Export Versus FDI with Heterogeneous Firms

By Elhanan Helpman, Marc J. Melitz, and Stephen R. Yeaple*

Multinational sales have grown at high rates over the last two decades, outpacing the remarkable expansion of trade in manufactures. Consequently, the trade literature has sought to incorporate the mode of foreign market access into the "new" trade theory. This literature recognizes that firms can serve foreign buyers through a variety of channels: they can export their products to foreign customers, serve them through foreign subsidiaries, or license foreign firms to produce their products.

Our work focuses on the firm's choice between exports and "horizontal" foreign direct investment (FDI). Horizontal FDI refers to an investment in a foreign production facility that is designed to serve customers in the foreign market.[1,2] Firms invest abroad when the gains from avoiding trade costs outweigh the costs of maintaining capacity in multiple markets. This is known as the proximity-concentration trade-off.[3] We introduce heterogeneous firms into a simple multicountry, multisector model, in which firms face a proximity-concentration trade-off. Every firm decides whether to serve a foreign market, and whether to do so through exports or local subsidiary sales. These modes of market access have different relative costs: exporting involves lower fixed costs while FDI involves lower variable costs.

Our model highlights the important role of within-sector firm productivity differences in explaining the structure of international trade and investment. First, only the most productive firms engage in foreign activities. This result mirrors other findings on firm heterogeneity and trade; in particular, the results reported in Melitz (2003).[4] Second, of those firms that serve foreign markets, only the most productive engage in FDI.[5] Third, FDI sales relative to exports are larger in sectors with more firm heterogeneity.

Using U.S. exports and affiliate sales data that cover 52 manufacturing sectors and 38 countries, we show that cross-sectoral differences in firm heterogeneity predict the composition of trade and investment in the manner suggested by our model. We construct several measures of firm heterogeneity, using different data sources, and show that our results are robust across all these measures. In addition, we confirm the predictions of the proximity-concentration trade-off. That is, firms tend to substitute FDI sales for exports when transport

[1] See Wilfred J. Ethier (1986), Ignatius Horstmann and James R. Markusen (1987), and Ethier and Markusen (1996) for models that incorporate the licensing alternative.

[2] We therefore exclude "vertical" motives for FDI that involve fragmentation of production across countries. See Helpman (1984, 1985), Markusen (2002, Ch. 9), and Gordon H. Hanson et al. (2002) for treatments of this form of FDI.

[3] See, for example, Horstmann and Markusen (1992), S. Lael Brainard (1993), and Markusen and Anthony J. Venables (2000).

[4] See also Andrew B. Bernard et al. (2003) for an alternative theoretical model and Yeaple (2003a) for a model based on worker-skill heterogeneity. James R. Tybout (2003) surveys the recent micro-level evidence on trade that has motivated these theoretical models.

[5] This result is loosely connected to the documented empirical pattern that foreign-owned affiliates are more productive than domestically owned producers. See Mark E. Doms and J. Bradford Jensen (1998) for the United States and Sourafel Girma et al. (2002) for the United Kingdom.

TABLE 1—PRODUCTIVITY ADVANTAGE OF MULTINATIONALS
AND EXPORTERS

Multinational	0.537
	(14.432)
Nonmultinational exporter	0.388
	(9.535)
Coefficient difference	0.150
	(3.694)
Number of firms	3,202

Notes: T-statistics are in parentheses (calculated on the basis of White standard errors). Coefficients for capital intensity controls and industry effects are suppressed.

costs are large and plant-level returns to scale are small. Moreover, the magnitude of the impact of our heterogeneity variables are comparable to the magnitude of the impact of the proximity-concentration trade-off variables. We conclude that intra-industry firm heterogeneity plays an important role in explaining international trade and investment.

As mentioned above, our model predicts that the least productive firms serve only the domestic market, that relatively more productive firms export, and that the most productive firms engage in FDI. We provide some evidence supporting this sorting pattern. We compute labor productivity (log of output per worker) for all firms in the COMPUSTAT database in 1996. We then regress this productivity measure on dummies for multinational firms (MNEs) and non-MNE exporters, controlling for capital intensity and 4-digit industry effects. Table 1 reports the resulting estimates for the productivity advantage of MNEs and non-MNE exporters over the remaining firms.[6] These results confirm previous findings of a significant productivity advantage of firms engaged in international commerce. In addition, they highlight a new prediction of our model: MNEs are substantially more productive than non-MNE exporters; the estimated 15-percent productivity advantage of multinationals over exporters is significant beyond the 99-percent level.

The remainder of this paper is composed of four sections. In Section I, we elaborate the model and we map the theoretical results into an empirical strategy. In Section II, we describe the data. We report and interpret the empirical findings in Section III, and we provide concluding comments in the closing section.

I. Theoretical Framework

There are N countries that use labor to produce goods in $H + 1$ sectors. One sector produces a homogeneous product with one unit of labor per unit output, while H sectors produce differentiated products. An exogenous fraction β_h of income is spent on differentiated products of sector h, and the remaining fraction $1 - \Sigma_h \beta_h$ on the homogeneous good, which is our numeraire. Country i is endowed with L^i units of labor and its wage rate is w^i.

Now consider a particular sector h that produces differentiated products. For the time being we drop the index h, with the implicit understanding that all sectoral variables refer to sector h. To enter the industry in country i, a firm bears the fixed costs of entry f_E, measured in labor units. An entrant then draws a labor-per-unit-output coefficient a from a distribution $G(a)$. Upon observing this draw, a firm may decide to exit and not produce. If it chooses to produce, however, it bears additional fixed overhead labor costs f_D. There are no other fixed costs when the firm sells only in the home country. If the firm chooses to export, however, it bears additional fixed costs f_X per foreign market. On the other hand, if it chooses to serve a foreign market via foreign direct investment (FDI), it bears additional fixed costs f_I in every foreign market. We think about f_X as the costs of forming a distribution and servicing network in a foreign country (similar costs for the home market are included in f_D). The fixed costs f_I include these distribution and servicing network costs, as well as the costs of forming a subsidiary in a foreign country and the duplicate overhead production costs embodied in f_D. The difference between f_I and f_X thus indexes plant-level returns to scale for the sector.[7] Goods that

---

[6] Our controls include the log of capital (book value net of depreciation) per worker, this variable squared, and 4-digit industry fixed effects. Controlling for material usage intensity does not change the results.

[7] Part of the cost difference $f_I - f_X$ may also reflect some of the entry costs represented by f_E, such as the initial cost of building another production facility.

are exported from country i to country j are subjected to melting-iceberg transport costs $\tau^{ij} > 1$. Namely, τ^{ij} units have to be shipped from country i to country j for one unit to arrive. After entry, producers engage in monopolistic competition.

Preferences across varieties of product h have the standard CES form, with an elasticity of substitution $\varepsilon = 1/(1 - \alpha) > 1$. These preferences generate a demand function $A^i p^{-\varepsilon}$ in country i for every brand of the product, where the demand level A^i is exogenous from the point of view of the individual supplier.[8] In this case, the brand of a monopolistic producer with labor coefficient a is offered for sale at the price $p = w^i a/\alpha$, where $1/\alpha$ represents the markup factor. As a result, the effective consumer price is $w^i a/\alpha$ for domestically produced goods—supplied either by a domestic producer or foreign affiliate with labor coefficient a—and is $\tau^{ji} w^j a/\alpha$ for imported products from an exporter from country j with labor coefficient a.

