WORKING PAPER 2009:2

David Greenaway

Joakim Gullstrand

Richard Kneller

Firm Heterogeneity and the Geography of International Trade

AgriFood economics centre

Revised draft, April 22nd 2009

Firm Heterogeneity and the Geography of International Trade

David Greenaway,
GEP, University of Nottingham

Joakim Gullstrand,
Lund University and GEP, University of Nottingham

Richard Kneller,
GEP, University of Nottingham

Abstract

An important contribution of heterogeneous-firm models is the need to consider the determinants of exports at the intensive and extensive margins. This paper exploits detailed firm data on the destination of exports and finds empirical evidence that trade flows are affected in a manner consistent with this theory. In addition we compare our results to those generated using data where information on firms or on destinations is not available. As Helpman, Melitz and Rubenstein (2008) demonstrate for the gravity model, and we additionally show for the firm export model, the relationship between firm and destination characteristics on trade volumes are biased upwards. However, we also find a second source of bias for these models from the interaction between destination and firm characteristics.

JEL codes:F10; F14
Keywords: exports, heterogeneous firms, gravity model

The authors acknowledge helpful comments from two anonymous Referees and From Steve Redding on an earlier draft. Financial support from The Leverhulme Trust, under Programme Grant F/00 114/AM is also acknowledged

I Introduction

Recent years have seen the heterogeneous firm model emerge as the principal model used to explain micro patterns of international trade (Melitz, 2003; Bernard, Eaton, Jensen and Kortum, 2003; Bernard, Redding and Schott 2007). Fundamental to the predictions of this class of models is the interaction between productivity differences across firms in the same industry and the up-front fixed costs incurred at each export market. Only the best firms can cover these costs and still make positive profits in the most physically and culturally distant markets (Chaney, 2008; Helpman *et al*., 2008).

The emergence of this new theory of international trade has led in turn to a re-evaluation of the dominant empirical model used to explain aggregate trade flows, the gravity model, where the volume of trade between two countries increases with their economic mass and declines with trade resistance (Tinbergen, 1962). The heterogeneous-firm model brings to the gravity model a need to consider the number of firms that export and to where. The first to consider this were Helpman, Melitz & Rubenstein (2008) (hereafter HMR). Using a two-stage estimation procedure that corrects for both sample selection bias, which occurs when there is zero trade between country-pairs, and for asymmetries in trade flows between countries, which arise because of differences in the fraction of firms that export, they find that current testing of the gravity model conflates the intensive margin of exporting with what is in fact adjustment at the extensive margin. Together these bias the relationships found for key variables in the gravity model such as distance by around one-third of its previous value.

In this paper we exploit detailed data on the destination of exports for each firm to understand the contribution that firm characteristics brings to a gravity model of international trade. Key predictions from the heterogeneous firm model that we seek to test include whether more productive firms serve a larger number of markets; links between firm exports and the size of, and distance from, foreign markets; and that between these same variables and total export sales by a firm to a particular market. Using detailed census data on the Swedish Food and Beverage sector from 1997 to 2002 we find empirical support for these predictions of the heterogeneous firm model.[1] Firms that are more productive and larger are more likely to serve foreign markets in particular if these markets are large and relatively close. They also export greater volumes to these markets.

In addition to the analysis of the importance of firm and destination characteristics in the gravity model we also compare our results to those generated when information on destinations or firm characteristics were not available. We find from this exercise that the availability of firm-destination export data does more than just add fine detail to the existing aggregate gravity, or firm level, modelling of export behaviour. As HMR (2008) demonstrate for the gravity model, and we additionally show for the firm export model, the relationship between firm and destination characteristics on trade volumes are biased upwards by standard methodologies. However, while the HMR (2008) correction deals with the bias that arises from conflating adjustment at the extensive and intensive margins it does not fully capture the effect of firm characteristics on the intensive margin. We find strong evidence that the effect of destination variables on firm exports are also dependent on the characteristics of firms (there is an interaction between these terms). This is a further source of (omitted variable) bias for the aggregate gravity and the firm-export model. Together these results suggest that firm-destination data may be important in understanding the effects of changes in trade policy on aggregate trade flows.

More generally in using firm-destination export data we contribute to a recent literature that has begun to exploit newly available data on the destination of trade by firms to reveal how the components of aggregate trade flows, such as varieties, quantities and unit values, respond to various characteristics of trading partners. Important work here includes the exploration of the anatomy of international trade by Eaton, Kortum and Kramarz (2004) for France and Bernard, Jensen and Schott (2006) for the US.

The remainder of the paper is organised as follows. Section II briefly refers to the theoretical literature on firm heterogeneity and exporting, setting out the hypotheses in which we are particularly interested. Section III explains our modelling framework. In Section IV we detail

---

[1] Focusing on a single sector has the important advantage that differences in the elasticity of substitution, which can generate different consequences for changes in trade costs on the margins of trade, are likely to be less important (Chaney, 2008). The disadvantage is that we must caution about generalising from our results.

the data sources and some stylised facts evident in the data, while in Section V we discuss our results for our empirical estimations. Finally, Section VI concludes.

II Heterogeneity and Firm Exports

A common starting point for the formulation of firm export sales is based on the 'new trade theory' perspective synthesised in Helpman and Krugman (1985). Using an assumption of identical preferences across countries implies that any demand effects on trade patterns are neutralised, and following an assumption of "love for variety", as in Krugman (1980), that consumers around the world always demand a product as soon it is produced. The pattern of trade is therefore solved once we determine where each product or variety is produced. This is ensured by monopolistic competition and differentiated products.[2] If we also assume a variable transport cost for exporting, implying price differences across countries, we derive the following export volume of firm f in country i to country j:

$$x_{ji}^{f} = \lambda Y_j \left[\frac{p_{jf}}{P_j} \right]^{1-\varepsilon}, \quad P_j = \left[\int_{l \in E_j} p_{jl}{}^{1-\varepsilon} dl \right], \tag{1}$$

where p_{jf} is the price of variety f in country j, Y_j is country j's income, ε is a constant demand elasticity, P_j is country j's ideal price index, λ is the utility function's distribution parameter across products, and E_j is the set of products available at market j. The price of f's products on market j depends on the demand elasticities, factor prices in j, and transport costs between production locations i and market j. This demand function is similar to the demand for "region i goods by region j consumers" as used in Anderson and Van Wincoop (2003) but for a single variety produced in i. The price index in market j depends on the costs of exporting from all locations to market j, and hence it is labelled the "multilateral trade resistance" variable in Anderson and Van Wincoop (2003). That is, any shift in trade costs between two partners affect an importer's propensity to import from all regions because of changes to relative prices.

---

[2] Other possibilities to determine specialisation patterns across countries are products differentiated across countries (Anderson, 1979; Anderson and Wincoop, 2004) or factor proportion/technology differences (Deardorff, 1998; Haveman and Hummels, 2004).

If we also assume, as in Helpman *et al* (2008) and Chaney (2008), that firms are heterogeneous, some are more efficient than others, the price of firm *f*'s variety to market *j* also depends negatively on its productivity level. The higher its productivity, the higher its export volumes. Finally, if all firms also face a fixed cost of exporting, firm *f* only serves market *j* as long as exporting is profitable, which is dependent upon its efficiency of production. This implies firms select themselves into export activities, and whether a firm elects to export to market *j* depends on its productivity level and the fixed costs of exporting to that particular market. In this setting we have the following firm-level export equation:

$$x_{jf}(a, a_f \geq a^j) = \lambda Y_j P_j^{\varepsilon-1}\left[m\tau_{ij}a_f\right]^{1-\varepsilon}, \tag{2}$$

where a_f is firm *f*'s productivity level, a^j is the productivity of the firm which is indifferent to exporting to market *j*, *m* is a constant mark-up (=$\varepsilon/[\varepsilon$-1]), and τ_{ij} is the variable transport cost. Although this firm-level equation is comparable to export equations of representative-firm models, it differs in several very important aspects because firms now select into exporting to market *j*. As shown in Melitz (2003) and Chaney (2008), the selection process is driven by the existence of sunk-export costs, and one firm may export to one country and not to another because the sunk costs differ across export destinations. The sunk-cost does not, however, affect the intensive margin of trade. This usefully implies that these costs may serve as an exclusion restriction when we correct for selection effects (see Helpman *et al,* 2008).

In addition, equation (2) implies that selection into exporting may affect the outcome of the firm's export volume to a particular market. Hence there is a risk of generating biased coefficients on both firm and country characteristics in a regression of the intensive margin. As in Helpman *et al* (2008) where the selection process is believed to bias the effect of trade resistance variables, the selection effect may also be important when multilateral firm-export volumes are considered. A firm's total export volumes may be affected by how much it sells to each market but also its success in selling to lots of destinations.

More generally models of this type can explain a number of stylised facts evident in the micro literature. The model is consistent for example, with evidence that not all firms within an industry export, that the extensive margin will vary across destinations - increasing in the size of the foreign market and decreasing with the fixed and variable costs of exporting - and that export behaviour is correlated with firm productivity. The most productive firms will serve the largest number of markets and the revenue earned in that market is proportional to its productivity. As described in HMR (2008) it is also capable of explaining why no trade may exist between pairs of countries (the zeros in the data) - no firm has a productivity level above the threshold value for that market a^j – but also asymmetries in the aggregate volume of trade between countries – depending on the productivity draws that firms' within each country receive. As we describe below these are all features of our data.[3]

III Empirical Specification

Heterogeneous firm models describe changes in aggregate export volumes using adjustment at the intensive margin, the volume of exports by each exporter, and at the extensive margin. The latter is made up of the decision for each firm to export and to which markets. The gravity model of bilateral trade estimated by HMR (2008) deals with the bias induced by these extensive margins on the determinants of trade flows such as distance. The gravity model they estimate is of the form:

$$m_{ij} = \beta_0 + \lambda_j + \chi_i - \gamma d_{ij} + w_{ij} + \Omega_{ij} + u_{ij}$$

Where i and j refer to countries, m refers to the log of trade, χ and λ are fixed importer and exporter effects respectively, d refers to distance between i and j and u_{ij}, is a random error term. Of the variables that are new relative to the standard gravity model, w_{ij} controls for the fraction of firms exporting from i to j and Ω_{ij} the is the inverse Mills ratio to control for unobserved country specific factors that lead to positive trade flows. These affect the coefficient on distance γ in opposite directions. Without w_{ij} the effects of trade variables on the volume of trade include changes in the number of firms exporting, which induces an upward bias in the coefficient on distance, whereas for Ω_{ij} the bias is predicted to be downward. The authors further derive the conditions necessary to provide consistent estimates of these key parameters and establish robustness to different estimation methods.

