NBER WORKING PAPER SERIES

GLOBALIZATION, MARKUPS AND U.S. WELFARE

Robert C. Feenstra
David E. Weinstein

Working Paper 15749
http://www.nber.org/papers/w15749

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2010

Previously circulated as "Globalization, Markups, and the U.S. Price Level." The authors thank Jessie Handbury, Morgan Hardy, Colin Hottman, and Philip Luck for excellent research assistance, along with Sam Kortum, Bob Staiger and seminar participants at the NBER, Stanford University, UCSD, and the University of Chicago for their comments. Financial support from the NSF is gratefully acknowledged. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Globalization, Markups and U.S. Welfare
Robert C. Feenstra and David E. Weinstein
NBER Working Paper No. 15749
February 2010
JEL No. E31,F12,F4

ABSTRACT

This paper is the first attempt to structurally estimate the impact of globalization on markups, and the effect of changing markups on welfare, in a monopolistic competition model. To achieve this, we work with a class of preferences that allow for endogenous markups and firm entry and exit that are especially convenient for empirical work – the translog preferences, with symmetry in substitution imposed across products. Between 1992 and 2005 we find the U.S. market experienced a series of changes that confirm a pro-competitive effect: import shares rose and U.S. firms exited, leading to an implied fall in markups, while product variety and welfare went up. We estimate the impacts of these effects on a national level, and find that U.S. welfare rose by as much as 0.86 percent between 1992 and 2005 as a result of these changes, with product variety contributing one-half of that total.

Robert C. Feenstra
Department of Economics
University of California, Davis
One Shields Avenue
Davis, CA 95616
and NBER
rcfeenstra@ucdavis.edu

David E. Weinstein
Columbia University, Department of Economics
420 W. 118th Street
MC 3308
New York, NY 10027
and NBER
dew35@columbia.edu

1. Introduction

A promise of the monopolistic competition model in trade was that it offered additional sources of the gains from trade, beyond those from comparative advantage (e.g. Krugman, 1979, and more recently Melitz, 2003). These additional sources include: 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] While the first two sources of gains have received recent empirical attention,[2] the promise of the third source – reduced markups – has not yet been realized. To be sure, there are estimates of reduced markups due to trade liberalization for a number of countries.[3] But these cases rely on dramatic liberalizations to identify the change in markups and are not tied in theory to the monopolistic competition model. The reason that this model is not used to estimate the change in markups is because of the prominence of the constant elasticity of demand (CES) system, with its implied constant markups. To avoid that case, the above authors do not specify the functional form for demand and instead rely on a natural experiment to identify the change in markups.

For these reasons, we do not have evidence beyond these case studies about how the broad process of globalization affects markups, and particularly no evidence on the impact of such markup reductions on U.S. welfare. This paper is the first attempt to structurally estimate the impact of globalization on markups, and the effect of changing markups on welfare, in a

[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.
[2] The consumer gains due to import variety have been estimated for the U.S. by Broda and Weinstein (2006). Gains due to the self-selection of efficient firms into exporting (as in Melitz, 2003) have been demonstrated for Canada by Trefler (2004) and for a broader sample of countries by Badinger (2007b, 2008). See also Head and Ries (1999, 2001) for Canada, and Tybout *et al* (1991, 1995) for Chile and Mexico.
[3] See Levinsohn (1993) for Turkey; Harrison (1994) for the Ivory Coast; Krishna and Mitra (1998) for India; Kim (2000) for Korea; Bottasso and Sembenelli (2001) for Italy; Konings, et al (2005) for Bulgaria and Romania; and Badinger (2007a) for European countries. Most recently, De Loecker, et al (2012) have found that markups in India increased following liberalization of trade in intermediate inputs.

monopolistic competition model.[4] To achieve this we work with a class of preferences that are relatively new to that literature – the translog preferences, with symmetry in substitution imposed across products.[5] These preferences have good properties for empirical work (Diewert, 1976): they can give a second-order approximation to an arbitrary expenditure function; and correspond to the Törnqvist price index, which is very close to price index formulas that are used in practice. Furthermore, these preferences prove to be highly tractable even as the range of import varieties changes, and because they are homothetic, we prefer them to the quadratic preferences of Melitz and Ottaviano (2008) to obtain endogenous markups in a general equilibrium setting.[6]

In the translog case the elasticity of demand is inversely related to a product's market share, so markups fall as more firms enter, which we call the *pro-competitive effect*. On the other hand, domestic firms may exit as foreign competition intensifies, offsetting some of this gain to consumers. This we will refer to as the *domestic exit effect*. Incorporating these two effects into the analysis allows us to estimate the impact of globalization on markups. Furthermore, this class of preferences also allows us to address three potential criticisms of Broda and Weinstein (2006): first, that CES preferences may overstate the gains from import variety because reservation prices are infinite;[7] second, that product space might become crowded resulting in diminishing returns from new varieties, which are not measured in the CES case; and third, that the CES

[4] See also Edmond, Midrigan and Xu (2012) and Yilmazkuday (2012).

[5] These preferences have been used previously in trade by Bergin and Feenstra (2009) and Rodriguez-Lopez (2011).

[6] The quadratic preferences used by Melitz and Ottaviano (2008) lead to linear demand curves with zero income elasticity, though country population can act as a demand shift parameter. Demand curves of this type and the associated markups are estimated in the industrial organization literature: see Bresnahan (1989) and the recent trade application by Blonigen *et al* (2007). Estimates from gravity equations, however, show that when population is used as one shift parameter, country income or income per capita is also needed (e.g. Bergstrand, 1989). So the zero income elasticity assumed in Melitz and Ottaviano (2008) is not sufficient for general equilibrium analysis, where it is often desirable to work with homothetic preferences: see, for example, Bilbiie, Ghironi, and Melitz (2012), who adopt the translog preferences when analyzing markups in a dynamic model. For these reasons, we find that the translog is a very attractive functional form to model variable markups. For other non-homothetic preferences that allow for variable markups see Behrens and Murata (2007), Behrens *et al* (2008) and Simonovska (2010).

[7] While the CES system has an infinite reservation price, the area under the demand curve is still bounded above (provided the elasticity of substitution exceeds unity). But it can be expected that the gains from new product varieties in the CES case might exceed the gain from other functional forms such as translog; see note 30.

model may overstate the gains because the welfare calculations typically assume that foreign entry results in no exit of domestic firms.

We shall address all these concerns using translog preferences, and surprisingly, obtain estimates of the gains from trade for the U.S. that are similar to those under CES, but with different sources. Our point estimate for the cumulative welfare gains to the U.S. from new varieties and decreased markups in our preferred specification is 0.86 percent over the period 1992 to 2005. While much of the previous literature has focused on the gains from varieties, our analysis suggests that variety gains only account for one-half of the aggregate gains. That effect incorporates the loss in variety due to the exit of domestic U.S. firms. The remainder of the welfare gain is due to the impact of new competitors on markups. Interestingly, our *combined* gains for the U.S. due to import variety and the pro-competitive effect are of the same magnitude as Broda and Weinstein's CES estimates in our preferred specification. This finding resonates well with the recent paper of Arkolakis, Costinot, Donaldson and Rodriguez-Clare (ACDR, 2012), who argue that with a broad class of preferences including translog, the formula for the gains from trade is the same as that obtained in the CES case analyzed by Arkolakis, Costinot, Rodriguez-Clare (ACR, 2012). But despite this apparent similarity between our results and theirs, the reasons for our findings are quite different.

Specifically, ACDR focus on the second reason mentioned above for gains from trade: the efficiency gains due to improved productivity, as the most efficient firms self-select into exporting. Under their assumptions that the distribution of firm productivity is Pareto with a support that is unbounded above, and the only parameter changing between equilibria are trade costs, they find that the efficiency gains are identical in the translog and CES cases. Furthermore, by construction in their model the first and third reasons for gains from trade mentioned above do not operate: trade brings about neither changes in product variety nor in the distribution of markups, both of which remain fixed as the costs of trade change. As shown in Feenstra (2014) and by numerical methods in ACDR, if the distribution of firm productivity is instead a Pareto with a *bounded* support, then the product variety and pro-competitive effects reappear. We focus exclusively on product variety and firm markups and ignore the impact of trade on productivity through the self-selection of firms into exporting. Additional efficiency gains due to improved productivity can certainly apply in our model, but we do not attempt to measure these.

In the next section, we present some features of the data that are used to infer product variety and markups. In section 3, we introduce the translog expenditure function and solve for the cost-of-living index in the presence of new and disappearing goods, which allow the gains from new products to be measured in section 4. Firms are introduced in section 5, where we solve for the pro-competitive effect of imports. To measure the change in markups we follow the general approach of the industrial organization literature (Bresnahan, 1989; Berry, 1994): markups are not observed directly because marginal costs are unknown, so we rely on estimates of the elasticity of demand to identify the markups. We do not have the data to measure those elasticities at the firm level, so we rely instead on Herfindahl indexes of concentration for both U.S. firms and countries exporting to the U.S. We show that these Herfindahl indexes, in conjunction with the share of each country in the U.S., provide theoretically consistent ways to measure the pro-competitive and variety effects. Significantly, we have been able to obtain these indexes for all countries selling to the U.S., by land or by sea, as described in Appendix A. In section 6 we discuss the procedure for estimating the system of demand and pricing equations, and results are presented in section 7. Section 8 concludes.

2. Data Preview

One of the dramatic changes that globalization has wrought on the U.S. economy is the declining share of U.S. demand supplied by plants located in the U.S. To see this, we define U.S. domestic supply as aggregate U.S. sales less exports for agricultural, mining, and manufacturing goods (see the Appendix A for detailed definitions of all of our variables). We define U.S. apparent consumption as domestic supply plus imports. Similarly, we define the U.S. suppliers' share of the U.S. market, as U.S. domestic supply divided by apparent consumption. Finally, we define each country's U.S. import share as the exports from that country to the U.S. divided by apparent consumption.

From Table 1 we see that the share of U.S. apparent consumption sourced domestically fell by a little more than 5 percentage points between 1992 and 1997 and by 9 percentage points between 1998 and 2005. This decline corresponds to an annual decline in the U.S. share of 1.4 percentage points per year in the early period and 1.7 percentage points per year in the later period. The flip side of this decline was an almost doubling of the import share. Interestingly, the growth of imports was not uniform across countries: depending on the time period, between one-half and two-thirds of the increase was due to increases in import shares from Canada, China, and Mexico – countries that were either growing rapidly or involved in free trade agreements.[8]

One possible explanation for the findings in Table 1 is that the rise in import penetration was confined to a few important sectors. We can examine whether this was the case by looking at more disaggregated data. In Figure 1, we plot the U.S. suppliers' share in 1997 or 2005 against its level in 1992 or 1998, for each HS 4-digit category. We also place a 45-degree line in the plot

[8] The switch in U.S. classification of output data from the SIC system to the NAICS in 1998 makes it difficult to compare sectoral output levels between 1997 and 1998. We therefore break our sample into two periods (1992-1997 and 1998-2005) to maintain consistent series. For the initial tables, we will present the raw data drawn from two subsamples, but we will present results for both sub-periods and the full sample in the results section.

so that one can easily see that the vast majority of sectors lie below the 45-degree line, meaning that import penetration was steadily expanding over this time period. This pattern establishes that the rise in import penetration, though quite pronounced in some sectors, was a general phenomenon that was common across many merchandise sectors.

To make this clear, it is convenient to work with Herfindahl indexes of market concentration, defined for each country selling to the U.S. We let i denote countries, j denote firms (each selling one product), k denote sectors and t denote time. Let s_{jt}^{ik} denote firm j's exports from country i to the U.S. in sector k, as a share of country i's total exports to the U.S. in that sector. Then the Herfindahl for country i is:

$$H_{it}^{k} = \sum_{j} \left(s_{jt}^{ik} \right)^{2} \quad . \tag{1}$$

The inverse of a Herfindahl can be thought of as the effective number of equally-sized exporters, or U.S. firms, in an industry. Thus, a Herfindahl of one implies that there is one firm in the industry and an index of 0.5 would arise if there were two equally-sized firms in the sector. Similarly, if we multiply the Herfindahl by the share of the country's suppliers in the market, one obtains the market share of a synthetic typical firm in the market. This Herfindahl index is a very useful statistic because in many demand systems, the markup of a firm is monotonically increasing in its market share, and this feature will also hold in our translog system.

We were able to obtain the Herfindahl indexes for most countries selling to the U.S., by land or by sea, in 1992 and 2005 (see the Appendix A). In Table 2, we present average Herfindahls at the HS 4-digit level for the U.S. and for the 10 major exporters to the U.S.[9] As one can see from the table, the average U.S. Herfindahl rose slightly over both sub-periods,

[9] For the U.S., we have adjusted the NAICS 6-digit Herfindahls from the Bureau of Economic Analysis data so that they match the HS 4-digit categories, and detail that procedure in the data Appendix.

indicating that increased foreign competition was likely associated with some exit of U.S. firms from the market. If we multiply this average Herfindahl by the share of each country i in the U.S. consumption of good k, s_{it}^{k} we can compute the *typical market share of a firm* from that country, $H_{it}^{k}s_{it}^{k}$. We report the weighted average of these per-firm market shares in the last column of Table 2, where the weights are based on the importance of each sector in total U.S. consumption. Table 2 reveals that the share in the U.S. market of a typical U.S. firm fell slightly in the first period and by about 8 percent in the second period. By contrast exporters to the U.S. appear to have gained market share in both periods. In other words, those U.S. firms that survived ended up with smaller market shares individually while some foreign firms gained market share. But the overall average of firm shares, shown by the final averages in the last column Table 2, decline in both periods.

