Where Do Firms Export, How Much, and Why?

Martina Lawless*
Central Bank and Financial Services Authority of Ireland

Karl Whelan†
University College Dublin

October 9, 2008

Abstract

The empirical finding that exporting firms are more productive on average than non-exporters has provoked a large theoretical literature based on models such as Melitz (2003), where more productive firms are more likely to overcome costs associated with trade. This paper provides a systematic empirical assessment of the productivity heterogeneity framework using an Irish dataset that includes information on destinations and firm characteristics such as productivity. We find a high degree of unpredictable idiosyncratic participation in export markets by firms, a relatively weak positive correlation between the extent of a firm's export market participation and its export sales, and a limited role for productivity in explaining firm exporting behavior. We illustrate the effect of firm heterogeneity on gravity regressions of aggregate trade flows and show how past exporting to a particular market has a strong impact on the current probability of exporting there.

*E-mail: martina.lawless@centralbank.ie. We would like to thank Forfás for providing the data used in this paper. The views expressed in this paper are our own, and do not necessarily reflect the views of the Central Bank and Financial Services Authority of Ireland or the ESCB.

†E-mail: karl.whelan@ucd.ie.

1. Introduction

Traditional trade theory focused on differences between countries as the principle mechanism behind trade, with all firms within a country treated as identical. In part, this focus reflected data limitations because only country-level trade statistics were available. Since the mid-1990s, however, empirical evidence on the exporting behavior of individual firms has provided significant insights that have had an important influence on how economists think about international trade. In particular, the findings of Andrew Bernard and J. Bradford Jensen (1995, 1999, 2004) that exporters are more productive than non-exporters has stimulated new theoretical research focused on the implications for international trade of heterogeneity in firm productivity. Particularly influential has been the work of Marc Melitz (2003), which provided a tractable model structure in which more productive firms have lower unit costs and so are more likely to overcome the fixed costs associated with trade that prevent other firms from exporting.[1]

Most of this literature has focused largely on whether firms are exporters or not. However, the heterogenous firm models presented by Helpman, Melitz, and Rubinstein (2008) and Chaney (2008) examine cases in which there are multiple export destinations and these generate strong predictions for where firms should export and how much they should sell in export markets. In particular, these models predict that firms should enter different export markets according to a pre-specified hierarchy, with only the most productive firms able to enter the least popular markets. Also, because these models predict that export sales depend positively on productivity, there should be a clear link between export participation and sales so that firms that participate in more export markets should also sell more within each individual market.

This paper provides an empirical assessment of the predictions of the heterogenous productivity framework for where firms export to and how much they sell in these markets. We do this using a panel dataset of Irish firms that combines information on firm characteristics such as productivity and sectoral information with a detailed description of exactly how each firm's exports are allocated across destinations. The fact that firms are tracked over time also means that we can assess the extent to which deviations from the model persist over time and how past export participation and sales—which do not feature in the Melitz framework—affect current export behavior.

The sylized predictions just noted—entry into export markets according to a hierarchy and a strict linkage between export participation and sales—are unlikely to fully match the data. Thus, we analyse the data using a more generalized framework incoporating random firm-country-specific shocks to trade costs and demand and in which firms differ systematically in terms of

[1] Bernard, Eaton, Jensen and Kortum (2003) was another important early paper to focus on the link between exporting and heterogeneity in firm productivity, though its analytical framework is somewhat more complicated and has been used less than the Melitz framework. As of August 2008, the IDEAS/Repec website lists Melitz (2003) as the sixth most-cited paper published in the last five years. See http://ideas.repec.org/top/top.ritem.nbcites.html.

productivity but also in terms of their fixed and variable trade costs. We illustrate how deviations from the stylized predictions can be interpreted in terms of their implications for the importance of various types of random shocks as well as the ways in which firms differ systematically.

Our findings can be grouped under six headings. First, we examine firm export participation using a Probit model for the dependent variable E_{ijt}—which equals one when firm i exports to destination j at time t and zero otherwise—with firm-year (D_{it}) and destination (D_j) dummies used as explanatory variables. In this regression, firm-year dummies capture productivity differentials and other potentially firm-specific factors that influence exporting decisions in a particular year, while country dummies capture the attractiveness of various destinations. In the absence of random firm-market elements, the data would be characterized by the hierarchy prediction and this model would have a perfect fit. We find that the model has a pseudo-R^2 of about 0.5, which means there is substantial randomness in firm-destination export matches: Many of the observed firm-country export matches cannot be explained by either the systematic properties of the firm or the properties of the destination.

Second, we explore the extent to which *observed* firm and destination characteristics can replicate the explanatory power of this benchmark dummy variable Probit model. We find mixed results. On the positive side, we find that a small number of country variables, such as distance, GDP and language, come very close to replicating the fit of the benchmark model incorporating country fixed effects. However, models using measures of firm productivity such as value-added per worker, and other firm characteristics such as average wages and sector dummies, achieve only about half of the fit of the benchmark model based on firm-specific effects. Thus, despite the widespread application of Melitz-style models which focus on productivity as the principle source of firm heterogeneity, our data suggest that productivity heterogeneity provides only a limited explanation for firm-level patterns of trade. This suggests that other firm-specific factors such as trade costs that vary across firms are playing an important role in determining observed patterns of firm trading.

Third, we examine the factors that determine *how much* exporting firms sell to various destinations and relate these results to findings about export participation. Again, we find that country variables such as distance and GDP do well in capturing the role played by country fixed effects in a benchmark model. Firm variables, such as value-added per worker, perform better than they do in explaining export participation but still fall well short of the benchmark model.

Fourth, we compare the results from our analyses of export sales and export participation. We examine the relationship between the estimated firm fixed effects in the export participation and export sales equation. Melitz-style models—in which a single summary factor explains both the decision to export and the amount that is sold should the firm choose to export—predict that these two sets of firm-level fixed effects should be perfectly correlated. While we find a positive correlation, there are still substantial deviations from this prediction. Our framework suggests this imperfect correlation reflects systematic deviations across firms in their fixed trade costs. We also find very little relationship between the random errors affecting firm export participation and the errors affecting export sales in individual markets. This implies that most of the randomness affecting export participation in individual markets reflects random variations across firms in their fixed trade costs related to market entry.

Fifth, we illustrate how aggregate gravity regressions, which estimate the effects of variables like distance on trade flows, combine two different effects—a direct effect on firm sales and a composition effect due to higher distance discouraging some firms from exporting at all. Empirical evidence on these two effects is provided. Our results support the findings of Helpman, Melitz, and Rubinstein (2008), using aggregate data, that the composition effect is substantively important.

Finally, we exploit the panel structure of our dataset to examine the impact of past export participation and sales on their current values. We estimate the effect that a firm's past participation in an export market has on its current participation that market. We find a significant, but moderately-sized effect: Participation by firm i in market j at time $t-1$ raises the probability of its participation in that market at time t by 0.5.

The existing literature on where firms export to is limited due to a lack of available datasets.[2] The closest comparison to our paper is Eaton, Kortum, and Kramarz (2007), which examines the hierarchy hypothesis using a cross-section of French firms and also presents a Melitz-style model augmented by additional random shocks. Our paper differs in four key ways. First, we assess our model using simple regression-based diagnostics, whereas Eaton, Kortum and Kramarz rely on comparing results from a model simulation with certain aspects of the data. Second, we explicitly link our firm-level regressions for export participation with the corresponding regressions for export sales and use the relationships between the coefficients in these regressions to assess the importance of variations in firm-specific fixed trade costs. Third, we use data on observable aspects of firms and markets to assess which are the key characteristics determining heterogeneity at the firm level and trade barriers at the country level. Fourth, we exploit a panel structure not available in the French data set.

The rest of the paper is organized as follows. Section 2 presents our model and Section 3 introduces the data. Section 4 presents our results on the factors that determine which countries firms export to. Section 5 analyzes the factors determining how much is exported. Section 6 presents our dynamic analysis of export participation and sales.

[2]Exceptions include Eaton, Kortum, and Kramarz (2004) for France, Lawless (2007) for Ireland, and Eaton, Eslava, Kugler and Tybout (2007) for Columbia.

2. A Firm-Level Model of Exporting

This section presents a model of firm export participation which incorporates the key features of Melitz (2003), namely firm heterogeneity and both fixed and variable trade costs. We first introduce a simple model, essentially the same as that presented by Helpman, Melitz, and Rubinstein (2008), in which the only random factor at the firm-level is productivity. We then present a more general model in which there are systematic and market-specific shocks to firm demand and trade costs. The remainder of the paper interprets our evidence on firm-level patterns of export behavior in the context of this more general model. Note that while our dataset is a panel, the model presented here applies to cross-sectional data. We will discuss the incorporation of a time element into the empirical analysis in later sections.

