Measuring the Gains from Trade under Monopolistic Competition

by

Robert C. Feenstra
Department of Economics,
University of California, Davis

Revised, September 2009

Abstract

Three sources of gains from trade under monopolistic competition are: (i) new import varieties available to consumers; (ii) enhanced efficiency as more productive firms begin exporting and less productive firms exit; (iii) reduced markups charged by firms due to import competition. The first source of gains can be measured as new goods in a CES utility function for consumers. We argue that the second source is formally analogous to the producer gain from new goods, with a constant-elasticity transformation curve for the economy. We suggest that the third source of gain can be measured using a translog expenditure function for consumers, which in contrast to the CES case, allows for finite reservation prices for new goods and endogenous markups.

For correspondence: rcfeenstra@ucdavis.edu, (530)752-7022

1. Introduction

One of the great achievements of international trade theory in the last three decades is the incorporation of the monopolistic competition model. The need to include increasing returns to scale in trade theory was recognized as early as Graham (1923; see also Ethier 1982), and in the Canadian context, by Eastman and Stykolt (1967) and Melvin (1969). Still, it was not until the formalization of the monopolistic competition model by Dixit and Stiglitz (1977), in parallel with Spence (1976) and Lancaster (1979), that a set of global equilibrium conditions that avoided the problems of large firms and multiple equilibria could be developed. That set of equilibrium conditions was first written down by Krugman (1979, 1980, 1981).[1]

There is no doubt that these developments have had important policy implications. For example, the simulation results of Harris (1984a,b) demonstrated large gains to Canada from free trade with the U.S., and were very influential in convincing policy makers to proceed with the Canada-U.S. free trade agreement in 1989; that agreement in turn paved the way for the North American free trade agreement in 1994. Subsequent empirical work for Canada by Trefler (2004), as well as Head and Ries (1999, 2001), confirmed the efficiency gains for Canada due to opening trade, though not in the manner predicted by Krugman's work. But a comprehensive empirical assessment of the gains from trade under monopolistic competition has not yet been made. The goal of this paper is to describe how these gains can be measured, using methods that draw heavily on duality theory from Diewert (1974, 1976).

The monopolistic competition model predicts three sources of gains from trade that are not present in traditional models. First, there are the consumer gains from having access to new import varieties of differentiated products. Those gains have recently been measured for the United States by Broda and Weinstein (2006), using the methods from Feenstra (1994), as described in section 2. Their approach assumes a constant elasticity of substitution (CES) utility function for consumers, in which case the import varieties are analogous to "new goods" in the utility function. We show that the gains from trade depend on the import share and the elasticity of substitution, as also found by Arkolakis *et al* (2008a).

The extension of the monopolistic competition model to allow for heterogeneous firms, due to Melitz (2003), leads to a second source of gains from the self-selection of more efficient firms into export markets. This activity drives out less efficient firms and therefore raises overall productivity. This self-selection of firms was demonstrated for Canada by Trefler (2004) following the free trade agreement with the U.S. We argue that this self-selection can still be interpreted as a gain from product variety, but now on the *export* side of the economy rather than for imports. Surprisingly, the consumer gains from new import varieties do not appear in this case, because they cancel out with disappearing domestic varieties. This finding, demonstrated in section 3, helps to explain the theoretical results of Arkolakis *et al* (2008b), where the gains from trade depend on the import share but are otherwise *independent* of the elasticity of substitution in consumption. Rather, the gains come from the production side of the economy, where the self-selection of firms leads to a constant-elasticity transformation curve between domestic and export varieties, with an elasticity depending on the Pareto parameter of productivity draws.

Third, the monopolistic competition model also allows for gains from a reduction in firm markups due to import competition. This third source of gains was stressed in Krugman (1979), but has been absent from much of the later literature due to the assumption of CES preferences, leading to constant markups. In section 4, I introduce the translog expenditure function into the monopolistic competition model, and summarize current research in Bergin and Feenstra (2009) and Feenstra and Weinstein (2009). I argue that the translog unit-expenditure function is tractable even as the number of product varieties is changing, as with monopolistic competition. It can be expected to lead to estimable formulas for the gains from product variety and the effect of imports on reducing markups. Conclusions are given in section 5.

2. Consumer Benefits from Import Variety

We start with the consumer gains from import variety. From a technical point of view, measuring the benefits of new import varieties is equivalent to the so-called "new goods" problem in index number theory. That has always been a favorite problem of Erwin Diewert's (see Diewert, 1987, p. 779), and arises because the price for a product before it is available is not observed, so we don't know what price to enter in an index number formula. The answer given many years ago by Hicks (1940) was that the relevant price of a product before it is available is the "reservation price" for consumers, namely, a price so high that demand is zero. Once the product appears on the market then it has a lower price, determined by supply and demand. The fall in the price from its "reservation" level to the actual price can be used in an index number formula to obtain the consumer gains from the appearance of that new good.

For the constant elasticity of substitution (CES) utility function, we immediately run into a problem with implementing this suggestion because the reservation price for any good is infinite: the demand curve approaches the vertical axis as the price approaches infinity.[2] But provided that the elasticity of substitution is greater than unity, then the area under the demand curve is bounded above, as shown in Figure 1, where the ratio of areas A/B = $1/(\sigma - 1)$ is easily calculated for a demand curve with elasticity σ. Thus, even with an infinite reservation price, there is a well-defined area of consumer surplus from having the new good available, depending on the elasticity of substitution (we will see the term $1/(\sigma-1)$ again in Theorem 2 below).

The second problem we run into is how to express these consumer gains when there is not just one but many new goods available. To address that case, we do not rely on consumer surplus to measure the welfare gain, as in Figure 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).