This data preview suggests that prior work on the impact of new varieties is likely to suffer from a number of biases. First, as foreign firms have entered the U.S. market there has been exit by U.S. firms, which serves to offset some of the gains of new varieties. Second, while U.S. Herfindahls rose, the Herfindahls of many of our largest suppliers fell. The fall in the firm-level import Herfindahls suggests that there may have been substantial variety growth that is not captured in industry level analyses. Finally, because the market shares of both U.S. firms and the average firm fell over this time period, the rise in foreign entry is likely to have depressed markups overall and therefore lowered prices.[10] Thus, estimates of the gains from new varieties

[10] While we infer the change in markups for U.S. firms using indirect evidence on their market shares and Herfindahl indexes, there is direct evidence from other countries to support the idea that trade liberalization is associated with falling markups. In addition to the work cited in our first paragraph, several studies have also shown that falls in domestic market share due to imports are associated with lower domestic markups. For example, Badinger (2007b) shows that when looking within the same manufacturing sectors, OECD countries that have higher import shares tend to have lower markups in their domestic industries. This result remains robust even when he instruments for import penetration exogenous factors related to openness such as distance and other gravity factors. Similarly, Konings et al (2005) find evidence that markups are positively related to Herfindahl indexes and

obtained from industry-level data using CES aggregators could either be too large if domestic exit is an important source of variety loss, or too small if foreign firm entry and market power losses are important unmeasured gains. We turn to quantifying these gains and losses in the next sections.

3. Translog Expenditure Function

We begin by considering a translog expenditure function defined over products denoted by i, and then will later introduce notation that allows us to consider countries, firms, and industries.[11] The translog function is defined over the universe of products, whose maximum number is denoted by the fixed number $\tilde{N}$. The expenditure needed to obtain a fixed utility (or the "ideal" price index) in period t is:

$$\ln e_t = \alpha_0 + \sum_{i=1}^{\tilde{N}} \alpha_i \ln p_{it} + \tfrac{1}{2} \sum_{i=1}^{\tilde{N}} \sum_{j=1}^{\tilde{N}} \gamma_{ij} \ln p_{it} \ln p_{jt}, \text{ with } \gamma_{ij} = \gamma_{ji}. \qquad (2)$$

Note that the restriction that $\gamma_{ij} = \gamma_{ji}$ is made without loss of generality.[12] We add the further restriction that the γ_{ij} coefficients for all goods are "symmetric," by which we mean:

$$\gamma_{ii} = -\gamma\left(\frac{\tilde{N}-1}{\tilde{N}}\right) < 0, \text{ and } \gamma_{ij} = \frac{\gamma}{\tilde{N}} > 0 \text{ for } i \neq j, \text{ with } i, j = 1,\ldots,\tilde{N}. \qquad (3)$$

These symmetry restrictions have not been used in past work dealing with estimating the translog function, but are essential in our monopolistic competition model.

---

that markups in concentrated sectors fall significantly when trade liberalization leads to higher levels of import penetration. This study provides additional evidence for a negative relationship between domestic market share and markups. Finally, Konings and Vandenbussche (2005) find that antidumping protection not only raises the markups of European firms, but does so even more when importers have high market shares. The fact that protection raises domestic markups more when domestic market shares are low provides additional support for the notion that firms with low market shares are likely to have relatively low markups.

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

[12] We also require $\sum_{i=1}^{\tilde{N}} \alpha_i = 1$, $\sum_{i=1}^{\tilde{N}} \gamma_{ij} = 0$ for the expenditure function to be homogeneous of degree one.

The share of each good in expenditure can be computed by differentiating (2) with respect to $\ln p_i$, obtaining:

$$
\begin{aligned}
s_{it} &= \alpha_i + \sum_{j=1}^{\tilde{N}} \gamma_{ij} \ln p_{jt} \\
&= \alpha_i - \gamma\left(\ln p_{it} - \overline{\ln p(t)}\right), \quad \text{with} \quad \overline{\ln p(t)} \equiv \sum_{j=1}^{\tilde{N}} \frac{1}{\tilde{N}} \ln p_{jt},
\end{aligned}
\tag{4}
$$

where the second line is obtained using the symmetry conditions in (3). Notice that the term $\overline{\ln p(t)}$ refers to the average over all the prices of *all goods* in period t. But that average will include goods that are not available to the consumer, in which case the appropriate price to use in the expenditure function is the reservation price, where demand equals zero. So we want to solve for the reservation prices for goods not available, and then re-express the expenditure and share equations so that they no longer depend explicitly on these reservation prices.

We let I_t denote the set of goods available each period, with number N_t. We can solve for the reservation price of the goods not available by setting the share equation in (4) equal to zero. More generally, we can invert the share equation to solve for the prices of all goods in terms of their shares:

$$
\ln p_{it} = \frac{\alpha_i - s_{it}}{\gamma} + \overline{\ln p(t)}, \tag{4'}
$$

which equals the reservation price of good i when $s_{it} = 0.$ It is convenient to substitute (4') back into the expenditure function (2) to obtain an expression for the cost-of-living in terms of shares. After some simplification, we obtain:[13]

$$
\ln e_t = \alpha_0 + \frac{1}{2\gamma} \sum_{i=1}^{\tilde{N}} (\alpha_i)^2 + \overline{\ln p(t)} - \frac{1}{2\gamma} \sum_{i=1}^{\tilde{N}} (s_{it})^2 . \tag{5}
$$

---

[13] See Appendix B.

While this expression is close to what we shall use to evaluate the cost-of-living, we still need to eliminate the reservation prices within the overall average $\overline{\ln p(t)}$. That is done by expanding this average as:

$$\overline{\ln p(t)} \equiv \sum_{j=1}^{\tilde{N}} \frac{1}{\tilde{N}} \ln p_{jt} = \frac{1}{\tilde{N}}\left(\sum_{j\in I_t} \ln p_{jt} + \sum_{j\notin I_t} \left(\frac{\alpha_j}{\gamma} + \overline{\ln p(t)} \right) \right),$$

where the reservation prices for goods $j \notin I_t$ in the final term are obtained from (4) with $s_{jt} = 0$. Since the number of goods available is $N_t = \tilde{N} - \sum_{j\notin I_t} 1$, it follows that:

$$\overline{\ln p(t)} = \frac{1}{N_t}\left(\sum_{j\in I_t} \ln p_{jt} + \sum_{j\notin I_t} \frac{\alpha_j}{\gamma} \right) = \overline{\ln p_t} + \frac{\alpha_t}{\gamma}, \tag{6}$$

where:[14] $\overline{\ln p_t} \equiv \frac{1}{N_t} \sum_{j\in I_t} \ln p_{jt}$ and $\alpha_t \equiv \frac{1}{N_t}\left(\sum_{j\notin I_t} \alpha_j \right) = \frac{1}{N_t}\left(1 - \sum_{j\in I_t} \alpha_j \right)$.

Notice that $\overline{\ln p_t}$ is the average of prices over the goods *actually available*, while α_t is inversely related to the range of goods available: as the set I_t expands, then the summation $\sum_{j\in I_t} \alpha_j$ rise and N_t also rises, so α_t falls. We can substitute these two terms into the expenditure function (5) to obtain:

$$\ln e_t = \alpha_0 + \frac{1}{2\gamma} \sum_{i=1}^{\tilde{N}} (\alpha_i)^2 + \overline{\ln p_t} + \frac{\alpha_t}{\gamma} - \frac{1}{2\gamma} \sum_{i=1}^{\tilde{N}} (s_{it})^2 .$$

Taking the change in the expenditure function the first two terms vanish and we obtain:

$$\ln\left(\frac{e_t}{e_{t-1}} \right) = \Delta \overline{\ln p_t} + \frac{\Delta \alpha_t}{\gamma} - \frac{1}{2\gamma} \Delta \sum_{i=1}^{N} s_{it}^2 . \tag{7}$$

Equation (7) is a useful expression for the change in the cost of living and consists of three terms. The first term is an unweighted average of the change in log-prices for available products, and will cancel out is subsequent formula, so we focus on the next two terms. The

---

[14] The final equality uses $\sum_{i=1}^{\tilde{N}} \alpha_i = 1,$ as in the previous footnote.

second term is a shift parameter $\Delta\alpha_t/\gamma$ where, as noted above, where the entry of new products will result in $\Delta\alpha_t < 0$ so that the cost of living falls. The third term is the change in the Herfindahl index of product shares.

To interpret the second term, denote the set of products available both periods by $\bar{I} \equiv I_t \cap I_{t-1}$, with $\bar{N} > 0$ elements. For simplicity, suppose that the prices of all products are unchanging. Then $\Delta s_{it} = \Delta\alpha_t$, and averaging over the $\bar{N}$ products we obtain:

$$\Delta\alpha_t = \frac{1}{\bar{N}}\sum_{i\in\bar{I}}\Delta s_{it} = \frac{1}{\bar{N}}\left[\left(1-\sum_{i\notin\bar{I}}s_{it}\right)-\left(1-\sum_{i\notin\bar{I}}s_{it-1}\right)\right] = -\frac{1}{\bar{N}}\left(\sum_{i\notin\bar{I}}s_{it}-\sum_{i\notin\bar{I}}s_{it-1}\right),$$

where we use the fact that the shares sum to unity each period. Thus, $-\Delta\alpha_t$ is directly related to the share of expenditure on new goods minus the share of expenditure on disappearing goods. Feenstra (1994) shows that those shares, in conjunction with the elasticity of substitution, determine the gains from variety in the CES case. To develop this comparison with the CES case more carefully, note that from Feenstra (1994) we can express the "ideal" CES price index as:

$$\ln\left(\frac{e_t}{e_{t-1}}\right) = \sum_{i\in\bar{I}} w_{it}\Delta\ln p_{it} + \frac{1}{(\sigma-1)}\left(\ln\sum_{i\notin\bar{I}}s_{it} - \ln\sum_{i\notin\bar{I}}s_{it-1}\right), \qquad (8)$$

where $\bar{I} \equiv I_t \cap I_{t-1}$ is the common set of goods in both periods. The first expression on the right of (8) is a weighted average of the log price changes, similar in spirit to the unweighted average appearing as the first term on the right of (7).[15] The second expression on the right of (8) is the welfare gain from new and disappearing goods in the CES case, which involves dividing the net change in expenditure on new and disappearing goods by the elasticity of substitution minus unity, $(\sigma - 1)$. As we have shown above, that expression is very similar to the second term $\Delta\alpha_t/\gamma$ on the right of (7), which likewise involves dividing the net change in expenditure on new and

[15] The weights w_{it} appearing in (8) are the Sato-Vartia weights defined in Feenstra (19994, pp. 158-159).

disappearing goods by the translog parameter γ. The expenditures on new and disappearing goods are measured in logs and levels, respectively, in the CES and translog cases, and the substitution parameters $(\sigma - 1)$ and γ correspondingly differ, but otherwise there is a high degree of similarity between the CES and translog expressions. The key difference between (7) and (8) is that in the translog case we have an *additional term* in (8), the change in the Herfindahl index of product shares, that has no analogue in the CES case. Notice that a fall in the Herfindahl index *raises* the cost-of-living index in (7), which is surprising because a fall in the Herfindahl indicates less concentration of shares, due to more new products for example. So why does the cost-of-living go up in that case? The answer illustrates a key difference between the translog and CES formulas.

Consider replacing one good having a high share parameter α_{1t-1} in period t-1 with two other goods having lower values of α_{2t} and α_{3t} in period t (while goods 4,…,N are unchanged). We choose these parameters so that the total share of spending on goods 1 and 2+3 are identical, $s_{1t-1} = s_{2t} + s_{3t}$ (with s_{4t}, …, s_{Nt} unchanged). Because there are two goods instead of one, the Herfindahl index is lower: $s_{1t-1}^2 + \sum_{i=4}^{N} s_{it-1}^2 > s_{2t}^2 + s_{3t}^2 + \sum_{i=4}^{N} s_{it}^2$, provided that $s_{2t}, s_{3t} > 0$. That would *raise* the cost of living in (6) because product space is more "crowded," so that goods are more substitutable.[16] In other words, as the number of varieties increases, we care less about each new variety. This explains why the Herfindahl index of product shares enters the cost-of-living index in a counter-intuitive way. Notice that this crowding effect does not occur in the CES case, because replacing one good with two new goods that have the same total expenditure would have no impact at all on consumer well being.

[16] This claim can be confirmed from the elasticity of demand in (17). With identical prices for all goods, so $s_{Nt} = 1/N$, this elasticity becomes $1 + \gamma(N - 1)$ which is increasing in N.

4. Consumer Gains from New Varieties

In order to implement the cost-of-living formula in (6), we estimate the shift parameter $\Delta\alpha_t$ as well as the translog parameter γ. Because we do not have firm-level data we shall rely on Herfindahl indexes of U.S. firms and of firms exporting to the U.S. To show how the aggregation over firms is done in the theory, we will henceforth let $i \in I_t$ denote the countries selling to the U.S. each period (including the U.S. itself) and $j \in J_{it}$ denote their firms. We shall assume that each firm sells one product in a given sector with its share denoted by s_{ijt}. The total import share from country i is $s_{it} \equiv \sum_{j \in J_{it}} s_{ijt}$, and we let $s^i_{jt} \equiv s_{ijt} / s_{it}$ denote firms' shares *within* the exports of country i. Then the Herfindahl index for country i is $H_{it} \equiv \sum_{j \in J_{it}} (s^i_{jt})^2$.

