The Home Market Effect and Bilateral Trade Patterns

By Gordon H. Hanson and Chong Xiang*

Abstract. We develop a monopolistic-competition model of trade with many industries to examine how home-market effects vary with industry characteristics. Industries with high transport costs and more differentiated products tend to be more concentrated in large countries than industries with low transport costs and less differentiated products. We test this prediction using a difference-in-difference gravity specification that controls for import tariffs, importing-country remoteness, home bias in demand, and the tendency for large countries to export more of all goods. We find strong evidence of home-market effects whose intensity varies across industries in a manner consistent with theory. (*JEL* F1, R1)

Much recent theoretical work in international trade is based on imperfect competition and increasing returns to scale. This includes models of intra-industry trade (Paul Krugman, 1979, 1981; Elhanan Helpman, 1981), multinational firms (Helpman, 1984; James R. Markusen, 1984), and economic geography (Krugman, 1991; Anthony J. Venables, 1996). These efforts have produced compelling explanations for why similar countries may gain from trade, why foreign direct investment tends to flow between rich countries, and why manufacturing activity tends to agglomerate spatially within countries.

For purposes of empirical work, however, these new trade theories pose a challenge. Their predictions for trade flows are similar to those of models based on comparative advantage (Helpman and Krugman, 1985; Donald R. Davis, 1995). This complicates testing trade theory (Helpman, 1999) and may account for why attempts to estimate the importance of imperfect competition and increasing returns for trade have yielded mixed results (Helpman, 1987; David Hummels and James Levinsohn, 1995; Peter Debaere, 2001).[1]

Recently, empirical researchers have begun to test new trade theories by exploiting the home-market effect, as derived by Krugman (1980). Standard treatments of the home-market effect (e.g., Helpman and Krugman, 1985; Robert C. Feenstra, 2003) are based on a monopolistic-competition model of trade that has one homogeneous-good industry (with zero fixed costs and zero transport costs) and one differentiated-product industry (with positive fixed costs and positive transport costs). The home-market effect is the tendency for the differentiated-product industry to concentrate in the large country, making it a net exporter of these goods.[2] The logic of the result is that fixed costs induce firms manufacturing differentiated products to locate their operations in a single country and transport costs make the economy with the larger market the optimal site. The homogeneous good is produced by the smaller country. The home-market effect implies a link between a country's market size and its exports that doesn't exist in trade models that are based solely on comparative advantage.

One approach to identify home-market effects uses the correlation between industry supply and industry demand across countries. In Krugman (1980), the demand for individual goods varies across markets because of differences in consumer preferences (e.g., German consumers prefer beer, French consumers prefer wine), leading production of a good to concentrate in markets with high levels of demand. Davis and David E. Weinstein (1999, 2003) find that for manufacturing industries in either OECD countries or Japanese regions industry production increases more than one-for-one with local demand for a good. Keith Head and John Ries (2001) find evidence of similar patterns of industry production and consumption in Canada and the United States. Both sets of results are interpreted as evidence of home-market effects. One potential problem with this approach is that it relies on industry demand shocks being uncorrelated with industry supply shocks. If this condition fails, estimation results may be inconsistent.

A second approach to estimate home-market effects is to examine how the income elasticity of exports varies across goods. In theory, the elasticity of export supply with respect to national income should be higher for differentiated products (James E. Rauch, 1999). Feenstra, Markusen, and Andrew K. Rose (1998) estimate gravity models of bilateral trade flows for an aggregate of differentiated products (which are primarily manufacturing industries) and for an aggregate of homogeneous products (which are primarily non-manufacturing industries). They find that the elasticity of export supply with respect to exporter GDP is higher in the former sample than in the latter, which they take as evidence of home-market effects.[3] One issue with the gravity model, raised by James E. Anderson and Eric van Wincoop (2003), is that some specifications do not control for the 'remoteness' of importing countries and so may be subject to omitted-variable bias.

Empirical work on home-market effects is based on a two-industry model, which is the standard setting in which these effects are derived. The industry presumed subject to home-market effects is manufacturing. In testing for home-market effects Davis and Weinstein (2003) pool across manufacturing industries,[4] and Feenstra, Markusen, and Rose (1998) aggregate over a large subset of manufacturing industries. In actuality, we know manufacturing is a composite of many differentiated-product industries. Do we expect that *all* of these industries will display home-market effects? Unfortunately, there is little theoretical work on this question.[5] But intuition would suggest that the strength of home-market effects will vary with industry characteristics, such as transport costs and the extent of production differentiation. If this is the case, then it is difficult to say whether existing empirical results are consistent with theory.

In this paper, we extend the monopolistic-competition model of trade to allow for a continuum of differentiated-product industries. In section two, we show that industries with high transport costs and low substitution elasticities (i.e., more product differentiation) tend to concentrate in the larger country and industries with low transport costs and high substitution elasticities (i.e., less product differentiation) tend to concentrate in the smaller country. This suggests that within manufacturing home-market effects will vary in a manner that is systematically related to industry characteristics.

Then, we develop a framework to test the predictions of our model based on a "difference-in-difference" gravity specification. This approach, presented in section three, has three steps: (i) we select pairs of countries that are likely to face common trade barriers in markets to which they export, (ii) we select two groups of industries, one with high transport costs and low substitution elasticities (high product differentiation) and one with low transport costs and high substitution elasticities (low product differentiation), and (iii) we examine whether, across exporter pairs, larger countries tend to have higher exports of high-transport cost, more-differentiated goods relative to their exports of low-transport cost, less-differentiated goods.

To see the value of this difference-in-difference gravity specification, consider exports by Belgium and Germany of steel (high transport cost, low substitution elasticity) and radios (low transport cost, high substitution elasticity). Our test of the home-market effect is, in essence, to see whether German-Belgian relative exports of steel are larger than German-Belgian relative exports of radios. The first "difference"–log German steel exports minus log Belgian steel exports–sweeps out of the estimation determinants of bilateral exports that are *specific to the importer* (e.g., importer remoteness, importer GDP, importer industry tariffs). This addresses the

Anderson-Van Wincoop (2003) criticism of the gravity model. By then subtracting off the second "difference"–log German radio exports minus log Belgian radio exports–we control for the tendency of larger countries to export more of all goods. The importance of this second difference is that, in a many industry world, theory has strong predictions about which industries will be *relatively* more concentrated in the larger country, but only weak predictions about which industries will concentrate in the larger country in *absolute* terms.[6]

To preview the empirical results, presented in section four, there is strong evidence of home-market effects. Larger countries have high exports of high-transport cost, more-differentiated goods relative to their exports of low-transport cost, less-differentiated goods. Importantly, we identify these effects only in the difference-in-difference gravity specification. When we look at industries individually – rather than match high-transport cost, more-differentiated goods against low-transport cost, less-differentiated goods – larger countries have higher exports of *all* goods, suggesting naively that all industries exhibit home-market effects. And when we pair industries randomly we find no evidence of home-market effects.

I. Theory and Empirical Specification

In this section, we extend the standard model of trade under monopolistic competition (Helpman and Krugman, 1985; Feenstra, 2003) to the case of a continuum of differentiated-product industries and then use the results from this framework to develop an estimation strategy for identifying home-market effects.

A. Home-Market Effects in a Many Industry World

There is a large country and a small country. Each has one factor of production, labor. The small country's labor endowment and wage are normalized to 1. The large country has labor endowment, $L > 1$, wage, w, and income, $Y = wL$.

There is a continuum of Dixit-Stiglitz monopolistically competitive industries indexed by $z \in [0, 1]$. Consumers have identical Cobb-Douglas preferences, where $\alpha(z)$ is the consumption share on industry z and $\int_0^1 \alpha(z)dz = 1$. For industry z, let $n(z)$ denote the number of product varieties, $\sigma(z)$ denote the elasticity of substitution between varieties, and $\tau(z) > 1$ denote the iceberg transport cost incurred in shipping one unit of output from one country to the other. Let $x(z) \equiv \tau(z)^{\sigma(z)-1}$ denote the effective trade cost for industry z. As is standard in models of monopolistic competition, the importance of trade costs in an industry depends on the degree of product differentiation, as captured by the elasticity of substitution (Richard Baldwin et al., 2003; Head and Theiry Mayer, 2003). We will assume that there is no international specialization at the industry level (although there is specialization at the product-variety level),[7] and that there are many industries with high values of σ and low values of x, and vice-versa. This latter assumption is necessary for there to be variation in the degree of production differentiation and effective trade costs across industries.

Given that the varieties of industry z are symmetric, let $q(z)$ denote the output per variety, $p(z)$ denote the price of each variety, and $c(z)$ denote the fixed labor requirement. We normalize the variable labor requirement for each variety to 1. It is well known that the price is a constant mark-up over marginal cost, and that, because free entry drives profits to zero, output is fixed and revenues are proportional to fixed costs:

$$(1) \qquad p(z) = \frac{\sigma(z)}{\sigma(z)-1}w, \quad q(z) = c(z)[\sigma(z) - 1], \quad \text{and} \quad p(z)q(z) = wc(z)\sigma(z)$$

With no sectoral specialization, both countries produce varieties of industry z and we can write the aggregate-shipments-equals-aggregate-expenditures condition as,

(2)
$$npq = \alpha Y\Gamma + \alpha\Gamma^*, \Gamma = \frac{np^{1-\sigma}}{np^{1-\sigma} + n^*(p^*)^{1-\sigma}x^{-1}}$$
$$n^*p^*q = \alpha Y(1-\Gamma) + \alpha(1-\Gamma^*), \Gamma^* = \frac{np^{1-\sigma}}{np^{1-\sigma} + n^*(p^*)^{1-\sigma}x}$$

where we drop the industry index z and use a "*" to denote small-country values.

Since the large country's income equals the sum of sales from all industries:

(3)
$$Y = \int_0^1 n(z)p(z)q(z)dz$$

We can substitute out $p(z)$ and $q(z)$ from (3) using (1), solve for $n(z)$ from (2) and then substitute the solution into (3) to obtain,

(4)
$$0 = \int_0^1 \alpha(z)g(z)dz, \qquad g(z) = [\frac{Y}{x(z)w^{\sigma(z)} - 1} - \frac{w^{\sigma(z)}}{x(z) - w^{\sigma(z)}}]$$

a solution to which always exists. Details on deriving (4) are in an appendix. Once we solve for w using (4), we can solve for all other endogenous variables using (1) and (2).

Lemma 1 Equation (4) has a unique solution of w, such that $1 < w < \min[x(z)]^{1/\sigma(z)}$.
Proof: See appendix.

Thus, wages are higher in the large country.

Following the literature (e.g., Helpman and Krugman, 1985), we define industry z to be subject to a home-market effect if $n(z)/n^*(z) > L$, or if the large country's share of varieties of z produced globally is greater than its share of world factor supplies. For analytical convenience, we state results on the home-market effect not in terms of $n(z)/n^*(z)$ but in terms of $g(z)$. As it turns out, industry z exhibits a home-market effect if and only if $g(z) > 0$. To see this, let $\tilde{n}(z) \equiv n(z)w$ and $h(z) \equiv \tilde{n}(z)/\tilde{n}^*(z)$, in which case we can show that $g(z) = (1+Y)[h(z)/(h(z)+1)$

$-$ $Y/(Y+1)$] (see appendix for details). It follows that $g(z) > 0$ if and only if $n(z)/n^*(z) > L$ (or, equivalently, $h(z) = \tilde{n}(z)/\tilde{n}^*(z) > Y$). Since $h(z)/(h(z)+1)$ represents the large country's share in world shipments of industry z and $Y/(Y+1)$ represents the large country's share in world income, $g(z)$ is increasing in the concentration of industry z in the large country.

