Robustness of the Extensive Margin in the Helpman, Melitz and Rubinstein (HMR) Model

Maxim, Belenkiy

UC Santa Cruz

July 2008

Robustness of the Extensive Margin in the Helpman, Melitz and Rubinstein (HMR) Model*

Maxim Belenkiy†

July, 2008
(Revised February, 2009)

Abstract

The HMR model extends the classical gravity model of trade to correct for the large number of zeros in the world trade matrix (export selection) and for the unobservable fraction of exporting firms (extensive margin). They find that, while omission of both of these corrections result in the biased estimates of the gravity model, the extensive margin correction is the most significant of the two when estimating the trade flows. I test the robustness of this conclusion by splitting the world trade data into OECD and non-OECD countries. The extensive margin should be both economically and statistically more significant for the OECD exporters, while export selection should play a larger in the trade flows for the non-OECD exporters. I find that the extensive margin is not significant for the OECD trade flows, but the export selection is important regardless of the exporter location. These findings call into question the conclusions of the HMR model. I posit and test possible hypothesis to explain them.

Keywords: trade flows, trade frictions, asymmetries, gravity model, estimation
JEL classification: F10, F12, F14

*I thank my advisor Phil McCalman for the invaluable guidance throughout the development of this paper; Marc Melitz and Yona Rubinstein for providing the data and suggestions. All errors and omissions remain mine.

†Maxim Belenkiy, Department of Economics, University of California Santa Cruz, Santa Cruz CA 95064, mbelenki@ucsc.edu

1 Introduction

The gravity model of trade is a workhorse model for the empirical estimation of the international trade flows. In its usual form, the gravity model predicts the volume of trade between two countries based on their economic sizes (often using GDP measurements). It has been also recognized that the measure of the economic size is proportional to the measures of "trade resistance" between the two countries[1]. Among others, these measures include: the geographic distance; a dummy for the common border and language and a dummy for the membership in a trade agreement.

With the development of the firm-level heterogeneity theory pioneered by Melitz (2003) the extended model by Helpman, Melitz and Rubenstein (HMR) (2008) allows to reconsider the statistical and economic significance of estimates in the gravity model. Since the Melitz (2003) model is capable of endogenously calculating the number of exporting firms in the market it becomes possible to decompose the trade flows into the intensive margin (the volume of trade per firm) and the extensive margin (the number of the exporting firms). Given the importance of the extensive margin on the theoretical grounds, the failure to control for it in the classical gravity model calls for questioning of its consistent estimation.

The underlying theory that is used to derive the classical gravity model treats each firm equally as productive, so that each firm can become an exporter. Recognizing that this outcome is strongly rejected by the data (50 percent of country pairs do not trade with each other) the HMR (2008) model links determinants of the trade flows between countries with the firm-level heterogeneity.Using Melitz (2003) framework, the HMR model bridges firm-level heterogeneity with country-level data by aggregating exports over varying distributions of firms that are productive enough to become exporters. Thus, without any firm-level data it becomes possible to separately control for the number of exporting firms as well for the volume of trade per exporting firm corrected for the non-random export selection through the characteristics of the marginal exporters to different destinations. Incorporating these controls allows to consistently estimate the gravity model. HMR finds that while omission of both of these corrections result in the biased estimates of the gravity model, the extensive margin correction is the most significant of the two when estimating the trade flows.

In this paper, I revisit the original HMR model, with an extension to test the robustness of the extensive margin in the HMR estimation. In particular I investigate whether it is still the case that the extensive margin remains both economically and statistically significant in correcting the firm-heterogeneity bias in the classical gravity model of trade when I split the world trade data such that the extensive margin must theoretically overwhelm the firm export selection. My methodology is similar to Hummels and Levinsohn (1995) in that I take the theoretical set up of the HMR model as given, while amending the empirical specification. Keeping the theoretical set-up unaltered, allows me to analyze the importance of the extensive margin at the firm level with the use of the country-level data. I depart from the symmetric trading world in the HMR model and

[1] Tinbergen (1962) was the first to recognize this proportionality.

consider the world consisting of two regions: North and South. Countries in the North are assumed OECD countries, while South countries are developing. This configuration allows for testing the importance the extensive margin of trade in the basic HMR model in the the two important ways.

First, I extend the original HMR empirical specification, by introducing region of export origin controls through interaction effects. These region-barrier interaction effects are aimed to capture differential effects of the trade barriers on the trade volumes for the Northern and the Southern exporters, which allows for a preliminary robustness check of the significance of the extensive margin relative to the export-selection with the world trade data split. Second, I divide the cross-section sample into four groups based on trading partner location pairs: North-North, North-South, South-North and South-South. For these location pairs, I estimate the original HMR model with no interaction effects and analyze the relative importance of the extensive margin to the non-random export selection in the relation to the theoretical predictions. On the theoretical grounds the extensive margin should be both economically and statistically more significant for the OECD exporters, while export selection should play a larger in the trade flows for the non-OECD exporters.

To preview my results, I find that the HMR estimation results give too much credit for the extensive margin in explaining biases in the standard gravity model. The extensive margin continues to be significant but its magnitude falls considerably, while the magnitude of the non-random selection rises. Importantly, when the trade data is split into four regions, I find that the extensive margin is not significant for the North-South trading partners, which contradicts theoretical predictions. However, the export-selection appears to be important regardless of the exporter region. Thus, while in the aggregate the extensive margin of trade is the main source of the biases in the classical gravity model, once the trade flows are split the significance and importance of the extensive marginal disappears. One of the possible explanations for this finding is that the extensive margin largely depends on the elasticity of substitution. For the Southern countries, that primarily export homogeneous varieties the elasticity of substitution between these varieties is high, making the extensive margin an unimportant determinant of the trade flows. For the Northern countries, my findings are puzzling.

This paper is organized as follows. In section 2, I discuss the inconsistencies in estimating the classical gravity model, describe the main features of the HMR model and present the model extension. In section 3, I describe the data used in my estimations. In section 4, I present all estimation results. Section 5, then concludes. I also include two appendices. In the Appendix A, I present the detailed derivation of the HMR model upon which I build my extension. In the Appendix B, I state the definitions of all the variables used in the estimations. These appendices are followed by the tables with estimation results and figures.

2 An Extension of the HMR Model

2.1 Inconsistencies in Estimating the Classical Gravity Model

As discussed by Anderson and van Wincoop (2004) the estimating gravity equation has the following general form:

$$x_{ij} = \alpha_1 y_i + \alpha_2 y_j + \sum_{m=1}^{M} \beta_m \ln(z_{ij}^m) + \varepsilon_{ij}, \tag{1}$$

where x_{ij} - volume of bilateral trade flows from j to i expressed in the natural logarithm; y_j, y_i - GDP of exporter j and importer i respectively and z_{ij}^m is a vector of the observable trade barriers. The estimate of β_m captures the effect of the intensive margin of trade - it predicts the negative effect of trade barriers on the trade volumes once j already exports to i.

Recently the estimation strategy of the gravity model has been challenged in the empirical trade literature. This concern stems from stylized trade data analysis: over fifty percent of all the bilateral trade volumes are zero. Moreover, while the underlying assumption of the classical gravity model is a symmetry in trade volumes between the trading partners ($x_{ij} = x_{ji}$), the data strongly rejects such assumption. There exists asymmetric one-way trading relations such that $x_{ij} > 0$ and $x_{ji} = 0$ or vice versa. If there are gains from trade, than why does such regularity appear? Traditionally, this issue was ignored by the researches, either by dropping the zero observations or by only considering the bilateral trade between the developed countries. However, the non-random nature of zeros in the trade matrix raises a concern of consistency in estimating β_m.

Ignoring the zeros in the trade matrix results in the inconsistent estimation of the gravity model (1) for the two reasons. First, dropping or ignoring the zeros in the trade matrix results in the selection bias. The selection bias is associated with unobserved (or not controlled for) trade-barriers that are correlated with the observed trade barriers in z_{ij}^m and are important in explaining the volumes of bilateral trade flows. Hence the countries with large unobservable trade barriers may not select into exporting. This explains the zeros in the trade matrix, but not for the random reasons. Second, given that a country-pair selects into exporting, the trade flows may be asymmetric. This can only happen if the fraction of exporters in these countries is different or potentially zero. Failure to control for the fraction of exporters (the extensive margin) results in heterogeneity bias. It confounds the effects of trade barriers on firm-level trade with their effects on the proportion of exporting firms.

2.2 Extensive Margin and Trade Volumes

The main contribution of the HMR model is to derive the measure of the extensive trade margin from the structural theoretical model. The HMR model[2] is an application of the Melitz (2003) model with few simplifications: no domestic production and no dynamics of entry and exit. If only

[2]The detailed derivation of the HMR model is provided in the Appendix A

some fraction of the firms in country j choose to export, choose to export, this can only happen if these firms can at least break even in terms their profits. While every firm in country j facing no fixed costs choose to serve its domestic market, the firms in country j that choose to export must be productive enough (or operate at a low enough unit cost a) to cover fixed trade-barrier costs f_{ij}. This set-up results in the zero-profit condition (A7) that implicitly defines the minimum unit cost cut-off a_{ij}.

The HMR framework can be best shown graphically. Figure A highlights how the variation in fixed export costs affects the selection of the firms in a country j into exporting[3]. The key departure from the traditional Melitz (2003) model is the use of truncated distribution of the unit costs with a cdf $G(a)$ that has the support $[a_{L,}a_H]$ such that $a_H > a_L > 0$. In this case the firm's productivity is $\phi^{1-\varepsilon} \equiv 1/a$, where ε is an elasticity of substitution. The truncated distribution insures that there are going to be mass of firms that will not be productive enough to export. Emprically the choice of such distribution implies zero trade flows for some exporting firms. The fraction of firms that choose to export is determined by the level of the fixed export costs f_{ij} and the export cut-off $\phi_{ij}^{1-\varepsilon}$. When the level of the fixed costs is as high as $f_{ij}^{'}$ (in the negative sense) no firms choose to export, since none of the firms are productive enough to cover fixed costs and make at least zero profits. In this case $\phi_{ij}^{(1-\varepsilon)\prime} < \phi_H^{1-\varepsilon}$: the least productive firm that find it profitable to export to country i has a unit cost above the support of $G(a)$. However, when the level of these costs is f_{ij}, some fraction of the firms in a country j will export. This fraction is implicitly determined through the bilateral trade volumes (A8) under the Melitz (2003) assumption that every firm produces exactly one variety l and it is shown by the shaded region in the Figure A. The expression for export volumes (A8) is the expected value of the fraction of all the firms that export from country j in the interval $[a_L, a_{ij}]$, where the unit cost a is drawn from a distribution with the CDF $G(a)$. With the assumed symmetric distribution, in the aggregate, the average firm in every country pair faces same probability of being selected into export market. However, it can also be the case that no firms will be productive enough to export from country j, but some firms will be able to export from country i resulting in one-way trade flows. Thus, the HMR framework can successfully capture the empirical regularities of the world trade data and provide a theoretical justification of the empirical importance of the extensive margin in trade.

The derivation of the extensive trade margin measure W_{ij}- fraction (possibly zero) of exporting firms requires an assumption on the functional form for CDF of $G(a)$. In the parametric form HMR selects the truncated Pareto distribution (A11) as a functional form of $G(a)$. While, they show that choice of this distribution is not specific to the results, the measure that controls for the fraction of exporters is based on the choice of this distribution and plays an important role in the estimating results. Using this distribution the export volumes can be written as in equation (A12). The measure of the fraction of the exporting firms is given by the expression for W_{ij}. It is the ratio of the bounded from the above productivity $G(a_{ij})$ that gives non-zero fraction of exporters. This

[3] Figure A shows the selection of firms into exporting in the country j. The export selection of the firms in the country i (not shown) is constructed similarly.

measure is derived from the solution to the firm's problem. Crucially it depends on the elasticity of substitution ε.