We re-writing the share equation in (4) and (6) using this new notation as:

$$s_{ijt} = (\alpha_{ij} + \alpha_t) - \gamma\left(\ln p_{ijt} - \overline{\ln p_t}\right), \qquad (9)$$

where $\alpha_t \equiv \frac{1}{N_t}(1 - \sum_{i,j} \alpha_{ij})$ is the shift parameter already discussed, and $\overline{\ln p_t} \equiv \frac{1}{N_t}\sum_{i,j} \ln p_{ijt}$ is the average log-price of all available goods in period t.[17] Multiplying the above equation by the firms' shares within the exports of a country, $s^i_{jt} \equiv s_{ijt} / s_{it}$, and summing over firms with $\sum_{j \in J_{it}} s^i_{jt} = 1$, we obtain:

$$H_{it} s_{it} = \alpha_{it} + \alpha_t - \gamma\left(\ln p_{it} - \overline{\ln p_t}\right), \quad i \in I_t,$$

where $\alpha_{it} \equiv \sum_{j \in J_{it}} s^i_{jt} \alpha_{ij}$ and $\ln p_{it} \equiv \sum_{j \in J_{it}} s^i_{jt} \ln p_{ijt}$ are weighted averages of the taste parameters and log prices, respectively. It is natural to model α_{it} as a country fixed effect plus an error term:

$$\alpha_{it} = \alpha_i + \varepsilon_{it}. \qquad (10)$$

[17] We use the summation $\sum_{i,j}$ as a shorthand for $\sum_{i \in I_t} \sum_{j \in J_{it}}$, i.e. to sum across all available products each period.

Substituting this equation above, we obtain the share equation,

$$H_{it}s_{it} = \alpha_i + \alpha_t - \gamma\left(\ln p_{it} - \overline{\ln p_t}\right) + \varepsilon_{it} \ , \quad i \in I_t \, . \tag{11}$$

In the next section we show how the translog parameter γ is obtained using an estimator that differences across time and across countries, thereby eliminating the fixed effects α_i and α_t in (11). Supposing that we have the translog parameter, we can move the prices to the left in (11) and therefore obtain an unbalanced panel from which we can estimate these fixed effects. But since we are only interested in $\Delta\alpha_t$, the simplest approach it to difference (11) for a non-empty set of countries $i \in \overline{I} \subseteq I_{t-1} \cap I_t$, thereby eliminating α_i, so that we are left with:

$$\Delta\left(\frac{\alpha_t}{\gamma} + \overline{\ln p_t}\right) = \frac{1}{\gamma}\Delta(H_{it}s_{it}) + \Delta \ln p_{it} - \Delta\varepsilon_{it} \ , \quad i \in \overline{I} \, . \tag{12}$$

By taking a weighted average of the terms on the right, we obtain the terms on the left, which is what we need to construct the cost-of-living index in (6). The weights that we shall use to form the average are the average shares shown in the Törnqvist index (9), restricted to the countries $i \in \overline{I}$. Specifically, define:

$$\overline{s}_{it} \equiv s_{it} + \frac{1}{\overline{N}}\left(1 - \sum_{j \in \overline{I}} s_{jt}\right), \ \text{for} \ i \in \overline{I} \, , \tag{13}$$

as the shares for the countries in $\overline{I}$, re-normalized to sum to unity. Then we apply the Törnqvist weights $\tfrac{1}{2}(\overline{s}_{it-1} + \overline{s}_{it})$ to (12) and sum over $i \in \overline{I}$ to obtain:

$$\Delta\left(\frac{\alpha_t}{\gamma} + \overline{\ln p_t}\right) = \sum_{i \in \overline{I}} \tfrac{1}{2}(\overline{s}_{it-1} + \overline{s}_{it})\left[\frac{1}{\gamma}\Delta(H_{it}s_{it}) + \Delta \ln p_{it} + \Delta\varepsilon_{it}\right]. \tag{14}$$

To measure the left-hand side precisely, we need all three terms in brackets on the right, but only the first two are observed. The final term on the right is the change in the weighted-average taste parameters for each country, $\Delta\varepsilon_{it} = \Delta\alpha_{it} \equiv \Delta\sum_{j \in J_{it}} s^i_{jt}\alpha_{ij}$. It is difficult to measure exact price indexes when taste parameters are changing, and this difficulty appeared in the CES case analyzed by Feenstra (1994). To overcome it, Feenstra assumed that there was a subset of countries $i \in \bar{I}$ for which the CES taste parameters were constant. In that case, it became possible to measure the exact CES price index for that subset of countries, while the remaining terms indicated the gains from new or disappearing countries as well as those with changing taste parameters. Taking the same approach here, we set $\Delta\varepsilon_{it} = 0$ in (14) for $i \in \bar{I}$, and substitute that expression into (6) to obtain the result:

Proposition 1

Suppose that there is a set of countries $i \in \bar{I} \subseteq I_{t-1} \cap I_t$, $\bar{I} \neq \varnothing$, with $\Delta\alpha_{it} = 0$. Then the Törnqvist cost-of-living index is:

$$\ln\left(\frac{e_t}{e_{t-1}}\right) = \sum_{i\in\bar{I}} \frac{1}{2}(\bar{s}_{it-1} + \bar{s}_{it}) \ln p_{it} - \underbrace{\left\{ -\sum_{i\in\bar{I}} \frac{1}{2\gamma}(\bar{s}_{it-1} + \bar{s}_{it}) \Delta(H_{it}s_{it}) + \frac{1}{2\gamma}\Delta\left(\sum_{i\in I_t} H_{it}s_{it}^2\right)\right\}}_{V_t}, \qquad (15)$$

where the final terms define the variety effect V_t. If $\Delta\alpha_{it} = 0$ for $i \in \bar{I} = I_{t-1} = I_t$, then:

$$V_t = -\frac{1}{2\gamma}\sum_{i\in\bar{I}} s_{it-1}s_{it}\Delta H_{it} \,. \qquad (16)$$

The first term on the right of (15) is the conventional Törnqvist cost-of-living index defined over the prices of those countries $i \in \bar{I}$ with constant taste parameters. The additional terms define V_t as the extra impact of product variety on the cost of living, with $V_t > 0$ indicating greater variety and lower cost of living. The variety effect itself is defined by two terms: the first is the effective firm shares $H_{it}s_{it}$, with falling shares indicating welfare gains. In Table 1, we saw that $H_{it}s_{it}$ fell for the United States and rose for some other exporters, but on average fell over both sample periods. This average decline in the typical market share of a firm indicates that there is increased spending on new countries and products, so we should expect to see a gain from variety. But that gain is potentially offset by the change in the *overall* Herfindahl appearing as the last term in (15), which is the share-weighted average of $\Delta(H_{it}s_{it})$ as shown in the last row of Table 1, and reflects crowding in product space. We note that this decomposition of the total welfare change in (15) is closer to the CES decomposition that we reported in (8), because in both cases the first term is the exact price index over the set of goods $i \in \bar{I} \subseteq I_{t-1} \cap I_t$, while the remaining terms reflect the cost-of-living reduction from new varieties: in the translog case in (15), those remaining terms include the change in the overall Herfindahl, reflecting crowding.

In (16) we provide a simplified expression for V_t that holds if there were no change in the set of countries selling to the United States, so that we could choose $\bar{I} = I_{t-1} = I_t$.[18] In the CES case considered by Feenstra (1994) and Broda and Weinstein (2006), there would then be no gains from import variety because there are no new countries selling to the U.S. But now we can use the Herfindahl indexes for each country to infer entry or exit of firms and products: falling Herfindahls due to entry will contribute towards gains with $V_t > 0$, whereas exit will contribute towards $V_t < 0$. We stress, however, that (16) is valid if and only if there is no change in the average taste parameters for all countries, $\Delta\alpha_{it} = 0$ for $i \in \bar{I} = I_{t-1} = I_t$.[19]

The remaining question is how the set of countries $\bar{I} \subseteq I_{t-1} \cap I_t$, for which the taste parameters are constant, will be chosen. Broda and Weinstein (2006) chose the set $\bar{I}$ as the

---

[18] The proof of (16) is provided in Appendix B.

[19] Notice that this assumption did not hold in the example we provided in section 3, where one good having a high share parameter α_{1t-1} was replaced with two other goods having lower values of α_{2t} and α_{3t} (while goods 4,…,N are unchanged). We chose those parameters so that the total share of spending on goods 1 and 2+3 were identical, $s_{1t-1} = s_{2t} + s_{3t}$, resulting in a *fall* in the Herfindahl index and a consumer loss. In contrast, a fall in the Herfindahl index in (16), with no change in taste parameters, will lead to consumer gains.

intersection of countries supplying in the first and last years of the sample. So all countries selling to the United States in the initial year, and not disappearing by the end of the sample, are presumed to have no expansion in product varieties *within* their exports to the United States. We can improve on this assumption by using information on the Herfindahl indexes of exporters to the U.S. in each sector. We shall interpret Herfindahl indexes that are changing by more than a tolerance as evidence of large changes in α_{it}, and exclude these countries from the set $\bar{I}$, as will be discussed in more detail in section 7.

5. Producers and the Pro-Competitive Effect

To develop a simple general equilibrium expression for welfare, suppose that firms selling to the U.S. each produce a single product and act as Bertrand competitors. The profit maximization problem for a firm j exporting from country i in period t is,

$$\max_{p_{ijt}>0} \; p_{ijt} x_{ij}(p_t, E_t) - C_{ij}[x_{ij}(p_t, E_t)] \;,$$

where $x_{ij}(p_t, E_t)$ denotes the demand arising from the translog system, with the price vector p_t and expenditure E_t, and $C_{ij}[x_{ij}(p_t, E_t)]$ is the cost of production. We denote the elasticity of demand by $\eta_{ijt} \equiv -\partial \ln x_{ij}(p_t, E_t) / \partial \ln p_{ijt}$, which from (9) is:

$$\eta_{ijt} = 1 - \left(\frac{\partial \ln s_{ijt}}{\partial \ln p_{ijt}} \right) = 1 + \frac{\gamma(N_t - 1)}{s_{ijt} N_t} \;. \qquad (17)$$

Then the optimal price can be written as the familiar markup over marginal costs:

$$\ln p_{ijt} = \ln C'_{ijt} + \ln\left(\frac{\eta_{ijt}}{\eta_{ijt} - 1} \right) = \ln C'_{ijt} + \ln\left[1 + \frac{s_{ijt} N_t}{\gamma(N_t - 1)} \right],$$

where $C'_{ijt} \equiv C'_{ij}[x_{ij}(p_t, E_t)]$ denotes marginal cost.

We do not have the firm-level data needed to measure the above expression directly. But we can aggregate it by taking a weighed average using the firm shares $s^i_{jt} \equiv s_{ijt} / s_{it}$ within the exports of country i. Then the geometric average of prices from country i is:

$$\ln p_{it} \equiv \sum_j s^i_{jt} \ln p_{ijt} = \ln C'_{it} + \sum_{j \in J_{it}} s^i_{jt} \ln\left[1 + \frac{s_{it} s^i_{jt} N_t}{\gamma(N_t - 1)}\right], \quad (18)$$

where $\ln C'_{it} \equiv \sum_j s^i_{jt} \ln C'_{ijt}$ are the average marginal costs. We are not able to measure (18) directly in the absence of firm-level information, but data on the Herfindahl indexes for each exporting country will allow us measure a first-order approximation to (18) around the point where the firm shares are equal. Recall that the inverse of the Herfindahl can be thought of as the effective number of equally-sized firms, each of which would then have the share $s^i_{jt} = H_{it}$. Since the first-order approximation to the log function is $\ln(1+x) \approx \ln(1+a) + \left(\frac{x-a}{1+a}\right)$ around x=a, we apply this formula to (18) around $s^i_{jt} = H_{it}$ to obtain:

$$\ln p_{it} \approx \ln C'_{it} + \sum_{j \in J_{it}} s^i_{jt} \left\{ \ln\left[1 + \frac{s_{it} H_{it} N_t}{\gamma(N_t - 1)}\right] + \left[\frac{\frac{s_{it} s^i_{jt} N_t}{\gamma(N_t - 1)} - \frac{s_{it} H_{it} N_t}{\gamma(N_t - 1)}}{1 + \frac{s_{it} H_{it} N_t}{\gamma(N_t - 1)}}\right] \right\} = \ln C'_{it} + \ln\left[1 + \frac{s_{it} H_{it} N_t}{\gamma(N_t - 1)}\right], \quad (19)$$

which is obtained since $H_{it} \equiv \sum_{j \in J_{it}} (s^i_{jt})^2$ and $\sum_{j \in J_{it}} s^i_{jt} = 1$. The average country markup appearing as the final term in (19) is *the markup for a firm with the average share* $s_{ijt} = s_{it} H_{it}$. In other words, we are ignoring the variation in firm sizes *within countries* when we compute the "average" markup for each country in (19). In the absence of having such firm-level information for all countries exporting to the United States, however, we believe that this average markup is the best that we can do.[20,21]

---

[20] In order to assess the importance of limitation of this approach, we simulated firm-level data to compare the true markup in (18) with the approximation in (19). We begin by assuming that γ=0.19 (which is our median estimate) and that the underlying firm sales distribution follows Zipf's Law, i.e. we used a Pareto distribution with a shape parameter of 1 and a minimum value of 10. Then we assumed that there were 500 firms in each sector and that every

To obtain an expression for welfare, suppose that labor is the only factor of production, with wage w_t. We make the key assumption that profits of firms are zero under monopolistic competition, so that welfare of the representative consumer is $W_t = w_t / e(p_t)$, or the real wage. The change in the real wage is $\Delta \ln W_t = \Delta \ln w_t - \Delta \ln e(p_t)$. We evaluate the cost-of-living index from (15), where the first term on the right is the conventional Törnqvist index defined over the average log-prices for each country. Those averages are in turn approximated as in (19). Combining these results, we have:

Proposition 2

The change in welfare of the representative consumer with translog preferences is:

$$\Delta \ln W = \Delta \ln w_t - \Delta \ln e(p_t) \approx \left[\Delta \ln w_t - \sum_{i \in \bar{I}} \tfrac{1}{2}(\bar{s}_{it-1} + \bar{s}_{it}) \Delta \ln C'_{it} \right] + P_t + V_t, \qquad (20)$$

where:

$$P_t \equiv -\sum_{i \in \bar{I}} \tfrac{1}{2}(\bar{s}_{it} + \bar{s}_{it-1}) \Delta \ln \left[1 + \frac{s_{it} H_{it} N_t}{\gamma (N_t - 1)} \right]. \qquad (21)$$

The approximation in (20) depends on the first-order approximation to prices in (19), and also assumes that the change in the taste parameters $\Delta\alpha_{it}$ for $i \in \bar{I}$ is small, as was assumed in Proposition 1. The term in brackets on the right in (20) is the change in wages relative to a Törnqvist index of marginal costs, for both domestic and foreign-produced goods. Its role in

firm in each sector except the first firm had a sales share drawn from this Pareto distribution. If we impose that the sector must have the actual Herfindahl index in the data, and the sales of all but the first firm are drawn from the Pareto distribution, then the Herfindahl index implicitly defines the sales of the first firm. We then used this simulated firm-level data to construct the true markup equation, i.e. the second term on the right of (18). We then calculated the percentage difference of the approximation, the last term on the right of (19) from the actual simulated values as: diff = [(actual – approx)/ (actual value)], where Jensen's inequality guarantees that diff ≤ 0. When we applied this method using the Herfindahl indexes for each sector and country selling to the U.S., we found that the median difference was -0.005 percent and the mean difference was -0.28 percent, indicating the first-order approximation holds quite closely.

