Globalization, Markups, and US Welfare

Robert C. Feenstra

University of California Davis and National Bureau of Economic Research

David E. Weinstein

Columbia University and National Bureau of Economic Research

This paper estimates the impact of globalization on markups, and the effect of changing markups on US welfare, in a monopolistic competition model. We work with symmetric translog preferences, which allow for endogenous markups and firm entry and exit, thereby changing product variety. We find that between 1992 and 2005, US import shares rose and US firms exited, leading to an implied fall in markups, while variety went up because of imports. US welfare rose by nearly 1 percent as a result of these changes, with product variety contributing one-half of that total and declining markups the other half.

I. 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 or improved productivity, and welfare gains due to reduced

We thank Jessie Handbury, Morgan Hardy, Colin Hottman, Philip Luck, and Vladimir Tyazhelnikov for excellent research assistance, along with Sam Kortum, Bob Staiger, and seminar participants at the NBER, Stanford University, University of California San Diego, and the University of Chicago for their helpful comments. Financial support from the National Science Foundation and the Sloan Foundation is gratefully acknowledged. Data are provided as supplementary material online.

Electronically published June 23, 2017
[*Journal of Political Economy*, 2017, vol. 125, no. 4]

markups.[1] While the first two sources of gains have received recent empirical attention, the promise of the third source—reduced markups—has not yet been realized.[2]

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 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.

Beyond these case studies, we have very little evidence about how the broad process of globalization affects markups (see Yilmazkuday 2013; De Blas and Russ 2015; Edmond, Midrigan, and Xu 2015) and no evidence on the impact of such markup reductions on US welfare in particular. This paper structurally estimates the impact of globalization on markups, and the effect of changing markups on US welfare, in a monopolistic competition model. We work with a class of preferences that are relatively new to that literature: translog preferences, with symmetry in substitution imposed across products.[4] 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 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.[5]

[1] These sources are not mutually exclusive. In Krugman (1979), e.g., the welfare gains due to reduced markups are identical to the gains from increasing returns to scale: as the scale of firms expands because of trade, the ratio of average to marginal cost and the ratio of price to marginal cost are both reduced.

[2] The consumer gains due to import variety have been estimated for the United States 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, de Melo, and Corbo (1991) and Tybout and Westbrook (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, Van Cayseele, and Warzynski (2005) for Bulgaria and Romania, and Badinger (2007a) for European countries. Most recently, De Loecker et al. (2016) have found that markups in India increased following liberalization of trade in intermediate inputs.

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

[5] 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

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. Incorporating these two effects into the analysis allows us to estimate the impact of globalization on markups. Translog preferences also allow us to address three potential criticisms of the study by Broda and Weinstein (2006), who estimated welfare gains from US imports using CES preferences: first, those preferences may overstate the gains from import variety because reservation prices are infinite;[6] second, those preferences do not measure the potential crowding of the product space resulting in diminishing returns from new varieties; and third, the CES 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. Surprisingly, we obtain estimates of the gains from trade for the United States that are similar to those under CES, but with different sources. Our point estimate for the cumulative welfare gains to the United States from new varieties and decreased markups in our preferred specification is 0.85 percent over the period 1992–2005. While much of the previous literature has focused on the gains from varieties, our analysis suggests that variety gains account for only one-half of the aggregate gains. That effect incorporates the loss in variety due to the exit of domestic US firms. The remainder of the welfare gain is due to the impact of new competitors on markups. Interestingly, our combined gains for the United States 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 et al. (2015), who argue that with a broad class of preferences including translog, the formula for the gains from trade is much the same as that obtained in the CES case analyzed by Arkolakis, Costinot, and Rodriguez-Clare (2012). But despite this apparent similarity between our results and theirs, the reasons for our findings are quite different.

Blonigen, Liebman, and Wilson (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; e.g., Bilbiie, Ghironi, and Melitz (2012) 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 nonhomothetic preferences that allow for variable markups, see Behrens and Murata (2007), Behrens et al. (2008), and Simonovska (2015).

[6] While the CES system has an infinite reservation price, the area under the demand curve is still bounded above (provided that 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 gains from other functional forms such as translog; see n. 26.

Specifically, Arkolakis et al. (2015) focus on the second reason mentioned in the opening paragraph 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 is trade costs, they find that the efficiency gains are identical in the translog and CES cases. By construction in their model, the first and third reasons for gains from trade mentioned above do not operate when preferences are homothetic: trade brings about changes neither 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), if the distribution of firm productivity is instead Pareto with a bounded support, then the product variety and pro-competitive effects reappear. In this paper we focus on product variety and firm markups and ignore the gain in productivity through the self-selection of firms into exporting. Efficiency gains due to improved productivity will 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 III, we describe the translog expenditure function and solve for the cost-of-living index in the presence of new and disappearing goods. Firms are introduced in Section IV, where we solve for the pro-competitive effect of imports and welfare gains in an open economy. 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. Those estimates can be obtained at the firm level for particular industries; for example, Atkin, Faber, and Gonzalez-Navarro (2015) use a translog structure to estimate markups from microdata for supermarkets in Mexico. It is difficult or impossible to obtain such firm-level data for all industries, however, so another approach must be taken to estimate the effect of trade on the entire economy. A contribution of this paper is to show how the translog demand and implied markups can be consistently aggregated to the industry level, in which case estimation relies on industry-level variables such as the Herfindahl indexes of firm concentration for each country selling to the United States. In Section V, we discuss the procedure for estimation, and results are presented in Section VI. Section VII presents conclusions, and several proofs are gathered in the Appendix.

II. Data Preview

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

From table 1 we see that the share of US apparent consumption sourced domestically fell by 6 percentage points between 1992 and 1997 and by 9 percentage points between 1998 and 2005. This decline corresponds to an average annual decline in the US share of 1.2 percentage points per year in the early period and 1.4 percentage points per year in the later period. The flip side of this decline was a 50 percent increase in the import share, from 0.2 to 0.3. 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.[7]

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 that was the case by looking at more disaggregated data. In figure 1, we plot the US suppliers' share in 1997 against its level in 1992, and likewise in figure 2 for 2005 compared to 1998, for each Harmonized System (HS) four-digit category. It can be easily seen that the vast majority of sectors lie below the 45-degree line, meaning that the domestic share declined quite broadly in both periods. 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 examine what has apparently happened to firm shares, it is convenient to work with Herfindahl indexes of market concentration, defined for each country selling to the United States. We let c denote countries, i denote firms (each selling one product), g denote sectors, and t denote time. Let s_{it}^{cg} denote firm i's exports from country c to the United States in sector g as a share of country c's total exports to the United States in that sector. Then the Herfindahl for country c is

[7] The switch in US classification of output data from the Standard Industrial Classification (SIC) system to the North American Industry Classification System (NAICS) in 1998 makes it difficult to compare sectoral output levels between 1997 and 1998. We therefore break our sample into two periods (1992–97 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 subperiods and the full sample in the results section.

TABLE 1
COUNTRY SHARES OF US TOTAL ABSORPTION

Country	1992 Share	1997 Share	Difference	Country	1998 Share	2005 Share	Difference
United States	.801	.745	−.056	United States	.781	.692	−.089
Total imports	.199	.255	.056	Total imports	.219	.308	.089
Including:				Including:			
Canada	.038	.052	.014	Canada	.043	.056	.013
Japan	.036	.035	.000	Japan	.030	.025	−.004
Mexico	.012	.024	.012	Mexico	.022	.031	.009
Germany	.010	.012	.002	China	.017	.041	.024
China	.010	.017	.007	Germany	.012	.015	.004

$$H_{ct}^{g} = \sum_{i} (s_{it}^{cg})^{2}. \tag{1}$$

The inverse of a Herfindahl can be thought of as the effective number of equally sized exporters, or US 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. If we multiply the Herfindahl by the share of country c within US sector g sales, s_{ct}^{g}, we obtain $H_{ct}^{g} s_{ct}^{g}$, which is interpreted as the share of a typical country c firm within the US market. This statistic will be very useful because in many demand