[3] As we have exports from Sweden we do not provide evidence on the last point.

The availability of firm-destination data allows us to side-step these complications and to directly observe both aspects of the extensive margins of trade. A possible sample selection bias remains however; as we continue to measure the effects of trade resistance on firm trade volumes only when trade flows by a firm to a given destination are positive. We therefore continue with the use of the Heckman (1979) correction for sample selection.

The benchmark specification of the gravity equation we use to describe the intensive margin of firm *f's* exports to market *j* is a reduced form of equation (2):

$$x_{fjt} = \alpha_0 + \sum_k \beta_k z_{kjft} + \sum_l \beta_l z_{lijt} + \gamma_j + \lambda \Phi_{ftj} + \varepsilon_{fjt}, \qquad (3)$$

where lower-case letters indicate logged variables, x_{fjt} is the export volume of firm *f* to importer *j*, z_{kjft} is a set of K explanatory firm-level variables, z_{lijt} is a set of L explanatory export-destination variables including bilateral trade resistance variables, γ_j is an export-destination effect, and Φ_{fj} is the mills ratio controlling for unobserved characteristics leading to export success.

Although the heterogeneous-firm model emphasises the primacy of productivity as the determinant of firm export decisions the empirical literature has found that a number of other firm characteristics are also important (Greenaway and Kneller, 2007). Within z_{kjft} we therefore include measures of firm size (measured by the number of employees), ownership status (foreign owned, owner of foreign firms or domestic) and alongside a measure of TFP.[4] In addition to market size we include in z_{lijt} aspects of trade barriers between countries. A country's propensity to import is affected by its trade relations with all countries, which underlines the importance of controlling for multilateral trade-resistance. One way to do this is to introduce time-invariant export destination effects, to take account of unobserved price indices effects.[5] In our sample this would make it impossible to estimate the effects of time-invariant bilateral effects of variables of interest such as distance. For this reason we use instead regional export-destination effects (the 19 regions are presented in Table A4 in the Appendix). We also include measures of distance, membership of the EU, whether the one of

[4] See Table A1 for variable definitions and sources.
[5] See Rose and Wincoop (2001). An alternative specification is to solve these price indices implicitly (as in Wincoop and Anderson, 2003).

the pair is English speaking or not, dummies for low and middle-income countries and the real exchange rate.

The impact of the selection into exporting is represented by the inverse Mills ratio, which is calculated having estimated a selection equation measuring the probability that firm *f* will export to market *j*. To specify this equation we draw on the existing empirical literature on the participation of firms into exporting. We define our selection equation as: [6]

$$\Pr(D_{jft} = 1 \mid observables) = \Phi\left(\sum_k \delta_k z_{kjft} + \sum_l \delta_l z_{lijt} + \delta_j D_{jft-1} \right), \qquad (4)$$
$$D_{jft} = 1 \mid x_{jft} > 0,$$

where z_{kjf} is a set of K explanatory firm-level variables, z_{lij} is a set of L explanatory country-level variables, δ_j and δ_j are estimated coefficients, D_{jft-1} is lagged export status (1 if the firm exported *t-1*, 0 otherwise), and δ_j is an estimation of the importance of sunk-cost of exporting (or the importance of last year's export decision on this year's).

According to the heterogeneous firm model, participation decisions are determined completely by a combination of sunk-costs and firm productivity. In the empirical counterpart to this, the set of firm characteristics has been extended to include factors such as firm size, age, human capital, relative capital-intensity (human as well as physical) and ownership. While there are differences in the exact methodology employed (the choice over logit or probit models and attempts to correct for bias from inclusion of lagged export status of the firm) results are for the most part robust. Our firm level controls include a measure of firm productivity, ownership (owned by a foreign firm or owner of foreign firms), size (number of employees), capital intensity (physical and human).[7] We would expect all to have a positive association with the margins of exporting. All these indicate whether a firm is successful or not on foreign markets, and hence we use lagged (one period) characteristics to avoid problems of endogeneity.

---

[6] This selection equation is similar to the parameterised reduced-form of export activity in Roberts and Tybout (1997) as well as in Bernard and Jensen (2004).
[7] See Table A1 in the Appendix for variable definitions.

The decision to export to a particular country also depends on the characteristics of the export destination, which uniquely we have the opportunity to analyse. We measure market size by including trading partners GDP, and include population of the importing country to capture the possibility that richer countries may spend a greater share of their income on tradables (Anderson and Van Wincoop, 2003). To capture the effects of bilateral trade resistance variables we include measures of distance, membership of the EU15, and a dummy indicating whether the importing country is low or middle-income. We also include exchange rate information, and in the selection equation consider exchange rate risk since a firm may avoid markets with high exchange rate fluctuations. (Our measure of exchange rate risks is the differences between the maximum and minimum exchange rate divided by its mean).

The exclusion restrictions needed to identify the sample-selection model are based on theoretical as well as empirical considerations. As discussed in Helpman et al (2008), the heterogeneous-firms model identifies possible variables that can be used as exclusion restrictions as the sunk-costs of exporting affect only the extensive margin. To capture the 'sunk-costs' of export market entry we therefore follow Roberts and Tybout (1997) and Bernard and Jensen (2004) and include the lagged export status of the firm in the selection equation. Bernard and Jensen (2004) also include the ratio of white collar to total employees as a proxy for work-force quality. We include such a measure alongside a measure of capital intensity.[8] We anticipate capital-intensive firms to be more likely to export as Sweden is relatively capital abundant.[9] We also include a measure of population in the selection equation to capture possible demand differences among trade partners. A larger population given the GDP level implies a lower per capita income, which is associated with a smaller demand for tradables. Hence it is less likely that firms enter markets with low per capita income. Finally, we follow evidence presented in Tenreyro (2007) that suggests that exchange rates volatility does not have a robust effect on exports to include it in the selection equation, although we control for the level of the exchange rate in both the regressions for the intensive and extensive margins.

IV Data and Summary Statistics

[8] In support of these variables as exclusion restrictions we find that these variables enter the firm-gravity model with insignificant coefficients.

Our firm-level data is provided by *Statistics Sweden* and consists of an unbalanced panel of 1,570 firms in the Food and Beverage sector covering the period 1997 until 2002. For all years we have detailed information for each firm on its output, choice of factor inputs, details of its ownership structure and most importantly on export volumes to each of 138 export destinations. A detailed definition of variables used can be found in Table 1. Combining firms, with destinations across time yields a total sample of 1.3 million observations.[10]

The stylised facts of firm export behaviour match those in other country and industry settings. For example, we find exporters are relatively rare in our data: only around 20 per cent of our firms export. Second, this export behaviour is strongly associated with firm characteristics such as the size of the firm. Around 16 per cent of firms with less than 50 employees export, while for bigger firms the figure is close to 80 per cent.[11] The majority of the data, around 70 per cent, consists of firms with less than 50 employees. Combining this with information on destinations we find that, because most firms export to a concentrated number of foreign markets, exporting is even rarer. According to the figures in Table 2 just 0.74 per cent (9,858 observations) of the available cells in the data contain positive trade flows.

The pattern of these firm-destinations over time is rather dynamic. Export durations are relatively short and there is a high degree of export exit. We find, for example, that around 23 per cent of all firm-destination export observations change their status to non-exports (about 1,800 observations) and around 0.2 per cent of all non-exporting observations (around 2,000 observations) change to exporting at some point during the sample window. This rate of exit is considerably greater than entry/exit rate of firms into and out of exporting more generally and implies a relatively low degree of persistence in export decisions.

Comparing the characteristics of firms and export destinations for cells in which trade flows are positive and those for which they are zero in Table 2 confirms our general expectations about export behaviour. The table shows for example, that the average distance between Sweden and cells for which trade flows are positive is 2,789km compared to 6,441km for

---

[9] Haveman and Hummels (2004) argued for that technology differences among countries could be used to identify trade partners and hence explain zero bilateral trade flows at an aggregated level.
[10] Note that we only consider firms that exist for at least three sequential years.
[11] There also exists a small number of self-employed within the data set. Around 4 per cent of these firms export.

firm-destinations for which no trade exists, while the average GDP for those same cells are $773bn and $195bn respectively. Firm exports are concentrated on nearby and large markets. These differences in mean are significant at the 1 per cent level. We similarly find that the markets that Swedish firms export to have (on average) large populations, are within the European Union, do not speak English as a first language and have a relatively high share of agricultural production. The latter may indicate that particular export destination has a relatively large food and beverage sector, which implies a relatively competitive environment and relatively high demand for intermediates. Again these differences in mean are all significant at the 1 per cent level or better.

The characteristics of firms across firm-destinations also differ significantly. Firms for which firm-destination combinations are non-zero are for example larger. The average size of firms for positive firm-destination exports is 436 compared to just 33 for non exporters, or if measured by the average annual sales, the respective figures are SK1,072 million and SK73 million. The table also suggests that they are also more productive, use more skill intensive production techniques, and are more likely to be part of a firm with an international ownership structure (measured by either their foreign or domestic MNE status).

