Theoretically-Consistent Parameterization of a Multi-sector Global Model with Heterogeneous Firms

by

Zeynep Akgul[1], Nelson Villoria, Thomas Hertel

April 2015

Abstract

Parameter selection in Computable General Equilibrium (CGE) models of international trade is important for trade policy analysis as parameter values affect trade and welfare responses to changes in trade barriers. This is particularly important if the CGE model incorporates firm heterogeneity because it introduces a new margin of adjustment and changes the interpretation of parameter estimates. A remaining obstacle to fully understanding the insights of trade in CGE models with firm heterogeneity is the lack of an appropriate set of parameter estimates consistent with the underlying theory. Our objective in this paper is to solve for structural parameters that are theoretically consistent with firm heterogeneity models. Specifically, we focus on the elasticity of substitution across varieties. We distinguish between the extensive and intensive margins of trade flows and estimate two gravity equations by using country-level data for the motor vehicles and parts industry. Our results show that the elasticity values that are consistent with the firm heterogeneity theory are considerably lower than Armington (1969) elasticities used in the GTAP model. This implies that current implementations of Melitz-type models which use elasticities of substitution estimated in the absence of firm heterogeneity will give overly large trade volume responses to policy reforms.

Keywords: Firm Heterogeneity, Melitz, Armington, Pareto Distribution, Elasticity of Substitution, Shape Parameter, Extensive Margin, Intensive Margin, Gravity Equation

JEL codes: F10, F12, F14

[1] Akgul is a Ph.D. candidate with the Department of Agricultural Economics, Villoria is a Research Assistant Professor of Agricultural Economics with the Center for Global Trade Analysis, Hertel is a Professor of Agricultural Economics and the Executive Director of the Center for Global Trade Analysis, Purdue University, West Lafayette, IN 47907, USA.

1. Introduction

Theoretical and empirical developments in the trade literature show that accounting for firm heterogeneity within an industry improves our understanding of how trade barriers affect trade flows and economic welfare by providing a new margin of adjustment through self-selection of firms into and out of markets. Due to this added explanatory power, firm heterogeneity theory has begun to be incorporated into computable general equilibrium models (CGE) (Akgul, et al. 2014, Balistreri, et al. 2011, Balistreri and Rutherford 2012, Dixon, et al. 2015, Zhai 2008). However, a remaining obstacle to fully harvesting the insights of trade in CGE models with firm heterogeneity is the lack of an appropriate set of estimates of the elasticities of substitution and productivity distribution parameters consistent with the underlying theory. Pinning down the structural parameters is paramount for policy analysis, as quantitative results heavily depend on parameter values. Our objective in this paper is to estimate elasticity of substitution that is consistent with existing shape parameters of Pareto distribution at the GTAP level of industry definition.

Chaney (2008) extends the seminal work of Melitz (2003) and demonstrates that in models with heterogeneous firms, changes in trade barriers affect both the volume of sales by existing exporters (i.e., the intensive margin of trade) as well as the number of firms in the export market (i.e., the extensive margin of trade) due to productivity differences across firms. An important finding in the literature is that the extensive margin is quantitatively very important in governing growth in international trade flows (Hummels and Klenow 2005, Yi 2003). As a result, estimates of the elasticity of substitution by models that ignore changes in the extensive margin are biased (Chaney 2008, Helpman, et al. 2008). This finding contrasts with the traditional Armington (1969) view of the world, whereby changes in trade barriers only affect the intensive margin of trade, which is governed by the elasticity of substitution across varieties, σ (Hillberry and Hummels 2013). However, in firm heterogeneity models there is an additional parameter of interest, namely the shape parameter of Pareto distribution, γ. The shape parameter is an inverse measure of heterogeneity in productivity across firms within an industry and it governs the supply-side effects of trade policies. In fact, the distribution of firm productivity significantly affects aggregate trade response to reduced trade costs as demonstrated in Chaney (2008), Bernard, et al. (2003) and di Giovanni and Levchenko (2013). Therefore, to work with a firm heterogeneity model, we need to have estimates of the shape parameter as well as the elasticity of substitution amongst varieties.

Empirical studies of international trade flows rely on gravity equations in order to estimate the structural parameters of trade models. Gravity models relate the volume of bilateral trade to distance and other determinants of trade. In a gravity model, the marginal effect of distance on trade volumes is given by $-\delta(\sigma-1)$, where δ is the distance elasticity of trade. Identification of $-\delta(\sigma-1)$ requires knowledge on either δ or σ. However, bringing in an additional parameter to reflect firm heterogeneity, i.e. γ, introduces further complexities in identifying the elasticity of substitution. Crozet and Koenig (2010) show that, in the firm-heterogeneity setting of Chaney (2008), the marginal effect of distance on the *probability of a bilateral trade flow taking place* is given by $-\delta\gamma$. Therefore, there are three parameters to estimate, i.e. $\delta, \sigma,$ and γ, which implies that an exogenous source of information is needed to identify all of them.

di Giovanni and Levchenko (2013) and Eaton, et al. (2011) circumvent this difficulty by imposing the prior values of σ on the model in order to calibrate the values of γ. This method has two drawbacks: (i) Often, estimates for σ are obtained from Armington-type models which are fundamentally inconsistent with firm heterogeneity theory. (ii) The resulting values for γ typically are not sector and region-specific and therefore do not capture the significant variation of heterogeneity along these dimensions. For example, the shape parameter estimates in Spearot (2015) show that electrical machinery is a more heterogeneous industry where productivity differences across firms is more pronounced, while petroleum refining is a much more homogeneous industry. Moreover, according to his estimates, even though electrical machinery is heterogeneous in the US, it is much more homogeneous in Chile. Not accounting for these drawbacks is likely to lead to biased estimates of parameters in the calibrated model.

A theory-consistent approach to estimating the shape parameter is offered by Crozet and Koenig (2010) and Spearot (2015). Both studies present estimates of γ at the product level in a firm heterogeneity model. The model in Crozet and Koenig (2010) is based on Chaney (2008), while the model in Spearot (2015) is based on Melitz and Ottaviano (2008). Even though Spearot

(2015) provides values for γ by industry and by region, he does not estimate elasticities that are consistent with γ. Only Crozet and Koenig (2010) have a rich enough dataset to identify both parameters. Interestingly, their estimates of the elasticity of substitution are lower when compared to the traditional Armington elasticity estimates in the GTAP model (Hertel, et al. 2003). Unfortunately, their estimates are of limited use for a global general equilibrium model because they are based only on French firms and cover a limited number of industries. Against this backdrop, our objective in this paper is to estimate a set of elasticities of substitution that are theoretically consistent with trade models considering firm heterogeneity.

To accomplish this, we extend the seminal work of Melitz (2003) to a multi-sector, multi-country model and build on Chaney (2008) to distinguish the intensive and extensive margins of trade. For our gravity estimations, we use bilateral trade data at the country level from GTAP Version 8.1 (Narayanan, et al. 2012) which covers the years 1995-2009. This makes sense, since our ultimate goal is to incorporate these parameters in a model based on the GTAP data set. In addition, we use the GeoDist and Gravity databases of CEPII (Mayer and Zignago 2011) which include bilateral data on several relevant variables such as distance, language, colonial link among others determinant of bilateral trade. The resulting dataset covers 113 countries over 1995-2006. In this paper, we focus on the motor vehicles and parts industry (MVH) of GTAP which, according to Spearot (2015)'s parameters, has one of the highest productivity dispersions across firms among manufacturing industries. Future research will extend this work to all of the GTAP sectors and regions.

Our estimation strategy merges the approach adopted by Helpman, et al. (2008) with the extensive margin specification used in Crozet and Koenig (2010). We distinguish between the intensive and extensive margins of trade which results in two estimating equations. The first equation estimates the probability of a bilateral trade taking place, while the second equation estimates the value of bilateral trade conditional on the choice to export. We refer to the first equation as the export participation equation and refer to the second equation as the gravity equation[2]. Following Crozet and Koenig (2010) in both equations we focus on the coefficient of

[2] In principle, both equations are gravity equations. However, we adopt this convention to distinguish the new margin of adjustment due to firm entry/exit from the traditional gravity equation that determines trade flows.

distance. In the export participation equation the distance coefficient is a combination of the distance and substitution elasticities. On the other hand, in the gravity equation, the distance coefficient is a combination of the shape parameter and substitution elasticity. This gives us two equations in three unknowns, whereupon we use the shape parameter estimates provided in Spearot (2015) to infer the theoretically consistent estimates of substitution elasticities.

Our estimation results show that the elasticity estimate consistent with firm heterogeneity for the motor vehicles and parts industry is considerably lower than the GTAP Armington elasticity (Hertel, et al. 2003). This implies that elasticities estimated in that traditional way were in fact picking up additional effects accruing from the supply-side heterogeneity in this framework. In summary, Armington elasticities are high when employed in the context of a firm heterogeneity model because they confound demand-side effects with the supply-side effects. This finding underlines the argument in Dixon, et al. (2015) about the observational equivalence between Armington and Melitz models. In particular, they argue that welfare implications of trade policies are similar in magnitude between these models if the Armington and Melitz elasticities are chosen such that trade responses are equal across model specifications. In such a scenario, Armington-based elasticities are higher than Melitz elasticities. This implies that using Armington elasticities in a firm heterogeneity model might lead to overestimated trade volumes and welfare effects.

By combining the theory-consistent econometric estimates with the firm heterogeneity model in a CGE setting, we believe that this work will pave the way for mainstream application of firm heterogeneity models in policy analysis within the global trade analysis community.

2. Background on Structural Parameters of the Firm Heterogeneity Model

Although Melitz (2003) does not impose any restrictions on productivity, the common approach in the firm heterogeneity literature is to assume that firms draw their productivity levels from a Pareto distribution[3]. There are two main reasons for choosing the Pareto distribution. First, the

[3] Well-known examples are Arkolakis (2010), Chaney (2008), Helpman, et al. (2008), Melitz and Ottaviano (2008), Eaton, et al. (2011) and Balistreri, et al. (2011) among many others. There are some alternative productivity distribution assumptions in the literature. Examples include lognormal distribution (Head, et al. 2014) and Fréchet

Pareto distribution is analytically tractable. As Chaney (2008) argues, an important property of Pareto distribution is its stability to truncation from below. As a result of this property, exporters, which are more productive and therefore at the upper tail of the distribution, are also Pareto distributed. Moreover, the same shape parameter that governs the distribution of domestic firms also governs that of exporters[4].

The second reason for favoring the Pareto distribution over alternatives is empirical. The Pareto distribution is a power law and provides a good fit for the observed size distribution of firms. Empirical support for this distribution is found for US firms (Axtell, 2011) and French firms (Eaton et al., 2011) among many others[5]. The Pareto assumption for firm sales is equivalent to assume that firm productivity is Pareto, though with a different shape parameter. Furthermore, the Pareto distribution predicts a linear relationship between the log of rank and the log of firm size (Crozet and Koenig, 2010). An ever-expanding body of empirical studies uses this property to consistently estimate shape parameters based on firm sales data. In particular, they estimate the Power Law exponent of firm sales given by $\gamma/(\sigma-1)$ to pin down γ and σ. However, since this expression is a combination of γ and σ, it is not possible to estimate the individual structural parameters in these studies.