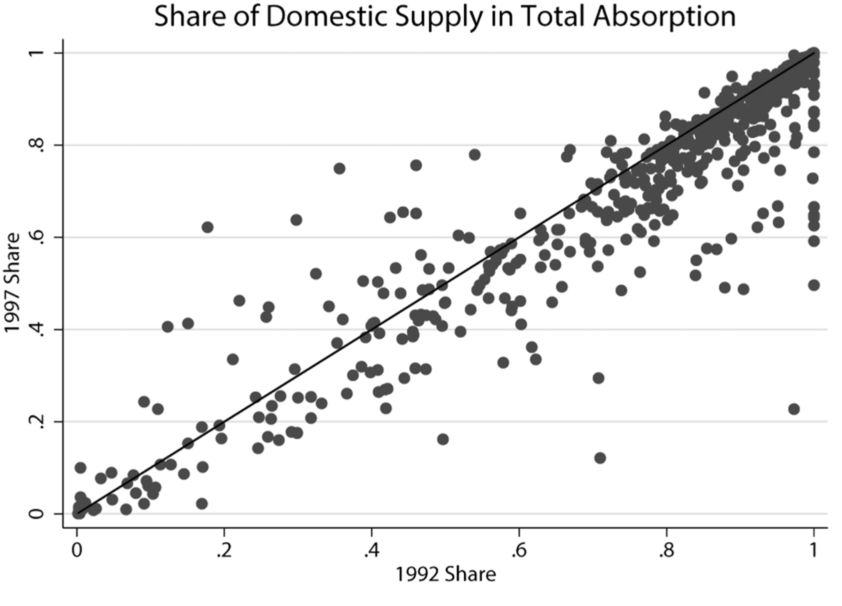

FIG. 1.—Share of domestic supply in total absorption, 1992 and 1995

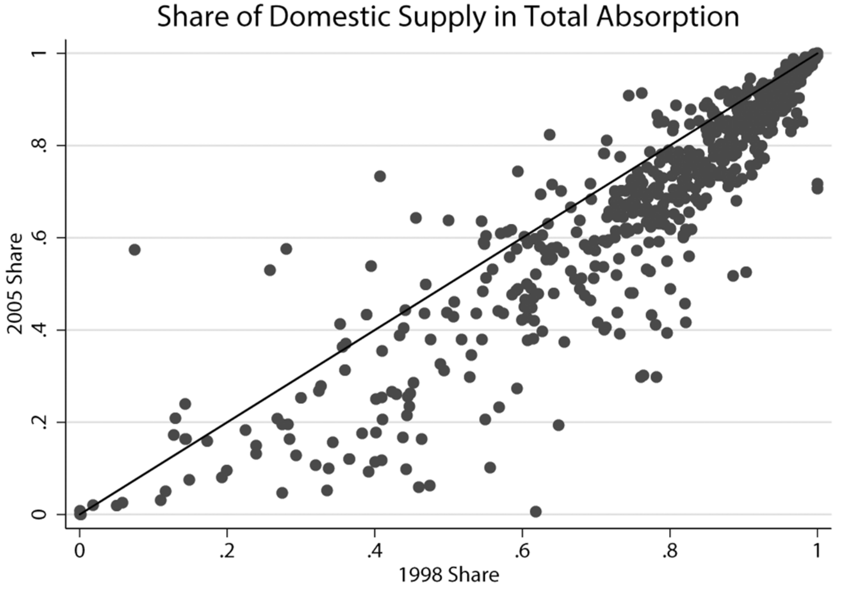

FIG. 2.—Share of domestic supply in total absorption, 1998 and 2005

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 United States, by land or by sea, in 1992 and 2005 (see online app. A). In table 2, we present average Herfindahls at the HS four-digit level for the United States and for the five major exporters to the United States.[8] As one can see from the table, the average US Herfindahl rose slightly over both subperiods, indicating that increased foreign competition was likely associated with some exit of US firms from the market. In contrast, the Herfindahl for China fell in both subperiods, so that it moves from being one of the most concentrated exporters to being the least concentrated of the top five. We report the weighted average of the per-firm market shares $H^g_{ct}s^g_{ct}$ in the last column of table 2, where the weights are based on the importance of each sector in total US consumption. Table 2 reveals that the share in the US market of a typical US firm fell slightly in the first period and by about 7 percent (from 0.139 to 0.129) in the second period. By contrast, exporters to the United States appear to have gained market share in both periods. In other words, those US firms that survived ended up with smaller market shares individ-

[8] For the United States, we have adjusted the NAICS six-digit Herfindahls from the Bureau of Economic Analysis data so that they match the HS four-digit categories and detail that procedure in online app. A.

TABLE 2
Herfindahl indexes by Country

Country	Herfindahl Index 1992	Herfindahl Index 1997	Herfindahl Index Difference	Weighted Average $H^g_{ct}s^g_{ct}$ 1992	Weighted Average $H^g_{ct}s^g_{ct}$ 1997	Weighted Average $H^g_{ct}s^g_{ct}$ Difference
United States	.147	.155	.008	.111	.111	−.001
Canada	.245	.252	.007	.011	.013	.003
Japan	.310	.313	.003	.009	.010	.001
Mexico	.393	.407	.014	.004	.009	.005
Germany	.358	.357	−.002	.003	.004	.001
China	.366	.293	−.073	.001	.002	.001
Weighted average	.160	.170	.010	.078	.069	−.009
Country	1998	2005	Difference	1998	2005	Difference
United States	.183	.189	.006	.139	.129	−.010
Canada	.249	.242	−.008	.011	.015	.003
Japan	.318	.331	.013	.008	.008	−.001
Mexico	.419	.403	−.016	.008	.010	.002
China	.280	.188	−.092	.002	.003	.001
Germany	.332	.335	.002	.003	.005	.001
Weighted average	.190	.191	.001	.090	.071	−.019

Note.—The Herfindahl index is the weighted average over sectors of the country's Herfindahl index of firms selling to the United States, where the weights correspond to the share of each HS four-digit sector g in US apparent consumption. The country share s^g_{ct} is defined to be the country's share of US apparent consumption in sector g. Then the reported values of $H^g_{ct}s^g_{ct}$ are the weighted average over sectors of the Herfindahl index in sector g in year t multiplied by that country's share of US apparent consumption and are interpreted as the market share of a typical firm from that country selling to the United States The last row reports a weighted average across sectors and also across all countries, using each country's share of US apparent consumption as weights.

ually while some foreign firms gained market share. But the overall average of firm shares, shown by the final weighted averages in the last column of table 2, declined in both periods.

This data preview suggests that prior work on the impact of new varieties (e.g., Broda and Weinstein 2006) is likely to suffer from a number of biases. First, as foreign firms have entered the US market, there has been exit by US firms, which serves to offset some of the gains of new varieties. Second, while the US Herfindahl rose, the Herfindahl of one of our largest suppliers, China, fell substantially. The fall in this Herfindahl suggests that there may have been substantial variety growth that is not captured in industry-level analyses. Finally, because the market shares of both US 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.[9] Thus, estimates of the gains from new varieties obtained from

[9] While we infer the change in markups for US 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 n. 3, several studies have also shown that falls in domestic market share

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 rest of the paper.

III. Consumer Gains with a 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.[10] The translog function is defined over the universe of products, whose maximum number is denoted by the fixed number N. The expenditure needed to obtain a fixed level of utility in period t, or the cost of living, is

$$\ln e_t = \alpha_0 + \sum_{i=1}^{N} \alpha_i \ln p_{it} + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \gamma_{ij} \ln p_{it} \ln p_{jt}, \quad \text{with } \gamma_{ij} = \gamma_{ji}. \tag{2}$$

Note that the restriction that $\gamma_{ij} = \gamma_{ji}$ is made without loss of generality.[11] We add the further restrictions that the γ_{ii} and γ_{ij} coefficients are "homogeneous" across goods, by which we mean

$$\gamma_{ii} = -\gamma \left(\frac{N-1}{N} \right) < 0, \quad \gamma_{ij} = \frac{\gamma}{N} > 0, \tag{3}$$

for $i \neq j$, with $i, j = 1, \dots, N$. These symmetry restrictions have not been used in past work dealing with estimating the translog function but are essential in our monopolistic competition model.

The share of each good in expenditure can be computed by differentiating (2) with respect to $\ln p_{it}$ and using (3) to obtain

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 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.