In general, the distribution of $g(z)$ across industries will depend on the distributions of $\alpha(z)$, $\sigma(z)$ and $\tau(z)$ across industries. We cannot obtain analytical solutions for the distribution of $g(z)$ without making assumptions about the distributions of these parameters. However, even absent such assumptions, we can compare values of $g(z)$ between two industries, z_0 and z_1. The results we can obtain are in terms of the factors that make some industry, z_1, *relatively* more concentrated in the large country than some other industry, z_0. That general results about home-market effects are limited to such pair wise industry comparisons has important implications for empirical work.

To motivate the results, note that the home-market effect reflects a trade-off between trade costs and production costs. It is clear that $g(z) > 0$ if and only if

(5) $$Y\left[\frac{x(z)}{w^{\sigma(z)}} - 1\right] > [x(z)w^{\sigma(z)} - 1]$$

Consider whether relocating from the small country to the large country might be profitable for a firm. The large country has a higher production cost (represented by $w^{\sigma(z)}$), but producing there offers savings in trade costs (represented by $x(z) = \tau(z)^{\sigma(z)-1}$) associated with having access to a larger market. The left-hand side of (5) summarizes the benefits of relocating to the large country. Meanwhile, the right-hand side of (5) summarizes the costs of relocation since supplying the small country from the large country results in higher trade and production costs (relative to supplying the small country from within its borders). Naturally, an industry will exhibit a home-market effect if and only if the benefits of the relocation outweigh the costs.[8]

For industries with higher trade costs and/or a lower substitution elasticity, relocating from the small to the large country yields larger savings in trade costs and smaller increases in production costs. Thus, high-x and low-σ industries are the ones that we expect to be most likely to show home-market effects.[9] To be rigorous:

(6) $$g(z) > 0 \Leftrightarrow \frac{Y}{w^{\sigma(z)}x(z)-1} > \frac{w^{\sigma(z)}}{x(z)-w^{\sigma(z)}} \Leftrightarrow Y > w^{2\sigma(z)} + \frac{w^{\sigma(z)}}{x(z)}(Y-1)$$

Since the right-hand side of (6) increases with $\sigma(z)$ and decreases with $x(z)$, we obtain:

Proposition 1 If $g(z_0) > (<) \ 0$ for some z_0, then $g(z_1) > (<) \ 0$ for all z_1 such that $x(z_1) \geq (\leq) \ x(z_0)$ and $\sigma(z_1) \leq (\geq) \ \sigma(z_0)$.

Proposition 1 states that if some industry z_0 exhibits a home-market effect, then so will all industries that have higher effective trade costs and lower elasticities of substitution. Furthermore, by looking at how $g(z)$ varies with $x(z)$ and $\sigma(z)$, we obtain the following:

Proposition 2 (a) If $x(z_1) = x(z_0)$ and $\sigma(z_1) < \sigma(z_0)$, then $g(z_1) > g(z_0)$. (b) If $\sigma(z_1) = \sigma(z_0)$ and $x(z_1) > x(z_0)$, then if $g(z_0) < 0$, $g(z_1) > g(z_0)$, and if $g(z_0) > 0$, then $g(z_1) > g(z_0)$ if both z_0 and z_1 are such that

$$1 - \left(\frac{Y[x(z)/w^{\sigma(z)}-1]}{x(z)w^{\sigma(z)}-1}\right)^2 \frac{w^{2\sigma(z)}}{Y} > 0.$$

Proof: See appendix.

Proposition 2 states that (a) for two industries with the same effective trade costs, the industry with the lower substitution elasticity (or the higher transport cost) will exhibit stronger home-market effects, and (b) for two industries with the same substitution elasticity, the industry with the higher effective trade costs will exhibit stronger home-market effects, provided that both industries' effective trade costs are not too high.[10]

Propositions 1 and 2 suggest an approach to identifying home-market effects empirically. Given that $\alpha(z)$, $\tau(z)$ and $\sigma(z)$ vary across industries, it is difficult to characterize the distribution of $g(z)$ across industries and thus to obtain general results for which industries will concentrate in which country. However, if we select high-τ and low-σ industries as "treatment industries" and low-τ and high-σ industries as "control industries", then by Propositions 1 and 2 the treatment industries will be more concentrated in larger countries than the control industries.[11]

To implement this approach, we need to derive precisely how $h(z) = \tilde{n}(z)/\tilde{n}^*(z)$ varies with relative country size. The solution for $h(z)$ is nonlinear in Y and the preference and technology parameters (equivalently, $n(z)/n^*(z)$ is nonlinear in L). For purposes of empirical work, we obtain approximations for $\ln[h(z)]$ that are linear in polynomials of $(Y\text{-}1)$ or $\ln(Y)$. Details on the derivation are in an appendix.

B. Empirical Specification

To specify the model empirically, we move from a world with a continuum of industries, one factor of production, and two countries to a world with a discrete number of industries, many factors and many countries. We assume that iceberg transport costs between countries j and k in industry m, τ_{mjk}, are a function of the distance between j and k, d_{jk}, such that $\tau_{mjk} = d_{jk}^{\gamma_m}$, where $\gamma_m>0$. Consider the demand by country k for varieties of m produced in country j. Given CES preferences and the symmetry of product varieties in preferences and technology, total sales in industry m by country j to country k equal,

$$(7) \qquad S_{mjk} = \alpha_m Y_k n_{mj} \left(\frac{P_{mjk}}{P_{mk}} \right)^{1-\sigma_m}$$

where P_{mjk} is the delivered (c.i.f.) price in country k of a good in industry m produced by country j and P_{mk} is the CES price index for industry m products in country k. Consider the variation in product prices across countries. P_{mjk} can be written as

(8) $$P_{mjk} = P_{mj} t_{mjk} (d_{jk})^{\gamma_m} = \left(\frac{\sigma_m}{\sigma_m - 1} \right) w_{mj} t_{mjk} (d_{jk})^{\gamma_m}$$

where P_{mj} is the f.o.b. price of a product in industry *m* manufactured in country *j,* t_{mjk} is one plus the ad valorem tariff in *k* on imports of *m* from *j,* w_{mj} is unit production cost in industry *m* and country *j,* and the second equality replaces P_{mj} with a markup over marginal cost.[12] We allow w_{mj} to vary across industries within a country because industries might require multiple factors of production and vary in their factor intensities.

To apply the logic of the home-market effect, compare country *j*'s exports of good *m* to country *k* with some other country *h*'s exports of good *m* to country *k.* Combining equations (7) and (8), these relative export sales are expressed as,

(9) $$\frac{S_{mjk}}{S_{mhk}} = \frac{n_{mj}}{n_{mh}} \left(\frac{w_{mj}}{w_{mh}} \right)^{1-\sigma_m} \left(\frac{d_{jk}}{d_{hk}} \right)^{(1-\sigma_m)\gamma_m} = \frac{\tilde{n}_{mj}}{\tilde{n}_{mh}} \left(\frac{w_{mj}}{w_{mh}} \right)^{-\sigma_m} \left(\frac{d_{jk}}{d_{hk}} \right)^{(1-\sigma_m)\gamma_m}$$

where $\tilde{n}_{mj} = n_{mj} w_{mj}$ and we assume countries *h* and *j* have common technology and face common tariffs in country *k.* Expressing sales in relative terms removes the price index and the tariff in country *k* from the expression. Since $\sigma_m > 1$, equation (9) shows that for industry *m* and destination market *k,* country *j*'s exports relative to country *h*'s exports are increasing in the relative number of product varieties produced in *j,* decreasing in relative production costs in *j,* and decreasing in relative distance from *j* to the destination market.

From theory, we do not know whether $\tilde{n}_{mj} / \tilde{n}_{mh}$ will be increasing or decreasing in Y_j/Y_h, the relative market size of the two countries. This relationship will depend in part on the distribution of preference and technology parameters across industries. From Propositions 1 and 2, what we can say is how $\tilde{n}_{mj} / \tilde{n}_{mh}$ will compare to the ratio of product shipments by the two countries *in some other industry o*, $\tilde{n}_{oj} / \tilde{n}_{oh}$, as long as we choose industries *m* and *o* appropriately. In particular, we choose industry m such that it has a low value of σ (high production differentiation) and a high value of γ (high transport costs), and choose industry o such that it has a high value of σ (low production differentiation) and a low value of γ (low transport costs). In what follows, we will refer to industry m as the "treatment" industry and industry o as the "control" industry. The ratio of relative sales of m versus o goods by countries j and h to country k is,

$$\text{(10)} \qquad \frac{S_{mjk} / S_{mhk}}{S_{ojk} / S_{ohk}} = \frac{\tilde{n}_{mj} / \tilde{n}_{mh}}{\tilde{n}_{oj} / \tilde{n}_{oh}} \frac{(w_{mj} / w_{mh})^{-\sigma_m}}{(w_{oj} / w_{oh})^{-\sigma_o}} (d_{jk} / d_{hk})^{(1-\sigma_m)\gamma_m - (1-\sigma_o)\gamma_o}$$

Propositions 1 and 2 suggest that the ratio $(\tilde{n}_{mj} / \tilde{n}_{mh})/(\tilde{n}_{oj} / \tilde{n}_{oh})$ will be increasing in the relative market size of the two countries.[13] In words, for two countries, j and h, the ratio of their relative exports of good m (high transport costs, low substitution elasticity) to their relative exports of good o (low transport costs, high substitution elasticity) will be higher the larger is the market of country j relative to country h.

To search for evidence of home-market effects empirically, we specify equation (10) in log terms using the following regression:

$$\text{(11)} \qquad \ln\left(\frac{S_{mjk} / S_{mhk}}{S_{ojk} / S_{ohk}}\right) = \alpha + \beta f(Y_j / Y_h) + \phi(\mathbf{X}_j - \mathbf{X}_h) + \theta \ln(d_{jk}/d_{hk}) + \varepsilon_{mojkh}$$

where $f()$ is an increasing function, Y_j/Y_h is relative exporter market size, $\mathbf{X}_l$ is a vector of control variables that determine relative production costs for industries m and o in country l, and ε_{mojkh} is an error term. The function $f()$ captures our approximation results on $\ln[\tilde{n}(z)/\tilde{n}^*(z)]$ being linear in polynomials of $(Y\text{-}1)$ or $\ln Y$. We try several alternative functional forms to find the specification of $f()$ that yields the best fit. Our test for home-market effects is whether $\beta > 0$, or whether larger countries export relatively more of high-transport cost, low-substitution elasticity goods.

In the estimation, we use national factor supplies to measure the vector X, which controls for industry production costs. In general equilibrium, national factor supplies map into national factor prices and these factor prices map into industry production costs. The latter mapping is likely to vary with industry factor intensity. We control for this in the estimation by allowing the vector of coefficients, ϕ, to vary across industries. This is clearly a reduced-form treatment of production costs, but one that is necessitated by a lack of detailed cross-national cost data on the industries in our sample.