Further evidence on the importance of country characteristics is contained in Figure 1, which shows the distribution of firms and exports across markets in 2002. From this there would appear to be a strong relationship between distance and exports on both aggregate and micro trade flows. The most important markets for Swedish exports, measured by their share of total Swedish exports in the food and beverage sector (x-axis) and the share of exporters that sell to that market (y-axis), are other Nordic countries (Norway, Denmark, Finland and Iceland) followed by Germany, US, UK and France. The effect of distance would appear to be quite strong. Beyond this relatively small number of countries the percentage of exporters that export to those destinations and their share in total Swedish exports declines very quickly. Most countries represent for less than 2 per cent of total exports from Sweden, while less than 2 per cent of firms export to these markets, such that it quickly becomes difficult to clearly identify countries within the graph. The figure also makes clear however, that it would be a gross over simplification to suggest that all Swedish exporter export to countries that are close. The Nordic countries account for more than 30 per cent of Sweden's exports but only 20 per cent of all exporting firms export there. The hierarchy of export destinations suggested by Chaney (2008) would appear to hold only in general in our data. Lawless (2009) reaches a similar conclusion using firm-destination export data for Ireland. Finally, the US is the obvious outlier in Figure 1, which suggests a strong role for market size on the intensive margin of trade. The US is the most important market for Swedish exports when countries are ranked by export volume, it accounts for around 30 per cent of all exports, but is served by only four per cent of all firms.

Figure 2 plots the number of firms exporting to a given market together with the average number of markets firms that serve that market export to. It reveals a pattern consistent with the expectation that successful firms export to a greater number and to more marginal markets. For example, in 2002 there are 180 firms that export to Norway. On average these 180 firms export to around 8 different markets. In comparison only four firms export to Argentina, who on average export to 40 different markets. Countries penetrated by many Swedish firms tend to attract firms that export to only a few markets. The extensive margin of trade declines as quickly in Sweden as in France (Eaton et *al.,* 2004).

In Figure 3 we provide further detail on the relationship between firm characteristics and export behaviour, the number of destinations each firm exports to and the productivity of that firm. This result is in line with the theoretical predictions. The relationship is, however, not one-to-one, a simple Poisson regression of TFP on the number of export destinations reveals that a unit increase in TFP increases the number of export destinations by 1.22. A more detailed inspection of Figure 3 also reveals that only a handful of firms export to a large number of markets, and one firm that exports to twice the number of destinations as the firm ranked second by number of export destinations. This exceptional exporter is a Swedish multinational. It accounts for between 20-30 percent of total export volumes each year, although it accounts for around three per cent of total sales within the industry. In comparison the next nine firms with the largest number of export destinations are also multinationals (Swedish or foreign) but on average account for around 2.5 of total exports and 2 per cent of total output in the Food and Beverage sector. Consistent with Bernard, Redding, Jensen and Schott (2007) total exports from Sweden are concentrated amongst a small number of firms.

V Gravity at the firm level

Our formal econometric results are presented in Tables 3 through to 5, with additional tests of robustness presented in tables A1 to A3 in the Appendix. We use a relatively large number of estimations in order to compare the results to those found when either information on firms or destinations are not available and to consider the sensitivity of our results to different estimation methodologies. In table 3 we present results exploiting all firm and destination information available in the dataset. Table 4 then replicates those found in the aggregate-trade literature, in that they omit firm information, and Table 5 the methodology found in the firm-trade literature, they ignore the information we have on the destination of exports. We also include in Table 4 a regression estimated using the HMR (2008) correction.

We begin by making some general comments on the results, comparing across all of the tables, before returning to focus in more detail on the results in Table 3. Within Table 3 our estimations of the firm-destination gravity model are reported as regressions 2 to 4, with regression 1 our estimate of the model for the extensive margin and used to correct for sample selection in regressions 2 and 4. Regression 2 forms our baseline model. Regression 3 is used to observe the effect on the results of not correcting for selection, while in regression 4 we attempt to control for possible interactions between firm and destination variables. As we show below the results from regression 4 are useful in explaining the differences between regression 2 and the HMR (2008) correction. All of the regressions control for fixed time, region and industry specific effects and use country-level clustered standard errors.

Comparing across Tables 3 and 4 we find that our results support the view of HMR (2008) that traditionally estimated gravity models conflate the effects of trade resistance on exports with the selection of firms into exporting. In regression 6 we report the effect of distance and market size in a standard gravity model, where we also control for any selection bias using regression 5 to model the selection process (although interestingly the inverse Mills ratio (lambda) is insignificant in these regressions). We do find, as expected, that increased GDP significantly increases the volume of exports from Sweden, while this declines with increased distance. The estimated elasticities for these variables are 0.68 and -0.74 respectively. In regression 2 we find the same relationship between the volume of trade for each firm and distance and market size, but the elasiticities are smaller. In regression 2 the estimated elasticity with respect to market size is 0.17 and -0.27 for distance. Consistent with the argument in HMR (2008) we find that the effect of distance on the number of firms that export to a given destination (regression 1) decreases aggregate trade flows independently of the effect that distance has on the firms' intensive margin of exports found in regression 2. This causes the elasticity of exporting with respect to distance to be overestimated in the aggregate model (regression 6).

An interesting question that naturally arises is how similar are the results from the firm gravity model to those found using the approach outlined in HMR (2008), which we present as regression 9. In our data the HMR (2008) correction reduces the size of the coefficient on distance (regression 9) by around a quarter of its original value of -0.74 in regression 6, close to the HMR (2008) estimate of one third. In the firm-gravity models (regression 2) the reduction in the estimated coefficient is closer to two-thirds however.[12]

This raises a question of what explains this difference relative to HMR (2008). Why is the HMR (2008) correction smaller than that suggested by the firm gravity model? Before we consider that question we first establish that the results from the aggregate model that we are comparing to, is not generated as a consequence of the methodology used and also note that the effect of the HMR correction on the distance parameter is not the same as that for other country characteristics such as market size.

To establish the robustness of our estimated elasticities in the aggregate gravity model we also include in Table 4 the estimates from a traditional gravity model excluding all zero-trade flows (regression 7). Again we find that our estimate of the elasticity on distance is close to the average found in the literature (Anderson and Van Wincoop, 2004).[13] The results are independent of methodologies.

Secondly we note that the correction on distance behaves differently to that for market size. While it would appear that the coefficients for both distance and market size in the HMR regression (regression 9) are pushed in the right direction compared to regression 6, they both become smaller in absolute magnitude, regression 2 suggests that this is underdone for distance but overdone for GDP. The effect of a change in market size on the intensive margin

[12] For both of these models we can confirm using likelihood-ratio tests that the coefficients on distance are significantly lower than those found in regression 6.

[13] We did also investigated whether the results were influenced by any exceptional export market (Figure 1 suggests that the U.S. may be an outlier) by excluding the U.S. This had no impact on our results except for a moderated but still significant effect of TFP on firm exports.

of trade is smaller in regression 9 than regression 2, and unlike these other regressions is not statistically significant.

According to the results from regression 4 (Table 3) the difference between the HMR (2008) regression and those for the firm gravity model would appear to be a consequence an omitted variable bias within regression 9: the effect of destination variables on firm exports are also dependent on the characteristics of firms. In regression 4 we capture these through the inclusion of interaction terms between TFP with GDP and the number of employees with distance.[14] We find that the volume of trade to a given destination responds to an increase in market size disproportionately for the most productive firms, whereas the negative effect of distance in stronger for large firms.[15] Thus the HMR (2008) approach does not fully adjust distance and over adjusts GDP.

The apparently contradictory signs on the coefficients for the interaction terms might reflect the incentive for large firms to serve the most distant markets by establishing affiliates in those locations along the lines of the model market-seeking FDI outlined in Helpman, Melitz and Yeaple (2004). More generally it would seem that for firms that are far above the threshold entry point in the heterogeneous firm model changes in market conditions are more important compared to firms closer to the threshold.[16] These effects are missing in the HMR (2008) model, although they might be captured when the model is extended to include interaction terms between country characteristics and the distribution of the productivity of firms. Motivation for this approach comes from Yeaple (2006) who has previously included a measure of the industry distribution of productivity in a model explaining intra-firm trade in US industries. Given the availability of firm-destination export data here we choose not to pursue the effectiveness of such a modelling approach here. Finally we note that the same interaction effects are insignificant in a selection equation of the same form as regression 1. The HMR correction would appear to consistently estimate the extensive margin, but fails to fully model the effect of firm characteristics on the intensive margin.

---

[14] According to a likelihood-ratio test these terms significantly improve the fit of the regression.

[15] We estimated a regression using TFP interacted with distance and GDP and employment interacted with distance and GDP. We found that only the TFP-GDP and employment-distance interactions were statistically significant at conventional levels.

[16] Consistent with this result, Gullstrand (2008) finds for the same sector that large and productive firms have lower persistence in their export decisions, while Lawless (2009) finds that once Irish firms export to more than three markets they are more likely to change their mix of export destinations from one period to the next as retain them.

Given that the standard gravity model inaccurately estimates the effect of destination characteristics such as market size on trade because of the failure to account for the interaction between firm and destination variables, it seems likely that the same bias will occur for firm characteristics such as productivity in a model that focuses solely on the firm export market participation decisions. That is, the existing firm-export literature will inaccurately estimate the effect of firm characteristics on the intensive margin of trade. We consider this question by comparing the results in regression 2 with those found in regressions 3, 11 and 12. Regression 3 uses firm-destination data but differs from regression 2 in not controlling for selection bias, regression 11 also considers selection bias but abstracts from the destination of export sales, and regression 12 ignores both selection bias and the destination of exports. Regressions 11 and 12 are specified to be similar to those found in Bernard and Jensen (2004), Greenaway and Kneller (2008) and others in this literature.

Our results confirm that the coefficients are indeed biased. Relative to regression 2 we find that the impact of firm characteristics on firm-export volumes are larger when information on the destination of exports is unavailable, with or without correcting for selection effects (regressions 11 and 12). The marginal effects, calculated at the mean of the independent variables, are reported in curly brackets within the table. According to these, the effect of a 1 per cent change in productivity or employment on trade volumes in regressions 11 and 12 are about twice as strong when we exclude information on export destinations.