2.1. A Simple Productivity-Based Hierarchy Model

We consider the export behavior of a set of firms, indexed by $i = 1, ..., N_f$, from the same country who can each export to a set of markets indexed by $j = 1, ..., N_m$. As in Melitz (2003), we assume that firm i produces a differentiated product using a Ricardian technology with cost-minimizing unit cost $\frac{c}{a_i}$, where the firm-specific productivity parameter a_i varies randomly across firms. Firm i faces a demand curve for its product in country j given by

$$Q_{ij} = \eta P_{ij}^{-\epsilon} Y_j^{\theta} \tag{1}$$

where P_{ij} is the price it charges in country j and Y_j is GDP in that country.[3]

In addition to production costs, there are two types of trade costs associated with exporting to country j. First, there are variable costs, which are modelled with the iceberg specification so that τ_j units have to be shipped from our country of interest to country j for one unit to arrive. These can be viewed as transport costs, tariffs, and the variable costs associated with marketing and distribution. Second, there are fixed costs F_j which are unrelated to export sales. These can be viewed as the bureaucratic paperwork costs associated with exporting, to marketing costs, and to the costs of running a wholesale and retail distribution chain. It is likely, of course, that some of these costs also increase with the scale of exports; however, what is important from a theoretical perspective is that at least *some* of them need to be incurred independent of the scale of export sales.

The assumptions about market structure and trade costs imply that the optimal selling price in

[3]This can be derived formally from the assumption of Dixit-Stiglitz preferences over all goods produced in all countries, in which case $\theta = 1$ and ω depends negatively on the overall price level in country j. Since we don't undertake any welfare analysis in the paper, we use this more *ad hoc* but less restrictive formulation instead of starting from utility functions.

country j for firm i (should it choose to produce there) is

$$P_{ij} = \frac{\epsilon}{\epsilon - 1} \frac{\tau_j c}{a_i} \tag{2}$$

Assuming firm i chooses to sell in this market, the value of its export sales will be

$$S_{ij} = P_{ij} Q_{ij} = \left(\frac{\epsilon - 1}{\epsilon} \frac{a_i}{\tau_j c} \right)^{\epsilon - 1} Y_j^{\theta} \tag{3}$$

The firm's export sales to a market depend positively on the firm's productivity level and on the destination market's GDP and depend negatively on the variable trade costs associated with that market. The profits generated by these sales are given by

$$\Pi_{ij} = \mu \left(\frac{a_i}{\tau_j c} \right)^{\epsilon - 1} Y_j^{\theta} - F_j \tag{4}$$

where $\mu = (\epsilon - 1)^{\epsilon - 1} \epsilon^{-\epsilon}$. These profits are positive as long as

$$a_i > \left(\frac{F_j}{\mu Y_j^{\theta}} \right)^{\frac{1}{\epsilon - 1}} \tau_j c \tag{5}$$

Re-written in terms of logs, the condition for firm i to export to country j become

$$\log a_i - \frac{1}{\epsilon - 1} \log \left(\frac{F_j}{\mu Y_j^{\theta}} \right) - \log (\tau_j c) > 0 \tag{6}$$

Thus, without any additional random elements affecting the export decision, export participation by firms can be explained using the following model. Let E_{ij} be a dummy equalling one if firm i exports to country j and zero otherwise. The model for this dummy variable becomes

$$E_{ij} = \begin{cases} 1 & \text{if} \quad \alpha + \delta_i + \delta_j > 0 \\ 0 & \text{if} \quad \alpha + \delta_i + \delta_j \leq 0 \end{cases} \tag{7}$$

where

$$\alpha = \frac{1}{\epsilon - 1} \log \mu - \log c \tag{8}$$
$$\delta_i = \log a_i \tag{9}$$
$$\delta_j = \frac{1}{\epsilon - 1} \log \left(\frac{Y_j^{\theta}}{F_j} \right) - \log (\tau_j) \tag{10}$$

This provides a compact formulation of the model's prediction of a strict hierarchy of export destinations. Trade costs and GDP combine to determine a "hurdle bar" for each export market and only firms with productivity levels above that bar will be observed exporting to that market. Technically, if this model was correct, a perfect fit could be obtained from estimating a Probit model for the combined $N_f N_m$ observations on E_{ij} by including firm- and destination-specific fixed effect dummies as explanatory variables.

2.2. A More General Model

The prediction of a strict hierarchy of export destinations is one that emerges from any model in which firms differ only across one dimension, in this case productivity. However, as Eaton, Kortum and Kramarz (2007) have demonstrated for French firms, and as can be confirmed for our dataset of Irish firms, deviations from the strict hierarchy are regularly observed in the data. Here, we generalize the model in a number of ways that are consistent with deviations from the hierarchy prediction. We start by generalizing the demand curve for firm i in country j to allow for preferences shocks. Specifically, we assume

$$Q_{ij} = \eta_{ij}^{Q} P_{ij}^{-\epsilon} Y_j^{\theta} \tag{11}$$

where $\log \eta_{ij}^{Q}$ is a zero mean random variable.

We also assume that trade costs have both firm-specific and market-specific elements to them. For instance, firms with a particularly small, light, or durable product may have systematically lower variable trade costs. Alternatively, some firms may face particularly high fixed trade costs in a specific market if, for instance, there was significant regulatory red tape related to its product in that country. A general model of these ideas is as follows:

$$F_{ij} = F_j \omega_i^F \eta_{ij}^F \tag{12}$$

$$\tau_{ij} = \tau_j \omega_i^{\tau} \eta_{ij}^{\tau} \tag{13}$$

In this case, both fixed and variable trade costs have an element that is market-specific (F_j and τ_j), an element that is firm-specific (ω_i^F and ω_i^{τ}) and an idiosyncratic element related to the firm-market combiniation (η_{ij}^F and η_{ij}^{τ}).

Under this formulation, profits are positive as long as

$$a_i > \left(\frac{F_{ij}}{\mu \eta_{ij}^{Q} Y_j^{\theta}} \right)^{\frac{1}{\epsilon - 1}} \tau_{ij} c = \bar{a}_{ij} \tag{14}$$

And the model for export participation becomes

$$E_{ij} = \begin{cases} 1 & \text{if} \quad \alpha + \gamma_i + \gamma_j + u_{ij} > 0 \\ 0 & \text{if} \quad \alpha + \gamma_i + \gamma_j + u_{ij} \leq 0 \end{cases} \tag{15}$$

where α is defined in equation (8) and

$$\gamma_i = \log\left(\frac{a_i}{\omega_i^\tau}\right) - \frac{1}{\epsilon - 1}\log \omega_i^F \tag{16}$$

$$\gamma_j = \frac{1}{\epsilon - 1}\log\left(\frac{Y_j^\theta}{F_j}\right) - \log(\tau_j) \tag{17}$$

$$u_{ij} = \frac{1}{\epsilon - 1}\left(\log \eta_{ij}^Q - \eta_{ij}^F\right) - \log \eta_{ij}^\tau \tag{18}$$

This implies a probabilistic model for firm participation in export markets such that

$$\text{Prob}\,(E_{ij} = 1) = \text{Prob}\ \left(u_{ij} > -\alpha - \gamma_i - \gamma_j\right) \tag{19}$$

If the combined idiosyncratic error term u_{ij} is normally distributed, then the statistical data generating process for the $N_f N_m$ observations on export participation is a Probit model with firm- and market-specific fixed effects whose fit will depend upon the importance of the various idiosyncratic shocks. In this case, the firm effect does not simply correspond to firm-level productivity but also includes terms related to systematic firm-specific factors influencing fixed and variable trade costs.