The estimation of W_{ij} amounts to the two-stage estimation procedure. Assuming normal distribution of the error term in the log-linear gravity model (A14), HMR estimate a Probit specification (A18). The residuals from this estimation are the predicted probabilities of the firm-level export selection. Using these probabilities HMR backs out the fraction of exporters by calculating the inverse CDF of the assumed normal distribution $\Phi(\bullet)$. The consistent estimate for W_{ij} (A19) depends on $\delta \equiv \sigma_\eta(k-\varepsilon+1)/(\varepsilon-1)$ where δ needs to be estimated. The final consistent estimating equation (A20) that controls for heterogeneity bias (through fraction of firms that choose to export) and selection bias (calculated using Mills ratio) is:

$$m_{ij} = \beta_0 + \lambda_j + \chi_i - \underbrace{\gamma d_{ij}}_{\text{Intensive Margin}} + \underbrace{\ln\{\exp[\delta(\widehat{z}^*_{ij} + \widehat{\overline{\eta}}^*_{ij})] - 1\}}_{\text{Extensive Margin}} + \underbrace{\beta_{u\eta}\widehat{\overline{\eta}}^*_{ij}}_{\text{Non-Random Selection}} + e_{ij}, \qquad (2)$$

where γ the elasticity of the variable trade barrier with a respect to trade volume m_{ij} between the exporter j and importer i; δ is non-linear parameter that measures the combined effect of the firm-level heterogeneity and non-random sample selection on trade volumes; $\beta_{u\eta}$ is a parameter controlling for non-random export selection and λ_j, χ_i are the exporter, importer fixed effects respectively.

The inference presented by HMR is based on the estimating equation (3) and merits detailed discussion. First, this equation controls for all discussed inconsistencies in estimating γ. Second, the extensive margin is controlled by the non-linear term $(\ln\{\exp[\delta(\widehat{z}^*_{ij} + \widehat{\overline{\eta}}^*_{ij})] - 1\})$ and requires the whole model to be estimated by the MLE. This measure is a linear combination of the selection effect $\widehat{\overline{\eta}}^*_{ij}$ and fraction of existing exporters $\widehat{z}^*_{ij}$. It is worth emphasizing that d_{ij} in the equation (2) also affects the estimate of δ. From the Figure A it is apparent that the lower variable trade barrier increases the productivity cutoff $\phi^{1-\varepsilon}_{ij,}$ meaning that for a given fixed export costs more firms are starting to export. The statistical significance, sign and magnitude of the extensive margin depends on the elasticity of substitution ε through the expression of δ. If the elasticity of substitution ε is high, than δ will be small. In the extreme case if $\varepsilon \to \infty \Rightarrow \delta \to 0$ and therefore the extensive margin will not be important in explaining the trade volumes. Moreover, for some values of ε, the estimate of δ may be statistically non-zero, but small enough so that it becomes negative.

If $\widehat{\delta}$ is significant, omitting the measure of the extensive margin bias the estimate of γ upwards (in the negative sense), since δ should have positive effect on export volumes, while correlation between the fraction of exporters and the variable trade barriers should be negative. Third, $\beta_{u\eta}$ captures non-random nature of zeros in the trade matrix. If this measure is omitted there should be a downward bias in the estimate of γ since the export countries with large observed trade costs are likely to have low unobserved trade costs. Also this measure does not depend on the elasticity of substitution ε and should be statistically significant in the estimation provided there is enough zero-trade relationships to indentify the selection equation (A18) in the first stage.

2.3 An Empirical Modification

The key result of the HMR model is showing that the omission of the extensive margin (heterogeneity bias) is most significant source of inconsistency in estimating γ in the gravity model (2) relative to controlling for the selection bias. However, since the estimator of the extensive margin $\widehat{\delta}$ depends on elasticity of substitution ε, the extent of the robustness of the HMR finding may be questionable if export composition of the trading partners is considered.

To account for differences in the export composition between the trading partners, I divide the set of all countries into two regions: the relatively skill-intensive North and the relatively labor-intensive South. On average, I expect the Northern countries to export differentiated varieties (manufactured products), while Southern countries to export homogeneous goods (agricultural and mineral products)[4]. Thus, once the regional effects are considered, the importance of the extensive margin becomes very sensitive to the value of elasticity of substitution ε[5]. In the regions where elasticity of substitution is low, the extensive margin should dominate the selection effect, while the opposite should emerge for the regions where elasticity of substitution is high. This reasoning is in line with Chaney (2008), who finds that higher elasticity of substitution magnifies the sensitivity of the intensive margin to changes in trade barriers, whereas it dampens the sensitivity of the extensive margin.

To test the robustness of the extensive margin, I modify the gravity specification (A14) to control for the regional response of trade barriers on trade flows through a refined measure of the intensive margin. This additional measure of the intensive margin can be introduced through region interaction terms in to the estimating gravity equation (A14). Let θ_j^r denote the indicator variable that is equal to one if the exporter j is a country from a region r and is zero otherwise. The estimating gravity equation then becomes:

$$m_{ij} = \beta_0 + \lambda_j + \chi_i - \beta_1 d_{ij} + \gamma\theta_j^r d_{ij} + w_{ij} + u_{ij}, \tag{3}$$

where β_1 captures the main impact of the trade barriers on the trade volumes, while γ (to be estimated) now captures an additional impact of the trade-barriers given a particular exporter region.

The sign of $\widehat{\gamma}$ in (3) depends on the exact region assignment for the interaction variable θ_j^r. If $\theta_j^r = 1$ when an exporter is from the skill-intensive region, I expect a positive estimated coefficient for γ. In this case, if the more skill-intensive technology lowers the marginal cost of production it will mitigate the negative impact of the trade-barriers on the firm-level exports. The opposite sign for γ should be expected if the firm is an exporter operates at a high marginal cost. Hence, the

[4] With the country-level data this is only a hypothesis.

[5] From the import demand elasticities reported by Broda and Weinsten (2006) I infer for example, that median elasticity of substitution in the food and vegetable sectors is 20 percent, while in various manufacturing sectors is 8 percent.

region interaction term proxies for the unobservable effect of technology differences among firms at the country-level once the firm-level heterogeneity has been partialed out. An omission of this proxy results in the biased estimation of w_{ij}. Since the estimate of w_{ij} depends on the elasticity of substitution ε, which in turn depends on the export composition of varieties, there is a non-zero correlation between the fraction of firms that choose to export and the regional response of change in the trade volumes to the change in the variable trade barrier. The direction of this bias depends on the choice of θ_j^r. When the skill-intensive region is picked, $\widehat{\delta}$ from $\widehat{\overline{w}}_{ij}$ will be overestimated[6]. My estimation results confirm the significance of the region interaction term with an expected sign.

3 Data

In my empirical analysis, I examine unidirectional bilateral exports for 158 countries that are split into the two regions: North and South over the 1972-1997 periods. The Northern countries are taken to be the OECD countries. Out of 158 countries in my sample there are 29 OECD countries. This leaves me with 129 developing/emerging (Southern) countries. The empirical framework discussed thus far only allows for cross-section analysis. To compare my results with the HMR, I estimate all the specifications for 1986. Hence for these specifications I drop five countries from the OECD sub-set since they were not part of the OECD block in 1986. In the additional robustness checks I estimate all the specifications for 1972 and 1996 modifying the OECD country set accordingly. The list of the OECD countries along with the accession dates is provided in the Table A1. The list of all the developing countries is provided in the Table A2.

The extended HMR model allows me to evaluate the impact of variable and fixed trade barriers on the export volumes at the country level controlling for the region of the exporter origin through the interaction of the trade barrier measures with the region indicator variable. The trade data comes from Feenstra's World Trade Flows, 1970-1992 and World Trade Flows, 1980-1997 constructed by the HMR. I use the same set of trade barriers as HMR. These trade barriers are constructed from the country level-data and come from the three sources: the Penn World Tables, the World Bank's World Development Indicators and the CIA'S World Factbook. Hence, my specification estimates can be straightforwardly compared to the results obtained by the HMR.

To gain an insight into importance of elasticity of substitution in the two-stage consistent estimating equation, I use the data on the import demand elasticities available from Broda and Weinstein (2006). This data set contains elasticities of imported products according to the six digit HS classification for 73 countries from 1994-2003. I calculate the mean of the import elasticities based on the product HS-code and use the median elasticity from each product group as a proxy for the elasticities of substitution.

The list of all the variables along with their description is provided in the Appendix B.

[6] If the regional response proxy is omitted $\widetilde{\delta} = \widehat{\delta} + \widehat{\gamma} corr(\theta_j^r d_{ij}, w_{ij})$. When an exporter is from the technologically advanced region ($\theta_j^r = 1$), $\widehat{\gamma} > 0$, $corr(\theta_j^r d_{ij}, w_{ij}) > 0$. Hence, the bias term is positive.

Preliminary Data Analysis

The expression for the trade volumes given the region of origin (3) suggests differences in number of trading partners regardless of the destination markets. In particular, for the skill-intensive North these differences quantitatively imply significantly fewer zeros in the trade matrix relative to the Southern exporters. To highlight the asymmetries in trade flows between the North (OECD) countries and the South (developing countries), I plot the distribution of country pairs based on the direction of trade. Figure 1 plots the fraction of country-pairs that engage in the trade in both directions (country j exports to country i and vice versa), in the one-way trade (only one of the trade partners exports) and in no trade (both trade partners do not export) for all the years in the sample given that exporter is one of the OECD countries. In Figure 2 the exporter is a developing country. Both figures[7] reveal an additional dimension in looking at the trade flow data: when considering the bilateral trade flows from the OECD countries, the fraction of the country pairs engaging in the two-way trade tremendously dominates that for the developing countries. Comparing these two figures to the original HMR calculation, where every country is symmetric, it appears that decomposition into developed and developing countries is fully subsumed in the World-World exports. Thus, when all the countries are treated symmetrically, the average effect of trade barriers on the trade volumes is obtained. It is overestimated for the developing countries, and it is underestimated for the developed countries.

In the regression analysis, I capture the differences among exporters by differences in responses of trade volumes to trade barriers given the exporter region of origin. Moreover, the extensive margin implicitly depends on effects of fixed trade barriers on the trade volumes. The preliminary significance of these interaction terms can be analyzed with fixed barrier mean-difference tests. The significant differences in the means between fixed trade barriers in the North and the South imply non-symmetric responses of these trade barriers on the exporter once the region of the exporter origin is accounted for. Table I provides the results of the mean-difference tests for all the fixed trade barriers that will be used in the main estimation. As most of these barriers are binary indicators, the means of these variables represent the fraction of exporters that face these barriers in the specific region. Consider, for example, the mean-difference test for the Language. This indicator variable takes the value of one if both the exporter j and the importer i speak same language. Thus, the exporters from the North have an average of twenty percent of the world trading partners who speak same language, while this fraction is around thirty percent for the exporters from the South. The difference in means for this trade barrier is significant at any level. Given that sharing same language should have a positive impact on trade volumes for the $j - i$ country pair, this test suggests that response of the trade volumes to the language trade barrier should be significantly different for the Northern exporter compared to the Southern exporter. The similar analysis applies to all other barriers. Across most of these barriers, I observe significant differences between the fixed trade barriers that an average exporter from a particular region faces.

[7] The distribution of direction of trade between OECD-OECD and Developing-Developing countries shows similar patterns.

4 Estimation Results

4.1 Traditional Gravity Empirics with the Region Controls

I begin by estimating the traditional gravity model to confirm theoretical predictions of the respective controls in the gravity regression. Importantly, I test whether the impact of the trade barriers on the trade volumes is lower for the Northern exporters as compared to the Southern. For example, the further apart the two country pairs are, the smaller the volume of trade, but less so if the exporter is a firm from the North. The extended empirical model (3) can be estimated with the following cross-sectional specification:

$$m_{ij} = \beta_0 + \lambda_j + \zeta_i - \beta_1 d_{ij} - \beta_2 \boldsymbol{\theta}_{ij} + \gamma_1 d_{ij} * North + \gamma_2 \boldsymbol{\theta}_{ij} * North + u_{ij}, \tag{4}$$

where m_{ij} is the log of the import volume of the trading partner i from the partner j; d_{ij} is a log of the variable trade barrier; $\boldsymbol{\theta}_{ij}$ is a vector of fixed trade barriers; $North$ is an indicator variable that is equal to 1 if an exporter is from the developed country (OECD) and is zero otherwise, and λ_j, ζ_i are the exporter and the importer fixed effects respectively. The coefficients on the interaction terms capture the differences in the export behavior between the firms in the both regions. The traditional gravity model uses data on the country pairs that trade at least in the one direction. As HMR, I take an alternative specification, by using the data on the unidirectional trade instead of constructing the symmetric trade flows for imports and exports for each country pair, but at the same time introduce the exporter and the importer fixed effects. The fixed effects capture underlying differences between trading partners that do not change in the given time period. This approach allows me to represent each country pair twice: once for exports from i to j and once for exports from j to i.

The results of the benchmark gravity estimation (4) for 1986 are reported in the column three of the Table II. There are only 11,146 non-censored observations out of 24,649 for the entire cross-section, reflecting no trade between many country pairs. All of the standard errors are clustered by country pairs to account for a bilateral trading partner relationship. For ease of comparison, the original HMR estimation results are shown in the first column. Even though the signs of the estimates are mostly the same in both specifications, the magnitudes differ substantially. As expected, the country j exports more to country i when they are closer to each other, speak the same language, are the members of the free trade agreements, share colonial ties, have the same legal system and share the same border, neither trading partner is an island nor land locked, and both share the same currency union. Interestingly, the sign on the common religion differs for both specifications, but in both it is not significant.