The comparison of regressions 2 and 3 also suggests selection bias may be important in regressions of this type. There is evidence of a significant correlation (indicated by lambda in Table 3) between the errors of the selection equation (regression 1) and the gravity equation (regression 2). That the correlation is negative suggests that the unobserved characteristics that make exporting by a firm to a given destination more likely, lowers the intensive margin of exporting. What might these unobserved characteristics be? Helpman *et al*. (2008) for example argue that the direction of bias from ignoring the extensive margins operates in opposite directions, although they find this is dominated by the upward bias from the endogenous number of exporters. Or consistent with the correlation that we find, Görg *et al.* (2008), using data on the products exported by Hungarian firms, find that on average firms export many products, most of which account for a very small share of total export sales. The average firm exports around 25 products (measured at the HS 6-digit level), but on average 20 of these account for less than 1 per cent of the total sales (and 15 less than 0.1 per cent). Firms appear capable of finding customers for many of their products, but often do not sell them in large quantities. Either way the results suggest that (unobserved) sunk costs may be less important for firm export decisions than is usually suggested by the heterogeneous firm model.

Focusing on the results in Table 3, of greater interest perhaps are the coefficients on distance and market size, where there exist no comparable regressions in the literature. In their meta analysis of gravity models (mostly country level) Disdier and Head (2008) report the average elasticity on distance is -0.9, with 90 per cent of estimates within the range -0.28 to -1.55. The estimated elasticity on distance of -0.27 in regression 2 lies just outside the bottom end of this range. Perhaps a better comparison comes from HMR (2008) who report an elasticity of around -1.1, as do Bernard *et al.* (2006) for their regressions of the number of exporters and number of products, the latter also using firm level export destination data. In regression 4 of course the effects of both GDP and distance on firm exports vary greatly with firm characteristics. Bringing together information on the distribution of employment with distance in our data then we find that the mean elasticity of distance on GDP is -0.26, close to that of course found without the interaction terms. Taking into account the distribution of employment across the sample the standard deviation for this elasticity is 0.24. For GDP the mean elasticity, controlling for the effects of TFP, is 0.11, with a standard deviation of 0.07. As Table 2 suggests the distribution of TFP in our data is somewhat more narrowly spread than employment.

Using the full range of values for employment in the data then we find that the elasticity for distance on export volumes ranges from -0.76 for the largest firm to 0.35 for the smallest. When we un-log the data (this a double-log regression) at no point does this suggest however that the effect of distance on trade volumes is positive. The effect of GDP has a slightly smaller range, although again passes through zero. Between the least and the most productive firm observation the effect of GDP on firm export volumes ranges from -0.36 to 0.38 (also using the log-log specification). In Figures 4-5 we shows the distribution of the elasticity of GDP and distance on firm exports, along with those for TFP and employment which we discuss below. The elasticity of trade volumes with respect to TFP is plotted in its own figure (Figure 5) due to a much greater variation for this variable. Figure 4 indicate that for distance the elasticity for most firms is towards the lower end of the distribution since firms in our sample mostly export to closer markets, such as the other Nordic countries. In the same manner the distribution for the GDP elasticitity is grouped towards the higher end since productive firms export more frequently.

Consistent with the heterogeneous firm model of international trade we also find a significant positive correlation between productivity and firm level exports in regression 2. More productive firms export in greater quantities to each market that they sell to. Comparing the magnitude of these two variables in regression 2, the results shows that the elasticity of TFP is larger than that for employment at 0.38 and 0.12 respectively. This difference is extenuated when we include the effects of the interaction terms in regression 4. In this regression at the mean level of GDP the elasticity of firm trade volumes to GDP is 0.49, with a standard deviation of 0.45, and is 0.10 for employment, with a standard deviation of 0.16. As suggested by Figure 4 the range for TFP is somewhat broad, ranging from -1.44 to 1.48.

Firm productivity would not appear to be capable of capturing all relevant characteristics of the firm however. We also find that the size of the firm is important; although ownership is not. Of interest to the literature on export platform FDI (see for example Yeaple, 2003 and Ekholm *et al.,* 2003) we find that foreign multinational firms behave in line with other firms when we correct for selection on a bilateral level (conditional on their productivity and size advantages). This is not the case in a multilateral setting since the results in regression 11 reveal that multinational firms export more even after we have controlled for productivity and size. In other words, multinationals export to more destinations but they do not tend to export higher volumes to each destination compared to domestically owned firms. This difference again probably reflects the upward bias on the coefficients measuring firm characteristics described above.

Finally, while our main focus is not the model for the extensive margin in regression 1, we note the similarity of our results to those found in the existing literature on sunk-costs of exporting at the firm level. We find that firms which are more productive, larger and more capital intensive (both physical and human) and are multinationals are more likely to export. We also find that previous export experience, usually interpreted as evidence of the sunk-costs of exporting (Bernard and Jensen, 2004), is an important determinant of export destination decisions at the firm level. As some attention has been given to the magnitude of coefficient on the lagged export status within the literature we do so here. According to our estimates a firm that exported last period is between 15-20 per cent more likely to export this period. The benefit of experience on where a firm exports to is much less important than that estimated for the decision to export or not suggested in existing studies. According to the results from regression 10 the marginal effect of experience on the decision to export or not in our data is 69 per cent.[17] For comparison the effects of persistence in similar countries has been estimated at 40 per cent for the US (Bernard and Jensen, 2004) and between 38 and 85 per cent (with 50 per cent seen as the most likely figure) for Germany (Bernard and Wagner, 1998), while Gullstrand (2008) shows for the Swedish food and beverage industry that a firm exporting last period is around 30-40 percent more likely to export this year. Connections between suppliers and their customers in foreign countries are much more transient, the sunk-costs of exporting are less important than existing estimates would suggest.

Differences for the effects of ownership gain exist between regression 1 and 10, but now for the extensive margin. In regression 10 we find no difference between domestically owned firms and foreign MNEs and a lower probability of exporting for Swedish multinationals. In contrast in regression 1 the coefficients on both foreign and Swedish MNEs are positive and statistically significant. As highlighted in the summary statistics, multinationals often export to a greater number of markets than domestically owned firms (the mean number of export destination are 12 and 4 respectively).

It is possible to model the country-specific determinants of trade only in regression 1. Of the country level variables again most are significant and in line with expectations. They show that Swedish firms are more likely to export when country *j* is closer to Sweden, and has a larger market size when measured by GDP. Conditional on these measures of market size and distance we also find it is less likely to trade with highly populated countries. This is consistent with the Balassa-Samuelson effect since poorer countries are expected to demand less tradables. In addition we find Swedish firms in the Food and Beverage industry are less likely to export to risky markets (measured by exchange rate volatility). The importance of local taste in the food industry and location of Sweden close to Norway (see Figure 1) accounts for the negative impact of EU15, low and middle-income countries, and English speaking countries on the probability of exporting.

---

[17] Given our interest in making the comparison with the persistence parameter in the firm-destination selection equation we make no attempt here to correct this estimate from the likely upward bias generated from the inclusion of a lagged dependent variable in the regression (Bernard and Jensen, 2004).

Robustness

An important implication of the heterogeneous-firms model is that it demonstrates clearly that exporters are a non-random sample and therefore there is a need to control for the selection process. If the selection process into exporting affects a firm's export volume, then we may draw erroneous inferences from the gravity equation. After using standard econometric techniques to correct for these biases, we found traditional approaches overestimated both trade resistance variables such as distance and firm characteristics. However, the standard techniques to control for selection biases do not consider the possibility of a time-invariant firm-destination effect. For that reason we test robustness of our results to the use of cross-section data and additionally the use of the panel data techniques outlined in Wooldridge (1995 and 2002).

The results are shown in Table A1. They are strongly consistent with the results already found. That is, the possibility of correcting for the selection process into exporting at firm-destination level seems to correct gravity elasticities downwards. An alternative approach is to use panel data models as in Wooldridge (1995 and 2002), in which linear projections of the fixed effects are used in both the selection and gravity equation. Results are presented in Table A2. They indicate that the between differences across firms are important for determining export volumes compared to marginal changes in firm size and TFP. If we include distance in the regression (regression 2 Table A2), then its effect becomes a compound of its own effect and the fixed effects. The elasticity of around -0.37 is still below that estimated without correction for selecting into exporting.[18]

As a final exercise we establish the robustness of our findings to the use of one-period lagged firm characteristics, used to reduce problems of simultaneity. Table A3 reports results when we use two-year (regression 1) and three-years lags (regression 3) for all firm characteristics in the selection and gravity equations. The destination characteristics are more or less unchanged independent of the lag structure, and the same is true when it comes to the importance of correcting for the selection process (see lambda in Table A3). The elasticities

[18] We explored the idea of omitted firm specific characteristics further by including a full set of firm-destination fixed effects (within estimation) without correcting for sample selection. In such a regression we found that only the productivity variable remains significant. Hence the relationship for the size of the firm appears to be driven primarily by variation between firms.

of firm characteristics in the intensive margin are just slightly smaller compared to those in regression 2 in Table 3, and the most significant change is that the significance of TFP falls with the number of time lags. Using three-time lags results in an elasticity of TFP that is significant at the 11 per cent level. In short we conclude that our results are robust.

VI Conclusions

This paper focuses on a key prediction from the heterogeneous firm model of international trade firm that the intensive and extensive margins of trade are affected by both the characteristics of the firm, and that of the foreign markets they serve such as their distance from the home country and their market size. This paper exploits detailed firm level data that includes information on the destination of exports for Sweden and finds strong empirical evidence that trade flows are affected in a manner consistent with this theory. Firms that are more productive and larger are more likely to serve markets that are large and relatively close and they export greater volumes to those markets.

We also show that the use of firm-destination data provides a more accurate assessment of the relationship between trade flows and firm and destination characteristics. Existing empirical methodologies bias these relationships because they conflate adjustment at the intensive and the extensive margins of trade and because they do not fully account for the interaction between firm and destination characteristics. The extensive margin includes which firms export and to which markets. Much of the existing literature, because it has information on firm characteristics but not destination, or because it has access to information on destinations of trade but not firms, cannot fully account for this adjustment. We also show however, that a further bias is explained by the interaction between firm and country characteristics. For firms that are far above the threshold entry point in the heterogeneous firm model changes in market conditions are more important than for firms that are closer to the entry threshold.