A key restriction on these parameter values in this context is the condition $\gamma > \sigma - 1$. This is described in Melitz (2003) and Chaney (2008) as the condition that ensures the firm size distribution has a finite mean. This is equivalent to saying $\frac{\gamma}{\sigma-1} > 1$. Therefore, the relative values of γ and σ become critical for quantitative outcomes such as export sales. Table 2 presents a summary of the structural parameters in the firm heterogeneity model.

---

distribution (Bernard, et al. 2003, Eaton and Kortum 2002). It is argued in Luttmer (2007) that a power law relationship is a better fit for firm size distribution compared to the lognormal distribution.

[4] There are new empirical findings that might challenge this proposition. di Giovanni, et al. (2011) argue that the shape parameter of firm size distribution is systematically different between exporters and non-exporters. Firm size distribution of exporters is more fat-tailed and has a lower shape parameter than non-exporters because they are more productive. This in turn implies that the Pareto shape parameter of productivity distribution is different between exporters and non-exporters given a constant elasticity of substitution for firm varieties.

[5] Size distribution of firms also follows a power law in the case of Japan (Fujiwara 2004, Okuyama, et al. 1999). See Gabaix (2008) for a full survey on power laws.

The value of the shape parameter determines price differences across firms in the industry. A small shape parameter implies a large dispersion of productivity among firms with low-productivity firms capturing a small share of the market. In this case new entrants charge higher prices compared to the existing exporters. On the other hand, in an industry with a large shape parameter, there is a big mass of low-productivity firms that represent a larger share of industry output. In this case, prices charged by new entrants are similar to the existing exporters. This supply-side heterogeneity is translated into export sales based on demand-side heterogeneity. A small elasticity of substitution means that consumers are willing to pay a premium for differentiated varieties which makes low productivity less of a disadvantage. Therefore, new entrants can capture a larger share of the market. However, a large elasticity of substitution increases the competition in the market and makes low productivity a bigger disadvantage. As a result, marginal firms capture a small share in the market. This discussion suggests that export sales by new entrants are largest when there is supply-side homogeneity (high γ) and demand-side heterogeneity (low σ) (Hillberry and Hummels 2013).

An opposite case is where $\frac{\gamma}{\sigma - 1} = 1$ which is known as the Zipf's Law. This yields a fat-tailed distribution of firm size where the infra-marginal firms in the industry are large and have a disproportionate share of overall sales compared to the small marginal firms. In that case, the welfare impact of trade is driven by infra-marginal firms rather than the marginal ones. Therefore, the contribution of the extensive margin to trade is found to be negligible (di Giovanni and Levchenko 2013)[6]. An implication of this finding is that quantitative results of trade cost reductions on trade flows and welfare are very sensitive to the firm size distribution and, by extension, very sensitive to the structural parameters of firm heterogeneity. This raises the stakes when it comes to obtaining reliable estimates of the Pareto parameters.

Even though there is a growing body of empirical work aimed at estimating structural parameters, there is still substantial uncertainty about the appropriate parameter values to use in the firm heterogeneity model. This is particularly true because of the challenges associated with

[6] This can be linked back to the discussion in Dixon, et al. (2015) about the offsetting effects of extensive margin and productivity on welfare in a tariff increase scenario.

the identification of two parameters using only one estimating equation, as mentioned above. A brief overview of parameter values used in the firm heterogeneity literature is provided in Table 3[7]. There are three key points that we can draw from this table.

First, empirical studies confirm that the Power Law exponent of firm size distribution is around 1 (Axtell 2001, di Giovanni, et al. 2011) and it is used in various studies to infer shape parameter values by relying on external sources for elasticities (di Giovanni and Levchenko 2013, Melitz and Redding 2013). Second, the shape parameter values that are calibrated using the Power Law exponent (di Giovanni and Levchenko 2013, Eaton, et al. 2011, Melitz and Redding 2013) or by other methods (Arkolakis, et al. 2008, Zhai 2008) are higher compared to the directly estimated values (Crozet and Koenig 2010, Spearot 2015). Using calibrated values of shape parameters would attribute lower productivity dispersion to the industry, while there could, in fact, be much higher productivity heterogeneity across firms. Therefore, we prefer to use the information contained in the shape parameter estimates instead of those from the calibration exercises.

Third, aggregation has a significant effect on parameter values. Estimates based on higher levels of aggregation are found to be higher than the ones based on lower levels of aggregation. This is because when we work with aggregated products, we fail to capture the variation across sectors and we settle on one parameter value to describe the entire industry. For example, in the two cases where the parameter values are estimated at a disaggregate level, for more than 30 sectors, the shape parameter estimates are found to show substantial variation in the range of 1.65-7.31 in Crozet and Koenig (2010) and 1.76-6.29 in Spearot (2015). On the other hand, estimates/calibrations that are at an aggregate industry level provide few values that are in the range of 3-7, on average (Arkolakis, et al. 2008, Balistreri, et al. 2011, Bernard, et al. 2003, Eaton and Kortum 2002, Eaton, et al. 2011, Zhai 2008). Similarly, the difference in aggregation is important for the elasticity values, as well. Elasticity estimates in Crozet and Koenig (2010) are in the range of 1.15-6.01, reflecting a wide range of demand-side heterogeneity compared to the more aggregated studies. In order to account for the variation across sectors, we prefer to

[7] This table is by no means a full review of the literature. The aim of this table is to present only a sample of the most relevant work to explore the mainstream approach in obtaining parameter estimates and to compare the values of key parameters used in these studies.

work at a disaggregated level of the manufacturing industry, focusing initially on the motor vehicles and parts.

Aggregation is extremely important in analyzing the extensive and intensive margin effects of trade flows, as well. Hillberry and Hummels (2013) argue that the extensive margin plays a larger role when one works with aggregated product lines. On the other hand, the impact of the intensive margin is more pronounced when we work with disaggregated product lines. Making this distinction is paramount in interpreting the results of any policy experiment.

3. Theoretical Model

We present a model of international trade with heterogeneous firms building on the theoretical model in Helpman et al. (2008) and Chaney (2008). We consider the world to be composed of R countries, where we index exporters by $r = 1, 2,\ldots, R$ and importers by $s = 1, 2,\ldots, R$. Every country produces and consumes differentiated as well as homogenous products. For the homogenous goods industry, we retain the traditional assumption of national product differentiation (Armington 1969) and the industry is characterized by perfect competition with constant returns to scale technology. On the other hand, we follow Melitz (2003) and assume that there are H differentiated industries indexed by $h = 1, 2, \ldots, H$. Each industry is composed of a continuum of firms where each firm produces a unique variety indexed by ω. Moreover, firms differ in their productivity levels and operate under monopolistic competition.

3.1.Consumers

We adopt a Dixit-Stiglitz treatment in the demand-side. In this setting, consumers are characterized by love-of-variety where they perceive each variety as a unique product and derive utility from that uniqueness. The utility function for the differentiated good h in country s, U_{hs}, is given by

$$U_{hs} = \left[\sum_{r} \int_{\omega_{hrs} \in \Omega_{hr}} q_{hrs} \left(\omega_{hrs} \right)^{\frac{\sigma_h - 1}{\sigma_h}} d\omega_{hrs} \right]^{\frac{\sigma_h}{\sigma_h - 1}} , \qquad (1.1)$$

where ω_{hrs} indexes the variety of good *h* imported by country *s* from the source country *r*, Ω_{hr} is the set of all varieties of good *h* available in country *r*, $q_{hrs}\left(\omega_{hrs}\right)$ is the quantity demanded by a representative consumer in country *s* of variety ω_{hrs} of good *h* imported from country *r* and $\sigma_h > 1$ is the elasticity of substitution between the varieties of good *h*.

Let P_{hs} be the price index of good *h* in country *s*, i.e. the dual price index of the Dixit-Stiglitz composite of demand in equation (1.1), which is given by

$$P_{hs} = \left[\sum_r \int_{\omega_{hrs}\in\Omega_{hr}} p_{hrs}\left(\omega_{hrs}\right)^{1-\sigma_h} d\omega_{hrs}\right]^{\frac{1}{1-\sigma_h}} , \tag{1.2}$$

where $p_{hrs}\left(\omega_{hrs}\right)$ is the price in country *s* of variety ω_{hrs} of good *h* imported from country *r* (gross of trade costs). Based on these demand and price aggregates, we can find the demand for each variety of good h shipped from country *r* to *s* to be as follows:

$$q_{hrs}\left(\omega_{hrs}\right) = \frac{p_{hrs}\left(\omega_{hrs}\right)^{-\sigma_h}}{P_{hs}^{1-\sigma_h}} Y_{hs} , \tag{1.3}$$

where Y_{hs} is the total expenditure in country *s* on industry *h* (equal to income in the relevant industry in country *s*)[8].

In this setting, there are N_{hr} varieties of good *h* produced in the exporting country *r*. A corollary to this is that there are N_{hr} active firms in industry *h* in country r. Each firm produces a unique variety and the varieties produced by firms in the exporting country *r* are distinct from the varieties produced by firms in the importing country *s*. Each country exports only a subset of its unique varieties because only some firms find it profitable to export into a given market. As a result, exports from country r to s includes only $N_{hrs} < N_{hr}$ varieties being shipped on the r-s trade route. This means that the total number of varieties of good h available to consumers in

[8] Please note that $Y_{hs} = P_{hs}U_{hs} = \int_{\omega_{hrs}\in\Omega_{hs}} p_{hrs}\left(\omega_{hrs}\right) q_{hrs}\left(\omega_{hrs}\right) d\omega_{hrs}$.

country s is N_{hs} domestic varieties plus $\sum_r N_{hrs}$ imported varieties. The following section gives more details as to why only some firms are able to export in this framework.

3.2.Producers

Producer behavior is based on Melitz (2003). As previously mentioned there are N_{hr} active firms in the monopolistically competitive industry h of country r each producing a different variety, ω, with different productivity, φ. Firms in industry h incur variable and fixed costs of production and of exporting. There are two types of fixed costs: sunk-entry costs to produce in the domestic market and fixed export costs to enter export markets. Fixed export costs are source-destination specific and are assumed to be identical across firms on the same bilateral trade route. There are two types of variable costs: marginal cost of production and transportation costs for export shipments. We adopt the standard assumption of 'iceberg' transportation costs, in which $\tau_{hrs} > 1$ units of good h must be shipped from country r in order for one unit of good h to arrive in country s.

The only type of cost that is firm-specific in this setting is the marginal cost of production which equals c_{hr} / φ_{hr} for an active firm in industry h of country r. Here, c_{hr} is the cost of the input bundle that is used for producing one unit of output in industry h of country r and φ_{hr} is the productivity of an active firm in industry h of country r which measures the amount of output produced by one bundle of input. Given the input bundle cost, let f_{hrs} measure the number of bundles that is used by firms in industry h to cover the fixed costs of exporting from country r to country s. Then, the fixed export costs on this particular bilateral trade route equals $c_{hr} f_{hrs}$.