2.3. Export Market Sales

Having derived the model determining whether a firm will choose to export to a particular market, we now look at the model's predictions for how much it will sell. From equation (3), firm i's export sales in market j, should it decides to sell there, will be

$$S_{ij} = \eta_{ij}^Q \left(\frac{\epsilon - 1}{\epsilon}\frac{a_i}{\tau_j \omega_i^\tau \eta_{ij}^\tau c}\right)^{\epsilon - 1} Y_j^\theta \tag{20}$$

Thus, the complete model determining export sales can be written as

$$\log S_{ij} = \begin{cases} \kappa + \beta_i + \beta_j + v_{ij} & \text{if} \quad \alpha + \gamma_i + \gamma_j + u_{ij} > 0 \\ 0 & \text{if} \quad \alpha + \gamma_i + \gamma_j + u_{ij} \leq 0 \end{cases} \tag{21}$$

where γ_i, γ_j and u_{ij} are defined in equations (16)-(18) and

$$\kappa = (\epsilon - 1)\log\left(\frac{\epsilon - 1}{\epsilon c}\right) \tag{22}$$

$$\beta_i = (\epsilon - 1)\log\left(\frac{a_i}{\omega_i^{\tau}}\right) \tag{23}$$

$$\beta_j = -(\epsilon - 1)\log(\tau_j) + \theta \log Y_j \tag{24}$$

$$v_{ij} = -(\epsilon - 1)\log\left(\eta_{ij}^{\tau}\right) + \log \eta_{ij}^{Q} \tag{25}$$

Two aspects of this model are worth emphasizing. First, the firm- and country-specific fixed effects in the export sales equation, β_i and β_j, are related to the corresponding fixed effects in the export participation equation as follows:

$$\beta_i = (\epsilon - 1)\gamma_i + \log \omega_i^F \tag{26}$$

$$\beta_j = (\epsilon - 1)\gamma_j + \log F_j \tag{27}$$

Thus, a comparison of the fixed effects in the participation and sales equations provides information about firm- and market-specific fixed trade costs. This is because fixed trade costs affect participation decisions but do not affect *ex post* export sales. So, for example, systematic differences across firms in their fixed trade costs will result in the pattern of firm effects in the export participation equation being different to corresponding pattern in the sales equation.

Second, to the extent that there are idiosyncratic firm-market elements to preferences and variable trade costs, the model exhibits the classic features of the Heckman (1979) sample selection model. Formally, this can be seen in the following relationship between the residual terms in the export participation and export sales equations:

$$v_{ij} = (\epsilon - 1) u_{ij} + \eta_{ij}^F \tag{28}$$

The regression for export sales is based on a selected sample, i.e. only those markets that firms were observed exporting to. Because export participation is positively correlated with u_{ij}, this means that the sample of export sales observations are likely to have a value of u_{ij} that is greater than zero on average. To the extent that the idiosyncratic random errors in the first and second stage are correlated, this selection problem will result in ordinary least squares estimates of (21) being biased. Heckman's solution to this problem is to include the inverse Mills ratio λ_{ij} derived from the first-stage regression, as this provides an unbiased estimate of $E(u_{ij}|E_{ij} = 1)$. Adding the inverse Mills ratio to the export sales regression implies a residual term which is a linear combination of η_{ij}^F and the sampling error $u_{ij} - E(u_{ij}|E_{ij} = 1)$ and so has a zero mean. A priori, however, we cannot be sure whether this sample selection bias is important. Equation (28) implies that if idiosyncratic firm-market shocks to fixed trade costs are subtantially larger than the other idiosyncratic shocks, then there will be very little correlation between the v_{ij} and u_{ij} and the Heckman adjustment will make little difference.

2.4. Aggregation and Gravity Regressions

Before moving on to our empirical analysis, we describe how this framework can be used to provide a new interpretation for the coefficients of aggregate “gravity” regressions for trade flows. To consider the model’s predictions for the behavior of aggregate trade flows in our data set, note that total export sales from our model country to market j will be

$$S_j = \left(\frac{\epsilon - 1}{\epsilon c \tau_j}\right)^{\epsilon-1} Y_j^{\theta} \left[\sum_{a_i > \bar{a}_{ij}} \eta_{ij}^{Q} \left(\frac{a_i}{\omega_i^{\tau} \eta_{ij}^{\tau}}\right)^{\epsilon-1}\right] \tag{29}$$

This can be summarized as

$$\log S_j = \nu + (\epsilon - 1)\log \tau_j + \theta \log Y_j + \log \Omega_j \tag{30}$$

where

$$\Omega_j = \sum_{a_i > \bar{a}_{ij}} \eta_{ij}^{Q} \left(\frac{a_i}{\omega_i^{\tau} \eta_{ij}^{\tau}}\right)^{\epsilon-1} \tag{31}$$

is a composite term that combines the firm-specific component of sales of all the firms able to sell in market j.

The literature on gravity regressions usually links τ_j with distance, via an assumption such as $(\epsilon - 1)\log \tau_j = \zeta \log d_j$. Replacing the unobservable τ_j with distance thus gives

$$\log S_j = \nu + \zeta \log d_j + \theta \log Y_j + \log \Omega_j \tag{32}$$

The modelling framework we’ve used here is similar to the one used by Helpman, Melitz and Rubinstein (2008) in a paper that examines aggregate bilateral trade flows in a large sample of countries. Their paper points to a potential bias in the estimation of this type of regression without controlling for unobserved terms such as the Ω_j term here. Within our framework, this bias works as follows. Suppose we were interested in estimating ζ, the effect distance has in reducing each firm’s export sales. Then direct estimation of (32) will be biased. This is because an increase in d_j will raise the productivity threshold $\bar{a}_{ij}$ for each firm and thus reduce the composite term Ω_j: This negative correlation between an unobserved error term and a right-hand-side variable will cause the estimated coefficient on distance to be biased downward. Conversely, the same argument suggests that the coefficient on GDP will be biased upwards. In the more standard case of aggregate data on a full set of bilateral trade flows (so the regressions is for $\log S_{ij}$ for i and j running over N different markets) the same argument applies.

Helpman, Melitz and Rubinstein have presented a method for solving this problem while still using data on aggregate trade flows. Their method uses the fact that many countries do not trade with each other at all and this leads to a form of sample selection bias of the form just discussed: At least one firm in country i must cross the threshold required to export to country j before i-to-j trade flows are recorded in the data and used in a regression. Within the Melitz framework, zeros and non-zeros in aggregate trade figures can be modelled using a similar Probit model to the firm-level one discussed above, only in this case the Probit relates to the most productive firm in an economy: If this firm is not productive enough to meet the threshold for exporting to country j, then no firm in the economy will be. Helpman, Melitz and Rubinstein show that the predicted probabilities from an aggregate Probit for the existence of trade flows at the country level can be used to construct an estimator of an unobserved term like the Ω_j term above, so that the inclusion of this term can deal with the bias due to firm-level heterogeneity.

Our dataset, however, allows us to directly estimate how distance affects firm-level export sales. Making the same assumption about the link between variable trade costs and distance, equation (21) becomes

$$\log S_{ij} = \begin{cases} \kappa + \beta_i + \zeta \log d_j + \theta \log Y_j + v_{ij} & \text{if } \alpha + \gamma_i + \gamma_j + u_{ij} > 0 \\ 0 & \text{if } \alpha + \gamma_i + \gamma_j + u_{ij} \leq 0 \end{cases} \tag{33}$$

which can be consistently estimated with firm-level data (subject to the potential bias due to sample selection). In Section 5, we provide such estimates and illustrate the role of firm-level heterogeneity by comparing these estimates with those based on aggregating over our data.

3. Data Set and Descriptive Statistics

We now turn to describing our dataset and presenting some descriptive statistics relating to the predictions of the model just described.

3.1. The Enterprise Ireland Survey

The data used in this paper come from a survey of Irish firms undertaken by Enterprise Ireland, a government agency charged with promoting indigenous Irish owned businesses.[4] Due to Enterprise Ireland's focus, the survey collects data on Irish-owned and predominantly exporting firms. Of the 751 firms in the sample, 676 exported at some point during the period covered by the dataset. The survey reports firm-level data on five years of exporting activity (2000-2004). Comparing the total exports of the firms covered by this survey to the Census totals from the Irish Central Statistics Office (2000-2004), the data cover approximately two-thirds of exports from Irish-owned firms.

The restriction to Irish-owned firms means that this dataset is not representative of Irish exports as a whole. In 2004, foreign-owned companies accounted for just over 90 per cent of the country's manufacturing exports (Central Statistics Office, 2004). This is primarily due to a history of economic policy focused on attracting export-platform foreign direct investment. However, it is clear that the Irish experience of FDI-dominated exports is a relatively uncommon pattern. As such, we believe that studying the export decisions and patterns of indigenous Irish firms is more likely to yield conclusions that apply more broadly across countries.

