

# DISCUSSION PAPER SERIES

No. 3741

## EXPORT VERSUS FDI

Elhanan Helpman, Marc J Melitz  
and Stephen R Yeaple

*INTERNATIONAL TRADE*



**C**entre for **E**conomic **P**olicy **R**esearch

[www.cepr.org](http://www.cepr.org)

Available online at:

[www.cepr.org/pubs/dps/DP3741.asp](http://www.cepr.org/pubs/dps/DP3741.asp)

# EXPORT VERSUS FDI

**Elhanan Helpman**, Tel Aviv University and CEPR  
**Marc J Melitz**, Harvard University and CEPR  
**Stephen R Yeaple**, University of Pennsylvania

Discussion Paper No. 3741  
February 2003

Centre for Economic Policy Research  
90–98 Goswell Rd, London EC1V 7RR, UK  
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999  
Email: [cepr@cepr.org](mailto:cepr@cepr.org), Website: [www.cepr.org](http://www.cepr.org)

This Discussion Paper is issued under the auspices of the Centre's research programme in **INTERNATIONAL TRADE**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as a private educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions. Institutional (core) finance for the Centre has been provided through major grants from the Economic and Social Research Council, under which an ESRC Resource Centre operates within CEPR; the Esmée Fairbairn Charitable Trust; and the Bank of England. These organizations do not give prior review to the Centre's publications, nor do they necessarily endorse the views expressed therein.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Elhanan Helpman, Marc J Melitz and Stephen R Yeaple

## ABSTRACT

### Export versus FDI\*

This Paper builds a multi-country, multi-sector general equilibrium model that explains the decision of heterogeneous firms to serve foreign markets either through exports or local subsidiary sales (FDI). These modes of market access involve different relative costs, some of which are sunk while others vary with sales volume (such as transport costs and tariffs). Relative to investment in a subsidiary, exporting involves lower sunk costs but higher per-unit costs. In equilibrium, only the more productive firms choose to serve the foreign markets and the most productive among this group will further choose to serve the overseas market via FDI. The Paper then explores several implications of the individual firms' decisions for aggregate export and FDI sales relative to the domestic and foreign market sizes. In particular, it is shown that firm-level heterogeneity is an important determinant of relative export and FDI flows.

We use the model to derive testable empirical predictions on the relative aggregate export and FDI sales in a given country for a given sector based both on relative costs and the extent of firm level heterogeneity in that sector. These predictions are tested on data of US affiliate sales and US exports in 38 different countries and 52 sectors. The comparative statics based on relative costs are very similar to those tested by Brainard (*AER* 1997) and are confirmed in our data: sector/country specific transport costs and tariffs have a strong negative effect on export sales relative to FDI. More importantly, our new predictions for the effects of firm-level heterogeneity on the relative export and FDI sales are also strongly supported by the data: more heterogeneity leads to significantly more FDI sales relative to export sales.

JEL Classification: F12, F14, F23 and L11

Keywords: FDI, size distribution and trade

Elhanan Helpman  
Eitan Berglas School of Economics  
Tel Aviv University  
IL-69978 Tel Aviv  
ISRAEL  
Tel: (972 3) 641 0194  
Fax: (972 3) 642 8074  
Email: helpman@post.tau.ac.il

Marc J Melitz  
Department of Economics  
Harvard University  
Littauer Center 120  
Cambridge, MA 02138-3001  
USA  
Tel: (1 617) 495 8297  
Fax: (1 617) 417 6536  
Email: mmelitz@harvard.edu

For further Discussion Papers by this author see:  
[www.cepr.org/pubs/new-dps/dplist.asp?authorid=100511](http://www.cepr.org/pubs/new-dps/dplist.asp?authorid=100511)

For further Discussion Papers by this author see:  
[www.cepr.org/pubs/new-dps/dplist.asp?authorid=151031](http://www.cepr.org/pubs/new-dps/dplist.asp?authorid=151031)

Stephen R Yeaple  
Department of Economics  
University of Pennsylvania  
3718 Locust Walk  
Philadelphia, PA 19104-6297  
USA  
Email: snyeapl12@econ.upenn.edu

For further Discussion Papers by this author see:  
[www.cepr.org/pubs/new-dps/dplist.asp?authorid=158590](http://www.cepr.org/pubs/new-dps/dplist.asp?authorid=158590)

\*The statistical analysis of firm-level data on US Multinational Corporations reported in this study was conducted at the International Investment Division, US Bureau of Economic Analysis, under arrangement that maintained legal confidentiality requirements. Views expressed are those of the authors and do not necessarily reflect those of the Bureau of Economic Analysis. Elhanan Helpman thanks the NSF for financial support. We also thank Daron Acemoglu, Roberto Rigobon, Yona Rubinstein and Dani Tsiddon for comments on an earlier draft, and Man-Keung Tang for excellent research assistance.

Submitted 24 December 2002

# 1 Introduction

Multinational sales have grown tremendously in the last two decades. Growth of these sales has even outpaced 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 service foreign buyers through a variety of channels: they can export their products to foreign customers, serve them through foreign subsidiaries by engaging in foreign direct investment (FDI), and license or contract with foreign firms to produce and sell their products.

Our work focuses on the firm’s choice between exports and “horizontal” FDI.<sup>1</sup> Horizontal FDI refers to investments in production facilities abroad that are designed to serve foreign customers. We therefore exclude “vertical” motives for FDI, that involve fragmentation of production across countries.<sup>2</sup> We follow the previous literature on horizontal FDI in assuming that foreign affiliate production is intended for the host country market.<sup>3</sup> However, we show in an appendix how our model can be extended to incorporate exports by foreign affiliates. This adds a new motive for FDI: the use of affiliates as “export platforms.” In all these cases, however, firms invest abroad when the gains from avoiding transport costs outweigh the costs of maintaining capacity in multiple markets. This is known as the proximity–concentration tradeoff.

We extend the proximity–concentration tradeoff literature by introducing intra-industry firm heterogeneity. We build a simple multicountry, multisector general equilibrium model that explains the decisions of heterogeneous firms to serve foreign markets through exports or local subsidiary sales. These modes of market access have different relative costs, some of which are sunk (such as entry costs) while others vary

---

<sup>1</sup>See Ethier (1986), Horstmann and Markusen (1987), and Ethier and Markusen (1996) for models that incorporate the licensing alternative.

<sup>2</sup>An example of vertical FDI is when a parent firm invests in a production facility in another country in order to produce inputs that will be shipped back to the parent for further processing. Or when a parent firm produces only inputs in the home country and it invests in an assembly facility in another country to which it ships these inputs. See Helpman (1984), Helpman (1985), Markusen (2002, Ch. 9), and Hanson, Mataloni and Slaughter (2002) for treatments of this form of FDI.

<sup>3</sup>See, for example, Horstmann and Markusen (1992), Brainard (1993), and Markusen and Venables (2000).

with sales (such as transport costs and tariffs). Relative to FDI, exporting involves lower sunk costs but higher per-unit costs.<sup>4</sup>

We show that firm heterogeneity in productivity plays an important role in explaining the structure of international commerce. 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 (2002).<sup>5</sup> Second, of those firms that serve foreign markets, only the most productive engage in FDI.<sup>6</sup> Third, the extent of intra-industry firm heterogeneity plays a key role in determining the volume of FDI sales relative to the volume of exports. Hence, we identify a new industry characteristic — the dispersion of productivity levels across firms — as a determinant of the composition of trade.

This allows us to derive new insights from the proximity–concentration tradeoff. First, standard proximity–concentration variables determine the productivity levels that a firm must achieve to make its international activities attractive. But these productivity levels do not fully determine the composition of trade; they have to be combined with information about the degree of dispersion of productivity levels across firms within every industry in order to forecast trade flows.<sup>7</sup> Second, our model avoids the knife-edge conditions associated with existing general equilibrium models of FDI that are based on representative firms. In a typical model from this family, exogenous

---

<sup>4</sup>Sunk costs associated with exporting allow the model to explain two important empirical patterns: the existence of substantial subsets of firms within every manufacturing sector that do not engage in any form of international commerce, and the existence of large numbers of foreign wholesale affiliates whose main activity is to redistribute the output manufactured by the parent firm. Although such firms are technically multinationals, the foreign affiliates do not duplicate the production process. In the context of our model, we characterize such firms as exporters who incur fixed distribution costs in the destination country. In our empirical work, we exclude the sales of these wholesale affiliates from our measure of FDI sales.

<sup>5</sup>See also Bernard, Eaton, Jenson and Kortum (2000) for an alternative theoretical model and Tybout (2002) for a survey of the empirical literature.

<sup>6</sup>This result is loosely connected to the documented empirical pattern that foreign-owned affiliates are more productive than domestically-owned producers. See Doms and Jensen (1998) for the U.S. and Girma, Thompson and Wright (2002) for the U.K.

<sup>7</sup>Our model formalizes the old idea that multinational firms must have some form of ownership advantage conferred by access to firm-specific intangible assets (for a discussion of this literature, see Markusen (1995)). In our model this intangible asset takes the form of a superior production technology. Our analysis takes this idea further, by allowing industry characteristics — such as transport costs — to govern the extent of an ownership advantage needed to induce a firm to become multinational and by positing a distribution of these assets within an industry.

industry characteristics mandate that either all firms invest abroad or that none does.<sup>8</sup> In contrast, in our model firm heterogeneity plays a central role in pinning down the number of firms that export and the number of firms that invest abroad. This provides a more appealing and realistic explanation for the concomitant use of exports and FDI sales. Finally, our model identifies a new “home market bias,” whereby the number of firms that locate their headquarters in a particular country rises disproportionately with the country’s size; while small markets are disproportionately served by affiliates of multinational companies and by exporters from other countries.

We test the predictions of the model on U.S. exports and affiliate sales data that cover 52 manufacturing industries and 38 countries. We show that the productivity dispersion measures help to predict the composition of trade and investment in the manner suggested by the model. Industries in which productivity levels vary highly across firms are characterized by a larger volume of FDI sales relative to exports. We show that these results are robust across several measures of productivity dispersion. In addition, we confirm the predictions of the proximity–concentration tradeoff. We find that firms tend to substitute FDI sales for exports when transport costs are relatively high and when plant-level returns to scale are relatively weak. We conclude that intra-industry firm heterogeneity plays an important role in determining the composition of international trade.