[21] It is possible that (19) overstates the average markup if the Herfindahl index of exporters is computed by bundling products within a single shipment, making firms appear bigger than they really are. We used the PIERS data, described in Appendix A, to check for the prevalence of trading companies, i.e. those whose names including "trading", "wholesale", "import" or "export", "group", or abbreviations of these terms. There were about 8,000 such firms, and they accounted for 5 percent of exports to the U.S. in 1992 and 7.5 percent in 2005.

welfare is similar to the real earnings of home factors in term of imported goods – or the "single factoral terms of trade" introduced by Viner (1937) – except that in (20) we have stripped out firm markups. Viner introduced this concept because it is highly relevant to the welfare gains from trade, and we agree. For example, this term captures the positive effect of import competition on firm selection and welfare, through forcing the exit of less-efficient firms and lowering average costs. That effect is the focus of ACR (2012) and ACDR (2012), but we do not attempt to measure it here.[22]

We focus on the remaining terms in welfare. The pro-competitive effect P_t in (21) once again depends on the change in $H_{it}s_{it}$, and from Table 1, that typical market share of a firm was on average falling over both sample periods. This decline provides the intuition for why we should expect to see a welfare gain from reduced markups as measured by P_t. A final approximation is needed to measure the pro-competitive effect, and that concerns the number of available product N_t which appears in (21). We shall consider two methods of making this approximation, the choice of which does not matter for the results.

One approach is to measure N_t by the inverse of the overall Herfindahl shown in (8), which would equal the number of synthetic, equally-sized firms in each industry. That approach is a lower-bound to the number of products, however, since firms might not be equally-sized and they might sell more than one product each. A second approach is to treat the term $N_t/(N_t – 1)$ appearing in (21) as close enough to unity to be ignored. This approach follows from a standard assumption in monopolistic competition models that firms ignore the impact of their prices on the overall price index. In the demand equation (9), the average log-prices $\overline{\ln p_t}$ acts like a price

[22] This term would be considerably more complex once we allow for multiple factors of production and realistic input-output linkages in the economy. For example, Feenstra et al (2013) argue that a portion of the productivity increase in the United States after the mid-1990's is due to improvements in the terms of trade, especially for high-technology goods. Measuring the full impact of trade on U.S. productivity is beyond the scope of this paper.

index for the market. If the firm does not consider the impact of its own price on this price index, then the elasticity of demand is,

$$\eta_{ijt} = 1 - \left.\frac{\partial \ln s_{ijt}}{\partial \ln p_{ijt}}\right|_{\overline{\ln p_t}} = 1 + \left(\frac{\gamma}{s_{ijt}}\right) .$$

Comparing this result with (17), we see that the two elasticities are the same if and only if N_t is sufficiently large so that $N_t/(N_t - 1) \approx 1$. We will experiment with both approaches when measuring N_t in order to obtain the pro-competitive effect.

Proposition 2 is our final expression for the change in welfare, and depends on the pro-competitive effect P_t and the variety effect V_t. Of course, these terms can differ across sectors, so we need to sum the welfare gains across all of the manufacturing sectors in our sample. We do this by adding an industry superscript k to P_t and V_t, then multiplying the welfare gain in each industry by the appropriate translog weight, and summing across all sectors,

$$\Delta \ln W_t^{PV} = \sum_k \tfrac{1}{2}\left(s_{t-1}^k + s_t^k\right)\left(P_t^k + V_t^k\right) , \tag{22}$$

where s_t^k is the share of sector k in U.S. absorption in period t.[23] Since the consumption of merchandise only accounts for about 19 percent of U.S. absorption, a drop in the price of manufactured goods translates into about a one-fifth as large gain in aggregate welfare. That is, we multiplied all our welfare calculations from (22) by 0.19 to present them as relative to the entire U.S. economy. We omit the sector k notation in the next section, though our estimation is done separately over each sector.

[23]While we had the appropriate shares of each sector within manufacturing for 1992 and 2005, we used data from the benchmark input-output tables to obtain the aggregate share of merchandise apparent consumption in total U.S. final demand. We computed this by using the 2002 use table from the benchmark input-output table before redefinitions. We set apparent consumption of merchandise equal to agricultural, mining, and manufacturing value added less exports plus imports. Total U.S. absorption was set equal to U.S. final goods demand.

6. Estimation

Our next task is to address how to estimate the translog parameter γ , which is obtained from the share equation (11). The important property of this share equation is that the parameter γ *does not* depend on the set of goods available, which only influences the time-effect α_t. We can expect that the prices appearing in (11) are endogenous, as in a conventional supply and demand system. For the CES case, Feenstra (1994) showed how this endogeneity could be overcome by specifying the supply equation and assuming that the demand and supply errors are uncorrelated: that moment condition was used to estimate the model parameters.[24] We will follow the same procedure in the translog case, as described next.

We use the first-order approximation to obtain the weighted average of prices in (19), which in practice are measured by the unit-value of imports from each source country and industry, and the price index within each industry for the United States. We further specify that the weighted average of marginal costs from each exporting country take on the iso-elastic form:

$$\ln C'_{it} = \omega_{i0} + \omega \ln\left(\frac{s_{it}E_t}{p_{it}}\right) + \delta_{it} ,$$

where the term $(s_{it}E_t / p_{it})$ reflects the total quantity exported from country i, and δ_{it} is an error term. Combining the above equation with (19), we obtain a modified pricing equation:

$$(1+\omega)\ln p_{it} = \omega_{i0} + \omega \ln s_{it} + \omega \ln E_t + \ln\left[1 + \frac{H_{it}s_{it}N_t}{\gamma(N_t - 1)}\right] + \delta_{it} . \qquad (23)$$

We see that the translog parameter γ appears in both the share equation (11) and the pricing equation (23): larger γ means that the goods are stronger substitutes and the markups are correspondingly smaller. It is also evident that the shares and prices are endogenously

---

[24] Identification of the model parameters from this moment condition depended on having heteroskedasticity in second-moments of the data, so this is an example of "identification through heteroskedasticity," as discussed more generally by Rigobon (2003).

determined: shocks to either supply δ_{it} or demand ε_{it} will both be correlated with shares s_{it} and prices p_{it}. To control for this endogeneity we will estimate these equations simultaneously using a similar methodology to that proposed in the CES case by Feenstra (1994) and extended by Broda and Weinstein (2006).

The first step in our estimation is to difference (11) and (23) with respect to country k and with respect to time, thereby eliminating the terms $\alpha_i + \alpha_t$ and the overall average prices $\overline{\ln p_t}$ appearing in the share equations and eliminating total expenditure $\ln E_t$. We also divide the share equation by γ and the pricing equation by $(1+\omega)$, and then express each equation in terms of its error term:

$$\frac{(\Delta\varepsilon_{it} - \Delta\varepsilon_{kt})}{\gamma} = \frac{[\Delta(H_{it}s_{it}) - \Delta(H_{kt}s_{kt})]}{\gamma} + (\Delta \ln p_{it} - \Delta \ln p_{kt}),$$

$$\frac{(\Delta\delta_{it} - \Delta\delta_{kt})}{(1+\omega)} = (\Delta \ln p_{it} - \Delta \ln p_{kt}) - \frac{\omega(\Delta \ln s_{it} - \Delta \ln s_{kt})}{(1+\omega)}$$
$$- \frac{1}{(1+\omega)}\left\{\Delta \ln\left[1 + \frac{H_{it}s_{it}N_t}{\gamma(N_t - 1)}\right] - \Delta \ln\left[1 + \frac{H_{kt}s_{kt}N_t}{\gamma(N_t - 1)}\right]\right\}.$$

We multiply these two equations together, and average the resulting equation over time, to obtain the estimating equation:

$$\bar{Y}_i = \frac{\omega}{(1+\omega)}\bar{X}_{1i} + \frac{\omega}{\gamma(1+\omega)}\bar{X}_{2i} - \left(\frac{1}{\gamma}\right)\bar{X}_{3i} + \frac{1}{(1+\omega)}\bar{Z}_{1i}(\gamma) + \frac{1}{\gamma(1+\omega)}\bar{Z}_{2i}(\gamma) + \bar{u}_i, \qquad (24)$$

where the over-bar indicates that we are averaging that variable over time, and:

$$Y_{it} \equiv (\Delta \ln p_{it} - \Delta \ln p_{kt})^2,$$

$$X_{1it} \equiv (\Delta \ln s_{it} - \Delta \ln s_{kt})(\Delta \ln p_{it} - \Delta \ln p_{kt}),$$

$$X_{2it} \equiv (\Delta \ln s_{it} - \Delta \ln s_{kt})[\Delta(H_{it}s_{it}) - \Delta(H_{kt}s_{kt})],$$

$$X_{3it} \equiv (\Delta \ln p_{it} - \Delta \ln p_{kt})[\Delta(H_{it}s_{it}) - \Delta(H_{kt}s_{kt})],$$

$$Z_{1it}(\gamma) \equiv \left\{ \Delta \ln\left[1 + \frac{H_{it}s_{it}N_t}{\gamma(N_t - 1)}\right] - \Delta \ln\left[1 + \frac{H_{it}s_{kt}N_t}{\gamma(N_t - 1)}\right]\right\}(\Delta \ln p_{it} - \Delta \ln p_{kt}),$$

$$Z_{2it}(\gamma) \equiv \left\{ \Delta \ln\left[1 + \frac{H_{it}s_{it}N_t}{\gamma(N_t - 1)}\right] - \Delta \ln\left[1 + \frac{H_{it}s_{kt}N_t}{\gamma(N_t - 1)}\right]\right\}(\Delta H_{it}s_{it} - \Delta H_{kt}s_{kt}).$$

and, $$u_{it} \equiv \frac{(\Delta\varepsilon_{it} - \Delta\varepsilon_{kt})(\Delta\delta_{it} - \Delta\delta_{kt})}{\gamma(1+\omega)}.$$

We shall assume that the error terms in demand and the pricing equation are uncorrelated, which means that the error term in (24) becomes small, $\bar{u}_i \to 0$ in probability limit as $T \to \infty$. That error term is therefore uncorrelated with any of the right-hand side variables as $T \to \infty$, and we can exploit those moment conditions by simply running OLS on (24). Feenstra (1994) shows that procedure will give us consistent estimates of γ and ω in a slightly simpler system, provided that the right-hand side variables in (24) are not perfectly collinear as $T \to \infty$. As in the CES case of Feenstra (1994), that condition will be assured if there is some heteroskedasticity in the error terms across countries i, so that the right-hand side variables in (24) are not perfectly collinear. More efficient estimates can be obtained by running weighted least squares on (24).

Before proceeding with the estimation, we need to address a number of data problems. First, while in principle we could estimate γ at the 10-digit level, our estimates would not be precise because often there are few countries exporting in a given 10-digit HS product. In order to make sure that we have enough data to obtain precise estimates, we assume that the γ's at the 10-digit level within an HS-4-digit sector are the same, and therefore pool the HS-10-digit goods within each 4-digit sector.[25]

[25] This approach means that a sector typically had 94 varieties – defined as a distinct (country,10-digit good) pair – when we estimate γ for an HS-4 sector.

A second complication arises because we have U.S. shipments data at the NAICS 6-digit level but we need to compute shares at the HS-10 digit level. Thus, we must allocate NAICS 6-digit production data to each HS 10-digit sector. In order to do this, we assume that the share of U.S. production in each HS 10-digit is the same as that of the U.S. in the NAICS 6-digit sector that contains it, as discussed in Appendix A.

A third complication arises because we use unit values of import prices from each source country rather than the geometric mean price, which introduces measurement error, especially for import flows that are very small. Broda and Weinstein (2006) propose a weighting scheme based on the quantity of imports at the HS-10 level. Unfortunately, we cannot implement precisely that scheme because the U.S. quantity indexes were defined at the NAICS 6-digit level and not at the HS-10 digit level. We therefore implement the Broda and Weinstein weighting scheme using the value of shipments instead of the quantity of shipments, since shipment values are likely to be highly correlated with shipment quantities across countries.