A firm from country i that remains in the industry will always serve its domestic market through domestic production. It may also serve a foreign market j. If so, it will choose to access this foreign market via exports or affiliate production (FDI). This choice is driven by the proximity-concentration trade-off: relative to exports, FDI saves transport costs, but duplicates production facilities and therefore requires higher fixed costs.[9] In equilibrium, no firm engages in both activities for the same foreign market.[10] We assume

$$(1) \qquad \left(\frac{w^j}{w^i}\right)^{\varepsilon - 1} f_I > (\tau^{ij})^{\varepsilon - 1} f_X > f_D.$$

These conditions will be clarified in the following analysis.
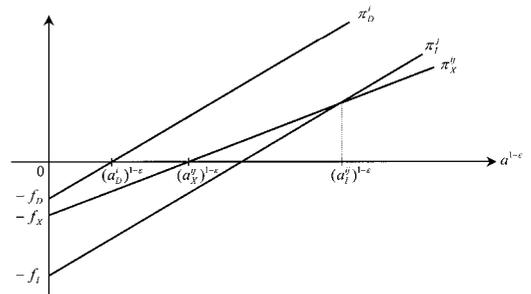

FIGURE 1. PROFITS FROM DOMESTIC SALES, FROM EXPORTS, AND FROM FDI

For expositional simplicity, assume for the time being unit wages in every country ($w^i = 1$).[11] Then, operating profits from serving the domestic market are $\pi_D^i = a^{1 - \varepsilon} B^i - f_D$ for a firm with a labor-output coefficient a, where $B^i = (1 - \alpha) A^i/\alpha^{1 - \varepsilon}$.[12] On the other hand, the *additional* operating profits from exporting to country j are $\pi_X^{ij} = (\tau^{ij} a)^{1 - \varepsilon} B^j - f_X$, and the additional operating profits from FDI in country j are $\pi_I^j = a^{1 - \varepsilon} B^j - f_I$. These profit functions are depicted in Figure 1 for the case of equal demand levels $B^i = B^j$.[13] In this figure, $a^{1 - \varepsilon}$ is represented on the horizontal axis. Since $\varepsilon > 1$, this variable increases monotonically with labor productivity $1/a$, and can be used as a productivity index. All three profit functions are increasing (and linear): more productive firms are more profitable in all three activities. The profit functions π_D^i and π_I^j are parallel, because we assumed $B^i = B^j$. However, profits from FDI are lower, as the fixed costs of FDI, f_I, are higher than the fixed costs of domestic production, f_D. The profit function π_X^{ij} is steeper than both π_D^i and π_I^j due to the trade costs τ^{ij}. Together with the first inequality in (1), these relationships imply that exports are more prof-

---

[8] As is well known, our utility function implies that $A^i = \beta E^i/[\int_0^{n^i} p^i(v)^{1 - \varepsilon} dv]$, where E^i is the aggregate level of spending in country i, n^i is the number (measure) of varieties available in country i and $p^i(v)$ is the consumer price of variety v.

[9] We exclude the possibility of exports by foreign affiliates. See, however, the Appendix of our working paper, Helpman et al. (2003), for a discussion of this possibility.

[10] In a dynamic model with uncertainty, an individual firm may choose to serve a foreign market through both exports and FDI. Rafael Rob and Nikolaos Vettas (2003) provide a rigorous treatment of this case.

[11] This will be the case so long as the numeraire good is produced in every country and freely traded.

[12] Note that the demand function $A^i p^{-\varepsilon}$ implies output $A^i (a/\alpha)^{-\varepsilon}$ when the price is a/α. Under these circumstances, costs are $\alpha A^i (a/\alpha)^{1 - \varepsilon}$, while revenue is $A^i (a/\alpha)^{1 - \varepsilon}$. Therefore, operating profits are $\pi_D^i = (1 - \alpha) A^i (a/\alpha)^{1 - \varepsilon} - f_D$.

[13] We thank Dani Tsiddon for proposing this figure. In the figure $f_X > f_D$, which is a sufficient condition for the second inequality in (1). Evidently, this inequality can also be satisfied when $f_X < f_D$, and we need only the inequality in (1) in order to ensure that some firms serve only the domestic market.

itable than FDI for low-productivity firms and less profitable for high-productivity firms. Moreover, there exist productivity levels at which exporters have positive operating profits that exceed the operating profits from FDI [since $(a_X^{ij})^{1-\varepsilon} > (a_I^{ij})^{1-\varepsilon}$], which ensures that some firms export to country j. In addition, the second inequality in (1) implies $(a_X^{ij})^{1-\varepsilon} > (a_D^i)^{1-\varepsilon}$, which ensures that some firms serve only the domestic market.

The least productive firms expect negative operating profits and therefore exit the industry. This fate befalls all firms with productivity levels below $(a_D^i)^{1-\varepsilon}$, which is the cutoff at which operating profits from domestic sales equal zero. Firms with productivity levels between $(a_D^i)^{1-\varepsilon}$ and $(a_X^{ij})^{1-\varepsilon}$ have positive operating profits from sales in the domestic market, but expect to lose money from exports and FDI. They choose to serve the domestic market but not to serve country j. The cutoff $(a_X^{ij})^{1-\varepsilon}$ is the productivity level at which exporters just break even. Higher-productivity firms can export profitably. But those with productivity above $(a_I^{ij})^{1-\varepsilon}$ gain more from FDI. For this reason, firms with productivity levels between $(a_X^{ij})^{1-\varepsilon}$ and $(a_I^{ij})^{1-\varepsilon}$ export while those with higher productivity levels build subsidiaries in country j, which they use as platforms for servicing country j's market. It is evident from the figure that the cutoff coefficients $(a_D^i)^{1-\varepsilon}$, $(a_X^{ij})^{1-\varepsilon}$, and $(a_I^{ij})^{1-\varepsilon}$ are determined by

$$(2) \qquad (a_D^i)^{1-\varepsilon} B^i = f_D, \ \forall i,$$

$$(3) \qquad (\tau^{ij} a_X^{ij})^{1-\varepsilon} B^j = f_X, \ \forall j \neq i,$$

(4)

$$[1 - (\tau^{ij})^{1-\varepsilon}] (a_I^{ij})^{1-\varepsilon} B^j = f_I - f_X, \ \forall j \neq i.$$

Free entry ensures equality between the expected operating profits of a potential entrant and the entry costs f_E. The form of this condition is reported in our working paper (Helpman et al., 2003). The free entry condition together with (2)–(4) provide implicit solutions for the cutoff coefficients a_D^i, a_X^{ij}, a_I^{ij}, and the demand levels B^i in every country. These solutions do not depend on the country-size variables L^i as long as productivity-adjusted wages w^i remain

equalized (the numeraire outside good is produced everywhere and freely traded). Moreover, we can also allow cross-country variations in the fixed-cost coefficients, as long as these variations do not lead some countries to stop producing the outside good. These generalizations are useful for empirical applications.

We report in our working paper general-equilibrium results for a special case in which countries only differ in size and trade costs per product are symmetric (implying $\tau^{ij} = \tau$ for $j \neq i$). These restrictions apply within each sector, so there can be arbitrary variations across sectors. Under these circumstances, (2)–(4) and free entry imply that, as long as countries do not differ too much in size, wages are the same everywhere, all countries share the same cutoffs $a_D^i = a_D$, $a_X^{ij} = a_X$, $a_I^{ij} = a_I$, and the same demand levels $B^i = B$. Larger countries attract a disproportionately larger number of entrants (relative to country size) and a larger number of sellers (hence, more product variety). We also show that larger markets are disproportionately served by domestically owned firms, i.e., the market share of domestically owned firms is larger in the home market of a larger country.

