addressed only briefly by Broda and Weinstein (2006), and we shall address it more fully in the next section.

2 1989 is omitted because West and East Germany unified then, making comparisons with later years difficult.

3 In addition, countries that are suspected of selling a changing range of product varieties within each HS good should be excluded from the set I, and instead included in the $\lambda\tau$ terms.

3. PRODUCTIVITY GAINS DUE TO THE SELECTION OF FIRMS

I mentioned above the problem that increased product variety of imports can be expected to drive out some domestic producers, so the overall gains have to take into account these opposed forces. The simplest CES model due to Krugman actually do not have this feature: as more import variety come in, the range of domestic varieties does not change. But that prediction seems unrealistic, so we want to go beyond that simple framework. Such a generalization is achieved by the monopolistic competition model due to Melitz (2003). That model has heterogeneous firms with differing productivities, which are random draws from a probability distribution. The firms learn their productivities only after paying a fixed cost to enter the market, and then they can decide whether to stay in the market and sell domestically, or pay an additional fixed cost and export. Like a conventional monopolistic competition model the Melitz model has zero profits in equilibrium, but in this case it is zero *expected* profits, computed before firm know their productivity draws. The export profits are in fact positive for those firms that stay in the market (except for the borderline firm that has zero profits).

It is known in the Melitz model that as import varieties come into a country, the range of domestic varieties are reduced. In fact, the reduction in domestic varieties can be strong enough that total product variety also falls. But regardless of whether total product variety rises or falls, I have shown in Feenstra (2010) that the welfare effects of the increased import variety and reduced domestic variety exactly cancel each other out: there is in fact no welfare gains due to import variety on the consumption side, because of the reduction in domestic varieties. So it would appear that the analysis of Broda and Weinstein, who measured welfare gains of just under 1% of GDP from increased import variety, might not stand up in a Melitz-style model, which is a very disturbing result.

Fortunately, it turns out that there is another source of gains in the Melitz model, which is still related to product variety, but now on the export side rather than the import side. This brings me to the second source of gains from trade in a monopolistic competition model, which are the productivity gains due to the self-selection of more efficient firms into exporting. In a Melitz-style model with heterogeneous firms, we capture the productivity gains from the selection of firms, with less efficient firms leaving the market and more efficient firms growing in scale. In fact, for the Canada-U.S. Free Trade Agreement, it is this self-selection of firms that appears to be most important in explaining the gains to Canadian productivity, as argued by Trefler (2004).

The question for us is: how can these productivity gains due to the self-selection of firms be measured? Trefler had access to firm-level data for Canada before and after the free trade agreement, in which case one would simply track which firms exited and which firms expanded in scale. But suppose, on the contrary, that we do not have any firm-level data. Then is there some method that will enable us to capture the gains from trade due to this productivity effect? It turns out that there is, and in fact, it is quite analogous to the gains from trade on the consumer side that I have already described. But now instead of focusing on the consumer gains from variety, we instead need to focus on the gains from *output* variety.

Output Variety in the CES Case

To theoretically demonstrate these gains in the Melitz (2003) model, we approach the problem in a somewhat unusual way: rather than focusing on consumer gains as in the last section we now ask whether there exist any *producer* gains due to output variety. This question can be answered by extending the range of values for the elasticity of substitution that we considered in the previous chapter. There we restricted out attention to $\sigma > 1$ in the utility and expenditure functions (2.1) and (2.2), but a wider range of values for this elasticity can be considered. In particular, if $\sigma < 0$ then

instead of obtaining convex indifference curves from (2.1) for a fixed level of U_t, we obtain a concave transformation curve as shown in Figure 3.1.[4] The parameter U_t in this case measures the resources devoted to production of the goods q_{it}, $i \in I_t$, and the elasticity of the transformation curve (measured as a positive number) is $-\sigma$.

To make this reinterpretation explicit, when $\sigma < 0$ we will denote its positive value by $\omega \equiv -\sigma$, which is the elasticity of transformation. Then we will rewrite (2.1) using labor resources L_t to replace utility U_t, obtaining:

$$L_t = \left(\sum_{i \in I_t} a_{it} q_{it}^{(\omega+1)/\omega} \right)^{\omega/(\omega+1)}, \; a_{it} > 0, \; \omega > 0. \tag{3.1}$$

The maximum revenue obtained using one unit of labor resources, dual to (2.9), is denoted by:

$$\psi(p_t, I_t) = \left[\sum_{i \in I_t} b_{it} p_{it}^{\omega+1} \right]^{1/(\omega+1)}, \qquad \mathrm{b_{it}} \equiv a_{it}^{-\omega}, \; \omega > 0. \tag{3.2}$$

With this reinterpretation, Theorem 2.2 continues to hold as:

$$\frac{\psi(p_t, I_t)}{\psi(p_{t-1}, I_{t-1})} = P_{SV}(p_{t-1}, p_t, q_{t-1}, q_t, I) \left(\frac{\lambda_t(I)}{\lambda_{t-1}(I)} \right)^{-1/(\omega+1)}, \tag{3.3}$$

where the exponent appearing on $(\lambda_t/\lambda_{t-1})$ is now negative. In other words, the appearance of "new outputs," so that $\lambda_t < 1$, will *raise revenue* on the producer side of the economy.

To understand where this increase in revenue is coming from, consider the transformation curve in Figure 3.1. If only good 1 is available, then the economy would be producing at the corner A, with revenue shown by the line AB. Then if good 2 becomes available to producers, the new equilibrium will be at point C, with an *increase* in revenue. This illustrates the benefits of output variety. In Figure 3.2 we illustrate the same idea in a partial equilibrium diagram, for a supply curve with constant elasticity ω. When the good becomes available for production, there is an effective price *increase* from the reservation price for producers (which is zero with a constant-elasticity supply curve) to the actual price. The gain in producer surplus is area C, and measured relative to total sales $C+D$, we can readily compute that $C/(C+D) = 1/(\omega+1)$.

While this reinterpretation of our earlier consumer model is mathematically valid, there is a problem in its application to international trade: the transformation curve between two outputs is often taken to be *linear* rather than strictly concave. That is the case in the Ricardian model, for example, or in the transformation curve (1.8) in Krugman's (1980) model. In that case, the gains from output variety would vanish. So the question arises as to whether the strictly concave case we illustrate in Figure 3.1 has any practical application?

We will now argue that a strictly concave transformation curve is indeed relevant, and in fact, arises in the generalization of the monopolistic competition model due to Melitz (2003), which allows for heterogeneous firms that differ in

4 Notice that the range cannot be considered, since then all goods are *essential* in (2.1), with a zero quantity for any single good resulting in zero utility. In that case the welfare gain from a new good is infinite.

their productivities ϕ. In the equilibrium with zero expected profits, only firms above some cutoff productivity ϕ^* survive; and of these, only firms with productivities above $\phi_x^* > \phi^*$ actually export. We will argue that the endogenous determination of these cutoff productivities leads to a strictly concave constant-elasticity transformation curve between domestic and export varieties, adjusted for the quantity produced of each. Furthermore, this interpretation of the Melitz (2003) model will allow us to make a precise calculation of the producer gains due to *export* variety.