Finally, as in Broda and Weinstein (2006), we also face the problem that only 86 percent of our estimates of γ had the right sign if we estimate them without constraints. If γ is less than zero, then this implies that markups are negative, and there is no equilibrium. Since we want to rule this out and because V is very sensitive to small values of γ, we impose $\gamma \geq 0.05$. To achieve this, we use a grid search procedure over γ and ω to minimize the sum of squared errors in (24). In this procedure we set an initial value for γ of 0.05 and increased it by 5 percent over the range [.05, 110].[26] Similarly, we set an initial ω of -5 and increased it by 0.1 over the range [-5, 15].

[26] In order to speed up the grid searches, in most specifications we increased the interval by 5 percent until 7.8 and then jumped to 109.9. We did this because we almost never found gammas between 7.8 and 109.9. Moreover, making this change did not qualitatively affect the results because all high gammas imply very small markups and variety effects.

7. Estimation Results

The coefficients γ^k for each sector k are obtained by estimating equation (24). Because we ultimately estimate over one thousand γ^k's, it is not possible to display all of them here. We display the sample statistics for γ^k and $1/\gamma^k$ in Table 3. The median γ is 0.19 and the average is 12. The large average γ^k is driven by the fact that their distribution is not symmetric and γ^k can take on very large values if goods are relatively homogeneous. It is difficult to have strong priors for what a reasonable value of γ^k should be. One possible benchmark is the implied markup. We can compute the markup for each industry by using equation (19). Based on this calculation the median estimated markup in our data is 0.30 (i.e. a 30% markup over marginal costs) in 2005. By comparison, Domowitz, Hubbard, and Petersen (1988) estimate markups across U.S. manufacturing and obtain an average markup of 0.37, which is a bit higher than ours but not dramatically different especially given the large differences in data and estimating procedures.

The markups in each sector depend on the value of the firm's market share as well. We can get some sense of the reasonableness of our estimates by looking at the most important sectors in U.S. absorption. In Table 4, we report the share of U.S. absorption from the ten largest sectors (with names not beginning with "other"), where we define the share to be the average share of absorption in 1992 and 2005. In the first column we report our estimate of the γ^k's. Based on this measure, we find the three sectors where the products are most heterogeneous and firms are likely to have the most market power are Aircraft and Spacecraft, Televisions, Video Cameras, and Receivers, and Private Motor Vehicles. In contrast, the most homogeneous sectors where firms are likely to have the least market power are Crude Petroleum, Natural Gas, and Cigarettes and Cigars. This pattern seems broadly sensible.

We next compute the variety and pro-competitive effects for each sector, and aggregate these across sectors as in (22). The key decision that must be made in measuring both these effects is the set of countries $i \in \bar{I}$, which we interpret as having small changes in the set of products and firms *within* each country (literally, in Proposition 1 we assume that the taste parameter for countries $i \in \bar{I}$ does not change at all). As mentioned earlier, in the CES case Broda and Weinstein (2006) used the set of countries supplying to the U.S. at both the beginning and end of the sample to compute the change in product variety between these dates. We can do better here because we also have information on the Herfindahl indexes of exporters and U.S. firms. We will suppose that if the Herfindahl index changes by more than some tolerance, that is evidence that the set of firms has changed sufficiently to exclude that country from the set of countries $\bar{I}$. For convenience, we refer to this set as "common" countries, that are supplying in the first and last periods and are judged to have little change in their exporting firms. In our baseline results reported in Table 5, we choose ±30 percent as the tolerance for the change in the Herfindahl index for each sector and country, and then will indicate how our results change for other tolerances.

Table 5 reports our welfare results. Because we had to drop the change between 1997 and 1998, we computed the aggregate welfare change between 1992 and 2005 as the sum of the gains from 1992-7, the gain from 1998-2005, and the geometric average annual gain over the two periods. Using the full sample our baseline Herfindahl cutoff of ±30 percent we find that the aggregate welfare gain over the time period was 0.86 percent.

The magnitude of this number is perhaps easiest to understand relative to Broda and Weinstein's (2006) estimates for the period 1990 to 2001. Those authors used a CES aggregator and obtained a gain to consumers of 0.8 percent over the 1990-2001 period. This result is almost exactly the same as what we obtain in the translog. However, they estimated the gain over a different and shorter sample period (eleven years vs. 13). If we multiply our estimates by 11/13, we find that implied aggregate gain due to varieties over the 11-year period in the translog case is 0.73. These results suggest that the translog functional form yields similar, but slightly smaller, variety gains as the CES.

One major advantage of the translog setup is that it allows us to examine the sources of these welfare gains. In particular, the translog specification indicates that markup declines, P_t, account for about half of the welfare gains – of the 0.86 percentage point welfare gain, 0.42 percent is due to markup declines. The fact that the gain due to new variety in the translog setup is approximately one-half of the gain obtained by using a CES formula is theoretically well-justified when there is only *one* new good,[27] and we are finding that the result applies here even with many new goods.

We can go further towards understanding the difference between the CES and translog by using equation (15) to decompose the total variety gains, V_t, into two components: the first term in the formula, V_{1t}, measures the impact of new varieties on welfare irrespective of crowding, and the second term, V_{2t}, provides the impact of crowding, i.e. the fact that consumers care less about varieties as more become available. The results in Table 5 indicate that crowding of the product space is an important offset of the variety gains: without it, the welfare gain due to import varieties would have been as large as what we obtain for the *total* welfare gains, including from reductions in markups. In other words, it is the crowding effect that makes the translog gain from variety less than the CES gains. But the reduced gains from variety are fully compensated by the extra gains due to reduced markups.

[27] Feenstra and Shiells (1994, Corollary 2) argue that with a single new good, and with the translog demand elasticity equal to the CES elasticity, the translog gain would be one-half as much as the CES gain.

We should check that the pro-competitive effect does not depend on our measurement of the total number of each goods in each sector, N_t in (21). So far we have used the inverse of the overall Herfindahl index in each sector to infer N_t. An alternative approach is to treat N_t as infinity so that $N_t/(N_t - 1) = 1$. We refer to this second case by W_{inf} and P_{inf} in the final columns of Table 5, and find that it reduces the pro-competitive effect P_t only slightly.

We can think of four other possible issues with the reliability of our estimates.[28] First, we should compute standard errors because of imprecision in the estimation of γ^k. Second, our results might be sensitive to our choice of Herfindahl cutoff. Third, our estimates might be heavily influenced by outliers of particular sectors. Finally, since the automobile sector is the largest sector and had some idiosyncratic factors, discussed below, we decided to also rerun our estimates without this sector. We deal with each of these concerns in turn.

In order to deal with the imprecision of our γ^k estimates, we bootstrapped each of the over one thousand γ^k's and ω^k's and then used these bootstrapped parameter values to compute the distribution of P_t^k, V_t^k, and total welfare $\Delta \ln W_t^{PV}$. This is computationally intensive, but ultimately we were able to compute 100 estimates of each γ^k and generate 5-95 percentile confidence bands.[29] The narrowness of these bands, reported in Table 5, indicates that our point estimates for the markup and variety effects are estimated with reasonable precision.

Second, we experimented with other cutoffs for the change in the Herfindahl indexes. The choice of what threshold to use in the definition of the "common" countries $i \in \bar{I}$ involves a tradeoff between two opposing forces. Classifying even the small movements in the Herfindahl

[28] A fifth issue, discussed in Appendix C, is whether the number of years in our sample (thirteen) is enough to avoid the small-sample bias in our estimator noted by Soderbery (2010) in the CES case. That issue is taken up in Appendix C, where we find that thirteen years is (just) enough to avoid significant bias.

[29] It took 10 days on an 8-processor SPARCstation.

index as evidence of firm entry and exit will reduce the number of countries included in the common set $\bar{I}$, thereby eliminating some sectors, which makes our welfare calculations very sensitive to Herfindahl and share movements in fewer remaining countries and sectors. On the other hand choosing a very wide band of allowable Herfindahl movements means that our estimates of welfare gains will be based on more countries and sectors, but we run the risk of erroneously *missing* some of the variety gains. As noted in equation (16) of Proposition 1, even if there is no change in the set of countries we can still use changing Herfindahls to infer gains or losses from variety, but that assumes that there is *no change* in the average taste parameters for countries. By leaving countries out of the common set $\bar{I}$, we do not need that assumption and are more accurately imputing welfare gains/losses due to the entry/exit of firms and products from those countries. We therefore considered a number of cutoff values for the Herfindahl indexes: movements of plus or minus 20, 30 (as already reported) and 40 percent.

We obtain very similar welfare gains if we use a ±40 percent cutoff, but the welfare gain is much smaller with the ±20 percent cutoff, because that estimate is excluding many more countries from the common set.[30] As a final robustness check for understanding the role played by the Herfindahl cutoff, we decided to examine the impact of completely shutting down this channel. In order to do this calculation, we set all the Herfindahl ratios equal to their 1992 values and only consider the welfare impacts coming from the entry and exit of countries selling to the U.S. The results from this calculation are presented in the "Constant HI" row of Table 5. Overall

---

[30] To get some intuition for whether these cutoff values are sensible, we can consider how much of the data we move from the common set of countries $\bar{I}$ to the excluded countries $i \notin \bar{I}$. We will split our estimation into two sub-periods (1992-1997 and 1998-2005) because of the change in industry definitions. If we follow the prior CES literature and assume that variety change is only measured when country import flows start or end, we find that on average 95 percent of the value of varieties – defined as a country/HS-10 digit good pair – available in the starting year are available in the last year in each period. If we reclassify countries in which the Herfindahl moved by more than ±40 percent as no longer common, we find that on average only 77 percent of varieties available in the first time period were available in the second period. Similarly, the share of commonly available varieties falls to 71 percent, 52 percent, and 30 percent as we move to 30 percent, 20 percent, and 10 percent cutoffs, respectively.

we obtain a welfare gain of 0.79 percent of GDP, which is quite similar to our baseline results, but focusing on this number misses some important differences. In particular, by assuming that there is no U.S. exit in response to new imported varieties, we obtain a very large pure variety effect as measured by V_{1t}, but the absence of exit also implies an offsetting crowding of the product space as measured by V_{2t}. Thus, on net, there is almost no variety gain, and virtually all of the welfare gain comes from decline in market power arising from the drop in the demand elasticities.

A third concern arises from the possible role played by outliers. In order to ensure that our results were robust to outliers, we dropped the sectors in the top 5th percentile of welfare gains and those in the lowest 5th percentile. In order to prevent the welfare gains from falling simply because we were summing across fewer sectors, we reweighted each sector's share in (22) so that the shares continued to sum to one. Overall, dropping the top and bottom 5th percentiles slightly lowered our aggregate welfare gain: taking the change in welfare from 0.86 percent of GDP to 0.64 percent. This indicates that outliers are not driving our results.

A fourth concern relates to the importance of the automobile sector. Automobiles is the largest sector in manufacturing, and although the point estimate for γ^k in this sector was not an outlier, the sector does exert a particularly large impact on the overall welfare gains. Between 1992 and 2005, there was enormous entry into this sector as Japanese car makers set up new plants (see Blonigen and Soderbery, 2010). This entry had two important impacts. First, the U.S. Herfindahl index declined sharply from 0.35 to 0.21, reflecting the large increase in the number of makers operating in the U.S. Second, the transplant of Japanese car makers to the U.S. was associated with a very large increase in U.S. automobile production: real output of autos made in the U.S. grew by 41 percent between 1992 and 1998, which contributed to a substantial increase in the share of U.S. consumption made domestically. That increase in the share and falling Herfindahl contribute to a large welfare gain from variety, V_t^k , from (15).

There are reasons to believe, however, that our welfare formula cannot accurately deal with the transplant of Japanese varieties to the United States: we have ignored multi-product firms, for example, and in the same way have assumed that the γ^k estimate for autos applies equally well to products across firms as to products within firms. That assumption clearly contradicts the theoretical literature on multi-product firms, which makes a strong distinction between consumer substitution of products within and between firms (see Allanson and Montagna, 2005, and Bernard, et al, 2011). For this reason we also computed the welfare gains after dropping the passenger vehicle sector. This gives the result shown in the "Excluding Auto Sector" row of Table 5, where we assume that the gains from variety in the automobile sector were same as in all other sectors. The welfare estimate of 0.49 is somewhat smaller than our baseline estimate, but not significantly different.

Finally, in the last two rows of the table, we can see the welfare gains in each of our sub-periods. Most of total welfare gain W_t appears to have accrued in the first sub-period, between 1992 and 1997. The reason for this finding is that in the second sub-period, from 1998 to 2005, there is substantial crowding in product space: the component V_{2t} subtracts a full 0.71 percent from any welfare gains, thereby canceling out the contributions of V_{1t} and the pro-competitive effect P_t in that period. Thus, while the markup gains were quite similar in both periods, the gains from variety collapse in the later period because of crowding in product space. This finding suggests that the gains from varieties may be declining as globalization progresses.

8. Conclusions

Using general additively separable preferences, Krugman (1979) demonstrated the reduction in markups that accompanies trade liberalization under monopolistic competition. That reduction in markups is not just a consumer gain, but is also a social gain: the reduction in markups in a zero-profit equilibrium indicates that the wedge between a firm's marginal and average cost is reduced, so that output is expanding and there are greater economies of scale. So the competition between firms from different countries is an important channel by which international trade leads to social gains.

Despite this insight, such a channel has received only limited attention in the empirical trade literature. We have argued that the reason for this gap in the literature is the common assumption of CES preferences, which leads to constant markups. So instead we must look to alternative preferences, and here we have adopted translog preferences. We have derived quite general formulas for the variety gains from new products with these preferences, and also the pro-competitive effect of new entry on reducing markups.