A. Exports Versus FDI Sales

We now consider the relative magnitude of exports and local FDI sales for a pair of countries i and j. Let s_X^{ij} be the market share in country j of country i's exporters and let s_I^{ij} be the market share in country j of affiliates of country i's multinationals. The relative size of these market shares is then

$$(5) \qquad \frac{s_X^{ij}}{s_I^{ij}} = \tau^{1-\varepsilon} \left[\frac{V(a_X)}{V(a_I)} - 1 \right]$$

in the symmetric case, where

$$(6) \qquad V(a) = \int_0^a y^{1-\varepsilon} dG(y).$$

Given our symmetry assumptions, this ratio is independent of i and j. That is, every country has the same relative sales of exporters and affiliates in every other country. This ratio rises

with the exporting cutoff coefficient a_X and declines with the FDI cutoff coefficient a_I. The cutoffs, in turn, are determined by the system of equilibrium conditions.

A rise in the export costs f_X or τ, or a decrease in the FDI costs f_I, all have similar impacts on the a_X and a_I cutoffs: they induce an increase in a_I and a decrease in a_X. The relative sales of exporters thus decline in all these cases. Recall that f_I encompasses both the country-level fixed costs embodied in f_X and the duplicate plant overhead costs represented by f_D. It is therefore natural to consider the effects of equivalent increases in f_I and f_X (representing higher country-level costs), and the effects of equivalent decreases in f_I and f_D (representing lower overhead plant costs, and hence smaller returns to scale). Again, we show that the a_I and a_X cutoffs move in the same directions as before, entailing a decrease in relative export sales.

These are sensible comparative statics predicting the cross-sectoral variation in relative exports sales. We expect the relative sales of exporters to be lower in sectors with higher transport costs or higher fixed country-level costs (even when the latter costs are also borne by multinational affiliates). We also expect them to be lower in sectors where plant-level returns to scale are relatively weak. These results show how the firm-level proximity-concentration trade-off results can be extended to sectors with heterogeneous firms that select different modes of foreign market access.

We now shift the focus to the role of firm-level heterogeneity in explaining the cross-sectoral variation in relative export sales. Note from (5) that the function $V(\cdot)$ directly impacts the relative sales (holding the cutoff levels fixed). Recall also that firm sales and variable profits are proportional to $a^{1-\varepsilon}$ in every market. $V(a)$ therefore captures (up to a multiplicative constant) the distribution of sales and variable profits among firms that make the same export or FDI decisions. It also captures the distribution of domestic sales and variable profits among all surviving firms. We think of $V(a)$ as summarizing firm-level heterogeneity in a sector. It is exogenously determined by the distribution of unit costs $G(a)$ and the elasticity of substitution ε.

In order to index differences in firm-level heterogeneity across sectors, we parametrize

$G(a)$. We use the Pareto distribution as a benchmark. When labor productivity $1/a$ is drawn from a Pareto distribution with the shape parameter k, the distribution of firm domestic sales, indexed by $V(a)$, is also Pareto, with the shape parameter $k - (\varepsilon - 1)$.[14] The shape parameter of the Pareto distribution offers a natural and convenient index of dispersion, which characterizes heterogeneity. Given our assumptions, the domestic sales of all firms with sales above any given cutoff are distributed Pareto with the same shape parameter $k - (\varepsilon - 1)$. A higher dispersion of firm productivity draws (lower k) or a higher elasticity of substitution ε, raise the dispersion of firm domestic sales and variable profits. We now investigate the impact of such changes in heterogeneity on the relative sales of exporters.

The Pareto distribution implies that $V(a_1)/V(a_2)$ equals $(a_1/a_2)^{k-(\varepsilon-1)}$ for every a_1 and a_2 in the support of the distribution of a. Relative export sales in (5) can then be written as[15]

$$(7) \quad \frac{s_X^{ij}}{s_I^{ij}} = \tau^{1-\varepsilon} \left[\left(\frac{a_X}{a_I} \right)^{k-(\varepsilon-1)} - 1 \right]$$

$$= \tau^{1-\varepsilon} \left[\left(\frac{f_I - f_X}{f_X} \frac{1}{\tau^{\varepsilon-1} - 1} \right)^{\frac{k-(\varepsilon-1)}{\varepsilon-1}} - 1 \right].$$

It follows that relative export sales decrease with decreases in k and increases in ε.[16] Thus, we expect sectors with higher levels of disper-

---

[14] The cumulative distribution function of a Pareto random variable X with the shape parameter k is given by

$$F(x) = 1 - \left(\frac{b}{x} \right)^k, \text{ for } x \geq b > 0,$$

where b is a scale parameter that bounds the support $[b, +\infty)$ from below. Log x is then distributed exponentially with a standard deviation equal to $1/k$. Any truncation from below of X is also distributed Pareto with the same shape parameter k. X has a finite variance if and only if $k > 2$. We therefore assume that $k > \varepsilon + 1$, which ensures that both the distribution of productivity draws and the distribution of firm sales have finite variances.

[15] Equations (3) and (4) are used in this derivation.

[16] Recall that $(f_I - f_X)/f_X(\tau^{\varepsilon-1} - 1)$ is greater than 1, by assumption; see (1).

sion in firm domestic sales—generated either by higher dispersion levels of firm productivity or by a higher elasticity of substitution—to have lower levels of relative export sales. This is a major implication of the model, which we test below.

B. *Testable Implications*

We focus our empirical work on the model's predictions concerning the determinants of the cross-sector and cross-country variation in relative export sales. This empirical analysis requires us to relax the symmetry assumptions imposed in the previous subsection and to allow for cross-country variation in wages, transport costs, and technology.

Consider the decisions of U.S. firms in sector h to serve country j via export sales versus affiliate sales. The equilibrium cutoff levels must satisfy:

$$(8) \qquad (\tau_h^{Uj} w^U a_{hX}^{Uj})^{1-\varepsilon_h} B_h^j = w^j f_X^j,$$

$$(9) \quad [(w^j)^{1-\varepsilon_h} - (w^U \tau_h^{Uj})^{1-\varepsilon_h}] (a_{hI}^{Uj})^{1-\varepsilon_h} B_h^j$$
$$= w^j (f_{hI}^j - f_X^j),$$

where w^U and w^j are the wage levels in the United States and country j, τ_h^{Uj} is the trade cost (transport and tariff) from the United States to country j in sector h, ε_h is the elasticity of substitution across varieties in sector h (common to all countries), B_h^j indexes the demand level for sector h in country j, and f_{hI}^j and f_X^j represent the fixed costs of doing FDI in and exporting to country j. These conditions replace (3) and (4). Note that f_{hI}^j is also indexed by sector h, since it includes plant setup and overhead production costs. On the other hand, the fixed exporting costs are common across sectors; they index particular characteristics of doing business in country j for U.S. firms. These costs would also be incurred by U.S. firms setting up affiliates in country j, so the difference $f_{hI}^j - f_X^j$ represents the overhead and setup production costs. Let $f_{hP} \equiv f_{hI}^j - f_X^j$ reference these costs. Equations (8) and (9) then imply:

$$(10) \quad \left(\frac{a_{hX}^{Uj}}{a_{hI}^{Uj}}\right)^{\varepsilon_h - 1} = \frac{f_{hP}}{f_X^j} [(\omega^j \tau_h^{Uj})^{\varepsilon_h - 1} - 1]^{-1},$$

where $\omega^j \equiv w^U/w^j$ indexes the U.S. wage relative to country j.[17]