To summarize, the advantage of the difference-in-difference specification is that (a) it removes from the estimation importing-country tariffs, price indices, and home bias effects, all of which are hard to measure, (b) for exporter pairs with similar production costs, it differences out of the estimation all determinants of relative exports, except relative distance and relative country size, and (c) it allows us to test for home-market effects in a world with many differentiated-product industries.

Estimation of equation (11) requires that we place restrictions on the set of industries and countries included in the sample and that we define the set of regressors. First, we choose pairs of exporting countries that face common trade policy barriers in the countries that import their goods. It is an added advantage if these country pairs have similar production costs, such that comparative advantage plays a small role in determining their relative exports (i.e., in (11), $(\mathbf{X}_j\text{-}\mathbf{X}_h) \cong \mathbf{0}$). We choose exporting country pairs that belong to a common preferential trade area and that have relatively similar average incomes. Second, we identify a set of high-transport cost, low-substitution elasticity industries and a set of low-transport cost, high-substitution elasticity industries. To do so, we use data on freight rates and estimation results on substitution elasticities from the trade literature.

II. Data and Estimation Issues

The data for the estimation come from several sources. For country exports by product, we use the World Trade Database for 1990 (Feenstra, Robert Lipsey and Charles Bowen, 1997). This source gives bilateral trade flows between countries for three- or four-digit SITC revision 2 product classes. At this level of industry classification (chemical fertilizers, woven cotton fabrics, gas turbines) product classes are better seen as industrial sectors than as individual product varieties, as is consistent with our model.

For data on GDP, we use the Penn World Tables. For country characteristics related to industry production costs, we use nonresidential capital per worker from the Penn World Tables, available land supply relative to the population and average education of the adult population from Robert J. Barro and Jong-Wha Lee (2000), and the average wage in low-skill industries (apparel and textiles) from the United Nations Industrial Development Organization (UNIDO) Industrial data base. For distance and other gravity variables (whether countries share a common border, whether countries share a common language), we use data from Jon Haveman (www.eiit.org). Table 1 gives summary statistics on the regression variables.

There are several estimation issues to be addressed. First, we need to select pairs of exporting countries under the constraint that both members of a pair face common trade policy barriers in importing countries. To ensure that exporters have diversified manufacturing industries (and are not specialized in commodities or low-skill goods), we limit the sample of exporters to OECD countries.[14] Within this group, we form country pairs from sets of countries that belong to a preferential trading arrangement of some kind. These include the members of the European Economic Community (now European Union),[15] Canada and the United States (U.S.-Canada Free Trade Area), and New Zealand and Australia (British Commonwealth). This yields a potential number of 107 exporter pairs per importer and industry. Figure 1 shows the cumulative distribution of log relative exporter GDP (in which the larger country of a pair is in the numerator) for all country pairs in the sample. There is considerable variation in country size. For 65 percent of the observations one country is at least 75 log points larger than the other.

Second, we need to choose the set of importers. One might presume that we should include all countries in the importer sample. A problem with this approach is that many small countries have zero imports from many of their bilateral trading partners. In many contexts, having the dependent variable take zero values can be addressed with standard techniques, such as the Tobit. In our case, however, the dependent variable is constructed from four separate export values (since it is a double log difference). Modeling the joint probability that two or more of these values are zero, as would be necessary to employ a Tobit-style estimator, is beyond the scope of this paper. Instead, we limit our sample to the 58 largest importing countries, which in 1990 accounted for 97 percent of world imports of manufacturing goods. Restricting the sample in this way greatly reduces the number of observations with zero export values.[16] Since the theory applies to importers on a case-by-case basis, there is no loss in focusing on large importers. To check the sensitivity of the results to zero values of trade, we also report results using samples of either the 15 largest importers (69 percent of 1990 world imports) or the 7 largest importers (52 percent of 1990 world imports).

Third, we need to identify industries with low transport costs and high substitution elasticities and industries with high transport costs and low substitution elasticities.[17] Since industry transport costs are often unobserved, we estimate these costs using data on freight rates for U.S. imports in Feenstra (1996). We take the implicit U.S. industry freight rate (insurance and freight charges/import value) and regress it on log distance to the origin country, where we allow both intercepts and slopes to vary across industries.[18] We then use the projected industry freight rate from these coefficient estimates (at median distance for our sample of importer-exporter countries) as the transport cost for an industry.[19] For the elasticity of substitution (σ) by industry, we draw on estimates in Hummels (1999), who uses data on bilateral trade flows, import tariffs, and transport costs to estimate a specification similar to (7). While others have estimated σ for all manufacturing industries or for select manufacturing industries (see Feenstra 2003 and Head and Mayer 2003 for surveys), Hummels' provides the only comprehensive estimates of σ by two-digit SITC industry of which we are aware.[20]

For the control group, we select industries with freight costs in the bottom two deciles of the industry distribution of freight costs and with substitution elasticities in the top three deciles of the industry distribution of this variable. For the treatment group, we select industries with freight rates in the top three deciles of the industry distribution of freight rates and with substitution elasticities in the bottom three deciles of the distribution of this variable. We use asymmetric cutoffs for high and low-transport cost industries because estimated freight rates have an asymmetric distribution with a long upper tail and a short lower tail, as seen in the deciles for freight costs shown in Table 3. Using symmetric cutoffs would result in having too many low-transport cost industries or too few high-transport cost industries. This criterion yields the following approximate cutoffs: for the treatment (high-transport cost, low-σ) industries, freight rates greater than 0.10 and substitution elasticities less than 4.5, and, for the control (low-transport cost, high-σ) industries, freight rates less than 0.05 and substitution elasticities greater than 7.5. We also present results for alternative industry-selection criteria.

Figure 2 shows the joint distribution of our estimates of three-digit industry freight rates and Hummels' estimates of two-digit industry substitution elasticities. Three-digit industry identifiers appear for treatment and control industries; other industries are indicated with zeros. Treatment industries occupy the upper-left of the distribution and control industries occupy the lower-right. There appears to be a weak negative relationship between industry freight rates and σ's.

A final estimation issue is how to measure market size. In equation (11), we use relative GDP to capture relative size for a pair of exporters. This is appropriate for a world with two countries, but with many countries, neighborhood effects may also affect industry location (Masahisa Fujita, Krugman, and Venables, 1999). The relative size of two economies may depend not just on their GDPs but also on the GDPs of their neighbors. In theory, the impact of neighborhood effects on the home-market effect is ambiguous and will depend on the distribution of GDP across countries. Consider Belgium and Sweden. In 1990, they had similar GDPs, but Belgium's immediate neighbors had a larger combined GDP than did Sweden's. One possible outcome is that Belgium's large neighbors create high demand for its goods, leading to a greater concentration of high-transport cost, low-σ industries in the country than in Sweden. A second possible outcome is that Belgium's large neighbors offer such attractive markets that they pull high-transport cost, low-σ industries out of Belgium, casting an *agglomeration shadow* over the country. Which outcome obtains will depend on realizations of taste and technology parameters and the distribution of national factor supplies and so is a question that can only be resolved empirically.

In recent literature, neighborhood effects are captured by a market-potential function, in which demand for a country's goods is a function of income in other countries weighted by transport costs to those economies.[21] Following Fujita, Krugman, and Venables (1999), we define the market potential for country i as the distance-weighted sum of GDP in other countries:

(12) $$MP_i = \sum_{l=1}^{J} Y_l d_{li}^{-\lambda}$$

Using market potential for country size is similar in spirit to Davis and Weinstein (2002), who test for home-market effects with a gravity-based measure of industry demand.[22]

We estimate equation (11) by matching industries from the first group in Table 2 to industries from the second group in Table 2. The high-transport cost, low-σ industries are the treatment group that theory suggests will be subject to home market effects; the low-transport cost, high-σ industries are the control group.

III. Estimation Results

A. Preliminary Results

Before we present the main estimation results, it is useful to consider a simple specification in which the dependent variable is for a pair of exporters log relative exports of a good. This is equivalent to taking the numerator of the regressand in equation (11) as the dependent variable and keeping the same independent variables. We want to see whether the results of this 'single-difference' specification are consistent with standard gravity models, which show an elasticity of bilateral exporters with respect to exporter GDP of about one. The results will also reveal whether the coefficient on relative exporter size is larger for treatment (high-transport cost, low-σ) industries than for control (low-transport cost, high-σ) industries, as would be consistent with home-market effects.

Table 4 shows single-difference gravity estimation results for the 21 treatment and 13 control industries in our sample. The regressors are, for a pair of exporters, log relative exporter size, dummy variables for whether an exporter and importer share a common border or a common language (in level differences for an exporter pair), log relative capital per worker, log relative land area per capita, log relative average schooling, and log relative wages in low-skill industries.[23] The variable of interest is log relative exporter size, which we measure as log relative exporter GDP.

Coefficients on relative exporter GDP are all positive and precisely estimated. Large countries export more of all kinds of goods, both those with high transport costs and those with low transport costs. This is consistent with results for standard gravity models, in which bilateral exports are increasing in exporter income.

More illuminating is to compare results on log relative exporter size for treatment and control industries. The average coefficient on exporter GDP is 1.48 for the treatment industries and 1.06 for the control industries. This is suggestive of home-market effects: larger exporters have high exports of high-transport cost, low-σ goods relative to their exports of low-transport cost, high-σ goods.

B. Pooled Industry Sample: Main Results

Table 5 shows estimation results for equation (11). The dependent variable is, for two countries, log relative exports of a treatment (high-transport cost, low-σ) industry minus log relative exports of a control (low-transport cost, high-σ) industry. We begin by pooling data across the 273 (21x13) treatment and control industry matches in the data. Pooling imposes the unrealistic assumption that coefficients are constant across industry matches, but is useful for gauging overall support for home-market effects. Below, we relax this restriction. The sample for each industry match is exports by 107 country pairs to 58 large importing countries. Each regression includes dummy variables for the industry match and adjusts standard errors to allow for correlation in the disturbances across observations of the same exporter pair.

The specifications in Table 5 differ in terms of their functional form. In Section I.A (and the appendix), we show that we can approximate for the log relative number of product varieties produced by two countries, the term $(\tilde{n}_{mj} / \tilde{n}_{mh}) / (\tilde{n}_{oj} / \tilde{n}_{oh})$ in equation (10), with an increasing function that is linear in polynomials of (Y_j/Y_h -1) or ln(Y_j/Y_h). This approximation result gives us the specification in equation (11), in which relative exports are a function of relative exporter size. In Table 5, we experiment with functional forms for *f*() in equation (11). Column 1 includes ln(Y_j/Y_h) and its square as regressors, column 2 includes (Y_j/Y_h -1) and its square as regressors, and columns 3 and 4 include ln(Y_j/Y_h) or (Y_j/Y_h -1) alone.