2.1 CES Utility Function

To address this problem, we will work with the non-symmetric CES function,

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

where $a_{it} > 0$ are tastes parameters that can change over time, and I_t denotes the set of goods available in period t at the prices p_{it}. The minimum expenditure to obtain one unit of utility is,

$$e(p_t, I_t) = \left[\sum_{i \in I_t} b_{it} p_{it}^{1-\sigma}\right]^{1/(1-\sigma)}, \ \sigma > 1, \ b_{it} \equiv a_{it}^{\sigma}. \tag{2}$$

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-expenditures:

<u>Theorem 1</u> (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:

$$\frac{e(p_t, I)}{e(p_{t-1}, I)} = P_{SV}(p_{t-1}, p_t, q_{t-1}, q_t, I) \equiv \prod_{i \in I} \left(\frac{p_{it}}{p_{it-1}}\right)^{w_i(I)}, \tag{3}$$

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,

$$w_i(I) \equiv \left(\frac{s_{it}(I) - s_{it-1}(I)}{\ln s_{it}(I) - \ln s_{it-1}(I)}\right) \Bigg/ \sum_{i \in I} \left(\frac{s_{it}(I) - s_{it-1}(I)}{\ln s_{it}(I) - \ln s_{it-1}(I)}\right). \tag{4}$$

The numerator in (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 (4) is needed to precisely measure the ratio of unit-expenditures in (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. In both cases, the geometric mean formula in (3) applies. The important point from Theorem 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:

<u>Theorem 2</u> (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 $\sigma > 1$:

$$\frac{e(p_t, I_t)}{e(p_{t-1}, I_{t-1})} = P_{SV}(p_{t-1}, p_t, q_{t-1}, q_t, I)\left(\frac{\lambda_t(I)}{\lambda_{t-1}(I)}\right)^{1/(\sigma-1)}, \tag{5}$$

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 (4), and the values $\lambda_t(I)$ and $\lambda_{t-1}(I)$ are constructed as:

$$\lambda_\tau(I) = \left(\frac{\sum_{i\in I} p_{i\tau}q_{i\tau}}{\sum_{i\in I_\tau} p_{i\tau}q_{i\tau}}\right) = 1 - \left(\frac{\sum_{i\in I_\tau, i\notin I} p_{i\tau}q_{i\tau}}{\sum_{i\in I_\tau} p_{i\tau}q_{i\tau}}\right), \quad \tau = t\text{-}1, t. \tag{6}$$

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 is a greater number of 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 (5), by an amount that depends on the elasticity of substitution.

The importance of the elasticity of substitution can be seen from Figure 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 consumption is at 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, which depends on the convexity of the indifference curve, or the elasticity of substitution. As the elasticity approaches infinity, so the varieties are perfect substitutes, then there is no welfare gain from having a second variety available (at the same price) and the bias in the conventional price index vanishes, since the final term in (5) goes to unity.

2.2 Krugman (1980) Model

Turning to the international trade application, we will suppose that the utility function in (1) applies to the purchases of a good from various source countries $i \in I_t$. That is, the elasticity of substitution we are interested in is the Armington (1969) elasticity between the source countries for imports. We refer to the source countries as providing *varieties* of the differentiated good, so the gains being measured in (5) are the gains from import variety. In this case, we can compare the formula in (5) with the gain from trade obtained in the model of Krugman (1980), as analyzed by Arkolakis *et al* (2008a).

In particular, suppose there are any number of countries, where the representative consumer in each has a CES utility function with elasticity $\sigma > 1$. Labor is the only factor of production and there is a single monopolistically competitive sector, with no other goods.[3] Firms face a fixed cost of f to manufacture any good, and an iceberg transport cost to sell it abroad, but no other fixed cost for exports. Then it is well known that with profit-maximization and zero profits through free entry, the output of each firm is fixed at the amount:[4]

$$q = (\sigma - 1)f\varphi, \tag{7}$$

where φ is the productivity of the firm, i.e. the number of units of output per unit of labor. With the population of L, the full-employment condition is then:

$$L = N[(q / \varphi) + f] = N\sigma f, \tag{8}$$

which determines the number of product varieties produced in equilibrium as $N = L / \sigma f$. This condition holds under autarky or trade, so opening a country to trade has *no impact* on the number of varieties produced within a country.

The gains from opening trade can be measured by the ratio of real wages under free trade and autarky. With labor as the only factor of production we can normalize wages at unity, so the gains from trade are simply measured by the drop in the cost of living, which is the inverse of (5). The "common" set of goods are those domestic varieties that are available both in autarky and under trade. Then the Sato-Vartia index

P_{SV} is just the change in the price of the domestic varieties, and with constant markups that is the change in home wages, which we have normalized to unity. So the gains from trade are measured by $(\lambda_t / \lambda_{t-1})^{-1/(\sigma-1)}$ in (5). The denominator of that ratio reflects the disappearance of domestic varieties, i.e. those varieties available in period t-1 but not in period t. As we have shown above, there are *no disappearing domestic varieties* in this model, so $\lambda_{t-1} = 1$. The numerator λ_t measures *the expenditure on the domestic varieties relative to total expenditure with trade, or one minus the import share.* The gains from trade are therefore $\lambda_t^{-1/(\sigma-1)}$, which is precisely the formula obtained by Arkolakis *et al* (2008a). While this formula is not too surprising, it will take on greater significance when we compare it to the results from the Melitz (2003) model, in the next section.

Broda and Weinstein (2006) measure these gains from trade for the U.S. They define a *good* as a 10-digit Harmonized System (HS) category, or before 1989, as a 7-digit Tariff Schedule of the United States (TSUSA) category. 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, using the GMM method from Feenstra (1994), which exploits heteroskedasticity across countries to identify this elasticity. Putting these together, they measure $(\lambda_t / \lambda_{t-1})^{-1/(\sigma-1)}$ for 30,000 goods available in the HS and TSUSA data. For the TSUSA data they used 1972 as the base year and measured the gains from new supply 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.[5] Aggregating over goods, they obtain an estimate of the gains from trade for the US 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 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. That is the correct way to measure the gains from new import varieties.[6]

Second, Broda and Weinstein (2006) did not incorporate any changes in the number of U.S. varieties into their estimation, nor 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 that is 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 not addressed by Broda and Weinstein (2006), and we shall begin to address it in the remainder of the paper.

3. Producer Benefits from Output Variety

The consumer benefits from product variety discussed in the previous section also apply to firms if they rely on differentiated intermediate inputs, and in particular, imported inputs. That is the case assumed in much of the endogenous growth literature, where the increased range of intermediate inputs fuels growth. The benefits to producers are conceptually the same the benefits that consumers enjoy from increased product variety, with the indifference curve shown in Figure 2 replaced by an isoquant.[7]

But there is another type of potential producer benefit, which arises from *output* variety. That benefit can be seen by expanding the range of elasticities σ that we considered in the utility and expenditure functions (1) and (2). While we have so far restricted out attention to $\sigma > 1$, if we consider $\sigma < 0$ then instead of obtaining convex indifference curves from (1) for a fixed level of U_t, we obtain a concave transformation curve as shown in Figure 3.[8] 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) equals $-\sigma$. This reinterpretation of (1) comes from Diewert (1976), who uses the general term "aggregator function" to refer to utility functions, production function, or transformation functions for an economy.