We index the level of U.S. firm heterogeneity across sectors using the Pareto benchmark. We assume that the productivity draws for U.S. firms in sector h are distributed Pareto with shape k_h^U, and therefore that the distribution of U.S. domestic sales indexed by $V_h^U(a)$ is also Pareto with shape $k_h^U - (\varepsilon_h - 1)$. The sales of U.S. exporters to country j relative to the U.S. affiliate sales in country j can then be written as

$$(11) \quad \frac{s_X^{Uj}}{s_I^{Uj}} = (\omega^j \tau_h^{Uj})^{1-\varepsilon_h} \left[\frac{V_h^U (a_{hX}^{Uj})}{V_h^U (a_{hI}^{Uj})} - 1\right]$$

$$= (\omega^j \tau_h^{Uj})^{1-\varepsilon_h} \left\{ \left[\frac{f_{hP}}{f_X^j} \frac{1}{(\omega^j \tau_h^{Uj})^{\varepsilon_h - 1} - 1}\right]^{\frac{k_h^U - (\varepsilon_h - 1)}{\varepsilon_h - 1}} - 1 \right\}.$$

Comparing (7) and (11) confirms that all our previously derived comparative statics remain valid in a cross section of both sectors and nonsymmetric countries: the proximity-concentration forces predict lower U.S. relative export sales for country-sector pairs with high transport costs τ_h^{Uj}, countries with high fixed costs f_X^j, and sectors with low plant-level returns to scale f_{hP}. As was previously the case, the extent of firm-level heterogeneity remains an important determinant of relative export sales. Sectors with higher productivity dispersion levels (lower k_h^U) or higher elasticities of substitution have lower relative export sales. We cannot separately measure k_h^U and ε_h. However, we can measure their difference $k_h^U - (\varepsilon_h - 1)$ under the Pareto assumption, because $1/[k_h^U - (\varepsilon_h - 1)]$ then indexes the measurable dispersion of firm size in sector h, and provides a convenient overall measure of differences in firm-level heterogeneity across sectors.

---

[17] We assume that $1/\tau^{Uj} < \omega^j < \tau^{jU}$, which ensures that some firms choose to locate in both county j and the United States, and we maintain assumption (1).

II. Data

To test our model, we need data that vary across sectors and countries. The required data fall into three categories: data on the composition of international trade, variables that represent the proximity-concentration trade-off, and indices of firm-level heterogeneity. We describe in this section our choice of these data. Unless otherwise noted, all of the data are for 1994.

A. *The Composition of International Commerce*

The biggest constraint on any analysis that considers the trade-off between exports and FDI sales is the dearth of internationally comparable measures of the extent of FDI across both industries and countries. Because the United States is one of a handful of countries that collects data on multinational affiliate sales, disaggregated by destination and industry, our study focuses on the composition of U.S. trade.

In the United States, the Bureau of Economic Analysis (BEA) collects census-type data on FDI. In its Benchmark Surveys, conducted every five years, the BEA collects affiliate-level data on a wide range of enterprise-level variables, including total affiliate sales. Affiliates are classified by their main line of business and assigned to one of 52 manufacturing sectors.[18] To make our FDI data comparable to the data for exports, we aggregated the firm-level multinational sales data to the level of the industry. The export data are more familiar and have been taken from Robert C. Feenstra (1997). The data have been concorded from 4-digit SITC industrial classifications into the BEA industry classifications.

Finally, we consider two separate samples of countries, which can be characterized as narrow and wide. The narrow sample consists of the 27 countries originally studied by Brainard (1997), while the wide sample includes 11 additional, smaller countries, which are typically less developed.[19] The benefit of the wider sample is

that it includes a larger and more diverse set of countries, while the drawback is that these countries are more likely to have fewer strictly positive levels of FDI, creating a concern about censoring.

B. *Proximity-Concentration Variables*

Our theoretical model predicts exports relative to FDI sales as a function of the costs of each activity: unit costs of exporting, fixed costs of exporting, and fixed costs of investment abroad. However, these costs are not easily quantified.

First consider unit costs of foreign trade. These costs can be due either to the costs of moving goods across borders, such as transport and insurance, or due to barriers to trade, such as tariffs. We proxy for them with the variables FREIGHT and TARIFF, which are ad valorem measures of freight and insurance costs, and trade taxes. FREIGHT is computed as the ratio of CIF imports into the United States to FOB imports, using the data in Feenstra (1997). TARIFF is calculated at the BEA industry/country-level from more finely disaggregated data. It is the unweighted average of tariffs across subindustries within the BEA industry. Trade taxes are taken from Yeaple (2003b), where the data are described in more detail.

While the unit costs of shipping goods are reasonably straightforward to measure, the same cannot be said for the fixed costs associated with exporting and FDI. In principle, these costs could vary by industry and country; but such measures do not exist in practice. To make progress, we maintain our assumption of a country-specific fixed cost that applies to both exports and FDI sales. As this cost is unobserved, country-specific, and common to all industries, we subsume its measure into a country fixed effect.

We therefore assume that any remaining cost associated with FDI stems from the cost of maintaining additional capacity. Given our

---

[18] See Table 1 in our working paper, Helpman et al. (2003), for this classification.

[19] The 27 countries in the narrow sample are Argentina, Australia, Austria, Belgium, Brazil, Canada, Chile, Denmark, France, Germany, Hong Kong, Ireland, Italy, Japan,

Mexico, Netherlands, New Zealand, Norway, Philippines, Singapore, South Korea, Spain, Sweden, Switzerland, Taiwan, United Kingdom, and Venezuela, while the 11 additional countries are Colombia, Finland, Greece, Indonesia, Israel, Malaysia, Peru, Portugal, South Africa, Thailand, and Turkey.

model, we cannot use a measure for a firm that is somehow "representative" of the sector. Thus, standard measures of plant-level fixed costs, such as the number of production workers at a plant of median size, are not appropriate. We seek a measure of such costs that is independent of firm size or productivity. We follow the model in choosing the number of nonproduction workers per establishment as reported in the 1997 Census of Manufacturing.[20] We calculate the average number of nonproduction workers at the North American Industry Classification System (NAICS) level.[21] Then, we compute the measure of plant-level fixed cost, FP, for every BEA sector as the average of these numbers within the BEA sector, weighted by the NAICS-level sales in the sector.

C. *Measures of Dispersion*

The most novel feature of our model is the relationship between the degree of intra-industry firm heterogeneity and the prevalence of subsidiary sales relative to export sales. To test this hypothesis, we require data that quantifies the extent of this heterogeneity across industries. As we cannot directly measure the dispersion of intra-industry productivity levels, we rely on guidance from the model to construct alternative measures of within-industry heterogeneity. According to our model, the dispersion of firm size within a sector captures the joint effect of the dispersion of firm productivity and the elasticity of substitution, which magnifies the effect of productivity differences across firms. Since the size distribution of firms is observable, we use its dispersion as a measure of firm-level heterogeneity.