References

Aw, B.Y., Chung, S., and Roberts, M.J. (2003). 'Productivity, Output, and Failure: A Comparison of Taiwanese and Korean Manufactures', *Economic Journal*, Vol. 113, pp. 485-510.

Anderson, J.E. and Van Wincoop, E. (2003). 'Gravity with Gravitas: A Solution to the Border Puzzle', *American Economic Review*, vol. 93, pp. 170-192.

Anderson, J.E., and E. van Wincoop (2004) 'Trade costs,' *Journal of Economic Literature*, Vol. 42, 691-751

Bernard, A. and Jensen, J.B. (2004). 'Why some firms export', *Review of Economics and Statistics*, vol. 86, pp.561-569.

Bernard, A. and Wagner, J. (1998). 'Export entry and exit by German firms' *Review of World Economics/Weltwirtschaftliches Archiv*, 2001, vol. 137, pp. 134-157.

Bernard, A. B., Eaton, J., Jensen, B. and Kortum, S. (2003), 'Plants and Productivity in International Trade', *American Economic Review*, vol. 93, pp. 1268-1290.

Bernard, A.B., Jensen, B., Redding, S. and Schott, P.K. (2007). Firms in International Trade" *Journal of Economic Perspectives,* Summer.

Bernard, A.B., Jensen, B. and Schott, P.K. (2006). 'Survivial of the Best Fit: Exposure to Low-Wage Countries and the (Uneven) Growth of U.S. Manufacturing Plants', *Journal of International Economics*, vol. 68, pp. 219-237.

Bernard, A.B., Redding, S.J., and Schott, P.K. (2007). 'Comparative Advantage and Heterogeneous Firms', *Review of Economic Studies*, Vol. 74, pp. 31-66.

Chaney, T. (2008). 'Distorted Gravity: Heterogeneous Firms, Market Structure and the Geography of International Trade', *American Economic Review* vol. 98, pp. 1707.

Disdier, A-C., and Head, K. (2008) 'The puzzling persistence of the distance effect on bilateral trade', *Review of Economics and Statistics*, Vol. 90, pp. 37-48.

Eaton, J., Kamarz, F. and Kortum, S. (2004). 'Dissecting Trade: Firms, Industries, and Export Destinations', *American Economic Review*, vol. 94, pp. 150-154.

Ekholm, K., Forslid, R. and Markusen J. R. (2003). 'Export-platform foreign direct investment', NBER Working Paper No. 9517.

Görg, H. Kneller, R. and Murakozy, B. (2008). 'What Makes a Successful Exporter?' GEP Working Paper 2008/01.

Greenaway, D., and Kneller, R. (2007). 'Firm Heterogeneity, Exporting, and Foreign Direct Investment', *Economic Journal*, vol. 117, pp. 134-161.

Greenaway, D., and Kneller, R. (2008). 'Exporting, Productivity and Agglomeration', *European Economic Review.* Vol. 52(5), pages 919-939

Gullstrand, J. (2008). 'Destination and firm-specific export costs', SLI-Working Paper 2008:1.

Haveman, J. and Hummels, D. (2004). 'Alternative hypotheses and the volume of trade: the gravity equation and the extent of specialisation', *Canadian Journal of Economics*, Vol. 37, pp. 199-218.

Heckman, J.J. (1979). 'Sample Selection Bias as a Specification Error', *Econometrica*, Vol. 47, pp. 153-162.

Helpman, E. and Krugman, P.R. (1985). *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy*, Cambridge: MIT Press.

Helpman, E., Melitz, M. and Rubenstein, Y. (2008). 'Estimating Trade Flows: Trading Partners and Trading Volumes', *Quarterly Journal of Economics, Vol. CXXIII(2), pp. 441-487.*

Helpman, E., Melitz, M. and Yeaple, S.R. (2004). 'Export versus FDI with Heterogeneous Firms', *American Economic Review*, Vol. 94, pp. 300-316.
Krugman, P.R. (1980). 'Scale Economies, Product Differentiation, and the Pattern of Trade', *American Economic Review*, vol. 70, pp. 950.
Lawless, M. (2009). 'Firm export dynamics and the geography of trade', *Journal of International Economics*, Vol. 77, pp. 245-254.
Levinsohn, J., and Petrin, A. (2003). 'Estimating Production Function Using Inputs to Control for Unobservables', *Review of Economic Studies*, Vol. 70, pp. 317-341.
Melitz, M. (2003). 'The impact of trade on intra-industry reallocations and aggregate industry productivity', *Econometrica*, Vol. 71, pp. 1695-1725.
Roberts, M. J. and Tybout J.R. (1997). 'The Decision to Export in Colombia: An Empirical Model of Entry and Sunk Costs', *The American Economic Review*, vol. 80, pp. 545-564.
Tinbergen, J. (1962). *Shaping the World Economy*, New York: Twentieth Century Fund.
Tenreyro, S. (2007). 'On the impact of nominal exchange rate volatility', *Journal of Development Economics*, Vol. 82, pp.485-508.
Wooldridge, J.M. (1995). 'Selection corrections for panel data models under conditional mean independence assumptions', *Journal of Econometrics*, Vol. 68, pp. 115-132.
Wooldridge, J.M. (2002). *Econometric analysis of cross section and panel data*, Cambridge, Mass.: MIT Press.
Yeaple, S. R. (2003). 'The complex integration strategies of multinationals and cross country dependencies in the structure of foreign direct investment', *Journal of International Economics*, Vol. 60, pp. 293-314.
Yeaple, S. R. (2006). 'Offshoring, Foreign Direct Investment, and the Structure of U.S. Trade', *Journal of the European Economic Association*, vol. 4, pp. 602-611.

Tables

Table 1: *Definitions and sources*

Characteristics	*Definition*
Firm level	
Productivity	Total factor productivity, multilateral index defined, as in Aw et al (2003), $\ln TFP_t^i = (q_t^i - \overline{q}_t) + \sum_{s=2}^{t}(\overline{q}_s - \overline{q}_{s-1}) - \{\sum_j \frac{1}{2}(\alpha_{tj}^i - \overline{\alpha}_{tj})(x_{tj}^i - \overline{x}_{tj}) + \sum_{s=2}^{t}\sum_j \frac{1}{2}(\overline{\alpha}_{sj} - \overline{\alpha}_{s-1j})(\overline{x}_{sj} - \overline{x}_{s-1j})\}$, where lower-case letters indicate natural logarithm of output (*q*) and inputs (*x*), bars indicate unweighted average over all firms (*i*) and hence the hypothetical firm used as a reference, and α is input-cost shares. Inputs (*j*) used are number of employees, capital stock and raw materials. An alternative productivity measure based on the method recommended by Levinsohn and Petrin (2003) does not change our results except for a larger coefficient on TFP in the gravity equation.
Capital stock	The capital stock is calculated by the perpetual method using book value the first year. Depreciation rate for equipment and for buildings are 0.1 and 0.05 respectively.
Employees	Number of employees in full-year-equivalents.
Physical capital intensity	Capital stock (see above) per employee.
Human capital intensity	The share of employees with university degree.
Foreign owned	One when more than 50 per cent of the firm is owned by a foreign firm, zero otherwise
Own foreign affiliations	One when the firm own affiliations in another country, zero otherwise
Market level	
Size	Gross domestic product (constant prices) from the World Development Indicators (World Bank)
Population	From the World Development Indicators (World Bank)
Distance	Kilometres, calculated with the great-circle distance formula based on longitudes and latitudes from the CSI's World Fact. The distance is calculated from Stockholm to the capital of the export destination.
Real exchange rate	Annual average from the Bank of Sweden (SEK/currency of the export destination) times the ratio of CPI (consumer price index) of the destination to the CPI of Sweden (from UNdata, see http://data.un.org). Currencies not available are replaced by the exchange rate of the USA.
Exchange rate risks	Exchange rate volatility is measured as the ratio of the difference between the highest and the lowest rate to the average rate.

Table 2: ***Descriptive figures (all firm-time-export destination observations)***

Variables	Mean of non-exporters # 1 314 791	Mean of exporters # 9 858
Destination characteristics		
Distance	6441 (4088)	2789 (3678)***
GDP (constant $US)	1.95e+11 (8.59e+8)	7.73e+11 (1.84e+12))***
Population (million)	38 (1.34e+8)	47 (1.32e+8)***
EU15 dummy	0.09 (0.29)	0.41 (0.49)***
English as first language	0.19 (0.39)	0.13 (0.34)***
Firm characteristics		
TFP	1.17 (0.54)	1.74 (0.53)***
Labour productivity [a]	365 (379)	694 (642)***
Sales (1,000 SEK)	73 649 (524419)	1 072 261 (1 887931)***
Share of high skill	0.03 (0.08)	0.06 (0.05)***
Foreign owned	0.03 (0.16)	0.30 (0.46)***
Own foreign	0.06 (0.23)	0.43 (0.49)***
Age [b]	3.34 (1.66)	3.45 (1.68)***
Employees	33 (183)	436 (775)***

Note: See Table 1 for variable definitions. Figures in parentheses are standard deviations. *** indicates that the mean is significantly different from non-exporters at the 1 per cent level of significance. [a] Ratio of firm value added to number of employees. [b] Number of years in sample.