The profit-maximizing price in a monopolistically competitive industry is a constant markup over marginal cost. Hence the delivered price in country s of the variety produced by a firm in country r with productivity φ is given by

$$p_{hrs}(\varphi) = \frac{\sigma_h}{\sigma_h - 1} \frac{\tau_{hrs} c_{hr}}{\varphi_{hrs}} \quad , \qquad (1.4)$$

where $\frac{\sigma_h}{\sigma_h - 1}$ is the markup that decreases with a larger elasticity of demand. If preferences are more homogeneous (large σ_h), the industry becomes more competitive and firms have to charge a lower markup for their respective varieties. Using the profit maximizing prices in equation (1.4) and utility maximizing level of sales in equation (1.3), the profit from exporting $q_{hrs}(\varphi)$ units of good h into country s is found to be

$$\pi_{hrs}(\varphi) = \frac{p_{hrs}(\varphi) q_{hrs}(\varphi)}{\sigma_h} - c_{hr} f_{hrs} = \left[\frac{\sigma_h}{\sigma_h - 1} \frac{\tau_{hrs} c_{hr}}{\varphi_{hrs} P_{hs}} \right]^{1-\sigma_h} Y_{hs} - c_{hr} f_{hrs} . \tag{1.5}$$

Firm export participation is determined by the potential profit to be made in each bilateral market based on equation (1.5). Firm profit increases with market size in the destination country (Y_{hs}), lower marginal costs (c_{hr} / φ_{hr}), and lower barriers to trade (τ_{hrs} and f_{hrs}). Productivity level of the firm plays a key role in determining the potential profit to be made on a particular trade route based on fixed costs associated with exporting. Particularly, destination-specific fixed export costs limit the number of exporters from source country r since only the firms with high productivity levels can cover fixed export costs and make positive profits in the export market. The cutoff productivity level of exporting is destination-specific and is determined by the zero profit condition on each bilateral trade route. The revenue made by the marginal exporting firm is just enough to cover total costs of exporting and determines the productivity threshold. Let the productivity threshold for firms in industry h to export from country r to s be φ^*_{hrs}, which is governed by the following equation

$$\varphi^*_{hrs} = \frac{\sigma_h}{\sigma_h - 1} \frac{\tau_{hrs} c_{hr}}{P_{hs}} \left[\frac{c_{hr} f_{hrs}}{Y_{hs}} \right]^{\frac{1}{\sigma_h - 1}} . \tag{1.6}$$

Firms that have a higher productivity level than φ^*_{hrs} will successfully export on the r-s route, while the rest of the firms, which have lower productivity levels than φ^*_{hrs}, will only supply the domestic market. This self-selection mechanism determines the number of firms in export markets which can differ across destinations. As mentioned above only a subset N_{hrs} firms out of the total N_{hr} firms are able to export into country s and the mass of firms in this subset depends on the productivity distribution in the industry.

We assume that firm productivity follows the Pareto distribution with support $\left[\varphi_{\min},\infty\right)$ and shape parameter γ_h that satisfies the condition $\gamma_h > \sigma_h - 1$. The associated density function, $g\left(\varphi\right)$, and cumulative distribution function, $G\left(\varphi\right)$, are then as follows:

$$g\left(\varphi\right)=\gamma\varphi_{\min}^{\gamma}\varphi^{-\gamma-1}, \qquad G\left(\varphi\right)=1-\left(\varphi_{\min}/\varphi\right)^{\gamma} , \tag{1.7}$$

where $\varphi_{\min}\in\left[1,\infty\right)$ is assumed in this paper.[9] Given the productivity distribution, $1-G\left(\varphi_{hrs}^{*}\right)$ measures the proportion of firms that have productivity levels higher than the threshold φ_{hrs}^{*}. Therefore, the fraction of active exporters to all firms in the industry N_{hrs}/N_{hr} equals $1-G\left(\varphi_{hrs}^{*}\right)$.[10]

3.3.Aggregate Trade Flows

The value of aggregate trade flows is the product of number of firms that sell in the destination market and the average revenue along the bilateral trade route. Let M_{hrs} be the total value of demand in destination country s for good h sourced in country r which is given by

$$M_{hrs}=\int_{\varphi_{hrs}^{*}}^{\infty} N_{hrs}\, p_{hrs}\left(\varphi\right) q_{hrs}\left(\varphi\right)\mu\left(\varphi\right)d\varphi, \tag{1.8}$$

where $\mu\left(\varphi\right)$ is the productivity distribution of successful firms in equilibrium, i.e. conditional distribution of $g\left(\varphi\right)$ on support $\left[\varphi_{hrs}^{*},\infty\right)$ as in Melitz (2003):

[9] Helpman et al. (2008) uses a truncated Pareto distribution by imposing upper and lower bounds to productivity. The reason for these bounds is to construct a model that can explain zero trade flows in the country level data with firm behavior. But, using a truncated Pareto distribution brings about nonlinearities into the model which we do not attempt to solve in this paper. For analytical tractability purposes we choose to impose only a lower bound for productivity. An implication of this assumption is that because there is a continuum of firms in the industry, there is a positive mass of exporters for all country pairs as noted in Head and Mayer (2014).

[10] This follows from $N_{hrs}=\int_{\varphi_{hrs}^{*}}^{\infty} N_{hr}\, g\left(\varphi\right)d\varphi$.

$$\mu(\varphi) = \begin{cases} \dfrac{g(\varphi)}{1 - G(\varphi^*)} & \text{if } \varphi \geq \varphi^* \\ 0 & \text{otherwise} \end{cases} \tag{1.9}$$

We simplify (1.8) by using optimal demand and price for good h given by equations (1.3) and (1.4). The simplified representation of bilateral trade flows is then given by

$$M_{hrs} = \begin{cases} \left[\dfrac{\sigma_h}{\sigma_h - 1} \dfrac{\tau_{hrs} c_{hr}}{P_{hs}} \right]^{1-\sigma_h} Y_{hs} N_{hr} V_{hrs} & \text{if } \varphi \geq \varphi^* \\ 0 & \text{otherwise,} \end{cases} \tag{1.10}$$

where V_{hrs} is defined as in Helpman et al. (2008) [11] :

$$V_{hrs} = \int_{\varphi_{hrs}^*}^{\infty} \varphi^{\sigma_h - 1} g(\varphi) d\varphi \tag{1.11}$$

Equation (1.10) can be thought of as a measure of the intensive margin because it takes export sales of all exporters into account in determining aggregate export sales on a particular trade route. Equation (1.10) also shows that bilateral trade flows increase with market size of the importer *s* (Y_{hs}), the mass of firms in the industry (N_{hr}), competition in the importing market (P_{hs}), reductions in barriers to trade (τ_{hrs}) and reductions in factor costs (c_{hr}). A quick look at equation (1.10) would suggest that the elasticity of trade with respect to reduced trade costs is $1 - \sigma_h$. This corresponds to the trade-cost elasticity in the Dixit-Stiglitz-Krugman monopolistic competition model. However, this is only part of the story. In fact, $1 - \sigma_h$ only represents the demand side effects of reduced trade barriers in a firm heterogeneity model. There are additional effects of trade cost reductions embedded in V_{hrs} which work through the supply side. In particular, V_{hrs} represents how self-selection of firms into export markets stimulate average

---

[11] V_{hrs} corresponds to the average productivity in the industry. In Melitz (2003), average productivity is defined as $\tilde{\varphi}_{hrs}(\varphi_{hrs}^*) = \left[\int_{\varphi_{hrs}^*}^{\infty} \varphi^{\sigma_h - 1} \mu(\varphi) d\varphi \right]^{\frac{1}{\sigma_h - 1}}$. Based on this definition, we have $V_{hrs} = \tilde{\varphi}_{hrs}^{\sigma_h - 1} \left[1 - G(\varphi_{hrs}^*) \right]$. Please note that since we define V_{hrs} as $V_{hrs} = \int_{\varphi_{hrs}^*}^{\infty} \varphi^{\sigma_h - 1} g(\varphi) d\varphi$, we can express M_{hrs} in terms of N_{hr} instead of the bilateral N_{hrs}.

productivity and thereby increase trade flows in the case of lower trade barriers. This mechanism introduces the supply side effects of trade cost changes into equation (1.10).

The combined effect of demand and supply side effects reveals that the trade-cost elasticity of trade flows in a firm heterogeneity model is different from that of a model with homogenous firms. In fact, Chaney (2008) shows analytically that trade-cost elasticity[12] is equal to the supply side parameter $-\gamma_h$ in a multi-country Melitz (2003) framework. This finding paved the way for subsequent empirical work that changed the interpretation of parameter estimates in gravity equations in the presence of heterogeneous firms.

3.4.Extensive, Intensive and Compositional Trade Margins

Many empirical studies in the gravity literature distinguish between two margins of adjustment to trade shocks: intensive and extensive margins. As trade costs fall, not only does the volume of sales from each exporter increase, i.e. intensive margin, but the set of exporters changes as well, i.e. extensive margin. As opposed to this two-way decomposition, Head and Mayer (2014) offer a three-way decomposition by arguing that an implicit margin is embedded in the conventional interpretation. Since new entrants are less productive than the existing exporters, sales of new entrants are lower than the average shipment prior to trade cost reductions. The margin of adjustment as a result of this difference in sales is referred to as the *compositional margin* by Head and Mayer (2014). The compositional margin is a part of the extensive margin in Chaney (2008) and Crozet and Koenig (2010), while it is included in the intensive margin in Bernard et al. (2007) and Hilberry and Hummels (2008). Needless to say, depending on how the compositional effects are assigned, the relative contribution of the intensive and extensive margins of trade will vary across otherwise identical studies. Therefore, it is appealing to break out this compositional effect.

Here we explicitly show the three-way decomposition of trade-cost elasticity. Trade flows in equation (1.8) can be written as the product of the number of exporters and average sales per

[12] Elasticity is defined as $\frac{\partial M_{hrs}}{\partial \tau_{hrs}} \frac{\tau_{hrs}}{M_{hrs}}$.

exporter, $M_{hrs} = N_{hrs} m_{hrs}(\tilde{\varphi})$ where average sales is defined as $m_{hrs}(\tilde{\varphi}) = \int_{\varphi^*_{hrs}}^{\infty} m(\varphi)\mu(\varphi)d\varphi$. Using the Leibniz rule, as in Chaney (2008), we obtain a decomposition of the trade-cost elasticity similar to the one in Head and Mayer (2014) as follows:

$$\frac{\partial \ln M_{hrs}}{\partial \ln \tau_{hrs}} = \underbrace{\frac{1}{m_{hrs}(\tilde{\varphi}_{hrs})}\int_{\varphi^*_{hrs}}^{\infty} \frac{\partial \ln m_{hrs}(\varphi)}{\partial \ln \tau_{hrs}} m_{hrs}(\varphi)\mu(\varphi)d\varphi}_{\text{intensive margin}} + \underbrace{\frac{\partial \ln N_{hrs}}{\partial \ln \tau_{hrs}}}_{\text{extensive margin}} + \underbrace{\left[\frac{m_{hrs}(\varphi^*_{hrs})}{m_{hrs}(\tilde{\varphi}_{hrs})} - 1\right] \frac{\partial \ln\left[1 - G(\varphi^*_{hrs})\right]}{\partial \ln \varphi^*_{hrs}} \frac{\partial \ln \varphi^*_{hrs}}{\partial \ln \tau_{hrs}}}_{\text{compositional margin}}. \tag{1.12}$$

The first component in equation (1.12) is the intensive margin which gives the adjustment in trade-cost elasticity due to changes in sales of the existing exporters. As mentioned before, the intensive margin effect is the same as in the traditional Armington model. The second component is the extensive margin, due to changes in the set of exporters. The third component is the compositional margin due to lower per firm sales of new entrants. In particular, the term $\left[\frac{m_{hrs}(\varphi^*_{hrs})}{m_{hrs}(\tilde{\varphi}_{hrs})} - 1\right]$ captures the difference between lower sales of new exporters and the average sales of the incumbents in the export market.