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

[11] We also require that $\Sigma_{i=1}^{N} \alpha_i = 1$ and $\Sigma_{i=1}^{N} \gamma_{ij} = 0$ for the expenditure function to be homogeneous of degree one.

$$s_{it} = \alpha_i + \sum_{j=1}^{N} \gamma_{ij} \ln p_{jt} = \alpha_i - \gamma[\ln p_{it} - \overline{\ln p(t)}],$$
$$\text{with } \overline{\ln p(t)} \equiv \sum_{j=1}^{N} \frac{1}{N} \ln p_{jt}. \tag{4}$$

Notice that the term $\overline{\ln p(t)}$ refers to the average over 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.[12] So we want to solve for the reservation prices for goods not available and then substitute these into the expenditure function and share equations to simplify these expressions. 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:

$$\ln p_{it} = \frac{\alpha_i - s_{it}}{\gamma} + \overline{\ln p(t)},$$

which equals the reservation price of good i when $s_{it} = 0$. It is convenient to substitute this solution for price back into the expenditure function (2) to obtain an expression for the cost of living in terms of shares. After some simplification, we obtain (see the Appendix)

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

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

$$\overline{\ln p(t)} \equiv \sum_{i=1}^{N} \frac{1}{N} \ln p_{it} = \frac{1}{N} \left\{ \sum_{i \in I_t} \ln p_{it} + \sum_{i \notin I_t} \left[\frac{\alpha_i}{\gamma} + \overline{\ln p(t)} \right] \right\},$$

where $I_t \subseteq \{1, \ldots, N\}$ denotes the set of goods available each period, with number $N_t \leq N$, and the reservation prices for goods $i \notin I_t$ are ob-

[12] It is worth pointing out how the demand equation given in (4) differs from that of the CES functional form. The CES demand system has almost the same functional form as the tranlog except that s_{it} is replaced by $\ln s_{it}$, while $\overline{\ln p(t)}$ is replaced by a CES price index, and the parameter γ instead appears as the elasticity of substitution minus one. For our purposes, however, the critical difference is the fact that the left-hand side of eq. (4) is expressed in levels in the translog case and not logs as in the CES case. This difference means that while a sufficiently high price can result in a good having no sales (i.e., a share of zero) in the translog case, the reservation price is infinite in the CES case.

tained from (4) with $s_{it} = 0$. Since the number of goods available is $N_t = N - \Sigma_{i \notin I_t} 1$, we can solve for $\overline{\ln p(t)}$ from the above equation as

$$\overline{\ln p(t)} = \frac{1}{N_t}\left(\sum_{i \in I_t} \ln p_{it} + \sum_{i \notin I_t} \frac{\alpha_i}{\gamma}\right) = \overline{\ln p_t} + \frac{\bar{\alpha}_t}{\gamma},$$

where

$$\overline{\ln p_t} \equiv \frac{1}{N_t} \sum_{i \in I_t} \ln p_{\mathrm{it}}$$

and[13]

$$\bar{\alpha}_t \equiv \frac{1}{N_t}\left(\sum_{i \notin I_t} \alpha_i\right) = \frac{1}{N_t}\left(1 - \sum_{i \in I_t} \alpha_i\right).$$

Notice that $\overline{\ln p_t}$ (as distinct from $\overline{\ln p(t)}$) is the average of log prices over the goods actually available, while $\bar{\alpha}_t$ is inversely related to the range of goods available: as the set I_t expands, the summation $\Sigma_{i \in I_t} \alpha_i$ rises and N_t also rises, so $\bar{\alpha}_t$ falls. We can substitute these two terms back into the share equation (4) and into the expenditure function (5) to obtain

$$\begin{aligned} s_{it} &= \alpha_i + \bar{\alpha}_t - \gamma\left(\ln p_{it} - \overline{\ln p_t}\right), \\ \ln e_t &= \alpha_0 + \frac{1}{2\gamma} \sum_{i=1}^{N} (\alpha_i)^2 + \overline{\ln p_t} + \frac{\bar{\alpha}_t}{\gamma} - \frac{1}{2\gamma} \sum_{i=1}^{N} (s_{it})^2. \end{aligned} \tag{6}$$

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

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

Equation (7) is a useful expression for the change in the cost of living and consists of three terms. The first term is the change in the average of log prices for available products. The second term is a shift parameter $\Delta\bar{\alpha}_t/\gamma$, where, as noted above, the entry of new products results in $\Delta\bar{\alpha}_t < 0$ so that the cost of living falls. The third term is the change in the Herfindahl index of product shares. Let us interpret the second and third terms while also comparing equation (7) to its analogue in the CES case.

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 such products $i \in \bar{I}$ are unchanging. Then $\Delta s_{it} = \Delta\bar{\alpha}_t$ for $i \in \bar{I}$ from (6). So averaging over the $\bar{N}$ products, we obtain

[13] The final equality uses $\Sigma_{i=1}^{N} \alpha_i = 1$, as in n. 11.

$$\begin{aligned}\Delta\bar{\alpha}_t &= \frac{1}{\bar{N}}\sum_{i\in\bar{I}}\Delta s_{it}\\ &= \frac{1}{\bar{N}}\left[\left(1-\sum_{i\notin\bar{I}}s_{it}\right)-\left(1-\sum_{i\notin\bar{I}}s_{it-1}\right)\right]\\ &= -\frac{1}{\bar{N}}\left(\sum_{i\notin\bar{I}}s_{it}-\sum_{i\notin\bar{I}}s_{it-1}\right),\end{aligned}$$

where we use the fact that the shares sum to unity each period. Thus, $-\Delta\bar{\alpha}_t$ is directly related to the share of expenditure on new goods minus the share of expenditure on disappearing goods, which is the last term appearing in parentheses above.

Feenstra (1994) shows that these same shares, in conjunction with the elasticity of substitution, also determine the gains from variety in the CES case. Specifically, from Feenstra (1994), we can express the CES price index as

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

where the weights w_{it} appearing in (8) sum to unity over $i \in \bar{I}$.[14] The first expression on the right of (8) is a change of the weighted average of the log price changes, similar in spirit to the unweighted average appearing as the first term on the right of (7). The second expression on the right of (8) is the welfare gain from new and disappearing goods in the CES case, which equals the net change in expenditure on new and disappearing goods divided by the elasticity of substitution minus unity $(\sigma - 1)$. That expression is very similar to the second term $\Delta\bar{\alpha}_t/\gamma$ on the right of (7), which likewise involves dividing the net change in expenditure on new and 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 (7) and (8). The key difference is that in the translog case we have an additional term in (7), which is the change in the Herfindahl index of product shares and 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? This effect is due to "crowding" in product space, as we now argue.

14 The weights w_{it} appearing in (8) are the Sato-Vartia weights defined in Feenstra (1994, 158–59).

Consider replacing one good having a high share parameter α_1 in period $t-1$ with two other goods having lower values of α_2 and α_3 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 $s_{4t}, \ldots, s_{Nt}$ unchanged). Then 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 raises the cost of living in (7) because product space is more "crowded," so that goods are more substitutable.[15] In other words, as the number of varieties increases, we care less about each new variety. This result explains why the Herfindahl index of product shares enters the cost-of-living index in a counterintuitive way. In the example we have constructed, there are no variety gains because the shares of new and disappearing goods are the same, $s_{1t-1} = s_{2t} + s_{3t}$, so that only the crowding effect remains, which lowers utility. In contrast, when there are only new goods, then it must be the case that the gains from variety are greater than the negative crowding effect, so that the consumer benefits. Finally, note 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 cancel out in the last term of (8) and therefore have no impact on consumer utility.

IV. Welfare Gains in an Open Economy

To develop an expression for welfare in an open economy, suppose that firms selling to the United States each produce a single product and act as Bertrand competitors. The profit maximization problem for a firm selling product i in period t is

$$\max_{p_{it}>0} p_{it} x_i(p_t, E_t) - \psi_i[x_i(p_t, E_t)],$$

where $x_i(p_t, E_t)$ denotes the demand arising from the translog system, with the price vector p_t and expenditure E_t, and $\psi_i[x_i(p_t, E_t)]$ is the cost of production. We denote the elasticity of demand by $\eta_{it} \equiv -\partial \ln x_i(p_t, E_t)/\partial \ln p_{it}$, which from (6) is

$$\eta_{it} = 1 - \left(\frac{\partial \ln s_{it}}{\partial \ln p_{it}}\right) = 1 + \frac{\gamma(N_t - 1)}{s_{it} N_t}. \tag{9}$$

[15] This claim can be confirmed from the elasticity of demand in (9), which rises as the share falls or as the number of goods N_t increases. The Herfindahl index also enters the AIDS welfare formula of Fajgelbaum and Khandelwal (2016) and the quadratic mean of order r formula analyzed by Feenstra (2014), again reflecting crowding.