Controlling for differences in regional responses through interaction terms agree with my initial predictions. Consider the estimate of distance together with its interaction term: this suggests that when the country is a Northern exporter, the magnitude of the negative effect of distance on trading volume is reduced to 0.98 ($-1.255 + 0.279$). These estimates suggest that even before controlling for unobserved heterogeneity among the firms, the coefficient on distance has declined by 1.3 times when interacted with the region indicator variable and this effect is highly significant. That is, the effect of distance on export volume for any exporter from developed country is 1.3 times smaller than the same effect for a developing country. Similar differences in magnitudes are obtained for other trade barriers. Thus, even through the benchmark traditional gravity estimation, it is apparent that the HMR model overestimates the effect of trade barriers on the trade volumes for the exporters in the developed countries, and it underestimates the same effect for the exporters in the developing countries. While this paper documents only the empirical differences in region asymmetries using symmetric HMR model, the reason for these differences may stem from the export composition in the both regions.

In addition to the basic gravity estimates, columns (2) and (4) in Table II report the marginal effects of estimating the export selection Probit model (A18) that I extend by adding region-barrier interaction terms similar to the model (4). These marginal effects are evaluated at the sample means and can be directly interpreted as probabilities of the firm selection into export market with differential effects of region-barrier indicator controls. While the Probit estimates are used in the two-stage consistent estimating method, they are reported here to verify that the trade barriers that affect the export volumes also affect the probabilities that exporter j exports to i in the same way. The reported probability estimates readily confirm this conjecture. Importantly, similar to the benchmark gravity estimation, controlling for the region of the exporter origin mitigates the negative effect of trade barriers on the probability of the firm export selection for the Northern firms as compared to the Southern firms. The notable exceptions are the border and the religion barriers. As in the original HMR estimation a common border raises the volume of trade but reduces the probability of trading. The opposite result appears for the religion: common religion reduces the export volumes, but less so for the Northern exporter while increases the probability of trading, but less so for the Northern exporter. Interestingly, the coefficient of the interaction term is highly significant in export volumes estimation and not significant at any level in the Probit estimation. Hence, it appears that the common religion strongly affects the formation of trading relationships when the exporter is a Southern country and not important when exporter is a Northern country. The HMR estimate of the religion barrier is an average effect for the World-World export selection[8]. For some trade barriers (colonial ties, common border, currency union and free trade agreement) the region interaction effects can be estimated for the export volume specification but not for the export selection. The separate selection effects of these trade barriers cannot be identified for the Northern exporters as very few of them share colonial ties, many share the common border and are the members of a currency union or a free trade agreement with their trading partners[9]. Similar to the HMR finding, the export selection equation extended with region-barrier interaction terms

[8] The HMR estimate for the common religion is 10 percent, while my estimate is 11 percent when exporter is a Southern country and (11-0.05) 10.95 percent when exporter is a Northern country, but this effect is not significant.

[9] The average fraction of the Northern exporters who share colonial ties with their trading partner is 3 percent ; common border is 1 percent; members of a currency union is 0.2 percent and members of a free trade agreement is 1.7 percent (See Table I)

appears to be important in correcting the selection bias in the traditional gravity model.

4.2 Extended Two-Stage Gravity Estimation

I now turn to the empirical specification and the results of the extended consistent estimation of the gravity model (3) using two-stage procedure that was outlined in the Section 2.2. This procedure amounts to estimating the Probit selection equation (A18). The residuals obtained from this estimation are then used to derive the controls[10] for the extensive margin ($\widehat{\overline{w}}^*_{ij}$) and the non-random selection ($\widehat{\overline{\eta}}^*_{ij}$)[11] in the gravity specification (2). Also in the Section 2.3, I set up an empirical extension of the model (A14) to include region-barrier interaction controls. Such extension results in the model (3). Combining specification (2) with (3) and using the log-linear definition of the fixed export costs (fixed trade barriers) (A16), I obtain the extended consistent specification for the gravity model (4):

$$m_{ij} = \beta_0+\lambda_j+\chi_i-\beta_1 d_{ij}-\beta_2\boldsymbol{\theta}_{ij}+\gamma_1 d_{ij}*North+\gamma_2\boldsymbol{\theta}_{ij}*North+\ln\{\exp[\delta(\widehat{z}^*_{ij}+\widehat{\overline{\eta}}^*_{ij})]-1\}+\beta_{u\eta}\widehat{\overline{\eta}}^*_{ij}+e_{ij}, \tag{5}$$

where all the parameters were defined previously. Since this model in non-linear in δ, I estimate it using Maximum Likelihood Estimator (MLE).

To avoid the reliance on the normality assumption of the unobserved trade cost, estimating the second stage model (5) requires an exclusion restriction. This restriction should be selected such that it provides the measure of the fixed trade costs that affect the probability of the export selection, but not the export volumes. In the previous studies that have used this two-stage method, few different variables were suggested to satisfy this restriction requirement. For example, in the original HMR paper, the authors use regulation costs and the common religion, while Manova (2006) uses an island as an excluded variable. I follow the HMR and use common religion as an excluded variable. In the original HMR estimation this variable significantly affects the probability of the export selection, but it is not important once such decision has been made (this variable is not correlated with second-stage estimated residuals). This means that religion is only a fixed cost hurdle that an exporter faces. Once the exporter overcomes this cost, it does not affect the export volumes through its relation with per-unit variable trade cost. When I introduce the region-barrier interaction terms, the justification of the common religion as a valid exclusion restriction becomes less obvious. While the coefficient on the common religion is significant in the Probit estimation, the region-religion interaction term is not significant implying that for the Northern exporter religion does not affect the probability of selection. Nonetheless, one can strongly reject the null hypothesis

[10] The detailed steps on how to derive these estimators are provided in the Appendix A3.2.

[11] Similar to HMR (2008) and Manova (2006) the first-stage Probit estimation results in the small number of exporter-importer pairs whose probability of trade $\widehat{\rho}_{ij}$ is indistinguishable from 1 or 0. In this case it is not possible to infer any differences in the latent variable that controls for the extensive margin - $\widehat{z}^*_{ij}$. I assign $\widehat{\rho}_{ij} = 0.9999999$ for the country pairs where $\widehat{\rho}_{ij} = 1$ and $\widehat{\rho}_{ij} = 0.0000001$ for the country pairs where $\widehat{\rho}_{ij}$ cannot be estimated or it is very close to zero. This transformation eliminates 3.1 percent of the non-censored country-pairs (out of 11,146).

of the joint significance for this variable pair. Hence, overall the common religion appears to be an important fixed cost variable that affects export selection. In addition, I test whether common religion is correlated with the residuals from the second stage estimation. I find that, while there is no correlation between the residuals and common religion, there is correlation of 0.18 between the residuals and the region-religion interaction term. Hence, the evidence for the validity of the religion as an exclusion restriction appears to be mixed. However, the overall effect of this variable seems to pass both requirements.

The results of estimating the model (5) are reported in the Table III. Columns (1) and (2) are the HMR estimates of the benchmark and the corrected gravity models without region of origin controls respectively. I re-estimate the original HMR model (A20) using Maximum Likelihood Estimator (MLE) instead of Non-Linear Least Squares. This approach provides more efficient estimation and as evident from the column (2) gives almost identical estimates to the ones found by HMR. The coefficient on distance drops by almost a third, while the magnitude of the other fixed barriers are either reduced by the order of the magnitude or become insignificant. Moreover, the controls for the extensive margin $\widehat{\overline{w}}_{ij}^*$ and the non-random selection $\widehat{\overline{\eta}}_{ij}^*$ are significant with the former by almost one and half times larger in the magnitude then the latter. The key implication of this estimation is the importance of the extensive margin (heterogeneity bias) over the non-random selection.

Next, I consider estimates of the gravity model where the region of the exporter origin is controlled for. Column (4) in the Table III reports these estimates. It is immediately evident that these estimates continue to support my argument: the negative/positive effect of the trade barriers on the export volumes is mitigated for the Northern (OECD) exporters, and is magnified for the Southern exporters. Importantly, the magnitude of the overall effects of the trade barriers, while still overestimated in the benchmark estimation with region controls (column (3)), does not drop by as much as reported by the HMR. For example the coefficient on distance using two-stage method is 1.2 times smaller then the distance coefficient in the benchmark estimation. In the original HMR estimation this coefficient is 1.4 times smaller (see columns (1) and (2)). This difference is picked by the region-barrier interaction term. Thus, when all countries treated as symmetric partners the extensive margin seems to play the key role in explaining the bias in estimating the trading volumes, but the magnitudes of the effects of the trade barriers are roughly the averages of the same estimates when Northern and Southern exporters are considered separately. Crucially, the relative importance of the firm-level heterogeneity and non-random export selection appears to be reversed. While, they are still both positive, it is now the selection measure $\widehat{\overline{\eta}}_{ij}^*$, that by is almost one and half times larger in the magnitude then the measure of the extensive margin $\widehat{\overline{w}}_{ij}^*$.

Recall that originally, HMR claim that the omission of the extensive margin from the gravity model results in overestimation of the elasticity of the trade barriers with a respect to trade volumes - coefficients β_1 and β_2 in (5), but ommitting the export-selection correction $\beta_{u\eta}\widehat{\overline{\eta}}_{ij}^*$ results in underestimation of these elasticities. However, with a more refined data analysis, while it appears that these elasticities are overestimated, the role of the extensive margin in explaining this fact is questionable. To gain some initial insight into the importance of the extensive margin when regional differences are accounted for, I decompose the biases into separate estimating equations. The results of this decomposition are provided in the Table IV. In the first two columns of the Table IV, I report the benchmark and MLE estimates of the extended consistent gravity model that I have estimated previously (see Table III). The last two columns give the estimates when I just control for firm-level heterogeneity (column (3)) and non-random export selection (column(4)). To estimate the latter, I apply simple linear correction $\widehat{z}^*_{ij} = \Phi^{-1}(\widehat{\rho}_{ij})$. This model can be estimated using OLS, as there is no non-linearity in the extensive margin estimator. To estimate the former, I use the two-step consistent Heckman sample selection model. In this case $\widehat{\overline{\eta}}^*_{ij}$ is the reported Mills Ratio. In both models, I continue to control for regional effects through region-barrier interaction terms.

It is evident from this decomposition that firm-level heterogeneity appears to explain almost all the biases in the standard gravity equation even when I control for regional differences among the exporters. This result is in-line with the original HMR findings. However, the estimate of the countries pairing into exporter-importer relationship $\widehat{\overline{\eta}}^*_{ij}$ is slightly larger than the estimate of the unobserved firm-level heterogeneity $\widehat{z}^*_{ij}$. Hence, it appears that, while within country variation in the fraction of exporters explains the biases in the standard gravity model, the non-random selection is equally as important. I now further split the data to determine the significance of the extensive margin based on the theoretical prediction: the extensive margin should be most important in explaining biases in the gravity model when the exporter in an OECD country.

4.3 Bias Decomposition Based on the Exporter-Importer Region of Origin

The main estimation discussed in the previous section has led to the following conjecture: the extensive margin is important in explaining the biases in the standard gravity model, but it is region dependent. That is, when I control for region of the exporter origin the magnitude of the impact of the extensive margin on export volumes is reduced substantially compared to the simple World-World trade estimation.

To further explore the significance of the extensive margin, I drop the region-barrier interaction terms and estimate the original HMR model (A20) on the sub-set of countries based on exporter-importer trading region. In my data set, I have the exporter and importer country code. This allows me to construct four exporter-importer trading zones: North-North, North-South, South-North and South-South. The definition for the Northern countries and Southern countries remains the same. I estimate the model (A20) using Maximum Likelihood Estimator (MLE), use religion as an excluded variable throughout all the specifications and perform bias decomposition. The statistical significance and relative magnitude of $\widehat{z}^*_{ij}$ to $\widehat{\overline{\eta}}^*_{ij}$ will indicate how important is the extensive margin relative to the selection into exporter-importer relationship. When the Northern country is an exporter, I expect the extensive margin to dominate. The exports from these countries are primarily differentiated manufactured goods and services. As the elasticity of substitution ε for these varieties is low, I expect the estimate of the extensive margin δ to be large and significant. Conversely, when the exporter is a Southern country, the export composition should consist primarily of the homogeneous agricultural and/or natural resource products. In this case, the elasticity of substitution ε should be high making $\widehat{\delta}$ small and insignificant. Thus, the elasticity estimates of the trade barriers with a respect to the trade flows in the gravity model, where the OECD (Northern) country is an exporter, must be biased upwards if the extensive margin correction is ommitted. Conversely, when the Non-OECD (Southern) country is an exporter these estimated should be biased downwards if the measure of the non-random export selection measure $\widehat{\overline{\eta}}^*_{ij}$. is ommitted. In the cross-regional trade, I expect some mixed results, but similarly in the North-South trade an extensive margin should prevail over the export-selection so that ommitting both corrections should result in the upwards bias in the elasticity estimates.