We follow Head and Mayer (2014) in simplifying equation (1.12) by applying the Pareto distribution and using the optimal demand and pricing equations. The resulting trade-cost elasticity of trade flows is identical to Chaney (2008).

$$\frac{\partial \ln M_{hrs}}{\partial \ln \tau_{hrs}} = \underbrace{(1-\sigma_h)}_{\text{intensive margin}} + \underbrace{(-\gamma_h)}_{\text{extensive margin}} + \underbrace{(\sigma_h - 1)}_{\text{compositional margin}} = -\gamma_h. \tag{1.13}$$

According to equation (1.13) the intensive margin depends only on the demand-side parameter σ_h and is equal to the trade-cost elasticity in a Krugman-type model with homogenous firms. Similarly, the compositional margin also depends on the demand-side parameter as sales of new entrants are also governed by the substitution elasticity. An important discussion in Head and

Mayer (2014) is that the intensive and compositional margins exactly offset each other due to the assumed Pareto distribution. This is in line with the discussion in Chaney (2008) even though his definition of the extensive margin includes the compositional part as well. He states that firm-level trade behaves in the same way as aggregate trade behaves in traditional models. As a result, the intensive and compositional margins affect the trade elasticity with the same magnitude, but in opposite direction. On the other hand, the extensive margin introduces supply-side effects through the shape parameter γ_h. In the end, what determines the final trade-cost elasticity of trade flows is the extensive margin in a firm heterogeneity model. An interesting point Chaney (2008) makes is that in firm heterogeneity models, the impact of trade barrier changes on trade flows is larger than that of the representative firm models. This due to the required condition $\gamma_h > \sigma_h - 1$ which shows that the quantitative importance of the extensive margin on trade flows is higher in a firm heterogeneity setting compared to the intensive margin effect.

In this paper we adopt the convention in Chaney (2008) and include the compositional margin within the extensive margin. Therefore, our definition of the extensive margin captures the combined effect of export sales per new exporter and the change in the set of exporters. Therefore, when we refer to the extensive margin in this paper, we refer to $-\gamma_h + \sigma_h - 1$.

4. Data

We use two data sources in this paper. Bilateral trade data comes from the Global Trade Analysis Project (GTAP) Version 8.1 (Narayanan, et al. 2012). This version includes 57 GTAP commodities and 134 GTAP regions of which 113 country titles are available. We use the time series bilateral trade data of this version that covers the period 1995 to 2009 with 2007 as the reference year. (Detailed information about data sources and variable definitions can be found in the data appendix.)

In this paper, we focus on the *motor vehicles and part*s sector, coded as *MVH* in GTAP. Therefore, we only use the trade data that is related to MVH. This choice is based on the information about the shape parameter estimates reported in Spearot (2015). Motor vehicles and parts is one of the most heterogeneous industries with respect to productivity in Spearot (2015) with a value of 1.79. On the other hand, the Armington elasticity in GTAP for motor vehicles and parts is 5.6 (Hertel, et al. 2003). A comparison of these values reveals that the condition for finite size distribution $\gamma > \sigma - 1$ is not satisfied for MVH if we stick to the Armington elasticity ($1.79 < 5.6 - 1$).

Trade barriers which are modeled as iceberg trade costs are not explicitly observed in the data. Therefore, a common approach in the gravity literature is to assume that iceberg trade cost is a function of many observable variables such as the physical distance between trading partners, sharing a common language, having a colonial relationship etc. We adopt the same approach and use the distance (***GeoDist)*** and gravity (***Gravity)*** databases of Centre d`Etudes Prospectives et d'Infiormations Internationales (CEPII) to obtain the information about gravity variables. ***GeoDist*** is CEPII's distance database developed by (Mayer and Zignago 2005) and it includes country-specific data for 225 countries and bilateral data for 224 country pairs. Further details about this database can be found in Mayer and Zignago (2011). In our paper, data on distance, contiguity, common language, colonial links and landlocked countries are obtained from *GeoDist*. In addition, we use CEPII's ***Gravity*** database based on Head, et al. (2010). This database covers an exhaustive set of variables for 224 countries for the period 1948 to 2006. In our paper, data on common legal origins, common currency, FTA and GATT/WTO membership are obtained from *Gravity*.

The time period considered in this paper is from 1995 to 2006 to match the time series of bilateral trade from GTAP and the gravity variables from CEPII. In particular, we drop the years 2007-2009 from the GTAP time series data and we drop the years 1948-1994 from the CEPII *Gravity* data. Our final dataset is obtained by merging GTAP data with CEPII data for motor vehicles and parts industry, 113 country titles and it covers the period from 1995 to 2006. The list of countries included in our dataset is presented in Table 1.

Table 4 provides descriptive statistics for the final sample of our dataset. We also tabulate the frequency of zero trade flows in the dataset by year in Table 5. Bilateral trade datasets are known to include large numbers of zeros even at the country level (Helpman, et al. 2008). Our dataset is no exception. As reported in Table 5, zero trade flows of motor vehicles and parts account for 77 per cent of the observations over the period 1995-2006. Large fraction of zeros in the dataset is known to cause sample selection bias in coefficient estimates in gravity equations where the dependent variable is log of trade flows. Since the logarithm of zero is undefined, zero observations are dropped from the sample in traditional OLS regressions. We follow the approach adopted in Helpman, et al. (2008) to control for sample selection.

Table 5 also shows that the fraction of zero trade flows diminished across years. While zero trade flows account for almost 81 per cent of the observations in 1995, this fraction reduces to 71 per cent in 2006. This reduction implies that there have been new bilateral trade routes created over the course of 12 years in motor vehicles and parts industry. We can interpret this as the reflection of extensive margin effect in the data resulting from firm entry and exit over the years.

5. Empirical Methodology

We follow the common practice of estimating the intensive and extensive margins of trade using a specification based on the gravity equation. Our empirical strategy draws on the work of Helpman et al. (2008) and Crozet and Koenig (2010). The empirical strategy in Helpman et al. (2008) is to develop a two-stage Heckman estimation procedure where they explicitly account for unobserved firm heterogeneity and sample selection bias to consistently estimate the gravity equation in a firm heterogeneity model. Similarly, we consider two equations. The first one is an export participation equation in which we estimate the effect of distance on the probability that a firm exports on the r-s route. The second one is a gravity equation in which we estimate the effect of distance on aggregate trade flows. We diverge from Helpman et al. (2008) on two fronts. First, we estimate these two equations separately, not simultaneously. Second, our latent variable definition for the first equation is different and follows Crozet and Koenig (2010).

Since we focus only on the motor vehicles and parts industry, we suppress the *h* subscript for industries in the rest of the paper. Moreover, we introduce a time subscript *t* to the variables that have different values across years.

5.1.Export Participation: Probit

The first equation we estimate is the probability of firm participation in export markets which captures entry/exit of firms, i.e. the extensive margin effect. Firm activity is not explicit in our dataset because we only observe trade flows at the country level. Helpman et al. (2008) use a latent variable in order to capture firm behavior in country level observations. Their latent variable is defined as the ratio of variable export profits to fixed export costs. According to this specification, positive trade flows are observed at the country level if and only if the latent variable is greater than one. However, this specification does not use the information implicit in the productivity distribution. As a result, the Pareto shape parameter does not appear in the export selection equation considered in Helpman et al. (2008). In this paper, we want to show the interaction between the shape parameter (γ) and the substitution elasticity (σ) which requires use of the productivity distribution. Therefore, we follow the latent variable definition in Crozet and Koenig (2010) and compare firm productivity with the productivity threshold in the export market. We now turn to the details of this approach.

A firm with productivity φ exports from country r to s if its productivity level passes the threshold level, i.e. $\varphi > \varphi^*_{rs}$. Let T_{rst} be an indicator variable where $T_{rst} = 1$ if the country r exports MVH to country s in year t, and zero otherwise. Then, the probability that a firm exports from r to s in year t is given by $\Pr\left(T_{rst} = 1\right) = \Pr\left(\varphi > \varphi^*_{rst}\right)$. When we apply the cumulative distribution function of the Pareto distribution we obtain

$$\Pr\left(T_{rst} = 1\right) = \Pr\left(\varphi > \varphi^*_{rst}\right) = \left(\varphi^*_{rst}\right)^{-\gamma}. \tag{1.14}$$

Substituting equation (1.6) into (1.14) and rearranging, we obtain the following equation for firm selection into export markets.

$$\Pr\left(T_{rst} = 1\right) = \left[\frac{\sigma}{\sigma-1}\frac{\tau_{rst}c_r}{P_s}\left(\frac{c_r f_{rst}}{Y_s}\right)^{\frac{1}{\sigma-1}}\right]^{-\gamma} = \left(\frac{\sigma}{\sigma-1}\right)^{-\gamma}\left(\frac{\tau_{rst}}{P_s}\right)^{-\gamma}\left(\frac{f_{rst}}{Y_s}\right)^{\frac{-\gamma}{\sigma-1}} c_r^{\frac{-\gamma\sigma}{\sigma-1}}. \tag{1.15}$$

We do not have information about the value of variable trade costs and bilateral fixed export costs in our dataset. Hence we follow the convention of imposing additional structure on variable and fixed costs. Variable trade costs are assumed to be a function of distance between countries and several other trade barriers as follows:

$$\tau_{rst} = D_{rs}^{\delta} \exp\left(-k\psi_{rst} - u_{rst}\right), \quad (1.16)$$

where D_{rs} is the distance between country r and s, δ is the distance elasticity of trade which is strictly positive, ψ_{rst} is a vector of trade impeding and trade facilitating variables and $u_{rst} \sim N\left(0, \sigma_u^2\right)$ captures unobserved trade costs that are i.i.d.

We follow Balistreri et al. (2011) and Helpman et al. (2008) and model fixed export costs as a combination of barriers imposed by importers only, by exporters only and by a county-pair specific bilateral cost. Let $f_{rst} \equiv \exp\left(\theta_r + \theta_s + \kappa\theta_{rs} - v_{rst}\right)$, where θ_r are fixed export costs common across destinations incurred by exporting country r, θ_s are the fixed trade barriers imposed by the importing country on all exporters, θ_{rs} are country-pair specific fixed trade barriers, and $v_{rst} \sim N\left(0, \sigma_v^2\right)$ captures unmeasured trade frictions. Helpman et al. (2008) notes that v_{rst} is i.i.d; however, they may be correlated with the unmeasured variable trade costs, u_{rst}.