Monopolistic Competition with Heterogeneous Firms

We outline here a two country version of the Melitz (2003) model that does not assume symmetry across the countries. We focus on the home country H, while denoting foreign variables with the superscript F. At home there is a mass of M firms operating in equilibrium. Each period, a fraction δ of these firms go bankrupt and are replaced by new entrants. Each new entrant pays a fixed cost of f_e to receive a draw φ of productivity from a cumulative distribution $G(\varphi)$, which gives rise to the marginal cost of w/φ, where w is the wage and labor is the only factor of production. Only those firms with productivity above a cutoff level φ^* find it profitable to actually produce (the cutoff level will be determined below). Letting M_e denote the mass of new entrants, then $[1-G(\varphi^*)]M_e$ firms successfully produce. In a stationary equilibrium, these should replace the firms going bankrupt, so that:

$$[1-G(\phi^*)]M_e = \delta M \quad . \tag{3.4}$$

Conditional on successful entry, the distribution of productivities for home firms is then:

$$\mu(\phi) = \begin{cases} \dfrac{g(\phi)}{[1-G(\phi^*)]} & if \;\; \phi \geq \phi^*, \\ \\ 0 & \text{otherwise}, \end{cases} \tag{3.5}$$

where $g(\phi) = \partial G(\phi)/\partial\phi$ is the density function.

Home and foreign consumers both have CES preferences that are symmetric over product varieties. Given home expenditure of wL, the revenue earned by a home firm from selling at the price $p(\varphi)$ is:

$$r(\phi) = p(\phi)q(\phi) = \left[\frac{p(\phi)}{P^H}\right]^{1-\sigma} wL \,, \sigma > 1 \,, \tag{3.6}$$

where $q(\varphi)$ is the quantity sold and P^H is the home CES price index. The profit-maximizing price from selling in the domestic market is the usual constant markup over marginal costs:

$$p(\phi) = \left(\frac{\sigma}{\sigma-1}\right)\frac{w}{\phi} \,. \tag{3.7}$$

Using this, we calculate variable profits from domestic sales as $r(\varphi) - (w/\varphi)q(\varphi) = r(\varphi)/\sigma$. The lowest productivity firm that just breaks even in the domestic market there satisfies the *zero-cutoff-profit* (ZCP) condition:

$$r(\phi^*)/\sigma = wf \quad \Rightarrow \quad q(\phi^*) = (\sigma - 1)f\phi^*, \tag{3.8}$$

where f is the fixed labor cost. Note that this cutoff condition for the marginal firm is identical to what is obtained in Krugman's (1980) model for all firms, as in (2.7).

While firms with productivities $\varphi \geq \varphi^*$ find it profitable to produce for the domestic market, only those with higher productivities $\varphi \geq \varphi_x^* > \varphi^*$ find it profitable to export. A home exporting firm faces the iceberg transport costs of $\tau \geq 1$ meaning that τ units must be sent in order for one unit to arrive in the foreign country. Letting $p_x(\varphi)$ and $q_x(\varphi)$ denote the price received and quantity shipped at the factory-gate, the revenue earned by the exporter is:

$$r_x(\varphi) = p_x(\varphi)q_x(\varphi) = \left[\frac{p_x(\varphi)\tau}{P^F}\right]^{1-\sigma} w^* L^*, \tag{3.9}$$

where P^F is the CES price index in the foreign country, and w^*L^* is foreign expenditure.

Again, the optimal export price is a constant markup over marginal costs:

$$p_x(\varphi) = \left(\frac{\sigma}{\sigma - 1}\right)\frac{w}{\varphi}. \tag{3.10}$$

The variable profits from export sales are therefore $r_x(\varphi) - (w/\varphi)q_x(\varphi) = r_x(\varphi)/\sigma$, so the ZCP condition for the exporting firm is:

$$r_x(\varphi_x^*)/\sigma = wf_x \quad \Rightarrow \quad q_x(\varphi_x^*) = (\sigma - 1)f_x\varphi_x^*, \tag{3.11}$$

where f_x is the additional fixed labor cost for exporting. Provided that $r_x(\varphi)/f_x < r(\varphi)/f$, which we assume is the case, then the cutoff productivity for the exporting firm will exceed that for the domestic firm, $\varphi_x^* > \varphi^*$. Then the mass of exporting firms is computed as:

$$M_x \equiv \int_{\varphi_x^*}^{\infty} M\mu(\varphi)d\varphi < M. \tag{3.12}$$

To close the model, we use the full employment condition and also zero expected profits for any entrant. The labor needed for domestic sales for a firm with productivity φ is $[q(\varphi)/\varphi + f]$, and for export sales is $[q_x(\varphi)/\varphi + f_x]$, so the full employment condition is:

$$L = M_e f_e + M\int_{\varphi^*}^{\infty}[q(\varphi)/\varphi + f]\mu(\varphi)\,d\varphi + M_x\int_{\varphi_x^*}^{\infty}[q_x(\varphi)/\varphi + f_x]\mu_x(\varphi)\,d\varphi\,, \tag{3.13}$$

where the distribution of productivities conditional on exporting is $\mu_x(\varphi) \equiv g(\varphi)/[1-G(\varphi_x^*)]$ if $\varphi \geq \varphi_x^*$, and zero otherwise. We can rewrite (3.13) by multiplying by w, and using the fact that $(w/\varphi)q(\varphi) = r(\varphi)(\sigma-1)/\sigma$, and likewise for exporters, to obtain:

$$\begin{aligned} wL &= w\left(M_e f_e + Mf + M_x f_x\right) + \left(\frac{\sigma-1}{\sigma}\right)\left[M\int_{\phi^*}^{\infty} r(\phi)\mu(\phi)\,d\phi + M_x\int_{\phi_x^*}^{\infty} r_x(\phi)\mu_x(\phi)\,d\phi\right] \\ &= w\left(M_e f_e + Mf + M_x f_x\right) + \left(\frac{\sigma-1}{\sigma}\right)wL, \end{aligned}$$

where the second line is obtained using the definition of *GDP*, with zero expected profits. It follows immediately that there is a *linear* transformation curve between the mass of entering, domestic and exporting firms, that is:

$$L = \sigma\left(M_e f_e + Mf + M_x f_x\right). \tag{3.14}$$

To obtain further results, we assume a Pareto distribution for productivities:

$$G(\phi) = 1 - \phi^{-\theta}\text{, with } \theta > \sigma - 1 > 0\,. \tag{3.15}$$

In that case, it can be shown that the number of entering firms is proportional to the labor force, $M_e = L(\sigma-1)/\sigma\theta f_e$, which was assumed by Chaney (2008), for example. So the transformation curve between domestic and export varieties is further simplified as:

$$L = \frac{\sigma\theta}{(\theta-\sigma+1)}\left(Mf + M_x f_x\right). \tag{3.16}$$

The fact that this transformation curve is linear between the mass of domestic and exported varieties is similar to that found in the Krugman (1980) model. But this fact does not tell us about the transformation curve between the economy's outputs, because we also need to take into account the *quantity* produced of each variety. In Krugman's model, the quantity produced by each firm is fixed. But in the Melitz (2003) model, only the zero-profit-cutoff firm has output identical to that in Krugman's model, and the cutoff productivity φ^* itself is endogenously determined. So to determine the transformation curve for the economy, we first need to determine the correct measure of output used to adjust the varieties M and M_x.