To quantify this dispersion measure across industries, we assume that the stochastic process that determines firm productivity levels is Pareto, with the shape of the distribution varying across industries. This assumption is convenient, because it suggests two conceptually

equivalent ways to measure dispersion. The first is to regress the logarithm of an individual firm's rank within the distribution on the logarithm of the firm's size. It can be shown that the estimated coefficient of such a regression is $k - (\varepsilon - 1)$, which is exactly the measure of dispersion that appears in the reduced form of the model.[22] The second method is to compute the standard deviation of the logarithm of firm sales, which—given our distributional assumption—is computationally equivalent to the slope of the conditional expectation of log rank on log size.[23]

While our distributional assumption yields a precise methodology for computing dispersion, the choice of data is more problematic. We require disaggregated data on the distribution of sales across firms. Unfortunately, we do not have access to such data on U.S. firms. As a result, we rely on two alternative sources.

First, we use the publicly available data from the 1997 U.S. Census of Manufacturing. However, these data are aggregated into ten different size categories, precluding the estimation of size dispersion measures using regression techniques. Nevertheless, we can compute the standard deviation of log sales by making an additional assumption: we assume that all establishments falling within the same size category have log sales equal to the mean for this category. We then treat each of the size categories in the many subindustries of the BEA industry classification as separate observations. Adopting this method, we calculate the standard deviation of log sales using the number of firms in each size category as weights.

Second, Bureau van Dijck Electronic Publishing has recently made available a large data set of European firms.[24] This database, named Amadeus, includes information on the consolidated sales, the national identity, and the main line of business of a large number of European

---

[20] This measure does not strictly conform to our modeling assumptions, because the number of nonproduction workers is not independent of establishment size.

[21] The new 6-digit North American Industrial Classification System replaces the 4-digit Standard Industrial Classification, but provides roughly the same level of industry aggregation.

[22] It is comforting that the distribution of firm size closely follows a Pareto distribution; see Robert L. Axtell (2001).

[23] While the two methods of calculation should be equivalent, in practice there are moderate to small differences in their values. We therefore calculate them both ways.

[24] This data set has been used by John W. Budd et al. (2002), who investigate international rent-sharing within multinational firms. We thank Matthew Slaughter for bringing this data set to our attention.
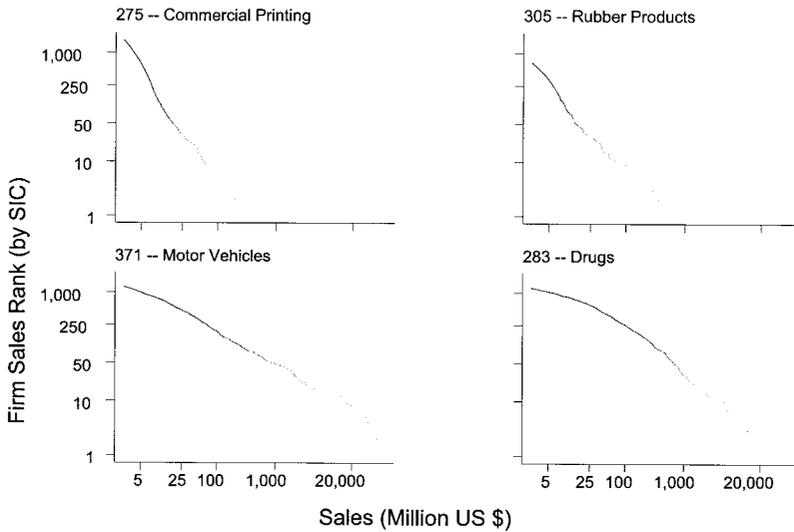
FIGURE 2. EMPIRICAL DISTRIBUTION OF FIRM SALES

firms. There are roughly 260,000 firms in this sample.

We compute each of our two measures of dispersion for every industry in two subsets of these data: all Western European firms and French firms only. We compute our firm dispersion measures using French firms only for two reasons. First, using data for multiple countries raises the issue of industrial composition. Within every BEA industry there are many sub-industries for which countries might produce different mixes. France's industrial structure is very similar to the United States, however, and so might share most of the same distributional aspects of firm characteristics. Second, French firms are highly overrepresented in the sample relative to all other Western European countries.[25] Our dispersion measures are based on a sample of 55,339 large Western European firms, and a subset of 15,148 French firms.[26]

The regression-based measures of dispersion provide a natural way of evaluating the cross-sectional variation in this variable relative to the measurement errors induced by the fitting of the Pareto distributions. Figure 2, which has been constructed from the sample of Western European firms, plots firm rank against firm sales in four sectors on the same log-log scale. In every plot the dispersion measure is represented by the slope of the regression line while its goodness of fit is represented by the deviation from linearity. Figure 3 quantifies this comparison by showing the 95-percent confidence intervals around the coefficients of dispersion, estimated as the slopes of the regression lines in these sectors. Evidently, these slopes are precisely estimated in all the sectors, with the exception of five outliers that we discuss below.[27]

There are four measures of dispersion calculated from the Amadeus data set and one measure calculated from the U.S. data. The

---

[25] Due to national differences in reporting requirements, no information on U.K. firms is available, and only an extremely limited number of German firms appear in the sample.

[26] Because small firms are underrepresented throughout the Amadeus database, we first drop firms with sales below a cutoff of U.S. $2.5 million per year. Note that, under the assumption of a Pareto size distribution, our measures of dispersion are invariant to the choice of lower bound cutoff. We computed the dispersion measures using several differ-

ent cutoffs. Any cutoff above U.S. $2.5 million yields a size distribution that is closely approximated by a Pareto distribution, and a dispersion measure that varies very little with the cutoff.

[27] As all 52 manufacturing sectors could not fit on one graph, only one of the seven food processing sectors (201—meat products) is represented. The coefficients and confidence intervals for the other six food processing sectors are very similar to the one represented.
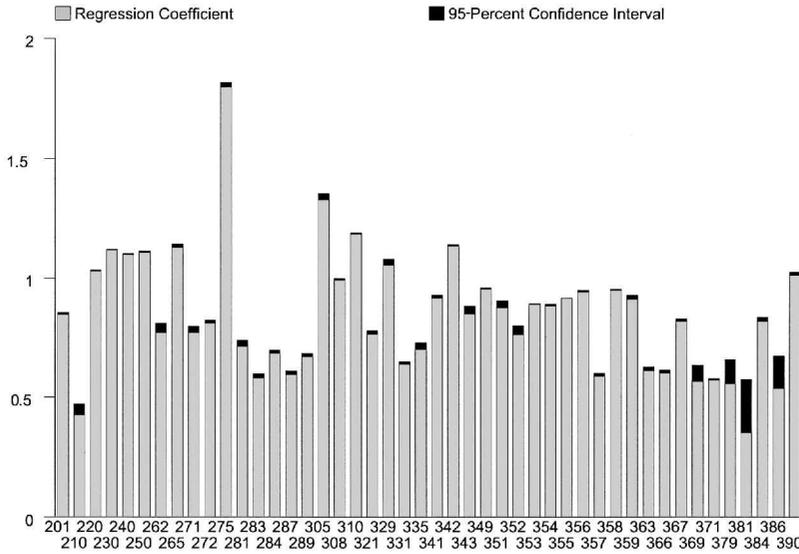
FIGURE 3. REGRESSION FIT TO THE PARETO DISTRIBUTION