Incorporating these definitions of variable and fixed export costs into equation (1.15) and taking logarithms of both sides we get the following probit equation:

$$\Pr\left(T_{rst} = 1\right) = \alpha_0 - \delta\gamma \ln D_{rs} + E_r + E_s + \alpha_4\theta_{rs} + \alpha_5\psi_{rst} + \eta_{rst} \ , \quad (1.17)$$

where $\alpha_0 = -\gamma \ln \frac{\sigma}{\sigma - 1}$, $\alpha_4 = \frac{\gamma\kappa}{1-\sigma}$, $\alpha_5 = k\gamma$, $E_r = \frac{\gamma\sigma}{1-\sigma} \ln c_r + \frac{\gamma}{1-\sigma}\theta_r$ is an exporter fixed effect which controls for the marginal cost (c_{hr}) and fixed cost (θ_{hr}) that are associated with the exporter, $E_s = \gamma \ln P_s + \frac{\gamma}{\sigma - 1} \ln Y_s + \frac{\gamma}{1-\sigma}\theta_s$ is an importer fixed effect which controls for market size and fixed costs associated with the importer and $u_{rst} + v_{rst} = \eta_{rst} \sim N\left(0, \sigma_u^2 + \sigma_v^2\right)$ is i.i.d[13]. We also add a year dummy that controls for the omitted variables which vary across years but common to all trade flows. The estimating equation, then, becomes

$$\Pr\left(T_{rst} = 1\right) = \alpha_0 - \delta\gamma \ln D_{rs} + E_r + E_s + E_t + \alpha_4 \theta_{rs} + \alpha_5 \psi_{rst} + \eta_{rst}, \tag{1.18}$$

where E_t is a year dummy. In our first step regression, we estimate the Probit equation in (1.18) for motor vehicles and parts industry. Since fixed export costs, captured by the variable θ_{rs}, only affect the probability of a bilateral trade taking place, we can use them as exclusion restrictions. We will turn to this again in the results section.

5.2. Trade Value: OLS

The second step in our empirical strategy is to estimate the value of export sales using the gravity equation. We use the aggregate sales of motor vehicles and parts from country r to country s that is governed by equation (1.8). Log linearizing equation (1.8) and using variable trade costs defined as (1.16), we obtain the following regression equation:

$$\ln M_{rst} = \lambda_0 - \delta\left(\sigma - 1\right) \ln D_{rs} + E_r + E_s + \lambda_4 \psi_{rst} + \ln \mathrm{V}_{rst} + u_{rst}, \tag{1.19}$$

where $\lambda_0 = \left(1-\sigma\right)\ln\sigma / \left(\sigma - 1\right)$, $\lambda_4 = k\left(\sigma - 1\right)$ $E_r = \left(1-\sigma\right)\ln c_r + \ln N_r$ is an exporter fixed effect which controls for the marginal cost (c_{hr}) and new varieties (N_r) that are associated with the exporter, $E_s = \left(\sigma - 1\right)\ln \mathrm{P}_s + \ln \mathrm{Y}_s$ is an importer fixed effect which controls for importer size and prices, and $u_{rst} \sim N\left(0, \sigma_u^2\right)$ is i.i.d.

Consistent estimation of Equation (1.19) requires two corrections as argued in Helpman, et al. (2008). The first correction requires adding a control variable into (1.19) for the sample selection bias. Omitting country pairs that have zero trade flows from the dataset might cause a correlation

[13] There is an implicit assumption we impose here. For simplicity, we assume that $\sigma_\eta^2 \equiv \sigma_u^2 + \sigma_v^2 = 1$. Helpman, et al. (2008) do not impose this restriction which means that all coefficient estimates in their Probit specification is normalized by σ_η .

between the unobserved u_{rst} and the explanatory variables. Therefore, we need a consistent estimate for $E\left[u_{rst} \mid ., T_{rst}=1\right]$. Following Helpman, et al. (2008) we define the consistent estimate as $E\left[u_{rst} \mid ., T_{rst}=1\right]=corr\left(u_{rst}, \eta_{rst}\right) \sigma_u \bar{\eta}_{rst}$ where $\bar{\eta}_{rst}=E\left[\eta_{rst} \mid ., T_{rst}=1\right]$. In order to be able use this in the gravity equation, we also need a consistent estimate of $\bar{\eta}_{rst}$. As is customary in the Heckman procedure, we obtain this consistent estimate from the inverse Mills ratio $\hat{\bar{\eta}}_{rst}=\frac{\phi\left(\hat{\rho}_{rst}\right)}{\Phi\left(\hat{\rho}_{rst}\right)}$, where $\hat{\rho}_{rst}$ be the predicted probability of trade between country r and s based on the estimated Probit equation in (1.18).

The second correction requires adding a control variable into (1.19) for the entry/exit of firms into export markets which is captured by the variable $\ln \mathrm{V}_{rst}$. Since firm productivity is not observed, we need a consistent estimate for $E\left[\ln \mathrm{V}_{rst} \mid ., T_{rst}=1\right]$. Here, we diverge from Helpman, et al. (2008) because our export participation is different from theirs. Instead, we use the relationship between $\ln \mathrm{V}_{rst}$ and $\ln \varphi_{rst}^{*}$ and apply the cumulative distribution function of firm productivity. The predicted value of our latent variable is $\hat{\rho}_{rst}=\left(\hat{\varphi}_{rst}^{*}\right)^{-\gamma}$. In log-linear form this is equivalent to $\ln \hat{\varphi}^{*}=\frac{1}{-\gamma} \ln \hat{\rho}_{rst}$. Using this condition and the definition in (1.11), a consistent estimate for $\ln \mathrm{V}_{rst}$ is given by the following:

$$\ln \hat{V}_{rst}=\ln \frac{\gamma}{\gamma-\sigma+1}+\frac{\gamma-\sigma+1}{\gamma} \ln \hat{\rho}_{rst} . \tag{1.20}$$

We use (1.20) to transform our gravity equation in (1.19) which gives

$$\ln M_{rst}=\beta_o-\delta(\sigma-1) \ln D_{rs}+E_r+E_s+E_t+\beta_4 \psi_{rst}+\beta_5 \ln \hat{\rho}_{rst}+\beta_6 \hat{\eta}_{rst}+\varepsilon_{rst} , \tag{1.21}$$

where $\beta_o=\lambda_0+\ln \frac{\gamma}{\gamma-\sigma+1}$, $\beta_4=\lambda_4$, $\beta_5=\frac{\gamma-\sigma+1}{\gamma}$ and $\beta_6=corr\left(u_{rst}, \eta_{rst}\right) \sigma_u$. We note that the new error term ε_{rst} satisfies the condition $E\left[\varepsilon_{rst} \mid ., T_{rst}=1\right]=0$. In our second step regression, we estimate the gravity equation in (1.21) for motor vehicles and parts.

Equation (1.17) delivers a combination of the distance elasticity and the shape parameter, $-\delta\gamma$, while equation (1.19) delivers a combination of the distance and demand elasticities, $-\delta(\sigma-1)$. However, estimates of $-\delta\gamma$ and $-\delta(\sigma-1)$ are not enough to identify three parameters separately. To circumvent this difficulty Crozet and Koenig (2010) estimate a third equation that governs the relationship between each firm's total factor productivity and its production by using firm level data. From this equation they obtain an estimate of $-\gamma+(\sigma-1)$, which facilitates the identification of three parameters in three equations. With country-level data we cannot determine the relationship between firm sales and their total factor productivity. Instead, we take the ratio of the two coefficients from the Probit and OLS equations which gives estimates of $\gamma/(\sigma-1)$. Incidentally, this fraction is the Power Law exponent of firm size distribution.

In order to solve for the elasticity of substitution in this fraction, we use the shape parameter estimates provided in Spearot (2015). This method delivers estimates of σ, which are conditional on \gamma, and therefore, consistent with the underlying firm heterogeneity theory; this estimate of \sigma is assumed to capture changes in trade flows coming from substitutability in consumption while \gamma captures the changes in trade flows taking into account the variation in productivity across industries and regions.

6. Results and Discussion

6.1.Estimation Results

Table 6 presents our estimation results. The first two columns give the regression results for a Probit model that determines the probability of firm export participation. Column (1) reports marginal effects evaluated at sample means, while estimates reported in (2) are parameter estimates. Column (3) gives our benchmark model which is a standard gravity equation estimated using ordinary least squares Column (4) reports estimation results for equation (1.21) where we include the variables X and Y which correct for sample selection as well as firm heterogeneity. All models in Table 6 include country-specific fixed effects as well as year dummies. Standard errors reported in all models are adjusted for clustering on country-pairs.

Our estimation results are in line with the gravity literature in general. In (1) and (3) distance is found to be statistically significant with an estimated coefficient around -1, which is consistent with the usual coefficient estimates in the gravity literature. Our results show that the rest of the explanatory variables are positive in both regressions. In particular, we find that the probability of exporting as well as the volume of exports increases between country-pairs when countries: (i) are closer to each other, (ii) are adjacent, (iii) are colonized by the same country, (iv) are both members of GATT/WTO, (v) are both in the same FTA, (vi) share a colonial link, or (vii) share a common legal system.

Our results show that the landlocked indicator is not significant either for export participation or for volume of exports. Moreover, the fact that two countries share the same currency is not significant for export participation, while it increases the volume of exports. To the contrary, the probability that two countries share a common language increases the probability of exporting while it is not significant for how much they trade. We attribute this to the fact that language is akin to a fixed export cost. Once the firm engages in trade, having a common language ceases to be a significant factor for trade volumes as the firm has already invested in the new language for marketing, legal work etc.

As discussed in Helpman, et al. (2008) estimating a two-stage model requires using an exclusion restriction that is correlated with the probability of export participation, but not correlated with the residuals in the second-stage gravity equation, as once a decision to export has been made, the exclusion restriction is no longer important for trade volumes. In our model, common language satisfies these requirements for a valid exclusion restriction. This is evident from the regression results in (1) and (3). As mentioned before, common language reduces fixed costs of exporting and thereby it is a significant factor in export participation, while it does not matter for trade volumes. The validity of common language as an excluded variable is also argued by Helpman, et al. (2008) who use common language as an alternative exclusion restriction and obtain similar results to the case where religion is used as an excluded variable.

Comparison of the benchmark model results in (3) with the corrected model results in (4) suggests that the coefficients for almost all trade barriers are underestimated in the benchmark model. These findings substantially differ from Helpman, et al. (2008). Their results suggest that the parameters in benchmark model are overestimated because the extensive margin effect and the country pairs that have zero trade flows are excluded in the gravity equation. In particular, they argue that ignoring sample selection introduces a downward bias, while ignoring firm heterogeneity introduces an upward bias. However, according to our results, not only sample selection control, but also firm heterogeneity control corrects for the downward bias. We attribute the difference in our results to the latent variable specification used for the Probit model. Basically, the control for firm heterogeneity captures the movements in productivity thresholds for export markets as trade barriers change. For example, a higher productivity threshold for exporting to a particular country forces low-productivity firms to exit the market which reduces the number of exporters on that bilateral route. Because of having fewer exporters in the market, aggregate trade flows for MVH declines. This is reflected as a significant and negative coefficient on the firm heterogeneity control variable reported in column (4).

While the variable correcting for firm heterogeneity is highly significant, the variable correcting for sample selection is barely significant in column (4). To further explore the effect of each correction in explaining aggregate trade flows in MVH, we estimate two more specifications each focusing on one of the corrections. Results of these regressions are reported in Table 7. In order to facilitate comparison across models we report the estimation results of the benchmark model in column (1) and the estimation results including both corrections in column (2). The results of sample selection correction are given in column (3), while the results of firm heterogeneity correction are given in column (4). We note that the coefficient estimate for sample selection is significant and enters positively contrary to (2). When we look at the firm heterogeneity correction in (4), we see that the coefficient estimates are slightly lower than (3) and similar to (2). Hence we see that firm heterogeneity correction dominates in (2) to the extent that sample selection changes sign and almost becomes insignificant for trade flows in the motor vehicles and parts industry. Overall, the coefficient of distance is robust to different specifications, which is reassuring, as this is what will give us the desired elasticities, to which we now turn[14].