The remainder of this paper is composed of three sections. In section 2 we elaborate the model and characterize its equilibrium. We then map the theoretical results onto an empirical strategy for testing our main hypotheses concerning the role played by firm heterogeneity in the proximity–concentration tradeoff between exports and FDI sales. In section 3 we describe the data. Finally, we report and interpret the empirical results in section 4.

---

<sup>8</sup>Only on the “knife-edge” can firms that export and firms that do FDI coexist.

## 2 Theoretical Framework

There are  $N$  countries that use labor to produce goods in  $H + 1$  sectors. One sector produces a homogeneous product while  $H$  sectors produce differentiated products. A fraction  $\beta_h$  of income is spent on differentiated products of sector  $h$  and a fraction  $1 - \sum_h \beta_h$  on the homogeneous good, which is our numeraire.

Country  $i$  is endowed with  $L^i$  units of labor. We take  $\sum_h \beta_h$  to be small enough and differences in  $L^i$  to be small enough so that the homogeneous product is produced in every country and wages are equalized across countries.<sup>9</sup> The homogeneous product is produced with one unit of labor per unit output. As a result, the common wage rate equals one.

Now consider a particular sector  $h$  that produces differentiated products. For the time being we drop the index  $h$ , letting it be understood 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, 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.<sup>10</sup> Goods that are exported from country  $i$  to country  $j$  are subjected to melting-iceberg transport

---

<sup>9</sup>We will show later the precise restrictions on the cross-country variation in  $L^i$  that are needed for this outcome. Our empirical work will not rely on this assumption.

<sup>10</sup>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.



costs  $\tau^{ij} > 1$ . Namely,  $\tau^{ij}$  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  $\varepsilon = 1/(1 - \alpha) > 1$ .<sup>11</sup> 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.<sup>12</sup> In this case, the brand of a monopolistic producer with labor coefficient  $a$  is offered for sale at the price  $p = a/\alpha$ , where  $1/\alpha$  represents the markup factor. As a result, the effective consumer price is  $a/\alpha$  for domestically produced goods — be they supplied by a domestic producer or a foreign affiliate (with the labor coefficient  $a$ ) — and is  $\tau^{ji} a/\alpha$  for imported products from country  $j$  (from exporters with the labor coefficient  $a$ ). Imported products are thus more expensive than domestically produced goods due to transport costs.

A firm from country  $i$  that remains in the industry will always serve its domestic market through domestic production. It may also serve any foreign market  $j$ . If so, it also chooses a channel to access this foreign market: exports via domestic production or local sales via affiliate production (FDI). This choice is driven by the proximity–concentration tradeoff: relative to exports, FDI saves transport costs, but duplicates production facilities and therefore requires higher fixed costs.<sup>13</sup> In equilibrium, no firm

---

<sup>11</sup>The utility function is

$$u = \left(1 - \sum_{h=1}^H \beta_h\right) \log z + \sum_{h=1}^H \frac{\beta_h}{\alpha_h} \log \left( \int_{v \in V_h} x_h(v)^{\alpha_h} dv \right),$$

where  $z$  is consumption of the homogenous good,  $x_h(v)$  is consumption of variety  $v$  from sector  $h$ , and  $V_h$  is the set of all potential varieties in sector  $h$ .

<sup>12</sup>As is well known, our utility function implies that

$$A^i = \frac{\beta E^i}{\int_0^{n^i} p^i(v)^{1-\varepsilon} dv},$$

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  $v$ .

<sup>13</sup>We exclude the possibility of exports by foreign affiliates. See, however, the appendix for a discussion of this possibility.

engages in both activities for the same foreign market.<sup>14</sup> We assume

$$f_I > (\tau^{ij})^{\varepsilon-1} f_X > f_D. \quad (1)$$

We shall clarify the role of these conditions in the following analysis.

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  $B^i = (1 - \alpha) A^i / \alpha^{1-\varepsilon}$ .<sup>15</sup> 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 in which the demand levels are the same in countries  $i$  and  $j$ .<sup>16</sup> 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 linear functions of this index. More productive firms are therefore more profitable in all three activities.

---

<sup>14</sup>In a dynamic model with uncertainty, an individual firm may choose to serve a foreign market through both exports and FDI. Rob and Vettas (2001) provide a rigorous treatment of this case.

<sup>15</sup>Note that the demand function  $A^i p^{-\varepsilon}$  implies output  $A^i (a/\alpha)^{-\varepsilon}$  when the price is  $a/\alpha$ . 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

$$\pi_D^i = (1 - \alpha) A^i (a/\alpha)^{1-\varepsilon} - f_D.$$

<sup>16</sup>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.

The slope of  $\pi_D^i$  equals  $B^i$  while the slope of  $\pi_I^j$  equals  $B^j$ . These profit functions are parallel to each other when the demand levels are the same in countries  $i$  and  $j$ . Profits from FDI are lower, however, because the fixed costs of FDI,  $f_I$ , are higher than the fixed costs of domestic production,  $f_D$ . The slope of  $\pi_X^{ij}$  equals  $(\tau^{ij})^{1-\varepsilon} B^j$ , which is smaller than the slope of  $\pi_I^j$ . Together with the first inequality in (1), these relationships imply that exports are more profitable 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, namely,  $(a_I^{ij})^{1-\varepsilon} > (a_X^{ij})^{1-\varepsilon}$ , which ensures that some firms export to country  $j$ . In addition, the second inequality in (1) implies that  $(a_X^{ij})^{1-\varepsilon} > (a_D^i)^{1-\varepsilon}$ , 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 the market in 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  $(a_I^j)^{1-\varepsilon}$  gain more from FDI. For this reason, firms with productivity levels between  $(a_X^{ij})^{1-\varepsilon}$  and  $(a_I^j)^{1-\varepsilon}$  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  $(a_I^j)^{1-\varepsilon}$  are determined by

$$(a_D^i)^{1-\varepsilon} B^i = f_D \quad \text{for all } i, \quad (2)$$

$$(\tau^{ij} a_X^{ij})^{1-\varepsilon} B^j = f_X \quad \text{for all } j \neq i, \quad (3)$$

$$(1 - \tau^{1-\varepsilon}) (a_I^j)^{1-\varepsilon} B^j = f_I - f_X \quad \text{for all } j \neq i. \quad (4)$$

Free entry ensures equality between the expected operating profits of a potential

entrant and the entry costs  $f_E$ . This condition can be expressed as<sup>17</sup>

$$V(a_D^i) B^i + \sum_{j \neq i} \left[ 1 - (\tau^{ij})^{1-\varepsilon} \right] V(a_I^{ij}) B^j + \sum_{j \neq i} (\tau^{ij})^{1-\varepsilon} V(a_X^{ij}) B^j - \left[ G(a_D^i) f_D + \sum_{j \neq i} G(a_I^{ij}) (f_I - f_X) + \sum_{j \neq i} G(a_X^{ij}) f_X \right] = f_E \quad \forall i, \quad (5)$$

where

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

Equations (2)-(5) 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. Evidently, these solutions do not depend on the country size variables  $L^i$ , as long as the variation in country size is not large enough to induce some countries to specialize in differentiated products. Moreover, it is easy to see that 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 our empirical application.

## 2.1 Solving the Full General Equilibrium Model: A Special Case

In order to build the intuition behind our model and its empirical predictions, we first examine a special case that exploits some symmetry across countries. Assume, for this purpose, that all fixed cost coefficients are the same in every country, that the distribution function  $G(\cdot)$  is the same in every country, and that transport costs per product are the same for every pair of countries. The latter assumption means that  $\tau^{ij} = \tau > 1$  for every  $j \neq i$ . These restrictions are within every sector, so that there

---

<sup>17</sup>The expected operating profits of a potential entrant are

$$\int_0^{a_D^i} (a^{1-\varepsilon} B^i - f_D) dG(a) + \sum_{j \neq i} \left\{ \int_{a_I^{ij}}^{a_X^{ij}} [(\tau^{ij} a)^{1-\varepsilon} B^j - f_X] dG(a) + \int_0^{a_I^{ij}} (a^{1-\varepsilon} B^j - f_I) dG(a) \right\} .$$

Using this expression and the definition of the function  $V(a)$  in (6), we obtain the free entry condition (5). Note that the expected operating profits can be smaller than the entry costs in some sectors, in which case no domestic firm would enter that industry. This is possible in a trading/investment equilibrium where consumers satisfy their consumption demand with foreign goods that are produced by either foreign exporters or by local subsidiaries of multinational corporations.

can be variations in these characteristics across sectors. Moreover, countries can differ in size.

Under these circumstances, the equilibrium system (2)-(5) implies the same cutoffs  $a_D^i = a_D$ ,  $a_X^{ij} = a_X$ ,  $a_I^{ij} = a_I$  and the same  $B^i = B$  for every  $i, j$ . They are the solution to

$$a_D^{1-\varepsilon} B = f_D, \quad (7)$$

$$(\tau a_X)^{1-\varepsilon} B = f_X, \quad (8)$$

$$(1 - \tau^{1-\varepsilon}) a_I^{1-\varepsilon} B = f_I - f_X, \quad (9)$$

$$\begin{aligned} V(a_D) B + (N - 1) (1 - \tau^{1-\varepsilon}) V(a_I) B + \tau^{1-\varepsilon} (N - 1) V(a_X) B \\ - [G(a_D) f_D + (N - 1) G(a_I) (f_I - f_X) + (N - 1) G(a_X) f_X] = f_E. \end{aligned} \quad (10)$$

Figure 2 describes the distribution of labor productivity that is induced by  $G(a)$ . It also describes the equilibrium cutoffs. The figure is the same for every country. The fraction of surviving firms is given by the area below the curve to the right of  $1/a_D$ . The area between  $1/a_D$  and  $1/a_X$  represents the fraction of entrants who serve only the domestic market. The fractions of entrants who export or invest in foreign countries are represented, respectively, by the area between  $1/a_X$  and  $1/a_I$ , and by the area above  $1/a_I$ .