Figure 1: ***Share of total number of exporting firms vs share of total exports, 2002*** **(World Bank country codes)**

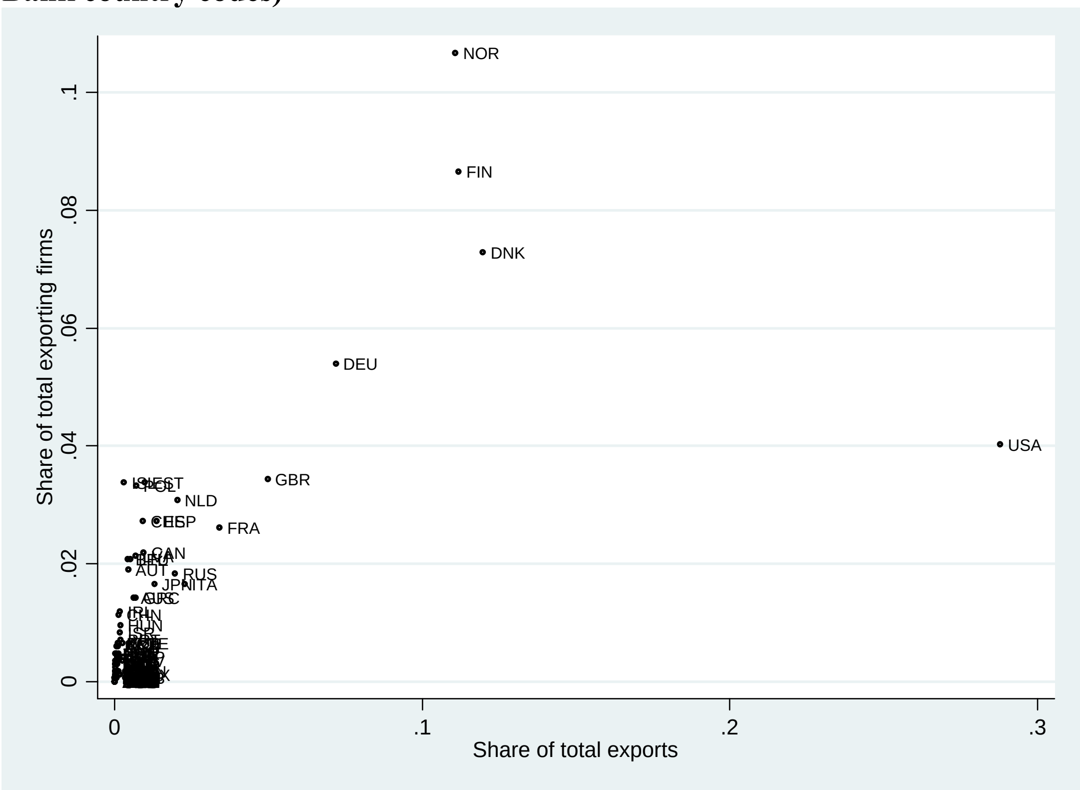


Figure 2: ***Number of firms in a market vs the average number of export destinations of the firms in that market, 2002*** **(World Bank country codes)**

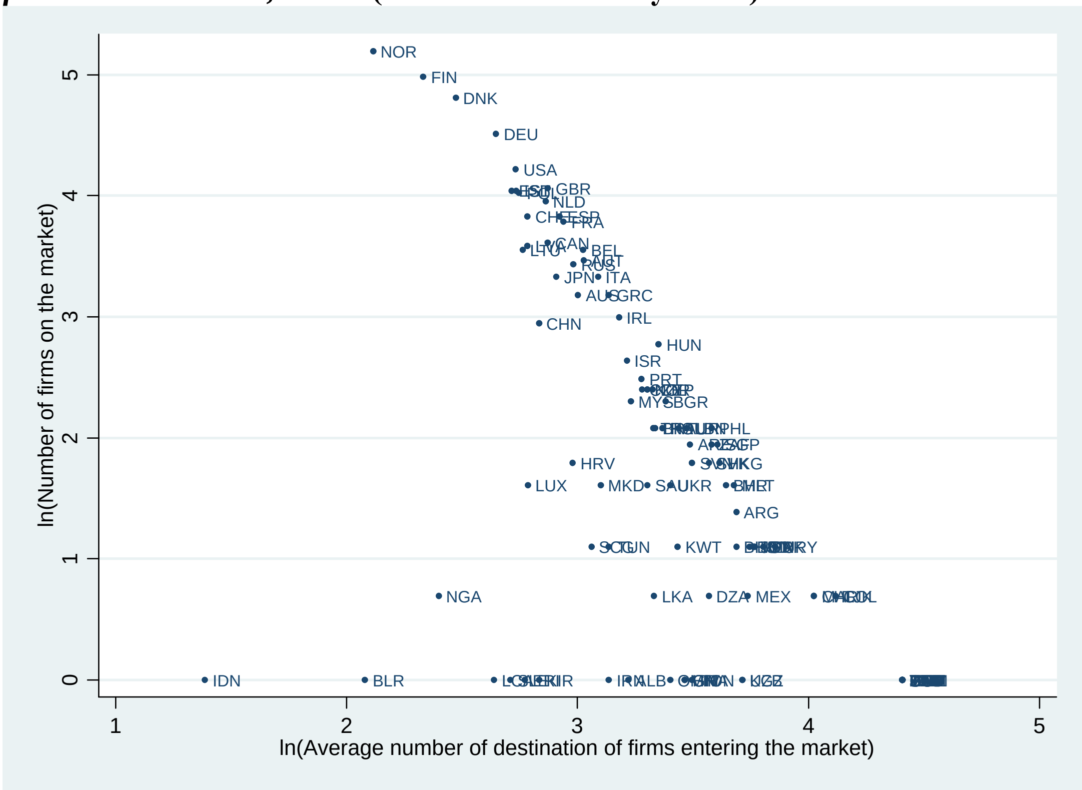


Figure 3: ***Total Factor productivity and number of export destinations at firm level***

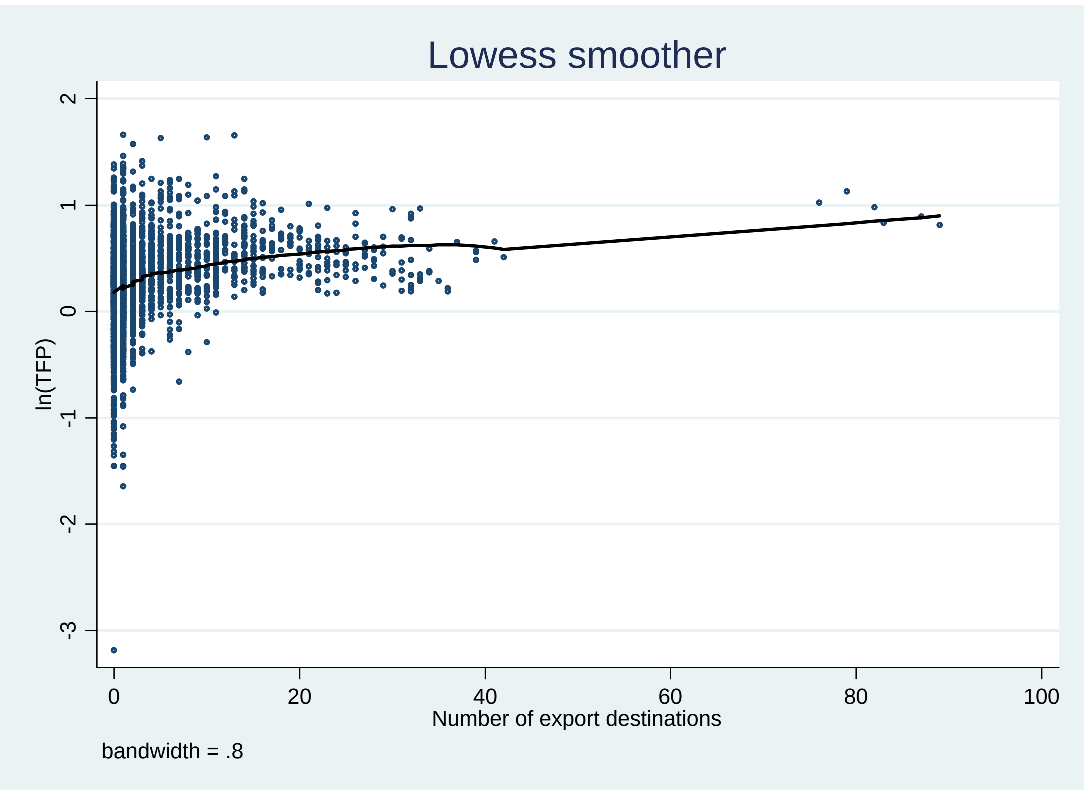


Table 3: ***Firm-level regression 1998-2002, bilateral trade flows*** [a]

Margin	Extensive	Intensive		
	Selection (probit) equation for export	Corrected regression due to selection	Excluding zeros (not corrected)	Corrected regression due to selection with interactions
Regression No.	*1*	*2*	*3*	*4*
Destination characteristics				
Ln(distance)	**-0.25** (*.00*)	**-0.27** (*.03*) {-0.51}	**-0.44** (*.00*)	0.33 (*.07*)
Ln(GDP)	**0.18** (*.00*)	**0.17** (*.00*) {0.34}	**0.27** (*.00*)	0.05 (*.50*)
Ln(population)	**-0.06** (*.03*)			
EU15 dummy	**-0.21** (*.00*)	**0.97** (*.00*) {0.76}	**0.86** (*.00*)	**0.93** (*.00*)
Low and middle-income dummy	**-0.17** (*.00*)	**0.25** (*.05*) {0.08}	-0.07 (*.62*)	0.21 (*.16*)
English speaking market	**-0.29** (*.00*)	**-0.65** (*.00*) {-0.92}	**-0.93** (*.00*)	**-0.72** (*.00*)
Ln(real exchange rate)	-0.003 (*.70*)	-0.01 (*.61*) {-0.01}	-0.02 (*.90*)	-0.01 (*.90*)
Exchange rate risk	**-0.05** (*.05*)			
Firm characteristics (all lagged one period)				
Lagged Export status	**2.51** (*.00*)			
Ln(TFP)	**0.31** (*.00*)	**0.38** (*.02*) {0.68}	**0.73** (*.00*)	**-5.71** (*.04*)
Ln(Employees)	**0.23** (*.00*)	**0.12** (*.00*) {0.33}	**0.28** (*.00*)	**1.04** (*.00*)
Foreign MNE	**0.14** (*.00*)	0.07 (*.57*) {0.21}	0.24 (*.09*)	0.67 (*.63*)
Swedish MNE	**0.11** (*.00*)	0.16 (*.11*) {0.51}	**0.29** (*.00*)	0.15 (*.16*)
Capital per labour	**0.30** (*.00*)			
Share of high skilled workers	**0.87** (*.00*)			
Interaction terms				
Ln(TFP)_Ln(GDP)				**0.23** (*.03*)
Ln(Employess)_Ln(distance)				**-0.12** (*.00*)
Constant	**-7.11** (*.00*)	3.19 (*.07*)	-0.15 (*.93*)	3.19 (*.07*)
Rho (correlation between 1st and 2nd step errors from MLE)		**-0.42** (*.02*)		**-0.42** (*.02*)
Lambda		**-0.99** (*.00*) [b]		**-0.99** (*.00*) [b]
Time dummies	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes
Industry dummies (3-digit)	Yes	Yes	Yes	Yes
Likelihood		-31 738		-31 667
R^2 (adjusted)			0.19	
Number obs.		1 087 666	8241	1 087 666

Notes: [a] See Table 1 for variable definitions. All figures within parentheses are p-values based on standard errors clustered around export destinations. The selection model in regression 1-2 is based on maximum-likelihood estimation, and figures within curly brackets are the marginal effects conditional on selected observations defined as $\partial E(Y \mid S^* > 0, X)/\partial X_k = \beta_k - \gamma_k \rho \sigma_\varepsilon \delta(-\mathbf{w}\boldsymbol{\gamma})$, where Y is export flows, S^* export selection variable, $\delta(.)$ a function of Mill's ratio, X and w are independents in the gravity and the selection equation respectively. [b] This is the coefficient of the Mill's ratio from a two-step Heckman regression.