[14] Silva and Tenreyro (2015) argue that the assumption of homoscedastic error terms adopted in Helpman, et al. (2008) causes misspecifications in their gravity model and might lead to biased and inconsistent estimators. In order

6.2.Elasticity of Substitution

Up to this point, we have largely followed on the heels of existing work. However, the main interest in this paper lies in obtaining substitution elasticities that are consistent with the underlying firm heterogeneity theory. Given the coefficient estimates reported in Table 6, we can now solve for the theoretically-consistent elasticities for use in global trade analysis. Table 8 reports these elasticities for the motor vehicles and parts industry under the benchmark and FH-SS specifications. A comparison with the associated GTAP Armington elasticity – dubbed ESUBM (Hertel, et al. 2003) -- is also presented in Table 8.

Column (1) reports the coefficient estimate of distance under the benchmark and FH-SS specifications, while column (2) reports the distance coefficient in Probit. The ratio of column (2) to column (1) gives a similar coefficient as the Power Law exponent of firm sizes. Values of this ratio under two specifications are reported in column (3) and found to be around 1. In particular, the ratio is 1.03 for the benchmark model and 0.84 for the FH-SS corrected model. These values are quite close to the Power Law exponent estimates found in the literature summarized in Table 3 (Axtell 2001, di Giovanni and Levchenko 2013, di Giovanni, et al. 2011, Melitz and Redding 2013). This suggests that empirical evidence about Power Law exponents for firm size is quite robust to the type of data used for estimation, as the value we obtain with country-level data is consistent with that obtained from the firm-level empirical studies.

Armed with empirically supported Power Law exponents, we move on to solve for the theoretically-consistent elasticities of substitution. The shape parameter estimate for motor vehicles and parts industry found in Spearot (2015) is reported in column (4) of Table 8. We use these estimates as data in the Power Law exponent to solve for our “Melitz” substitution elasticities which are reported in column (5). Elasticity values are found to be quite close across our specifications, 2.74 for the benchmark model and 3.13 for the corrected model. Both are substantially lower than the GTAP Armington elasticity of 5.60 and both satisfy the key parameter restriction of the model ($\gamma > \sigma - 1$).

to control for the heteroskedasticity in the data Silva and Tenreyro (2015) suggest using Poisson Pseudo-Maximum Likelihood Estimator (PPML). As a robustness check, we estimate the gravity equation in (1.19) using PPML. Regression results show that the coefficient estimate of distance is -0.71 and is significant at 1 per cent level.

It is important to note that even when the firm heterogeneity and sample selection corrections are not applied, the elasticity implied by the theory is lower than the Armington elasticity used in the GTAP model. This finding is consistent with our arguments and deserves further discussion. Even though the benchmark model does not take sample selection and firm heterogeneity into account, it still gives us a "Melitz" elasticity in this framework. This is because we are complementing the estimates found in the benchmark model with the estimates found in the Probit model to infer those elasticities. Although we do not use the Probit predictions in the trade-flow equation, we still use the information about export participation through the Power Law exponent. On the other hand, GTAP Armington elasticities are estimated based only on the trade-flow equation with an Armington structure; thereby, they do not contain any information about firm entry/exit behavior. As a result, we can say that when used with the Probit model, even the benchmark elasticity removes the supply-side effect captured in the GTAP Armington elasticity. In fact, when we compare it with the corrected model, we see that the benchmark case gives a lower elasticity estimate which implies that it removes more than the supply-side effects. That is to say the appropriate elasticity for the firm heterogeneity model lies somewhere between the benchmark elasticity and the GTAP Armington elasticity.

6.3.Implications and Limitations

So, what is the economic significance of finding a lower elasticity of substation between varieties for use in global economic analyses? To answer this question we should recall the effect of parameter choice on the extensive margin. Based on the definition in Chaney (2008) the extensive margin captures the contributions to trade flows of both the change in the number of exporters and their respective export volumes. As you may recall this corresponds to the familiar form $\gamma-(\sigma-1)$. This is where the choice of structural parameters becomes the key to policy implications. The extensive margin is less responsive to trade barriers when the elasticity is high, while the intensive margin is more responsive. Therefore, the choice of structural parameters will determine the trade response[15] as well as the welfare response to policy changes through micro and macro mechanisms in the model. The most relevant mechanisms in this context are changes

[15] We should recall from equation (1.13) that aggregate trade volume only depends on the shape parameter, not the elasticity of substitution. However, the relative contributions of the intensive and extensive margins of trade depend on the value of elasticity and they have significant policy implications.

in average productivity through the self-selection of firms into export markets and changes in consumer utility through the availability of new varieties. Both of these mechanisms primarily depend on the parameter choice.

Finding a lower elasticity means that the demand-side is more heterogeneous in the firm heterogeneity model for motor vehicles and parts than we thought it was based on the Armington elasticity. Since consumer preferences are more heterogeneous there is more room for new exporters in the MVH market to invest in differentiating their varieties. Therefore, marginal firms can markup their prices against large infra-marginal firms in the market. It should be noted that there is also significant supply-side heterogeneity in the MVH market. Spearot's (2015) shape estimate is 1.79 for MVH is one of the lowest shape parameter values within the aggregate manufacturing industry. This implies that infra-marginal firms have a disproportionate share of the overall activity in this market and marginal firms are much less productive compared to the incumbents. As noted in previous discussions and in Table 2, having a low productivity is less of a disadvantage when preferences are more heterogeneous (low elasticity). Even though marginal firms charge slightly higher prices than the incumbents, consumers are willing to pay a premium for new varieties. However, with a higher elasticity, marginal firms would have lost their market power and would be subsumed by the large and productive infra-marginal firms. So moving from the higher substitution elasticities used previously in GTAP-based studies of firm heterogeneity to the lower values suggested by this study represents an important change.

The take-away from this discussion is that the relative value of the shape parameter and the elasticity of substitution have important consequences for trade and welfare responses in a firm heterogeneity model. In a sense, quantitative outcomes are driven by the Power Law exponent of firm size. For example, as mentioned before, di Giovanni and Levchenko (2013) show that welfare impact of the extensive margin of trade is negligible when the Power Law exponent is 1, i.e. when the firm size distribution converges to Zipf's Law. In fact, they compare welfare gains from reductions in fixed and variable costs when the Power Law exponent is 1 to the case when it equals 2. They show that when the Power Law exponent equals 1, welfare gains from reductions in fixed costs are an order of magnitude lower and welfare gains from reductions in variable costs are an order of magnitude higher compared to the case when the Power Law exponent is 2.Quantitative outcomes are not the only policy implications we are interested in. Parameter choice also matters for analyzing the dominant mechanisms in bringing about the changes in trade flows and welfare.

The empirical evidence provided in this paper is a first step towards a more informed discussion about the theoretically-consistent parameterization of the firm heterogeneity model for purposes of global economic analysis. Of course our results are conditional on our choice of the shape parameter value as well as the underlying model specification. To be more specific, it is fair to argue that some of the variation in our Melitz elasticities is the result of using Spearot (2015) shape estimates, which are lower than the inferred shape parameters in the mainstream literature. In our view, this issue is an econometric one that requires firm-level data to estimate both parameters simultaneously. Future work should focus on separate identification of key firm heterogeneity parameters and provide confidence intervals to those estimates for systematic sensitivity analysis. However, we do not attempt to solve that problem in this paper because it does not contribute more to our main argument: Armington elasticities are not appropriate in a firm heterogeneity model and can result in significant overestimation of the trade response to policy reforms.

7. Concluding Remarks

In this paper we discuss a theoretically-consistent way to parameterize the firm heterogeneity model with a focus on the elasticity of substitution across varieties. The current CGE literature relies on Armington elasticities and infers shape parameters based on these elasticities. However, Armington elasticities are not appropriate in a firm heterogeneity model. In fact, their interpretation and the underlying econometric specification for their estimation are different in a Melitz (2003) framework. Particularly, the traditional gravity equation that delivers Armington elasticities do not control for the impact of firm self-selection into export markets which is the main micro mechanism for productivity and variety induced gains from trade. In the absence of firm behavior the resulting coefficient estimates confound the demand-side effects with the supply-side effects. This indicates overestimated elasticities which pick up part of the supply-side heterogeneity governed by the shape parameter. The resulting parameter set used in the current CGE literature is, then, an overestimated Armington elasticity with an inferred shape parameter that does not capture the substantial variation across industries.

In this paper we distinguish between the intensive and extensive margins of trade flows to obtain theoretically-consistent elasticities. In particular, we estimate two equations: an export participation equation and a gravity equation that governs bilateral trade flows. Since we use country level data, we impose further information in order to identify the elasticities. Specifically, we use the shape parameter estimates provided in Spearot (2015) which shows the variation of heterogeneity across industries and regions. Our results show that GTAP Armington elasticities are significantly higher than the elasticity estimates that are theoretically consistent with the Melitz (2003) model.

This paper provides an informed discussion about the theoretically-consistent parameterization of firm heterogeneity models in a CGE setting. Since we work with country-level data, separate identification of parameters is not feasible. Therefore, we rely on external shape parameter estimates. Our future research agenda is to identify elasticities and shape parameters separately by utilizing firm-level data. We, then, will be able to test the observational equivalence between an Armington-based model with a Melitz (2003) model in a CGE setting.

We believe that combining theory-consistent econometric evidence with the firm heterogeneity model in a CGE framework will lead the way for mainstream application of firm heterogeneity models in the GTAP community.

References

AKGUL, Z., N. B. VILLORIA, and T. W. HERTEL (2014): "Introducing Firm Heterogeneity into the Gtap Model," *In review for the GTAP Technical Paper Series*: Purdue University.

ARKOLAKIS, C. (2010): "Market Penetration Costs and the New Consumers Margin in International Trade," *Journal of Political Economy,* 118, 1151-1199.

ARKOLAKIS, C., S. DEMIDOVA, P. J. KLENOW, and A. RODRIGUEZ-CLARE (2008): "Endogenous Variety and the Gains from Trade," *American Economic Review*, 98, 444-450.

ARMINGTON, P. S. (1969): "Theory of Demand for Products Distinguished by Place of Production," *International Monetary Fund Staff Papers*, 16, 159-178.

AXTELL, R. L. (2001): "Zipf Distribution of Us Firm Sizes," *Science*, 293, 1818-1820.

BALISTRERI, E. J., R. H. HILLBERRY, and T. F. RUTHERFORD (2011): "Structural Estimation and Solution of International Trade Models with Heterogeneous Firms," *Journal of International Economics*, 83, 95-108.

BALISTRERI, E. J., and T. F. RUTHERFORD (2012): "Computing General Equilibrium Theories of Monopolistic Competition and Heterogeneous Firms," in *Handbook of Computable General Equilibrium Modeling*, ed. by P. B. Dixon, and D. W. Jorgenson.

BERNARD, A. B., J. EATON, J. B. JENSEN, and S. KORTUM (2003): "Plants and Productivity in International Trade," *American Economic Review*, 93, 1268-1290.

CHANEY, T. (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 98, 1707-1721.

CROZET, M., and P. KOENIG (2010): "Structural Gravity Equations with Intensive and Extensive Margins," *Canadian Journal of Economics-Revue Canadienne D Economique*, 43, 41-62.

DI GIOVANNI, J., and A. A. LEVCHENKO (2013): "Firm Entry, Trade, and Welfare in Zipf's World," *Journal of International Economics*, 89, 283-296.