Having solved for the cutoffs and the demand level  $B$ , we can then determine the number of entrants in every country as a function of country size. To characterize the number of entrants in country  $i$  (in sector  $h$ ), note that

$$B = \frac{1 - \alpha}{\alpha^{1-\varepsilon}} A = \frac{(1 - \alpha) \beta E^i}{\alpha^{1-\varepsilon} \int_0^{n^i} p^i(v)^{1-\varepsilon} dv}, \quad (11)$$

where  $E^i$  is the aggregate level of spending in country  $i$ ,  $p^i(v)$  is the consumer price of variety  $v$  in country  $i$ , and  $n^i$  is the number of brands available to consumers in country  $i$ . Since there are no pure profits in equilibrium, spending equals labor income:  $E^i = L^i$ . It then follows from (11) that the numbers of entrants in country  $i$ ,  $n_E^i$ ,  $i = 1, 2, \dots, N$ ,

are the solution of the linear system

$$\begin{bmatrix} v_D & v_{IX} & \cdots & v_{IX} \\ v_{IX} & v_D & \ddots & \vdots \\ \vdots & \ddots & \ddots & v_{IX} \\ v_{IX} & \cdots & v_{IX} & v_D \end{bmatrix} \begin{bmatrix} n_E^1 \\ n_E^2 \\ \vdots \\ n_E^N \end{bmatrix} = \frac{(1-\alpha)\beta}{B} \begin{bmatrix} L^1 \\ L^2 \\ \vdots \\ L^N \end{bmatrix},$$

where  $v_D = V(a_D) > v_{IX} = V(a_I) + \tau^{1-\varepsilon} [V(a_X) - V(a_I)]$ .<sup>18</sup> So long as the differences in the  $L^i$ 's are not too large, the number of entrants that solve this system is positive in every country; it is given by

$$n_E^i = \frac{(1-\alpha)\beta}{B \det(v)} \left\{ [(N-1)v_{IX} + v_D] L^i - v_{IX} \sum_j L^j \right\}, \quad (12)$$

where  $\det(v)$  is the determinant of the matrix  $v$  that has  $v_D$  as the diagonal elements and  $v_{IX}$  as the off-diagonal elements. Since  $v_D > v_{IX} > 0$ , this determinant must be positive. Evidently,  $n_E^i$  is positive when all countries are of equal size. We assume that the differences in size are small enough so that  $n_E^i > 0$  for every country.<sup>19</sup>

---

<sup>18</sup>Recall that all brands that are produced in country  $i$ , by domestic firms or by foreign subsidiaries, have a consumer price of  $a/\alpha$  when the producer's labor cost is  $a$  per unit output, and all imported brands have a consumer price of  $\tau a/\alpha$  when the exporter's labor cost is  $a$  per unit output. Therefore,

$$\begin{aligned} \int_0^{n^i} p^i(v)^{1-\varepsilon} dv &= n_E^i \int_0^{a_D} \left(\frac{a}{\alpha}\right)^{1-\varepsilon} dG(a) + \sum_{j \neq i} n_E^j \left[ \int_0^{a_I} \left(\frac{a}{\alpha}\right)^{1-\varepsilon} dG(a) + \int_{a_I}^{a_X} \left(\frac{\tau a}{\alpha}\right)^{1-\varepsilon} dG(a) \right] \\ &= \frac{v_D n_E^i + v_{IX} \sum_{j \neq i} n_E^j}{\alpha^{1-\varepsilon}} \end{aligned}$$

<sup>19</sup>Namely, we assume that

$$\frac{L^i}{\sum_j L^j} > \frac{v_{IX}}{(N-1)v_{IX} + v_D} \quad \forall i.$$

Since  $v_D > v_{IX}$ , the right-hand side of this inequality is smaller than  $1/N$ , and therefore the inequality is satisfied when  $L^i$  is the same in all countries.

This assumption is not essential, however. Without it, the number of entrants is positive for the largest countries and zero for the smaller countries. The arguments that follow then apply only to the set of countries with positive entry.

### 2.1.1 Home Market Effects

Equation (12) implies that more firms enter in larger countries and that  $n^i/L^i > n^j/L^j$  for  $L^i > L^j$ ; that is, in a cross-country comparison, the number of entrants rises more than proportionately with country size. Since the cutoff coefficients  $a_l$ ,  $l = D, X, I$ , and the distribution function  $G(\cdot)$  are the same in all countries, the number of firms that exit, the number of firms that serve only the domestic market, the number of firms that export, and the number of firms that invest in foreign countries are all proportional to the number of entrants. In addition, the demand level coefficient  $B$  is also the same in all countries. Therefore aggregate sales of country- $i$ -based firms are proportional to  $n_E^i$ . Moreover, their sales in the domestic market are proportional to  $n_E^i$ , their exports are proportional to  $n_E^i$ , and foreign sales of their multinational subsidiaries are also proportional to  $n_E^i$ . It follows that larger countries have proportionately larger sales in each one of these categories.

Now define  $n_O^i = G(a_D)n_E^i$  as the number of active firms based in country  $i$  and  $n_B^i = G(a_X) \sum_{j \neq i} n_E^j + G(a_D)n_E^i = G(a_X) \sum_{j=1}^N n_E^j + [G(a_D) - G(a_X)] n_E^i$  as the number of firms doing business in country  $i$ . The ratio  $n_O^i/n_B^i$  is then higher in larger countries. That is, the larger a country the larger the number of its active firms relative to the number of firms that operate in the country. Also note that  $n_B^i$  is larger in larger countries, which implies that consumers in larger countries enjoy a larger product variety range and higher welfare.<sup>20</sup>

Next consider relative market shares. Let  $\sigma_D^i$  be the market share of domestic firms in country  $i$ , let  $\sigma_X^i$  be the market share of foreign exporters in country  $i$ , and let  $\sigma_I^i$  be the market share of foreign multinationals in country  $i$  (these shares must sum to

---

<sup>20</sup>The country-level price index  $\left( \int_0^{n^i} p^i(v)^{1-\varepsilon} dv \right)^{1/(1-\varepsilon)}$  is lower in bigger countries. This result is driven by the effect of country size on variety range.

1). Then

$$\begin{aligned}\sigma_D^i &= \frac{BV(a_D) n_E^i}{(1-\alpha)\beta L^i}, \\ \sigma_X^i &= \frac{B\tau^{1-\varepsilon} [V(a_X) - V(a_I)] \sum_{j \neq i} n_E^j}{(1-\alpha)\beta L^i}, \\ \sigma_I^i &= \frac{BV(a_I) \sum_{j \neq i} n_E^j}{(1-\alpha)\beta L^i}.\end{aligned}$$

It follows that the larger is country  $i$ , the larger the market share of its firms in the domestic market, and the smaller the market share of foreign exporters and foreign multinationals. Moreover, the market shares of foreign exporters and foreign multinationals are equi-proportionately smaller, because  $\sigma_X^i/\sigma_I^i$  is independent of country size.

### 2.1.2 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

$$\frac{s_X^{ij}}{s_I^{ij}} = \tau^{1-\varepsilon} \left[ \frac{V(a_X)}{V(a_I)} - 1 \right]. \quad (13)$$

Given our symmetry assumptions on technologies and international transaction costs, this ratio is independent of  $i$  and  $j$ : 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 (7)-(10). This system can therefore be used to assess the consequences of changes in the costs of international transactions  $f_I$ ,  $f_X$ , and  $\tau$ .

A rise in the export costs  $f_X$  or  $\tau$ , 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  $a_X$ .<sup>21</sup> 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, system (7)-(10) can be used to 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 tradeoff results can be extended to sectoral levels. In a departure from the previous literature, however, these predictions are now based on the aggregation across heterogeneous firms that select different modes of foreign market access (FDI or exports).

We now shift the focus to the role of firm-level heterogeneity in explaining the cross-sectoral variation in relative export sales. Note from (13) 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  $\varepsilon$ , which magnifies

---

<sup>21</sup>Given (7)-(9), it can be shown that shifts in the cutoffs  $a_l$ ,  $l = D, X, I$ , have no first-order effect on equation (10). Therefore (10) can be used to directly calculate the shifts in  $B$  in response to changes in any of the exogenous parameters. An increase in  $\tau$  therefore induces an increase in  $B$  and a decrease in  $\tau^{1-\varepsilon}B$ . The direction of the change in the cutoffs  $a_X$  and  $a_I$  is then immediate from (8) and (9). Similarly, an increase in  $f_X$  will induce an increase in  $B$  and a decrease in  $B/f_X$ ; a decrease in  $f_I$  will induce a decrease in  $B$  and an increase in  $B/(f_I - f_X)$ . The effects of these changes on the cutoffs are then once more directly obtained from (8) and (9).

differences in firm-level outcomes that are induced by  $G(a)$ .

In order to index differences in firm-level heterogeneity across sectors, we parametrize  $V(a)$  by parametrizing the distribution  $G(a)$ . 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  $k - (\varepsilon - 1)$ .<sup>22</sup> The shape parameter of the Pareto distribution offers a natural and convenient index of dispersion, which we will use to characterize 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  $\varepsilon$ , 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 productivity-draws. Relative export sales in (13) can then be written as<sup>23</sup>

$$\begin{aligned} \frac{s_X^{ij}}{s_I^{ij}} &= \tau^{1-\varepsilon} \left[ \left( \frac{a_X}{a_I} \right)^{k-(\varepsilon-1)} - 1 \right] \\ &= \tau^{1-\varepsilon} \left[ \left( \frac{f_I - f_X}{f_X} \frac{1}{\tau^{\varepsilon-1} - 1} \right)^{\frac{k-(\varepsilon-1)}{\varepsilon-1}} - 1 \right]. \end{aligned} \quad (14)$$

Comparative statics on (14) predict that relative export sales decrease with decreases in  $k$  and increases in  $\varepsilon$ .<sup>24</sup> Thus, we expect sectors with higher levels of dispersion in

---

<sup>22</sup>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.

<sup>23</sup>Equations (8) and (9) are used in this derivation.