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 (1) using labor resources L_t to replace utility U_t, obtaining:

$$L_t = \left(\sum_{i \in I_t} a_{it} q_{it}^{(\omega+1)/\omega} \right)^{\omega/(\omega+1)}, \quad a_{it} > 0, \ \omega > 0. \tag{9}$$

The maximum revenue obtained using one unit of labor resources, dual to (9), is then:

$$e(p_t, I_t) = \left[\sum_{i \in I_t} b_{it} p_{it}^{\omega+1} \right]^{1/(\omega+1)}, \qquad b_{it} \equiv a_{it}^{-\omega}, \ \ \omega > 0. \tag{10}$$

With this reinterpretation, Theorem 2 continues to hold as:

$$\frac{e(p_t, I_t)}{e(p_{t-1}, I_{t-1})} = P_{SV}(p_{t-1}, p_t, q_{t-1}, q_t, I) \left(\frac{\lambda_t(I)}{\lambda_{t-1}(I)} \right)^{-1/(\omega+1)}, \tag{11}$$

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. 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 4 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.[9] 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 (8) in Krugman's (1980) model. In that case, the gains from output variety vanish. So the question arises as to whether the strictly concave case we illustrate in Figure 3 has any practical application?

We will now argue that the case of a strictly concave transformation curve is indeed relevant, and in fact, arises in the generalization of the monopolistic competition model due to Melitz (2003). Melitz assumes that labor is the only factor of production, but he allows firms to differ in 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 $\varphi_x^* > \varphi^*$ 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.[10]

3.1 Melitz (2003) model

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:

$$[1-G(\varphi^*)]M_e = \delta M \quad . \tag{12}$$

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

$$\mu(\varphi) = \begin{cases} \dfrac{g(\varphi)}{[1-G(\varphi^*)]} & \text{if} \quad \varphi \geq \varphi^*, \\ \\ 0 & \text{otherwise,} \end{cases} \tag{13}$$

where $g(\varphi) = \partial G(\varphi)/\partial\varphi$ 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:

$$r(\varphi) = p(\varphi)q(\varphi) = \left[\frac{p(\varphi)}{P^H}\right]^{1-\sigma} wL, \qquad \sigma > 1, \tag{14}$$

where $q(\varphi)$ is the quantity sold and P^H is the home CES price index.[11] The profit-maximizing price from selling in the domestic market is the usual constant markup over marginal costs:

$$p(\varphi) = \left(\frac{\sigma}{\sigma-1}\right)\frac{w}{\varphi}. \tag{15}$$

Using this, we can 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:

$$r(\varphi^*)/\sigma = wf \quad \Rightarrow \quad q(\varphi^*) = (\sigma-1)f\varphi^*, \tag{16}$$

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, in (7), for all firms.

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:

$$r_x(\varphi) = p_x(\varphi)q_x(\varphi) = \left[\frac{p_x(\varphi)\tau}{P^F}\right]^{1-\sigma} w^*L^*, \tag{17}$$

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

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

$$p_x(\varphi) = \left(\frac{\sigma}{\sigma-1}\right)\frac{w}{\varphi}. \tag{18}$$

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:

$$r_x(\varphi_x^*)/\sigma = wf_x \quad \Rightarrow \quad q_x(\varphi_x^*) = (\sigma-1)f_x\varphi_x^*, \tag{19}$$

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:

$$M_x \equiv \int_{\varphi_x^*}^{\infty} M\mu(\varphi)d\varphi < M. \tag{20}$$

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:

$$L = M_e f_e + M \int_{\varphi^*}^{\infty} [q(\varphi)/\varphi + f]\mu(\varphi)d\varphi + M_x \int_{\varphi_x^*}^{\infty} [q_x(\varphi)/\varphi + f_x]\mu_x(\varphi)d\varphi , \qquad (21)$$

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 (21) by multiplying by w, and using the fact that $(w/\varphi)q(\varphi) = r(\varphi)(\sigma - 1)/\sigma$, and likewise for exporters, to obtain:

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

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:

$$L = \sigma\left(M_e f_e + Mf + M_x f_x\right). \qquad (22)$$

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

$$G(\varphi) = 1 - \varphi^{-\theta}, \text{ with } \theta > \sigma - 1 > 0 . \qquad (23)$$

In that case, it can be shown (see the Appendix) 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:

$$L = \frac{\sigma\theta}{(\theta - \sigma + 1)}\left(Mf + M_x f_x\right). \tag{24}$$

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, in (7). 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, as in (6), so the transformation is *also* linear in the quantity produced by any groups of firms. 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 (14) we obtain:

$$r(\varphi) = A_d\, q(\varphi)^{\frac{\sigma-1}{\sigma}}, \text{ where } A_d \equiv P^H\left(\frac{wL}{P^H}\right)^{\frac{1}{\sigma}}. \tag{25}$$

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:

$$r_x(\varphi) = A_x q_x(\varphi)^{\frac{\sigma-1}{\sigma}}\text{, where } A_x \equiv \left(\frac{P^F}{\tau}\right)\left(\frac{\tau_i w^* L^*}{P^F}\right)^{\frac{1}{\sigma}}. \tag{26}$$

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

$$wL = A_d M \int_{\varphi^*}^{\infty} q(\varphi)^{\frac{\sigma-1}{\sigma}} \mu(\varphi)d\varphi + A_x M_x \int_{\varphi_x^*}^{\infty} q_x(\varphi)^{\frac{\sigma-1}{\sigma}} \mu_x(\varphi)d\varphi\,. \tag{27}$$

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 and A_x as given, are precisely the monopolistic competition equilibrium conditions. So the quantities appearing in this expression 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 (15), 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:

$$\tilde{\varphi} \equiv \left[\int_{\varphi^*}^{\infty} \varphi^{(\sigma-1)} \mu(\varphi)d\varphi\right]^{1/(\sigma-1)}, \tag{28}$$

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:

$$\tilde{M} \equiv M q(\tilde{\varphi})^{(\sigma-1)/\sigma} \quad \text{and} \quad \tilde{M}_x \equiv M_x q_x(\tilde{\varphi}_x)^{(\sigma-1)/\sigma}. \tag{29}$$