In all regressions, relative exports are increasing in relative exporter GDP. This implies that larger countries export more of high-transport cost, low-σ goods relative to their exports of low-transport cost, high-σ goods and is consistent with a home-market effect as stated in Propositions 1 and 2. The square terms on relative exporter size are statistically insignificant (as are higher order polynomials) and we drop them in later regressions. The coefficient estimate on ln(Y_j/Y_h) is precisely estimated in all specifications and the coefficient estimates on ($Y_j/Y_\mathrm{h} - 1$) is precisely estimated when its square is excluded as a regressor. Since the specification with log relative exporter GDP is closest to the standard gravity model, we adopt column 3 as our preferred specification. In this specification, log relative exporter GDP has a coefficient estimate of 0.42. This implies that if one exporter is 10 percent larger than another exporter, then the larger country will on average have export shipments of high-transport cost, low-σ goods that are 4.2 percent higher than the shipments of the smaller country, where these values are normalized by the two countries' relative shipments of low-transport cost, high-σ goods.

Coefficient estimates on other regressors are consistent with results from gravity model estimation. Relative exports of high-transport cost, low-σ goods are decreasing in relative distance and higher for neighboring countries. This suggests, quite sensibly, that exports of high-transport goods are more sensitive to distance than are exports of low-transport goods. That coefficients on log relative capital stocks, log relative average education levels, and log relative land area are positive suggests that relative exports of high-transport cost, low-σ goods are also higher for countries that are more abundant in physical capital, human capital, and/or land. This makes sense if low-σ goods (e.g., iron, steel, non-metallic minerals) tend to be intensive in the use of capital relative to high-σ goods (e.g., consumer electronics, machinery).

C. Pooled Industry Sample: Additional Results

In Table 6, we continue with the pooled sample of industry matches and experiment with adding additional regressors, changing the sample of importer countries, and redefining the selection criterion for treatment and control industries. Column 1 includes the market potential variable as a regressor (see note 22). While relative GDP remains positive and precisely estimated, market potential is negative and precisely estimated. This is consistent with an agglomeration shadow: all else equal, home-market effects appear to be weaker in countries that have larger neighbors. Large countries may pull high-transport cost, low-σ industries out of their smaller neighbors, leaving such industries less concentrated in these countries than in other small countries. However, unreported results suggest that the results on market potential are sensitive to which industries are included in the sample. If we drop the 5 treatment industries in SITC 66 (non-metallic minerals), the coefficient on market potential rises from -1.02 to -0.69 and is no longer statistically significant (the coefficient on log relative exporter GDP falls slightly to 0.38 and remains highly significant). While there is some evidence of an agglomeration shadow, it appears to be confined to a subset of the treatment industries.

Columns 2 and 3 in Table 6 restrict the sample of importing countries to be either the 15 or 7 largest importers. In either case, log relative exporter GDP remains positive and statistically significant, with coefficient estimates of 0.42 to 0.49. These samples contain fewer zero trade values than the sample of 58 large importers. That results are similar for these different samples of importers suggests that zero industry trade values are not having excessive influence on the results.

Columns 4-6 in Table 6 change the selection criterion for treatment and control industries. The original treatment group was industries with freight rates in the top 30 percent and substitution elasticities in the bottom 30 percent. We now impose a more-restrictive selection criterion for the treatment industries, requiring freight rates to be in the top 20 percent and substitution elasticities in the bottom 15 percent. This leaves 9 treatment industries (666, 678, 625, 676, 677, 672, 673, 661, 662). The original control group was industries with freight rates in the bottom 20 percent and substitution elasticities in the top 30 percent. We also impose a more-restrictive selection criterion for the control industries, requiring freight rates to be in the bottom 15 percent and substitution elasticities in the top 15 percent (where we again use asymmetric cutoffs on freight rates for treatment and control industries). This leaves 6 control industries (541, 752, 761, 764, 762, 759). In column 4, we match the more-restrictive groups of treatment and control industries. For comparison, in column 5 we match the more-restrictive treatment industries to the original control industries and in column 6 we match the original treatment industries to the more-restrictive control industries. In all three regressions, the coefficient on log relative exporter GDP is positive and precisely estimated, ranging in value from 0.43 to 0.46. For the pooled industry sample, the results on home-market effects are unaffected by making the industry-selection criterion more restrictive.

So far, we have seen that the evidence of home-market effects is robust to changes in specification, the sample of importers, and the industry-selection criterion. The coefficient estimate on log relative exporter GDP is positive for industry matches in which theory suggests it *should* be positive. By extension, the logic of Propositions 1 and 2 suggests that were we to estimate equation (11) on a sample of randomly matched industries, the coefficient on relative exporter GDP should be zero. To verify this, we perform an additional exercise, which is reported in the final two columns of Table 6. First, we construct a data set of all 561 matches of the 34 industries (21 treatment, 13 control) in our sample. Second, we randomize which industry is the "treatment" industry (industry *m,* which is in the numerator of the dependent variable in equation (11)) and which industry is the "control" industry (industry *o,* which is in the denominator of the dependent variable in (11)). Third, we estimate equation (11) on this sample of industry matches with randomly selected treatment industries. Fourth, we repeat steps two and three 1,000 times (yielding 1,000 coefficient estimates on relative exporter GDP).

Table 6 reports the mean and standard error of coefficient estimates on relative exporter GDP for these 1,000 regressions. Column 7 shows results for 1,000 repetitions on all industry matches. Column 8 shows results where we only match original treatment industries to original control industries and randomize which industry is the "treatment" industry (i.e., which industry is in the denominator of the dependent variable). In either case, the mean coefficient estimate on relative exporter GDP is effectively zero (less than 0.00003) with relatively large standard errors (greater than 0.00006). For randomly matched industry pairs, we find no evidence of home-market effects.

To summarize the results of this section, for a pooled sample of treatment and control industry matches we find strong evidence of home-market effects. For a range of specifications of equation (11), log relative exporter GDP is positive and precisely estimated. Larger countries have higher exports of high-transport cost, low-σ industries than they do of low-transport cost, high-σ industries.

D. Industry-by-Industry Samples

We also estimate equation (9) separately for each of the 273 (21x13) treatment and control industry matches in the data;[24] i.e. we allow the coefficients to vary across industry matches. The dependent variable remains, for a pair of countries, log relative exports of a treatment industry minus log relative exports of a control industry and the independent variables are as in Tables 4-6. For expositional ease we summarize in Table 7 the coefficient estimates on log relative exporter GDP for all the 273 regressions as well as for the 4 subsets of regressions with more- and less-restrictive treatment industries matched to more- and less-restrictive control industries. The detailed results for each industry match are available upon request. The sample is exports by 107 country pairs to the 58 largest importing countries. Results using the sample of either the 7 or 15 large importers are very similar to those we report below.

Of the 273 regressions the coefficient on relative exporter GDP is positive in 234 (86 percent) of the cases and positive and statistically significant at the 10 percent level in 159 (58 percent) of the cases. Sample sizes for the industry-by-industry regressions are much smaller than those for the pooled industry regressions in Tables 5 and 6, and there is a corresponding loss in the precision of coefficient estimates. Among the 4 subsets of regressions, evidence for home-market effects is strongest for the matches of more-restrictive treatment and control industries. These are the matches for which the theoretical case for home-market effects is clearest. For these more-restrictive industry matches, the coefficient estimate on relative exporter GDP is positive in all cases and statistically significant at the 10 percent level in 74 percent of the cases. Results are similar for matches between less-restrictive treatment industries and more-restrictive control industries. Coefficient estimates on relative exporter GDP are positive in 99 percent of cases and statistically significant at the 10 percent level in 75 percent of cases.

Evidence of home-market effects is weakest for matches involving the less-restrictive control industries. Of these industries, there are two, Camera Supplies (SITC 882) and Watches and Clocks (SITC 885), for which there is little or no evidence of home-market effects. Matches involving these two industries account for nearly all of the negative coefficient estimates on relative exporter GDP in the 273 regressions. In the remaining industries, evidence of home-market effects is much stronger. Table 8 summarizes results for these remaining 231 (21x11) industry matches. Of these matches, the coefficient estimate on relative exporter GDP is positive in 98 percent of cases and positive and statistically significant in 68 percent of cases.

E. Discussion

Previous empirical literature tests for home-market effects either by (a) comparing gravity-model estimation results for aggregates of differentiated- and homogeneous-product industries, or (b) estimating the cross-country correlation between the supply of industry output and the demand for industry output. The distinctive feature of our approach is that we compare the correlation between relative exports and relative country size in explicitly chosen industry pairs, as is consistent with our theoretical results on home-market effects in a many-industry world. To see how these empirical approaches differ, we compare our results with representative results from previous literature.

Our approach is similar in spirit to Feenstra, Markusen, and Rose (1998), which falls under approach (a). They find that for the subset of manufacturing identified by Rauch (1999) as differentiated-product industries, the elasticity of bilateral exports with respect to exporter GDP is 1.07 (in 1990 for a sample of OECD countries). For homogeneous goods, which primarily include non-manufacturing industries, this elasticity is 0.38. To compare their results to ours, take the difference of these two coefficient estimates, which approximates a single-difference gravity regression in which the dependent variable is relative bilateral exports (for a single exporter) of differentiated and homogenous products. This single-difference coefficient estimate is 0.69, which is about two-thirds larger than our double-difference estimate of 0.42 for the pooled sample of industries (column (3), Table 5). The Feenstra-Markusen-Rose results are qualitatively similar to ours, though their coefficient estimates on exporter GDP are somewhat larger. In contrast to our approach, their approach does not control for importing-country remoteness or tariffs and cannot detect variation in the strength of home-market effects across differentiated-product industries.

In one example of approach (b), Davis and Weinstein (2002) find evidence consistent with home-market effects for many industries, including food products, textiles, leather, and wood products. These are industries with high-transport costs and large estimated substitution elasticities. Our theoretical model suggests that they are poor candidates for either treatment or control industries. Based on our approach, we would interpret evidence of home-market effects for these industries as at best neutral support for the proposition that increasing returns influence trade patterns. Davis and Weinstein (2003) find evidence inconsistent with home-market effects for other industries, including paper and pulp, chemicals, non-metallic minerals, machinery, and transportation equipment. Of this group, our selection criterion identifies pulp and paper and non-metallic minerals as good candidates for treatment industries and we find evidence of home-market effects in both cases. While a full evaluation of competing approaches to test for home-market effects is beyond the scope of this paper, it is clear that the strategy one takes to identify home-market effects matters for what one finds.

IV. Conclusion

In this paper, we develop a monopolistic competition model of trade with many industries and show the conditions under which home-market effects obtain. We then test the predictions of this model using a difference-in-difference gravity specification.

In our theoretical model, we show that the nature of home-market effects depends on the number of differentiated-product industries. In a world with many such industries, home-market effects take the form of industries with higher transport costs and more differentiated products (lower substitution elasticities) being more concentrated in large countries than industries with lower transport costs and less differentiated products (higher substitution elasticities). In a world with just two industries, the former type of industry concentrates in the larger country in absolute and not just relative terms.

On the basis of our theoretical results, we develop a difference-in-difference gravity specification. We test whether, across country pairs, the larger country tends to have higher exports of goods in "treatment" industries (high-transport costs, low substitution elasticities) than of goods in "control" industries (low-transport costs, high substitution elasticities). The difference-in-difference gravity specification sweeps out of the regression the effects of import tariffs, home bias in demand, importing-country remoteness, and the tendency for larger countries to export more of all goods. Across all treatment and control industry matches, we find strong evidence of home-market effects: the log ratio of relative treatment to control industry exports is increasing in relative exporter GDP. Across individual matches of treatment and control industries we find some variation in the strength of these effects. When we match industries randomly (rather than on the basis of our theoretical selection criterion) the correlation between relative exports and relative exporter GDP shrinks to zero.