The estimates of these bias decompositions are provided in Tables V (A-D) for each trade zone combination respectively. In each of the tables I report the number of censored observations along with total number of underlying observations. When the number of censored and total observations is nearly the same, it is impossible to identify the selection equation. As a result, the relative importance of the extensive margin to the non-random selection cannot be determined. I encounter such situation when the North-North trading relations are being considered (Table V (A)). Since the number of censored observations is six, both $\widehat{z}^*_{ij}$ and $\widehat{\overline{\eta}}^*_{ij}$ are insignificant as partial effects of these variables cannot be determined. Thus, even though I expect the extensive margin to be most significant in this trading relationship, I cannot convincingly conclude so.

First, I find that while $\widehat{\delta}$ is insignificant in each second-stage trading region estimations, the sample selection correction $\widehat{\overline{\eta}}^*_{ij}$ is always significant except for North-North estimation for the reasons discussed earlier. Since the measure of the extensive margin is a non-linear combination of imputed probabilities of export selection and non-random sample selection correction, it appears that the significance of the extensive margin in the HMR estimation is driven by non-random sample selection rather than firm-level heterogeneity. Second, the puzzling outcome occurs with the North-South trading relation. While the number of censored observations is fairly high to identify the selection equation, the estimate of the extensive margin is not significant in the simple linear correction model ($\widehat{z}^*_{ij}$). That is at the country level firm-level heterogeneity plays no role in explaining trade flows from the Northern exporters. The estimates of $\widehat{z}^*_{ij}$ and $\widehat{\overline{\eta}}^*_{ij}$ for the South-North and the South-South trading relations are in-line with my initial analysis. In the bias decomposition for the South-North, the extensive margin estimate is significant, but has a negative sign. This can be explained by high elasticity of substitution for the exported varieties. For the South-South, the extensive margin is not significant, while non-random selection effect dominates. Interestingly, the coeffient on distance in the benchmark model (without export-selection correction) is smaller, than the same estimate when this correction is applied[12]. Thus, as early predicted, when the export-selection plays most important role in explaining the trade flows, omission of this correction leads to downward bias in the bencmark model. More generally, these findings highlight that the extensive

[12] In Table V(D) these estimates are -1.214 and -1.414 respectively

margin of trade that corrects for the firm-level heterogeneity is only significant and important only for the aggregate trade-flows. Once the trading relations are more refined, the effect of extensive margin on the trade flows disappears.

The table below summarizes these findings:

Extensive Margin vs. Non-Random Export Selection

Region-Pair	Extensive Margin	Export Selection	Data Support?
N-N	✓	—	Inconclusive
N-S	✓	✓	No
S-N	✓	✓	Partial
S-S	—	✓	Yes

In this table the second and the third columns identify the relative importance of the extensive margin to the non-random export selection that should hold in theory. For example, the check mark and the dash for North-North trading relation means that the extensive margin should explain all the biases in the standard gravity model, while export selection should not play a role. For the South-South trading relation the importance of these controls is reversed. In the cross-region trade both extensive margin and export selection should play some role with extensive margin slightly dominating in the North-South trade. The last column indicates if these relationships hold in the data. For the North-South trade (Table V(B)), while coeffiecient on distance seems to be slightly overestimated in the benchmark model as compared to the two-stage model estimates, it appears that extensive margin correction cannot explain this upward bias as it is not significant in any of the specifications, which as I mentioned earlier is a puzzling result.

4.4 Robustness of the Extensive Margin over Time

One of the main deficiencies of the Melitz framework is inability to endogenize the technology accumulation over time. Thus, it is not possible to estimate a structural gravity model on the panel with many countries over some span of years. While the main set of results obtained in this paper came from estimating the cross-sectional regressions for 1986, the question arises how the importance of extensive margin changes over time.

To address this question, I estimate the main extended model (5) for 1972 and 1996[13]. The results of these estimations are provided in the Table IV. The magnitudes and signs of the respective coefficients on the trade barriers and region-barrier interaction terms remain nearly the same as for original estimation for 1986. Similar to the 1986 cross-section, for the Northern exporter the negative impact of the trade barriers on trade volumes is mitigated, while the positive impact is amplified. Thus, the effects of trade barriers on trading volumes seem to be robust over time.

[13] 1972 is the year when the data sample begins. While, I have the data through 1997, this year coincide with Asian financial crisis. To avoid any perverse results that can be associated with this year, I select 1996.

The key interest lies in the estimates of the extensive margin $\widehat{\delta}$ and the non-random selection $\widehat{\overline{\eta}}^*_{ij}$. If the technological accumulation matters, than over time the relative importance of the extensive margin and export selection should reverse their roles. As more developing countries acquire new skills through increasing demand for education or spillovers from multinationals, the firms in these countries considerably reduce their unit-costs and are able to export mode differentiated varieties. Thus, over time the composition of exports in the developing countries should begin to resemble that for the developed countries where firm-level heterogeneity seems to play the key role in explaining the biases in the standard gravity model of trade. The results reported in the Table IV strongly support this analysis. Interestingly, not only the roles of firm-level heterogeneity and export selection have reversed from 1972 to 1996, their magnitudes have changed roughly by the order of two. That is when I estimate the standard gravity model for 1972, I find that the failure to control for export selection biases the estimates of the trade barriers in the standard gravity model with much smaller effect of the extensive margin. For the 1996 estimation the opposite outcome holds. From the bias decomposition analysis for 1986, the relative importance of the extensive margin to the export selection lies somewhere in between.

With the fall of the unit-costs over time, the probability of the export selection in the developing countries should rise as more exporters would be able to overcome fixed export costs. To check if this holds true, I select France and Paraguay as two representative countries from the North and the South respectively according to their median distance to all other trading partners. The average probabilities of export selection for 1972 and 1996 can be inferred from the residuals obtained by estimating the Probit specification (A18) for each of the cross-sections. I find that for the Paraguay the average probability of the export selection was 0.29 in 1972 and 0.38 in 1996. Even though the increase is not very considerable it reflects the export trends in Paraguay[14]. Finally, I plot the residual probabilities of export selection for France and Paraguay against the export volumes for 1972, 1986 and 1996. The respective Figures 3 (a-c) show that, while for France the probability of export selection is close to 1 for all the years, these probabilities seems to converge to France especially for 1996. This result further confirms the change in the importance of the extensive margin which is associated with a change in the export composition over time.

5 Conclusion

This paper builds on the HMR model to determine how robust the importance of the extensive margin (number of exporting firms) in explaining the biases in the standard gravity model. I apply the HMR methodology on the more refined world-trade data set, where I split the data by the regions of the exporter origins. The motivation for doing this is to check if the predictions of the HMR model continue to hold for the trade flows that theoretically must favor the extensive margin correction over the export-selection in explaining the trade flows.

[14] According to the World Bank report (2007) for example, in 1986 Paraguay exported \$519 million worth of the manufactured goods, while in 2006 this figure was \$1,500 millions - an almost 2 percent increase.

My estimation results confirm that elasticity estimates in the benchmark gravity model are overestimated regardless of the data split as shown by HMR. However, unlike HMR I find that the extensive margin correction cannot fully explain this upward bias when the trade between OECD and non-OECD countries is considered. One of the explanation for this finding is the effect of the elasticity of substitution in determining the signficance of the extensive margin. For the countries that primarily export differentiated products, ommitting the extensive margin correction results in the upwards bias in the elasticity of trade barriers with a respect to trade volumes (the intensive margin), while for the countries that export homogeneous products, the intensive margin estimates are biased downwards when not corrected for the firm selection. I partially confirm these predictions by the bias decomposition in the estimation of the gravity model given the trade-relation regions. Importantly, contrary to the theoretical predictions, I find it puzzling that extensive margin is insignificant for the North-South trade.

The cross-section estimations of the extended gravity model reveal the change in the magnitude of the extensive margin relative the export selection. This measure was four times larger in the estimation for 1996 compared to 1972. Such significant increase can be attributed to the change in the export composition over time. Perhaps, through the technology accumulation over time the developing countries are switching from exporting homogeneous goods to differentiated products so that their export patterns resemble the developed countries. However, a separate study is required to explore this linkage. The present framework fails to capture this time dimension, as it requires endogenizing the growth in technology accumulation over time.

The HMR framework appears to be elegant in the implementation. It provides a bridge between the "new trade" theory and econometric estimations. With a relatively simple extension, I was able to obtain additional important insights that challenge the conclusions of the HMR model. A more detailed indutsry-level study is needed, however to fully explore what indentifies the extensive margin in the trade flows.

A The HMR Model

Note: Most of the contents in the Appendix A is taken from the original paper by Helpman, Melitz and Rubinstein. It is reported here for the reference purpose only.

A.1 Set Up

The theoretical part of the HMR model is slightly modified and simplified version of the Melitz(2003) model. Consider a world with J countries, that are indexed by $j = 1, 2..., J$ and unpartioned by any region. There is a set of varieties l available for a consumption in every country j that is denoted by $B_{j.}$ The demand for each variety is derived from the CES utility function that is common to the every country j:

$$U_j = \left[\int_{l \in B_j} q_j(l)^{\alpha} dl \right]^{1/\alpha}, 0 < \alpha < 1, \tag{A1}$$

where $q_j(l)$ is its consumption of the product l and the parameter α determines the elasticity of substitution across products. This constant elasticity can be defined as $\varepsilon = \frac{1}{1-\alpha}$ and it is the same in every country. Given the parameter restrictions on α, $\varepsilon > 1$.

Let Y_j be the income of the country j, which is equal to some expenditure level such that $U_j \equiv Y_j$. This notation gives the following budget constraint:

$$Y_j = \int_{l \in B_j} p_j(l) q_j(l) dl, \tag{A2}$$

where $p_j(l)$ is the price of product l in any country j. Maximizing (A1) subject to (A2), the demand for the product l in any country j is

$$q_j(l) = \frac{p_j(l)^{-\varepsilon} Y_j}{P_j^{1-\varepsilon}} \text{ with } P_j = \left[\int_{l \in B_j} p_j(l)^{1-\varepsilon} dl \right]^{1/(1-\varepsilon)}, \tag{A3}$$

where $p_j(l)$, Y_j, ε (constant elasticity of substitution) are defined as above and P_j is the country's j an ideal price index.

A.2 Production and Trade Volumes

As in standard Melitz(2003) model, in any country j there is a continuum of firms of measure N_j each producing a differentiated variety l in a monopolistically competitive environment. Additionally, the varieties produced by the firms in country j are distinct from the varieties produced by the firms in country i, for $i \neq j$. Hence there are $\sum_{j=1}^{J} N_j$ products in the world economy.

To participate in the domestic and the export production firms in any country j bear variable and fixed costs. The variable cost is assumed to be a production cost which is a combination of the country specific cost c_j and per-unit firm-specific marginal cost a. The inverse of this marginal cost $(1/a)$ represents the firm productivity level that is different across firms in the same country. Thus, the firm with the lowest marginal cost a is the most productive. Given this notation, each firm in country j is producing a variety l using cost-minimizing combination of inputs $c_j a$. To determine how productive a firm j is, there is a cumulative distribution function of the marginal costs $G(a)$ with the support $a_H > a_L > 0$. This distribution is common to all countries.

When producing for the domestic market, the HMR model assumes that any firm j bears only variable production cost $c_j a$ and no fixed costs. Denote any fixed cost as f_{ij}, where i is any foreign country and j is any domestic country. The assumption of zero fixed costs to produce for the domestic market means that $f_{jj} = 0$. If a firm in the country j decides to enter the export market, it bears non-zero fixed cost $f_{ij} > 0$ and per-unit variable 'ice-berg' type transport cost[15] $\tau_{ij} > 1$, such that τ_{ij} units of any variety l must be shipped from a country j for one unit of this variety to arrive to a country i.