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

$$\ln p_{it} = \ln\psi'_{it} + \ln\left(\frac{\eta_{it}}{\eta_{it}-1}\right) = \ln\psi'_{it} + \ln\left[1 + \frac{s_{it}N_t}{\gamma(N_t-1)}\right], \quad (10)$$

where $\psi'_{it} \equiv \psi'_i[x_i(p_t, E_t)]$ denotes marginal cost.

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 $W_t = w_t/e_t$, or the real wage. The proportional change in welfare is $\Delta\ln W_t = \Delta\ln w_t - \Delta\ln e_t$. We evaluate the cost-of-living index from (7) using the prices in (10), which gives

$$\begin{aligned}\Delta\ln W &= \Delta\ln w_t - \Delta\ln e_t \\ &= \left[\Delta\ln w_t - \Delta\left(\frac{1}{N_t}\sum_{i\in I_t}\ln\psi'_{it}\right)\right] \\ &\quad -\Delta\left\{\frac{1}{N_t}\sum_{i\in I_t}\ln\left[1+\frac{s_{it}N_t}{\gamma(N_t-1)}\right]\right\} - \left[\frac{\Delta\bar{\alpha}_t}{\gamma} - \frac{1}{2\gamma}\Delta\left(\sum_{i\in I_t}s_{it}^2\right)\right].\end{aligned} \quad (11)$$

The first term in brackets on the right in (11) is the change in wages relative to an average of marginal costs, for both domestic and foreign-produced goods. Its role in 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 (11) 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 Arkolakis et al. (2012, 2015), but we do not attempt to measure it here. The second term on the right of (11) is the change in average firm markups, with a reduction in markups indicating a rise in welfare.[16] The third term is the gains from consumer variety that we solved for in the previous section and includes the negative effect of crowding in product space as captured by the change in the Herfindahl index.

Our focus will be on measuring the final two terms—the pro-competitive and product variety effects. In order to implement formula (11), we must

[16] As mentioned in n. 1, the reduction in markups is more than just redistribution from firms to consumers because—with zero profits—a reduced ratio of price to marginal cost corresponds to a reduced ratio of average cost to marginal cost, so the firm is taking greater advantage of economies of scale.

recognize that we do not have data at the level of individual products or individual firms, but instead observe Herfindahl indexes by country of origin for exporters and domestic firms selling to the United States, as well as other industry aggregates. We therefore need to reexpress this formula in terms of weighted averages rather than simple averages, so that larger countries and industries receive more weight. The typical weights used for a translog expenditure function are the average of expenditure shares over periods $t-1$ and t. In the remainder of this section we show how to reexpress the welfare formula (11) using Herfindahl indexes and these average expenditure shares, and in the next section we show how to estimate the parameter γ for each sector.

Let I_{ct} denote the set of products exported from country c to the United States each period, and let $c \in C_t$ denote the set of supplying countries (including the United States itself). We assume that each firm from country c sells one product $i \in I_{ct}$ in a given sector, with its share denoted by s_{ict}, so that i also indexes firms. The total import share from country c is $s_{ct} \equiv \sum_{i \in I_{ct}} s_{ict}$, and we let $s^c_{it} \equiv s_{ict}/s_{ct}$ denote firms' shares within the total sales of country c. Then the Herfindahl index of firms exporting to the United States from country c is $H_{ct} \equiv \sum_{i \in I_{ct}} (s^c_{it})^2$. The total Herfindahl index over all firms, which appears as the last term in (11), is measured by

$$\sum_{c \in C_t} \sum_{i \in I_{ct}} s^2_{ict} = \sum_{c \in C_t} \sum_{i \in I_{ct}} (s^c_{it})^2 s^2_{ct} = \sum_{c \in C_t} H_{ct} s^2_{ct} .$$

In other words, by summing the country Herfindahl indexes times the square of country shares, we obtain the overall Herfindahl index.

The Herfindahl indexes can be used to reexpress the share equation. Using our new notation for countries c and products i, the share equation in (6) becomes

$$s_{ict} = \alpha_{ic} + \bar{\alpha}_t - \gamma(\ln p_{ict} - \overline{\ln p_t}), \tag{12}$$

where

$$\bar{\alpha}_t \equiv \frac{1}{N_t}\left(1 - \sum_{c,i} \alpha_{ic}\right)$$

is the shift parameter already discussed and

$$\overline{\ln p_t} \equiv \frac{1}{N_t} \sum_{c,i} \ln p_{ict}$$

is the average log price of all available goods in period t.[17] Multiplying the above equation by the firms' shares within the total sales of a country,

[17] We use the summation $\sum_{c,i}$ as shorthand for $\sum_{c \in C_t} \sum_{i \in I_{ct}}$, i.e., to sum over all available products and countries in each period.

$s^c_{it} \equiv s_{ict}/s_{ct}$, and summing over firms from country c (with $\Sigma_{i \in I_{ct}} s^c_{it} = 1$), we obtain

$$H_{ct} s_{ct} = \alpha_{ct} + \bar{\alpha}_t - \gamma\left(\ln p_{ct} - \overline{\ln p_t}\right),$$

where $\alpha_{ct} = \Sigma_{i \in I_{ct}} s^c_{it} \alpha_{ic}$ and $\ln p_{ct} \equiv \Sigma_{i \in I_{ct}} s^c_{it} \ln p_{ict}$ are weighted averages over country products of the taste parameters and log prices, respectively.

It is natural to model α_{ct} as a country fixed effect plus an error term:

$$\alpha_{ct} = \alpha_c + \varepsilon_{ct}.$$

Substituting this equation above, we obtain the share equation

$$H_{ct} s_{ct} = \alpha_c + \bar{\alpha}_t - \gamma(\ln p_{ct} - \overline{\ln p_t}) + \varepsilon_{ct}. \tag{13}$$

Notice that the dependent variable of (12), the share of product i from country c, has been replaced in (13) by the country c Herfindahl index times the country c share. Interpreting the country c Herfindahl as the inverse of the number of firms, then its product with the country c share becomes an estimate of the expenditure share on the output of a typical firm (as discussed in Sec. II). Using this $H_{ct}s_{ct}$ term as the dependent variable in (13) allows the share equation to be estimated without firm-level data. In addition, it will give us a method to identify the change in shift parameter $\bar{\alpha}_t$, which appears in the welfare expression (11). Differencing (13), multiplying by $\frac{1}{2}(\bar{s}_{ct-1} + \bar{s}_{ct})/\gamma$ and summing over $c \in \bar{C}$, we obtain

$$\Delta\left(\overline{\ln p_t} + \frac{\bar{\alpha}_t}{\gamma}\right) = \sum_{c \in \bar{C}} \frac{1}{2}\left(\bar{s}_{ct-1} + \bar{s}_{ct}\right)\left[\Delta \ln p_{ct} + \frac{1}{\gamma}\Delta(H_{ct}s_{ct}) - \frac{\Delta\varepsilon_{ct}}{\gamma}\right].$$

Substituting this expression into (7) yields

$$\begin{aligned} \Delta \ln e_t = {} & \sum_{c \in \bar{C}} \frac{1}{2}\left(\bar{s}_{ct-1} + \bar{s}_{ct}\right)\left[\Delta \ln p_{ct} + \frac{1}{\gamma}\Delta(H_{ct}s_{ct}) - \frac{\Delta\varepsilon_{ct}}{\gamma}\right] \\ & - \frac{1}{2\gamma}\Delta\left(\sum_{c \in C_t} H_{ct} s^2_{ct}\right), \end{aligned} \tag{14}$$

as we shall use below.