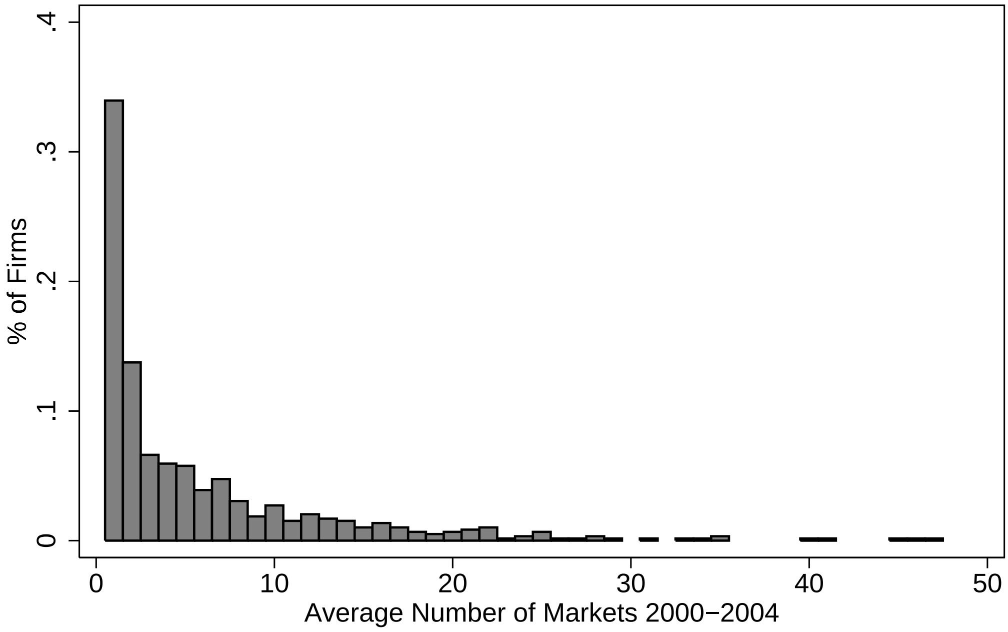
The Enterprise Ireland survey records information on a number of firm characteristics such as employment, sales, inputs, and exporting activity. More importantly for our analysis, the survey records detailed information on exports to 53 individual markets and is a panel, so that individual firms can be followed over time. Taken together, these features make the Enterprise Ireland dataset a particularly valuable tool for assessing the heterogeneous-firm approach to trade theory outlined in the previous section.

The other datasets used in this literature each have had some, but not all, of the features of the Enterprise Ireland survey. For instance, the French dataset used by Eaton, Kortum and Kramarz (2004) has comprehensive figures on export destinations but is a single cross-section and does not report many additional firm characteristics. The US Annual Survey of Manufacturers has been used in key papers such as Bernard and Jensen (1999, 2004) to establish differences

[4] A separate agency, the Industrial Development Agency, is responsible for attracting foreign direct investment and promoting foreign-owned businesses. The data from the Enterprise Ireland survey were made available to us by Forfás, which is the Irish national policy advisory board for enterprise, trade and technology.

Figure 1

Distribution of Firms by Market Coverage



between exporters and non-exporters but does not decompose exports by destinations, while the US Census Bureau's Linked-Longitudinal Firm Trade Transaction Database, used by Bernard, Jensen, Redding and Schott (2007), contains very detailed transactions-level data on exports by product and destination but has no additional data on firm characteristics. Thus, we believe the Enterprise Ireland dataset, while relatively modest in scale, is uniquely suited to addressing the predictions of models such as that presented in the previous section.

3.2. Descriptive Statistics

In the next few sections, we will formally assess the fit of the model developed in the previous section by estimating equation (15) for export participation and equation (21) for sales. However, it is useful to start with a few summary statistics which point in the direction of the results from our more formal analysis. Figure 1 and Table 1 provide summary information on export market participation by the firms, with the figure illustrating differences across firms in the number of export markets and the table breaking down some of this information by firm size, as measured by numbers of employees.

The principal message from these figures is that there is a very wide variation in the level of engagement in international trade among exporters. Consistent with the findings of Eaton, Kortum and Kramarz (2004) for France and of Bernard, Jensen and Schott (2006) for US firms, most firms export to only a small number of markets, with over one-third exporting to a single market. However, some firms export to many destinations: The average number of markets exported to over the five-year period was 5.9, with a median of 2.8. This finding of a highly skewed distribution has also been reported in previous studies. Only 17% of the firms in this paper export to more than 10 markets and just 3% to more than 25. Eaton, Kortum and Kramarz (2004) found approximately 20% of firms exporting to more than 10 markets and reported 1.5% exporting to over 50.

The Melitz framework explains variations in export participation as being the result of differences across firms in productivity levels. Table 2 provides a preliminary analysis of this idea, comparing value-added per employee for firms with different levels of export market coverage. The table shows that the relationship between a firm's number of export markets and its level of value added per worker is very weak, with a positive correlation only showing up clearly in the fact that the small number of firms that export to at least six markets are clearly at the high end of the productivity distribution.

More consistent with the model is the fact that firms that export sales to the UK increase steadily with the number of export markets that a firm participates in. Average sales per export market, however, *does not* increase as firms add more markets. This is because the additional markets tend to be more marginal markets with lower GDP and higher trade costs. Informally, one can see this from Table 3 which lists the average number of firms in our sample that participate in each of the 53 export destinations covered by the survey. The table certainly suggests that GDP, distance from Ireland, and sharing a common language appear to be important factors determining the number of firms that choose to participate in an export market. We provide a more formal analysis of this issue in the next section.

Table 1: Market Coverage and Firm Size (Average 2000-2004)

		Firm Employment					
	All Firms	1-24	25-49	50-99	100-249	250-499	500+
Average Markets	5.93	4.70	4.87	5.93	8.05	12.29	9.88
Median Markets	2.80	2.00	2.00	3.20	5.40	9.20	7.10
% Exporting to 1 Market	0.34	0.43	0.40	0.28	0.23	0.16	0.13
% Exporting to 2-5 Markets	0.33	0.30	0.35	0.41	0.27	0.14	0.31
% Exporting to 6-10 Markets	0.15	0.14	0.11	0.15	0.23	0.26	0.20
% Exporting to 11-25 Markets	0.14	0.11	0.12	0.13	0.22	0.32	0.30
% Exporting to $>$ 25 Markets	0.03	0.02	0.02	0.03	0.05	0.12	0.07

Table 2: Firm Characteristics and Market Coverage (Average 2000-2004, EUR'000s)

Markets	Employment	Value-Added per Emp.	Exports	Sales per Market	UK Sales
1	55	49	1978	1978	1878
2	55	50	2681	1341	2191
3	106	45	5995	1998	4482
4	71	42	4771	1193	2627
5	85	48	6375	1275	3986
6-10	121	61	10979	1391	5073
11+	166	106	29095	1509	8611

Table 3: Average Number of Exporters by Destination, 2000-2004

	Exporters		Exporters
UK	584	Saudi Arabia	40
USA	228	Hong Kong	36
Germany	213	Hungary	38
France	210	China	39
Netherlands	183	S. Korea	31
Italy	144	Taiwan	32
Spain	136	India	35
Belgium	139	Brazil	23
Sweden	122	New Zealand	33
Denmark	110	Malaysia	31
Portugal	76	Egypt	26
Switzerland	87	Philippines	21
Japan	75	Argentina	19
Norway	74	Kuwait	23
Canada	71	Mexico	24
Austria	69	Lebanon	17
Finland	78	Nigeria	22
Poland	61	Slovak R.	14
Australia	65	Slovenia	19
South Africa	56	Jordan	17
Greece	59	Thailand	20
Russia	43	Pakistan	17
Israel	53	Chile	15
Turkey	41	Algeria	7
Czech R.	46	Morocco	8
UAE	44	Tunisia	5
Singapore	40		

4. Where Do Firms Export?

Our analysis of the patterns underlying firm exports by destination begins with estimation of equation (19) in which firm and market dummies are used as proxies for the underlying systematic factors determining export participation. We then examine the extent to which observed variables capture the systematic patterns determining firm-level exporting propensity and the attractiveness of countries as export destinations.

4.1. Assessing The Hierarchy Hypothesis

Our model predicts that, in the absence of random firm-market-specific shocks, then we would observe a strict hierarchy of export markets. In other words, that we would observe our sample of firms entering into markets according to the pattern suggested by Table 3: Firms should enter the UK first, then the US, then Germany, and so on. However, only a small fraction of the firms in our sample conformed to this prediction. In 2004, 97% of firms that exported to only one market did indeed export to the UK. However, only 32% of the firms that exported to two markets choose the UK and US and once one goes beyond firms that exported to more than two markets, hardly any chose these markets in accordance with the strict hierarchy implied by Table 3. In this sense, it is clear that hierarchy-based models fit the data very poorly. However, this metric is somewhat harsh. For instance, the fourth most popular market is France, and a firm that exports to three markets and chooses the UK, US and France is hardly deviating significantly from the hierarchy hypothesis.