To simplify this expression for GDP further, we note that a property of the Pareto distribution is that an integral like (28) is always a constant multiple of the lower bound of integration. That is:

$$\tilde{\varphi} = \left[\frac{\theta}{(\theta-\sigma+1)}\right]^{1/(\sigma-1)} \varphi^*, \tag{30}$$

as obtained by evaluating the integral in (28), 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:

$$(\varphi^*)^{-\theta} = \frac{\delta\sigma\theta f_e}{L(\sigma-1)} M. \tag{31}$$

Gathering together these results, we can use $q(\tilde{\varphi}) = (\tilde{\varphi}/\varphi^*)^{\sigma} q(\varphi^*)$ to 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 second equality uses (30) and the ZCP condition $q(\varphi^*) = (\sigma-1)f\varphi^*$, and the third follows from (31), where $k_1 > 0$ 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 (24), we obtain a concave transformation curve between $\tilde{M}$ and $\tilde{M}_x$, with elasticity $\omega \equiv \frac{\theta\sigma}{(\sigma-1)} - 1 > 0$:

$$L = k_2 f_e^{1/(\omega+1)} \left(\tilde{M}^{\frac{\omega+1}{\omega}} f^{1+\frac{(\omega+1)(\sigma-1)}{\omega\sigma}} + \tilde{M}_x^{\frac{\omega+1}{\omega}} f_x^{1+\frac{(\omega+1)(\sigma-1)}{\omega\sigma}} \right)^{\omega/(\omega+1)}, \qquad (32)$$

where $k_2 > 0$ 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 $\omega \equiv \frac{\theta\sigma}{(\sigma-1)} - 1 > 0,$ just like in (9) as we initially asserted. Our earlier results in Theorems 1 and 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, analogous to (10):

<u>Theorem 3</u> (Feenstra and Kee, 2008)

Assume that the distribution of firm productivity is Pareto, as in (23). Then maximizing GDP subject to the transformation curve (32) results in $e(A_d, A_x)L$, where:

$$w = e(A_d, A_x) \equiv \frac{1}{k_2 f_e^{1/(\omega+1)}} \left[A_d^{\omega+1} f^{1-\frac{\theta}{(\sigma-1)}} + A_x^{\omega+1} f_x^{1-\frac{\theta}{(\sigma-1)}} \right]^{\frac{1}{(\omega+1)}}. \qquad (33)$$

The function $e(A_d, A_x)$ 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 (33) 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 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 5. 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:

$$\begin{aligned} \frac{w_t / P_t^H}{w_{t-1} / P_{t-1}^H} &= \frac{e(A_{dt}, A_{xt})}{e(A_{dt-1}, 0)} \left(\frac{P_t^H}{P_{t-1}^H} \right)^{-1} \\ &= \left(\frac{A_{dt}}{A_{dt-1}} \right) \left(\frac{R_{dt}}{w_t L_t} \right)^{\frac{-1}{\omega+1}} \left(\frac{P_t^H}{P_{t-1}^H} \right)^{-1} \\ &= \left(\frac{w_t / P_t^H}{w_{t-1} / P_{t-1}^H} \right)^{\frac{1}{\sigma}} \left(\frac{R_{dt}}{w_t L_t} \right)^{\frac{-1}{\omega+1}} \end{aligned} \tag{34}$$

where the first line follows from wages in Theorem 3; the second line follows from Theorem 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 (25).

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

<u>Theorem 4</u> (Arkolakis, *et al,* 2008b)

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

$$\frac{w_t / P_t^H}{w_{t-1} / P_{t-1}^H} = \left(\frac{R_{dt}}{w_t L_t} \right)^{\frac{-1}{\omega+1}\left(\frac{\sigma}{\sigma-1}\right)} = \left(\frac{R_{dt}}{w_t L_t} \right)^{-\frac{1}{\theta}}, \tag{35}$$

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 $\frac{-1}{(\sigma-1)}$ in that case with $-\frac{1}{\theta}$ in (35). This result is precisely the result derived by Arkolakis *et al* (2008b), 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 5, 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 $\frac{-1}{(\omega+1)}$, which is a straightforward application of Theorem 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 (34), 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 (35). Through these two channels, the gains equal one minus the import (or export) share with the exponent $-\frac{1}{\theta}$. This exponent exceeds $\frac{1}{(\omega+1)} = \frac{(\sigma-1)}{\theta\sigma}$ in absolute value, which appears on the last term in (34), so the feedback effect through real wages amplifies the gains from trade. But this exponent is less than $\frac{1}{(\sigma-1)}$ in absolute value, since $\theta > (\sigma-1)$, so the gain are *less than* in the Krugman (1980) model.[12]

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. Baldwin and Forslid (2009) argue that the total number of product varieties falls with trade liberalization, whereas Arkolakis *et al* (2008b) show that it can rise or fall. 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. As we now show, this ratio turns out to be *unity: the gains due to new import varieties are exactly offset for reduced domestic varieties*. Therefore, the production-side gains we have already identified in Theorem 4 are all that is available.

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

$$P^H = \left[\int_{\varphi^*}^{\infty} p(\varphi)^{1-\sigma} M\mu(\varphi)d\varphi + \int_{\varphi_x^{F*}}^{\infty} p^F(\varphi)^{1-\sigma} M^F \mu^F(\varphi)d\varphi \right]^{\frac{1}{1-\sigma}}, \tag{36}$$

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). The average prices of domestic goods appearing in (36) are:

$$\left[\int_{\varphi^*}^{\infty} p(\varphi)^{1-\sigma} M\mu(\varphi)d\varphi \right]^{\frac{1}{1-\sigma}} = \left(\frac{\sigma}{\sigma-1} \right)\left(\frac{w}{\tilde{\varphi}} \right) M^{\frac{-1}{\sigma-1}}, \tag{37}$$

which uses the prices (15) together with the definition of average productivity in (28).