The same technique of relying on the Herfindahl indexes can be used to estimate the markups, up to a first-order approximation. The pricing equation (10) is modified by adding the subscript c so that prices, marginal costs, and shares become p_{ict}, ψ'_{ict}, and s_{ict}, respectively. We aggregate this expression by taking a weighted average using the firm shares $s^c_{it} \equiv s_{ict}/s_{ct}$ within the total exports of country c. Then the geometric average of prices from country c is

$$\ln p_{ct} \equiv \sum_{i\in I_{ct}} s^c_{it} \ln p_{ict} = \ln \psi'_{ct} + \mu_{ct},$$
$$\mu_{ct} \equiv \sum_{i\in I_{ct}} s^c_{it} \ln\left[1 + \frac{s^c_{it} s_{ct} N_t}{\gamma(N_t - 1)}\right], \tag{15}$$

where $\ln\psi'_{ct} \equiv \Sigma_i s^c_{it} \ln\psi'_{ict}$ is the average of marginal costs and the final term in (15) is the average markup, denoted by μ_{ct}. We are not able to measure the average markup directly in the absence of firm-level information, but data on the Herfindahl indexes for each country will allow us to measure a first-order approximation to it around the point at which the firm shares are equal. In that case the shares would be measured by the Herfindahl index, $s^c_{it} = H_{ct}$, and the average markup is approximated as[18]

$$\ln\mu_{ct} \approx \sum_{i\in I_{ct}} s^c_{it}\left\{\ln\left[1 + \frac{H_{ct} s_{ct} N_t}{\gamma(N_t - 1)}\right] + \left[\frac{\dfrac{s^c_{it} s_{ct} N_t}{\gamma(N_t - 1)} - \dfrac{H_{ct} s_{ct} N_t}{\gamma(N_t - 1)}}{1 + \dfrac{s_{ct} H_{ct} N_t}{\gamma(N_t - 1)}}\right]\right\} \tag{16}$$
$$= \ln\left[1 + \frac{H_{ct} s_{ct} N_t}{\gamma(N_t - 1)}\right].$$

The right-hand side of this expression is the log markup for a firm with the average share $s_{ict} = H_{ct} s_{ct}$. In other words, we are ignoring the variation in firm sizes within countries when we compute the first-order approximation in (16).[19] This approach shows how we can estimate the pricing equation even in the absence of having firm-level data for all countries

[18] The first-order approximation to the log function is $\ln(1 + x) \approx \ln(1 + a) + [(x - a)/(1 + a)]$ around $x = a$, so we apply this formula to the average markup around $s^c_{it} = H_{ct}$ in (16), using $H_{ct} \equiv \Sigma_{i\in I_{ct}} (s^c_{it})^2$ and $\Sigma_{i\in I_{ct}} s^c_{it} = 1$ to obtain the second line.

[19] In order to assess the accuracy of this first-order approximation, we simulated firm-level data to compare the true markup μ_{ct} with the approximation in (16). We begin by assuming that $\gamma = 0.19$ (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 one and a minimum value of 10. Then we assumed that there were 500 firms in each sector and that every 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 the simulated firm-level data to construct the true markup equation, μ_{ct}. We then calculated the percentage difference of the approximation, the last term on the right of (16), 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 United States, we found that the median difference was −0.005 percent and the mean difference was −0.28 percent, indicating that the first-order approximation holds quite closely.

selling to the United States, using the industry-by-country Herfindahl indexes instead.[20]

We now combine the previous results to obtain an expression for welfare that we can implement with available data. Recall that the error term in the share equation (13) is the weighted change in the taste parameters for each country, $\Delta\varepsilon_{ct} = \Delta\alpha_{ct} \equiv \Delta\Sigma_{i\in I_{ct}} s^c_{it}\alpha_{ic}$. It is challenging to measure exact price indexes when all taste parameters are changing, and to overcome this problem in the CES case, Feenstra (1994) assumed that there was a subset of countries for which there were no changes in taste parameters. In our framework, the taste parameters might change as a result of a differing set of goods being exported by a country. Analogous to the CES case, we assume that there is a subset of countries $c \in \bar{C} \subseteq C_t \cap C_{t-1}$ for which the set of goods exported by each and their taste parameters α_{ic} do not change, so that $\Delta\varepsilon_{ct} = 0$. We define the expenditure shares over these countries as

$$\bar{s}_{ct} \equiv s_{ct} + \frac{1}{|\bar{C}|}\left(1 - \sum_{c\in\bar{C}} s_{ct}\right) \quad \text{for} \quad c \in \bar{C}, \tag{17}$$

where $|\bar{C}|$ is the number of countries in $\bar{C}$, so that $\bar{s}_{ct}$ sums to unity over them.

Then substituting (15) and (16) back into (14), and using $\Delta\varepsilon_{ct} = 0$ for $c \in \bar{C}$, we compute the change in welfare $\Delta\ln W_t = \Delta\ln w_t - \Delta\ln e_t$ as

$$\Delta\ln W \approx \left[\Delta\ln w_t - \sum_{c\in\bar{C}} \frac{1}{2}(\bar{s}_{ct-1} + \bar{s}_{ct})\Delta\ln\psi'_{ct}\right] + P_t + V_t, \tag{18}$$

where

$$P_t \equiv -\sum_{c\in\bar{C}} \frac{1}{2}(\bar{s}_{ct} + \bar{s}_{ct-1})\Delta\ln\left[1 + \frac{H_{ct}s_{ct}N_t}{\gamma(N_t - 1)}\right], \tag{19}$$

and

$$V_t \equiv -\sum_{c\in\bar{C}} \frac{1}{2\gamma}(\bar{s}_{ct-1} + \bar{s}_{ct})\Delta(H_{ct}s_{ct}) + \frac{1}{2\gamma}\Delta\left(\sum_{c\in C_t} H_{ct}s^2_{ct}\right). \tag{20}$$

[20] It is possible that (15) 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 online app. A, to check for the prevalence of trading companies, i.e., those whose names include "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 United States in 1992 and 7.5 percent in 2005.

The approximation in (18) follows directly from the first-order approximation to markups. The first bracketed term on the right of (18) includes a share-weighted average of the change in marginal costs for countries in the set $\bar{C}$. That expression has much the same interpretation as the change in the simple average of costs in (11); that is, it reflects the self-selection of more efficient firms, and we do not attempt to measure it. Our focus is on the next two terms.

The pro-competitive effect P_t in (19) depends on the change in $H_{ct}s_{ct}$, which is the change in the expenditure share on the output of a typical firm from country c. From table 1, we see that the typical market share was falling, on average, 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. Turning to the product variety effect V_t in (20), it can be compared to the final bracketed term appearing in (11). The theoretical shift parameter $\Delta\alpha_t$ appearing there has been replaced in (20) by the weighted average of the drop in typical firm shares $\Delta(H_{ct}s_{ct})$ for countries within the set $\bar{C}$. The final term appearing in (20) is the change in the overall Herfindahl index, which reflects crowding and is unchanged from (11).

Both of the terms appearing in the variety effect (20) depend on the change in the country Herfindahl indexes H_{ct}. To understand how these terms might sum up, consider a simplified case in which there is no change in the set of countries selling to the United States, so that we can choose $\bar{C} = C_{t-1} = C_t$. 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 United States. For the translog case, however, we show in the Appendix that if $\Delta\alpha_{ct} = 0$ for $c \in \bar{C} = C_{t-1} = C_t$, then there are still potential variety gains of

$$V_t = -\frac{1}{2\gamma}\sum_{c\in\bar{C}} s_{ct-1}s_{ct}\Delta H_{ct}. \tag{21}$$

In this case, the changes in the Herfindahl indexes for each country are used to infer entry or exit of firms and products: falling Herfindahls due to entry will contribute toward gains with $V_t > 0$, whereas exit will contribute toward losses with $V_t < 0$.

The remaining question is how the set of countries $\bar{C} \subseteq C_{t-1} \cap C_t$, for which the set of goods exported by each and their taste parameters α_{ic} do not change, will be chosen. Broda and Weinstein (2006) conservatively chose the set $\bar{C}$ as the intersection of countries supplying in the first and last years of the sample. So for each sector, only those countries that were not exporting at all to the United States in the first or last year of the sample were excluded from the set $\bar{C}$ and were therefore used to measure the change in the product variety of exports to the United States. We can improve on this assumption by using information on the Herfindahl indexes for exporters and domestic firms selling to the United States in each sector. We shall interpret any country Herfindahl indexes that are changing by more than a specified tolerance as evidence to exclude those countries from the set $\bar{C}$, as will be discussed in more detail in Section VI.

Expressions (18)–(20) are the formulas we implement for the change in welfare for each sector. Of course, we need to sum the welfare gains across all of the manufacturing sectors in our sample. We do this by adding a sector superscript g to P_t and V_t, then multiplying the welfare gain in each industry by the appropriate translog weight, and summing across all sectors,

$$\Delta \ln W_t^{PV} = \sum_g \frac{1}{2}(s_{t-1}^g + s_t^g)(P_t^g + V_t^g), \tag{22}$$

where s_t^g is the share of sector g in US absorption in period t.[21] Since the consumption of merchandise accounts for only about 19 percent of US 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 US economy. We omit the sector g notation in the next section, though our estimation is done separately over each four-digit HS sector.