With this in mind, a more formal way to assess the hierarchy prediction is from the fit of the Probit model implied by equation (19) for export participation. One sense in which our dataset does not match the model is that our sample is a panel, so that each firm is observed over five years. Because the firm's underlying productivity is likely to be changing each year, the Melitz model would suggest that its position relative to the various threshold bars could also change each year. As such, rather than use a single firm dummy, we assess the hierarchy hypothesis using a Probit regression for E_{ij} with firm-year (D_{it}) and country (D_j) dummies as explanatory variables. In this sense, for now, we are treating firm i at time t as essentially a different firm from firm i at time $t-1$, i.e. treating our dataset as a repeated cross-section rather than as a panel. In Section 6, we will move on to explicitly utilize the panel element of the dataset.

Restricting our sample to firms that are observed exporting to at least one market, our dataset yields 158,586 firm-market observations of ones and zeros to explain. Table 4 reports two different measures of fit from Probit regressions with firm-year and country fixed effects to explain this series. Because these regressions have large numbers of explanatory variables, rather than report the traditional pseudo-R^2, we report McFadden's adjusted pseudo-R^2, defined as

$$\text{Adj. Pseudo}R^2 = 1 - \frac{\log L^{mod} - K}{\log L^{int}} \tag{34}$$

where $\log L^{mod}$ is the model's log-likelihood, $\log L^{int}$ is the log-likelihood for a model featuring only an intercept and K is the number of explanatory variables: This adjustment works in a similar fashion to the traditional adjustment associated with the $\bar{R}^2$. In addition, because the vast majority (about 89%) of the observations on E_{ij} are zeros, we also assess the model based on the fraction of observed firm-market combinations that it "predicts correctly" in the sense of generating a predicted probability of over 0.5.

The use of dummy variables for each firm-year is designed to pick up the effects of systematic variation at the firm level on export participation; this includes productivity differences across firms but also other factors such as the nature of the product and or other firm-specific variables that affect trade costs. The approach of using firm-year effects is made possible because we have observations for each firm in fifty-three destinations in each year. A model based on firm-year effects alone has an adjusted pseudo-R^2 of 0.17 and only predicts 8% of the observed firm-market pairings. Country dummies are designed to pick up all factors associated with market size and destination-specific trade costs; including these dummies alone also gives an adjusted pseudo-R^2 of 0.20 and predicts 16% of the observed pairings. Combining these two sets of dummy variables produces a model of export participation with an adjusted pseudo-R^2 of 0.45 and which predicts 38% of the observed firm-market pairings.

These results show that, as would be expected, systematic differences across firms and markets account for much of the observed pattern of export participation by firms. However, the data still show a very significant idiosyncratic element: Many of the observed firm-market export matches cannot be explained by either the systematic properties of the firm or the properties of the market. Within our theoretical framework, equation (18) shows that these random elements can be interpreted as the combination of random firm-market shocks to preferences, variables trade costs, and fixed trade costs. In the next section, we show how export sales regressions provide evidence that these random shocks largely relate to random firm-market elements in fixed trade costs.

4.2. Explaining Firm- and Country-Specific Effects

Having established how well the observed patterns of export participation by the firms in our dataset can be explained by firm and market effects, we now examine how much of this explanatory power can be ascribed to *observable* variables related to firm and markets. In particular, the

Table 4: Measures of Fit for Dummy Variable Probit Models of Export Participation

Export Dummy	Adjusted Pseudo R^2	Exporters Predicted
Firm-Year Effect	0.17	0.08
Country Effect	0.20	0.16
Firm-Year & Country Effects	0.45	0.38
Observations	158586	

Notes: Probit regressions of E_{ijt} on dummy variables for firm-year D_{it} and country D_j. Exporters Predicted column reports the percentage of firms exporting to a particular destination that the model assigns a predicted probability of over 0.5 (i.e. the percentage of export-destination pairs the model determines correctly).

substantial theoretical literature based on Melitz-style models has emphasized differences across firms in productivity as the key determinant of differences in export participation. Our dataset contains a direct measure of productivity, value-added per employee, as well as a couple of variables that may be useful proxies such as sales per employee, wages per employee and sector information.

Table 5 addresses this question by repeating the Probit regressions with country and year dummies but replacing the firm-year dummies with data on firm characteristics and sector dummies. If the Melitz model is correct in its assumption that the main source of heterogeneity across firms comes from differences in productivity, these characteristics should go a long way toward explaining the variation captured by the firm dummy.

As expected, all three variables were positively and significantly associated with export participation. However, in terms of the fit of the model, the observed firm characteristics fall well short of explaining the amount of systematic variation in the data. Relative to the fit of the benchmark model (with its adjusted pseudo-R^2 of 0.45), the best fit using firm characteristics is an adjusted pseudo-R^2 of 0.26 in columns (3) and (4), using sector dummies and either value-added per employee or sales per employee. When one considers that regressions using country dummies alone have an adjusted pseudo-R^2 of 0.20, it is clear that these observable proxies for productivity are doing little to explain the systematic differences in export participation across firms. Recall, that the less formal calculations reported in Table 2 also suggested a limited explanatory role for productivity. Thus, firm-specific factors other than productivity must play a key role in explain-

Table 5: Firm Characteristics and Export Participation

Dependent Variable: Export Dummy X_{ij}					
	(1)	(2)	(3)	(4)	(5)
Country Dummy	Yes	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes
Sector Dummy	Yes	No	Yes	Yes	Yes
Ln VA per Employee		0.20	0.21		
		(0.005)	(0.006)		
Ln Sales per Employee				0.33	
				(0.007)	
Ln Wage per Employee					0.39
					(0.011)
Adjusted $PseudoR^2$	0.24	0.22	0.26	0.26	0.25
Export Markets Predicted	0.16	0.24	0.24	0.17	0.16
Observations	158586	144492	144492	158546	158546

Notes: Probit coefficients reported. Standard Errors in parentheses. Exporters Predicted column reports the percentage of firms exporting to a particular destination that the model assigns a predicted probability of over 0.5 (i.e. the percentage of export-destination pairs the model determines correctly).

Table 6: Destination Country Characteristics and Export Participation

Dependent Variable: Export Dummy X_{ij}			
	(1)	(2)	(3)
Firm-Year Dummy	Yes	Yes	Yes
Ln GDP	0.41	0.39	0.34
	(0.004)	(0.005)	(0.005)
Ln Distance	-0.69	-0.57	-0.53
	(0.006)	(0.007)	(0.007)
Ln GDP per Capita		0.47	0.32
		(0.010)	(0.010)
English Dummy			1.01
			(0.021)
Adjusted $PseudoR^2$	0.36	0.39	0.41
Export Markets Predicted	0.25	0.28	0.31
Observations	158029	158029	158029

Notes: Probit coefficients reported. Standard Errors in parentheses. Exporters Predicted column reports the percentage of firms exporting to a particular destination that the model assigns a predicted probability of over 0.5 (i.e. the percentage of export-destination pairs the model determines correctly).

ing where individual firms export to. Our model suggests that systematic variations in fixed and variable trade costs across firms are potential candidates for explaining the remainder of the fit obtained using firm dummies.

One possible reaction to these results is that it is highly unlikely that a single firm characteristic could match the explanatory power of a model featuring a large number of dummy variables. And, indeed in simulations of our model, we have found that models featuring firm dummies can have a considerably higher psudo-R^2 values than "true" models in which firms only differ systematically across one variable (for instance, productivity). However, our use of an adjusted measure of fit is designed to deal with this problem. Simulation evidence shows that dummy variable models will obtain values for the adjusted pseudo-R^2 statistic that are approximately the same as the pseudo-R^2 for the true restricted model.

Attempts to replicate the fit of the benchmark model using observed *country* characteristics are considerably more successful than the performance of the models based on observed firm characteristics. Table 6 keeps the firm-year fixed effects but replaces the country dummy with a small number of observed country characteristics commonly used in the literature on gravity equations to explain aggregate trade flows.[5] All of the coefficients on the country characteristics have the signs expected from aggregate gravity regressions. Distance has a negative relationship with trade participation while GDP and GDP per capita, measuring market size and wealth respectively, are positive.[6] Sharing a common language is also positively associated with export participation at the firm level.[7]

The noteworthy feature of these results is how well this small group of variables can essentially replicate the fit of the benchmark model. Replacing the 53 country dummy with a small number of observable characteristics constitutes a significant restriction on the benchmark model. However, it turns out that very little explanatory power is lost from this restriction. The most restrictive model, including only GDP and distance, gives a pseudo-R^2 of 0.36, compared to 0.45 when country fixed effects were used. Adding GDP per capita and a dummy for common language give us an adjusted pseudo-R^2 of 0.41, meaning that almost all of the systematic element related to the market can therefore be explained with only these four variables.