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 (37), along with the fact the all imported goods are new. Applying Theorem 2 gives rise to the following ratio of unit-expenditures:

$$\frac{P_t^H}{P_{t-1}^H} = \left(\frac{w_t / \tilde{\varphi}_t}{w_{t-1} / \tilde{\varphi}_{t-1}} \right) \left(\frac{R_{dt} / w_t L_t}{M_t / M_{t-1}} \right)^{\frac{1}{\sigma-1}}. \tag{38}$$

The first term appearing on the right of (38) 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. 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, or one minus the share of spending on new imported varieties. The denominator of the second term is λ_{t-1} in Theorem 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 (38), 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 (16). The second expression appearing in (16) is $q(\varphi^*) = (\sigma - 1) f \varphi^*$, which is familiar from the Krugman model – see (7). We will combine this with the first expression appearing in (16), $r(\varphi^*)/\sigma = wf$, which can be rewritten using the inverse demand curve in (25), 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 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:

$$\left(\frac{\varphi_t^*}{\varphi_{t-1}^*}\right) = \left(\frac{w_t / P_t^H}{w_{t-1} / P_{t-1}^H}\right), \tag{39}$$

so that the increase in real wages reflects the increase in the ZCP productivities. From (30) the ratio of ZCP productivities equals the ratio of average productivities, $(\tilde{\varphi}_t / \tilde{\varphi}_{t-1})$, then comparing (38) with (39) we immediately see that $M_t / M_{t-1} = R_{dt} / w_t L_t$, as we intended to show.

These results from the Melitz (2003) model obviously challenge the empirical finding of Broda and Weinstein (2008), who treated domestic varieties as unchanged. In the next section, we consider an alternative framework to CES that allows for changes in domestic varieties as well as changes in the markups charged by firms. Changing markups have already been introduced in theory by Melitz and Ottaviano (2008), using a quadratic utility function with an additively separable numeraire good, leading to linear demand functions. As useful as that framework is, its zero income elasticities suggest that in empirical application it is best suited for partial equilibrium analysis. We will consider instead a translog expenditure function, which has income elasticities of unity and price elasticities that are not constant.

Before turning to the translog case, we conclude by noting that the gains from trade in the Melitz (2003) model have been estimated on the production side of the economy. Intuitively, movements along the transformation curve in Figure 5 due to greater export variety will be associated with higher GDP and productivity. That hypothesis is strongly confirmed empirically by Feenstra and Kee (2008). They analyze 48 countries exporting to the U.S. over 1980–2000, and find that average export variety to the United States increases by 3.3% per year, so it nearly doubles over these two decades. That total increase in export variety is associated with a cumulative 3.3% productivity improvement for exporting countries, i.e. after two decades, GDP is 3.3% higher than otherwise due to growth in export variety, on average. That estimate is *greater* than the welfare gains for the U.S. found by Broda and Weinstein (2006), which was that after 30 years, real GDP was 2.6% higher than otherwise due to growth in import variety. Of course, because the U.S. has a low import share we might expect to find greater gains to exporters, but these results still demonstrate that the gains on the production side of the economy can be substantial.

4. Translog Expenditure Function

We turn now to consider a translog unit-expenditure function. 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:[13]

$$\ln e = \alpha_0 + \sum_{i=1}^{\tilde{N}} \alpha_i \ln p_i + \tfrac{1}{2} \sum_{i=1}^{\tilde{N}} \sum_{j=1}^{\tilde{N}} \gamma_{ij} \ln p_i \ln p_j \text{, with } \gamma_{ij} = \gamma_{ji} \text{ and } \alpha_i > 0. \qquad (40)$$

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:

$$\sum_{i=1}^{\tilde{N}} \alpha_i = 1, \quad \text{and} \quad \sum_{i=1}^{\tilde{N}} \gamma_{ij} = 0\,. \qquad (41)$$

The share of each good in expenditure is obtained by differentiating (40) with respect to $\ln p_i$, obtaining:

$$s_i = \alpha_i + \sum_{j=1}^{\tilde{N}} \gamma_{ij} \ln p_j\,. \qquad (42)$$

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,…,N, while $s_j = 0$ for j=N+1,…, $\tilde{N}$. Then for the latter goods, we set $s_j = 0$ within the share equations (42), and use these $(\tilde{N} - N)$ equations to solve for the reservation prices $\tilde{p}_j$, j=N+1,…, $\tilde{N}$, in terms of the observed prices p_i, i=1,…,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), 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:

$$\gamma_{ii} = -\gamma\left(\frac{\tilde{N}-1}{\tilde{N}}\right) < 0, \text{ and } \gamma_{ij} = \frac{\gamma}{\tilde{N}} > 0 \quad \text{for } i \neq j, \text{ with } i, j = 1,\ldots,\tilde{N}. \tag{43}$$

It is readily confirmed that the restrictions in (43) satisfy the homogeneity conditions (41), and also guarantee that the reservation prices are finite. Because $\tilde{N}$ is a fixed number, (43) 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 (43) 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 as in (41), so there are $\tilde{N}-1$ free α_i parameters.[14] In addition, we have the free parameter α_0 in (40) as well as $\gamma > 0$ in (43), 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 (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 (43); 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…, $\tilde{N}$. In fact, both the CES function in (1) and the translog in (40) 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 (43) is shown by the following result:

<u>Theorem 5</u> (Feenstra, 2003; Bergin and Feenstra, 2009)

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

$$\ln e = a_0 + \sum_{i=1}^{N} a_i \ln p_i + \tfrac{1}{2} \sum_{i=1}^{N}\sum_{j=1}^{N} b_{ij} \ln p_i \ln p_j, \tag{44}$$

where: $b_{ii} = -\gamma \dfrac{(N-1)}{N} < 0$, and $b_{ij} = \dfrac{\gamma}{N} > 0$ for $i \neq j$ with i, j = 1,…,N, (45)

$$a_i = \alpha_i + \tfrac{1}{N}\left(1 - \textstyle\sum_{i=1}^{N} \alpha_i\right), \text{ for i = 1,…,N,} \tag{46}$$

$$a_0 = \alpha_0 + \left(\frac{1}{2\gamma}\right)\left\{\sum_{i=N+1}^{\tilde{N}} \alpha_i^2 + \left(\frac{1}{N}\right)\left(\sum_{i=N+1}^{\tilde{N}} \alpha_i\right)^2\right\}. \tag{47}$$

Notice that the expenditure function in (44) looks like a conventional translog function defined over the goods i=1,…,N, while the symmetry restrictions continue to hold in (45), 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 (46), 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,…,N. The final term a_0, appearing in (47), 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 5 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.[15] 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 5, 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 (47), 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* (2008b). 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.