Consistent with new trade theory, our results suggest that imperfect competition and increasing returns to scale affect patterns of trade between countries. Country size matters for industrial specialization. Our results identify a set of industry characteristics that predict which industries will be relatively concentrated in which countries. Of additional value, the approach we develop in this paper is widely applicable as a test for home-market effects. It is grounded in general-equilibrium trade theory and, within the confines of the monopolistic competition model, imposes few restrictions on the number of industries or on parameter values. Also, our approach has modest data requirements and can be applied to data sets that are publicly available.

Appendix

A. Derivation of equation (4).

By (1), $p(z)/p^*(z) = w/w^* = w$ (w^* is normalized to 1) and $q(z) = q^*(z)$. Since $\tilde{n}(z) \equiv n(z)w$ and $\tilde{n}^*(z) \equiv n^*(z)w^* = n^*(z)$, (2) can be simplified as:

(A1) $$\tilde{n}(z)c(z)\sigma(z) = \alpha(z)Y\frac{\tilde{n}(z)}{\tilde{n}(z)+\tilde{n}^*(z)w^{\sigma(z)}x(z)^{-1}} + \alpha(z)\frac{\tilde{n}(z)}{\tilde{n}(z)+\tilde{n}^*(z)\cdot w^{\sigma(z)}x(z)}$$

$$\left(\tilde{n}(z)+\tilde{n}^*(z)\right)c(z)\sigma(z) = \alpha(z)(Y+1)$$

Thus,

(A2) $$\tilde{n}(z) = \frac{Yx(z)^2 - w^{\sigma(z)}(1+Y)x(z)+1}{x(z)^2-(w^{\sigma(z)}+w^{-\sigma(z)})x(z)+1}\frac{\alpha(z)}{c(z)\sigma(z)}$$

Plugging (A2) into (3):

$$Y = \int_0^1 \frac{Yx(z)^2 - w^{\sigma(z)}(1+Y)x(z)+1}{x(z)^2-(w^{\sigma(z)}+w^{-\sigma(z)})x(z)+1}\alpha(z)dz$$

$$\Leftrightarrow \int_0^1 \alpha(z)\left(\frac{Yx(z)^2 - w^{\sigma(z)}(1+Y)x(z)+1}{x(z)^2-(w^{\sigma(z)}+w^{-\sigma(z)})x(z)+1} - Y\right)dz = 0$$

$$\Leftrightarrow \int_0^1 \frac{Y(x(z)-w^{\sigma(z)}) - w^{\sigma(z)}(w^{\sigma(z)}x(z)-1)}{(x(z)-w^{\sigma(z)})(x(z)-w^{-\sigma(z)})}\alpha(z)dz = 0$$

$$\Leftrightarrow 0 = \int_0^1 \alpha(z)g(z)dz,\ g(z) = \left[\frac{Y}{x(z)w^{\sigma(z)}-1} - \frac{w^{\sigma(z)}}{x(z)-w^{\sigma(z)}}\right]$$

B. Derivation of $g(z) = (1+Y)[h(z)/(h(z)+1) - Y/(Y+1)]$.

Since, $h(z)/(h(z)+1) = \tilde{n}(z)/(\tilde{n}(z)+\tilde{n}^*(z))$, by (A1) and (A2),

$$(1+Y)\left[\frac{h(z)}{h(z)+1} - \frac{Y}{Y+1}\right]$$

$$= \frac{Yx(z)^2 - w^{\sigma(z)}(1+Y)x(z)+1}{x(z)^2 - (w^{\sigma(z)} + w^{-\sigma(z)})x(z)+1} - Y$$

$$= \frac{Y(x(z)-w^{\sigma(z)}) - w^{\sigma(z)}(w^{\sigma(z)}x(z)-1)}{(x(z)-w^{\sigma(z)})(x(z)-w^{-\sigma(z)})} = g(z).$$

C. Proof of **Lemma 1.**

Denote the right-hand side of (4) by $R(w)$. It is easy to verify that if $R'(w)$ exists, $R'(w) < 0$ for all values of w and Y.

(C1) It is easy to verify that $R(1) > 0$. If $w > 1$ and w rises towards $\min[x(z)]^{1/\sigma(z)}$, $R(w)$ approaches $-\infty$. (Notice that $x(z)w^{\sigma(z)} - 1 > x(z) - w^{\sigma(z)}$ since $w > 1$.) This establishes that (4) always has a solution such that $1 < w < \min[x(z)]^{1/\sigma(z)}$.

(C2) It is easy to verify that $R() > 0$ if $w > \max[x(z)]^{1/\sigma(z)}$. If $\min[x(z)]^{1/\sigma(z)} < w < \max[x(z)]^{1/\sigma(z)}$, $R(w)$ is ill defined since $\exists\, z$ such that $x(z) - w^{\sigma(z)} = 0$. Thus, the solution in (C1) is the unique solution for $w > 1$.

(C3) It is easy to verify that $R(0) = 0$. As w rises from 0 towards $\min[x(z)]^{-1/\sigma(z)}$, $R(w)$ falls towards $-\infty$. When w falls from 1 towards $\max[x(z)]^{-1/\sigma(z)}$, $R(w)$ rises towards $+\infty$. If $\min[x(z)]^{-1/\sigma(z)} < w < \max[x(z)]^{-1/\sigma(z)}$, $R(w)$ is ill defined since $\exists\, z$ such that $x(z)w^{\sigma(z)} - 1 = 0$. Thus no $w < 1$ can solve (4).

D. Proof of **Proposition 2.**

(D1) Since $Y/[x(z)w^{\sigma(z)} - 1]$ decreases with $\sigma(z)$ and $w^{\sigma(z)}/[x(z) - w^{\sigma(z)}]$ increases with $\sigma(z)$, $g(z)$ decreases with $\sigma(z)$. This establishes part (1).

(D2) The statement is true if holding $\sigma(z)$ constant, $g(z)$ increases with $x(z)$. Drop the index z. Let $g_1 \equiv Y/(q_1 w - 1), q_1 \equiv xw^{\sigma-1}, g_2 \equiv 1/(q_2/w - 1), q_2 \equiv xw^{1-\sigma}$. Then $dq_1 = q_1 dx$, $dq_2 = q_2 dx$, and $wq_1 - 1 > q_2/w - 1 > 0$. Normalize dx to 1. Then:

(A3) $$dg = dg_1 - dg_2 = \frac{[dq_2]/w}{(q_2/w-1)^2} - Y\frac{w[dq_1]}{(wq_1-1)^2}$$

$$\propto (wq_1 - 1)^2[dq_2] - Y(q_2 - w)^2[dq_1] = (wq_1 - 1)^2 q_2 - Y(q_2 - w)^2 q_1$$

It is easy to verify that $Y = g_1(wq_1 - 1)$, $1/g_2 = (q_2 - w)/w$ and $w^{2\sigma} - wq_1 = wq_1(w/q_2 - 1)$.

(A4) $$\therefore -(q_2 - w)^2 Y q_1 = q_2(wq_1 - 1)(w^{2\sigma} - wq_1)\frac{g_1}{g_2}$$

Plug (A4) into (A3), we have:

(A5) $$dg \propto (wq_1 - 1)^2 q_2 - Y(q_2 - w)^2 q_1$$

$$= q_2[wq_1 - 1][(wq_1 - 1) + (w^{2\sigma} - wq_1)\frac{g_1}{g_2}]$$

$$\propto \frac{Y}{g_1}[1 - (\frac{g_1}{g_2})^2\frac{w^{2\sigma}}{Y}] \propto 1 - (\frac{Y[x(z)/w^{\sigma(z)} - 1]}{x(z)w^{\sigma(z)} - 1})^2\frac{w^{2\sigma(z)}}{Y}$$

By Proposition 1, $g()$ increases with $x(z)$ around z^* such that $g(z^*) = 0$, and so $dg > 0$ when $g_1 = g_2$, and (A5) is positive. Then if $g(z_0) < 0$ so that $g_1(z_0) < g_2(z_0)$, (A5) remains positive, and $dg > 0$ for z_0. If $g(z_1) > 0$, then clearly $g(z_1) > g(z_0)$; if $g(z_1) < 0$, then $g()$ also increases with $x(z)$ at z_1, and so $g(z_1) > g(z_0)$. On the other hand, if $g(z_0) > 0$ and (A5) is positive for both z_0 and z_1, then $dg > 0$ for both industries, and $g(z_1) > g(z_0)$.

E. Approximation of $h(z)$ as linear in polynomials of $\ln Y$ and Y-1.

Combining equation (4) and the definition of $g(z)$, we obtain

(A6) $$\frac{h(z)}{h(z)+1}=\frac{Y}{Y+1}(1+\frac{1}{x(z)w^{\sigma(z)}-1}+\frac{w^{\sigma(z)}}{x(z)-w^{\sigma(z)}})-\frac{w^{\sigma(z)}}{x(z)-w^{\sigma(z)}}$$

$$=\frac{Y}{Y+1}\frac{x(z)^2-1}{(x(z)-w^{\sigma(z)})(x(z)-w^{-\sigma(z)})}-\frac{w^{\sigma(z)}}{x-w^{\sigma(z)}}$$

If $Y = 1$, then $h(z) = 1$ for all z (i.e., if the 2 countries are of equal size, then they produce the same numbers of varieties for each good). Around $h_0 = 1$, a first-order Taylor approximation of $h(z)/(1+h(z))$ yields $h(z)/(1+h(z)) \approx h_0/(h_0+1) + [h(z) - h_0]/(1 + h_0)^2 = (h(z)+1)/4$. Likewise, around $Y = 1$, $Y/(Y+1) \approx (Y+1)/4$. Plugging these into (A6):

(A7) $h(z) \approx aY + b$

$$a=\frac{x(z)^2-1}{(x(z)-w^{\sigma(z)})(x(z)-w^{-\sigma(z)})},\ b=\frac{x(z)^2-1}{(x(z)-w^{\sigma(z)})(x(z)-w^{-\sigma(z)})}-\frac{4w^{\sigma(z)}}{x-w^{\sigma(z)}}-1$$

Take the log of (A7) and then approximate $\ln(Y+b/a)$ using a first-order Taylor approximation around $Y=1$ so that $\ln(Y+b/a) \approx \ln(a+b)-\ln a+a(Y-1)/(a+b)$. Thus:

(A8) $\ln h(z) \approx \ln(a+b) + \frac{a}{a+b}(Y-1)$

Let $\ln(a+b) = H(Y, z)$ (the dependence of a and b on Y is through w). Then around $Y = 1$, $\ln(a+b) \approx H(1, z) + H'(1,z)(Y-1)$. Since both $H(1, z)$ and $H'(1, z)$ are independent of Y, denote them as constants c_1 and c_2. Thus, $\ln(a+b) \approx c_1 + c_2(Y-1)$. Likewise, $a/(a+b) \approx c_3 + c_4(Y-1)$. Plugging these two expressions into (A8), we have:

(A9) $\ln h(z) \approx c_1 + c_5(Y-1) + c_4(Y-1)^2,\ c_5 = c_2 + c_3$

Since the c's depend on z but not Y, equation (A9) is an approximation of $\ln h(z)$ by the polynomials of $Y - 1$. To use the polynomials of $\ln Y$ instead, note that around $Y = 1$, $\ln Y = \ln(Y - 1 + 1) \approx Y - 1$. Thus (A9) becomes:

(A10) $\ln h(z) \approx c_1 + c_5 \ln Y + c_4(\ln Y)^2,\ c_5 = c_2 + c_3$

Higher-order approximations can be obtained analogously.