V. Estimation

Our next task is to address how to estimate the translog parameter γ, which appears in the share equation (13). In practice, the prices p_{ct} in (13) are measured by the unit value of imports from each source country and sector, or by the price index for each sector in the United States. We further specify that the marginal costs from each exporting country, or from the United States, take on the log-linear form with elasticity $\omega \geq 0$:

$$\ln \psi'_{ct} = \omega_{c0} + \omega \ln\left(\frac{s_{ct}E_t}{p_{ct}}\right) + \delta_{ct}, \tag{23}$$

where $s_{ct}E_t/p_{ct}$ is the total quantity sold by country c, E_t is US expenditure, and δ_{ct} is an error. Combining the above equation with (15) and (16), we obtain a modified pricing equation:

[21] 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 US 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 US absorption was set equal to US final goods demand.

$$(1 + \omega)\ln p_{ct} = \omega_{c0} + \omega \ln s_{ct} + \omega \ln E_t + \ln\left[1 + \frac{H_{ct} s_{ct} N_t}{\gamma(N_t - 1)}\right] + \delta_{ct}. \quad (24)$$

Prices are jointly determined by (24) and the share equation (13), repeated here:

$$H_{ct} s_{ct} = \alpha_c + \bar{\alpha}_t - \gamma(\ln p_{ct} - \overline{\ln p_t}) + \varepsilon_{ct}.$$

Notice that the per-firm share $H_{ct}s_{ct}$ appearing as the dependent variable in this share equation determines the markup in the pricing equation (24). This means that any shock to the share through the error term ε_{ct} will influence the markups and therefore the price, so that the price appearing in the share equation is correlated with that error ε_{ct}. Adding to this endogeneity problem, if marginal costs depend on quantity so that $\omega > 0$, then the share also appears directly within the pricing equation. To control for this endogeneity, we will estimate these equations simultaneously using a generalized method of moments approach similar 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 (13) and (24) with respect to country k and with respect to time, thereby eliminating the terms $\alpha_c + \bar{\alpha}_t$ along with the average prices $\overline{\ln p_t}$ and expenditure $\ln E_t$ appearing in these equations. 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:

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

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

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

$$\begin{aligned}\bar{Y}_c = {} & \frac{\omega}{1+\omega}\bar{X}_{1c} + \frac{\omega}{\gamma(1+\omega)}\bar{X}_{2c} - \left(\frac{1}{\gamma}\right)\bar{X}_{3c} + \frac{1}{1+\omega}\bar{Z}_{1c}(\gamma) \\ & + \frac{1}{\gamma(1+\omega)}\bar{Z}_{2c}(\gamma) + \bar{u}_c,\end{aligned} \quad (25)$$

where the overbar indicates that we are averaging that variable over time, and

$$
\begin{aligned}
Y_{ct} &\equiv (\Delta \ln p_{ct} - \Delta \ln p_{kt})^2, \\
X_{1ct} &\equiv (\Delta \ln s_{ct} - \Delta \ln s_{kt})(\Delta \ln p_{ct} - \Delta \ln p_{kt}), \\
X_{2ct} &\equiv (\Delta \ln s_{ct} - \Delta \ln s_{kt})[\Delta(H_{ct} s_{ct}) - \Delta(H_{kt} s_{kt})], \\
X_{3ct} &\equiv (\Delta \ln p_{ct} - \Delta \ln p_{kt})[\Delta(H_{ct} s_{ct}) - \Delta(H_{kt} s_{kt})], \\
Z_{1ct}&(\gamma) \\
&\equiv \left\{\Delta \ln\left[1 + \frac{H_{ct} s_{ct} N_t}{\gamma(N_t - 1)}\right] - \Delta \ln\left[1 + \frac{H_{kt} s_{kt} N_t}{\gamma(N_t - 1)}\right]\right\}(\Delta \ln p_{ct} - \Delta \ln p_{kt}), \\
Z_{2ct}&(\gamma) \\
&\equiv \left\{\Delta \ln\left[1 + \frac{H_{ct} s_{ct} N_t}{\gamma(N_t - 1)}\right] - \Delta \ln\left[1 + \frac{H_{kt} s_{kt} N_t}{\gamma(N_t - 1)}\right]\right\}[\Delta(H_{ct} s_{ct}) - \Delta(H_{kt} s_{kt})],
\end{aligned}
$$

and

$$
u_{ct} \equiv \frac{(\Delta \varepsilon_{ct} - \Delta \varepsilon_{kt})(\Delta \delta_{ct} - \Delta \delta_{kt})}{\gamma(1 + \omega)} \quad \text{for} \ \ c \in \bar{C} \subseteq C_{t-1} \cap C_t, \, c \neq k.
$$

We assume that the error terms in demand and the pricing equation are uncorrelated, which means that the error term in (25) becomes small, $\bar{u}_c \to 0$ in probability limits 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 ordinary least squares on (25). Feenstra (1994) shows that that procedure will give us consistent estimates of γ and ω in a simpler CES system, provided that the right-hand-side variables in (25) are not perfectly collinear as $T \to \infty$.[22] That condition is assured in the CES case of Feenstra (1994) if there is some heteroskedasticity in the error terms across countries c, so that the right-hand-side variables in (25) are not perfectly collinear. More efficient estimates can be obtained by running weighted least squares on (25).

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 HS 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 four-digit sector are the same and therefore pool the HS 10-digit goods within each four-digit sector.[23]

[22] 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).

[23] 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 four-digit sector.

A second complication arises because we have US shipments data at the NAICS six-digit level, but we need to compute shares at the HS 10-digit level. Thus, we must allocate NAICS six-digit production data to each HS 10-digit sector. This allocation is achieved by assuming that the share of US production in each HS 10-digit sector is the same as that of the United States in the NAICS six-digit sector that contains it, as discussed in online 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-digit level. Unfortunately, we cannot implement precisely that scheme because the US quantity indexes were defined at the NAICS six-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.

A fourth issue arises in the measurement of the number of available products N_t, which appears in the pro-competitive effect (19) and in the estimating equations. We shall consider two methods of measuring N_t. One approach is to measure N_t by the inverse of the overall Herfindahl, 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 $N_t/(N_t - 1)$ 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 (12), the average log prices $\overline{\ln p_t}$ acts like a simple price index for the sector. If the firm does not consider the impact of its own price on this price index, then the elasticity of demand is

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

Comparing this result with (9), 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.

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 markups are negative and there is no equilibrium. To avoid that outcome and because V_t in (20) is very sensitive to small values of γ, we impose $\gamma \geq 0.05$. To satisfy that constraint, we use a grid search procedure over γ and ω to minimize the sum of squared errors in (25). In this procedure we set an initial value for γ of 0.05 and increased it by 5 percent over the range [0.05, 110].[24] Similarly, we set an initial ω of -5 and increased it by 0.1 over the range [-5, 15].

VI. Estimation Results

A. *Parameter Estimates and Markups*

In order to obtain our welfare estimates, we first need to estimate the γ^g parameters for the expenditure function. The coefficients γ^g for each four-digit HS sector g are obtained by estimating equation (25). Because we ultimately estimate over 1,000 γ^g's, it is not possible to display all of them here. We display the sample statistics for γ^g and $1/\gamma^g$ in table 3. The median γ^g is 0.19 and the average is 12. The large average γ^g is driven by the fact that their distribution is not symmetric and γ^g can take on very large values if goods are relatively homogeneous.[25]

While we do not have strong priors for what a reasonable value of γ^g should be, the markup these parameter estimates imply is consistent with prior studies. To see this fact, we compute the markup for each industry by using equation (15). On the basis of this calculation, the median estimated markup in our data is 0.30 (i.e., a 30 percent markup over marginal costs) in 2005. By comparison, Domowitz, Hubbard, and Petersen (1988) estimate markups across US 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 US absorption. In table 4, we report the share of US absorption from the 10 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^g's. On the basis of this measure, we find that the three sectors in which the products are most heterogeneous and firms are likely to have the most market power are plastics, aircraft,

[24] 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 γ's between 7.8 and 109.9. Moreover, making this change did not qualitatively affect the results because all high γ's imply very small markups and variety effects.

[25] We allow for marginal costs that depend on quantity in our estimation, with the elasticity ω^g. We have found that this formulation results in fewer outlying estimates of the demand parameter. If we constrain ω^g to be zero, then the median estimate of γ^g across sectors is reduced only slightly (from 0.19 to 0.18), but the mean estimate of γ^g increases because of some large, positive outliers.