Table 4: ***Country-level analysis correcting for zero-trade flows*** [a]

	Maximum likelihood specification		Excluding zeros		HMR-specification
	Selection equation	Gravity equation	OLS	Within	Gravity equation
Regression No.	*5*	*6*	*7*	*8*	*9*
Ln(distance)	-0.01 (*.96*)	**-0.74** (*.00*) {-0.74}	**-0.74** (*.00*)		**-0.57** (*.00*)
Ln(GDP)	**0.73** (*.00*)	**0.68** (*.00*) {0.67}	**0.66** (*.00*)	0.10 (*.76*)	0.06 (*.83*)
Ln(population)	**-0.33** (*.00*)				
EU15 dummy	**4.52** (*.00*)	-0.49 (*.32*) {-0.51}	-0.45 (*.37*)		
Low & middle-income dum	-0.03 (*.70*)	**-1.60** (*.00*) {-1.60}	**-1.60** (*.00*)		
Real exchange rate	-0.31 (*.85*)	-0.01 (*.88*) {-0.01}	-0.01 (*.88*)	-0.12 (*.38*)	0.02 (*.85*)
Exchange rate risk	-0.11 (*.54*)				
Constant	-10.6 (*.25*)	-1.25 (*.69*)	-0.77 (*.80*)		16.39 (*.07*)
Rho (correlation between 1st and 2nd step errors from MLE)		0.09 (*.34*)			
Correcting for zeros with Mill's ratio (lambda)		0.58 (*.16*) [b]			-12.87 (*.09*)
Correcting for share of exporting firms [c]					-5.31 (*.24*)
Correcting for share of exporting firms, squared [c]					**1.97** (*.02*)
Correcting for share of exporting firms, cubic [c]					**-0.19** (*.00*)
Time dummies	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	No	No
Fixed destination effects	No	No	No	Yes	No
Likelihood / R^2 (adjusted)		-1 491	0.73		-1 242
Number obs.		893	626	626	716

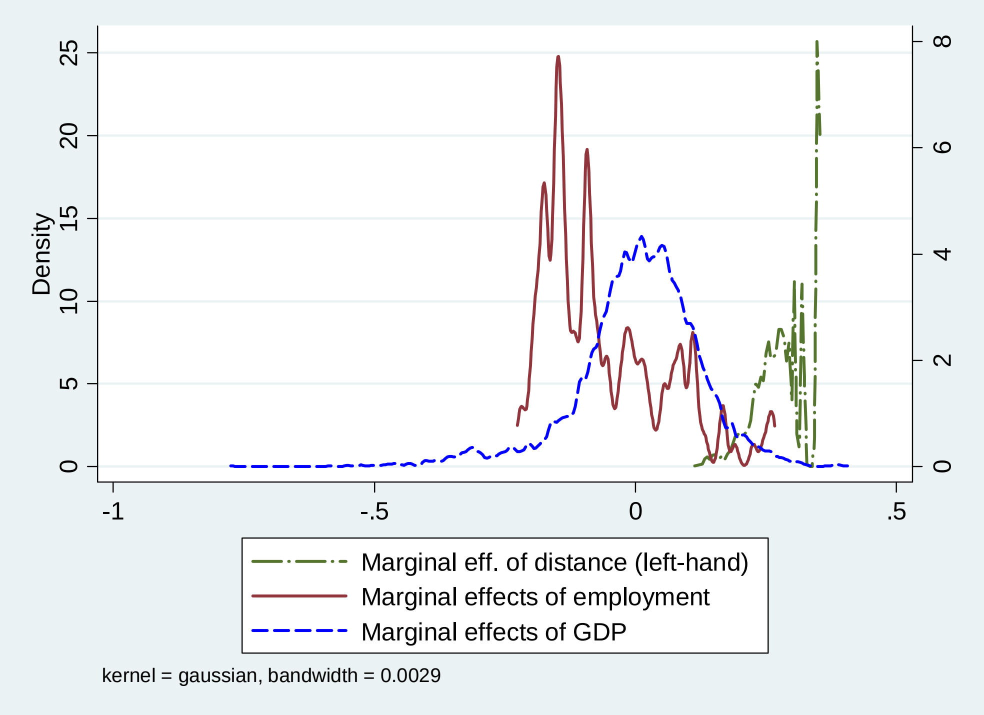
Note: [a] See Table 1 for variable definitions. All figures within parentheses are p-values based on standard errors clustered around export destination. The selection model in regression 1 is based on maximum-likelihood estimation, and figures within curly brackets are the marginal effects conditional on selected observations defined as $\partial E(Y \mid S^{*} > 0, X) / \partial X_k = \beta_k - \gamma_k \rho \sigma_\varepsilon \delta(-\mathbf{w\gamma}),$ where Y is export flows, S^* export selection variable, $\delta(.)$ a function of Mill's ratio, X and w are independents in the gravity and the selection equation respectively. [b] This is the coefficient of the Mill's ratio from a two-step Heckman regression. [c] We use the alternative estimation technique in Helpman et al (2008) in order to control for firm heterogeneity. Hence we incorporate not only the Mill's ratio but also a polynomial (cubic) of $\hat{z} = \Phi^{-1}(\hat{p}) + \hat{n},$ where p is the predicted probability of a bilateral export flow and n the Mill's ratio.

Table 5: ***Firm- level regression 1998-2002, multilateral trade flows*** [a]

Margin	Extensive	Intensive		
	Selection (probit) equation for export	Corrected regression due to selection	Excluding zeros	Within estimation excluding zeros
Regression No.	10	11	12	13
Lagged Export status	**2.35** (*.00*)			
Ln(TFP)	**0.18** (*.00*)	**0.90** (*.00*) {1.17}	**1.63** (*.00*)	0.25 (*.26*)
Ln(Employees)	**0.25** (*.00*)	**0.37** (*.00*) {0.74}	**0.60** (*.00*)	0.24 (*.17*)
Foreign MNE	0.32 (*.11*)	**1.23** (*.00*) {1.74}	**1.48** (*.00*)	
Swedish MNE	**-0.19** (*.03*)	**1.08** (*.00*) {0.79}	**0.14** (*.00*)	
Capital per labour	**0.27** (*.00*)			
Share of high skilled workers	**0.71** (*.00*)			
Constant	**-3.32** (*.00*)	**6.17** (*.00*)	**3.85** (*.00*)	
Rho (correlation between 1st and 2nd step errors from MLE)		**-0.73** (*.00*)		
Lambda		**-1.77** (*.00*) [b]		
Time dummies	Yes	Yes	Yes	No
Industry dummies (3-digit)	Yes	Yes	Yes	No
Fixed firm effects	No	No	No	Yes
Likelihood		-4 142		
R^2 (adjusted)			0.41	
R^2 (within)				0.05
Number obs.		7 206	1 334	1 334

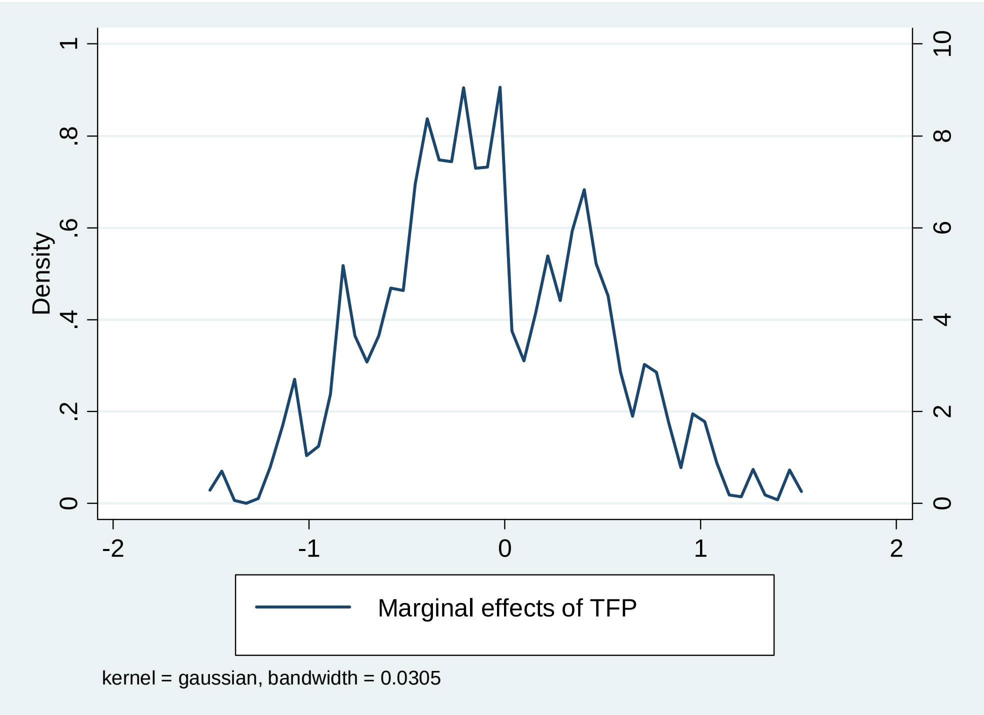
Notes: [a] See Table 1 for variable definitions. All figures within parentheses are p-values based on standard errors clustered around firms. The selection model in regression 1-2 is based on maximum-likelihood estimation, and figures within curly brackets are the marginal effects conditional on selected observations defined as $\partial E(Y \mid S^* > 0, X) / \partial X_k = \beta_k - \gamma_k \rho \sigma_\varepsilon \delta(-\mathbf{w}\boldsymbol{\gamma}),$ where Y is export flows, S^* export selection variable, $\delta(.)$ a function of Mill's ratio, X and w are independents in the gravity and the selection equation respectively. [b] This is the coefficient of the Mill's ratio from a two-step Heckman regression.