TABLE 2—CORRELATIONS BETWEEN ALTERNATIVE MEASURES OF DISPERSION

	U.S. std. dev.	Europe std. dev.	France std. dev.	Europe reg. coeff.	France reg. coeff.	FP	RD	KL
U.S. s.d.	1							
Europe s.d.	0.507	1						
France s.d.	0.567	0.895	1					
Europe reg.	0.526	0.959	0.919	1				
France reg.	0.541	0.973	0.905	0.984	1			
FP	0.455	0.621	0.508	0.652	0.624	1		
RD	0.134	0.445	0.354	0.438	0.475	0.498	1	
KL	0.129	0.585	0.500	0.507	0.523	0.515	0.365	1

correlations between these measures are shown in Table 2 (along with our measure of plant-level fixed costs, FP, and the industries' capital-labor ratio, KL, and R&D intensity, RD). The table shows that all four measures from Amadeus are highly correlated with one another, as one might expect. The table also shows that the U.S.-based measure of dispersion is positively correlated with the measures of dispersion calculated from the European data, except that this correlation is not as high as the correlations among the four measures of dispersion that were calculated from the European data. There are at least two reasons why this might be so. First, the method of calculation is very differ-ent: the European measures are computed from actual firm-level data while the American measure is calculated from semiaggregated establishment-level data. Given the differences in methods of calculation, one might argue that the correlations are surprisingly high. Second, there exists an aggregation problem. If the composition of output varies across countries according to comparative advantage, then within each BEA industry the product mix of goods produced in the United States may differ from the mix produced in Europe. For this reason the European and American dispersion measures cannot be perfectly aligned.

III. Specifications and Results

Our aim is to estimate a linearized version of (11) that relates the logarithm of relative sales to our measure of firm-size dispersion, the logarithm of our proxy for plant fixed costs, the logarithms of transport and tariff costs, and a set of country dummies that we use to control for the differences in f_X and ω across countries. Of course, this linearization precludes any structural interpretation of the estimated parameters. Our goal is limited to testing whether the central tendencies in the data are consistent with the partial derivatives implied by (11), and to assessing the economic significance of the magnitudes associated with the estimated coefficients.

We consider several variants of the basic specification in order to raise the level of confidence in the results. Given the critical importance of the size distribution of firms, we report results corresponding to each one of the five measures of dispersion in firm size. We also report results for both samples of countries: narrow and wide. Finally, we explore the sensitivity of the results to alternative assumptions that incorporate other determinants of relative sales not captured by equation (11).

We begin the analysis by considering a specification that controls for sectoral capital and R&D intensities.[28] The results across specifications for our two samples and five measures of dispersion are shown in Table 3. The columns correspond to different measures of dispersion, beginning with the U.S. standard deviation of log sales, proceeding to the European and French-only standard deviation measures, and ending with the estimated distribution parameters for the European and the French-only sample, respectively. Country fixed effects are not reported.

First consider the narrow sample of relatively large countries, studied by Brainard (1997). The coefficients on FREIGHT and TARIFF are negative and statistically significant in each one of the five specifications. These results are consistent with Brainard (1997). In addition, the co-

efficient of FP is positive and significant. We therefore confirm the predictions of the proximity-concentration trade-off: firms substitute FDI sales for exports when the costs of international trade are relatively high and the returns to scale are relatively small.

Next consider the effects of dispersion. The estimated coefficients on the various dispersion measures are all negative and statistically significant. Industries in which firm size is highly dispersed are associated with relatively more FDI sales relative to exports, precisely as the model predicts. None of these results changes significantly when the set of countries is expanded to include the 11 smaller countries (the wide country sample).[29]

Although all measures of dispersion yield coefficients that are statistically significant, the choice of dispersion measure has a noticeable impact on the results. The measures that were derived by fitting a Pareto distribution to the distribution of firm size, yield substantially lower coefficients and higher standard errors than the nonparametric dispersion measures, i.e., the standard deviations of log sales. This pattern is driven, in large part, by five sectors that exhibit the largest differences between the measurement of dispersion by means of the shape of a Pareto distribution and by means of the standard deviation, for both Western European and French firms.[30] These sectors have the lowest number of firms in the data, and they yield—without exception—the poorest fits to the Pareto distribution, as measured by R-squares. We believe that in these cases the nonparametric measures (the standard deviations) better describe the levels of dispersion within the sectors. Dropping these five outliers from the sample and reestimating the equations, we find that the two different ways of measuring dispersion yield much more similar results. Af-

---

[28] In Helpman et al. (2003), we also report estimates without these controls. The two sets of estimates do not differ by much, and the measures of dispersion are highly significant in both cases.

[29] The magnitude of the coefficients on virtually all dispersion measures are lower in the wider sample. One possible explanation is that attenuation bias has affected the magnitudes of the coefficients. Another explanation is that the process generating FDI in the smaller developing countries is somewhat different from the process generating FDI in the larger developed countries.

[30] The five outliers consist of the following sectors: 210—tobacco, 369—other electronics, 379—other transport equipment, 381—scientific and measuring equipment, and 386—optical and photographic equipment.

TABLE 3—EXPORTS VERSUS FDI

Narrow sample (N = 961)

	U.S. std. dev.	Europe std. dev.	France std. dev.	Europe reg. coeff.	France reg. coeff.
FREIGHT	−1.040	−0.959	−1.019	−0.935	−0.944
	(−7.392)	(−6.749)	(−7.328)	(−6.526)	(−6.594)
TARIFF	−0.365	−0.512	−0.421	−0.545	−0.539
	(−2.644)	(−3.636)	(−3.917)	(−3.781)	(−3.775)
FP	1.177	0.932	0.927	0.947	0.934
	(10.159)	(7.827)	(8.059)	(7.453)	(7.450)
DISPERSE	−2.343	−2.153	−2.061	−1.503	−1.491
	(−8.374)	(−5.250)	(−6.664)	(−4.535)	(−4.470)
KL	−0.868	−0.495	−0.456	−0.628	−0.626
	(−7.790)	(−4.529)	(−4.256)	(−5.876)	(−5.859)
RD	−0.104	0.007	0.007	0.006	−0.002
	(−2.197)	(0.150)	(0.144)	(0.125)	(−0.047)
R^2	0.373	0.340	0.364	0.332	0.334

Wide sample (N = 1,175)

	U.S. std. dev.	Europe std. dev.	France std. dev.	Europe reg. coeff.	France reg. coeff.
FREIGHT	−1.011	−0.935	−0.960	−0.915	−0.919
	(−7.968)	(−7.246)	(−7.714)	(−7.040)	(−7.053)
TARIFF	−0.241	−0.384	−0.306	−0.411	−0.407
	(−1.876)	(−2.964)	(−2.457)	(−3.073)	(−3.057)
FP	1.133	0.861	0.868	0.867	0.848
	(10.428)	(7.719)	(7.994)	(7.318)	(7.243)
DISPERSE	−2.248	−1.866	−1.833	−1.284	−1.215
	(−8.611)	(−4.919)	(−5.982)	(−4.132)	(−3.924)
KL	−0.793	−0.454	−0.412	−0.569	−0.576
	(−7.483)	(−4.347)	(−3.982)	(−5.574)	(−5.636)
RD	−0.086	0.017	0.021	0.015	0.007
	(−1.914)	(0.367)	(0.446)	(0.326)	(0.153)
R^2	0.338	0.305	0.325	0.298	0.298

Notes: T-statistics are in parentheses (calculated on the basis of White standard errors). Constant and country dummies are suppressed.

ter dropping the outliers, all the dispersion measures yield negative coefficients that are significant beyond the 99-percent confidence level.