[5]See Disdier and Head (2006) for a very useful summary of this literature.

[6]Data on GDP and population were obtained from the Penn World Tables (Heston, Summers and Aten, 2006). Distance is calculated as straight-line distances in kilometers based on the latitude and longitude of the capital cities. Available at www.indo.com/distance. The distance calculation is done using the 'geod' program, which is part of the 'PROJ' system available from the U.S. Geological Survey.

[7]We use a dummy variable equalling one if country j has English as (one of) its official language(s) and zero otherwise. From http://en.wikipedia.org/wiki/List_of_countries_where_English_is_an_official_language.

5. How Much Do Firms Export?

We now examine how much firms sell in different export markets once they participate. Equation (21) from our model implies that the log of export sales can be modelled in a parallel fashion to the model for export participation, though in this case a standard linear regression model can be used. Here we report results for regressions in which the dependent variable is $\log S_{ijt}$, the log of real export sales of firm i in country j in year t, where the Irish Wholesale Price Index has been used as the price deflator.

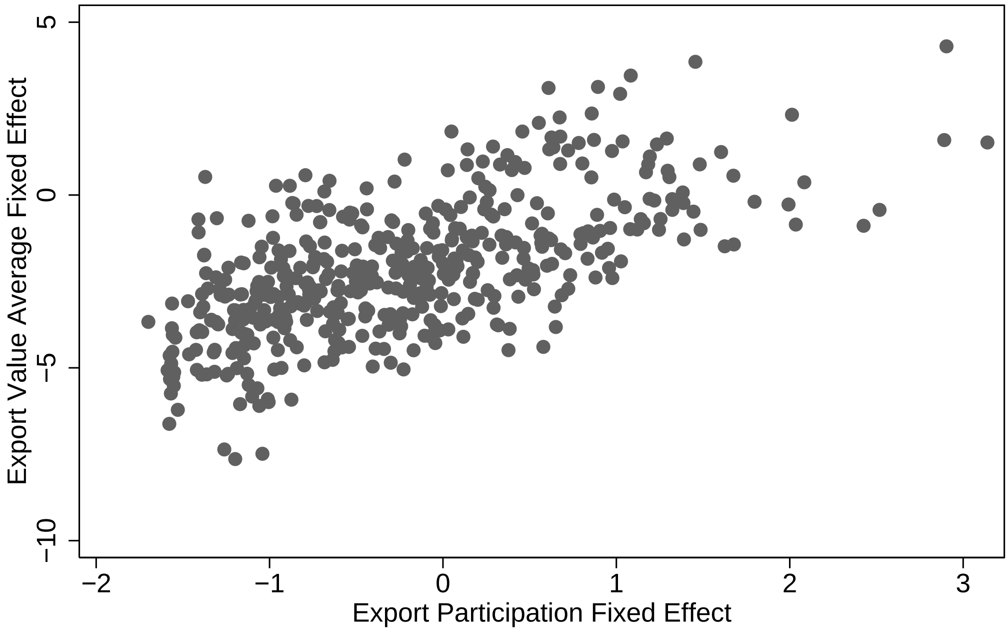
Goodness-of-Fit: Table 7 reports fit statistics for dummy-variable-based models. A model using firm-year dummy variables alone has an $\bar{R}^2$ of 0.38. (Again, the use of the adjusted R^2 is important in this case because the firm-level dummies can explain a lot of variation due to their ability to fit random sampling error.) A model using country dummies alone has an $\bar{R}^2$ of 0.16 while putting all the firm-year and country dummies together produces a model with an $\bar{R}^2$ of 0.57. Again, paralleling our findings for export participation, the data on export sales still suggest a considerable role for randomness related to specific firm-market matches that cannot be explained by either the characteristics of the firm or the characteristics of the market.

Table 7: Measures of Fit for Dummy Variable Models of Export Values

Ln(Export Value)	Adjusted R^2
Firm-Year Effect	0.38
Country Effect	0.16
Firm-Year & Country Effects	0.57

Notes: OLS regressions of $\log S_{ijt}$ on dummy variables for firm-year D_{it} and country D_j.

Figure 2
Firm Fixed Effects for Export Participation and Export Value



Relationships Between Dummy Variable Coefficients: Turning to the dummy variable coefficients themselves, recall from equations (26) and (27), that our model suggests there is likely to be a relationship between these coefficients and their equivalents from the Probit regressions. Specifically, to the extent that systematic differences across firms are limited to productivity and variable cost differences, then the two sets of firm dummies should be highly correlated (with only sampling error explaining deviations). In contrast, to the extent that systematic differences across firms in fixed trade costs are important, then the firm effects from the first stage will be less correlated with those from the second stage. Similar arguments apply to the country dummies: To the extent that systematic differences across countries are limited to GDP and variable trade costs, then the two sets of estimates should be highly correlated.

Figure 3

Country Fixed Effects for Export Participation and Export Value

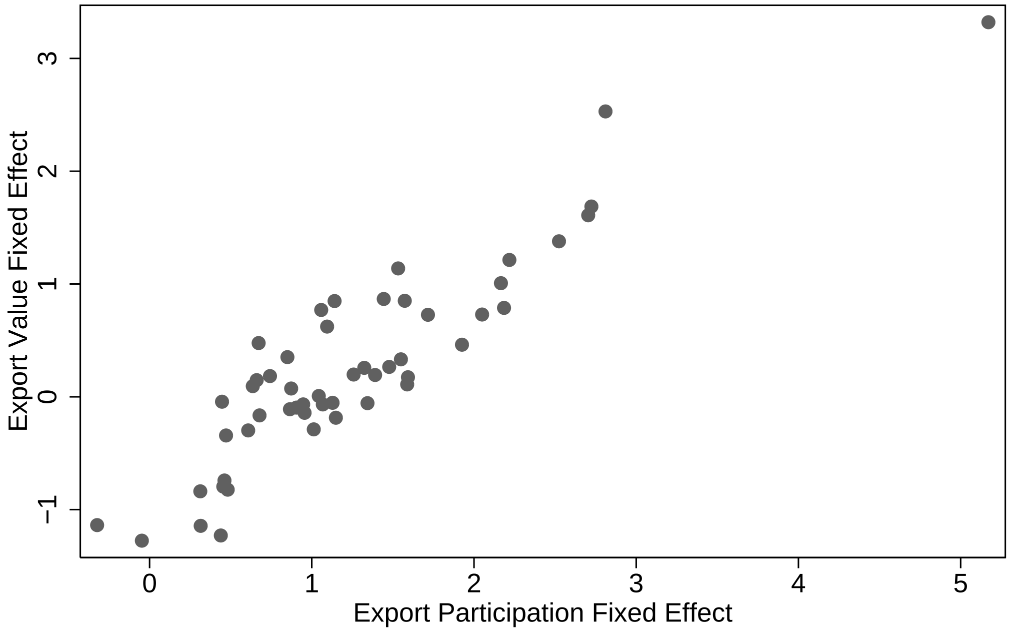


Figure 2 illustrates the correlations between firm-level fixed effects from the participation and sales regressions.[8] Because some firms have identical patterns of export participation and thus the same Probit coefficients, the figure shows the average value of the firm effect in the sales regression for each separate value for the firm effect in the participation equation. While the chart shows a reasonably strong positive relationship, consistent with a correlation of 0.6, there is still plenty of random variation around this pattern, suggesting that some firms may have systematically low or high fixed trade costs. Figure 3 shows the corresponding graph for countries, graphing the 53 Probit country fixed effects for export participation against the country fixed effects for export sales. This relationship is clearly much stronger (correlation of 0.91) suggesting little systematic variation across countries in fixed trade costs.

Observable Firm and Country Variables: As with the Probit regressions, we next turn to the question of which observable variables these firm and country dummies correspond to. Table 8 reports the results obtained from replacing the firm dummies with data on firm characteristics and sector dummies. The firm characteristic variables do better here than they did in explaining export

[8] To simplify the chart and to reduce the influence of temporary sampling errors, we have used a single firm dummy over the five years for each firm.

participation. Comparisons of Tables 4 and 5 show that the adding firm characteristics to a model with country dummies resulted in only a small improvement in the fit of the participation models. In this case, however, we see that a model that adds sales per employee and a sector dummy raises the $\bar{R}^2$ to 0.36, relative to a fit of only 0.16 for a model with country dummies alone. That said, these models still fall a long way short of the fit obtained by the firm fixed effects model. Perhaps most surprising is the poor performance as an explanatory variable of value added per worker, which produces a low fit of $\bar{R}^2 = 0.20$.