DI GIOVANNI, J., A. A. LEVCHENKO, and R. RANCIÈRE (2011): "Power Laws in Firm Size and Openness to Trade: Measurement and Implications," *Journal of International Economics*, 85, 42-52.

DIXON, P. B., M. JERIE, and M. T. RIMMER (2015): "Modern Trade Theory for Cge Modelling: The Armington, Krugman and Melitz Models," *GTAP Technical Paper*, No. 36.

EATON, J., and S. KORTUM (2002): "Technology, Geography, and Trade," *Econometrica*, 70, 1741-1779.

EATON, J., S. KORTUM, and F. KRAMARZ (2011): "An Anatomy of International Trade: Evidence from French Firms," *Econometrica*, 79, 1453-1498.

ERGUL, E. (2010): "Firm Heterogeneity and Political Economy of International Trade," University of Illinois at Urbana Champaign.

FUJIWARA, Y. (2004): "Zipf Law in Firms Bankruptcy," *Physica a-Statistical Mechanics and Its Applications*, 337, 219-230.

GABAIX, X. (2008): "Power Laws in Economics and Finance," *National Bureau of Economic Research*, 14299.

HEAD, K., and T. MAYER (2014): "Gravity Equations: Workhorse,Toolkit, and Cookbook," 4, 131-195.

HEAD, K., T. MAYER, and J. RIES (2010): "The Erosion of Colonial Trade Linkages after Independence," *Journal of International Economics*, 81(1), 1-14.

HEAD, K., T. MAYER, and M. THOENIG (2014): "Welfare and Trade without Pareto," *American Economic Review*, 104, 310-316.

HELPMAN, E., M. MELITZ, and Y. RUBINSTEIN (2008): "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics*, 123, 441-487.

HERTEL, T., D. HUMMELS, M. IVANIC, and R. KEENEY (2003): "How Confident Can We Be in Cge-Based Assessments of Free Trade Agreements?," GTAP Working Paper No.26.

HILLBERRY, R., and D. HUMMELS (2013): "Chapter 18 - Trade Elasticity Parameters for a Computable General Equilibrium Model," in *Handbook of Computable General Equilibrium Modeling*, ed. by B. D. Peter, and W. J. Dale: Elsevier, 1213-1269.

HUMMELS, D., and P. J. KLENOW (2005): "The Variety and Quality of a Nation's Exports," *American Economic Review*, 95, 704-723.

LUTTMER, E. G. J. (2007): "Selection, Growth, and the Size Distibution of Firms," *The Quarterly Journal of Economics*, 122, 1103-1144.

MAYER, T., and S. ZIGNAGO (2005): "Market Access in Global and Regional Trade," *CEPII Working Paper 2*.

MAYER, T., and S. ZIGNAGO (2011): "Notes on Cepii's Distances Measures: The Geodist Database," *CEPII Working Paper 2011-25*.

MELITZ, M., and S. J. REDDING (2013): "Firm Heterogeneity and Aggregate Welfare," *National Bureau of Economic Research*, Working Paper 18919.

MELITZ, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71, 1695-1725.

MELITZ, M. J., and G. I. P. OTTAVIANO (2008): "Market Size, Trade, and Productivity (Vol 75, Pg 295, 2008)," *Review of Economic Studies*, 75, 985-985.

NARAYANAN, G. B., A. AGUIAR, and R. MCDOUGALL, Editors (2012): *Global Trade, Assistance, and Production: The Gtap 8 Data Base*. Center for Global Trade Analysis: Purdue University.

OKUYAMA, K., M. TAKAYASU, and H. TAKAYASU (1999): "Zipf's Law in Income Distribution of Companies," *Physica A*, 269, 125-131.

SILVA, J., and S. TENREYRO (2015): "Trading Partners and Trading Volumes: Implementing the Helpman-Melitz-Rubinstein Model Empirically," *Oxford Bulletin of Economics and Statistics*, 77, 93-105.

SPEAROT, A. (2015): "Unpacking the Long Run Effects of Tariff Shocks: New Structural Implications from Firm Heterogeneity Models.," Working Paper, University of California - Santa Cruz.

YI, K. M. (2003): "Can Vertical Specialization Explain the Growth of World Trade?," *Journal of Political Economy*, 111, 52-102.

ZHAI, F. (2008): "Armington Meets Melitz: Introducing Firm Heterogeneity in a Global Cge Model of Trade," *Journal of Economic Integration*, 23, 575-604.

Appendix: Data

This section defines the variables used in the empirical analysis and describes the data sources. We used two sources to obtain the data. The bilateral trade flows are from the GTAP Data Base Version 8.1 (https://www.gtap.agecon.purdue.edu/databases/v8/default.asp) (Narayanan et al., 2012). This version includes 57 GTAP commodities and 134 GTAP regions of which 113 country titles are available. We use the time series bilateral trade data of this version that covers the period 1995 to 2009 with 2007 as the reference year.

The gravity data have been obtained from the CEPII distance and gravity databases (http://www.cepii.fr/CEPII/en/bdd_modele/bdd.asp). ***GeoDist*** is CEPII's distance database developed by Mayer and Zignago (2005). In our paper, data on distance, contiguity, common language, colonial links and landlocked countries are obtained from *GeoDist*. There are two files available in this database: a country-specific dataset `geo_cepii.xls` (`geo_cepii.dta`) which includes geographical variables for 225 countries and a dyadic dataset `dist_cepii.xls` (`dist_cepii.dta`) which includes different measures of bilateral distances between 224 countries. The content of these files and details about the variables included in these files are explained in Mayer and Zignago (2011). ***Gravity*** is CEPII's gravity database `gravity_cepii` (`gravdata_cepii.dta`) based on Head et al. (2010). This database covers an exhaustive set of variables for 224 countries for the period 1948 to 2006. Details about the sources used in creating this database are explained in Head et al. (2010). In our paper, data on common legal origins, common currency, FTA and GATT/WTO membership are obtained from *Gravity*.

The time period considered in this paper is from 1995 to 2006 to match the time series of bilateral trade from GTAP and the gravity variables from CEPII. In particular, we drop the years 2007-2009 from the GTAP time series data and we drop the years 1948-1994 from the CEPII *Gravity* data. Our final dataset is obtained by merging GTAP data with CEPII data for motor vehicles and parts industry, 113 country titles and it covers the period from 1995 to 2006.

All of the variables used in our empirical work are summarized below with details about the respective data sources resorted to obtain them. To facilitate comparison with the gravity literature we adopt the convention in Helpman et al. (2008) for several of the variable definitions.

Bilateral Trade: is the bilateral trade between exporter r and importer s in millions of US dollars. We use GTAP database for information about bilateral trade flows. In particular, we use value of export sales evaluated at world (FOB) prices which corresponds to 'VXWD' in GTAP. The dependent variable in our empirical work is value of export sales in logs.

Distance: the population-weighted bilateral distance between the biggest cities of exporter r and importer s in kilometers. For distance, we use the `dist_cepii` file of the CEPII *GeoDist* database. This uses city level data to evaluate the geographic distribution of population inside each country. There are two population-weighted distance measures reported in this database. We use the one named as '`distw`' which is calculated by setting the sensitivity of trade flows to bilateral distance as 1. We use log(distance) in the regression equations.

Contiguity: a dummy variable that equals one if exporter r and importer s are adjacent countries, i.e. are contiguous, and zero otherwise. For contiguity, we use the `dist_cepii` file of the CEPII *GeoDist* database. The name of the variable in the database is '`contig`'.

Common Colony: a dummy variable that equals one if exporter r and importer s have had a common colonizer after 1945. For common colony, we use the `dist_cepii` file of the CEPII *GeoDist* database. The name of the variable in the database is '`comcol`'.

Colonial Link: a dummy variable that equals one if exporter r and importer s have ever had a colonial link, and zero otherwise. For colonial link, we use the `dist_cepii` file of the CEPII *GeoDist* database. The name of the variable in the database is '`colony`'.

Common Language: a dummy variable that equals one if exporter r and importer s share a common official language, and zero otherwise. For common language, we use the `dist_cepii` file of the CEPII *GeoDist* database. The name of the variable in the database is '`comlang.off`'.

Landlocked: a dummy variable that equals one if both exporter r and importer s are landlocked countries, i.e. no direct access to sea, and zero otherwise. For landlocked countries, we use the `geo_cepii` file of the CEPII *GeoDist* database. The name of the variable in the database is 'landlocked'. This database is country specific; therefore, landlocked is defined as a dummy variable that equals one if the particular country is landlocked. We define a new dummy variable for our purposes using the country-specific information available in geo_cepii.

Common Legal Origins: a dummy variable that equals one if exporter r and importer s share a common legal origin, and zero otherwise. For common legal origins, we use the gravity_cepii file of the CEPII *Gravity* database. The name of the variable in the database is 'comleg'.

Common Currency: a dummy variable that equals one if exporter r and importer s use the same currency, and zero otherwise. The data on currency unions come from the gravity_cepii file of the CEPII *Gravity* database. The name of the variable in the database is 'comcur'.

GATT/WTO Membership: a dummy variable that equals one if both exporter r and importer s are GATT/WTO members, and zero otherwise. The data on GATT/WTO membership comes from the gravity_cepii file of the CEPII *Gravity* database. This database has separate information about the GATT/WTO membership of exporter r (gatt_o) and importer s (gatt_d). Therefore, we define a new dummy variable to incorporate the membership information on both countries which matches our definition above.

FTA/RTA: a dummy variable that equals one if both exporter r and importer s belong to the same regional trade agreement, and zero otherwise. FTA data comes from the gravity_cepii file of the CEPII *Gravity* database. The name of the variable in the database is 'rta'.