Appendix

Using $L = \sigma\left(M_e f_e + Mf + M_x f_x\right)$ and the full employment condition, we have that:

$$\left(\frac{\sigma-1}{\sigma}\right)L = M\int_{\varphi^*}^{\infty}[q(\varphi)/\varphi]\mu(\varphi)d\varphi + M_x\int_{\varphi_x^*}^{\infty}[q_x(\varphi)/\varphi]\mu_x(\varphi)d\varphi\,,$$

Evaluating these integrals:

$$\begin{aligned}
\int_{\varphi^*}^{\infty}[q(\varphi)/\varphi]\mu(\varphi)d\varphi &= \int_{\varphi^*}^{\infty}\frac{q(\varphi^*)}{\varphi}\left(\frac{\varphi}{\varphi^*}\right)^{\sigma}\mu(\varphi)d\varphi \\
&= \frac{q(\varphi^*)}{\varphi^*}\int_{\varphi^*}^{\infty}\left(\frac{\varphi}{\varphi^*}\right)^{\sigma-1}\frac{\theta\varphi^{-\theta-1}}{(\varphi^*)^{-\theta}}d\varphi \\
&= \frac{q(\varphi^*)}{\varphi^*}\frac{\theta}{(\sigma-\theta-1)}\left(\frac{\varphi}{\varphi^*}\right)^{\sigma-\theta-1}\Bigg|_{\varphi^*}^{\infty} \\
&= f\,\frac{(\sigma-1)\theta}{(\theta-\sigma+1)},
\end{aligned}$$

where the first line uses $q(\varphi) = (\varphi/\varphi^*)^{\sigma}q(\varphi^*)$ and the last line uses $q(\varphi^*)/\varphi^* = (\sigma-1)f$.

Likewise, $$\int_{\varphi_x^*}^{\infty}[q_x(\varphi)/\varphi]\mu_x(\varphi)d\varphi = f_x\,\frac{(\sigma-1)\theta}{(\theta-\sigma+1)}.$$

Substituting these in to the full employment condition above we obtain:

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

from which it follows that $M_e = L(\sigma-1)/\sigma\theta f_e$.

Footnotes

Lead footnote: Prepared as a "State of the Art" lecture for the Canadian Economics Association meetings, Toronto, May 29-30, 2009. The author thanks Costas Arkolakis and the anonymous referees for helpful comments. Financial support was provided by the National Science Foundation grant SES-0317699.

[1] See also the early contributions of Dixit and Norman (1980, chapter 9), Lancaster (1980) and Helpman (1981); these various approaches were integrated by Helpman and Krugman (1985).

[2] Feenstra (2006) shows that an infinite reservation price leads to a well-behaved limit for the "quadratic mean of order r" index number formula of Diewert (1976), providing an alternative proof of Theorem 1 below.

[3] In particular, we are ruling out having a second sector, as sometimes introduced into this model to obtain a "home market" effect; see Krugman (1980, section III).

[4] See Arkolakis *et al* (2008a, p. 3).

[5] 1989 is omitted because West and East Germany unified then, making comparisons with later years difficult.

[6] 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 λ_τ terms.

[7] One the empirical side, papers that assess the importance of differentiated intermediate inputs include Broda, Greenfield and Weinstein (2006), Goldberg *et al* (2008), and the earlier contributions of Feenstra *et al* (1999) and Funke and Ruhwedel (2000a,b, 2001).

[8] Notice that the range $0 \le \sigma \le 1$ cannot be considered, since then all goods are *essential*

in (1), with a zero quantity for any single good resulting in zero for the entire CES aggregate. In that case the welfare gain from a new good is potentially infinite, as we find in (5) as $\sigma \to 1$.

[9] Diewert (1987, p. 499) notes that when a new good becomes available for a producer, we should use its reservation prices in the period before, when the good is not available.

[10] Interestingly, Bergstrand (1985, 1989) and Bergstrand and Baier (2001) assume a constant elasticity of transformation curve between firm outputs of each variety. The derivation here can be viewed as a micro-foundation of such an approach.

[11] Equation (14) follows from the standard CES demand function,

$q(\phi) = [p(\phi)/P^H]^{-\sigma}(wL/P^H)$, where we define the price index in (36) below.

[12] As discussed in note 3, we did not allow for a second sector in the Krugman model, which would influence the gains from trade. Likewise, Balistreri, Hillberry and Rutherford (2009) show how the gains from trade in the Melitz model are heavily influenced when a second sector is added.

[13] The translog direct and indirect utility functions were introduced by Christensen, Jorgenson and Lau (1975), and the expenditure function in (40) was proposed by Diewert (1976, p. 122).

[14] Feenstra (2003) adds an additional symmetry restriction on the α_i parameters, but Bergin and Feenstra (2009) show that Theorem 5 below can be obtained without that restriction.

[15] For other functional forms that allow for endogenous markups see Behrens *et al* (2007, 2008) and Simonovska (2009).

References

Arkolakis, Costas, Svetlana Demidova, Peter J. Klenow, and Andrés Rodríguez-Clare (2008a), 'Welfare Analysis of the Krugman Model of Trade,' unpublished Appendix to Arkolakis *et al* (2008b)

Arkolakis, Costas, Svetlana Demidova, Peter J. Klenow, and Andrés Rodríguez-Clare (2008b), 'Endogenous Variety and the Gains from Trade,' *American Economic Review*, May, 444-450

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

Baier, Scott L. and Jeffrey H. Bergstrand (2001) "The Growth of World Trade: Tariffs, Transport Costs, and Income Similarity,' *Journal of International Economics,* 53(1), 1-28.

Baldwin, Richard E. and Rikard Forslid (2009) 'Trade Liberalization with Heterogeneous Firms,' *Review of Development Economics,* forthcoming.

Balistreri, Edward J., Russell H. Hillberry and Thomas F. Rutherford (2009) 'Trade and Welfare: Does Industrial Organization Matter?' mimeo, Colorado School of Mines, University of Melbourne and ETH Zurich.

Behrens, Kristian, and Yasusada Murata (2007) 'General Equilibrium Models of Monopolistic Competition: A New Approach,' *Journal of Economic Theory*, 136, 776-787

Behrens, Kristian, Giordano Mion, Yasusada Murata, and Jens Südekum (2008) 'Trade, Wages and Productivity,' Institute for the Study of Labor, Discussion paper no. 3682

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

Bergstrand, Jeffrey H. (1985) ‘The Generalized Gravity Equation, Monopolistic Competition, and the Factor-Proportions Theory in International Trade,’ *Review of Economics and Statistics*, 71(1), February, 143-153.