To get a sense of the economic significance of the estimated coefficients on our dispersion measures, we have calculated standardized coefficients—also known as "beta" coefficients—for all the independent variables. They are reported in Table 4 for the narrow sample, along with the sample means and standard deviations. A beta coefficient is defined as the product of the estimated coefficient and the standard deviation of its corresponding

TABLE 4—"BETA" COEFFICIENTS: NARROW SAMPLE WITH CONTROLS

	Mean	Standard deviation	"Beta" coefficient
Dependent variable	−0.595	2.375	
FREIGHT	1.863	0.653	−0.271
TARIFF	2.015	1.020	−0.205
FP	3.321	0.785	0.325
U.S. s.d.	1.749	0.316	−0.312
Europe s.d.	1.198	0.276	−0.250
France s.d.	1.224	0.375	−0.325
Europe reg.	1.260	0.333	−0.210
France reg.	1.257	0.336	−0.211

independent variable, divided by the standard deviation of the dependent variable. It converts the regression coefficients into units of sample standard deviations.[31] These beta coefficients suggest that each one of the five measures of dispersion has a comparable impact to each one of the standard proximity-concentration variables.[32] For instance, a one standard deviation decline in an industry's freight costs raises the logarithm of the ratio of exports to FDI sales by 27 percent of a standard deviation; and a one standard deviation decline in the dispersion measures induce comparable changes in the dependent variable, with an average impact of 26 percent across the dispersion measures. The impact of tariffs is lower while the impact of returns to scale is higher. Taken as a whole, these results suggest that firm-level heterogeneity adds an important dimension to the observed trade-off between exports and FDI sales.

These results strongly support the theoretical model's predicted link between firm-level heterogeneity and the ratio of exports to FDI sales. Nevertheless, these results have to be interpreted with caution, because they may also reflect—at least to some degree—variations in industry characteristics that are not captured by our parsimonious model. This problem is partly taken care of by our control of cross-industry variations in capital and R&D intensities. Note that both these variables represent characteristics of an industry's technology that are not captured by our model.[33] Furthermore, as shown in Table 2, these measures of technology are correlated with all the different dispersion measures, although the correlations with the U.S.-data-based dispersion measure are rather weak.[34] Table 3 suggests that R&D intensity is not a useful predictor of exports relative to FDI

sales, while capital intensity is; more capital-intensive sectors export less relative to FDI sales. These results are interesting, but our theoretical model offers no guidance concerning their interpretation.[35]

Of course, differences in capital intensity may not be the only other source of variation across sectors that affects exports relative to FDI sales. In order to address the possibility that some other unmeasured characteristics of sectors fall into this category, we estimate the previous specification (with the capital and R&D intensity controls) adding random industry effects. A benefit of this estimation strategy is that it allows for efficient estimation in the presence of common components in the residuals that might be induced by unmeasured industry characteristics. To validate this specification, we need to assume that these unmeasured industry characteristics are uncorrelated with our right-hand-side variables. This is a strong assumption. We feel, however, that it is most likely to hold for our dispersion measures, which are the focus of the empirical analysis.[36]

The results are reported in Table 5. As could be predicted, the standard errors have increased. But so have the point estimates of the impact of dispersion on exports relative to FDI sales. Importantly, however, the coefficients for all the dispersion measures remain highly significant. On the other hand, the magnitude of the coefficients on FREIGHT and TARIFF are greatly reduced, and the coefficients on TARIFF are no longer significant. These results support our earlier conclusion that the economic significance of firm heterogeneity compares favorably with the significance of the standard proximity-concentration trade-off variables in explaining the export to FDI sales ratio.

Another robustness check addresses the potential interdependence of the residuals across countries, which may exist even after we con-

---

[31] See Jeffrey M. Wooldridge (2003, Sec. 6.1) for a further description of this transformation.

[32] In the case of FREIGHT, TARIFF, and FP, the coefficients are averaged across the five specifications.

[33] We have restricted our choice of controls to the measurable characteristics of sectors that are outside the scope of the model, and we have excluded attributes that are predicted to endogenously respond to changes in the model's exogenous variables.

[34] Capital intensity is measured as the industry's aggregate capital to labor ratio (from the NBER productivity database) and R&D intensity is measured as the ratio of R&D expenditures to sales (from a 1978 FTC survey).

[35] Pehr-Johan Norback (2001) finds that R&D-intensive firms tend to export rather than engage in FDI when technology transfer costs are high, but not when these costs are low. This suggests that we need a more detailed model in order to study the role of R&D in the proximity-concentration trade-off.

[36] The inclusion of industry fixed effects would eliminate the need for this assumption, but would also preclude any identification of sector-level characteristics, such as our dispersion measures.

TABLE 5—EXPORTS VERSUS FDI—RANDOM EFFECTS

	U.S. std. dev.	Europe std. dev.	France std. dev.	Europe reg. coeff.	France reg. coeff.
Narrow sample (N = 961)					
FREIGHT	−0.430	−0.398	−0.428	−0.397	−0.397
	(−2.554)	(−2.344)	(−2.533)	(−2.336)	(−2.334)
TARIFF	−0.113	−0.127	−0.105	−0.136	−0.133
	(−0.922)	(−1.033)	(−0.857)	(−1.107)	(−1.085)
FP	1.376	1.132	1.096	1.154	1.137
	(5.145)	(4.128)	(4.233)	(4.107)	(4.093)
DISPERSE	−2.623	−2.763	−2.445	−2.031	−1.991
	(−4.897)	(−3.459)	(−4.761)	(−3.098)	(−3.180)
KL	−1.106	−0.613	−0.570	−0.757	−0.758
	(−4.652)	(−2.238)	(−2.168)	(−2.891)	(−2.896)
RD	−0.002	0.126	0.116	0.133	0.119
	(−0.020)	(1.029)	(0.970)	(1.081)	(0.972)
R^2	0.352	0.316	0.342	0.307	0.308
Wide sample (N = 1,175)					
FREIGHT	−0.331	−0.322	−0.328	−0.320	−0.320
	(−2.296)	(−2.230)	(−2.278)	(−2.215)	(−2.215)
TARIFF	−0.004	−0.018	−0.004	−0.022	−0.021
	(−0.038)	(−0.155)	(−0.035)	(−0.193)	(−0.187)
FP	1.361	1.110	1.081	1.127	1.103
	(4.123)	(3.455)	(3.475)	(3.377)	(3.369)
DISPERSE	−2.518	−2.559	−2.265	−1.864	−1.786
	(−3.824)	(−2.733)	(−3.706)	(−2.398)	(−2.424)
KL	−1.069	−0.599	−0.561	−0.734	−0.739
	(−3.660)	(−1.871)	(−1.789)	(−2.373)	(−2.408)
RD	0.006	0.123	0.116	0.129	0.115
	(0.042)	(0.862)	(0.811)	(0.894)	(0.805)
R^2	0.312	0.274	0.297	0.267	0.267

Notes: T-statistics are in parentheses. Constant and country dummies are suppressed.

trol for country fixed effects. This type of interdependence pattern could be created by the ability of affiliates to re-export a portion of their production to a third country. In this case, a firm's decision to operate an affiliate in one country, say Belgium, would not be independent from its decision to locate affiliates in other neighboring European countries. In the Appendix to our working paper, Helpman et al. (2003), we show that the predicted link between firm-level heterogeneity within sectors and exports relative to FDI sales is theoretically consistent with an extended version of the model that explicitly allows for re-exports by affiliates. However, the pattern of interdependence may be particularly strong among the overrepresented and highly integrated economies of Western Europe.