Figure 4: ***Distribution of marginal effects of employment, GDP and distance on firm exports***



Note: These marginal effects do not incorporate indirect effects from the selection equation.

Figure 5: ***Distribution of the marginal effects of TFP on firm exports***



Note: These marginal effects do not incorporate indirect effects from the selection equation.

Appendix

Tables

Table A1: *Gravity equation 2001, different levels* [a]

Intensive margin	Firm-level bilateral flows 1	Firm-level multilateral flows 2	Country-level flows 3	Country-level flows (HMR) 4
Destination characteristics				
Ln(distance)	**-0.38** (*.00*) {-0.68}		**-0.81** (*.00*) {-0.81}	**-0.47** (*.59*)
Ln(GDP)	**0.14** (*.00*) {0.34}		**0.66** (*.00*) {0.65}	**0.18** (*.78*)
EU15 dummy	**0.93** (*.00*) {0.65}		-0.90 (*.10*) {-0.85}	n.a.
Low & middle-income dum	0.12 (*.46*) {-0.05}		**-2.12** (*.00*) {-2.12}	n.a.
English speaking market	-0.45 (*.12*) {-0.38}		n.a.	n.a.
Ln(real exchange rate)	-0.02 (*.59*) {-0.02}		-0.03 (*.64*) {-0.03}	-0.11 (*.48*)
Firm characteristics (all lagged one period)				
Ln(TFP)	**0.41** (*.04*) {0.63}	**0.93** (*.03*) {1.02}		
Ln(Employees)	**0.18** (*.00*) {0.37}	**0.44** (*.00*) {0.73}		
Foreign owned	-0.34 (*.03*) {-0.28}	0.66 (*.11*) {1.63}		
Swedish multinational	0.03 (*.85*) {0.21}	**1.12** (*.00*) {0.89}		
Constant	**5.02** (*.01*)	**4.24** (*.03*)	1.49 (*.67*)	
Rho (correlation between 1st and 2nd step errors from MLE)	**-0.42** (*.00*)	**-0.66** (*.00*)	0.09 (*.71*)	
Correcting for zeros with Mill's ratio	**-0.89** (*.00*) [b]	**-1.37** (*.00*) [b]	0.36 (*.72*) [b]	-14.82 (*.46*)
Correcting for share of exporting firms [c]				-6.06 (*.62*)
Correcting for share of exporting firms, squared [c]				2.41 (*.25*)
Correcting for share of exporting firms, cubic [c]				**-0.36** (*.02*)
Regional dummies	Yes	No	Yes	
Industry	Yes	Yes	No	
Likelihood	-6 467	-861	-232	
Nobs.	222 486	1 474	151	

Note: [a] See Table 1 for variable definitions. All figures within parentheses are p-values based on standard errors clustered around export destination. The selection model in regression 1-3 is based on maximum-likelihood estimation, and figures within curly brackets are the marginal effects conditional on selected observations defined as $\partial E(Y \mid S^* > 0, X) / \partial X_k = \beta_k - \gamma_k \rho \sigma_\varepsilon \delta(-\mathbf{w\gamma})$, where Y is export flows, S^* export selection variable, $\delta(.)$ a function of Mill's ratio, X and w are independents in the gravity and the selection equation respectively. [b] This is the coefficient of the Mill's ratio from a two-step Heckman regression. [c] We use the alternative estimation technique in Helpman et al (2008) in order to control for firm heterogeneity. Hence we incorporate not only the Mill's ratio but also a polynomial (cubic) of $\hat{z} = \Phi^{-1}(\hat{p}) + \hat{n}$, where p is the predicted probability of a bilateral export flow and n the Mill's ratio.

Table A2: *Firm-level regression 1998-2002, bilateral trade flows (Wooldridge 1995-approach)* [a]

Margin	Intensive	
	Without time-invariant variables [b]	With time-invariant variables
	1	2
Ln(distance)		**-0.37** (*.00*) [*.00*]
Ln(GDP)	**0.96** (*.00*) [*.00*]	**0.92** (*.00*) [*.00*]
EU15 dummy		**0.66** (*.00*) [*.00*]
English speaking market		0.08 (*.66*) [*.35*]
Ln(real exchange rate)	**0.46** (*.00*) [*.00*]	**0.50** (*.00*) [*.00*]
Ln(TFP)	0.25 (*.06*) [*.22*]	0.18 (*.16*) [*.32*]
Ln(Employees)	0.17 (*.07*) [*.14*]	**0.15** (*.00*) [*.18*]
Foreign owned		-0.19 (*.21*) [*.26*]
Swedish multinational		**0.27** (*.00*) [*.01*]
Mill's ratio 1998	**-1.72** (*.00*) [*.00*]	**-1.52** (*.00*) [*.00*]
Mill's ratio 1999	**-1.78** (*.00*) [*.00*]	**-1.56** (*.00*) [*.00*]
Mill's ratio 2000	**-1.69** (*.00*) [*.00*]	**-1.50** (*.00*) [*.00*]
Mill's ratio 2001	**-1.13** (*.00*) [*.00*]	**-0.97** (*.00*) [*.00*]
Mill's ratio 2002	**-1.08** (*.02*) [*.00*]	**-0.94** (*.00*) [*.00*]
Fixed effects are assumed to be a linear projection of the average of all explanatory variables.		
Constant	**3.51** (*.00*) [*.00*]	**5.03** (*.00*)
R^2 (adjusted)	0.22	0.27
Nobs.	8 358	8 358

Notes: [a] See Table 1 for variable definitions. P-values based on robust standard errors, clustered around export destination (not corrected for sample selection), in parentheses as well as bootstrapped standard errors in brackets (500 replications). The selection equation is excluded in order to save space.. [b] This approach is based on Wooldridge (1995 and 2002) and uses a Mundlak specification of the fixed effects, which implies that fixed effects are assumed to be a linear projection of the average of all variables included in the model (both in the selection equation and in the gravity equation). The selection process is a standard probit on yearly basis including capital intensity (both human and physical), population, lagged export dummy, and interaction terms between TFP, distance and GDP.

Table A3: ***Gravity equation at firm-level regression 1998-2002, bilateral trade flows, with different lags*** [a]

Margin	Intensive	
	2-years lag	3-year lag
	1	2
Ln(distance)	**-0.29** (*.02*)	**-0.25** (*.05*)
Ln(GDP)	**0.13** (*.00*)	**0.14** (*.00*)
EU15 dummy	**0.97** (*.00*)	**0.98** (*.00*)
Low and middle-income dummy	0.18 (*.19*)	0.25 (*.06*)
English speaking market	**-0.54** (*.02*)	**-0.59** (*.03*)
Ln(real exchange rate)	-0.004 (*.83*)	-0.004 (*.85*)
Ln(TFP)	**0.31** (*.05*)	0.29 (*.11*)
Ln(Employees)	**0.10** (*.00*)	**0.09** (*.00*)
Foreign owned	0.08 (*.59*)	0.09 (*.59*)
Swedish multinational	0.08 (*.47*)	0.19 (*.86*)
Constant	**2.28** (*.28*)	**4.24** (*.03*)
Rho (correlation between 1st and 2nd step errors from MLE)	**-0.39** (*.00*)	**-0.41** (*.00*)
Correcting for zeros with Mill's ratio	**-0.85** (*.00*) [b]	**-0.93** (*.00*) [b]
Regional dummies (see Appendix)	Yes	Yes
Industry dummies (3-digit level)	Yes	Yes
Likelihood	-27 142	-20 825
R^2 (adjusted)		
R^2 (within)		
Nobs.	850 684	613 702

Notes: [a] See Table 1 for variable definitions. All figures within parentheses are p-values based on standard errors clustered around export destinations. The selection model is based on maximum-likelihood estimation. [b] This is the coefficient of the Mill's ratio from a two-step Heckman regression.

Table A4: ***Regions***

South America (reg 1)	Northern Africa (reg 6)	Eastern Asia (reg 11)	Western Asia (reg 16)
Oceania (reg 2)	Middle Africa (reg 7)	South-Eastern Asia (reg 12)	Eastern Europe (reg 17)
Western Africa (reg 3)	Southern Africa (reg 8)	Sothern Europe (reg 13)	Northern Europe (reg 18)
Central America (reg 4)	Northern America (reg 9)	Southern Asia (reg 14)	Western Europe (reg 19)
Eastern Africa (reg 5)	Caribbean (reg 10)	Central Asia (reg 15)	

Note: These regions are based on UN's regional coding. Number of countries in each region in paranthesis

Kort om AgriFood Economics Centre

AgriFood Economics Centre utför kvalificerade samhällsekonomiska analyser inom livsmedels-, jordbruks- och fiskeområdet samt landsbygdsutveckling. Verksamheten är ett samarbete mellan Sveriges Lantbruksuniversitet och Lunds Universitet och syftar till att ge regering och riksdag vetenskapligt underbyggda underlag för strategiska och långsiktiga beslut.

Alla publikationer kan beställas kostnadsfritt via http://www.agrifood.se/

AgriFood Economics Centre
Scheelevägen 15 D
Box 730, 220 07 Lund
www.agrifood.se
e-post: info@agrifood.lu.se

AgriFood
economics centre