References

Amiti, Mary. “Inter-Industry Trade in Manufactures: Does Country Size Matter?” *Journal of International Economics*, April 1998, 44(2), pp. 231-56.

Anderson, James E. and van Wincoop, Eric. “Gravity with Gravitas: A Solution to the Border Puzzle.” National Bureau of Economic Research (Cambridge, MA) Working Paper No. 8079, January 2001.

Barro, Robert J. and Lee, Jong-Wha. “International Data on Educational Attainment Updates and Implications.” National Bureau of Economic Research (Cambridge, MA) Working Paper No. 7911, September 2000.

Baldwin, Richard; Forslid, Rikard; Martin, Philippe; Ottaviano, Gianmarco and Robert-Nicoud, Frederic. *Public Policies and Economic Geography*. Princeton: Princeton University Press, 2003.

Brulhart, Marius and Trionfetti, Federico. “A Test of Trade Theories when Expenditure is Home Biased.” Mimeo, Universite de Lausanne, 2001.

Davis, Donald R. “The Home Market, Trade, and Industrial Structure.” *American Economic Review*, December 1998, 88(5), pp. 1264-76.

Davis, Donald R. and Weinstein, David E. "Economic Geography and Regional Production Structure: An Empirical Investigation." *European Economic Review*, February 1999, 43(2), pp. 379-407.

Davis, Donald R. and David E. Weinstein. "Market Access, Economic Geography and Comparative Advantage: An Empirical Assessment.” *Journal of International Economics*, January 2003, 59(1), pp. 1-24.

Debaere, Peter. “Testing ‘New’ Trade Theory without Gravity: Reinterpreting the Evidence.” Mimeo, University of Texas, 2001.

Dixit, Avinash and Stiglitz, Joseph. “Monopolistic Competition and Optimum Product Diversity.” *American Economic Review*, June 1977, 67(3), pp. 297-308.

Eaton, Jonathan and Tamura, Akiko. “Bilateralism and Regionalism in Japanese-U.S. Trade and Direct Foreign Investment Patterns.” *Journal of the Japanese and International Economies*, December 1994, 8(4), pp. 478-510.

Ellison, Glen and Glaeser, Edward L. “Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach.” *Journal of Political Economy,* October 1997, 105(5), pp. 889-927.

Evenett, Simon J. and Keller, Wolfgang. “On Theories Explaining the Success of the Gravity Equation.” *Journal of Political Economy*, April 2002, 110(2), pp. 281-316.

Feenstra, Robert C. “U.S. Imports, 1972-1994: Data and Concordances.” National Bureau of Economic Research (Cambridge, MA) Working Paper No. 5515, March 1996.

Feenstra, Robert C. *Advanced International Trade: Theory and Evidence*. Princeton: Princeton University Press, 2003.

Feenstra, Robert C.; Lipsey, Robert and Bowen, Charles. “World Trade Flows, 1970-1992, with Production and Tariff Data.” National Bureau of Economic Research (Cambridge, MA) Working Paper No. 5910, January 1997.

Feenstra, Robert C.; Markusen, James R. and Rose, Andrew K. “Understanding the Home Market Effect and the Gravity Equation: The Role of Differentiating Goods.” National Bureau of Economic Research (Cambridge, MA) Working Paper No. 6804, November 1998.

Fujita, Masahisa; Krugman, Paul and Venables, Anthony J. *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge, MA: MIT Press, 1999.

Hanson, Gordon. "Market Potential, Increasing Returns, and Geographic Concentration." Mimeo, University of California at San Diego, 2001.

Hanson, Gordon. "Scale Economies and the Geographic Concentration of Industry," *Journal of Economic Geography*, July 2001, 1(3), pp. 255-76.

Head, Keith and Ries, John. "Increasing Returns versus National Product Differentiation as an Explanation for the Pattern of US-Canada Trade." *American Economic Review*, September 2001, 91(4), pp. 858-76.

Head, Keith, and Mayer, Theiry. "The Empirics of Agglomeration and Trade." Mimeo, University of British Columbia, 2003.

Helpman, Elhanan. "International Trade in the Presence of Product Differentiation, Economies of Scale, and Monopolistic Competition: A Chamberlin-Heckscher-Ohlin Approach." *Journal of International Economics* August 1981, 11(3), pp. 305-40.

Helpman, Elhanan. "A Simple Theory of Trade with Multinational Corporations." *Journal of Political* Economy, June 1984, 92(3), pp. 451-71.

Helpman, Elhanan. "Imperfect Competition and International Trade: Evidence from Fourteen Industrial Countries." *Journal of the Japanese and International Economies*, March 1987, 1(1), pp. 62-81.

Helpman, Elhanan. "The Structure of Foreign Trade." *Journal of Economic Perspectives*, Spring 1999, 13(2), pp. 121-144.

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

Holmes, Thomas J. and Stevens, John J. "The Home Market and the Pattern of Trade: Round Three." Federal Reserve Bank of Minneapolis Staff Report No. 304, 2002.

Hummels, David. "Towards a Geography of Trade Costs." Mimeo, University of Chicago, 1999.

Hummels, David and Levinsohn, James. "Monopolistic Competition and International Trade: Reconsidering the Evidence." *Quarterly Journal of Economics*, August 1995, 110(3), pp. 799-836.

Krugman, Paul. "Increasing Returns, Monopolistic Competition and International Trade." *Journal of International Economics*, November 1979, 9(4), pp. 469-79.

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

Krugman, Paul. "Increasing Returns and Economic Geography." *Journal of Political Economy*, June 1991, 99(3), pp. 483-99.

Krugman, Paul and Venables, Anthony J. "How Robust Is the Home Market Effect?" Mimeo, Massachusetts Institute of Technology and London School of Economics, 1999.

Markusen, James R. "Multinationals, Multi-Plant Economies, and the Gains from Trade." *Journal of International Economics*, May 1984, 16(3-4), pp. 205-26.

Rauch, James E., "Networks versus Markets in International Trade," *Journal of International Economics,* June 1999, 48(1), pp. 7-37.

Redding, Stephen and Venables, Anthony J. "Economic Geography and Global Development." Mimeo, London School of Economics, 2002.

Trionfetti, Federico. "On the Home Market Effect: Theory and Empirical Evidence." Centre for Economic Performance Working Paper No. 987, 1998.

Trionfetti, Federico. "Using Home-Biased Demand to Test Trade Theories."

Weltwirtschaftliches Archiv May 2001, 137(3), pp. 404-26.

Venables, Anthony J. "Equilibrium Locations of Vertically Linked Industries." *International Economic Review*, May 1996, 37(2), pp. 341-60.

Weder, Rolf. "Comparative Home-Market Advantage: An Empirical Analysis of British and American Exports." Mimeo, University of Basel, 1998.

[*]Hanson: IR/PS 0519, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0519 (e-mail: gohanson@ucsd.edu, phone: 858-822-5087, fax: 858-534-3939). Xiang: Purdue University, Economics, KRAN, 403 West State Street, West Lafayette, IN 47907-2056 (e-mail: cxiang@mgmt.purdue.edu, phone: 765-494-4499, fax: 765-496-1778). The authors thank two anonymous referees, Richard Baldwin, John Bound, Alan Deardorff, Robert Feenstra, David Hummels, and seminar participants at the CEPR, the Fall 2002 Midwest International Meetings, the University of Michigan, Purdue University, UCSD, and the University of Florida for valuable suggestions.

[1] On the observational equivalence of scale economies and comparative advantage, see also Glen Ellison and Edward L. Glaeser (1997) and Hanson (2001). On trade and scale economies, see also Simon J. Evenett and Wolfgang Keller (2002).

[2] There is debate about the robustness of the home-market effect in Krugman (1980). Davis (1998) finds that with one differentiated-product industry (with positive fixed costs), one homogeneous-product industry (with zero fixed costs), and identical industry transport costs the home-market effect disappears. Krugman and Venables (1999) counter this result by showing that the home-market effect exists as long as some homogenous goods have low transport costs or some differentiated goods have zero fixed costs.

[3] To be precise, they find that for an aggregate of differentiated-product industries the elasticity of bilateral exports with respect to the exporting country's GDP exceeds the elasticity of bilateral exports with respect to the importing country's GDP, but for homogeneous-product industries the reverse is true.

[4] For regressions on individual industries, they find that about half exhibit a home-market effect.

[5] Mary Amiti (1998) and Krugman and Venables (1999) examine the case of two monopolistically competitive, differentiated-product industries. Thomas J. Holmes and John J. Stevens (2002) depart from the monopolistic-competition framework and allow for a continuum of industries that have common transport costs and varying returns to scale. They find that goods with strong scale economies are subject to a home-market effect and that goods with weak scale economies are non-traded. See Feenstra, Markusen, and Rose (1998) for results on home-market effects in a two-sector, reciprocal dumping model of trade.

[6] An additional advantage of our approach is that by using national income, distance, and other gravity variables as regressors we lessen concerns about simultaneity in the estimation.

[7] This assumption is not as onerous as it may seem. Consider two examples. First, suppose Y is large. It might seem that sectoral specialization would occur as in Helpman and Krugman (1985). However, in their model, non-specialization implies that the relative wage equals one, so that relative goods prices cannot adjust as Y changes. In our model, w is free to adjust as Y changes, so that relative goods prices are also free to adjust and specialization at the sectoral level doesn't occur. Second, suppose industry z_0 has a high effective trade cost. It might seem that exporting z_0 from the large country to the small country is unprofitable, so that z_0 is not traded. However, ice-berg transport costs tend to keep export sales unchanged, so that quantities remain positive even when prices are high. Thus, z_0 will be traded.

[8] This is also the intuition for Lemma 1. Suppose $w \leq 1$. Then relocating to the large country offers not only savings in trade costs but *lower* production costs, in which case every firm has an incentive to relocate. Clearly, this cannot be an equilibrium.

[9] Previous studies have focused on the roles of σ and τ and found the effects of σ on industry concentration in the large country to be non-monotonic: holding τ constant, both very high and very low values of σ work against home-market effects (Krugman and Venables, 1999). This is because σ affects both production and trade costs. The savings in trade costs are too small for sectors with very low σ, and the extra production costs too large for those with very high σ. However, if we hold x constant (i.e., if we vary τ with σ such that x is unchanged) lowering σ always raises industry concentration in the large country.