Bergstrand, Jeffrey H. (1989) ‘The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence,’ *Review of Economics and Statistics*, 67(3), August, 474-481.

Broda, Christian and David E. Weinstein (2006) ‘Globalization and the Gains from Variety,’ *Quarterly Journal of Economics*, May, 121(2), 541-585

Broda, Christian, Joshua Greenfield and David E. Weinstein (2006) “From Groundnuts to Globalization: A Structural Estimate of Trade and Growth,” NBER working paper no. 12512

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

Christensen, Laurits R., Dale W. Jorgenson and Lawrence J. Lau (1975) ‘Transcendental Logarithmic Utility Functions,’ *American Economic Review* 65, 367-383.

Diewert, W. Erwin (1974) ‘Applications of Duality Theory’, in *Frontiers of Quantitative Economics*, Volume II, M. Intriligator and D. Kendrick, eds. (Amsterdam: North-Holland) 106-171

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

Diewert, W. Erwin (1980) ‘Exact Index Numbers,’ in John Eatwell, Murray Milgate, and Peter Newman, eds. *The New Palgrave: A Dictionary of Economics*, 2, 767-779

Diewert, W. Erwin (1987) ‘Aggregation Problems in the Measurement of Capital,’ in Dan Usher, ed. *The Measurement of Capital* (Chicago: University of Chicago Press) 423-528

Dixit, Avinash and Victor Norman (1980) *Theory of International Trade.* Cambridge University Press

Dixit, Avinash K. and Joseph E. Stiglitz (1977) ‘Monopolistic Competition and Optimum Product Diversity,’ *American Economic Review*, 67(3), June, 297-308

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

Ethier, Wilfred (1982) ‘Decreasing Costs in International Trade and Frank Graham’s Argument for Protection,’ *Econometrica*, 50, 1243-1268

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. (2006) ‘New Evidence on the Gains from Trade,’ *Review of World Economics/Weltwirtschaftliches Archiv*, December, 142(4), 617-641

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., Dorsati Madani, Tzu-Han Yang and Chi-Yuan Liang (1999) ‘Testing Endogenous Growth in South Korea and Taiwan,’ *Journal of Development Economics*, 60, 317-341

Feenstra, Robert C. and David E. Weinstein (2009) ‘Globalization, Competition, and the U.S. Price Level,’ UC Davis and Columbia University, in process

Funke, Michael and Ralf Ruhwedel (2001a) ‘Product Variety and Economic Growth: Empirical Evidence from the OECD Countries,’ *IMF Staff Papers*, 48(2), 225-242

Funke, Michael and Ralf Ruhwedel (2001b) ‘Export Variety and Export Performance: Evidence from East Asia,’ *Journal of Asian Economics,* 12, 493-505

Funke, Michael and Ralf Ruhwedel (2002) ‘Export Variety and Export Performance: Empirical Evidence for the OECD Countries,’ *Review of World Economics/Weltwirtschaftliche Archiv,*138(1), 97-114

Goldberg, Pinelopi K., Amit Khandelwal, Nina Pavcnik and Petia Topalova (2008) “Imported Intermediate Inputs and Domestic Product Growth: Evidence from India,” NBER working paper no. 14416

Graham, Frank (1923) ‘Some Aspects of Protection Further Considered,’ *Quarterly Journal of Economics,* 37, 199-227

Harris, Richard (1984a) ‘Applied General Equilibrium Analysis of Small Open Economies with Scale Economies and Imperfect Competition,’ *American Economic Review,* 74(5), December, 1016-32

Harris, Richard (1984b) *Trade, Industrial Policy, and Canadian Manufacturing.* Toronto: Ontario Economic Council

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

Head, Keith C. and John Ries (2001) ‘Increasing Returns Versus National Product Differentiation as an Explanation for the Pattern of US-Canada Trade,’ *American Economic Review*, 91(4), September, 858-876

Helpman, Elhanan (1981) ‘International Trade in the Presence of Product Differentiation, Economics of Scale and Monopolistic Competition: A Chamberlin-Heckscher-Ohlin Approach,’ *Journal of International Economics*, 11, 305-340

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

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

Lancaster, Kelvin (1979) *Variety, Equity and Efficiency* (New York: Columbia University Press)

Lancaster, K. (1980) ‘Intra-Industry Trade under Perfect Monopolistic Competition,’ *Journal of International Economics,* 10, May, 151-175

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

Melvin, James R. (1969) ‘Increasing Returns to Scale as a Determinant of Trade,’ Canadian Journal of Economics, 2(3), August, 389-402

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

Simonovska, Ina (2009) ‘Income Differences and the Prices of Tradables,’ University of California, Davis.

Spence, Michael (1976) ‘Product Selection, Fixed Costs, and Monopolistic Competition,’ *Review of Economic Studies*, 43(2), June, 217-235