<sup>24</sup>Recall that  $(f_I - f_X)/f_X (\tau^{\varepsilon-1} - 1)$  is greater than 1, by assumption; see (1).

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 our model. It highlights the importance of firm-level heterogeneity, which we will test.

## 2.2 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 section 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:

$$\left(\tau_h^{Uj} w^U a_{hX}^{Uj}\right)^{1-\varepsilon_h} B_h^j = w^j f_X^j, \quad (15)$$

$$\left[ (w^j)^{1-\varepsilon_h} - \left(w^U \tau_h^{Uj}\right)^{1-\varepsilon_h} \right] \left(a_{hI}^{Uj}\right)^{1-\varepsilon_h} B_h^j = w^j (f_{hI}^j - f_X^j), \quad (16)$$

where  $w^U$  and  $w^j$  are the wage levels in the U.S. and country  $j$ ,  $\tau_h^{Uj}$  is the trade cost (transport and tariff) from the U.S. to country  $j$  in sector  $h$ ,  $\varepsilon_h$  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 (8) and (9). 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. (15) and (16) then imply:

$$\left(\frac{a_{hX}^{Uj}}{a_{hI}^{Uj}}\right)^{\varepsilon_h-1} = \frac{f_{hP}}{f_X^j} \left[ \left(\omega^j \tau_h^{Uj}\right)^{\varepsilon_h-1} - 1 \right]^{-1}, \quad (17)$$

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

We further assume the following conditions on relative wages and trade costs:

- $w^U \tau_h^{Uj}/w^j < (1 + f_{hP}/f_X^j)^{1/(\varepsilon_h-1)} = (f_{hI}^j/f_X^j)^{1/(\varepsilon_h-1)}$ , which ensures that there exist U.S. firms that prefer export to FDI in country  $j$ ;
- $w^U \tau_h^{Uj}/w^j > 1$ , which ensures that there exist firms that choose to locate in country  $j$ ; and
- $w^j \tau_h^{jU}/w^U > 1$ , which ensures that there exist firms that choose to locate in the U.S.<sup>25</sup>

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

$$\begin{aligned} \frac{S_X^{Uj}}{S_I^{Uj}} &= \left(\omega^j \tau_h^{Uj}\right)^{1-\varepsilon_h} \left[ \frac{V_h^U(a_{hX}^{Uj})}{V_h^U(a_{hI}^{Uj})} - 1 \right] \\ &= \left(\omega^j \tau_h^{Uj}\right)^{1-\varepsilon_h} \left\{ \left[ \frac{f_{hP}}{f_X^j} \frac{1}{\left(\omega^j \tau_h^{Uj}\right)^{\varepsilon_h-1}} - 1 \right]^{k_h^U - (\varepsilon_h - 1)} - 1 \right\}. \end{aligned} \quad (18)$$

Comparing (14) and (18) confirms that all our previously derived comparative statics remain valid in a cross-section of both sectors and non-symmetric countries: the proximity–concentration forces predict lower U.S. relative export sales for country-sector pairs with high transport costs  $\tau_h^{Uj}$ , 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

---

<sup>25</sup>The relative wage  $w^U/w^j$  must be measured in effective units of labor (adjusted for productivity and human capital differences). In our sample of countries the differences in productivity adjusted relative wages are small. In any case, our second and third conditions ensure that the relative wages are bounded by transport costs.

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  $\varepsilon_h$ . 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  $h$ .

### 3 Data

To test our multisector, multicountry model, we require data that varies in both of these dimensions. The data required fall into roughly three categories: data on the composition of international commerce across countries and sectors, measures for key variables affecting the proximity–concentration tradeoff, and indices capturing differences in the extent of firm-level heterogeneity across sectors. In this section, we describe our choice of data in this order. Unless otherwise noted, all of the data described below are for the single year 1994.

#### 3.1 The Composition of International Commerce

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

In the United States, the organization that collects census-type data on FDI is the Bureau of Economic Analysis (BEA). 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 classifications, which are shown in table 1. To make this FDI data comparable to the data for exports, we aggregated the firm-level multinational sales data to the industry level. Our export data are more familiar and have been taken from Feenstra (1997). The data have been concorded from 4-digit

SITC industrial classifications into the BEA industry classifications shown in table 1.

Finally, we consider two separate samples of countries, which can roughly be characterized as narrow and wide. The narrow sample consists of the 27 countries originally considered by Brainard (1997) while the wide sample includes 11 additional countries, which are smaller and typically less developed. The country coverage is shown in table 2. 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 some concern about censoring.

### **3.2 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. These costs are not easily quantified, however. We therefore need to discuss our choice of proxies in some detail.

First consider unit costs of foreign trade. These costs can be either due to the costs of moving goods across borders, such as transport and insurance, or due to barriers created by destination-country governments, such as tariffs. We proxy for them with the variables FREIGHT and TARIFF, respectively, where FREIGHT is an ad-valorem measure of freight and insurance costs and TARIFF is an ad-valorem measure of trade taxes. FREIGHT is computed as the ratio of CIF imports into the United States to FOB imports, which is calculated from the data presented 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 sub-industries within the BEA industry. Trade taxes are taken from Yeaple (2000), 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 foreign direct investment. In principle, these costs could vary by both industry and country, but such measures do not exist in practice. To make progress, we begin by assuming that there is a country-specific fixed cost associated with any form of commerce involving that country. This country-specific fixed cost applies to both exports and FDI sales.

Having assumed that this measure is unobserved, country-specific, and yet common to all industries, we subsume this measure into a country fixed effect.

We assume that any remaining cost associated with FDI stems from the cost of maintaining additional capacity. The difficulty associated with choosing a proxy for plant-level fixed costs is that there is no such thing as a representative firm in our model. It is therefore important to find a measure of plant-level fixed costs that is independent of any particular firm's size or level of productivity. This means that standard measures of these costs, such as the number of production workers at a plant of median size, are not appropriate. Instead, we follow the model in choosing the number of non-production workers per establishment as reported in the 1997 Census of Manufacturing.<sup>26</sup> We calculate the average number of non-production workers at the NAICS level.<sup>27</sup> Then we compute our 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 this sector.

### 3.3 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. Everything else equal, international commerce should be skewed toward FDI sales and away from exports, in industries with greater productivity heterogeneity across firms. To test this hypothesis, we require data that quantifies the extent of productivity dispersion by industry. These measures are difficult to construct, because we have no data on the intra-industry distribution of productivity. We therefore rely on guidance from the model to construct a suitable measure of within-industry heterogeneity.

According to the model, differences in firm size reflect differences in productivity, because more productive firms sell more. This mapping — from the distribution of productivity to the distribution of firm size — depends, however, on the elasticity of

---

<sup>26</sup>This measure does not strictly conform to our modeling assumptions, because the number of non-production workers is not independent of establishment size.

<sup>27</sup>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.

substitution among products within an industry. Fortunately, our analysis shows that this coefficient can be recovered from data on the size distribution of firms, which are available, and which we use in the following analysis.

To quantify the extent of dispersion within an industry, 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.<sup>28</sup> 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.<sup>29</sup>

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. We have no access to these data for U.S. firms, however. As a result, we rely on two alternative sources.

First, we use the publicly available data from the 1997 U.S. Census of Manufacturing. Unfortunately, these data report only the number of establishments that fall into 10 different size-categories, which makes it impossible to estimate size dispersion measures by regressing log rank on log sales. Nevertheless, by making additional assumptions one can compute the inverse of the standard deviation of log sales. In particular, assume that all establishments that fall within the same size-category have log sales equal to the center of this category. Then one can treat each of the size-categories in the many subindustries of the BEA industry classification as separate observations. Adopting this method, we calculate the inverse of the standard deviation of log sales using the number of firms in each size-category as weights.

---

<sup>28</sup>It is comforting that the distribution of firm size closely follows a Pareto distribution; see Axtell (2001).

<sup>29</sup>While the two methods of calculation should be equivalent in practice, there are moderate to small differences in the measures. We therefore calculate them both ways.



Second, Bureau van Dijck Electronic Publishing has recently made available a large data set of European firms.<sup>30</sup> This database, named Amadeus, includes information on the consolidated sales, the national identity, and the main line of business by industry of a large number of European firms.<sup>31</sup> 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 subindustries for which countries might produce different mixes. France's industrial structure is very similar to the U.S., 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.<sup>32</sup> Our dispersion measures are based on a sample of 55,339 large Western European firms, and a subset of 15,148 French firms.<sup>33</sup>

The regression-based measure of dispersion provides a natural way of evaluating the cross-sectional variation in this variable relative to the measurement errors induced by fitting the Pareto distribution. Figure 3, which has been constructed from the sample of Western European firms, plots firm rank against firm sales in 4 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 4 quantifies this comparison by showing the 95% confidence interval around the coefficient of dispersion, estimated as the slope of the regression line in each one of the sectors. Evidently, these slopes are precisely estimated in all the sectors, with the

---

<sup>30</sup>This data set has recently been used by Budd, Konings and Slaughter (2002), who investigate international rent-sharing within multinational firms. We thank Matthew Slaughter for bringing this data set to our attention.

<sup>31</sup>Both Western and Eastern European firms are represented.

<sup>32</sup>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.

<sup>33</sup>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 different 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.

exception of five outliers that we shall discuss below.<sup>34</sup>

There are 4 measures of dispersion calculated from the Amadeus data and one measure calculated from the U.S. data.<sup>35</sup> The correlations between these measures are shown in table 3 (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 4 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 4 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 different: the European measures are computed from actual firm-level data while the American measure is calculated from semi-aggregated 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 U.S. may differ from the mix produced in Europe. For this reason the European and American dispersion measures cannot be perfectly aligned.

## 4 Specifications and Results

Our aim is to estimate a linearized version of (18) 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  $\omega$  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

---

<sup>34</sup>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 sectors are very similar to the one represented.

<sup>35</sup>All the dispersion measures are reported in an appendix table.

the partial derivatives implied by (18) 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 (18).