To determine the appropriate measure of quantity, it is convenient to invert the demand curve and treat revenue as a function of quantity, so from (3.6) we obtain:

$$r(\phi) = A_d\, q(\phi)^{\frac{\sigma-1}{\sigma}}, \text{ where } A_d \equiv P^H \left(\frac{wL}{P^H}\right)^{\frac{1}{\sigma}}. \tag{3.17}$$

We introduce the notation A_d as shift parameter in the demand curve facing home firms for their domestic sales. It depends on the CES price index P^H, and also on domestic expenditure wL.

Likewise, export revenue can be written as:

$$r_x(\phi) = A_x q_x(\phi)^{\frac{\sigma-1}{\sigma}}, \text{ where } A_x \equiv \left(\frac{P^F}{\tau}\right)\left(\frac{\tau_i w * L *}{P^F}\right)^{\frac{1}{\sigma}}. \tag{3.18}$$

Integrating domestic and export revenue over firms, we obtain *GDP*:

$$wL = A_d M \int_{\phi^*}^{\infty} q(\phi)^{\frac{\sigma-1}{\sigma}} \mu(\phi) d\phi + A_x M_x \int_{\phi_x^*}^{\infty} q_x(\phi)^{\frac{\sigma-1}{\sigma}} \mu_x(\phi) d\phi. \tag{3.19}$$

Thus, in order to measure *GDP* the mass of domestic and export varieties are multiplied by the quantities shown above. Feenstra and Kee (2008) demonstrate that the first-order conditions for maximizing GDP subject to the resource constraint for the economy, taking A_d and A_x as given, are precisely the monopolistic competition equilibrium conditions. So the quantities appearing in (3.19) are the "right" way to adjust the mass of domestic and export varieties.

We can simplify these quantities by noting that CES demand, combined with constant-markup prices in (3.7), imply that the quantity sold equals $q(\varphi) = (\varphi / \tilde{\varphi})^{\sigma} q(\tilde{\varphi})$ for any choice of reference productivity $\tilde{\varphi}$. We follow Melitz (2003) in specifying $\tilde{\varphi}$ as *average* productivity:

$$\tilde{\varphi} \equiv \left[\int_{\varphi^*}^{\infty} \varphi^{(\sigma-1)} \mu(\varphi) d\varphi\right]^{1/(\sigma-1)}, \tag{3.20}$$

and likewise for the average productivity $\tilde{\varphi}_x$ for exporters, computed using φ_x^* and μ_x. It follows that GDP simply equals $(A_d \tilde{M} + A_x \tilde{M}_x)$, using the adjusted mass of varieties:

$$\tilde{M} \equiv M q(\tilde{\varphi})^{(\sigma-1)/\sigma} \quad \text{and} \quad \tilde{M}_x \equiv M_x q_x(\tilde{\varphi}_x)^{(\sigma-1)/\sigma}. \tag{3.21}$$

To simplify the expression for GDP further, we note that a property of the Pareto distribution is that an integral like (3.20) is always a constant multiple of the lower bound of integration. That is:

$$\tilde{\varphi} = \left[\frac{\theta}{(\theta - \sigma + 1)}\right]^{1/(\sigma-1)} \varphi^*, \tag{3.22}$$

as obtained by evaluating the integral in (3.20), which is finite provided that $\theta > \sigma - 1$. The cutoff productivity φ^* is in turn related to the mass of firms by $[1 - G(\varphi^*)]M_e = \delta M$, and using the mass of entering firms $M_e = L(\sigma - 1)/\sigma\theta f_e$ and the Pareto distribution, it follows that:

$$(\varphi^*)^{-\theta} = \frac{\delta\sigma\theta f_e}{L(\sigma - 1)} M. \tag{3.23}$$

Gathering together these results, we can compute that the adjusted mass of domestic varieties is:

$$\tilde{M} = M\left(\frac{\tilde{\varphi}}{\varphi^*}\right)^{\sigma-1} q(\varphi^*)^{\frac{\sigma-1}{\sigma}} = \frac{\theta M}{(\theta - \sigma + 1)}[(\sigma - 1) f \varphi^*]^{\frac{\sigma-1}{\sigma}} = k_1 f^{\frac{\sigma-1}{\sigma}} M^{1-\frac{\sigma-1}{\theta\sigma}} \left(\frac{f_e}{L}\right)^{-\frac{(\sigma-1)}{\theta\sigma}},$$

where the first equality follows from (3.21) and $q(\tilde{\varphi}) = (\tilde{\varphi}/\varphi^*)^{\sigma} q(\varphi^*)$; the second equality uses (3.22) and the ZCP condition $q(\varphi^*) = (\sigma - 1) f \varphi^*$; and the third follows from (3.23), where $k1 > 0$ depends on the parameters θ, σ and δ. Thus, the adjusted mass of domestic varieties is an increasing but *nonlinear* function of the mass M.

A similar expression holds for exports, but replacing f, M, and $\tilde{M}$ with f_x, M_x, and $\tilde{M}_x$. Solving for M and M_x and substituting these into the linear transformation curve (3.16), we obtain a concave transformation curve between $\tilde{M}$ and $\tilde{M}_x$, with elasticity $\omega \equiv \frac{\theta\sigma}{(\sigma-1)} - 1 > 0$:

$$L = k_2 f_e^{1/(\omega+1)} \left(\tilde{M}^{\frac{\omega+1}{\omega}} f^{1+\frac{(\omega+1)(\sigma-1)}{\omega\sigma}} + \tilde{M}_x^{\frac{\omega+1}{\omega}} f_x^{1+\frac{(\omega+1)(\sigma-1)}{\omega\sigma}}\right)^{\omega/(\omega+1)}, \tag{3.24}$$

where $k2 > 0$ again depends on the parameters θ, σ and δ.

Summing up, from the Melitz (2003) model we have obtained a constant-elasticity transformation curve, with elasticity $\omega \equiv \frac{\theta\sigma}{(\sigma-1)} - 1 > 0$. Our earlier results in Theorems 2.1 and 2.2 continue to apply to this transformation curve. In particular, consider the problem of maximizing $(A_d\tilde{M} + A_x\tilde{M}_x)$ subject to this transformation curve. This Lagrangian problem leads to the following solution:

Theorem 3.1 (Feenstra and Kee, 2008)

Assume that the distribution of firm productivity in Pareto, as in (3.15). Then maximizing *GDP* subject to the transformation curve (3.24) results in *GDP*= $\psi(A_d, A_x)L$, where:

$$w = \psi(A_d, A_x) \equiv \frac{1}{k_2 f_e^{1/(\omega+1)}}\left[A_d^{\omega+1} f^{1-\frac{\theta}{(\sigma-1)}} + A_x^{\omega+1} f_x^{1-\frac{\theta}{(\sigma-1)}}\right]^{\frac{1}{(\omega+1)}}. \tag{3.25}$$

The function $\psi(A_d, A_x)$ is the revenue earned with $L=1$ on the transformation curve, and equals wages. Note that the exponents appearing on the fixed costs f and f_x in (3.25) are obtained as $-[\omega + (1+\omega)\frac{(\sigma-1)}{\sigma}] = 1 - \frac{\theta}{(\sigma-1)} < 0$. This expression also appears as the exponent on fixed costs in the gravity equation of Chaney (2008).