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

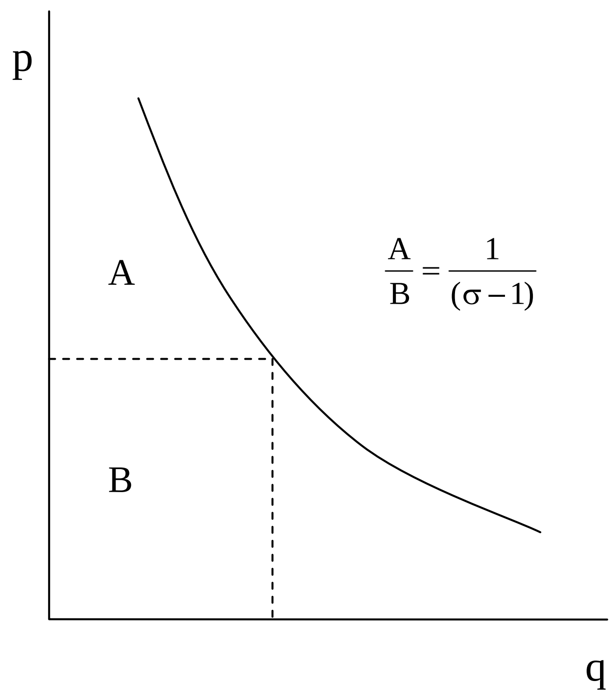

Figure 1 : CES Demand

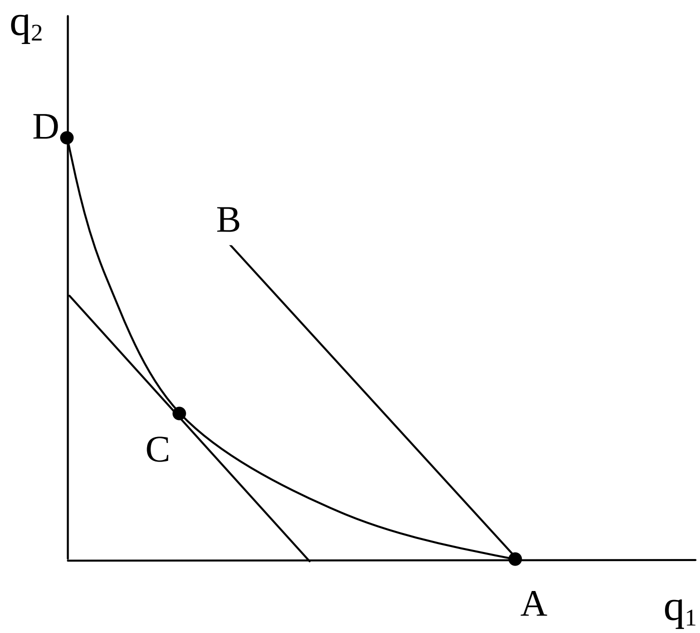

Figure 2: CES Indifference Curve

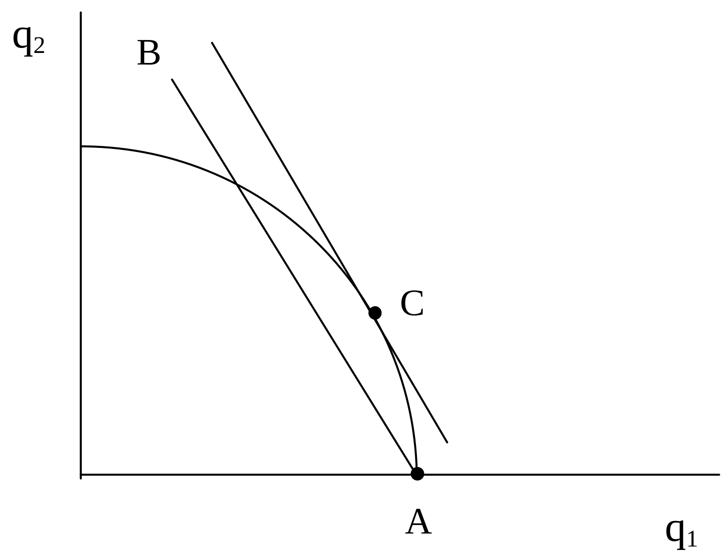

Figure 3: Constant-Elasticity Transformation Curve

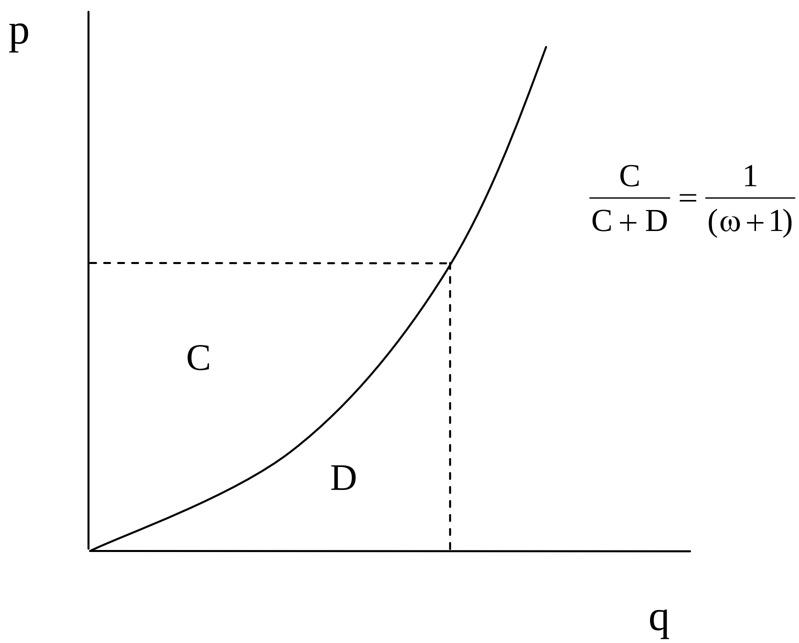

Figure 4: Constant-Elasticity Supply Curve

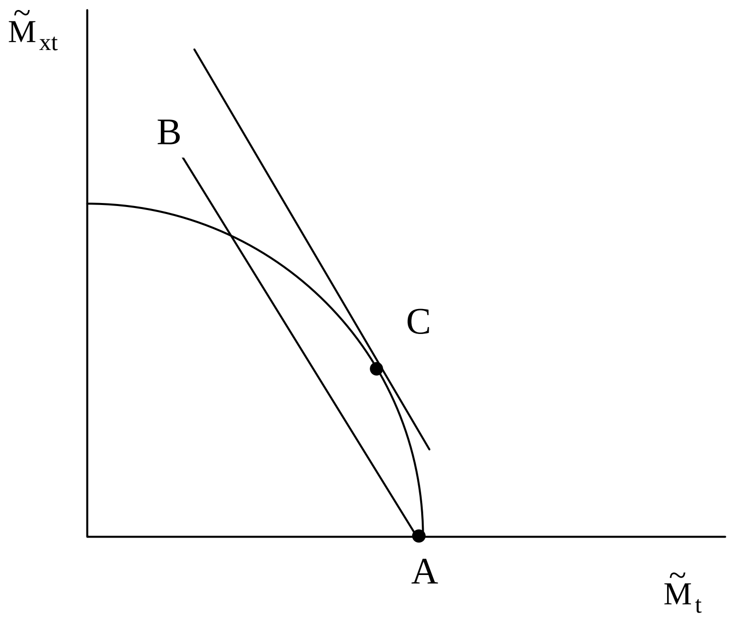

Figure 5: Constant-Elasticity Transformation Curve in Melitz (2003)