To address this concern, we treated all the Western European countries as a single aggregate unit and reestimated our specification with the industry controls (capital and R&D intensities) and industry random effects. We found that all the dispersion measures remain highly significant. As could be predicted, the point estimates on the dispersion measures were slightly lower, which reflects the fact that the smaller developing countries now receive a greater weight in the sample.[37]

[37] See Table 8 of our working paper for these estimates.

TABLE 6—EXPORTS VERSUS FDI—ADDITIONAL ROBUSTNESS RESULTS
(Clustered standard errors and IV specifications)

	OLS			IV		
	Narrow sample	Wide sample	Aggregated Europe	Narrow sample	Wide sample	Aggregated Europe
FREIGHT	−1.040	−1.011	−1.001	−1.218	−1.118	−1.053
	(−3.997)	(−4.437)	(−4.464)	(−3.862)	(−4.366)	(−4.545)
TARIFF	−0.365	−0.241	−0.077	−0.188	−0.124	−0.016
	(−1.611)	(−1.081)	(−0.304)	(−0.706)	(−0.478)	(−0.056)
FP	1.177	1.133	1.086	1.609	1.457	1.344
	(4.876)	(4.472)	(4.166)	(3.605)	(3.311)	(3.182)
DISPERSE (U.S.)	−2.343	−2.248	−2.150	−4.321	−3.681	−3.198
	(−3.689)	(−3.655)	(−3.349)	(−2.606)	(−2.248)	(−2.217)
KL	−0.868	−0.793	−0.848	−0.938	−0.848	−0.890
	(−3.032)	(−2.513)	(−2.593)	(−2.855)	(−2.496)	(−2.599)
RD	−0.104	−0.086	−0.087	−0.158	−0.127	−0.121
	(−0.851)	(−0.687)	(−0.691)	(−1.233)	(−0.985)	(−0.968)
R^2	0.373	0.338	0.314	0.328	0.315	0.315
N	961	1,175	678	961	1,175	678

Notes: In the IV specifications, the U.S. dispersion measure is instrumented using all four European dispersion measures. All *T*-statistics (in parentheses) are computed from standard errors that are heteroskedasticity consistent and adjusted for clustering by industry. Constant and country dummies are suppressed.

Our final robustness check addresses sources of endogeneity bias in the dispersion measures, including measurement error. To address these concerns, we instrument the U.S. dispersion measure using all four European dispersion measures. We also use a different method to control for the potential correlation of the residuals within sectors by adjusting the standard errors for clustering (within sectors).[38] These specifications are reported in Table 6 for all previously discussed country samples (narrow, wide, and aggregated Europe). Instrumenting the U.S. dispersion measure significantly increases the magnitude of both the estimated coefficient and its standard error. However, as in all the previous specifications, the effect of dispersion on relative exports and FDI sales remains statistically significant.

Finally, we briefly report a number of other robustness checks. One potential complication arises from the fact that firms engage in intra-firm trade in intermediate inputs. This trade does not appear in our model, but is of sufficient size in a number of industries to be of concern.

---

[38] Under our assumptions on the source of this potential correlation in the residuals—unmeasured sectoral characteristics—the previously reported random-effects coefficients are the efficient estimators.

We found that netting out the value of these imports from our FDI sales data had no appreciable impact on the dispersion coefficients, although it had a small impact on the size of the FREIGHT and TARIFF coefficients. In other specifications, we included the four-firm concentration ratio as a control, in order to assess whether our measures of firm heterogeneity offer information in excess of this crude measure of concentration. We found that controlling for concentration reduces the point estimates of the coefficients on the dispersion measures, but that this decline is rather small.

IV. Conclusion

We have developed in this paper a model of international trade and investment in which firms can choose to serve their domestic market, to export, or to engage in FDI in order to serve foreign markets. Every industry is populated by heterogeneous firms, which differ in productivity levels. As a result, firms sort according to productivity into different organizational forms. The least productive firms leave the industry, because, if they stay, their operating profits will be negative no matter how they organize. Other low-productivity firms choose to serve only the domestic market. The remaining firms serve the

domestic market as well as foreign markets. Their mode of operation in foreign markets differs, however. The most productive firms in this group choose to invest in foreign markets while the less productive firms choose to export. This sorting pattern is confirmed by previous empirical work and by our own estimates.

Our model embodies standard elements of the proximity-concentration trade-off in the theory of horizontal foreign direct investment. As a result, it predicts that foreign markets are served more by exports relative to FDI sales when trade frictions are lower or economies of scale are higher. To these factors, our model adds a role for the within-sector heterogeneity of productivity levels. This heterogeneity induces a size distribution of firms, which affects the ratio of exports to FDI sales.

Using data on exports and FDI sales of U.S. firms in 38 countries and 52 industries, we estimated the effects of trade frictions, economies of scale, and within-industry dispersion of firm size, on exports versus FDI sales. The results support the theoretical predictions. In particular, they show a robust cross-sectoral relationship between the degree of dispersion in firm size and the tendency of firms to substitute FDI sales for exports. The size of this effect is of the same order of magnitude as trade frictions. We therefore conclude that we have identified a new element—namely, within-sectoral heterogeneity—that plays an important role in the structure of foreign trade and investment.

REFERENCES

Axtell, Robert L."Zipf Distribution of U.S. Firm Sizes." *Science,* September 2001, *293,* pp. 1818–20.

Bernard, Andrew B.; Eaton, Jonathan; Jenson, J. Bradford and Kortum, Samuel."Plants and Productivity in International Trade." *American Economic Review,* September 2003, *93*(4), pp. 1268–90.

Brainard, S. Lael."A Simple Theory of Multinational Corporations and Trade with a Trade-Off Between Proximity and Concentration." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 4269, February 1993.

———. "An Empirical Assessment of the Proximity-Concentration Trade-off Between Multinational Sales and Trade." *American Economic Review,* September 1997, *87*(4), pp. 520–44.

Budd, John W.; Konings, Jozef and Slaughter, Matthew J."International Rent Sharing in Multinational Firms." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 8809, February 2002.

Doms, Mark E. and Jensen, J. Bradford."Comparing Wages, Skills, and Productivity between Domestically and Foreign-Owned Manufacturing Establishments in the United States," in Robert E. Baldwin, Robert E. Lipsey, and J. David Richardson, eds., *Geography and ownership as bases for economic accounting.* NBER Studies in Income and Wealth, Vol. 59. Chicago: University of Chicago Press, 1998, p. 235–55.

Ethier, Wilfred J."The Multinational Firm." *Quarterly Journal of Economics,* November 1986, *101*(4), pp. 805–33.

Ethier, Wilfred J. and Markusen, James R."Multinational Firms, Technology Diffusion and Trade." *Journal of International Economics,* August 1996, *41*(1–2), pp. 1–28.

Feenstra, Robert C."U.S. Exports, 1972–1994: With State Exports and Other U.S. Data." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 5990, April 1997.

Girma, Sourafel; Thompson, Steve and Wright, Peter W."Why Are Productivity and Wages Higher in Foreign Firms?" *Economic and Social Review,* Spring 2002, *33*(1), pp. 93–100.

Hanson, Gordon H.; Mataloni, Raymond J. and Slaughter, Matthew J."Vertical Specialization in Multinational Firms." Mimeo, University of California, San Diego, 2002.

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

———. "Multinational Corporations and Trade Structure." *Review of Economic Studies,* July 1985, *52*(3), pp. 443–57.

Helpman, Elhanan; Melitz, Marc J. and Yeaple, Stephen R."Export Versus FDI." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 9439, January 2003.

Horstmann, Ignatius and Markusen, James R."Licensing versus Direct Investment: A