[10] To see the logic behind (b), suppose σ is the same for all industries, and $x(z_0)$ is very high (so that $x(z_1)$ is even higher). Then the left-hand side of the inequality in Proposition 2 approaches $1 - Y/w^{2\sigma(z_0)} < 0$ and Proposition 2 fails to hold. The intuition for this is that, although industries with very high effective trade costs (such as z_1) are likely to show home-market effects (by Proposition 1), goods in these industries are not much traded, making them likely to be less concentrated in the large country than industries with intermediate values of effective trade costs (such as z_0). See note 11.

[11] By Proposition 1, industries with very high trade costs are likely to show home-market effects, but since goods in these industries are not much traded, they might be less concentrated in the large country than industries with intermediate values of trade costs. In our empirical work, we restrict our control industries to the bottom 20 percent of the industry distribution of freight costs.

[12] For analytical ease, we assume that the markup of price over marginal cost is a multiplicative function of production costs, tariffs, and transport costs.

[13] To be rigorous, Propositions 1 and 2 show that the ratio of relative exports exceeds 1 when country *j* is larger than country *h*. Consider a sequence of experiments, each involving the continuum-good model of Section I.A and a randomly drawn value for relative country size. Record the logs of the relative exports ratios of industry *m* and *o* in all these experiments. These ratios are positive for cases with $\ln(Y_j/Y_h) > 0$ (i.e. $Y_j > Y_h$) and negative for cases with $\ln(Y_j/Y_h) < 0$. A plot of the log of the relative exports ratios against $\ln(Y_j/Y_h)$ would reveal an upward-sloping relationship. Thus, for a regression that is across country pairs for fixed industry pairs such as ours, Propositions 1 and 2 suggest that the relative exports ratio is increasing in relative market size.

[14] Debaere (2002) finds that the gravity model fits better for OECD countries than for non-OECD countries. This suggests that drawing our sample of exporters from rich countries may aid in finding support for home-market effects. Results may be weaker for a sample of poor-country exporters.

[15] The European exporter countries are Austria, Belgium-Luxembourg, Denmark, Finland, France, Germany, Great Britain, Greece, Italy, Ireland, the Netherlands, Norway, Portugal, Spain, and Sweden.

[16] For this sample, 85 percent of the observations have non-zero values for all four components of the dependent variable. To preserve information on zero trade values, we follow Jonathan Eaton and Akiko Tamura (1994) and assume countries with zero bilateral imports of a good actually import minute quantities, which we set to one. The results are unaffected by dropping observations that contain zero trade values from the sample.

[17] Theory would suggest using effective trade costs, $x=\tau^{\sigma-1}$, to select industries. Actual transport costs depend on both unit shipping costs and distance shipped. It is difficult to translate these two dimensions into a scalar measure of τ, as would be necessary to calculate x. Proposition 2 states that when comparing low σ to high σ industries, x should be at least as large in the former as in the latter. To achieve this, we limit low σ industries to those with high freight rates and high σ industries to those with low freight rates.

[18] We allow intercepts to vary across three-digit SITC industries and slopes on distance to vary across two-digit SITC industries. We estimate this regression on four samples of U.S. imports: with U.S. border countries and without U.S. border countries (with two different cutoffs on minimum exporter size). We use the mean of coefficient estimates across these four samples to estimate the freight rate for an industry.

[19] Median distance for the country pairs in our sample is 8008 km. Our results are robust to using instead a distance value of 4004 km to construct estimates of freight rates.

[20] Hummels' (1999) estimate of σ for all manufacturing is similar to other results. We use Hummels' OLS estimates. He also reports IV estimates, which yield higher values of σ for the industries in our sample.

[21] For other applications of market potential see Hanson (2001) and Stephen Redding and Venables (2002).

[22] Using coefficient estimates on the distance variable from the gravity model in Hummels (1999), we set λ equal to 0.92. These estimates of the gravity distance coefficient are similar to results in many other empirical studies. We obtain similar results when using other values of λ. In the specifications in which we control for market potential, we include both log relative exporter GDP and log relative exporter market potential as regressors (where we define exporter market potential excluding own-country GDP).

[23] Since the data include observations on relative income, distance, and other variables for given exporter pairs across many importing countries, we correct the standard errors to allow for correlation in the errors across observations that share the same exporter pair.

[24] Since the regressors do not vary across industries, there is no gain to estimating equation (9) jointly across pairs of treatment and control industries (OLS is just as efficient as GLS).

Table 1: Summary Statistics

Variable	Mean	Standard Deviation
Exports	-0.005	3.978
GDP	0.271	1.562
Market Potential	-0.124	0.400
Distance	0.067	0.500
Common Language	0.029	0.279
Common Border	-0.013	0.192
Capital per Worker	-0.172	0.501
Wage in Low-Skill Industries	-0.046	0.504
Land Area/Population	0.120	1.525
Average Education	-0.065	0.311

Notes: Data are for 1990. Trade data are by country and three-digit SITC industry. All variables are differences in log values for pairs of exporting countries (except for common language and common border, which are level differences in dummy variables). The exporter pairs are Australia-New Zealand, Canada-United States, and all pair wise combinations of the set, Austria, Belgium-Luxembourg, Denmark, Finland, France, Germany, Great Britain, Greece, Italy, Ireland, the Netherlands, Norway, Portugal, Spain, and Sweden. The importing countries are the 58 largest importers.

Table 2: Industry Freight Costs and Substitution Elasticities (σ)

Control Industries: Low Transport Costs, High σ

SITC	Industry	Freight Rate	σ
541	Pharmaceuticals	0.0315	9.53
752	Computers	0.0333	11.02
761	Televisions	0.0364	9.44
884	Optical Lenses	0.0405	8.13
764	Audio Speakers	0.0407	9.44
762	Radios	0.0408	9.44
759	Computer Parts	0.0420	11.02
514	Nitrogen Compounds	0.0475	7.50
881	Cameras	0.0477	8.13
751	Office Machines	0.0481	11.02
882	Camera Supplies	0.0488	8.13
885	Watches and Clocks	0.0490	8.13
726	Printing Machinery	0.0495	8.52

Treatment Industries: High Transport Costs, Low σ

SITC	Industry	Freight Rate	σ
671	Pig Iron	0.1010	3.53
621	Rubber and Plastics	0.1037	3.57
674	Iron Sheets	0.1099	3.53
679	Iron Castings	0.1118	3.53
665	Glassware	0.1119	2.65
663	Mineral Mfg.	0.1135	2.65
666	Pottery	0.1229	2.65
678	Iron Tubes	0.1310	3.53
642	Paper Products	0.1313	4.25
812	Sanitary and Plumbing	0.1317	4.40
625	Tires	0.1321	3.57
676	Steel Rails	0.1368	3.53
641	Paper and Paperboard	0.1368	4.25
677	Iron Wire	0.1380	3.53
672	Iron Ingots	0.1404	3.53
635	Wood Manufacturing	0.1420	3.99
673	Iron Bars	0.1557	3.53
821	Furniture	0.1573	3.64
634	Wood Panels	0.1594	3.99
661	Cement	0.2117	2.65
662	Clay	0.2721	2.65

Table 3: Quantiles for Industry Freight Costs

Percentile	Freight Rate
10	0.0405
20	0.0512
30	0.0624
40	0.0729
50	0.0829
60	0.0922
70	0.1037
80	0.1144
90	0.1318

Notes: See text for details on how freight rates are constructed.

Table 4: Single-Difference Gravity Estimation Results

Low-Transport Cost Industries	Relative GDP	High-Transport Cost Industries	Relative GDP	High-Transport Cost Industries	Relative GDP
541	0.835	**671**	1.505	**677**	1.455
Pharmaceuticals	(5.32)	Pig Iron	(7.73)	Iron Wire	(8.31)
752	1.097	**621**	1.411	**672**	1.360
Computers	(6.42)	Rubber, Plastics	(12.55)	Iron Ingots	(21.26)
761	1.155	**674**	2.390	**635**	1.174
Televisions	(12.68)	Iron Sheets	(11.75)	Wood Mfg.	(22.14)
884	1.217	**679**	1.432	**673**	1.841
Optical Lenses	(6.49)	Iron Castings	(15.68)	Iron Bars	(10.92)
764	1.188	**665**	1.521	**821**	1.380
Audio Speakers	(9.47)	Glassware	(12.02)	Furniture	(25.50)
762	0.997	**663**	1.259	**634**	1.202
Radios	(10.46)	Mineral Mfg.	(11.77)	Wood Panels	(8.02)
759	0.914	**666**	1.471	**661**	1.195
Computer Parts	(6.33)	Pottery	(20.00)	Cement	(10.61)
514	1.053	**678**	1.842	**662**	1.692
Nitrogen	(6.60)	Iron Tubes	(15.73)	Clay	(12.84)
881	1.132	**642**	1.555		
Cameras	(6.51)	Paper Products	(12.78)		
751	0.849	**812**	1.226		
Office Machines	(3.82)	Plumbing Fixtures	(20.05)		
882	1.709	**625**	1.355		
Camera Supplies	(12.24)	Tires	(14.07)		
885	1.637	**676**	1.214		
Watches, Clocks	(15.47)	Steel Rails	(8.10)		
726	0.032	**641**	1.654		
Printing Machin.	(12.89)	Paper, Paperboard	(7.16)		

Notes: This table shows coefficient estimates on log relative exporter GDP from regressions in which the dependent variable is log relative industry exports for a pair of countries. T-statistics (calculated from standard errors that have been adjusted for correlation of the errors across observations that share the same pair of exporting countries) are shown in parentheses. Coefficient estimates on other regressors (see text) are suppressed. For each industry, the sample is relative bilateral exports by 107 country pairs to 58 large importing countries (5115 observations).

Table 5: Difference-in-Difference Gravity Equation, Pooled Sample of Industries

Independent Variables	(1)	(2)	(3)	(4)
ln(GDP)	0.420		0.420	
	(3.46)		(3.45)	
$\ln(GDP)^2$	0.026			
	(0.72)			
GDP – 1		0.105		0.104
		(1.71)		(4.40)
$(GDP - 1)^2$		0.000		
		(-0.03)		
Distance	-0.273	-0.264	-0.275	-0.264
	(-5.01)	(-4.95)	(-5.07)	(-4.99)
Common Language	-0.420	-0.346	-0.422	-0.345
	(-3.39)	(-2.65)	(-3.34)	(-2.81)
Common Border	0.888	0.811	0.893	0.811
	(10.33)	(8.91)	(10.13)	(9.13)
Capital/Worker	1.697	1.819	1.699	1.819
	(4.62)	(4.53)	(4.69)	(4.49)
Wage in Low-Skill Industries	-1.897	-1.730	-1.901	-1.729
	(-8.37)	(-7.33)	(-8.35)	(-8.00)
Area/Population	0.253	0.160	0.243	0.159
	(2.43)	(1.97)	(2.32)	(2.12)
Average Education	-3.090	-3.492	-3.139	-3.496
	(-7.03)	(-7.95)	(-7.22)	(-8.88)
Constant	-0.260	-0.308	-0.191	-0.306
	(-1.80)	(-2.13)	(-1.51)	(-2.44)
R Squared	0.077	0.074	0.077	0.074

Notes: This table shows estimation results for the specification in equation (11), in which the dependent variable is, for a pair of countries, log relative exports in a treatment industry minus log relative exports in a control industry. GDP is the GDP ratio for a country pair. Other variables are expressed as differences (Common Language, Common Border) or log differences (all other variables) for a country pair. T-statistics (calculated from standard errors that have been adjusted for correlation of the errors across observations that share the same pair of exporting countries) are in parentheses. The sample is exports by 107 country pairs to 58 importing countries pooled across the 273 treatment-control industry matches in the data (N=1,396,395).