We can now apply Theorem 2.2 to compute the gain from trade. Denoting autarky by $t-1$, the economy is at the corner of the transformation curve with $A_{xt-1} = \tilde{M}_{xt-1} = 0$, as illustrated by point A in Figure 3.3. Using t to denote the trade situation, under free trade we have $A_{xt} > 0 \; and \; \tilde{M}_{xt} > 0,$ as at point C. We can therefore evaluate the gain from trade as the ratio of real wages in trade and under autarky:

$$\begin{aligned}
\frac{w_t / P_t^H}{w_{t-1} / P_{t-1}^H} &= \frac{\psi(A_{dt}, A_{xt})}{\psi(A_{dt-1}, 0)}\left(\frac{P_t^H}{P_{t-1}^H}\right)^{-1} \\
&= \left(\frac{A_{dt}}{A_{dt-1}}\right)\left(\frac{R_{dt}}{w_t L_t}\right)^{\frac{-1}{\omega+1}}\left(\frac{P_t^H}{P_{t-1}^H}\right)^{-1} \\
&= \left(\frac{w_t / P_t^H}{w_{t-1} / P_{t-1}^H}\right)^{\frac{1}{\sigma}}\left(\frac{R_{dt}}{w_t L_t}\right)^{\frac{-1}{\omega+1}}
\end{aligned} \tag{3.26}$$

where the first line follows from wages in Theorem 2.1; the second line follows from Theorem 2.2, using the domestic "price" A_d as the common good available both periods, with spending on domestic goods in period t of $R_{dt} \equiv A_{dt}\tilde{M}_t$; and the third line follows directly from the definition of A_d in (3.17).

We use this equation to solve for the ratio of real wages, obtaining the result:

Theorem 3.2 (Arkolakis, et al, 2012)

The gains from trade in the Melitz (2003) model are:

$$\frac{w_t / P_t^H}{w_{t-1} / P_{t-1}^H} = \left(\frac{R_{dt}}{w_t L_t}\right)^{\frac{-1}{\omega+1}\left(\frac{\sigma}{\sigma-1}\right)} = \left(\frac{R_{dt}}{w_t L_t}\right)^{-\frac{1}{\theta}}, \tag{3.27}$$

where the final equality is obtained because $\omega \equiv \frac{\theta\sigma}{(\sigma-1)} - 1$, so $\frac{1}{(\omega+1)}\left(\frac{\sigma}{\sigma-1}\right) = \frac{1}{\theta}$.

Note that the ratio of domestic expenditure R_{dt} to total income $w_t L_t$ is equal to one minus the import share, so this formula is identical to the gains from trade in the Krugman (1980) model, except that we replace the exponent $-1/(\sigma-1)$ in that case with $-1/\theta$ in (3.27). This result is precisely the result derived by Arkolakis *et al* (2012), and remarkably, the elasticity of substitution σ does not enter the formula at all (except insofar as it affects the import share). Our derivation gives some intuition as to where this simple formula comes from. Namely, the movement from a corner of the transformation curve A in Figure 3.3, with exports equal to zero, to an interior position like C, gives rise to gains equal to one minus the import (or export) share with the exponent $-1/(\omega+1)$, which is a straightforward application of Theorem 2.2 on the production side of the economy. We might interpret these gains as due to *export* variety. These gains are shown in the second line of (3.26), and reflect the increase in wages due to the productivity improvement as the exporting firms drive out less productive domestic firms. But in addition, this productivity improvement drives down prices, and therefore further increase *real* wages: that is shown as we substitute for the endogenous value of A_d, and thereby solve for real wages in (3.27). Through these two channels, the gains equal one minus the import (or export) share with the exponent $-1/\theta$, which exceeds $-1/(\omega+1) = -(\sigma-1)/\theta\sigma$ in absolute value.

But what about any *further* gain due to *import* variety? Now we must be careful, because the Melitz model leads to the exit of domestic firms and therefore a reduction in domestic varieties, which must be weighted against the increase in import variety. But simply counting the total number of varieties is not the right way to evaluate the welfare gains: instead, we need to take the ratio $(\lambda_t / \lambda_{t-1})^{-1/(\sigma-1)}$ on the consumption side of the economy, as in Theorem 2.2. As we now show, this ratio turns out to be *unity: the gains due to new import varieties are exactly offset by reduced domestic varieties*. Therefore, the production-side gains we have already identified in Theorem 3.2 are all that is available.

To obtain this result, we use the CES price index for the Melitz model:

$$P^H = \left[\int_{\varphi^*}^{\infty} p(\varphi)^{1-\sigma} M\mu(\varphi) d\varphi + \int_{\varphi_x^{F*}}^{\infty} p^F(\varphi)^{1-\sigma} M^F \mu^F(\varphi) d\varphi \right]^{\frac{1}{1-\sigma}}, \tag{3.28}$$

where φ_x^{F*} denotes the zero-profit-cutoff for the foreign exporters, with prices $p^F(\varphi)$. This CES price index is conceptually identical to what we referred to as the unit-expenditure function in (2.2). The average price of domestic goods appearing in (3.28) is:

$$\left[\int_{\varphi^*}^{\infty} p(\varphi)^{1-\sigma} M\mu(\varphi) d\varphi \right]^{\frac{1}{1-\sigma}} = \left(\frac{\sigma}{\sigma-1} \right) \left(\frac{w}{\tilde{\varphi}} \right) M^{\frac{-1}{\sigma-1}}, \tag{3.29}$$

which uses the prices (3.7) together with the definition of average productivity in (3.20).

When comparing autarky (denoted by $t-1$) with free trade (denoted by t), we need to take into account the changing price of domestic goods and their changing variety, as in (3.29), along with the fact the all imported goods are new. Applying Theorem 2.2 gives rise to the following ratio of unit-expenditures:

$$\frac{P_t^H}{P_{t-1}^H} = \left(\frac{w_t / \tilde{\varphi}_t}{w_{t-1} / \tilde{\varphi}_{t-1}}\right)\left(\frac{R_{dt} / w_t L_t}{M_t / M_{t-1}}\right)^{\frac{1}{\sigma-1}}. \tag{3.30}$$

The first term appearing on the right of (3.30) is just the change in the average price of domestic goods, reflecting the change in wages and in average productivity. The aggregate domestic good is available in both periods, so the first term reflects the Sato-Vartia index P_{SV} over the "common" good in Theorem 2.2. The numerator of the second term on the right is the spending on domestic goods relative to total spending in period t; this equals λ_t in Theorem 2.2, or one minus the share of spending on new imported varieties. The denominator of the second term is λ_{t-1} in Theorem 2.2, and reflects the reduction in the number of domestic varieties, $M_t < M_{t-1}$.