With the monopolistic competition, the firms choose price $p_j(l)$ of a variety l to maximize profits, taking demand (A3) as given. Thus, any firm j solves the following problem:

$$\max_{p_j(l)} \Pi = p_j(l) q_j(l) - c_j a \tau_{ij} q_j(a) - f_{ij}. \tag{A4}$$

The solution to the problem (A4) yields the following expression for the delivery price of variety l from exporter j to importer i:

$$p_j(l) = \tau_{ij} \frac{c_j a}{\alpha}. \tag{A5}$$

This is a standard mark-up pricing equation, with the mark-up $1/\alpha$ diminishing in the elasticity of demand α, adjusted for per-unit transportation cost τ_{ij}. Substitution of (A5) and (A3) into (A4) yields the operating profits from sales into a country i that are associated with this price level:

$$\pi_{ij}(l) = (1-\alpha) \left(\frac{\tau_{ij} c_j a}{\alpha P_i} \right) Y_i - c_j f_{ij}. \tag{A6}$$

While the assumption of zero domestic fixed costs $f_{jj} = 0$ implies that every firm will produce in the domestic market ($\pi_{jj}(l) > 0$), only a fraction $G(a_{ij})$ of all firms in country j will choose to export. The export participation cut-off a_{ij} can be implicitly found from the zero-profit condition such that $\pi_{ij}(l) = 0$. Hence this cut-off defines the minimum level of productivity or alternatively the maximum marginal cost required for an exporter in a country j to at least break-even:

$$(1-\alpha) \left(\frac{\tau_{ij} c_j a_{ij}}{\alpha P_i} \right) Y_i = c_j f_{ij}. \tag{A7}$$

The bilateral trade volumes regardless of any region can be expressed as:

$$V_{ij} = \begin{cases} \int_{a_L}^{a_{ij}} a^{1-\varepsilon} dG(a) & \text{for } a_{ij} \geq a_L \\ 0 & \text{otherwise} \end{cases}. \tag{A8}$$

Substitution of the pricing rule (A5) and the trade volume expression (A8) into the demand function (A3) yields an expression for the value of a country i's imports from a country j:

$$M_{ij} = \left(\frac{\tau_{ij} c_j}{\alpha P_i} \right)^{1-\varepsilon} Y_i N_j V_{ij}. \tag{A9}$$

Whenever $a_{ij} \leq a_L$, this trade volume is zero since $V_{ij} = 0$. Finally, using the definition (A8), (A5) and (A3), the ideal price index in a country i is:

$$P_i^{1-\varepsilon} = \sum_{j=1}^{J} \left(\frac{\tau_{ij} c_j}{\alpha} \right)^{1-\varepsilon} N_j V_{ij}. \tag{A10}$$

[15] Since there are no transportation costs to deliver to the domestic market $\tau_{jj} = 1$

Equations (A7)-(A10) provide mapping from the income levels Y_i, the number of the firms N_i, the unit costs c_i, the fixed costs f_{ij}, and the transportation costs τ_{ij} to the bilateral trade flows M_{ij}.

A.3 Empirical Framework

Assume that $G(a)$ follows Pareto-truncated distribution with the following CDF:

$$G(a) = \frac{(a^k - a_L^k)}{(a_H^k - a_L^k)}, k > \varepsilon - 1, [a_{L,}a_H]. \tag{A11}$$

As theoretical implications of the model require, this CDF can capture the case of the zero trade flows such that $a_{ij} < a_{L,}$ $(V_{ij} = M_{ij} = 0)$ as well as an asymmetric trade flows where $M_{ij} \neq M_{ji}$ for some $i - j$ country pairs. Differentiating (A11) with respect to a^k (A8) becomes:

$$V_{ij} = \frac{k a_L^{k-\epsilon+1}}{(k-\epsilon+1)(a_H^k - a_L^k)} W_{ij}, \tag{A12}$$

where $W_{ij} = \max\left\{\left(\frac{a_{ij}}{a_L}\right)^{k-\epsilon+1} - 1, 0\right\}$ and a_{ij} is determined from the zero-profit condition (A7). Both V_{ij} and W_{ij} are monotonic functions of the proportion of exporters from j to i.

Log-Linearizing (A9) the estimating gravity model can written as:

$$m_{ij} = (\epsilon - 1)\ln\alpha - (\epsilon - 1)\ln c_i + n_j + (\epsilon - 1)p_i + y_i + (\epsilon - 1)\ln\tau_{ij} + v_{ij.} \tag{A13}$$

The variable costs that affect the firm-level exports are captured by the logarithm of 'ice-berg' type cost τ_{ij} and are the same for any exporter regardless of the region. These costs are stochastic due to an i.i.d. unmeasured trade frictions u_{ij} which are country-pair specific. Letting $\tau_{ij}^{\epsilon-1} \equiv D_{ij}^{\gamma} e^{u_{ij}}$, where D_{ij} represents symmetric distance between i and j and $u_{ij} \sim N(0, \sigma_u^2)$ the estimating gravity equation (A13) becomes:

$$m_{ij} = \beta_0 + \lambda_j + \chi_i - \gamma d_{ij} + w_{ij} + u_{ij}, \tag{A14}$$

where $\lambda_j = -(\epsilon - 1)\ln c_j + n_j^r$ is exporter fixed effects and $\chi_i = (\epsilon - 1)p_i + y_i$ is importer fixed effects.

A.3.1 Firm Export Selection

Denote the latent variable Z_{ij} to be the ratio of the variable export profits of the most productive firm (with productivity $\frac{1}{a_L}$) to the fixed export costs for exports from j to i:

$$Z_{ij} = \frac{(1-\alpha)\left(P_i \frac{\alpha}{c_j \tau_{ij}}\right)^{\varepsilon-1} Y_i a_L^{1-\varepsilon}}{c_j f_{ij}}. \tag{A15}$$

Assume that f_{ij} are stochastic fixed costs due to unmeasured i.i.d friction v_{ij}˜$N(0, \sigma_v^2)$ that may be correlated with u_{ij} and are defined as follows:

$$f_{ij} \equiv \exp(\phi_{EX,j} + \phi_{IM,i} + \mu\phi_{ij} - v_{ij}), \tag{A16}$$

where $\phi_{IM,i}$ is a fixed trade barrier imposed by the importing country, $\phi_{EX,j}$ is a measure of fixed export costs common across all export destinations and ϕ_{ij} is an observed measure of any additional country-pair specific fixed trade costs[16]. With this assumption the latent variable Z_{ij} in (A15) can now be expressed as:

$$z_{ij} \equiv \ln(Z_{ij}) = \gamma_0 + \xi_j + \zeta_i - \gamma d_{ij} - \mu\phi_{ij} + \eta_{ij}, \tag{A17}$$

where $(\varepsilon - 1)\ln \tau_{ij} \equiv \gamma d_{ij} - u_{ij}$; $\eta_{ij} \equiv u_{ij} + v_{ij}\tilde{}N(0, \sigma_u^2 + \sigma_v^2)$ is i.i.d. but correlated with an error term u_{ij} in the gravity model (A14); $\xi_j = -\varepsilon \ln c_{ij} + \phi_{EX,j}$, $\zeta_i = (\varepsilon - 1)p_i + y_i - \phi_{IM,i}$ are exporter and importer fixed effects respectively. Even though z_{ij} is unobserved, it is positive whenever j exports to i i.e. there is non-zero value of the export volumes in the bilateral trade matrix and it is zero otherwise.

To obtain the export selection equation, define the indicator variable $T_{ij} = 1$ if the country j exports to country i regardless of the region of the exporter origin and zero otherwise. Let ρ_{ij} be the probability that the country j exports to the country i conditional on the observed variables. The export selection equation is the following Probit specification:

$$\rho_{ij} = \Pr(T_{ij} = 1|\text{observed variables}) = \Phi(\gamma_0^* + \xi_j^* + \zeta_i^* - \gamma^* d_{ij} - \mu^*\phi_{ij}), \tag{A18}$$

where $\Phi(\bullet)$ is a CDF of the unit-normal distribution, and every starred coefficient represents the original coefficient divided by σ_η.

To obtain the consistent estimate of W_{ij} , let $\widehat{\rho}_{ij}$ be the predicted probability of exports from j to i that can be obtained from the estimated residuals in the Probit equation (A18). Given the vector of these predicted probabilities, the estimated fraction of exporting firms can be backed out by taking an inverse of the unit-normal CDF $\Phi(\bullet)$ - $\widehat{z}_{ij}^* = \Phi^{-1}(\widehat{\rho}_{ij})$. A consistent estimate for W_{ij} is:

$$W_{ij} = \max\{(Z_{ij})^\delta - 1, 0\}, \tag{A19}$$

where $\delta \equiv \sigma_\eta(k - \varepsilon + 1)/(\varepsilon - 1)$ and δ needs to be estimated.

A.3.2 Consistent Estimation of the Gravity Model

There are two requirements to obtain consistent estimate of γ in the gravity specification (A14). There should be a control variable for endogenous number of exporters (via w_{ij}) $E[w_{ij}|., T_{ij} = 1]$ and a control variable for selection of a country into the trading partner $E[u_{ij}|., T_{ij} = 1]$. Both of these terms depend on $\widehat{\overline{\eta}}_{ij}^* \equiv E[\eta_{ij}^*|., T_{ij} = 1]$. Also $E[u_{ij}|., T_{ij} = 1] = corr[(u_{ij}, \eta_{ij}), (\frac{\sigma_u}{\sigma_\eta})\overline{\eta}_{ij}^*]$. Since $\overline{\eta}_{ij}^*$ has a CDF of the unit-normal distribution, a consistent estimate $\widehat{\overline{\eta}}_{ij}^*$ can be obtained from the inverse Mills ratio: $\widehat{\overline{\eta}}_{ij}^* = \frac{\phi(\widehat{z}_{ij}^*)}{\Phi(\widehat{z}_{ij}^*)}$, or estimated from Heckman procedure available from any statistical package provided a valid exclusion restriction. Finally $\widehat{\overline{z}}_{ij}^* \equiv \widehat{z}_{ij}^* + \widehat{\overline{\eta}}_{ij}^*$ is a consistent estimate for $E[z_{ij}^*|., T_{ij} = 1]$ and $\widehat{\overline{w}}_{ij}^* \equiv \ln\{\exp[\delta(\widehat{z}_{ij}^* + \widehat{\overline{\eta}}_{ij}^*)] - 1\}$ is a consistent estimate for $E[w_{ij}|., T_{ij} = 1]$ from (A19). Hence the consistent estimating gravity model is now given by:

$$m_{ij} = \beta_0 + \lambda_j + \chi_i - \gamma d_{ij} + \ln\{\exp[\delta(\widehat{z}_{ij}^* + \widehat{\overline{\eta}}_{ij}^*)] - 1\} + \beta_{u\eta}\widehat{\overline{\eta}}_{ij}^* + e_{ij}, \tag{A20}$$

where $\beta_{u\eta} \equiv corr[(u_{ij}, \eta_{ij}), (\frac{\sigma_u}{\sigma_\eta})]$ and e_{ij} is i.i.d. distributed error term satisfying $E[e_{ij}|., T_{ij} = 1] = 0$. Since (A20) is non-linear in δ, I estimate it using MLE (unlike the HMR who use the NLS).

[16] See Appendix B for the list of such costs

B Description of the Main Variables

Note: The data used for this paper is identical to the HMR's paper. Below are the definitions of all the variables.

Dependent Variables

- *trade* volume - Unidirectional value of trade volumes between the $i-j$ country pair (in logs).
- trade - a binary variable which is equal to one if *trade volume* is non-zero,and is zero otherwise.

Explanatory Variables

Variable Trade Barrier

- *distance* - the symmetric distance between the importer's i and the exporter's j capitals (in logs).

Fixed Trade Barriers

- *common* border - a binary variable which is equal to one if the importer i and the exporter j share same physical border, and is zero otherwise.
- *island* - a binary variable which is equal to one if the importer i and the exporter j are both islands, and is zero otherwise.
- *landlocked* - a binary variable which is equal to one if the importer i and the exporter j have both no coastline or direct access to the sea, and zero otherwise.
- *colonial ties* - a binary variable which is equal to one if the importer i had ever colonized the exporter j or vice versa, is zero otherwise.
- *currency union* - a binary variable that is equal to one if the importer i and the exporter j use same currency or if within the country pair money was interchangeable at 1:1 exchange rate for an extended period of time (see Rose (2000), Glick and Rose (2002) and Rose (2004)), and is zero otherwise.
- *legal system* - a binary variable that is equal to one if the importer i and the exporter j share the same legal origin, and is zero otherwise.
- *religion* - (% Protestants in country i $\cdot$ % Protestants in country j) + (% Catholics in country i $\cdot$ % Catholics in country j) + (% Muslims in country i $\cdot$ % Muslims in country j) .
- *FTA* - a binary variable that is equal to one if the importer i and the exporter j belong to a common regional trade agreement, and is zero otherwise.
- *language* - a binary variable that is equal to to one if the importer i and the exporter j speak the same language, and is zero otherwise.