The estimated declines in market share of incumbent firms in the wake of the tremendous amount of entry of foreign countries into U.S. markets, as well as more exporters within those countries, drives our measure of the welfare gains. This entry has been offset to some degree by crowding of the product space as measured by declining Herfindahl indexes. Nevertheless, we find that the exit from the U.S. market has been less than the new entry, in the sense that the demand for the typical incumbent firm's output fell, so that the *per-firm* share of surviving U.S. firms fell in many sectors. That feature of the data drives our estimates of the fall in markups, which is the pro-competitive effect of globalization. In our benchmark results, we find that the variety gain from globalization for the U.S. in the translog case is one-half of that found by Broda and Weinstein (2006) in the CES case, but that the total welfare gain is the same size.

As noted in the Introduction, we have not attempted to measure an important source of welfare gain that can operate in our model: the efficiency gains that come from the self-selection of more efficient firms into exporting. That source of gains is the focus of ACDR (2012), but by construction in their model, the variety and pro-competitive effects do not operate when the distribution of firm productivity is Pareto with unbounded support: in that case, variety in each country and the distribution of markups are not affected by reductions in trade costs. If instead the distribution is Pareto with a bounded support, so that productivity does not exceed some maximum value, then the product variety and pro-competitive effects operate once again, along with the self-section of more efficient firms into exporting (Feenstra, 2014). We have not attempted to measure this latter effect, and that is an important area for further research.

Appendix A: Data

The dataset used for this project contains quantity, value, and price information aggregated at the HS-10 digit level, as well as HS-4 digit level Herfindahl Indexes, for the U.S. and all countries exporting to the U.S. for every year from 1992 to 2005.

One challenge in piecing together this dataset was calculating the amount of U.S. absorption produced in the U.S. We begin with the identity that the U.S. supply of U.S. absorption is equal to the difference between U.S. production and exports. We obtained data on industry-level production from the Bureau of Economic Analysis at www.bea.gov and export data from http://www.internationaldata.org/. Unfortunately, the BEA production data are classified according to the SIC system for years 1992 to 1997 and according to the NAICS system for years 1998 to 2005, while the trade data is at the HS-10 digit level. Addressing this complication required a two-step process: the first step was to adjust the BEA production data so that the data are on the NAICS level for all years within the sample. The second step was to use our import/export data (containing both NAICS and HS-10 digit codes) and our newly created NAICS level production data to infer domestically produced absorption at the HS-10 digit level, as described below.

It is not easy to concord SIC and NAICS categories because there is not always a one-to-one mapping between the two. To deal with this issue, we first used a NAICS-SIC concordance from the BEA to convert the SIC data to the NAICS level. The absence of a one-to-one mapping meant that sometimes we would observe large jumps in a NAICS category derived from SIC data from 1997 relative to 1998. In order to deal with this problem, we used a "bridging dataset", from the U.S. Department of Commerce, containing SIC level values for both 1997 and 1998. This enabled us to construct a ratio between the actual NAICS output levels and the

NAICS levels that we constructed from the SIC data for 1998. We then multiplied all of NAICS data that was constructed from the SIC data by this ratio. If a SIC sector did not match any NAICS sector we dropped the observations prior to 1998 in the estimation. We also dropped all changes between 1997 and 1998 in the regressions where we estimated γ, so that concordance problems would not affect our estimates.

After our BEA data was brought to the NAICS level, we use it, along with our import and export data, to calculate HS-10 digit level U.S. domestic supply. We begin with the identity that U.S. supply for the domestic market at the NAICS level – denoted by k – equals U.S. production at the NAICS level less U.S. exports: $\text{Supply}_t^k = \text{Production}_t^k - \text{Exports}_t^k$. Using the NAICS import data, we can compute the share of U.S. supply in apparent consumption according to the following formula: $\text{Share}_t^k = \text{Supply}_t^k / (\text{Supply}_t^k + \text{Imports}_t^k)$. By assuming that the U.S. share of a NAICS code is equal to that of the U.S. share in a corresponding HS-10 code, we calculate supply at the HS-10 digit level using the following formula:

$$\text{Share}_t^k = \text{Share}_t^{HS10} \quad \text{and} \quad \text{Share}_t^{HS10} = \frac{\text{Supply}_t^{HS10}}{\text{Supply}_t^{HS10} + \text{Imports}_t^{HS10}}$$

$$\Rightarrow \text{Supply}_t^{HS10} = \frac{\text{Share}_t^k}{\left(1 - \text{Share}_t^k\right)} \text{Imports}_t^{HS10}.$$

We next needed to merge in data for Herfindahl indexes for domestic firms and exporters to the U.S. For land shipments from Canada, we purchased Herfindahl indexes at the 4-digit Harmonized system (HS) level, for 1996 and 2005, from Statistics Canada. These Canadian Herfindahl indexes were constructed from firm-level export data to the U.S.

For land shipments from Mexico, the Herfindahl indexes were constructed using data sourced from the *Encuesta Industrial Anual* (Annual Industrial Survey) of the *Instituto Nacional* *de Estadistica y Geografia*. This data contains firm-level exports for 205 CMAP94 categories for 1993 and 2003. We also obtained the export Herfindahl for 232 categories at the HS-4 level. These categories cover the most important Mexican export sectors.

For all other major exporters to the U.S., we computed these Herfindahls for *sea* shipments from PIERS (www.piers.com), for 1992 and 2005. PIERS collects data from the bill of lading for every container that enters a U.S. port. The median country exports about 80 percent of its goods by sea. Thus for the typical country in our sample, the sea data covers a large fraction of their exports. Although purchasing the disaggregated data is prohibitively expensive, we were able to obtain information on shipments to the U.S. for the 50,000 largest exporters to the U.S., for 1992 and 2005. For each exporter and year, we obtained the estimated value, quantity and country of origin of the top five HS-4 digit sectors in which the firm was active. We also obtained this data for the top ten HS-4 digit sectors for the largest 250 firms in each year.

Let $i \in I_t$ denote the countries exporting to the U.S. each period and $j \in J_{it}$ denote their firms.[31] The total import share from country i in sector k is $s_{it}^k \equiv \sum_{j \in J_{it}} s_{ijt}^k$, and we let $s_{jt}^{ik} \equiv s_{ijt}^k / s_{it}^k$ denote firms' shares *within* the exports of country i. Then with the Herfindahl index $H_{it}^k \equiv \sum_{j \in J_{it}} (s_{jt}^{ik})^2$, the sum of squared firm shares from country i is:

$$\sum_{j \in J_{it}} (s_{ijt}^k)^2 = \sum_{j \in J_{it}} (s_{jt}^{ik})^2 (s_{it}^k)^2 = H_{it}^k (s_{it}^k)^2 . \qquad \text{(A1)}$$

Thus, the sum of squared shares can be measured by using the Herfindahl indexes and overall shares (squared) from those countries. If we further summed over all countries i, we would obtain what we will call the *overall Herfindahl index* (in each sector):

$$\sum_{i \in I_t} \sum_{j \in J_{it}} (s_{ijt}^k)^2 = \sum_{i \in I_t} \sum_{j \in J_{it}} (s_{jt}^{ik})^2 (s_{it}^k)^2 = \sum_{i \in I_t} H_{it}^k (s_{it}^k)^2 . \qquad \text{(A2)}$$

[31] These sets depend on the sector k, but we omit that notation for simplicity.

The PIERS data has a number of limitations relative to other firm level data sets. The first is relatively minor: we do not have the universe of exporters but only the largest ones. This turns out not to be a serious problem because the aggregate value of these exporters is typically within 5 percent of total sea shipments. Thus, smaller exporters are unlikely to have a qualitatively important impact on our results. A second potential problem is that PIERS sometimes lists trading companies as exporters. Sometimes these trading companies may actually be exporters, however this does not appear to be a big problem in our data since companies that have the words "trading," "exports," "Imports" or variants of these words in their name (such as "exporting") only accounted for 5 percent of exports in 1992 and 7.5 percent of exports in 2005. Thus, we think that the vast majority of exporters in our sample are actual exporters.[32]

A larger problem is that the PIERS data only comprises sea shipments and thus we have no information in these data on land and air shipments. This means that we have to adjust our Herfindahl indexes to take into account land and air shipments. The Herfindahl of country i's exports in sector k can be written as

$$H_{it}^{k} = H_{it}^{kSea}\left(\frac{V_{it}^{kSea}}{V_{it}^{kTotal}}\right)^2 + H_{it}^{kNon-Sea}\left(1-\frac{V_{it}^{kSea}}{V_{it}^{kTotal}}\right)^2, \qquad \text{(A3)}$$

where V_{it}^{kSea} $\left(V_{it}^{kTotal}\right)$ denotes the value of sea (total) shipments and $H_{it}^{kNon-Sea}$ is the Herfindahl for non-sea exporters, which is defined analogously as the sea Herfindahl. We do not have a measure of $H_{it}^{kNon-Sea}$, but theory does place bounds on the size of the Herfindahl since the true index must be contained in the following set, obtained with $H_{it}^{kNon-Sea}$ = 1 or 0:

$$\left[H_{it}^{kSea}\left(\frac{V_{it}^{kSea}}{V_{it}^{Total}}\right)^2, H_{it}^{kSea}\left(\frac{V_{it}^{kSea}}{V_{it}^{Total}}\right)^2 + \left(1-\frac{V_{it}^{kSea}}{V_{it}^{kTotal}}\right)^2\right].$$

[32] It's not clear whether we should delete these companies or not because many of them , such as "Kerala Cashew Exports" may serve as the exporters of firms that do not export independently.

For most sectors the share of sea shipments in total shipments is quite high, so these bounds are quite tight. In the analysis we assume that $H_{it}^{kSea} = H_{it}^{kNon-Sea}$, but our results do not change qualitatively if we assume that $H_{it}^{kNon-Sea} = 0$.[33]

For the U.S. Herfindahls, we rely on data from the Census of Manufactures at the NAICS 6-digit level. Unfortunately, this is more aggregate than the 4-digit HS level at which we have the foreign export Herfindahl indexes. Accordingly, we need to convert the U.S. Herfindahl indexes from the NAICS 6-digit level to the HS 4-digit level. Abusing our above *country* notation, let $i \in I_k$ denote a 4-digit sector i *within* the NAICS code k. Then the Herfindahl for 4-digit sector i is $H_{it}^{k} \equiv \sum_{j \in J_i} (s_{jt}^{i})^2$, where s_{jt}^{i} is the share of firm $j \in J_i$ in sector i. We see that the overall Herfindahl in NAICS code k is:

$$\sum_{i \in I_k} H_{it}^{k} \left(s_{it}^{k}\right)^2 = \sum_{i \in I_k} \sum_{j \in J_i} (s_{jt}^{i})^2 (s_{it}^{k})^2 = \sum_{j \in J_k} (s_{jt}^{k})^2 \equiv H_t^k, \tag{A4}$$

where s_{it}^{k} is the share of 4-digit HS sector i within NAICS sector k, and $s_{jt}^{k} = s_{jt}^{i} s_{it}^{k}$ is the share of product j within the NAICS sector, $j \in J_k$. In words, the inner-product of the Herfindahl firm indexes and the squared sector shares, on the left of (A4) is exactly the right way to aggregate these indexes to obtain an *overall Herfindahl for the good k in question,* on the right of (A4). One of the problems that we faced is that we know H_t^k but not H_{it}^k. A solution can be obtained by assuming that H_{it}^k is equal across all 4-digit sectors $i \in k$, in which case we solve for H_{it}^k as:

$$H_{it}^{k} = H_t^k / \sum_{i \in I_k} (s_{it}^{k})^2 . \tag{A5}$$

[33] One can see this from a simple example. Our median sea Herfindahl is 0.6 and our median share of sea shipments is 0.8. This means that the true Herfindahl ranges from .38 to .42 and our estimate would be 0.41. Nevertheless, we are implicitly assuming that goods shipped by air and goods shipped by sea are not the same. We justify this assumption because it costs substantially more to ship goods by air, and thus the mode of shipment is likely to differentiate the goods in some important ways.

In other words, the 4-digit HS Herfindahl is estimated by dividing the 6-digit NAICS Herfindahl by the corresponding Herfindahl index of 4-digit HS shares within the 6-digit sector. This simple solution assumes that the 4-digit HS Herfindahl indexes are constant within a sector, but is the best that we can do in the absence of additional data.