We now show that $M_t / M_{t-1} = R_{dt} / w_t L_t$ in (3.30), so the reduction in the number of domestic varieties just cancels with share of spending on new imported varieties, and there are no further consumption gains. This result is obtained from the ZCP condition for domestic firms, in (3.8). The second expression appearing in (3.8) is $q(\varphi^*) = (\sigma - 1) f \varphi^*$. We will combine this with the first expression appearing in (3.8), $r(\varphi^*) / \sigma = wf$, which can be rewritten using the inverse demand curve in (3.17), to obtain:

$$\left[\frac{A_{dt}\, q(\varphi_t^*)^{\frac{\sigma-1}{\sigma}}}{A_{dt-1}\, q(\varphi_{t-1}^*)^{\frac{\sigma-1}{\sigma}}}\right] = \left(\frac{w_t}{w_{t-1}}\right).$$

Using the definition $A_d \equiv P^H (wL / P^H)^{1/\sigma}$, we can readily simplify this expression as:

$$\left[\frac{q(\varphi_t^*)}{q(\varphi_{t-1}^*)}\right] = \left(\frac{w_t / P_t^H}{w_{t-1} / P_{t-1}^H}\right).$$

Now using the ZCP condition that $q(\varphi^*) = (\sigma - 1) f \varphi^*$, we immediately obtain:

$$\left(\frac{\varphi_t^*}{\varphi_{t-1}^*}\right) = \left(\frac{w_t / P_t^H}{w_{t-1} / P_{t-1}^H}\right), \tag{3.31}$$

so that the increase in real wages reflects the increase in the ZCP productivities. From (3.22) the ratio of ZCP productivities equals the ratio of average productivities, $(\tilde{\varphi}_t / \tilde{\varphi}_{t-1})$, then comparing (3.30) with (3.31) we immediately see that $M_t / M_{t-1} = R_{dt} / w_t L_t$, as we intended to show.

The finding that there are no additional consumption gains from variety in the Melitz (2003) model is discussed explicitly by di Giovanni and Levchenko (2010), who argue that if the distribution of firm size follows Zipf's Law then the

extensive margin of imports accounts for a vanishing small portion of the total gains from trade. Their model differs from our discussion above because firms also use differentiated intermediate inputs, but they still assume a Pareto distribution for productivities. This assumption implies that the distribution of firms by size follows a power distribution, which correspond to Zipf's Law as $\theta \rightarrow (\sigma - 1)$. That is the case where they find that the extensive margin of imports has a vanishing contribution to the gains from trade. In comparison, our results above are more general because we show that the extensive margin of imports has a welfare contribution that just cancels with the reduced extensive margin of domestic goods, and this result holds for all values of $\theta > (\sigma - 1)$, and not just the limiting case $\theta \rightarrow (\sigma - 1)$.

4. PRO-COMPETITIVE GAINS FROM TRADE

Let me turn now to the third and final source of gains from trade in a monopolistic competition model, which is the reduced markups of firms when they are faced with import competition. Those gains were emphasized in the first published article on this topic by Krugman (1979), but have been absent in much of the later literature due to its use of CES preferences, in which case markups are themselves fixed. There are now some alternatives to the use of CES preferences, such as the quadratic utility function of Melitz and Ottaviano (2008) which leads to linear demand curves. In that case, markups certainly change as firms are faced with import competition. But the quadratic utility function with linear demand curves also has income elasticities of zero, so it is really most appropriate for partial-equilibrium analysis; and in fact, has been used in the industrial organization literature. Is there some other utility function we could use that is homothetic, or non-homothetic in an interesting way, and still can be used in a monopolistic competition model? It turns out that there is, and is it the favorite functional form of Erwin Diewert's: the translog cost or expenditure function.

Many of you here will be familiar with the translog function from your work on production or consumer demand systems. But it takes some work before we can use these functions in a monopolistic competition model, with new goods coming in and out. In particular, the reservation prices of goods are no longer infinity, but need to be solved for explicitly. Hauseman (1997, 1999)did exactly that in his analysis of honey-nut cheerios and cellular phones, and determined the welfare gains of these new products by solving for their reservation prices before they were available. But that approach become intractable when we have many new sources of supply for differentiated products through international trade. It becomes difficult to keep track of all those reservation prices. fortunately, there is a solution to that problem that makes the translog—or AIDS—function tractable even in a monopolistic competition model. let me conclude my talk by describing this solution and some of its implications.

In a monopolistic competition model we need to be explicit about which goods and available and which are not, so let $\tilde{N}$ denote the maximum number of goods conceivably available, which we treat as *fixed*. The translog unit-expenditure function (Diewert ,1976) is defined as:

$$\ln e = \alpha_0 + \sum_{i=1}^{\tilde{N}} \alpha_i \ln p_i + \tfrac{1}{2} \sum_{i=1}^{\tilde{N}} \sum_{j=1}^{\tilde{N}} \gamma_{ij} \ln p_i \ln p_j \text{, with } \gamma_{ij} = \gamma_{ji} \text{ and } \alpha_i > 0. \tag{4.1}$$

Note that the restriction that $\gamma_{ij}=\gamma_{ji}$ is made without loss of generality. To ensure that the expenditure function is homogenous of degree one, we add the conditions that:

$$\sum_{i=1}^{\tilde{N}} \alpha_i = 1 \text{, and } \sum_{i=1}^{\tilde{N}} \gamma_{ij} = 0 .$$

The share of each good in expenditure is obtained by differentiating (4.1) with respect to $\ln p_i$, obtaining:

$$s_i = \alpha_i + \sum_{j=1}^{\tilde{N}} \gamma_{ij} \ln p_j . \tag{4.2}$$

These shares must be non-negative, of course, but we will allow for a subset of goods to have zero shares because they are not available for purchase. To be precise, suppose that $s_i > 0$ for $i=1,\ldots,N$, while $s_j=0$ for $j=N+1,\ldots,\tilde{N}$. Then for the latter goods, we set $s_j=0$ within the share equations (4.2), and use these $(\tilde{N}-N)$ equations to solve for the reservation prices $\tilde{p}_j, j=N+1,\ldots,\tilde{N}$, in terms of the observed prices p_i, $i=1,\ldots,N$.

Solving for the reservation prices introduces a level of complexity that did not arise in the CES case, where reservation prices are infinite: in the expenditure function (2.2), an infinite reservation price raised to the negative power $(1-\sigma)$ simply vanishes. To solve for finite reservation prices in the translog case, it is essential to simplify the translog by imposing the additional "symmetry" requirements:

$$\gamma_{ii} = -\gamma\left(\frac{\tilde{N}-1}{\tilde{N}}\right) < 0, \; and \; \gamma_{ij} = \frac{\gamma}{\tilde{N}} > 0 \quad for \; i \neq j, \text{ with } i,j=1,\ldots,\tilde{N}. \tag{4.3}$$

It is readily confirmed that the restrictions in (4.3) satisfy the above homogeneity conditions, and also guarantee that the reservation prices are finite. Because $\tilde{N}$ is a fixed number, (4.3) simply says that the Γ matrix has a negative constant on the diagonal, and a positive constant on the off-diagonal, chosen so that the rows and columns sum to zero.