We begin the analysis by considering the raw specification in which we do not attempt to control for any variables that might affect the tradeoff between exporting and FDI sales. The results across specifications for our two samples and five measures of dispersion are shown in table 4. 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 coefficient of FP is positive and significant. We therefore confirm the predictions of the proximity–concentration tradeoff: 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).<sup>36</sup>

---

<sup>36</sup>The magnitude of the coefficients on virtually all dispersion measures are lower in the wider sample.

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 data on 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.<sup>37</sup> 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. After dropping the outliers, all the dispersion measures yield negative coefficients that are significant beyond the 99% 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 5 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 independent variable, divided by the standard deviation of the dependent variable. It converts the regression coefficients into units of sample standard deviations.<sup>38</sup> 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

---

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.

<sup>37</sup>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.

<sup>38</sup>See Wooldridge (2003, Section 6.1) for a further description of this transformation.

variables.<sup>39</sup> For instance, a one standard deviation increase in an industry’s freight cost is generally associated with a third of a standard deviation increase in the logarithm of the ratio of exports to FDI sales. A one standard deviation increase in the dispersion measures induce similar (though, on average, slightly smaller) changes in the dependent variable. The impact of tariffs and returns to scale are, in turn, similar, although somewhat smaller than the impact of the dispersion measures. Taken as a whole, these results suggest that firm-level heterogeneity adds an important dimension to the observed tradeoff between exports and FDI sales.

These results strongly support the theoretical model’s predicted link between firm-level heterogeneity and the ratio of exports relative 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. Cross-industry variations in capital and R&D intensity may, for example, contribute to the observed variations in relative export and FDI sales. Note that both these variables represent characteristics of an industry’s technology that are not captured by our model.<sup>40</sup> Furthermore, as shown in table 3, these measures of technology are correlated with all the different dispersion measures, although the correlations with the U.S.-data-based dispersion measure is rather weak.<sup>41</sup> We therefore rerun our previous specification, controlling for both capital and R&D intensities.

Table 6 reports the results. Evidently, all the dispersion measures remain highly significant. As was the case in the baseline specification, the measurement problems associated with the regression-based dispersion measures affect the results for these variables (the magnitude of the coefficients is significantly lower). When the five outlier sectors are removed from the sample, however, the difference in the coefficients shrinks considerably while all coefficients remain significant beyond the 99% confidence level.

---

<sup>39</sup>In the case of FREIGHT, TARIFF, and FP, the coefficients are averaged across the five specifications.

<sup>40</sup>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.

<sup>41</sup>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).

The results also suggest that R&D intensity is not a useful predictor of the extent of exports versus FDI sales, while industries that are capital intensive are much more likely to substitute FDI for export sales. These results are interesting, but we will not discuss them further, because our theoretical model offers no guidance concerning their interpretation.

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 estimated 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 our empirical analysis.<sup>42</sup>

The results are reported in table 7. 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 effect of the standard proximity–concentration variables in the export versus FDI sales tradeoff.

Another robustness check addresses the potential interdependence of the residuals across countries, which may exist even after we control for country fixed effects. This type of interdependence pattern could be created by the ability of affiliates to reexport a portion of their production to a third country. In this case, a firm’s decision to

---

<sup>42</sup>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.

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, 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 reexports 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 treat all the Western European countries as a single aggregate unit and reestimate our specification with the industry controls (capital and R&D intensity) and industry random effects. The results are reported in table 8. Once again, all the dispersion measures remain highly significant. As could be predicted, the point estimates on the dispersion measures are slightly lower, which reflects the fact that the smaller developing countries now receive a greater weight in the sample.

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 of the 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).<sup>43</sup> These specifications are reported in table 9 for all previously described 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 that we performed, but have chosen not to report in detail (in order to save space). 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. We found that netting out the value of these imports from

---

<sup>43</sup>Under our assumptions on the source of this potential correlation in the residuals — unmeasured sector characteristics — the previously reported random effects coefficients are the efficient estimators.

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 4-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.

## 5 Conclusion

We have developed in this paper a model of international trade in which firms can choose to serve the domestic market, to export, or to engage in FDI in order to serve foreign markets. Every industry is populated by heterogenous 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 they cannot generate positive operating profits 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 the group choose to invest in foreign markets while the less productive firms choose to export.

Our model embodies standard elements of the proximity–concentration tradeoff 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-sectoral heterogeneity of the productivity levels of firms, which induces a size distribution of firms that also affects exports versus FDI sales.

Using data on exports versus 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. 2001. "Zipf Distribution of U.S. Firm Sizes." *Science* 293:1818–1820.
- Bernard, Andrew B., Jonathan Eaton, J. Bradford Jenson and Samuel Kortum. 2000. "Plants and Productivity in International Trade." NBER Working Paper No. 7688.
- Brainard, S. Lael. 1993. "A Simple Theory of Multinational Corporations and Trade with a Trade-Off Between Proximity and Concentration." NBER Working Paper No. 4269.
- Brainard, S. Lael. 1997. "An Empirical Assessment of the Proximity-Concentration Trade-off between Multinational Sales and Trade." *American Economic Review* 87:520–44.
- Budd, John W., Jozef Konings and Matthew J. Slaughter. 2002. "International Rent Sharing in Multinational Firms." NBER Working Paper No. 8809.
- Doms, Mark E. and J. Bradford Jensen. 1998. "Comparing Wages, Skills, and Productivity between Domestically and Foreign-Owned Manufacturing Establishments in the United States". In *Geography and ownership as bases for economic accounting*, ed. Robert E. Baldwin, Robert E. Lipsey and J. David Richardson. NBER Studies in Income and Wealth, vol. 59, University of Chicago Press.
- Ethier, Wilfred J. 1986. "The Multinational Firm." *The Quarterly Journal of Economics* 101:805–33.
- Ethier, Wilfred J. and James R. Markusen. 1996. "Multinational Firms, Technology Diffusion and Trade." *Journal of International Economics* 41:1–28.
- Feenstra, Robert C. 1997. "U.S. Exports, 1972-1994: With State Exports and Other U.S. Data." NBER Working Paper No. 5990.
- Girma, Sourafel, Steve Thompson and Peter W. Wright. 2002. "Why are Productivity and Wages Higher in Foreign Firms?" *The Economic and Social Review* 33:93–100.
- Hanson, Gordon H., Raymond J. Mataloni and Matthew J. Slaughter. 2002. "Vertical Specialization in Multinational Firms." Mimeo, University of California, San Diego.
- Helpman, Elhanan. 1984. "A Simple Theory of International Trade with Multinational Corporations." *Journal of Political Economy* 92:451–71.
- Helpman, Elhanan. 1985. "Multinational Corporations and Trade Structure." *The Review of Economic Studies* 52:443–57.
- Horstmann, Ignatius and James R. Markusen. 1987. "Licensing versus Direct Investment: A Model of Internalization by the Multinational Enterprise." *Canadian Journal of Economics* 20:464–81.

- Horstmann, Ignatius and James R. Markusen. 1992. "Endogenous Market Structures in International Trade." *Journal of International Economics* 32:109–29.
- Markusen, James R. 1995. "The Boundaries of Multinational Enterprises and the Theory of International Trade." *Journal of Economic Perspectives* 9:169–89.
- Markusen, James R. 2002. *Multinational Firms and the Theory of International Trade*. MIT Press.
- Markusen, James R. and Anthony J. Venables. 2000. "The Theory of Endowment, Intra-industry and Multi-national Trade." *Journal of International Economics* 52:209–34.
- Melitz, Marc J. 2002. "The Impact of Trade on Aggregate Industry Productivity and Intra-Industry Reallocations." NBER Working Paper No. 8881.
- Rob, Rafael and Nikolaos Vettas. 2001. "Foreign Direct Investment and Exports with Growing Demand." CEPR Discussion Paper No. 2786.
- Tybout, James R. 2002. "Plant and Firm-Level Evidence on New Trade Theories". In *Handbook of International Economics*, ed. James Harrigan. Basil-Blackwell. Forthcoming.
- Wooldridge, Jeffrey M. 2003. *Introductory Econometrics*. 2 ed. South-Western.
- Yeaple, Stephen R. 2000. "The Determinants of U.S. Outward Foreign Direct Investment: Market Access versus Comparative Advantage." Mimeo, University of Pennsylvania.

## Appendix

In this appendix, we discuss the effects of firm heterogeneity on relative exports and FDI sales when firms have an incentive to export from foreign affiliates to other countries. Since our interest in this extension is motivated by our empirical work, we focus the analysis on U.S. firms and their strategies to reach consumers in other countries. For this purpose we do not need to construct a full general equilibrium model.

Consider an industry with  $n_E^U$  entrants from the U.S. A fraction  $G(a_D^U)$  of these firms remain active and serve their U.S. domestic market, where  $a_D^U$  is the solution to  $a_D^i$  from equation (2) for  $i = U$ . Firms with lower productivity levels exit. We normalize the U.S. wage to 1.

Now suppose that every U.S. firm can also sell its products in  $M$  foreign markets; call them European countries. All these countries are symmetrical, with the same market demand level  $B$ . For expositional simplicity, we consider the case where the European wage is equal to the U.S. wage, and hence also equal to 1.<sup>44</sup> We assume that transport costs from the U.S. to each one of the European countries is  $\tau^U$  while transport costs between every pair of European countries is  $\tau$ , where  $\tau < \tau^U$ . This last assumption provides the rationale for exports from European subsidiaries to other countries in Europe. If transport costs across Europe were the same as between Europe and the U.S., then American firms would not have any incentive to export from a European subsidiary to another European country.

As in the main text, an American firm with labor coefficient  $a$  makes profits

$$M \left[ (a\tau^U)^{1-\varepsilon} B - f_X \right]$$

from exporting to the European countries. It follows that exports are profitable for all firms with  $a \leq a_X^U$ , where  $a_X^U$  satisfies

$$(a_X^U \tau^U)^{1-\varepsilon} B = f_X. \tag{A1}$$

---

<sup>44</sup>Relaxing this assumption would not qualitatively change any of the ensuing results.