Recall from our discussion in Section 2.3 that this regression displays elements of the Heckman sample selection problem, and thus the coefficient estimates may be biased. To address this issue, Table 9 thus repeats the regressions from Table 8, but this time adding the standard "Heckman correction," i.e. the inverse Mills ratio ω_{ij}, derived from a Probit for export selection, with the two equations estimated jointly via maximum likelihood. Recall from equation (28) that v_{ij} (the residual in the export sales equation) is correlated with u_{ij} (the idiosyncratic element of the export participation equation) and ω_{ij} provides an unbiased estimator of the expected value of u_{ij} contingent on exporting being observed.[9] The results show that the addition of the inverse Mills ratio does not change the reported fit statistics (rounded to two-digits) at all and has little effect on the estimated coefficients.

Taken together, the results in Table 8 and 9 provide further support for a conclusion reached from our analysis of export participation: Despite the emphasis placed on differences in productivity as the key factor distinguishing firms in the large literature following Melitz (2003), our evidence points to a fairly limited role for productivity as the source of systematic firm differences in export participation and subsequent export sales. One potential criticism of this conclusion could be that our proxies for firm productivity—value-added per employee, sales per employee, average wage, and sector dummies—are perhaps poor proxies for the true underlying differences in firm productivity assumed in the Melitz model. While this criticism undoubtedly has some truth to it. That said, comparisons of data from Enterprise Ireland survey with corresponding figures from the Irish Census of Industrial Production suggest that the survey appears to be a reliable one, so we have little reason to believe the underlying data on value added or sales are problematic.

[9]While technically one can identify the model using an inverse Mills ratio obtained from a first-stage regression with the same list of explanatory variables as the second-stage regression, in practice this tends to produce a ω_{ij} that is highly correlated with the other explanatory variables. For this reason, it is standard to have a different list of explanatory variables in the first-stage. The regressions in Table 9 do this by including a firm dummy in the selection equation but firm characteristics in the sales regression.

Table 8: Firm Characteristics and Export Values

Dependent Variable: Ln Exports exp_{ij}					
	(1)	(2)	(3)	(4)	(5)
Country Dummy	Yes	Yes	Yes	Yes	Yes
Year Dummy	Yes	Yes	Yes	Yes	Yes
Sector Dummy	Yes	No	Yes	Yes	Yes
Ln VA per Employee		0.45	0.40		
		(0.016)	(0.016)		
Ln Sales per Employee				0.88	
				(0.018)	
Ln Wage per Employee					0.94
					(0.037)
Adjusted R^2	0.28	0.20	0.31	0.36	0.30
Observations	18226	16525	16525	18225	18222

Notes: Standard Errors in parentheses.

Table 9: Firm Characteristics: Full Selection Specification

	Probit Selection	*Dependent Variable: Ln Exports* exp_{ij} (1)	(2)	(3)	(4)	(5)
Firm-Year Dummy	Yes	No	No	No	No	No
Country Dummy	Yes	Yes	Yes	Yes	Yes	Yes
Sector Dummy	No	Yes	No	Yes	Yes	Yes
Ln VA per Employee			0.42	0.35		
			(0.016)	(0.017)		
Ln Sales per Employee					0.86	
					(0.020)	
Ln Wage per Employee						0.82
						(0.038)
λ_{ij}		-0.58	-0.20	-0.33	-0.08	-0.38
		(0.034)	(0.037)	(0.037)	(0.035)	(0.035)
Adjusted R^2	0.45	0.29	0.20	0.31	0.36	0.30
Observations	158029	18226	16525	16525	18225	18222

Notes: Standard Errors in parentheses.

Table 10: Destination Country Characteristics and Export Values

	Dependent Variable: Ln Exports exp_{ij}			Heckman Two-Step	
	(1)	(2)	(3)	Participation	Ln Exports
Country Dummy	No	No	No	Yes	No
Firm-Year Dummy	Yes	Yes	Yes	Yes	Yes
Ln GDP	0.58	0.56	0.49		0.49
	(0.010)	(0.010)	(0.011)		(0.011)
Ln Distance	-0.52	-0.47	-0.39		-0.39
	(0.014)	(0.014)	(0.014)		(0.014)
Ln GDP per Capita		0.35	0.20		0.20
		(0.026)	(0.026)		(0.026)
English Dummy			1.11		1.11
			(0.044)		(0.044)
λ_{ij}					9.65
					(8.80)
Adjusted R^2	0.52	0.53	0.55	0.45	0.55
Observations	18226	18226	18226	158029	18226

Notes: Standard Errors in parentheses.

Table 10 reports results obtained from replacing the country dummies in the benchmark export sales model with the same set of observed characteristics used in Table 6 for export participation. As in that case, and unlike the case for firm characteristics, this small number of country characteristics essentially mimics the fit of the benchmark dummy variable model, with the model based on four variables having an $\bar{R}^2$ of 0.55, just a touch below the fixed effects model. Thus, the combination of GDP, GDP per capita, distance, and an English language dummy, together explain all of the systematic information that was previously captured by 53 separate dummy variables. As with the firm characteristics regressions, the addition of the inverse Mills ratio has essentially no effect on the fit of the regression or the size of the model's coefficients, suggesting that the bias associated with the selection problem in not an important one.

Because the good fit of this regression suggests that it is less likely to be mis-specified than the firm-characteristic regressions in Table 9, it is perhaps more appropriate in this case to provide a structural interpretation of the role of the Heckman adjustment. The inverse Mills ratio adds little to the fit of this regression and is not statistically significant. As we noted in Section 2.3, within the context of our model, this result can be interpreted as evidence that most of the idiosyncratic variation in firm export participation in export markets reflects random variation in their country-specific fixed trade costs. These influence export decisions but then have little influence on subsequent export sales.

Composition Effects in Aggregate Gravity Regressions: The results in Table 10 are unusual in comparison with the large literature on gravity regressions because they estimate the effects of distance and other trade frictions on exports using firm-level data rather than aggregate trade flows. As discussed above in Section 2.4, firm-level heterogeneity means that the coefficients of aggregate regressions are likely to combine two effects: The intensive margin effect estimated in Table 10 (the effect on individual firms' export sales) and extensive margin effect (the effect due to changing the number of firms and type of firms exporting).

By aggregating our data over all our firms (thus reducing our sample from 18,266 firm-level observations to 252 country-level observations) we can demonstrate the magnitudes of these two sets of effects. As described earlier, because variables such as distance and GDP have an impact of the same sign on both the intensive and extensive margins, we would expect that the magnitude of coefficients in an aggregate gravity regression should be larger than in the firm-level regression in Table 10. The results from the aggregate gravity regression reported in Table 11 confirm that this is the case, with each of the coefficients on our four explanatory variables in Table 11 being larger in magnitude than the comparable coefficients in Table 10.

For example, our results show an aggregate elasticity of Irish exports with respect to distance of -0.51.[10] This can be composed into an effect of -0.39 due to the intensive margin on firm sales and an additional effect of -0.12 due to the extensive margin. While the magnitude of our distance elasticity is lower, our conclusion that 24% of this effect is due to an extensive margin effect is similar to Helpman, Melitz and Rubinstein's (2008) conclusion—reached using a very different aggregate methodology—that 30% of their estimated distance elasticity of -1.17 was due to this margin.

Table 11: Country-Level Gravity Model

Dep. Variable: $LnExports_{ij}$	
Ln GDP	0.74
	(0.05)
Ln Distance	-0.51
	(0.08)
Ln GDP per Capita	0.53
	(0.09)
English Dummy	1.21
	(0.24)
Adjusted R^2	0.61
Observations	252

Notes: Standard Errors in parentheses.

6. Dynamics of Exporting at the Firm Level

Up to now, our approach has been to test the generalized Melitz framework by treating our panel essentially as a repeated cross-section. Because each firm can have a different productivity level each year, we adopted the approach of treating firms as though they are a new firm each year. Another reason we adopted this approach is because the model itself is a static one and a firm's past decisions have no direct effect on the present. However, even within this framework, the

[10]This is lower than the median distance elasticity of -0.9 found by Disdier and Head (2008) in a meta-analysis of 103 gravity model papers. However, they reported that 90% of estimates were between -0.28 and -1.55, so our result is well within the standard range.

firm's past track record in exporting could have lasting *indirect* effects if, for instance, they had an influence on the trade costs associated with exporting to various markets. For example, if fixed trade costs have a partially "sunk" element to them, then past participation in an export market may reduce the fixed costs associated with participation this period and thus raise the probability of continuing participation.[11] Similarly, a high volume of sales in the past may produce cost-saving efficiencies that reduce variable trade costs today. With these considerations in mind, we generalize our framework to examine the effect of a firm's exporting history on its current position.