Interaction Terms

Both variable and fixed trade barrier variables are interacted with a following variable:

- *North* - a binary variable that is equal to one if the exporter j belongs to a group of the OECD countries[17], and is zero otherwise.

[17] See Table A1 for the list of these countries

References

Anderson, J. E. and Van Wincoop, E. (2004), “Trade costs”, NBER Working Paper No. W10480.

Anderson, J. E. and Wincoop, E. v. (2003), “Gravity with gravitas: A solution to the border puzzle”, *The American Economic Review* , Vol. 93, pp. 170–192.

Broda, C. and Weinstein, D. E. (2006), “Globalization and the gains from variety”, *Quarterly Journal of Economics* , Vol. 121, pp. 541–585.

Chaney, T. (2008), “Distorted gravity: Heterogeneous firms, market structure and the geography of international trade”, *American Economic Review* . forthcoming.

Demidova, S. A. (2006), “Productivity improvements and falling trade costs: Boon or bane?”, Working Paper.

Falvey, R., Greenaway, D. and Yu, Z. (2006), “Extending the melitz model to asymmetric countries”, University of Nottingham Research Paper No. 2006/07.

Glick, R. and Rose, A. K. (2002), “Does a currency union affect trade? the time-series evidence”, *European Economic Review* , Vol. 46, pp. 1125–1151.

Helpman, E. (1987), “Imperfect competition and international trade: Evidence from fourteen industrial countries”, *Journal of the Japanese and International Economies* , Vol. 1, pp. 62–81.

Helpman, E. (1999), “The structure of foreign trade”, *The Journal of Economic Perspectives* , Vol. 13.

Helpman, E., Melitz, M. J. and Yeaple, S. R. (2004), “Export versus fdi with heterogeneous firms”, *The American Economic Review* , Vol. 94.

Helpman, E., Melitz, M. and Rubinstein, Y. (2008), “Estimating trade flows: Trading partners and trading volumes”, *Quarterly Journal of Economics* , Vol. 123, pp. 441–487.

Hummels, D. and Levinsohn, J. (1995), “Monopolistic competition and international trade: Reconsidering the evidence”, *The Quarterly Journal of Economics* , Vol. 110.

Johnson, R. C. (2007), “Trade and prices with heterogeneous firms”, Job Market Paper.

Manova, K. B. (2006), “Credit constraints, heterogeneous firms and international trade”, Working Paper.

Melitz, M. J. (2003), “The impact of trade on intra-industry reallocations and aggregate industry productivity”, *Econometrica* , Vol. 71, pp. 1695–1725.

Rose, A. K. (2000), “One money, one market: the effect of common currencies on trade”, *Economic Policy* , Vol. 15, pp. 7–46.

Rose, A. K. (2004), “Do we really know that the wto increases trade?”, *The American Economic Review* , Vol. 94, pp. 98–114.

Tinbergen, J. (1962), *Shaping the World Economy*, New York: The Twentieth Century Fund.

Wooldridge, J. M. (2002), *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: The MIT Press.

Table I - The Descriptive Effects of the Trade Barriers by Region of Origin

Year=1986	North (total = 4,082)		South (total=20,724)		
Fixed Barrier	Mean	Variance	Mean	Variance	t-Test
Language	0.204	0.403	0.302	0.459	(13.843)***
FTA	0.017	0.128	0.004	0.061	(-6.296)***
Colonial	0.030	0.172	0.006	0.075	(-9.057)***
Religion	0.199	0.275	0.227	0.284	(5.774)***
Legal	0.249	0.433	0.393	0.488	(18.986)***
Border	0.015	0.122	0.018	0.132	1.255
Island	0.326	0.469	0.373	0.484	(5.748)***
Land Locked	0.278	0.448	0.269	0.444	-1.104
Currency Union	0.002	0.049	0.011	0.104	(7.992)***

Notes: t-test compares mean differences; unequal pair variance is assumed; *** significant at 1%

Table II - Benchmark Gravity and Firm Export Selection by Region

Year 1986	(1)	(2)	(3)	(4)
COEFFICIENT	HMR m_{ij}	Probit-HMR T_{ij}	m_{ij}	Probit T_{ij}
Distance	-1.176***	-0.263***	-1.255***	-0.265***
	(0.028)	(0.0096)	(0.034)	(0.0095)
Distance * North			0.279***	0.0616
			(0.051)	(0.042)
Language	0.176***	0.113***	0.131**	0.106***
	(0.056)	(0.015)	(0.066)	(0.015)
Language*North			0.205**	0.0995**
			(0.095)	(0.049)
FTA	0.759***	0.494***	0.838***	0.505***
	(0.17)	(0.017)	(0.17)	(0.018)
Colonial	1.299***	0.128	1.232***	0.0843
	(0.098)	(0.094)	(0.15)	(0.10)
Colonial*North			-0.0325	
			(0.19)	
Religion	0.102	0.104***	-0.0147	0.111***
	(0.091)	(0.023)	(0.10)	(0.024)
Religion*North			0.329***	-0.0506
			(0.12)	(0.068)
Legal	0.486***	0.0384***	0.456***	0.0328***
	(0.045)	(0.012)	(0.055)	(0.012)
Legal*North			0.163*	0.0885**
			(0.084)	(0.044)
Border	0.458***	-0.148***	0.554***	-0.154***
	(0.12)	(0.036)	(0.15)	(0.036)
Border*North			-0.501*	
			(0.26)	
Island	-0.391***	-0.136***	-0.315***	-0.110***
	(0.11)	(0.029)	(0.12)	(0.029)
Island*North			-0.213**	-0.209***
			(0.096)	(0.036)
Landlocked	-0.561***	-0.0717*	-0.499***	-0.0547
	(0.16)	(0.038)	(0.17)	(0.039)
Landlocked*North			-0.168	-0.0722
			(0.10)	(0.048)
Currency Union	1.364***	0.190***	1.258***	0.200***
	(0.25)	(0.044)	(0.27)	(0.045)
Observations	11,146	24,649	11,146	24,463
R^2	0.71	0.587	0.71	0.584

Notes: Exporter and Importer Fixed Effects; North =1 if exporer j is from OECD country; Marginal Effects at sample means are reported for Probit; Robust Standard Errors are in parenthesis, Pseudo R^2 is reported for Probit

Table III - The Consistent Gravity Model Estimation with Region Controls

Year 1986	(1)	(2)	(3)	(4)
COEFFICIENT	HMR Benchmark	HMR MLE	Benchmark	MLE
Distance	-1.176***	-0.807***	-1.255***	-0.992***
	(0.028)	(0.032)	(0.034)	(0.038)
Distance * North			0.279***	0.207***
			(0.051)	(0.049)
Language	0.176***	0.0272	0.131**	0.0454
	(0.056)	(0.055)	(0.066)	(0.064)
Language*North			0.205**	0.0899
			(0.095)	(0.092)
FTA	0.759***	0.394***	0.838***	0.577***
	(0.17)	(0.13)	(0.17)	(0.14)
Colonial	1.299***	1.008***	1.232***	1.071***
	(0.098)	(0.095)	(0.15)	(0.15)
Colonial*North			-0.0325	0.648
			(0.19)	(0.41)
Legal	0.486***	0.390***	0.456***	0.375***
	(0.045)	(0.044)	(0.055)	(0.054)
Legal*North			0.163*	0.104
			(0.084)	(0.081)
Border	0.458***	0.827***	0.554***	0.737***
	(0.12)	(0.12)	(0.15)	(0.14)
Border*North			-0.501*	0.333
			(0.26)	(0.42)
Island	-0.391***	-0.175*	-0.315***	-0.237**
	(0.11)	(0.11)	(0.12)	(0.11)
Island*North			-0.213**	0.127
			(0.096)	(0.095)
Landlocked	-0.561***	-0.449***	-0.499***	-0.456***
	(0.16)	(0.16)	(0.17)	(0.16)
Landlocked*North			-0.168	-0.0526
			(0.10)	(0.10)
Currency Union	1.364***	1.030***	1.258***	1.071***
	(0.25)	(0.24)	(0.27)	(0.27)
δ (from $\widehat{\overline{w}}_{rij}^{*}$)		0.699***		0.481***
		(0.051)		(0.050)
$\widehat{\overline{\eta}}_{rij}^{*}$		0.406***		0.658***
		(0.059)		(0.054)
Observations	11,146	11,146	11,146	11,146
R^2	0.71		0.71	

Notes: Exporter and Importer Fixed Effects; r = North/South
Religion is excluded variable and not reported; Robust Standard Errors are in parenthesis with country pair clustering *significant at 10%; ** significant at 5%; *** significant at 1%

Table IV - Bias Decompositon With Region Controls

Year 1986 COEFFICIENT	(1) Benchmark	(2) MLE	(3) Firm Heterogeneity	(4) Heckman Selection
Distance	-1.255***	-0.992***	-1.056***	-1.318***
	(0.034)	(0.038)	(0.043)	(0.032)
Distance * North	0.279***	0.207***	0.165***	0.343***
	(0.051)	(0.049)	(0.052)	(0.056)
Language	0.131**	0.0454	0.0378	0.174***
	(0.066)	(0.064)	(0.067)	(0.060)
Language*North	0.205**	0.0899	0.176*	0.222**
	(0.095)	(0.092)	(0.095)	(0.106)
FTA	0.838***	0.577***	0.697***	0.842***
	(0.17)	(0.14)	(0.15)	(0.161)
Colonial	1.232***	1.071***	1.122***	1.269***
	(0.15)	(0.15)	(0.15)	(0.187)
Colonial*North	-0.0325	0.648	2.678***	-0.076
	(0.19)	(0.41)	(0.49)	(-0.29)
Legal	0.456***	0.375***	0.420***	0.452***
	(0.055)	(0.054)	(0.055)	(0.049)
Legal*North	0.163*	0.104	0.130	0.188**
	(0.084)	(0.081)	(0.084)	(0.092)
Border	0.554***	0.737***	0.677***	0.510***
	(0.15)	(0.14)	(0.15)	(0.132)
Border*North	-0.501*	0.333	2.153***	-0.483*
	(0.26)	(0.42)	(0.54)	(0.273)
Island	-0.315***	-0.237**	-0.230*	-0.364***
	(0.12)	(0.11)	(0.12)	(0.112)
Island*North	-0.213**	0.127	-0.0517	-0.208**
	(0.096)	(0.095)	(0.099)	(0.103)
Landlocked	-0.499***	-0.456***	-0.458***	-0.519***
	(0.17)	(0.16)	(0.17)	(0.166)
Landlocked*North	-0.168	-0.0526	-0.145	-0.133
	(0.10)	(0.10)	(0.10)	(0.120)
δ (from $\widehat{\overline{w}}^*_{rij}$)		0.481***		
		(0.050)		
$\widehat{\overline{\eta}}^*_{rij}$		0.658***		0.341***
		(0.054)		(0.059)
$\widehat{\overline{z}}^*_{rij}$			0.303***	
			(0.043)	
Observations	11,146 (24,806)	11,146 (24,806)	11,146 (24,806)	11,146 (24,806)
R^2	0.71		0.71	

Notes: Exporter and Importer Fixed Effects; r = North/South;Currency Union not reported
Religion is excluded variable; Robust standard errors with country pair clustering and total # of underlying observations are in parenthesis; *significant at 10%; ** significant at 5%; *** significant at 1%

Table V(A) - Bias Decomposition With Exporter-Importer Regions:N-N (Year=1986)

Variables	Benchmark	MLE	Firm Heterogeniety	Heckman Selection
Distance	-1.083***	-1.086***	-1.086***	-1.086***
	(0.074)	(0.086)	(0.074)	(0.074)
Language	-0.0151	0.001	0.010	0.010
	(0.14)	(0.19)	(0.14)	(0.14)
FTA	-0.104	-0.133	-0.133	-0.133
	(0.12)	(0.21)	(0.12)	(0.12)
Colonial	0.517*	0.499	0.499*	0.499*
	(0.28)	(0.34)	(0.28)	(0.28)
Religion	-0.0831			
	(0.29)			
Legal	0.678***	0.686***	0.686***	0.686***
	(0.10)	(0.13)	(0.10)	(0.10)
Border	0.0508	0.0394	0.039	0.0394
	(0.19)	(0.23)	(0.19)	(0.19)
Island	-0.133	-0.130	-0.130	-0.130
	(0.43)	(0.37)	(0.42)	(0.42)
Land Locked	-1.093***	-1.111***	-1.111***	-1.111***
	(0.33)	(0.38)	(0.34)	(0.34)
δ(from $\widehat{\overline{w}}^*_{rij}$)		1.218		
		(9.96)		
$\widehat{\overline{z}}^*_{rij}$			0.027	
			(0.029)	
$\widehat{\overline{\eta}}^*_{rij}$		1.378		-0.052
		(7.89)		(0.057)
Observations	643(649)	643(649)	643(649)	643(649)
R^2	0.88		0.88	