There are now other strategies a firm can use to serve foreign markets: it can form a subsidiary in each one of them, which is dedicated to serving the local market; or it can form a subsidiary in a single country and use it as an export platform to the remaining  $M - 1$  countries. The former strategy yields profits

$$M (a^{1-\varepsilon} B - f_I)$$

while the latter yields profits

$$(a^{1-\varepsilon} B - f_I) + (M - 1) [(a\tau)^{1-\varepsilon} B - f_x] = a^{1-\varepsilon} B [1 + (M - 1) \tau^{1-\varepsilon}] - [f_I + (M - 1) f_x].$$

It follows from these functional forms that the profit function from U.S. exports is the flattest among the three when presented in a figure such as figure 1 in the text, while the profit function from dedicated FDI is the steepest. Therefore, there may now exist three regions defined by the cutoffs  $a_X^U > a_{IX}^U > a_I^U$ . Firms with  $a$  above  $a_X^U$  do not serve the European markets; firms with  $a$  between  $a_X^U$  and  $a_{IX}^U$  export from the U.S. to Europe; and firms with  $a$  between  $a_{IX}^U$  and  $a_I^U$  invest in one European country and export from that country to the other European countries; finally, firms with  $a < a_I^U$  invest in dedicated facilities in every European country (similarly to the unique FDI option in the main text). The cutoff coefficient  $a_{IX}^U$  and  $a_I^U$  satisfy

$$M \left[ (a_{IX}^U \tau^U)^{1-\varepsilon} B - f_x \right] = (a_{IX}^U)^{1-\varepsilon} B [1 + (M - 1) \tau^{1-\varepsilon}] - [f_I + (M - 1) f_x], \quad (\text{A2})$$

$$(a_{IX}^U)^{1-\varepsilon} B [1 + (M - 1) \tau^{1-\varepsilon}] - [f_I + (M - 1) f_x] = M \left[ (a_{IX}^U)^{1-\varepsilon} B - f_I \right]. \quad (\text{A3})$$

There are obvious restrictions which ensure that  $a_X^U > a_{IX}^U > a_I^U$ . We assume that they are satisfied.

>From (A1) and (A2) we obtain

$$\left( \frac{a_X^U}{a_{IX}^U} \right)^{\varepsilon-1} = \left( \frac{f_I - f_x}{f_x} \right) \frac{(\tau^U)^{1-\varepsilon}}{1 + (M - 1) \tau^{1-\varepsilon} - M (\tau^U)^{1-\varepsilon}} \quad (\text{A4})$$

while from (A2) and (A3) we obtain

$$\left(\frac{a_X^U}{a_I^U}\right)^{\varepsilon-1} = \left(\frac{f_I - f_X}{f_X}\right) \frac{(\tau^U)^{1-\varepsilon}}{1 - \tau^{1-\varepsilon}}. \quad (\text{A5})$$

The assumption of equal market demands across all  $M$  countries entails a symmetric equilibrium outcome in which the distribution of affiliates in each of the  $M$  countries is identical. This will be satisfied if the U.S. firms who only have one foreign affiliate (with costs between  $a_I^U$  and  $a_{IX}^U$ ) choose the location of this affiliate among the  $M$  identical countries at random. Given this feature, the number of U.S. firms exporting to any one of the  $M$  countries relative to the number of U.S. firm with affiliates located in this country, is

$$\frac{n_X^U}{n_I^U} = \frac{G(a_X^U) - G(a_{IX}^U)}{\frac{1}{M}G(a_{IX}^U) + \frac{M-1}{M}G(a_I^U)}.$$

Similarly, the market share of U.S. exporters to a European country  $i$  relative to the *local* market share of affiliates of U.S. multinationals that are located in country  $i$ , is

$$\frac{s_X^U}{s_I^U} = \frac{(\tau^U)^{1-\varepsilon} [V(a_X^U) - V(a_{IX}^U)]}{\frac{1}{M}V(a_{IX}^U) + \frac{M-1}{M}V(a_I^U)}. \quad (\text{A6})$$

When  $G(\cdot)$  is Pareto,  $n_X^U/n_I^U$  and  $s_X^U/s_I^U$  depend on the cutoffs only through the ratios described in (A4) and (A5). We can therefore substitute (A4) and (A5) into these equations to obtain  $n_X^U/n_I^U$  and  $s_X^U/s_I^U$  as functions of the model's parameters, as we did in the main text. It then follows that, more dispersion in firm productivity levels — as measured by the shape parameter of the Pareto distribution — reduces in country  $i$  the number and market share of firms that export from the U.S. to country  $i$  relative to the number and market share of affiliates of U.S. multinationals that are located in country  $i$ . (A4), (A5), and (A6) also imply that increases in trade costs  $\tau^U$  — holding either  $\tau$  or  $\tau/\tau^U$  fixed — will also lead to a decrease in the relative number and market share of U.S. exporters (versus FDI). Note that in the definition of the sales in country  $i$  of affiliates of U.S. multinationals we do not include exports from affiliates located in other European countries.