Appendix B:

Proof of (5):

Substituting (4') into (2) and using (3) we obtain:[34]

$$
\begin{aligned}
\ln e_t &= \alpha_0 + \sum_{i=1}^{\tilde{N}} \alpha_i \left(\frac{\alpha_i - s_{it}}{\gamma} + \overline{\ln p(t)} \right) + \frac{\gamma}{2} \sum_{i=1}^{\tilde{N}} \left(\frac{\alpha_i - s_{it}}{\gamma} + \overline{\ln p(t)} \right)^2 \\
&\quad + \frac{\gamma}{2\tilde{N}} \sum_{i=1}^{\tilde{N}} \sum_{j=1}^{\tilde{N}} \left(\frac{\alpha_i - s_{it}}{\gamma} + \overline{\ln p(t)} \right) \left(\frac{\alpha_j - s_{jt}}{\gamma} + \overline{\ln p(t)} \right) \\
&= \alpha_0 + \frac{1}{\gamma} \sum_{i=1}^{\tilde{N}} \alpha_i (\alpha_i - s_{it}) + \overline{\ln p(t)} + \frac{1}{2\gamma} \sum_{i=1}^{\tilde{N}} (\alpha_i - s_{it})^2 - \frac{\gamma \tilde{N}}{2} \left(\overline{\ln p(t)} \right)^2 \\
&\quad + \frac{\gamma}{2} \overline{\ln p(t)} \sum_{i=1}^{\tilde{N}} \left(\frac{\alpha_i - s_{it}}{\gamma} + \overline{\ln p(t)} \right) \\
&= \alpha_0 + \frac{1}{2\gamma} \sum_{i=1}^{\tilde{N}} (\alpha_i)^2 + \overline{\ln p(t)} - \frac{1}{2\gamma} \sum_{i=1}^{\tilde{N}} (s_{it})^2 - \frac{\gamma \tilde{N}}{2} \left(\overline{\ln p(t)} \right)^2 + \frac{\gamma \tilde{N}}{2} \left(\overline{\ln p(t)} \right)^2 \\
&= \alpha_0 + \frac{1}{2\gamma} \sum_{i=1}^{\tilde{N}} (\alpha_i)^2 + \overline{\ln p(t)} - \frac{1}{2\gamma} \sum_{i=1}^{\tilde{N}} (s_{it})^2 .
\end{aligned}
$$

Proof of (16):

The only expression in Propositions 1 or 2 that is not already derived in the text is (16). To obtain this result, substitute the shares from (9) into the definition of V_t in (15). Note that any difference $\Delta(x_t y_t)$ can be expressed as $\Delta(x_t y_t) = \frac{1}{2}(x_{t-1} + x_t)\Delta y_t + \frac{1}{2}(y_{t-1} + y_t)\Delta x_t$. Using (9) and this result, we can simplify V_t when $i \in \overline{I} = I_{t-1} = I_t$ as:

[34] We thank Sam Kortum for this derivation.

$$V_t = -\sum_{i\in\bar{I}} \frac{1}{2\gamma}\left[s_{it-1} + \frac{1}{\bar{N}}\left(1 - \sum_{j\in\bar{I}} s_{jt-1}\right) + s_{it} + \frac{1}{\bar{N}}\left(1 - \sum_{j\in\bar{I}} s_{jt}\right)\right]\Delta(H_{it}s_{it}) + \frac{1}{2\gamma}\Delta\left(\sum_{i\in I_t} H_{it}s_{it}^2\right)$$

$$= -\frac{1}{2\gamma}\sum_{i\in\bar{I}}(s_{it-1}+s_{it})[\tfrac{1}{2}(H_{it-1}+H_{it})\Delta s_{it} + \tfrac{1}{2}(s_{it-1}+s_{it})\Delta H_{it}] - \frac{1}{2\gamma\bar{N}}\sum_{j\notin\bar{I}}(s_{jt-1}+s_{jt})\sum_{i\in\bar{I}}\Delta(H_{it}s_{it})$$

$$+\frac{1}{2\gamma}\Delta\left(\sum_{i\notin\bar{I}} H_{it}s_{it}^2\right) + \frac{1}{2\gamma}\sum_{i\in\bar{I}}[\tfrac{1}{2}(H_{it-1}+H_{it})\Delta s_{it}^2 + \tfrac{1}{2}(s_{it-1}^2+s_{it}^2)\Delta H_{it}]$$

$$= -\frac{1}{2\gamma\bar{N}}\sum_{j\notin\bar{I}}(s_{jt-1}+s_{jt})\sum_{i\in\bar{I}}\Delta(H_{it}s_{it}) + \frac{1}{2\gamma}\Delta\left(\sum_{i\notin\bar{I}} H_{it}s_{it}^2\right) - \frac{1}{2\gamma}\sum_{i\in\bar{I}} s_{it-1}s_{it}\Delta H_{it}.$$

Thus, for $\bar{I} = I_{t-1} = I_t$ we are left with only the last term, as shown in (16).

Appendix C: Monte Carlo Simulation

One concern with our methodology stems from the fact that we rely on asymptotic properties that arise as the number of observations approaches infinity, but we are working with a sample of only thirteen years. In this Appendix, we use a Monte Carlo simulation to determine whether this is a problem. Our approach is quite similar to that used in Soderbery (2010).

We begin by reproducing the estimating equation (24), but without averaging over time:

$$Y_{it} = \frac{\omega}{(1+\omega)}X_{1it} + \frac{\omega}{\gamma(1+\omega)}X_{2it} - \left(\frac{1}{\gamma}\right)X_{3it} + \frac{1}{(1+\omega)}Z_{1it}(\gamma) + \frac{1}{\gamma(1+\omega)}Z_{2it}(\gamma) + u_{it}\,. \quad \text{(C1)}$$

If we average equation (C1) over time we obtain the estimating equation used in the paper:

$$\bar{Y}_i = \frac{\omega}{(1+\omega)}\bar{X}_{1i} + \frac{\omega}{\gamma(1+\omega)}\bar{X}_{2i} - \left(\frac{1}{\gamma}\right)\bar{X}_{3i} + \frac{1}{(1+\omega)}\bar{Z}_{1i}(\gamma) + \frac{1}{\gamma(1+\omega)}\bar{Z}_{2i}(\gamma) + \bar{u}_i\,. \quad \text{(C2)}$$

Our next task was to generate a synthetic dataset to use in our simulations that matched the moments in actual data. We decided that in this synthetic data, we would work with a sector whose estimated γ equals the median estimated γ. Since HS sector 2902 (Styrene) had the median γ, we decided to base our exercise using the data and parameter from this sector and then see how sensitive they were to estimation problems. We therefore conducted our simulation using a ω equal to 1.52, which is the same as the estimated ω in this sector. We also took the values of H_{it}, s_{it}, H_{kt}, s_{kt} and N_t terms from the actual data for this sector. We then set the error term equal to zero in equation (C1) and solved for $\Delta lnp_{it} - \Delta lnp_{kt}$ so that equation (C1) fit exactly.

Once we had generated data that fit the model exactly, we needed to also generate a series of errors to test how efficient our estimation procedure was. In order to do this we first had to match the variance of the error terms that we actually observed. To obtain the variance estimates, we generated the residuals of equation (C1) after imposing our chosen parameter values and using the actual, rather than synthetic data. The sample variance of the residuals for each country-HS-10 sector provided the variance estimate for each country-HS-10 sector.

We then added an error term (u_{it}) to each simulated country-HS-10 sector observation whose variance was drawn from a normal distribution with zero mean and a variance equal to the one we obtained for that country-HS-10 sector. In principle, this would let us also solve for the underlying demand and supply shocks, but it turns out that we only need to know their product in order to do the simulation.

With the simulated data constructed, the observations were then averaged across time within country-HS-10 sectors in order to transform observations for equation (C2). We then estimated the parameters of the model using the same technique as in the paper. This provided the first estimate (and iteration of the Monte Carlo). For each subsequent iteration, we discarded the previous error terms (u_{it}) and drew new error terms from the normal distribution to add to each observation. We then estimated the parameters of the model again. This process was repeated for each sample of a given time period for a total of 200 iterations.

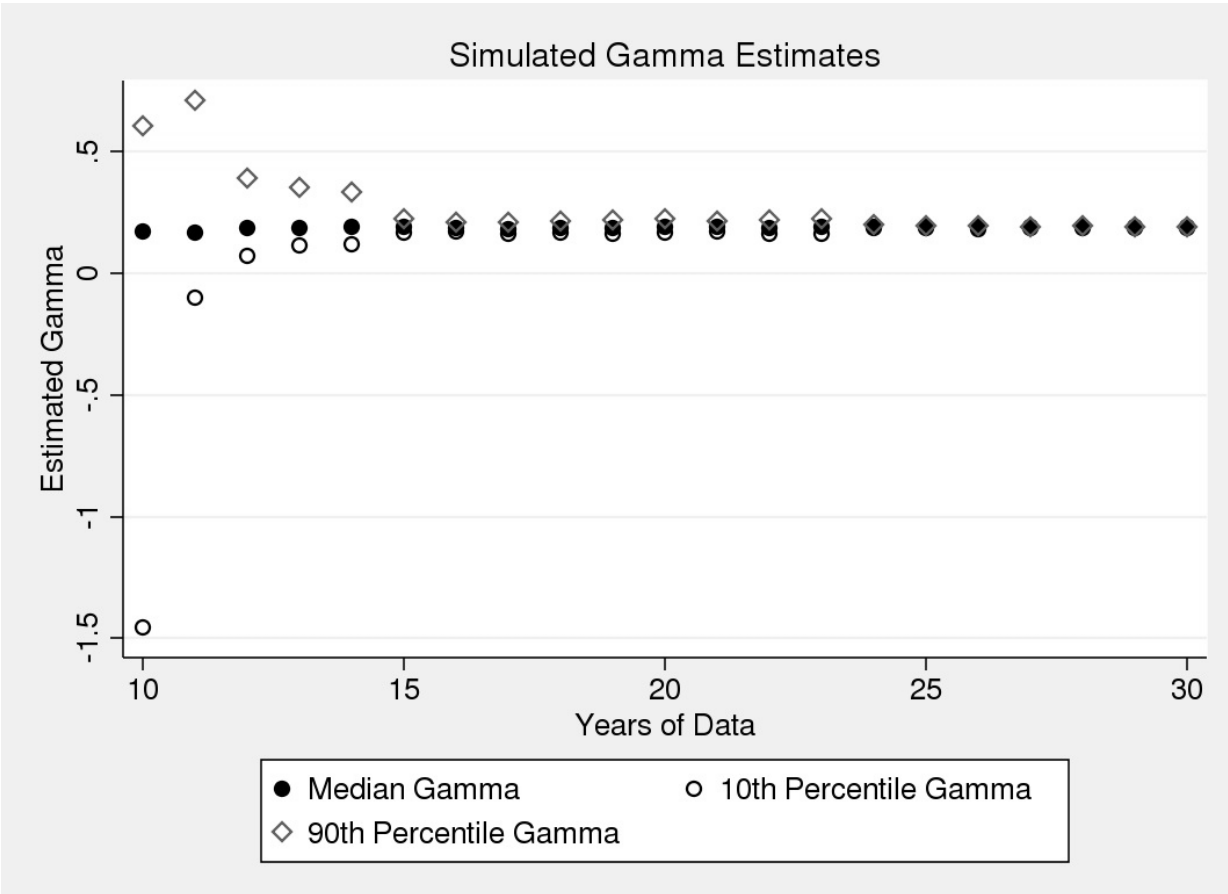
The simulation results for samples that were less than thirteen years were based on using the actual H_{it}, s_{it}, H_{kt}, s_{kt} and N_t values for each sector from the corresponding years in the data. . For each additional year above thirteen, we independently re-sampled one set of values of H_{it}, s_{it}, H_{kt}, s_{kt} and N_t from the actual data. This re-sampled data was then appended to the actual data from 1992-2005, an error term was added, and the estimation proceeded as outlined above.

Simulation results

The simulation results for the structural parameter γ are presented in Figure C1. The median estimated gamma is shown, along with the 10^{th} and 90^{th} percentile gammas. The results show a rapid convergence to the correct value of gamma beginning after approximately 12 years and almost complete convergence within 24 years. These results are similar to the results in Soderbery (2010).

The simulation can also be used to predict the percent of gammas estimated to be negative by the unconstrained WLS. The estimation used data for 14 years, from 1992-2005. However, because the change from 1997 to 1998 was dropped, the actual estimation used the equivalent of 13 years of data. In the simulation, 2.2% of the γ estimates were negative. By contrast 14% of the estimates of γ in the paper were actually negative. The higher percentage of estimated γ's relative to simulated γ's may reflect the fact that we simulated estimation assuming the true γ equaled the median gamma. If we had simulated the data using a smaller γ (and indeed half of all γ's are less than the median), we would have obtained more negative gamma estimates. Thus, the fact that our simulation produces few negative γ estimates when we constrain γ to be the median, seems basically in line with what our estimation procedure produces.

Figure C1: Estimated Gamma from Monte Carlo Simulation



References

Allanson, Paul and Catia Montagna, 2005, “Multiproduct Firms and Market Structure: An Explorative Application to The Product Life Cycle” *International Journal of Industrial Organization* 23, 587-597.

Arkolakis, Konstantinos, Arnaud Costinot and Andres Rodriguez-Clare, 2012, “New Trade Models, Same Old Gains? *American Economic Review*, 102(1), 94-130.

Arkolakis, Konstantinos, Arnaud Costinot, David Donaldson and Andres Rodriguez-Clare, 2012, “The Elusive Pro-Competitive Effect of Trade,” March 16, available at (2/7/13): http://www.econ.yale.edu/~ka265/research/ACDR/ACDRdraft.pdf ; and later draft of June 18, 2012.

Badinger, Harald, 2007a, “Has the EU’s Single Market Programme Fostered Competition? Testing for a Decrease in Markup Ratios in EU Industries,” *Oxford Bulletin of Economics and Statistics*, 69 (4): 497-519

Badinger, Harald, 2007b, “Market size, Trade, Competition and Productivity: Evidence from OECD Manufacturing Industries. *Applied Economics*, 39 (11): 2143 - 2157.

Badinger, Harald, 2008. “Trade Policy and Productivity,” *European Economic Review*, 52, 867-891.

Behrens Kristian and Yasusada Murata, 2007, “General Equilibrium Models of Monopolistic Competition: A New Approach,” *Journal of Economic Theory*, 13, 776-787.

Behrens, Kristian, Giordano Mion, Yasusada Murata, Jens Südekum, 2008, “Trade, Wages, and Productivity,” Institute for the Study of Labor, September.

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

Bergstrand, Jeffrey, H., 1989, ‘The Generalized Gravity Equation, Monopolistic Competition, and the Factor-Proportions Theory in International Trade,’ *The Review of Economics and Statistics*, 71(1), February, 143-153.

Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen and Samuel Kortum, 2003, “Plants and Productivity in International Trade,” *American Economic Review*, 93(4), 1268-1290.

Bernard, Andrew, Stephen Redding and Peter K. Schott, 2011, “Multi-Product Firms and Trade Liberalization.” *The Quarterly Journal of Economics*, 126(3), 1271-1318.