Notes: Exporter and Importer Fixed Effects; r = North/South;Currency Union not reported
Religion is excluded variable; Robust standard errors with country pair clustering and total # of underlying observations are in parenthesis; *significant at 10%; ** significant at 5%; *** significant at 1%

Table V(B) - Bias Decomposition With Exporter-Importer Regions:N-S (Year=1986)

Variables	Benchmark	MLE	Firm Heterogeneity	Heckman Selection
Distance	-1.351***	-1.272***	-1.368***	-1.364***
	(0.064)	(0.077)	(0.065)	(0.056)
Language	0.0997	0.0675	0.132	0.128
	(0.089)	(0.12)	(0.089)	(0.087)
FTA	1.529***	0.714	1.471***	1.513***
	(0.45)	(0.55)	(0.46)	(0.370)
Colonial	1.289***	0.270	1.205***	1.296***
	(0.14)	(0.40)	(0.28)	(0.154)
Religion	0.385**			
	(0.17)			
Legal	0.650***	0.554***	0.660***	0.658
	(0.074)	(0.094)	(0.075)	(0.067)
Border	0.423	-0.302	0.400	0.443
	(0.35)	(0.44)	(0.38)	(0.285)
Island	-0.504***	-0.261	-0.503***	-0.493***
	(0.17)	(0.23)	(0.17)	(0.169)
Land Locked	-0.560***	-0.459*	-0.553***	-0.549***
	(0.21)	(0.27)	(0.21)	(0.195)
δ(from $\widehat{\overline{w}}^*_{rij}$)		0.0005		
		(0.0021)		
$\widehat{\overline{z}}^*_{rij}$			-0.011	
			(0.029)	
$\widehat{\overline{\eta}}^*_{rij}$		0.797***		-0.017
		(0.073)		(0.140)
Observations	2922(3432)	2922(3432)	2922(3432)	2922(3432)
R^2	0.79		0.79	

Notes: Exporter and Importer Fixed Effects; r = North/South;Currency Union not reported
Religion is excluded variable; Robust standard errors with country pair clustering and total # of underlying observations are in parenthesis; *significant at 10%; ** significant at 5%; *** significant at 1%

Table V(C) - Bias Decomposition With Exporter-Importer Regions:S-N (Year=1986)

Variables	Benchmark	MLE	Firm Heterogeneity	Heckman Selection
Distance	-1.351***	-1.108***	-1.239***	-1.238***
	(0.064)	(0.11)	(0.084)	(0.076)
Language	0.0997	0.116	0.172	0.187
	(0.089)	(0.16)	(0.12)	(0.117)
FTA	1.529***	0.181	0.802*	1.048**
	(0.45)	(0.71)	(0.44)	(0.482)
Colonial	1.289***	-0.200	0.623*	1.054***
	(0.14)	(0.50)	(0.32)	(0.199)
Religion	0.385**			
	(0.17)			
Legal	0.650***	0.668***	0.792***	0.790***
	(0.074)	(0.13)	(0.094)	(0.091)
Border	0.423	-0.148	0.314	0.579
	(0.35)	(0.57)	(0.38)	(0.364)
Island	-0.504***	-0.0826	-0.239	-0.222
	(0.17)	(0.32)	(0.23)	(0.230)
Land Locked	-0.560***	-0.624	-0.648**	-0.621**
	(0.21)	(0.39)	(0.32)	(0.277)
δ(from $\widehat{\overline{w}}^*_{rij}$)		0.052		
		(0.22)		
$\widehat{\overline{z}}^*_{rij}$			-0.065*	
			(0.039)	
$\widehat{\overline{\eta}}^*_{rij}$		0.787***		0.273*
		(0.10)		(0.160)
Observations	2529 (3432)	2529 (3432)	2529 (3432)	2529 (3432)
R^2	0.73		0.73	

Notes: Exporter and Importer Fixed Effects; r = North/South;Currency Union not reported
Religion is excluded variable; Robust standard errors with country pair clustering and total # of underlying observations are in parenthesis; *significant at 10%; ** significant at 5%; *** significant at 1%

Table V(D) - Bias Decomposition With Exporter-Importer Regions:S-S (Year=1986)

Variables	Benchmark	MLE	Firm Heterogeneity	Heckman Selection
Distance	-1.219***	-1.086***	-0.983***	-1.415***
	(0.042)	(0.19)	(0.16)	(0.052)
Language	0.301***	0.273*	0.191*	0.427***
	(0.087)	(0.14)	(0.11)	(0.087)
FTA	2.006***	1.540**	1.427***	2.431***
	(0.31)	(0.60)	(0.51)	(0.309)
Colonial	0.972**	1.306	1.412***	0.664
	(0.38)	(1.60)	(0.49)	(1.128)
Religion	-0.0379			
	(0.14)			
Legal	0.246***	0.201**	0.231***	0.223***
	(0.071)	(0.095)	(0.071)	(0.068)
Border	0.601***	0.664***	0.739***	0.475***
	(0.16)	(0.25)	(0.19)	(0.164)
Island	-0.216	-0.207	-0.160	-0.234
	(0.16)	(0.23)	(0.17)	(0.162)
Land Locked	-0.710	-0.651	-0.752*	-0.643
	(0.44)	(0.60)	(0.44)	(0.420)
δ(from $\widehat{\overline{w}}^*_{rij}$)		0.222		
		(0.42)		
$\widehat{\overline{z}}^*_{rij}$			0.331	
			(0.22)	
$\widehat{\overline{\eta}}^*_{rij}$		0.781***		0.730***
		(0.29)		(0.115)
Observations	5052 (17292)	5062 (17292)	5052 (17292)	5052 (17292)
R^2	0.60		0.60	

Notes: Exporter and Importer Fixed Effects; r = North/South;Currency Union not reported
Religion is excluded variable; Robust standard errors with country pair clustering and total # of underlying observations are in parenthesis; *significant at 10%; ** significant at 5%; *** significant at 1%

Table VI - Sensitivity of the Gravity Estimates with Region Controls

	(1)	(2)	(3)	(4)
Year	1972	1972	1996	1996
COEFFICIENT	Benchmark	MLE	Benchmark	MLE
Distance	-1.293***	-1.034***	-1.358***	-1.054***
	(0.039)	(0.045)	(0.031)	(0.035)
Distance * North	0.357***	0.302***	0.238***	0.089**
	(0.057)	(0.057)	(0.045)	(0.043)
Language	0.278***	0.150*	0.244***	0.096*
	(0.080)	(0.079)	(0.058)	(0.057)
Language*North	0.186*	0.155	0.142	0.098
	(0.11)	(0.11)	(0.087)	(0.083)
FTA	0.359	0.737**	0.686***	0.271***
	(0.38)	(0.33)	(0.13)	(0.10)
Colonial	1.557***	1.505***	1.091***	0.528***
	(0.19)	(0.19)	(0.15)	(0.15)
Colonial*North	-0.316	-0.926***	-0.0003	0.022
	(0.23)	(0.23)	(0.20)	(0.20)
Legal	0.369***	0.318***	0.295***	0.237***
	(0.066)	(0.064)	(0.047)	(0.046)
Legal*North	0.309***	0.355***	0.0136	0.038
	(0.095)	(0.093)	(0.072)	(0.069)
Border	0.489***	0.499***	0.631***	0.773***
	(0.15)	(0.15)	(0.13)	(0.13)
Border*North	-0.553**	-0.344	-0.405	2.917***
	(0.25)	(0.23)	(0.27)	(0.52)
Island	-0.290**	-0.186	-0.456***	-0.322***
	(0.15)	(0.14)	(0.11)	(0.11)
Island*North	-0.320***	-0.202*	0.0562	0.251***
	(0.11)	(0.11)	(0.086)	(0.083)
Landlocked	-0.566***	-0.477**	-0.699***	-0.633***
	(0.21)	(0.21)	(0.14)	(0.14)
Landlocked*North	-0.308**	-0.376***	0.001	0.178**
	(0.13)	(0.13)	(0.089)	(0.087)
Currency Union	1.879***	1.662***	1.311***	0.663***
	(0.23)	(0.23)	(0.25)	(0.25)
δ (from $\widehat{\overline{w}}^*_{rij}$)		0.242***		0.761***
		(0.063)		(0.049)
$\widehat{\overline{\eta}}^*_{rij}$		0.465***		0.452***
		(0.079)		(0.065)
Observations	9,711	9,711	12,795	12,795
R^2	0.66		0.76	

Notes: Exporter and Importer Fixed Effects; r = North/South
Religion is excluded variable and not reported; Robust Standard Errors are in parenthesis with country pair clustering *significant at 10%; ** significant at 5%; *** significant at 1%

Table A1 - List of the OECD (Northern) Countries

Country	Year of Accession	Country	Year of Accession
AUSTRALIA	1971	KOREA	1996
AUSTRIA	1961	MEXICO	1994
BELGIUM-LUX	1961	NETHERLANDS	1961
CANADA	1961	NEW ZEALAND	1973
CZECH REPUBLIC	1995	NORWAY	1961
DENMARK	1961	POLAND	1996
FINLAND	1969	PORTUGAL	1961
FRANCE	1961	SLOVAK REPUBLIC	2000
GERMANY	1961	SPAIN	1961
GREECE	1961	SWEDEN	1961
HUNGARY	1996	SWITZERLAND	1961
ICELAND	1961	TURKEY	1961
IRELAND	1961	UNITED KINGDOM	1961
ITALY	1962	UNITED STATES	1961
JAPAN	1964		

Source: Organisation for Economic Co-operation and Development (OECD), www.oecd.org

Table A2 - List of the Developing/Emerging (Southern) Countries

AFGHANISTAN	COTE D IVOIRE	ISRAEL	PAKISTAN	UNTD ARAB EM
ALBANIA	CUBA	JAMAICA	PANAMA	UNTD RP TNZ
ALGERIA	CYPRUS	JORDAN	PAPUA N.GUIN	URUGUYA
ANGOLA	DJIBOUTI	KENYA	PARAGUAY	VENEZUELA
ARGENTINA	DOMINICAN RP	KIRIBATI	PERU	VIETNAM
BAHAMAS	ECUADOR	KOREA DPR	PHILLIPINES	WESTERN SAHA
BAHRAIN	EGYPT	KUWAIT	QATAR	YEMEN
BANGLADESH	EL SALVADOR	LAOS	REUNION	ZAIRE
BARBADOS	EQ. GUINEA	LEBANON	ROMANIA	ZAMBIA
BELIZE	ETHIOPIA	LIBERIA	RWANDA	ZIMBABWE
BENIN	FIJI	LIBYA ARAB	SAUDI ARABIA	
BERMUDA	FM USSR	MADAGASCAR	SENEGAL	
BHUTAN	FM YUGOSLAVI	MALAWI	SEYCHELLES	
BOLIVIA	FRENCH GUIAN	MALAYSIA	SIERRA LEONE	
BRAZIL	GABON	MALDIVES	SINGAPORE	
BRUNEI	GAMBIA	MALI	SOLOMON ISLD	
BULGARIA	GHANA	MALTA	SOMALIA	
BURKINA FASO	GREENLAND	MAURITANIA	SOUTH AFRICA	
BURUNDI	GUADELOUPE	MAURITIUS	SRI LANKA	
CAMBODIA	GUATEMALA	MONGOLIA	ST KITTS NEV	
CAMEROON	GUINEA	MOROCCO	SUDAN	
CAYMAN ISLDS	GUINEA-BISSA	MOZAMBIQUE	SURINAM	
CENTRAL AFR.	GUYANA	MYANMAR	SYRN ARAB RP	
CHAD	HAITI	NEPAL	TAIWAN	
CHILE	HONDURAS	NETH ANTILLE	THAILAND	
CHINA	HONG KONG	NEW CALEDONI	TOGO	
COLOMBIA	INDIA	NICARAGUA	TRINIDAD-TOB	
COMOROS	INDONESIA	NIGER	TUNISIA	
CONGO	IRAN	NIGERIA	TURKS CAICOS	
COSTA RICA	IRAQ	OMAN	UGANDA	

Source: The HMR Data Set

Figure A – Theoretical Determination of the Bilateral Export Volumes in the HMR Model