Starting with export participation, we define a dummy variable E_{ijt} to equal one if firm i exports to country j at time t and zero otherwise. We estimate the effect of past exporting behavior using a Probit model of the form:

$$\text{Prob}\left(E_{ijt}=1\right)=F\left(\beta, D_{it}, D_{j}, E_{ij,t-1}\right) \tag{35}$$

where D_{it} is a firm-year dummy capturing the firm's current characteristics, D_j is a country dummy, and β is a vector of parameters. The left-hand column of Table 12 shows that the marginal effect of having $E_{ij,t-1}=1$ is precisely estimated at about one-half. In other words, independent of other factors related to firm i's current export participation or the features of market j, the firm's participation in that market last period raises the probability that it will sell there this period by 0.51. This shows that there is a substantial correlation across years (but within firms) in the deviations from the static hierarchy model. When compared with the fit measures reported in Table 4, it is clear that incorporating an effect of lagged market participation produces a much better empirical fit. The adjusted pseudo-R^2 rises from 0.45 to 0.65 and the fraction of observed firm-market pairings predicted by the model (in the sense of a predicted probability greater than a half) rises from 0.38 to 0.80.

While the marginal effect of past participation is sizable, this estimate also tells us that deviations from a hierarchy model are also likely to be relatively transitory. Consider, for instance, the case of firm that the hierarchy model predicts should have little or no probability of exporting to a particular market. If the firm is observed exporting to that market, these estimates predict that, *ceteris paribus*, the firm will have a one-half chance of exporting there one year later, a one-quarter chance two years later, and only a one-eighth chance three years later. This shows that while the strict hierarchy model provides a relatively limited fit for the observed data on firm export participation by market, the deviations from the model's predictions tend to be relatively transitory.

[11] In referring here to costs being partially sunk on a market-by-market basis, we have quite a different model in mind than the well-known sunk costs model of Roberts and Tybout (1997) which describes sunk costs relating to entering and exiting exporting altogether.

Table 12: Dynamic Model of Export Participation

Dependent Variable	Export Dummy	Export Value
Lagged Export Dummy	0.51	
	(0.009)	
Lagged Export Value		0.81
		(0.006)
Firm-Year & Country Effect	Yes	Yes
Adjusted Pseudo R^2	0.65	
Export Markets Predicted	0.80	
Non-Exporting Predicted	0.98	
Adjusted R^2		0.86
Observations	122293	12188

Notes: Marginal effects from probit regression reported. Standard Errors in parentheses. Exporters Predicted column reports the percentage of firms exporting to a particular destination that the model assigns a predicted probability of over 0.5 (i.e. the percentage of export-destination pairs the model determines correctly).

Finally, the right-hand-side of Table 12 reports the results from adding a lagged dependent variable to our export sales regressions, i.e. regressing S_{ijt} on $S_{ij,t-1}$, D_{it}, and D_j. A comparison with Table 7 shows that the fit for the export sales model rises from 0.57 to 0.86. The size of the lagged dependent variable effect is also quite large at 0.81. This suggests that there is more persistence in the variable cost and preference factors that affect sales than in the fixed trade cost factors that only influence market entry.

7. Conclusions

The growing literature on heterogeneous firms and exporting patterns has been an important recent development in international trade theory, with Melitz (2003) being the most influential contribution. While the underlying fact that exporters are more productive than non-exporters is well-established, there has until now been little systematic empirical analysis of the predictions of the heterogenous-firm models for firm-level patterns of trade across destinations. Our paper has provided such an assessment using a panel survey of Irish firm which, thanks to its combination of data on export destinations and firm characteristics, provides a useful testing ground for these models.

A key stylized prediction of models based on firm differences in productivity is that export market participation should show a distinct "hierarchy" pattern, with firms entering markets according to a specified order with the number of markets entered dependent on the firm's level of productivity. Our analysis sheds light on two aspects of this prediction. First, we show that no single firm factor can explain the observed data on export market participation. The hierarchy prediction falls well short of explaining the observed pattern of firm-export market combinations in our data, with substantial amounts of random heterogeneity evident. Second, we show that to the extent that there are systematic firm-level factors determining the extent of export participation, these factors are poorly correlated with the various measures of, and proxies for, firm productivity available in our dataset. It appears that other factors, for instance systematic differences across firms in trade costs, may explain much of the observed variation.

A second prediction of Melitz-style models is that the same factor that determine the extent of a firm's export participation (productivity) should also determine its relative amount of export sales. We find some evidence in favor of this idea—firm fixed effects in participation and export sales equations are positively correlated—but also evidence for considerable random variation unexplained by this hypothesis. Our generalized model points to systematic differences across firms in their fixed trade costs—which affect entry decisions but not sales—as a possible explanation for this finding.

Our paper also makes a number of new contributions to the existing literature gravity relationships for trade flows. We show that a small number of country variables, such as distance and GDP, do an excellent job of capturing the systematic factors explaining how many firms will choose to export to a particular destination as well as how they will sell there. We also provide calculations to illustrate how the coefficients in aggregate gravity regressions combine two different effects: An intensive margin related to the effect on individual firms' export sales and an extensive margin effect related to changing the number of firms and type of firms exporting).

Finally, we illustrate some elements of the dynamics over time in firm exporting patterns by destination. We show that lagged exporting activity has a significant effect on a firm's current exporting profile. Most notably, previous participation in a particular market raising the probability of current participation in that market by about 0.50.

The results here suggest a number of avenues for further research. Identifying which observable factors determine systematic differences across firms in their extent of export participation and export sales, would appear to be an important research topic, as would the development of dynamic versions of the Melitz framework which can then be tested with panel data of the type used here. More practically, we believe that empirical models of the type estimated in this paper can be used effectively for various types of scenario and policy analysis. For instance, the model could be used to project future changes in export participation and sales related to specific projections for GDP growth across a range of countries. Alternatively, the model could be used to assess the implications of potential structural changes such as a reduced effect of distance on trade costs.

References

[1] Bernard, Andrew and J. Bradford Jensen (1995). "Exporters, Jobs and Wages in US Manufacturing: 1976-1987" *Brookings Papers on Economic Activity: Microeconomics*, pages 67-112

[2] Bernard, Andrew and J. Bradford Jensen (1999). "Exceptional Exporter Performance: Cause, Effect or Both?" *Journal of International Economics*, 47, 1-25.

[4] Bernard, Andrew and J. Bradford Jensen (2004). "Why Do Some Firms Export?" *Review of Economics and Statistics*, 86, 561-569.

[4] Bernard, Andrew, Jonathan Eaton, J. Bradford Jensen and Samuel Kortum (2003). "Plants and Productivity in International Trade," *American Economic Review*, 93, 150-152

[5] Bernard, Andrew, J. Bradford Jensen, Stephen Redding, and Peter Schott (2007). "Firms in International Trade," *Journal of Economic Perspectives*, 21, 105-130.

[6] Central Statistics Office (2000-2004). *Census of Industrial Production*, Dublin

[7] Chaney, Thomas (2008). "Distorted Gravity: Heterogeneous Firms, Market Structure and the Geography of International Trade," *American Economic Review*, 98, 1707-1721.

[8] Disdier, Anne-Celia and Keith Head (2008). "The Puzzling Persistence of the Distance Effect on Trade," *Review of Economics and Statistics*, 90, 3748.

[9] Eaton, Jonathan, Marcela Eslava, Maurice Kugler and James Tybout (2007). "Export Dynamics in Colombia: Firm-Level Evidence" *NBER Working Paper* No. 13531

[10] Eaton, Jonathan, Samuel Kortum and Francis Kramarz (2004). "Dissecting Trade: Firms, Industries, and Export Destinations," *American Economic Review*, 94, 150-152.

[11] Eaton, Jonathan, Samuel Kortum and Francis Kramarz (2007). "An Anatomy of International Trade: Evidence from French Firms", working paper, CREST.

[12] Heckman, James (1979). "Sample Selection Bias as a Specification Error," *Econometrica*, 47, 153-161.

[13] Helpman, Elhanan, Marc Melitz and Yona Rubinstein (2008). "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics*, 123, 441-487.

[14] Lawless, Martina (2007). "Firm Export Dynamics and the Geography of Trade," Working Paper 2/RT/07, Central Bank and Financial Services Authority of Ireland.

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

[16] Roberts, Mark and James Tybout (1997). "The Decision to Export in Columbia: An Empirical Model of Entry with Sunk Costs," *American Economic Review*, 87, 545-564