TABLE 3
PARAMETER ESTIMATES

Statistic	Value	Standard Deviation
	Distribution of γ Estimates	
Mean	11.90	1.75
Median	.19	.01
	Distribution of $1/\gamma$ Estimates	
Mean	8.06	.32
Median	5.27	.34

cell phones, and passenger motor vehicles. In contrast, the most homogeneous sectors in which firms are likely to have the least market power are crude petroleum, natural gas, and cigarettes and cigars. This pattern seems broadly sensible.

B. *Welfare Estimates*

We 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 $c \in \bar{C}$, which we interpret as having small changes in the set of products and their taste parameters for each country (literally, in [18]–[21] we assumed that the set of products and taste parameters for countries $c \in \bar{C}$ do not change at all). As mentioned earlier, in the CES case, Broda and Weinstein (2006) used the set of countries supplying to the United States 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, by country of origin, for exporters and domestic firms selling to the United States. We will suppose that if the

TABLE 4
GAMMA VALUES FROM SECTORS WITH HIGH SHARES OF US DOMESTIC ABSORPTION

HS4	γ	Average Share of Total Absorption
Passenger motor vehicles	.14	.07
Parts and accessories for nonpassenger motor vehicles	.39	.05
Crude petroleum	.76	.04
Automatic data processing machines	.18	.03
Nonmilitary aircraft	.06	.02
Cartons, boxes, cases, bags, and other packing containers	.25	.02
Cell phones	.07	.01
Cigarettes	1.41	.01
Plastics	.05	.01
Natural gas	1.11	.01

Herfindahl index changes by more than some tolerance, that is evidence that the set of products has changed sufficiently to exclude that country from the set $\bar{C}$. For convenience, we refer to this set as "common" countries, which 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 estimates. Because we had to drop the change between 1997 and 1998 because of the transition from SIC to NAICS codes in the United States, we computed the aggregate welfare change between 1992 and 2005 as the sum of the gains from 1992–97, the gain from 1998–2005, and the geometric average annual gain over the two periods. Using the full sample, for our baseline Herfindahl cutoff of ±30 percent we find that the aggregate welfare gain over the time period was 0.85 percent.

The magnitude of this number is perhaps easiest to understand relative to Broda and Weinstein's (2006) estimates for the period 1990–2001. 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 (11 years vs. 13). If we multiply our estimates by 11/13, we find that the 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 gains similar to, but slightly smaller than, those of 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 0.85 percent welfare gain, one-half or 0.42 percent is due to markup declines. The fact that the gain due to new varieties 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,[26] and we are finding that the result applies here even with many new varieties.

We can go further toward understanding the difference between the CES and translog by using equation (20) 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, that is, the fact that

[26] Feenstra and Shiells (1997, 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.

TABLE 5
Markup and Variety Welfare Gains as a Percentage of GDP

Herfindahl Range	W	P	V	V_1	V_2	W_{inf}	P_{inf}
	Baseline Results (Herfindahl Range = ±30%, All Sectors, 1992–2005)						
±30%	.85	.41	.44	.92	−.49	.83	.39
5%–95% confidence interval	(.37, 1.22)	(.28, .66)	(.05, .65)	(.24, 1.72)	(−1.24, −.01)	(.34, 1.19)	(.25, .63)
	Robustness (Differing Herfindahl Ranges, Sectors, and Years)						
±20%	.27	.25	.02	.68	−.66	.24	.22
±40%	.77	.39	.38	.87	−.49	.74	.37
Constant HI	.79	.66	.13	1.67	−1.54	.81	.68
±30%: trimming top/bottom 5%	.63	.41	.22	1.02	−.79	.62	.39
±30%: excluding auto sector	.49	.23	.26	.51	−.25	.46	.20
±30%: 1992–97	.83	.18	.65	.40	.25	.80	.15
±30%: 1998–2005	−.01	.21	−.22	.49	−.71	.00	.22

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

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.

C. *Sensitivity*

We should check that the pro-competitive effect does not depend on our measurement of the total number of goods in each sector, N_t in (19). 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.[27] First, we should compute standard errors because of imprecision in the estimation of γ^g. 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^g estimates, we bootstrapped each of the over 1,000 γ^g's and ω^g's and then used these bootstrapped parameter values to compute the distribution of P_t^g, V_t^g, and total welfare $\Delta \ln W_t^{PV}$. This procedure is computationally intensive, but ultimately we were able to compute 100 estimates of each γ^g and generate 5th–95th percentile confidence bands. 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 $c \in \bar{C}$ involves a trade-off between two opposing forces. Classifying even the small movements in the Herfindahl index as evidence of firm entry and exit will reduce the number of countries in-

[27] A fifth issue is whether the number of years in our sample (13) is enough to avoid the small-sample bias in our estimator noted by Soderbery (2010, 2015) in the CES case. That issue is taken up in online app. B, where from a Monte Carlo analysis we find that 13 years is (just) enough to avoid significant bias.

cluded in the common set $\bar{C}$, 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 movement 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 (21), 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 $\bar{C}$, 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 ±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.[28] 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 consider only the welfare impacts coming from the entry and exit of countries selling to the United States. The results from this calculation are presented in the "constant HI" row of table 5. Overall 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 US 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 the decline in market power arising from the drop in the demand elasticities.

[28] 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 $\bar{C}$ to the excluded countries $c \notin \bar{C}$. We will split our estimation into two subperiods (1992–97 and 1998–2005) because of the change in industry definitions. If we follow the prior CES literature and assume that variety change is measured only 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 is 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.

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 0.85 percent of GDP to 0.63 percent. This finding 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^g 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 carmakers set up new plants (see Blonigen and Soderbery 2010). This entry had two important impacts. First, the US Herfindahl index declined sharply from 0.35 to 0.21, reflecting the large increase in the number of makers operating in the United States. Second, the transplant of Japanese carmakers to the United States was associated with a very large increase in US automobile production: real output of autos made in the United States grew by 41 percent between 1992 and 1998, which contributed to a substantial increase in the share of US consumption made domestically. That increase in the share and falling Herfindahl contribute to a large welfare gain from variety, V_t^g, from (20).

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 multiproduct firms, for example, and in the same way have assumed that the γ^g estimate for autos applies equally well to products across firms as to products within firms. That assumption clearly contradicts the theoretical literature on multiproduct firms, which makes a strong distinction between consumer substitution of products within and between firms (see Allanson and Montagna 2005; Bernard, Redding, and Schott 2011). For this reason we also computed the welfare gains after dropping the passenger vehicle sector, giving 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 the 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 subperiods. Most of total welfare gain W_t appears to have accrued in the first subperiod, between 1992 and 1997. The reason for this finding is that in the second subperiod, 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.

VII. 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 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. 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. Here we have adopted translog preferences, which relax this restriction on markups. 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 US markets, as well as more exporters within those countries, drive 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 US market has been less than the new entry, in the sense that the demand for the typical incumbent firm's output fell. 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 United States 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. Our estimates do not incorporate the efficiency gains due to the self-selection of more efficient firms into exporting, as we have explained, and quantifying that effect is an important area for further research.

Appendix

Proof of (5)

Substituting $\ln p_{it} = [(\alpha_i - s_{it})/\gamma] + \overline{\ln p(t)}$ into (2), and using (3) and $\Sigma_{i=1}^{N}\alpha_i = \Sigma_{i=1}^{N} s_{\text{it}} = 1$, we obtain[29]

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

Proof of (21)

Assuming $\bar{C} = C_{t-1} = C_t$, the shares $\bar{s}_{ct}$ become s_{ct}, so we can rewrite V_t in (20) as

$$V_t = -\sum_{c\in\bar{C}}\frac{1}{2\gamma}(s_{ct-1} + s_{ct})\Delta(H_{ct}s_{ct}) + \frac{1}{2\gamma}\Delta\left(\sum_{c\in\bar{C}}H_{ct}s_{ct}^2\right).$$

Note that a 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 this result, we can simplify V_t as

$$\begin{aligned}
V_t &= -\frac{1}{2\gamma}\sum_{c\in\bar{C}}(s_{ct-1} + s_{ct})\left[\frac{1}{2}(H_{ct-1} + H_{ct})\Delta s_{ct} + \frac{1}{2}(s_{ct-1} + s_{ct})\Delta H_{ct}\right] \\
&\quad + \frac{1}{2\gamma}\sum_{c\in\bar{C}}\left[\frac{1}{2}(H_{ct-1} + H_{ct})\Delta s_{ct}^2 + \frac{1}{2}(s_{ct-1}^2 + s_{ct}^2)\Delta H_{ct}\right] \\
&= -\frac{1}{2\gamma}\sum_{c\in\bar{C}}s_{ct-1}s_{ct}\Delta H_{ct}.
\end{aligned}$$

[29] We thank Sam Kortum for this derivation.