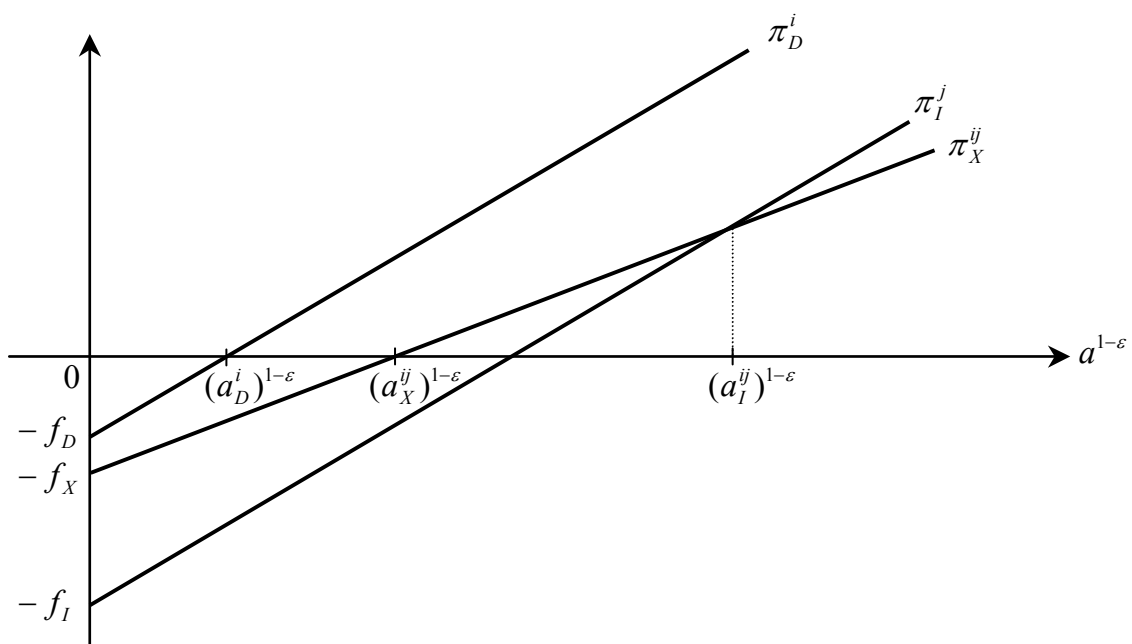


Figure 1: Profits from domestic sales, from exports and from FDI

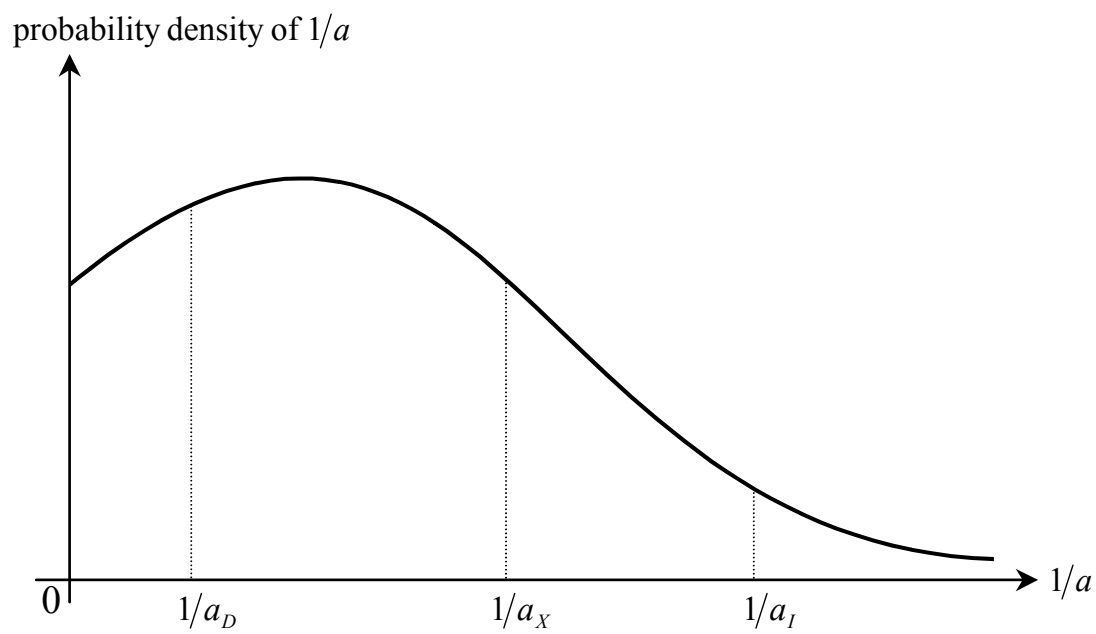


Figure 2: Probability density of labor productivity

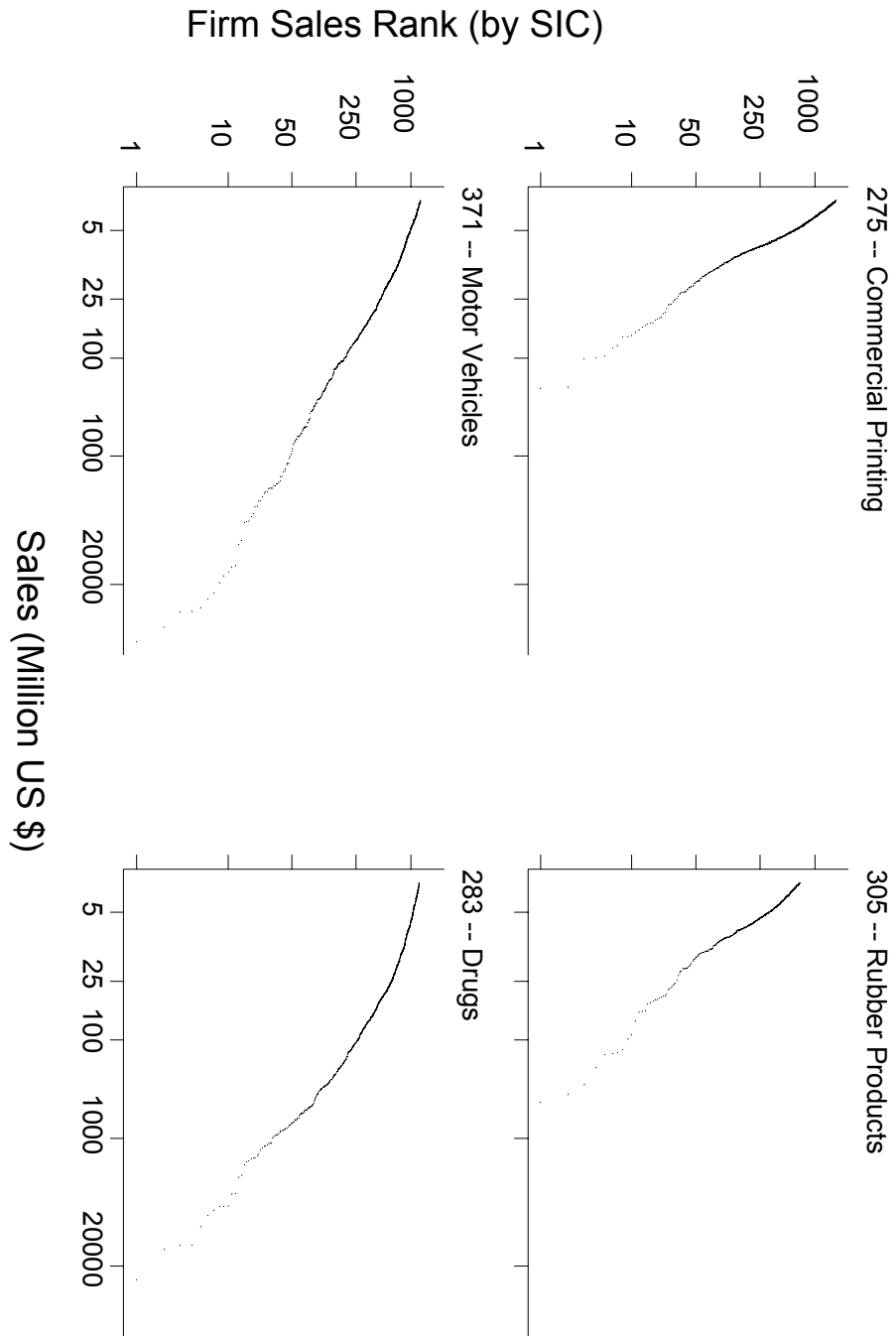


Figure 3: Empirical Distribution of Firm Sales



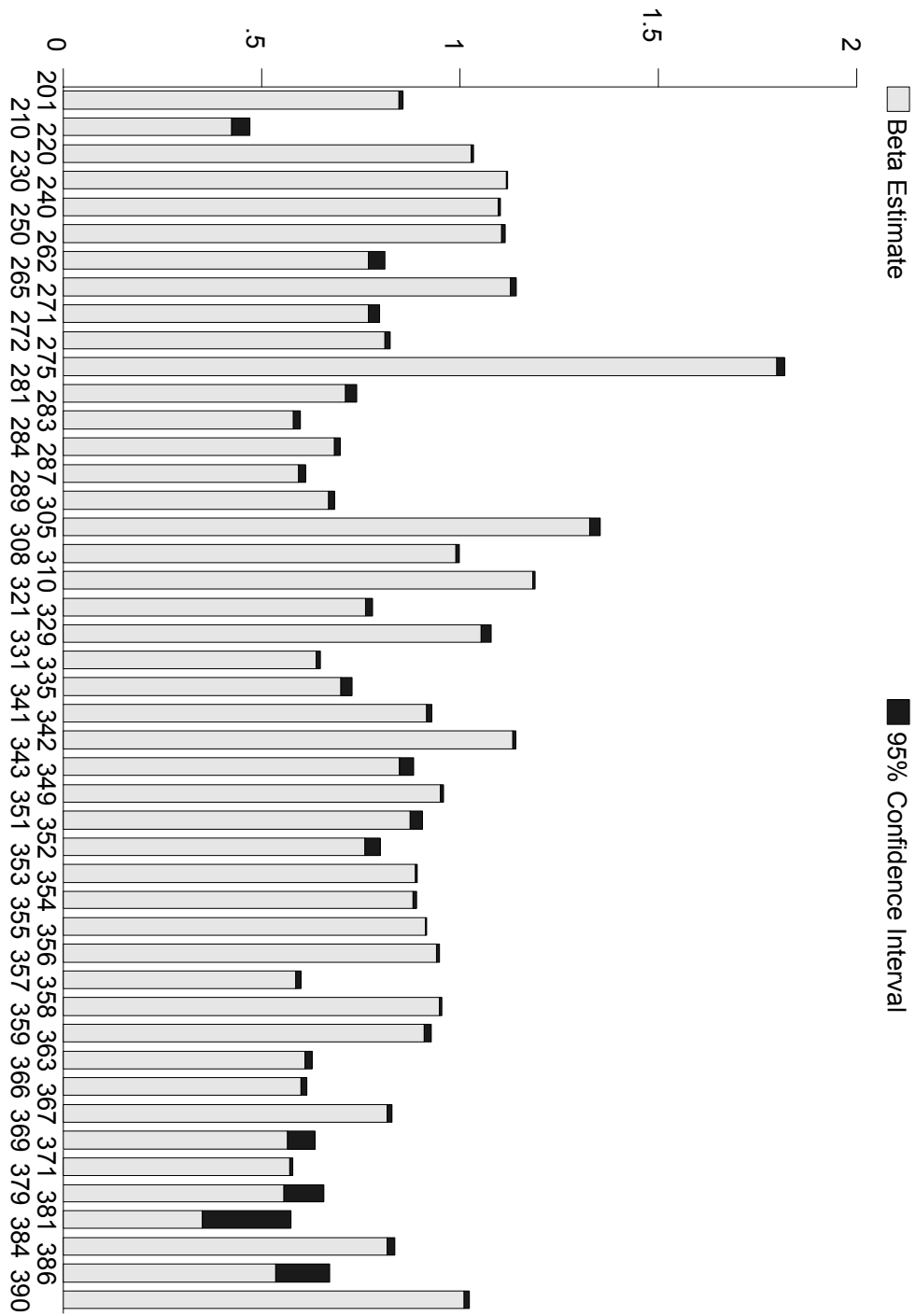


Figure 4: Regression Fit to the Pareto Distribution

Table 1: BEA 3-Digit Manufacturing Sectors

201	Meat Products	329	Stone, Minerals, and Ceramics
202	Dairy Products	331	Ferrous metals
203	Vegetables and Preserves	335	Non-Ferrous metals
204	Grain Mill Products	341	Metal Cans, Fabricated Metal
205	Bakery Products	342	Cutlery
208	Beverages	343	Heating and Plumbing Equipment
209	Other Food	349	Metal Services
210	Tobacco	351	Engines and Turbines
220	Textiles	352	Farm Machinery
230	Apparel	353	Construction Machinery
240	Wood and Lumber	354	Metalworking Machinery
250	Furniture	355	Special Industrial Machinery
262	Pulp and Paper	356	General Industrial Machinery
265	Processed Paper	357	Computers
271	Newsprint	358	Refrigeration Equipment
272	Other publishing	359	Other Industrial Equipment
275	Commercial Printing	363	Household Appliances
281	Industrial Chemicals	366	Audio, Video, Communications Equipment
283	Drugs	367	Electronic Components
284	Soap and Cleansing Products	369	Other Electronics
287	Agricultural Chemicals	371	Motor Vehicles
289	Other Industrial Chemicals	379	Other Transport Equipment
305	Rubber	381	Scientific and Measuring Equipment
308	Miscellaneous Plastics	384	Medical Equipment
310	Leather	386	Optical and Photographic Equipment
321	Glass	390	Miscellaneous Manufacturers

Table 2: Countries (by continent)

Austria*	Argentina*
Belgium*	Brazil*
Denmark*	Canada*
France*	Chile*
Finland	Colombia
Germany*	Mexico*
Greece	Peru
Ireland*	Venezuela*
Italy*	
Netherlands*	Australia*
Norway*	New Zealand*
Portugal	Hong Kong*
Spain*	Indonesia
Sweden*	Japan*
Switzerland*	Malaysia
Turkey	Philippines*
United Kingdom*	Singapore*
	South Korea*
Israel	Taiwan*
South Africa	Thailand

\* Indicates Narrow Sample

Table 3: Correlation Between Alternative Measures of Dispersion

	U.S. Std. Dev.	Europe Std. Dev.	France Std. Dev.	Europe Reg. Coef	France Reg. Coef	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

Table 4: Exports versus FDI - Baseline Results

Narrow Sample (N=962)					
	U.S. Std. Dev.	Europe Std. Dev.	France Std. Dev.	Europe Reg. Coeff.	France Reg. Coeff.
FREIGHT	-1.241 (-9.018)	-1.146 (-8.489)	-1.193 (-9.049)	-1.173 (-8.582)	-1.176 (-8.609)
TARIFF	-0.354 (-2.506)	-0.512 (-3.732)	-0.410 (-3.203)	-0.555 (-3.901)	-0.542 (-3.855)
FP	0.580 (6.011)	0.794 (6.747)	0.767 (7.006)	0.731 (5.845)	0.702 (5.675)
DISPERSE	-2.052 (-6.903)	-2.727 (-7.024)	-2.396 (-8.273)	-1.935 (-5.667)	-1.899 (-5.545)
$R^2$	0.318	0.327	0.352	0.309	0.310