Berry, Steven T., 1994, “Estimating Discrete-Choice Models of Product Differentiation,” *Rand Journal of Economics*, 25, 242-246.

Bilbiie, Florin, Fabio Ghironi, and Marc J. Melitz, 2012, “Endogenous Entry, Product Variety, and Business Cycles,” *Journal of Political Economy*, 120(2), 304-345.

Blonigen, Bruce A., Benjamin H. Liebman, and Wesley W. Wilson, 2007, “Trade Policy and Market Power: The Case of the U.S. Steel Industry, NBER working paper no. 13761.

Blonigen, Bruce A. and Soderbery, Anson, 2010, “Measuring the Benefits of Product Variety with an Accurate Variety Set,” *Journal of International Economics*, 2(82), 168-180.

Bottasso, Anna and Alessandro Sembenelli, 2001, "Market Power, Productivity and the EU Single Market Program: Evidence from a panel of Italian firms," *European Economic Review*, 45, 167-186.

Bresnahan, Timothy F., 1989, "Empirical Studies of Industries with Market Power," in Richard Schmalensee and Robert Willig, eds., *Handbook of Industrial Organization, Volume 2*. Amsterdam: North-Holland, 1011-57.

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

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

Diewert, W. E., 1976, "Exact and Superlative Index Numbers," *Journal of Econometrics*, 4, 115-145.

De Blas, Beatriz and Katheryn Niles Russ, 2012, "Understanding Markups in the Open Economy," Universidad Autonoma de Madrid and University of California, Davis.

De Loecker, Jan, Pinelopi K. Goldberg, Amit K. Khandelwal, Nina Pavcnik, 2012, "Prices, Markups and Trade Reform," NBER working paper no. 17925.

Domowitz, Ian, R. Glenn Hubbard, and Bruce C. Petersen, 1988. "Market Structure and Cyclical Fluctuations in U.S. Manufacturing," *The Review of Economics and Statistics*, 70(1), February, 55-66.

Edmond, Chris, Virgiliu Midrigan and Daniel Yi Xu, 2012, "Competition, Markups, and the Gains from International Trade," NBER working paper no. 18041.

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

Feenstra, Robert C., 2014,"Restoring the Product Variety and Pro-competitive Gains from Trade with Heterogeneous Firms," NBER Working Paper 19833, 2014.

Feenstra, Robert C. and Clinton Shiells , 1997, "Bias in U.S. Import Prices and Demand," in Timothy Bresnahan and Robert Gordon, eds. *New Products: History, Methodology and Applications*, NBER and Univ. of Chicago Press, 1997, 249-276.

Feenstra, Robert C., Benjamin Mandel, Marshall B. Reinsdorf and Matthew J. Slaughter, 2013, "Effects of Terms of Trade Gains and Tariff Changes on the Measurement of U.S. Productivity Growth," *American Economic Journal: Economic Policy*, 2013, 5(1), 59–93.

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

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 U.S.-Canada Trade," *American Economic Review*, 91(4), September, 858-876.

Kim, Euysung, 2000, "Trade liberalization and Productivity Growth in Korean Manufacturing Industries: Price Protection, Market Power, and Scale Efficiency," *Journal of Development Economics*, 62, 55-83.

Krishna, Pravin and Devashish Mitra, 1998, "Trade liberalization, Market discipline and Productivity Growth: New Evidence from India," *Journal of International Economics*, 56, 447-462.

Konings, Jozef, Patrick Van Cayseele, and Frederic Warzynski, 2005, "The Effects of Privatization and Competitive Pressure on Firms' Price-Cost Margins: Micro Evidence from Emerging Economies," *Review of Economics and Statistics*, 87, 124-134.

Konings, Jozef and Hylke Vandenbussche, 2005, "Antidumping protection and markups of domestic firms," *Journal of International Economics*, 65, 151-165.

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

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

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

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

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

Rodriguez-Lopez, Antonio, 2011, "Prices and Exchange Rates: A Theory of Disconnect," *Review of Economic Studies*, 78(3), July, 1135-1177.

Simonovska, Ina, 2010, "Income Differences and Prices of Tradables," NBER working paper no. 16233.

Soderbery, Anson, 2010, "Investigating the Asymptotic Properties of Elasticity of Substitution Estimates," *Economics Letters*, 2(190), 57-62.

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

Tybout, James, R., Jaime de Melo, and Vittorio Corbo, 1991, "The Effects of Trade Reform on Scale and Technical Efficiency: New Evidence from Chile," *Journal of International Economics*, 31(3/4), 231-250.

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

Viner, Jacob, 1937, *Studies in the Theory of International Trade*. New York: Harper and Brothers Publishers.

Yilmazkuday, Hakan, 2012, "Markups versus Trade Costs," Florida International University.

Table 1

Share of U.S. Total Absorption in 1992 or 1998

	1992	1997			1998	2005	
Country	**Share**	**Share**	**Diff**	**Country**	**Share**	**Share**	**Diff**
United States	0.801	0.745	-0.056	United States	0.781	0.692	-0.089
Total Imports	0.199	0.255	0.056	Total Imports	0.219	0.308	0.089
Including:				*Including:*			
Canada	0.038	0.052	0.014	Canada	0.043	0.056	0.013
Japan	0.036	0.035	0.000	Japan	0.030	0.025	-0.004
Mexico	0.012	0.024	0.012	Mexico	0.022	0.031	0.009
Germany	0.010	0.012	0.002	China	0.017	0.041	0.024
China	0.010	0.017	0.007	Germany	0.012	0.015	0.004

Table 2
Herfindahl index in 1992 or 1998

	Herfindahl Index			Weighted Average $H_{it}S_{it}$		
Country	**1992**	**1997**	**Diff**	**1992**	**1997**	**Diff**
United States	0.147	0.155	0.008	0.111	0.111	-0.001
Canada	0.245	0.252	0.007	0.011	0.013	0.003
Japan	0.310	0.313	0.003	0.009	0.010	0.001
Mexico	0.393	0.407	0.014	0.004	0.009	0.005
Germany	0.358	0.357	-0.002	0.003	0.004	0.001
China	0.366	0.293	-0.073	0.001	0.002	0.001
Weighted Ave.	0.160	0.170	0.010	0.078	0.069	-0.009
Country	**1998**	**2005**	**Diff**	**1998**	**2005**	**Diff**
United States	0.183	0.189	0.006	0.139	0.129	-0.010
Canada	0.249	0.242	-0.008	0.011	0.015	0.003
Japan	0.318	0.331	0.013	0.008	0.008	-0.001
Mexico	0.419	0.403	-0.016	0.008	0.010	0.002
China	0.280	0.188	-0.092	0.002	0.003	0.001
Germany	0.332	0.335	0.002	0.003	0.005	0.001
Weighted Ave.	0.190	0.191	0.001	0.090	0.071	-0.019

Notes:
The Herfindahl Index is the weighted average of the country's Herfindahl Index, where the weights correspond to the share of each HS-4 sector in U.S. apparent consumption. “Share” s_{it} is defined to be the country's share of U.S. apparent consumption. The “Weighted Average $H_{it}s_{it}$” is the weighted average of the Herfindahl Index in sector i in year t multiplied by that country's share of U.S. apparent consumption; the weights are the same as before. The last row reports a weighted average across all countries using each country's share of U.S. apparent consumption as weights. Thus, the number shown in the last row and column of each panel is $\Sigma_i H_{it}(s_{it})^2$, averaged across sectors.

Table 3

Distribution of γ Estimates		
Statistic	Value	Standard Deviation
Mean	11.90	1.75
Median	0.19	0.01

Distribution of $1/\gamma$ Estimates		
Statistic	Value	Standard Deviation
Mean	8.06	0.32
Median	5.27	0.34

Table 4

Gamma Values From Sectors with High Shares of Domestic Absorption

Hs4	γ	Average Share of Total Absorption
Passenger motor vehicles	0.14	0.07
Parts and accessories for non-passenger motor vehicles	0.39	0.05
Crude petroleum	0.76	0.04
Automatic data processing machines	0.18	0.03
Non-military aircrafts	0.06	0.02
Cartons, boxes, cases, bags and other packing containers	0.25	0.02
Cell phones	0.07	0.01
Cigarettes	1.41	0.01
Plastics	0.05	0.01
Natural Gas	1.11	0.01

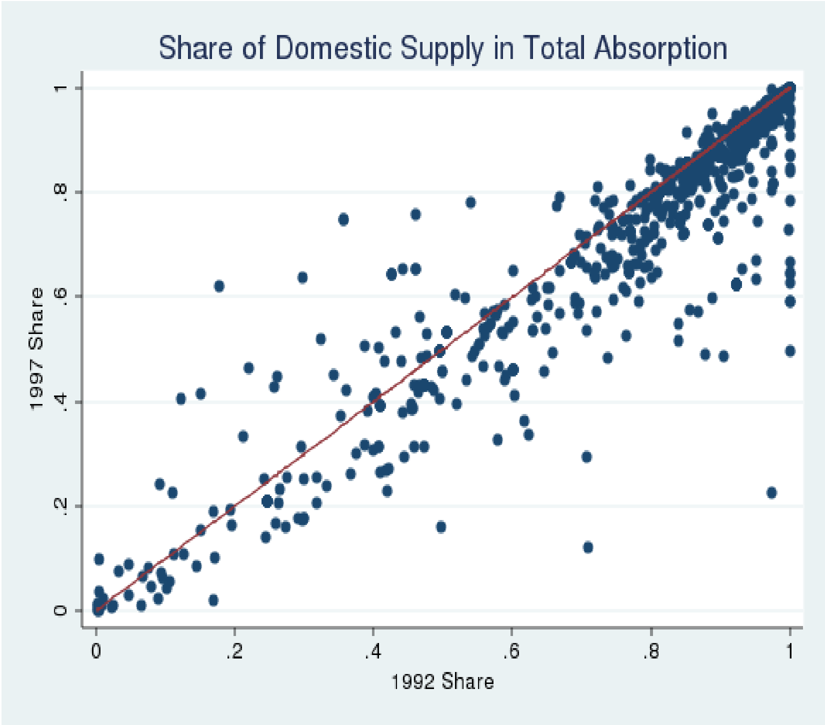
Table 5

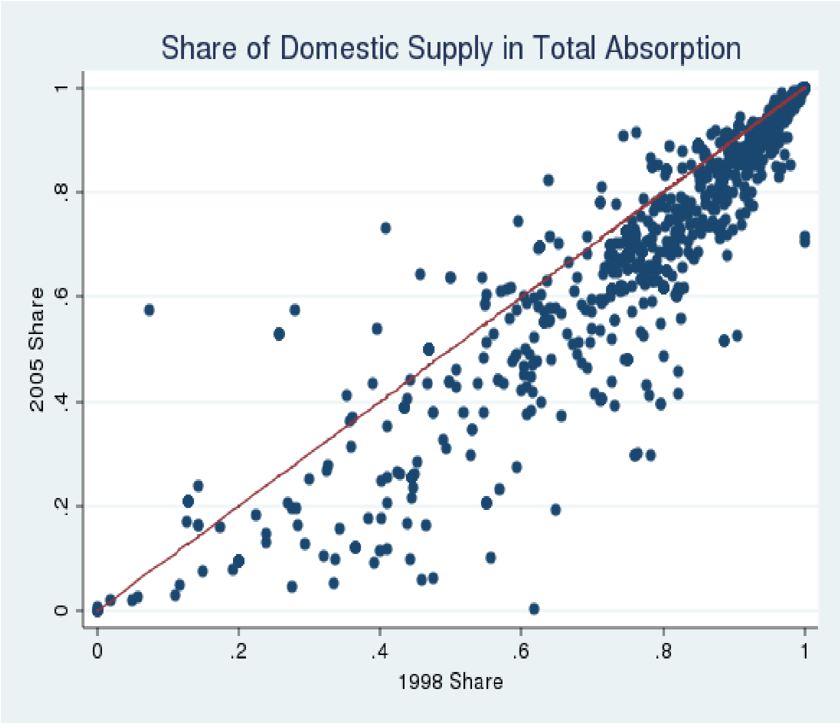
Markup and Variety Welfare Gains as a Percent of GDP

Baseline Results (All Sectors, 1992–2005, Herfindahl Range = ± 30%)							
Hefindahl Range	W	P	V	V_1	V_2	W_{inf}	P_{inf}
± 30%	0.86	0.42	0.44	0.92	-0.48	0.83	0.39
5-95% Confidence Interval	(0.37, 1.22)	(0.28, 0.66)	(0.05, 0.65)	(0.24, 1.72)	(-1.24, -0.01)	(0.34, 1.19)	(0.25, 0.63)
Robustness							
± 20%	0.27	0.25	0.02	0.68	-0.66	0.24	0.22
± 40%	0.78	0.40	0.38	0.87	-0.49	0.74	0.37
Constant HI	0.79	0.66	0.13	1.67	-1.54	0.81	0.68
± 30% (Trimming Top/Bottom 5%)	0.64	0.41	0.23	1.02	-0.79	0.62	0.39
± 30% (Excluding Auto Sector)	0.49	0.23	0.26	0.51	-0.25	0.46	0.20
± 30% (1992–1997)	0.82	0.17	0.65	0.40	0.25	0.80	0.15
± 30% (1998–2005)	0.01	0.24	-0.22	0.49	-0.71	0.00	0.22

Note: $W = P + V$, $W_{inf} = P_{inf} + V$, $V = V_1 + V_2$

Figure 1


Share of Domestic Supply in Total Absorption
1997 Share
0
.2
.4
.6
.8
1
1992 Share
0
.2
.4
.6
.8
1


Share of Domestic Supply in Total Absorption
2005 Share
0
.2
.4
.6
.8
1
1998 Share
0
.2
.4
.6
.8
1