References

Allanson, Paul, and Catia Montagna. 2005. "Multiproduct Firms and Market Structure: An Explorative Application to the Product Life Cycle." *Internat. J. Indus. Org.* 23:587–97.

Arkolakis, Konstantinos, Arnaud Costinot, David Donaldson, and Andres Rodriguez-Clare. 2015. "The Elusive Pro-competitive Effects of Trade." Working Paper no. 21370, NBER, Cambridge, MA.

Arkolakis, Konstantinos, Arnaud Costinot, and Andres Rodriguez-Clare. 2012. "New Trade Models, Same Old Gains?" *A.E.R.* 102 (1): 94–130.

Atkin, David, Benjamin Faber, and Marco Gonzalez-Navarro. 2015. "Retail Globalization and Household Welfare: Evidence from Mexico." Working Paper no. 21176, NBER, Cambridge, MA.

Badinger, Harald. 2007a. "Has the EU's Single Market Programme Fostered Competition? Testing for a Decrease in Markup Ratios in EU Industries." *Oxford Bull. Econ. and Statis.* 69 (4): 497–519.

———. 2007b. "Market Size, Trade, Competition and Productivity: Evidence from OECD Manufacturing Industries." *Appl. Econ.* 39 (11): 2143–57.

———. 2008. "Trade Policy and Productivity." *European Econ. Rev.* 52:867–91.

Behrens, Kristian, Giordano Mion, Yasusada Murata, and Jens Südekum. 2008. "Trade, Wages, and Productivity." Manuscript (September), Inst. Study Labor, Bonn.

Behrens, Kristian, and Yasusada Murata. 2007. "General Equilibrium Models of Monopolistic Competition: A New Approach." *J. Econ. Theory* 13:776–87.

Bergin, Paul, and Robert C. Feenstra. 2009. "Pass-Through of Exchange Rates and Competition between Fixers and Floaters." *J. 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." *Rev. Econ. and Statis.* 71 (February): 143–53.

Bernard, Andrew B., Stephen Redding, and Peter K. Schott. 2011. "Multi-product Firms and Trade Liberalization." *Q.J.E.* 126 (3): 1271–1318.

Berry, Steven T. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *Rand J. Econ.* 25:242–46.

Bilbiie, Florin, Fabio Ghironi, and Marc J. Melitz. 2012. "Endogenous Entry, Product Variety, and Business Cycles." *J.P.E.* 120 (2): 304–45.

Blonigen, Bruce A., Benjamin H. Liebman, and Wesley W. Wilson. 2007. "Trade Policy and Market Power: The Case of the U.S. Steel Industry." Working Paper no. 13761, NBER, Cambridge, MA.

Blonigen, Bruce A., and Anson Soderbery. 2010. "Measuring the Benefits of Product Variety with an Accurate Variety Set." *J. Internat. Econ.* 82 (2): 168–80.

Bottasso, Anna, and Alessandro Sembenelli. 2001. "Market Power, Productivity and the EU Single Market Program: Evidence from a Panel of Italian Firms." *European Econ. Rev.* 45:167–86.

Bresnahan, Timothy F. 1989. "Empirical Studies of Industries with Market Power." In *Handbook of Industrial Organization*, vol. 2, edited by Richard Schmalensee and Robert Willig, 1011–57. Amsterdam: North-Holland.

Broda, Christian, and David E. Weinstein. 2006. "Globalization and the Gains from Variety." *Q.J.E.* 121 (May): 541–85.

Christensen, Laurits R., Dale W. Jorgenson, and Lawrence J. Lau. 1975. "Transcendental Logarithmic Utility Functions." *A.E.R.* 65:367–83.

De Blas, Beatriz, and Katheryn Niles Russ. 2015. "Understanding Markups in the Open Economy." *American Econ. J.: Macroeconomics* 7(2): 157–80.

De Loecker, Jan, Pinelopi K. Goldberg, Amit K. Khandelwal, and Nina Pavcnik. 2016. "Prices, Markups and Trade Reform." *Econometrica* 84 (2): 445–510.

Diewert, W. Erwin. 1976. "Exact and Superlative Index Numbers." *J. Econometrics* 4:115–45.

Domowitz, Ian, R. Glenn Hubbard, and Bruce C. Petersen. 1988. "Market Structure and Cyclical Fluctuations in U.S. Manufacturing." *Rev. Econ. and Statis.* 70 (February): 55–66.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu. 2015. "Competition, Markups, and the Gains from International Trade." *A.E.R.* 105 (10): 3183–3221.

Fajgelbaum, Pablo D., and Amit Khandelwal. 2016. "Measuring the Unequal Gains from Trade." *Q.J.E.* 131 (3): 1113–80.

Feenstra, Robert C. 1994. "New Product Varieties and the Measurement of International Prices." *A.E.R.* 84 (March): 157–77.

———. 2014. "Restoring the Product Variety and Pro-competitive Gains from Trade with Heterogeneous Firms." Working Paper no. 19833, NBER, Cambridge, MA.

Feenstra, Robert C., and Clinton Shiells. 1997. "Bias in U.S. Import Prices and Demand." In *The Economics of New Goods*, edited by Timothy F. Bresnahan and Robert J. Gordon, 249–76. Chicago: Univ. Chicago Press (for NBER).

Harrison, Ann E. 1994. "Productivity, Imperfect Competition and Trade Reform: Theory and Evidence." *J. Internat. Econ.* 36 (February): 53–73.

Head, Keith C., and John Ries. 1999. "Rationalization Effects of Tariff Reductions." *J. Internat. Econ.* 47 (April): 295–320.

———. 2001. "Increasing Returns versus National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade." *A.E.R.* 91 (September): 858–76.

Kim, Euysung. 2000. "Trade Liberalization and Productivity Growth in Korean Manufacturing Industries: Price Protection, Market Power, and Scale Efficiency." *J. Development Econ.* 62:55–83.

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." *Rev. Econ. and Statis.* 87:124–34.

Konings, Jozef, and Hylke Vandenbussche. 2005. "Antidumping Protection and Markups of Domestic Firms." *J. Internat. Econ.* 65:151–65.

Krishna, Pravin, and Devashish Mitra. 1998. "Trade Liberalization, Market Discipline and Productivity Growth: New Evidence from India." *J. Internat. Econ.* 56:447–62.

Krugman, Paul R. 1979. "Increasing Returns, Monopolistic Competition, and International Trade." *J. Internat. Econ.* 9:469–79.

Levinsohn, James. 1993. "Testing the Imports-as-Market-Discipline Hypothesis." *J. Internat. Econ.* 35 (August): 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." *Rev. Econ. Studies* 75 (1): 295–316.

Rigobon, Roberto. 2003. "Identification through Heteroskedasticity." *Rev. Econ. and Statis.* 85 (July): 777–92.

Rodriguez-Lopez, Antonio. 2011. "Prices and Exchange Rates: A Theory of Disconnect." *Rev. Econ. Studies* 78 (July): 1135–77.

Simonovska, Ina. 2015. "Income Differences and Prices of Tradables." *Rev. Econ. Studies* 82 (4): 1612–56.

Soderbery, Anson. 2010. "Investigating the Asymptotic Properties of Import Elasticity Estimates." *Econ. Letters* 109 (2): 57–62.

———. 2015. "Estimating Import Supply and Demand Elasticities: Analysis and Implications." *J. Internat. Econ.* 96 (May): 1–17.

Trefler, Daniel. 2004. "The Long and Short of the Canada-U.S. Free Trade Agreement." *A.E.R.* 94 (September): 870–95.

Tybout, James R., Jaime de Melo, and Vittorio Corbo. 1991. "The Effects of Trade Reform on Scale and Technical Efficiency: New Evidence from Chile." *J. Internat. Econ.* 31 (3/4): 231–50.

Tybout, James R., and M. Daniel Westbrook. 1995. "Trade Liberalization and Dimensions of Efficiency Change in Mexican Manufacturing Industries." *J. Internat. Econ.* 39 (1/2): 53–78.

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

Yilmazkuday, Hakan. 2013. "Markups versus Trade Costs." *SSRN Electronic J.* 01/2013. doi:10.2139/ssrn.2144037.