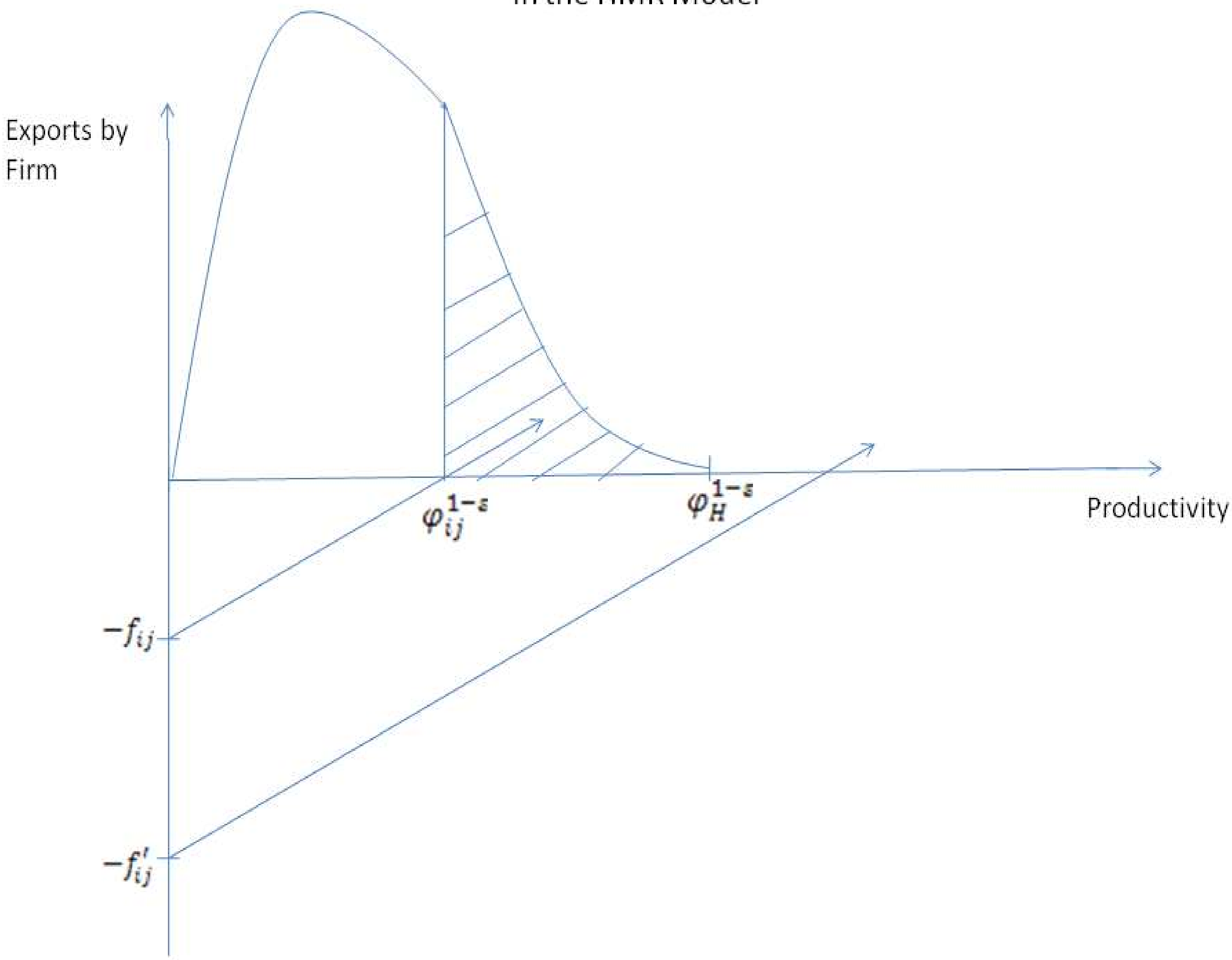


Figure 1 - Distribution of Country Pairs Based on Direction of Trade and Destination Markets - OECD(28 Countries) to World

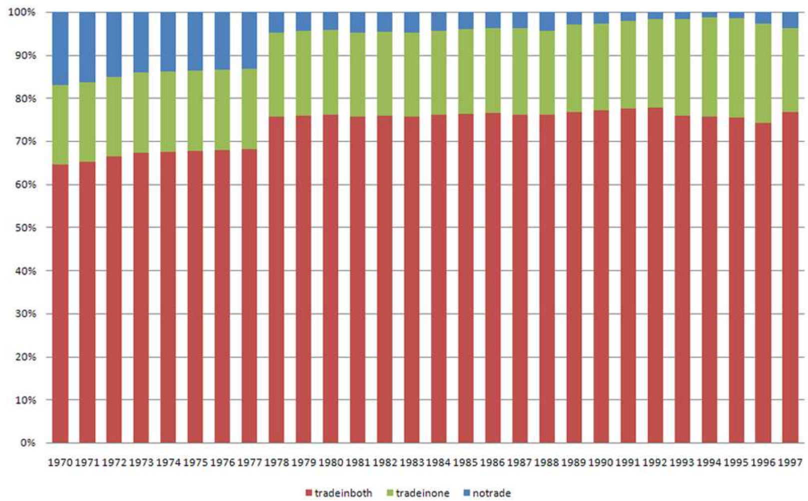
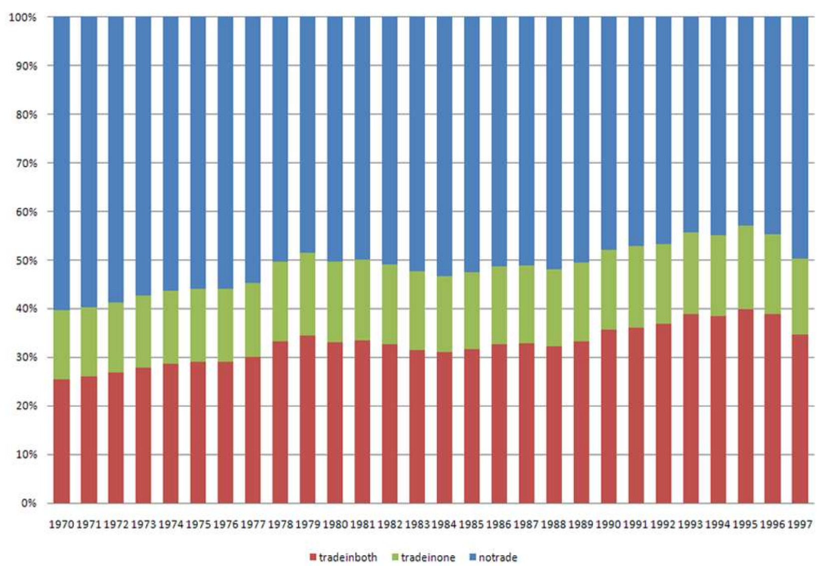


Figure 2 - Distribution of Country Pairs Based on Direction of Trade and Destination Markets - Non-OECD(130 Countries) to World



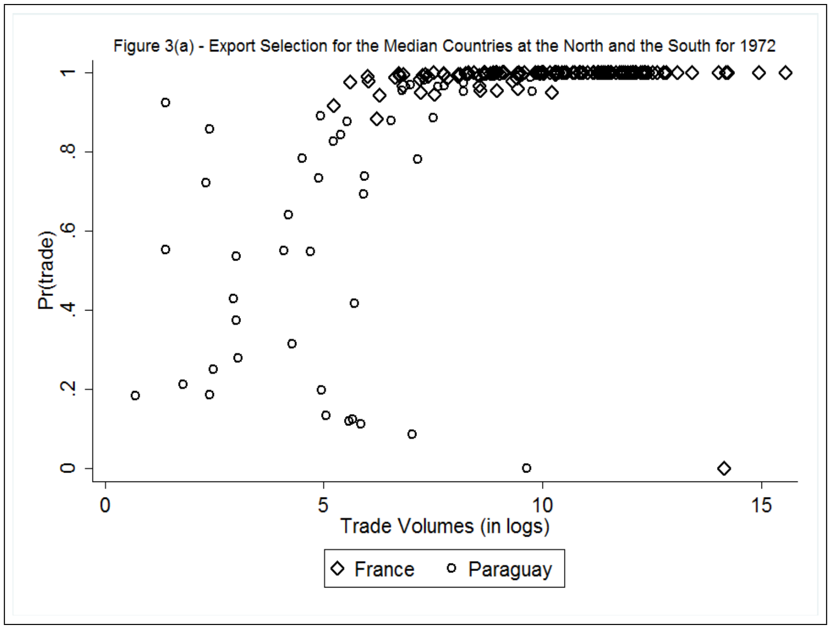


Figure 3(a) - Export Selection for the Median Countries at the North and the South for 1972
Pr(trade)
0
.2
.4
.6
.8
1
0
5
10
15
Trade Volumes (in logs)
France
Paraguay

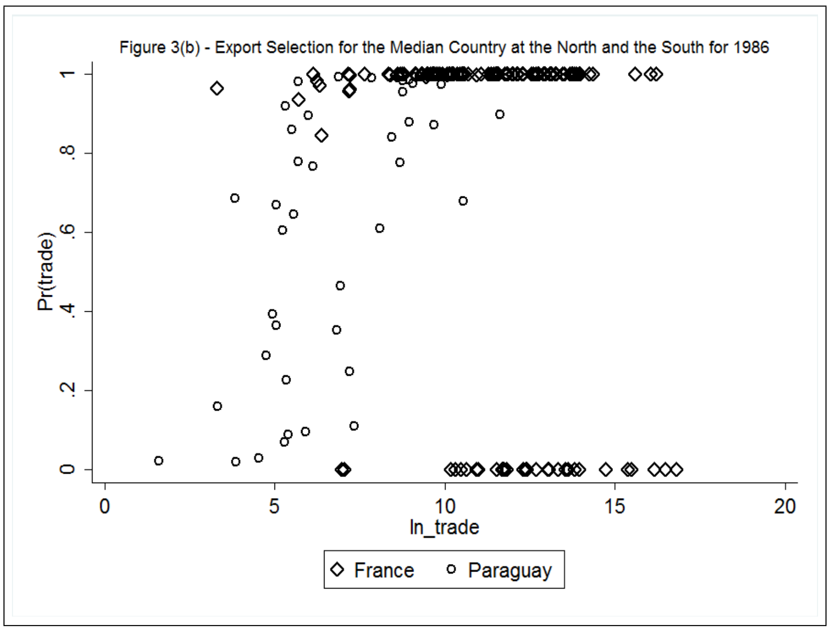


Figure 3(b) - Export Selection for the Median Country at the North and the South for 1986
Pr(trade)
0
.2
.4
.6
.8
1
0
5
10
15
20
ln_trade
France
Paraguay

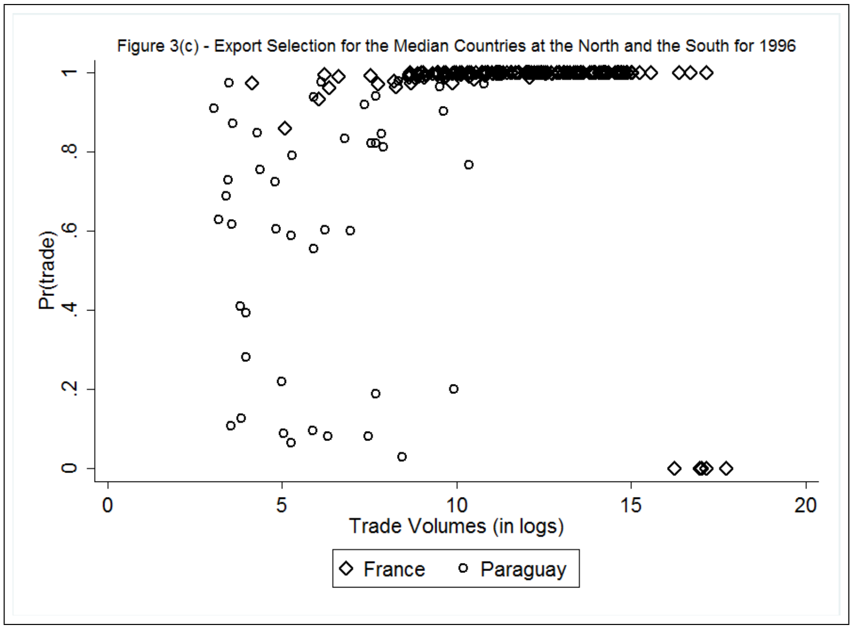


Figure 3(c) - Export Selection for the Median Countries at the North and the South for 1996
Pr(trade)
0
.2
.4
.6
.8
1
0
5
10
15
20
Trade Volumes (in logs)
France
Paraguay