Table 6: Difference-in-Difference Gravity Equation, Additional Results

Regressors	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
GDP	0.333	0.420	0.493	0.461	0.429	0.452	0.00003	0.00002
	(2.98)	(2.96)	(3.45)	(4.08)	(3.80)	(3.77)	(0.00007)	(0.00010)
Market	-1.019							
Potential	(-2.53)							
Distance	-0.308	0.195	0.519	-0.249	-0.363	-0.161		
	(-6.18)	(2.08)	(4.75)	(-4.06)	(-7.25)	(-2.59)		
Common	-0.324	-0.618	-1.334	-0.433	-0.471	-0.384		
Language	(-2.74)	(-4.09)	(-6.85)	(-3.60)	(-3.50)	(-3.33)		
Common	0.818	1.250	1.357	0.896	0.919	0.870		
Border	(10.33)	(8.52)	(9.10)	(8.02)	(9.25)	(8.56)		
Capital/	2.009	2.70	2.076	1.398	1.493	1.604		
Worker	(5.21)	(6.70)	(5.17)	(4.34)	(4.37)	(4.86)		
Low-Skill	-1.833	-2.260	-2.422	-2.654	-2.360	-2.195		
Wage	(-9.06)	(-7.99)	(-9.66)	(-12.09)	(-10.85)	(-9.98)		
Area/	0.006	0.189	0.258	0.037	0.085	0.195		
Population	(0.04)	(1.50)	(1.99)	(0.37)	(0.89)	(1.84)		
Average	-3.080	-2.697	-2.446	-4.914	-3.950	-4.103		
Education	(-7.70)	(-5.60)	(-5.37)	(-12.01)	(-9.95)	(-9.71)		
Constant	-0.206	-0.156	-0.186	-0.190	-0.160	-0.220		
	(-1.90)	(-1.06)	(-1.30)	(-1.62)	(-1.37)	(-1.78)		
N	1396395	344526	162981	276210	598455	644490		
R Squared	0.082	0.112	0.132	0.141	0.1	0.108	--	--

Notes: This table shows results in which the sample or specification is modified relative to the regression in column (3) of Table 5, which we refer to as the preferred regression. See notes to Table 5 for other details on the estimation. Column (1) adds log relative market potential to the preferred regression. Columns (2) and (3) change the sample in the preferred regression by restricting importers to be the 15 (column 2) or 7 (column 3) largest importing countries. Columns (4), (5), and (6) change the sample in the preferred regression by restricting treatment-control industry matches to be more restrictive treatment to more restrictive control industries (4), more restrictive treatment to all control industries (5), or all treatment to more restrictive control industries (6). Columns (7) and (8) change the sample and specification in the preferred regression by randomly matching industries (where industry pairs are either the full set of possible industry matches (7) or the initial treatment-control industry matches (8)), rerunning the difference-in-difference gravity regression, and then repeating the exercise 1,000 times.

Table 7: Summary of Industry-by-Industry Regression Results

	No. of Cases	Share of Regressions with $\beta>0$	Share of Regressions with p-value<0.1
All Industry Matches	273	0.857	0.582
More-restrictive Treatment-More-restrictive Control	54	1.000	0.741
Less-restrictive Treatment-More-restrictive Control	72	0.986	0.750
More-restrictive Treatment-Less-restrictive Control	63	0.762	0.460
Less-restrictive Treatment-Less-restrictive Control	84	0.726	0.429

Notes: This table summarizes the coefficient estimates on log relative exporter GDP for the specification in column (3) of Table 5, in which we re-estimate the regression separately for each of the 273 treatment-control industry matches in the data (N=5115 for each industry pair). It shows, for all industry matches or subsets of industry matches, shares of regressions with a positive coefficient estimate and with a positive coefficient estimate that is statistically significant at the 10 percent level. The more restrictive treatment industries are 666, 678, 625, 676, 677, 672, 673, 661 and 662. The less restrictive treatment industries are 671, 621, 674, 679, 665, 663, 642, 812, 641, 635, 821 and 634. The more restrictive control industries are 541, 752, 761, 764, 762 and 759. The less restrictive control industries are 884, 514, 881, 751, 882, 885 and 726.

Table 8: Summary of Regressions without Industries 882 and 885

	No. of Cases	Share of Regressions with $\beta>0$	Share of Regressions with p-value<0.1
All industries	231	0.978	0.680
More-restrictive Treatment-More-restrictive Control	54	1.000	0.741
Less-restrictive Treatment-More-restrictive Control	72	0.986	0.750
More-restrictive Treatment-Less-restrictive Control	45	0.956	0.644
Less-restrictive Treatment-Less-restrictive Control	60	0.967	0.567

Notes: This table summarizes regression results in a way that is similar to Table 7, except that all the matches involving industries 882 and 885 have been dropped from the sample. See notes to Table 7 for other details.

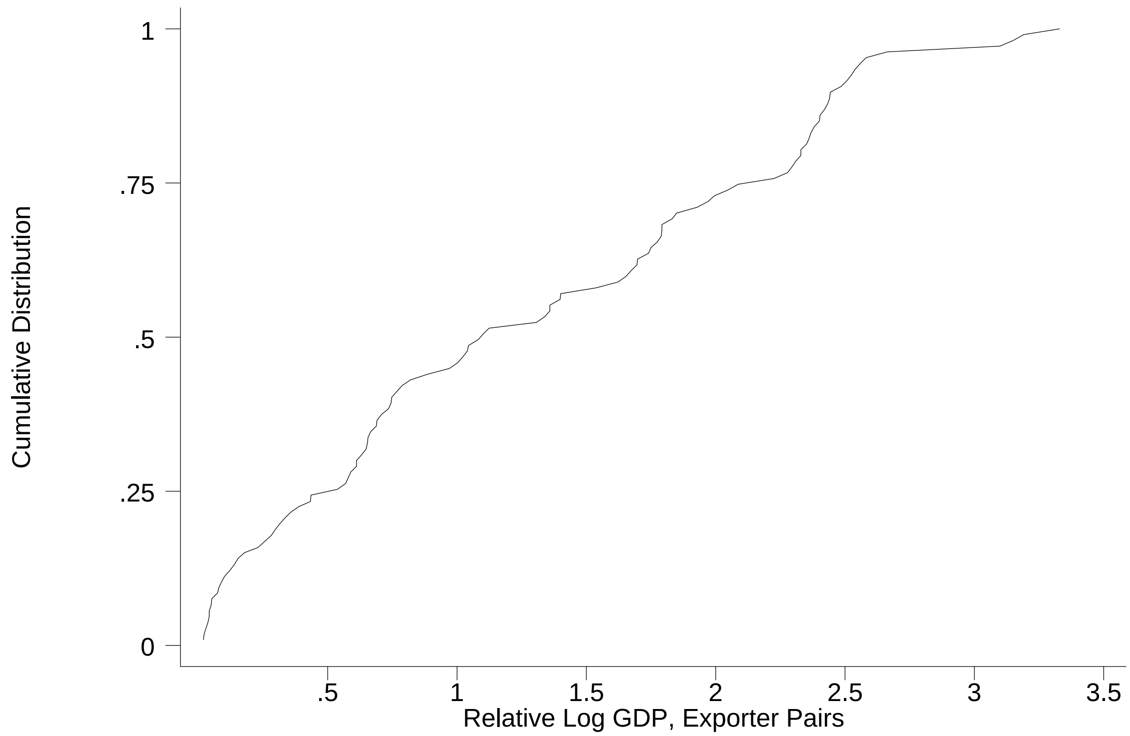

Figure 1: Distribution of Relative Exporter GDP

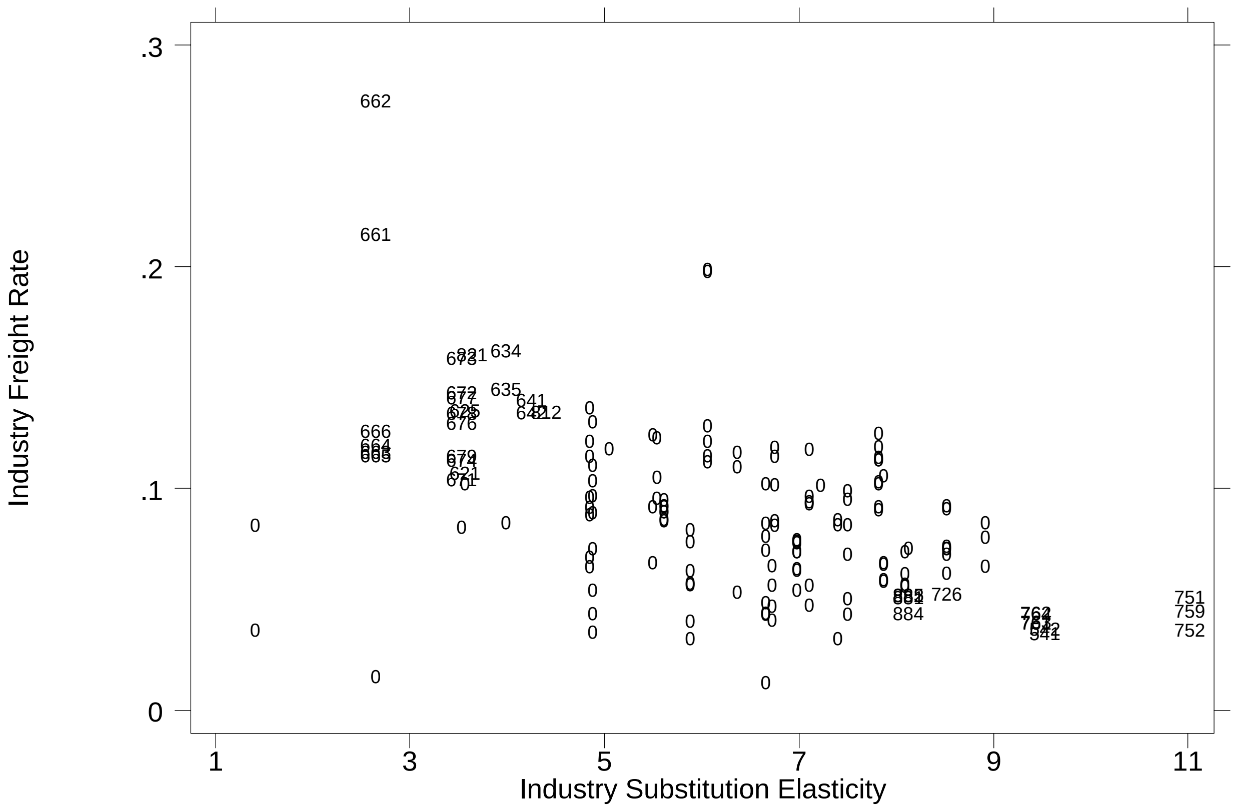

Figure 2: Distribution of Industry Freight Rates and Substitution Elasticities