The restrictions in (4.3) are not familiar from the translog literature, but are essential to solve for reservation prices for goods not available. Note that we have *not* restricted the $\alpha_i > 0$ parameters, though they must sum to unity, so there are $\tilde{N}-1$ free α_i parameters. In addition, we have the free parameter α_0 in (4.1) as well as $\gamma > 0$ in (4.3), so there are a total of $\tilde{N}+1$ free parameters in this "symmetric" translog function. That is the same number of free parameters in our "non-symmetric" CES function (2.1), where we allowed for $\tilde{N}$ parameters $a_i > 0$ (possibly changing over time) along with the elasticity $\sigma > 1$. So in describing the translog case as "symmetric" we are comparing it to the empirical version that does not use (4.3); while in describing the CES function as "non-symmetric" we are comparing it to the theoretical version in monopolistic competition models that assumes $a_i \equiv 1$, $i=1\ldots,\tilde{N}$. In fact, both the CES function in (2.1) and the translog in (4.1) have the same number of free parameters, or degree of symmetry, which we have chosen to be tractable in a monopolistic competition framework.

The usefulness of the symmetric restrictions in (4.3) is shown by the following result:

Theorem 4.1 (Feenstra, 2003; Bergin and Feenstra, 2009)

Using the symmetry restrictions (43), suppose that only the goods $i = 1,\ldots,N$ are available, so the reservation prices $\tilde{p}_j$ for $j = N+1,\ldots,\tilde{N}$ are used. Then the unit-expenditure function equals:

$$\ln e = a_0 + \sum_{i=1}^{N} a_i \ln p_i + \tfrac{1}{2} \sum_{i=1}^{N}\sum_{j=1}^{N} b_{ij} \ln p_i \ln p_j, \tag{4.4}$$

$$\text{where: } b_{ii} = -\gamma \frac{(N-1)}{N} < 0, \;\; and \;\; b_{ij} = \frac{\gamma}{N} > 0 \;\; for \; i \neq j \;\; \text{with } i,j = 1,\ldots,N, \tag{4.5}$$

$$a_i = \alpha_i + \tfrac{1}{N}\left(1 - \sum_{i=1}^{N} \alpha_i\right), \text{ for } i = 1,\ldots,N, \tag{4.6}$$

$$a_0 = \alpha_0 + \left(\frac{1}{2\gamma}\right)\left\{\sum_{i=N+1}^{\tilde{N}} \alpha_i^2 + \left(\frac{1}{N}\right)\left(\sum_{i=N+1}^{\tilde{N}} \alpha_i\right)^2\right\}. \tag{4.7}$$

Notice that the expenditure function in (4.4) looks like a conventional translog function defined over the goods $i = 1,\ldots,N$, while the symmetry restrictions continue to hold in (4.5), but are defined now using the number of available goods N, which can change over time. As N grows, for example, we find that the price elasticity of demand also grows because goods are closer substitutes. To interpret (4.6), it implies that each of the coefficient α_i is increased by the same amount to ensure that the coefficients a_i sum to unity over $i = 1,\ldots,N$. The final term a_0, appearing in (4.7), incorporates the coefficients α_i of the unavailable products. If the number of available products N rise, then a_0 falls, indicating a welfare gain from increasing the number of available products.

Theorem 4.1 is a promising start towards using the translog function in monopolistic competition models. For theoretical work, this result is all that is needed and it shows that the translog system can join the quadratic preferences used by Melitz and Ottaviano (2008) as being tractable alternatives to the CES case. Furthermore, both the translog and quadratic preferences allow for endogenous markups. The real advantage of the translog unit-expenditure function is on empirical grounds. As argued by Diewert (1976), it has a number of convenient properties: it is obtained from homothetic preferences, provides a second-order approximation to an arbitrary expenditure function, and corresponds to the Törnqvist price index, which is very close to price index formulas that are used in practice.

Feenstra and Weinstein (2009) develop an alternative formula for the welfare gain from new products, beyond Theorem 4.1, that depends on the observable expenditure shares on goods and can therefore be implemented. The terms appearing in the formula for the welfare gain are analogous to those appearing in (4.7), but using observable expenditure shares in place of α_i: the welfare gain from new products depends on the sum of squared shares, and on the square of the sum of shares, of new products. The sum of squared product shares—or Herfindahl indexes—also determine the average markups charged by firms in each market. Increased shares of imports and reduced U.S. shares can lead to reduced U.S. markups, and also contribute to variety gains. For these reasons, the translog case offers a promising theoretical and empirical framework to assess the gains from import variety and the effect of imports on reducing markups.

5. CONCLUSIONS

This paper is about measurement: how to measure the gains from trade that arise in the monopolistic competition model. The CES functional form, introduced into the monopolistic competition model by Dixit and Stiglitz (1977) and adopted by Krugman (1980, 1981) and later literature, is just as convenient in empirical work as it is in theory. Using this functional form, Feenstra (1994) showed how the gains from new product varieties depend on their expenditure share as well as on the elasticity of substitution. The expenditure on new imported products, or more precisely, on new *source countries* for imports, are available from highly disaggregate trade statistics. In addition, estimates of the elasticity of substitution between source countries for imports can be obtained using the same disaggregate trade statistics over time, as described in Feenstra (1994). Broda and Weinstein (2006) applied these methods to import data for the United States, and find that the gains from new source countries for imports can be substantial: by 2001, these gains amount to 2.6% of U.S. GDP.

Recently, attention has shifted in the monopolistic competition literature to the production side of the economy. Whereas Dixit and Stiglitz (1977) and Krugman (1980, 1981) relied on the "symmetry" assumption that all firms are identical, Melitz (2003) was able to introduce heterogeneity in the productivity of firms. This framework allows firms to have stochastic draws of productivity, but still imposes that firm profits are zero *ex ante*, as required by free entry into the industry. This extension to the monopolistic competition model is well-grounded in empirical observations: it allows for only a subset of firms in the industry—the more efficient firms—to be exporters. In the Canadian context, Trefler (2004) showed that the exit of less-efficient firms led to a substantial increase in average industry productivity following the Canada-U.S. free trade agreement, which supports the Melitz model.

We have explored the industry-level implications of the Melitz model, and found that it leads to a concave, constant-elasticity transformation curve between domestic and export varieties, adjusting for the appropriate quantity of each. Analogous to the CES results on the consumer side, the gains from trade depend on the share of revenue devoted to exports and on the elasticity of transformation, which itself depends on the elasticity of substitution and on the Pareto parameter for productivity draws. Remarkably, once we take into account the general equilibrium increase in spending following trade liberalization (i.e. endogeneity of the shift parameters A_{dt}), then the gains from trade simplify so that they depend on the share of revenue devoted to exports (or equivalently, imports), and on the Pareto parameter. This confirms the very simple formula for the gains from trade found by Arkolakis *et al* (2012). All these gains come from the production side of the economy, and there are *no further* gains from product variety on the consumption side: the gains from import varieties just cancel with the losses from reduced domestic varieties. That results follows from having an "interior solution" where only a fraction of the domestic firms are exporters: if all firms exported or no firms exported in some industries, then we would expect to again see consumption gains from variety as in Krugman (1980, 1981).