Table 1. List of Countries

List of Countries				
Albania	Côte d'Ivoire	Iran	Namibia	South Africa
Argentina	Croatia	Ireland	Nepal	Spain
Armenia	Cyprus	Israel	Netherlands	Sri Lanka
Australia	Czech Republic	Italy	New Zealand	Sweden
Austria	Denmark	Japan	Nicaragua	Switzerland
Azerbaijan	Ecuador	Kazakstan	Nigeria	Taiwan
Bahrain	Egypt	Kenya	Norway	Thailand
Bangladesh	El Salvador	Korea	Oman	Togo
Belarus	Estonia	Kuwait	Pakistan	Tunisia
Belgium - Lux.	Ethiopia	Kyrgyzstan	Panama	Turkey
Benin	Finland	Laos	Paraguay	Uganda
Bolivia	France	Latvia	Peru	Ukraine
Botswana	Georgia	Lithuania	Philippines	United Arab Emirates
Brazil	Germany	Luxembourg	Poland	United Kingdom
Bulgaria	Ghana	Madagascar	Portugal	United Rep. of Tanzania
Burkina Faso	Greece	Malawi	Qatar	United States of America
Cambodia	Guatemala	Malaysia	Russian Federation	Uruguay
Cameroon	Guinea	Malta	Rwanda	Venezuela
Canada	Honduras	Mauritius	Saudi Arabia	Viet Nam
Chile	Hong Kong	Mexico	Senegal	Zambia
China	Hungary	Mongolia	Singapore	Zimbabwe
Colombia	India	Morocco	Slovakia	
Costa Rica	Indonesia	Mozambique	Slovenia	

Table 2. Structural Parameters in the Firm Heterogeneity Model

Parameters	Low	High
σ	Demand side is heterogeneous	Demand side is homogeneous
	Differentiated varieties	Homogeneous varieties
	Marginal firms capture a big share of the market	*Infra-marginal* firms capture a big share of the market
	Low productivity is not a disadvantage	Low productivity is a disadvantage
γ	Supply side is heterogeneous	Supply side is homogeneous
	Productivity is more dispersed across firms	Productivity is less dispersed across firms
	Infra-marginal firms have a big share of industry output	*Marginal* firms have a big share of industry output
	Producing homogenous varieties is a disadvantage	Producing homogeneous varities is not a disadvantage

Table 3. Overview of Structural Parameter Values in the Firm Heterogeneity Literature

Author (Year)	Country (Period)	Sectors	Model	Dist.	Power Law Exponent, ϒ/(σ-1)	Shape parameter, ϒ	Substitution elasticity, σ
Axtell (2001)	US firm-level data (1997)	-	-	Zipf	1.06 (estimated)		
Eaton and Kortum (2002)	Cross-section data of 19 OECD Countries (1990)	Manufacturing	Ricardian model	Fréchet	-	8.28 , 12.86 (based on prices) 3.6 (based on wages)	-
Bernard et al., BEJK (2003)	US plant-level data (1992) for 47 importers	Manufacturing	Eaton and Kortum (2002)	Fréchet	-	3.6 (calibrated)	3.79 (calibrated)
Arkolakis et al. (2008)	Costa Rican imports from 111 exporters (1986-1992)	-	-	Pareto	-	5.3 (calibrated)	6.0
Zhai (2008)	Cross-section data	11 aggregate sectors	Melitz (2003)	Pareto	-	5.17 , 6.20, 7.75 (calibrated)	4.3, 5.0, 6.0 (calibrated)
Crozet and Koenig (2010)	France, panel data at the firm-level (1986-1992)	34 manufacturing sectors	Chaney (2008)	Pareto	-	[1.65-7.31] (estimated) mean: 3.09	[1.15-6.01] (estimated) mean: 2.25
Balistreri et al. (2011)	Cross-section (2001)	7 aggregate sectors	Melitz (2003)	Pareto	-	3.924, 4.582, 5.171 (estimated)	3.8 (BEJK, 2003)
Eaton et al. (2011)	French firm-level data for 113 importers (1986)	Manufacturing	Melitz (2003)	Pareto	2.46 (estimated)	4.87 (implied)	2.98 (method in BEJK(2003))
di Giovanni et al. (2011)	Fench firm-level data (2006)	25 tradeable sectors	Melitz (2003)	Pareto	[0.362 - 1.011] (exporters) [0.470 -1.663] (non-exporters) 1.06 (domestic sales)	-	-
di Giovanni and Levchenko (2013)	Cross-section data of 50 largest economies	-	Melitz (2003) Eaton et al. (2011)	Pareto	1.06 (di Giovanni et al., 2011)	5.3 (implied)	6 (Anderson and van Wincoop, 2004)
Melitz and Redding (2013)	US	-	Melitz (2003)	Pareto	1.42 (empirical evidence)	4.25 (implied)	4 (BEJK, 2003)
Spearot (2015)	Cross-section	39 sectors	Melitz and Ottaviano (2008)	Pareto	-	[1.76-6.29] (estimated)	-

Notes: This table is ordered based on each paper's publish date. Empirical methods followed in these studies include : Axtell (2001) uses a Power Law specification for firm sizes to estimate the Power Law exponent; Eaton and Kortum (2002) uses the Method of Moments Estimator to estimate the shape parameter of Fréchet distribution based on trade and prices as well as based on trade and wages; BEJK (2003) calculates the parameter values by matching the productivity and size advantage of exporters in the simulated data with that of the empirical data; Arkolakis et al. (2008) uses the Feenstra (1994) Ratio to adjust the standard import price index for changing varieties and calibrates the shape parameter; in Zhai (2008) the shape parameter is calibrated to profit ratio in total markup, while the elasticity is calibrated to the markup ratio; Crozet and Koenig (2010) use a structural estimation method to estimate both parameters; Balistreri et al. (2011) uses a structural estimation method where they fit trade flows subject to equilibrium conditions in the model; Eaton et al. (2011) uses a Simulated Method of Moments estimator to estimate the Power Law exponent; di Giovanni et al. (2011) estimate the Power Law exponent by using log rank-log size regression; di Giovanni and Levchenko (2013) use the Power Law exponent in di Giovanni et al. (2011) with the elasticity in Anderson and van Wincoop (2004) to infer the shape parameter; Melitz and Redding (2013) calibrate the shape parameter to match the Power Law exponent for firm revenue; Spearot (2015) estimates a structural trade growth equation using Maximum Likelihood to find the shape parameter.

Table 4. Summary Statistics of the Dataset, Motor Vehicles and Parts, 113 Countries, 1995-2006

Variable	Obs	Mean	Std.Dev.	Min	Max
Exports (millions $US)	151,872	48.35	809.4	0	65206
Distance (km)	151,872	7467	4346	131.8	19781
Contiguity	151,872	0.026	0.158	0	1
Common Language	151,872	0.112	0.315	0	1
Common Colony	151,872	0.062	0.241	0	1
Colonial Link	151,872	0.017	0.128	0	1
FTA/RTA	151,872	0.091	0.287	0	1
Common Legal Origins	151,872	0.299	0.458	0	1
Common Curency	151,872	0.010	0.098	0	1
GATT /WTO Membership (both)	151,872	0.714	0.452	0	1
Landlocked	151,872	0.040	0.196	0	1

Table 5. Zeros in the Motor Vehicles and Parts Industry, 113 Countries

Year	Frequency of Zeros	Fraction of Zeros (%)
1995	10,246	80.96
1996	10,165	80.32
1997	10,094	79.76
1998	10,010	79.09
1999	10,040	79.33
2000	10,028	79.24
2001	9,942	78.56
2002	9,943	78.56
2003	9,458	74.73
2004	9,293	73.43
2005	9,074	71.70
2006	8,996	71.08
Pooled	117,289	77.23

Table 6. Gravity Estimation Results for Motor Vehicles and Parts (MVH in GTAP), 113 Countries, for years 1995-2006

	Export Participation (Probit)		Export Value (OLS)	
Variables	Marginal Effects (1)	Coefficients (2)	Benchmark (3)	FH-SS (4)
log(Distance)	-0.093***	-0.941***	-0.914***	-1.121***
	(0.003)	(0.031)	(0.038)	(0.056)
Contiguity	0.043***	0.435***	0.548***	0.635***
	(0.011)	(0.116)	(0.105)	(0.118)
Common Colony	0.067***	0.675***	0.767***	0.960***
	(0.009)	(0.087)	(0.140)	(0.237)
Colonial Link	0.058***	0.582***	0.405***	0.581***
	(0.012)	(0.121)	(0.108)	(0.120)
Landlocked	0.015	0.154	0.023	-0.074
	(0.01)	(0.104)	(0.138)	(0.188)
Common Legal Origins	0.023***	0.235***	0.224***	0.323***
	(0.004)	(0.037)	(0.047)	(0.051)
Common Currency	0.018	0.179	0.374***	0.166
	(0.023)	(0.227)	(0.113)	(0.120)
GATT /WTO Membership (both)	0.025***	0.256***	0.383***	0.532***
	(0.004)	(0.045)	(0.056)	(0.066)
FTA/RTA	0.041***	0.410***	0.681***	0.728***
	(0.005)	(0.050)	(0.062)	(0.073)
Common Language	0.021***	0.212***	0.0565	
	(0.006)	(0.059)	(0.077)	
Sample Selection (η_hat)				-0.435*
				(0.227)
Firm Heterogeneity (lnρ_hat)				-0.183***
				(0.033)
Observations	151,872		34,583	28,355
R^2	0.672		0.699	0.721

Notes: Probit reports both the marginal effects at sample means and coefficient estimates, Benchmark is an OLS specification of a traditional gravity equation without any corrections, FH-SS is an OLS specification of a gravity equation with both firm heterogeneity (FH) and sample selection (SS) corrections, SS only correct for the sample selection bias, and FH only corrects for the firm heterogeneity bias. Each model includes importer, exporter and year fixed effects. R^2 in Probit corresponds to pseudo-R^2. Robust standard errors with country-pair clustering are reported in parentheses. ***p<0.01, **p<0.05, *p<0.1

Table 7. Corrections in the Gravity Equation for Motor Vehicles and Parts (MVH in GTAP), 113 Countries, for years 1995-2006

	Export Value (OLS)			
Variables	Benchmark (1)	FH-SS (2)	SS (3)	FH (4)
log(Distance)	-0.914***	-1.121***	-1.267***	-1.126***
	(0.038)	(0.056)	(0.051)	(0.056)
Contiguity	0.548***	0.635***	0.629***	0.628***
	(0.105)	(0.118)	(0.109)	(0.118)
Common Colony	0.767***	0.960***	1.046***	0.960***
	(0.140)	(0.237)	(0.157)	(0.238)
Colonial Link	0.405***	0.581***	0.696***	0.590***
	(0.108)	(0.120)	(0.117)	(0.120)
Landlocked	0.023	-0.074	0.030	-0.070
	(0.138)	(0.188)	(0.146)	(0.188)
Common Legal Origins	0.224***	0.323***	0.358***	0.325***
	(0.047)	(0.051)	(0.045)	(0.051)
Common Currency	0.374***	0.166	0.110	0.150
	(0.113)	(0.120)	(0.123)	(0.120)
GATT /WTO Membership (both)	0.383***	0.532***	0.504***	0.536***
	(0.056)	(0.066)	(0.053)	(0.066)
FTA/RTA	0.681***	0.728***	0.819***	0.737***
	(0.062)	(0.073)	(0.061)	(0.073)
Common Language	0.0565			
	(0.077)			
Sample Selection (η_hat)		-0.435*	1.375***	
		(0.227)	(0.071)	
Firm Heterogeneity (lnρ_hat)		-0.183***		-0.098***
		(0.033)		(0.028)
Observations	34,583	28,355	34,583	28,355
R^2	0.699	0.721	0.729	0.721

Notes: Benchmark is an OLS specification of a traditional gravity equation without any corrections, FH-SS is an OLS specification of a gravity equation with both firm heterogeneity (FH) and sample selection (SS) corrections, SS only correct for the sample selection bias, and FH only corrects for the firm heterogeneity bias. Each model includes importer, exporter and year fixed effects. R^2 in Probit corresponds to pseudo-R^2. Robust standard errors with country-pair clustering are reported in parentheses. ***p<0.01, **p<0.05, *p<0.1

Table 8. Elasticities of Substitution between Varieties of Different Sources

	Models	Probit	Power Law Exponent	Shape Parameter Spearot (2015)	Melitz Elasticity	GTAP Armington Elasticity
	(1)	(2)	(3)	(4)	(5)	(6)
	$-\delta(\sigma-1)$	$-\delta\gamma$	$\gamma/(\sigma-1)$	γ	σ	ESUBM
Benchmark	-0.91	-0.94	1.03	1.79	2.74	5.60
FH-SS	-1.12	-0.94	0.84	1.79	3.13	5.60

Notes: Probit reports coefficient estimates, Benchmark is an OLS specification of a traditional gravity equation without any corrections, FH-SS is an OLS specification of a gravity equation with both firm heterogeneity (FH) and sample selection (SS) corrections.