Wide Sample (N=1176)					
	U.S. Std. Dev.	Europe Std. Dev.	France Std. Dev.	Europe Reg. Coeff.	France Reg. Coeff.
FREIGHT	-1.184 (-9.545)	-1.094 (-8.950)	-1.107 (-9.411)	-1.114 (-8.981)	-1.115 (-8.979)
TARIFF	-0.241 (-1.851)	-0.398 (-3.106)	-0.311 (-2.570)	-0.434 (-3.297)	-0.424 (-3.234)
FP	0.579 (6.421)	0.734 (6.662)	0.727 (6.975)	0.669 (5.773)	0.630 (5.491)
DISPERSE	-1.968 (-7.267)	-2.370 (-6.587)	-2.123 (-7.408)	-1.667 (-5.259)	-1.574 (-4.978)
$R^2$	0.293	0.294	0.316	0.279	0.278

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

Table 5: “Beta” Coefficients

	Mean	Standard Deviation	“Beta” Coefficient
Dependent Var.	-0.595	2.375	
FREIGHT	1.863	0.653	-0.328
TARIFF	2.015	1.020	-0.204
FP	3.321	0.785	0.236
U.S. S.D.	1.749	0.316	-0.273
Europe S.D.	1.198	0.276	-0.317
France S.D.	1.224	0.375	-0.378
Europe Reg.	1.260	0.333	-0.271
France Reg.	1.257	0.336	-0.269

Table 6: Exports Versus FDI - Controls

Narrow Sample (N=961)

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

Wide Sample (N=1175)

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

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

Table 7: Exports versus FDI - Random Effects  
(With Controls)

Narrow Sample (N=961)

	U.S. Std. Dev.	Europe Std. Dev.	France Std. Dev.	Europe Reg. Coeff.	France Reg. Coeff.
FREIGHT	-0.430 (-2.554)	-0.398 (-2.344)	-0.428 (-2.533)	-0.397 (-2.336)	-0.397 (-2.334)
TARIFF	-0.113 (-0.922)	-0.127 (-1.033)	-0.105 (-0.857)	-0.136 (-1.107)	-0.133 (-1.085)
FP	1.376 (5.145)	1.132 (4.128)	1.096 (4.233)	1.154 (4.107)	1.137 (4.093)
DISPERSE	-2.623 (-4.897)	-2.763 (-3.459)	-2.445 (-4.761)	-2.031 (-3.098)	-1.991 (-3.180)
KL	-1.106 (-4.652)	-0.613 (-2.238)	-0.570 (-2.168)	-0.757 (-2.891)	-0.758 (-2.896)
RD	-0.002 (-0.020)	0.126 (1.029)	0.116 (0.970)	0.133 (1.081)	0.119 (0.972)
$R^2$	0.352	0.316	0.342	0.307	0.308

Wide Sample (N=1175)

	U.S. Std. Dev.	Europe Std. Dev.	France Std. Dev.	Europe Reg. Coeff.	France Reg. Coeff.
FREIGHT	-0.331 (-2.296)	-0.322 (-2.230)	-0.328 (-2.278)	-0.320 (-2.215)	-0.320 (-2.215)
TARIFF	-0.004 (-0.038)	-0.018 (-0.155)	-0.004 (-0.035)	-0.022 (-0.193)	-0.021 (-0.187)
FP	1.361 (4.123)	1.110 (3.455)	1.081 (3.475)	1.127 (3.377)	1.103 (3.369)
DISPERSE	-2.518 (-3.824)	-2.559 (-2.733)	-2.265 (-3.706)	-1.864 (-2.398)	-1.786 (-2.424)
KL	-1.069 (-3.660)	-0.599 (-1.871)	-0.561 (-1.789)	-0.734 (-2.373)	-0.739 (-2.408)
RD	0.006 (0.042)	0.123 (0.862)	0.116 (0.811)	0.129 (0.894)	0.115 (0.805)
$R^2$	0.312	0.274	0.297	0.267	0.267

Z-statistics in parentheses. Constant and country dummies are suppressed.

Table 8: Exports versus FDI - Aggregated Europe  
(With Controls and Random Effects)

Wide Sample (N=680)					
	U.S. Std. Dev.	Europe Std. Dev.	France Std. Dev.	Europe Reg. Coeff.	France Reg. Coeff.
FREIGHT	-0.493 (-2.882)	-0.485 (-2.810)	-0.487 (-2.837)	-0.485 (-2.806)	-0.483 (-2.795)
TARIFF	-0.005 (-0.030)	-0.044 (-0.254)	-0.032 (-0.183)	-0.058 (-0.335)	-0.052 (-0.299)
FP	1.216 (4.683)	0.926 (3.713)	0.922 (3.874)	0.929 (3.620)	0.928 (3.663)
DISPERSE	-2.242 (-4.351)	-1.845 (-2.595)	-1.748 (-3.765)	-1.301 (-2.217)	-1.322 (-2.371)
KL	-0.989 (-4.339)	-0.636 (-2.560)	-0.588 (-2.447)	-0.740 (-3.105)	-0.737 (-3.102)
RD	-0.023 (-0.208)	0.076 (0.692)	0.073 (0.671)	0.078 (0.706)	0.071 (0.648)
$R^2$	0.314	0.273	0.289	0.268	0.270

Z-statistics in parentheses. Constant and country dummies are suppressed.



Table 9: Exports versus FDI - Additional Robustness Results  
(Clustered Standard Errors and IV specifications)

	OLS			IV		
	Narrow Sample	Wide Sample	Agg. Europe	Narrow Sample	Wide Sample	Agg. Europe
FREIGHT	-1.040 (-3.997)	-1.011 (-4.437)	-1.001 (-4.464)	-1.218 (-3.862)	-1.118 (-4.366)	-1.053 (-4.545)
TARIFF	-0.365 (-1.611)	-0.241 (-1.081)	-0.077 (-0.304)	-0.188 (-0.706)	-0.124 (-0.478)	-0.016 (-0.056)
FP	1.177 (4.876)	1.133 (4.472)	1.086 (4.166)	1.609 (3.605)	1.457 (3.311)	1.344 (3.182)
DISPERSE (US)	-2.343 (-3.689)	-2.248 (-3.655)	-2.150 (-3.349)	-4.321 (-2.606)	-3.681 (-2.248)	-3.198 (-2.217)
KL	-0.868 (-3.032)	-0.793 (-2.513)	-0.848 (-2.593)	-0.938 (-2.855)	-0.848 (-2.496)	-0.890 (-2.599)
RD	-0.104 (-0.851)	-0.086 (-0.687)	-0.087 (-0.691)	-0.158 (-1.233)	-0.127 (-0.985)	-0.121 (-0.968)
$R^2$	0.373	0.338	0.314	0.328	0.315	0.315
$N$	961	1175	678	961	1175	678

In the IV specifications, the US 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.

Table A1: Dispersion Measures

Sector	Standard Deviation			Reg. Coef.		Sector	Standard Deviation			Reg. Coef.	
	U.S.	Eur.	Fra.	Eur.	Fra.		U.S.	Eur.	Fra.	Eur.	Fra.
201	2.15	1.16	1.13	1.18	1.19	329	1.48	0.90	0.89	0.94	0.95
202	1.82	1.43	1.49	1.47	1.45	331	1.88	1.52	1.45	1.56	1.58
203	2.03	1.12	1.17	1.14	1.12	335	1.49	1.34	1.20	1.40	1.38
204	1.36	1.17	1.19	1.19	1.20	341	1.92	1.05	1.21	1.09	1.06
205	1.96	1.11	1.06	1.15	1.18	342	1.71	0.86	0.81	0.88	0.90
208	1.94	1.29	1.33	1.32	1.32	343	1.84	1.09	1.27	1.16	1.14
209	1.69	1.16	1.12	1.19	1.21	349	1.62	1.03	0.92	1.05	1.03
210	2.55	2.00	2.97	2.24	2.33	351	2.56	1.08	1.05	1.12	1.11
220	1.84	0.96	0.99	0.97	0.96	352	1.81	1.20	1.07	1.28	1.30
230	1.57	0.89	0.91	0.89	0.90	353	1.67	1.11	1.07	1.12	1.11
240	1.53	0.90	0.81	0.91	0.92	354	1.40	1.11	1.04	1.13	1.14
250	1.72	0.89	0.94	0.90	0.89	355	1.61	1.08	0.99	1.09	1.09
262	1.25	1.19	1.07	1.27	1.27	356	1.66	1.05	1.15	1.06	1.08
265	1.30	0.86	1.05	0.88	0.85	357	2.00	1.64	1.46	1.69	1.73
271	1.66	1.22	1.26	1.28	1.28	358	1.85	1.04	0.99	1.05	1.05
272	1.42	1.19	1.00	1.22	1.29	359	1.32	1.06	1.11	1.09	1.07
275	1.34	0.55	0.52	0.55	0.54	363	2.45	1.55	1.55	1.62	1.65
281	1.31	1.32	1.45	1.38	1.33	366	2.02	1.59	1.60	1.65	1.66
283	2.11	1.62	1.55	1.68	1.77	367	1.91	1.18	1.27	1.22	1.20
284	1.87	1.41	1.45	1.45	1.43	369	1.91	1.49	1.54	1.65	1.83
287	1.64	1.59	1.58	1.66	1.65	371	2.16	1.71	1.78	1.74	1.76
289	1.40	1.43	1.47	1.48	1.43	379	1.70	1.39	1.45	1.65	1.57
305	1.62	0.73	0.71	0.75	0.75	381	2.33	1.50	1.88	2.16	2.00
308	1.56	0.99	0.99	1.01	1.01	384	1.74	1.17	1.11	1.21	1.22
310	1.73	0.84	0.88	0.84	0.85	386	1.79	1.34	1.75	1.66	1.37
321	1.24	1.25	0.96	1.30	1.30	390	1.50	0.97	0.97	0.98	0.99