The final topic we have discussed is the gains from trade due to reduced markups charged by firms, as in Krugman (1979). It is worth emphasizing that these are social gains and not just a transfer from firms to consumers. In Krugman (1979), reduced markups combined with zero profits in equilibrium imply that firms are moving down their average cost curves, taking greater advantage of economies of scale. So the reduction in consumer prices due to reduced markups do not come at the expense of firms profits. In order to measure these gains we must move beyond the CES case, however, where markups are constant. In theory, the quadratic utility function used by Melitz and Ottaviano (2008) offers a very useful form of endogenous markets. Because this utility function uses an additively separable numeraire good, all other products all have income elasticities of zero. On empirical grounds, we recommend instead

the translog unit-expenditure function, which corresponds to homothetic preferences (income elasticities of unity). Unlike the CES case, goods then have finite reservation prices that must be solved for. Feenstra (2003) and Bergin and Feenstra (2009) show how this expenditure function, when simplified to allow for some "symmetry" across goods, has a convenient solution for the reservation prices that can be substituted back into the expenditure function, obtaining a tractable form even as the number of goods varies. Feenstra and Weinstein (2009) are making use of this functional form to estimate the impact of globalization on markets and product variety in the U.S. market. It can be expected that applications to many other countries will follow, thereby allowing us measure this third source of gains from trade due to monopolistic competition.

FIGURE 2.1. CES DEMAND

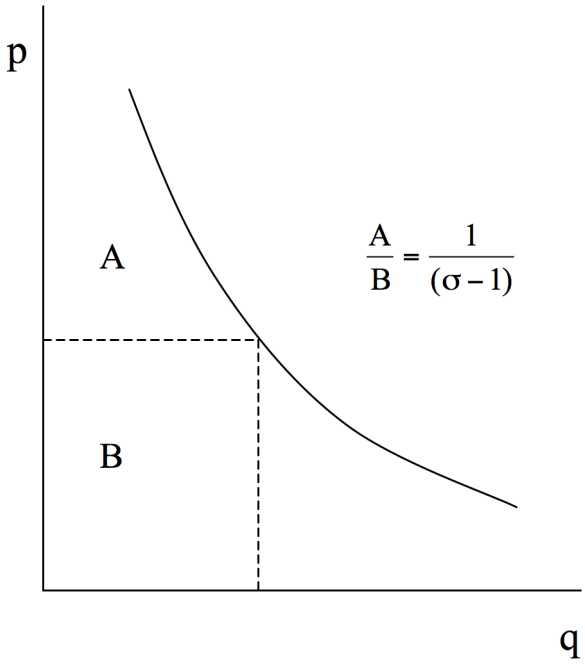


FIGURE 2.2. CES INDIFFERENCE CURVE

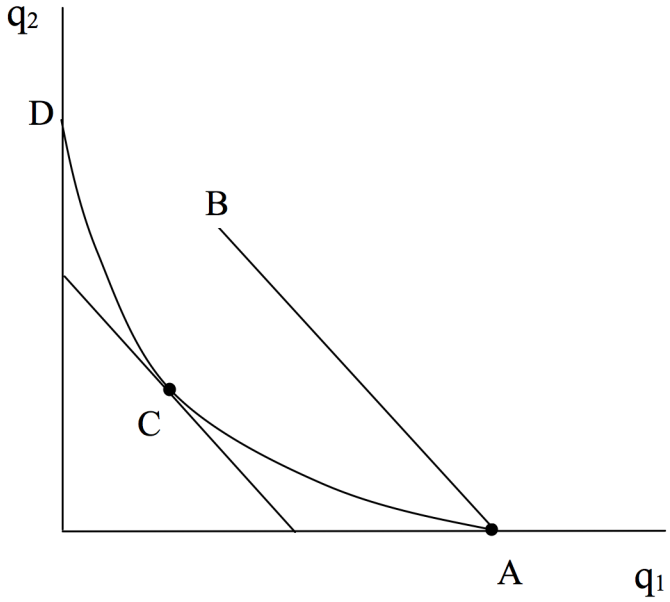


FIGURE 3.1. CONSTANT-ELASTICITY TRANSFORMATION CURVE

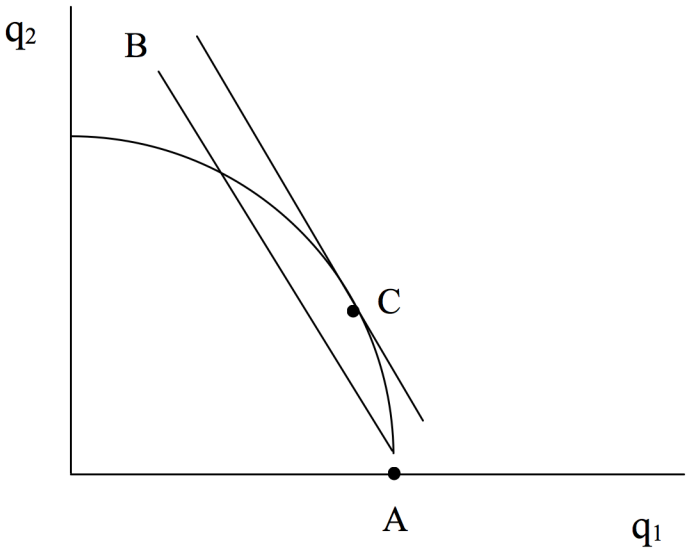
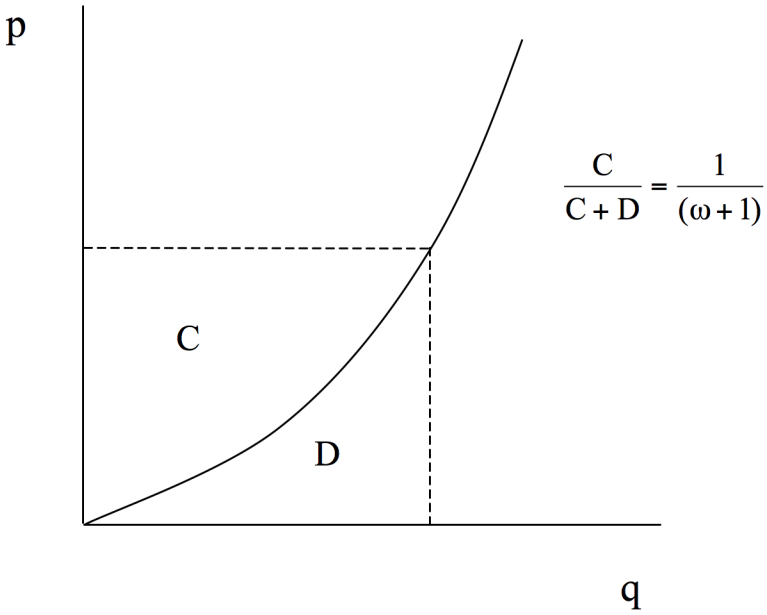


FIGURE 3.2. CONSTANT-ELASTICITY SUPPLY CURVE



$$\frac{C}{C+D} = \frac{1}{(\omega+1)}$$

References

Arkolakis, Costas, Arnaud Costinot and Andrés Rodríguez-Clare (2012), "New Trade Models, Same Old Gains?," American Economic Review, 102(1), February, 94-130.

Armington, Paul S. (1969) "A Theory of Demand for Products Distinguished by Place and Production," IMF Staff Papers, 16, 159-178.

Bergin, Paul R. and Robert C. Feenstra (2009) "Pass-through of Exchange Rates and Competition Between Fixers and Floaters," Journal of Money, Credit and Banking, February, 41(s1), 35-70.

Chaney, Thomas (2008) "Distorted Gravity: The Intensive and Extensive Margins of International Trade," American Economic Review, 98(4): 1707–21.

Diewert, W. Erwin (1976) "Exact and Superlative Index Numbers," Journal of Econometrics, 4, 115-145.

Eastman, Harry C. and Stefan Stykolt (1967) The Tariff and Competition in Canada. Toronto: Macmillan.

Eaton, Jonathan and Samuel Kortum (2002) "Technology, Geography and Trade," Econometrica, 70(5), September, 1741-1780.

Feenstra, Robert C. (1994) "New Product Varieties and the Measurement of International Prices," American Economic Review, 84(1), March, 157-177.

Feenstra, Robert C. (2003) "A Homothetic Utility Function for Monopolistic Competition Models, Without Constant Price Elasticity," Economic Letters, 78, 79-86

Feenstra, Robert C. (2009) "Paul Krugman, Recipient of the 2008 Nobel Prize in Economics: An Appreciation," Challenge: The Magazine of Economic Affairs, January-February, 97-107.

Feenstra, Robert C. (2010) Product Variety and the Gains from International Trade. MIT Press.

Feenstra, Robert C. and Hiau Looi Kee (2008) "Export Variety and Country Productivity: Estimating the Monopolistic Competition Model with Endogenous Productivity," Journal of International Economics, 74(2), 500-518.

Feenstra, Robert C. and Andrew K. Rose (1993) "Trade With Mexico and Water Use in California Agriculture," in Peter Garber, ed., The Mexico-U.S. Free Trade Agreement, MIT Press, 1993, 189-218.

Feenstra, Robert C. and David Weinstein (2010) "Globalization, Competition, and the U.S. Price Level," NBER Working Paper no. 15749.

Giovanni, Julian di and Andrei A. Levchenko (2010) "Firm Entry, Trade and Welfare in Zipf's world," NBER Working Paper No. 16313.

Hausman, Jerry A. (1997) "Valuation of New Goods under Perfect and Imperfect Competition," in Timothy F. Bresnahan and Robert J. Gordon, eds., The Economics of New Goods, NBER Studies in Income and Wealth, no. 58, Chicago: University of Chicago, 209-237.

Hausman, Jerry A. (1999) "Cellular Telephone, New Products and the CPI," Journal of Business and Economics Statistics, 17(2), 188-194.

Head, Keith C. and John Ries (1999) "Rationalization effects of tariff reductions," Journal of International Economics, 47(2), April, 295-320.

Hicks, John R. (1940) "The Valuation of Social Income," Economica, 7, 105-24.

Krugman, Paul R. (1979) "Increasing Returns, Monopolistic Competition and International Trade," Journal of International Economics, 9, 469-479.

Krugman, Paul R. (1980) "Scale Economies, Product Differentiation, and the Pattern of Trade," American Economic Review, 70, December, 950-959.

Krugman, Paul R. (1981) "Intra-Industry Specialization and the Gains from Trade," Journal of Political Economy, 89, 959-973.

Leamer, Edward E. (1981) "Is it a Demand Curve, or is it a Supply Curve? Partial Identification 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 Giancarlo I. P. Ottaviano (2008) "Market Size, Trade, and Productivity," Review of Economic Studies, 75(1), 295-316.

Rigobon, Roberto (2003) "Identification Through Heteroskedasticity," Review of Economics and Statistics, 85(4), July, 777-792.

Sato, Kazuo (1976) "The Ideal Log-Change Index Number," Review of Economics and Statistics 58, May, 223-228.

Trefler, Daniel (2004) "The Long and Short of the Canada-U.S. Free Trade Agreement," American Economic Review, 94(4), September, 870-895.

Vartia, Y.O. (1976) "Ideal Log-Change Index Numbers", Scandinavian Journal of Statistics 3, 121-126.

DISCUSSION

MUNISAMY GOPINATH
OREGON STATE UNIVERSITY

Professor Feenstra is a pioneer in assessing welfare gains from trade, especially in the context of new trade theory and the emerging heterogeneous-firms models of trade. Three sources of gains from trade are commonly found in the literature: consumer gains from the increase in imported varieties; producer gains from self-selection and the ensuing changes in extensive and intensive margins; and disciplining of mark-ups from import competition. In this paper, Professor Feenstra elegantly measures the consumer gains from the increase in the number of imported varieties by adapting the Sato-Vartia index to the Krugman model. The added advantage of this index is that it yields similar results under alternative weighting schemes. However, Professor Feenstra notes that previous attempts to measure consumer gains did not account for self-selection of firms in domestic and export markets and the accompanying loss in domestic varieties. He then shows that gains from trade accrue on the producer side, despite loss of domestic varieties, with an exposition of Melitz's (2003) trade model with heterogeneous firms. With additional algebra, the gains from trade in Melitz's model are captured in the Feenstra framework for a combined assessment of consumer and producer gains. Then, Professor Feenstra shows how a translog expenditure function can be employed to assess gains from import varieties and falling mark-ups. Finally, empirical results from Feenstra and Weinstein (2009), based on the translog method, yields gains from trade in the form of declining merchandise prices by 5.4 percent and consumer prices by 1.0 percent during 1992-2005.

The gains from trade demonstrated by Professor Feenstra are critical as academics and free-traders continue to make the case for increased integration of global markets. Nevertheless, the results should be viewed in a broader context. For instance, the definition of a variety can be tricky especially at the merchandise and consumer levels. Importing a shirt collar does not directly add to shirt varieties in a country. Moreover, the presence of multinational firms and the production-sharing/outsourcing may overestimate the number of varieties arriving at the U.S. border. Estimates of gains at the national level are useful, but the political economy of U.S. trade policy requires some more attention to the distribution of gains within this large country. Are there regions where self-selection leads to loss of jobs and net welfare, despite consumer gains through price declines? Does the ensuing spatial adjustment of firms and workers impose additional unaccounted costs in the assessment of welfare gains from trade? The emerging new economic geography literature addresses some of these questions, but micro data availability and access (consistent firm- or plant-level data across and within countries) remain major obstacles to these efforts.