Which Gravity? A comparison approach using finite mixture modelling

Martin Bresslein* and Jan Pablo Burgard†

August 8, 2014

Abstract

In this paper, we employ finite mixture modelling to compare two theoretical models that result in a structural gravity equation. Obtaining exporter-sector fixed effects from a gravity estimation, we calculate the probabilities that an observation is consistent with one demand-side and one supply-side model by regressing the fixed effects on the underlying theoretical variables suggested by the two models. This procedure lets us infer on the models' performance. We find that both models explain variation in the data quite well. Also, a clear sectoral clustering structure in aligning observations to the two models is revealed.

PRELIMINARY AND INCOMPLETE, PLEASE DO NOT CITE!

Keywords: International Trade; Gravity; Model Comparison; Finite Mixture Modelling.

*Corresponding author: Department of Economics, Trier University, Universitätsring 15, 54286 Trier, Germany, telephone: +49-651-201-2741, e-mail: *mbress@uni-trier.de*.

†Department of Economic and Social Statistics, Trier University.

1 Introduction

The gravity model has been a workhorse model in the international trade literature for the past decades. Much progress has been made from the beginning of its use in the 1960s. Today, Gravity Models can be described as a class of models featuring a certain structure, exporter-specific characteristics, importer-specific characteristics, and pair-specific characteristics (Keith Head and Thierry Mayer, 2014). Whereas the gravity model started as a model borrowed from physics without any theoretical foundation in economic theory, in the past decades, it has been derived from various supply and demand systems. On the demand side, this entails, e.g. CES utility (James E. Anderson and Eric Van Wincoop, 2003), the love-of-variety approach (Simon P Anderson, Andre De Palma and Jacques Franois Thisse, 1992), quadratic utility functions (Marc J. Melitz and Gianmarco I. P. Ottaviano, 2008), and translog utility functions (Dennis Novy, 2013). On the supply side, e.g. a Ricardian-type framework (Jonathan Eaton and Samuel Kortum, 2002), an endowment framework (Anderson and Wincoop, 2003), or a heterogeneous firms framework (Thomas Chaney, 2008) have been used.

Gravity models have also been applied to evaluate various policy-relevant issues, from the effects of economic integration agreements (see e.g. Peter Egger and Mario Larch (2011) and Scott L. Baier, Jeffrey H. Bergstrand and Michael Feng (2014)) over the effects of monetary unions (see e.g. Reuven Glick and Andrew K. Rose (2002)) to the effects of colonial ties or cultural aspects (see e.g. Keith Head, Thierry Mayer and John Ries (2010)).

Estimations of these models have been performed at various levels of disaggregation with respect to trade data, yielding a wide range of results with varying goodness-of-fit. However, even at an aggregate level, the different specifications of supply and demand imply important assumptions and restrictions that compete with each other in terms of appropriateness.

Our goals are, on the one hand, to examine how the underlying theories perform when taken to data and, more importantly, to explore a model comparison approach recently introduced into the political science literature by Kosuke Imai and Dustin Tingley (2012), which lets us compare the appropriateness of different models using finite mixture modelling. Finite mixture modelling is not new in itself. However, until very recently, it has not been applied as a tool for model comparisons. It has also barely been used in the gravity literature. To our knowledge, only

Maureen B.M. Lankhuizen, Thomas de Graaff and Henri L.F. de Groot (2012) have applied finite mixture modelling in a gravity model context so far, clustering 4-digit industry-level data in terms of distance effects.

Finite mixture modelling is a flexible approach in the sense that it allows for the inclusion of more than two competing theories.[1] It therefore enables us to allocate each observation of a gravity dataset to one specific model by modelling the probability that an observation is consistent with a particular model as a function of the respective theory-consistent variables. It includes the possibility that the models are non-nested. The resulting likelihood function can be estimated using the Expectation-Maximization algorithm (see A. P. Dempster, N. M. Laird and D. B. Rubin (1977)). Afterwards, we can evaluate the overall performance of the different models by comparing the expected sample proportion of observations consistent with each model.

For now, we revert to comparing two gravity models, one demand-side derivations and one supply-side derivation. The demand-side derivation is the constant elasticity of substitution monopolistic competition model ("Dixit-Stiglitz-Krugman", CES MC DSK, see e.g. Shang-Jin Wei (1996)), while the supply-side derivation is a heterogeneous firms model with a log-concave demand system by Costas Arkolakis, Arnaud Costinot, Dave Donaldson and Andrés Rodríguez-Clare (2012). Both models yield a structurally similar aggregate or industry-level gravity equation that can be estimated consistently using exporter- and importer-fixed effects as well as some bilateral terms proxying trade costs. Our analysis consists of two steps. First, we estimate a gravity model as specified above. We then use the estimated exporter-fixed effects, which represent shifts in trade shares due to the underlying theory variables, as the dependent variable in the second step, regressing them on proxies for the underlying theory variables. Employing finite mixture modelling, we cannot only estimate the second-step equations simultaneously, but also predict the share of observations consistent with each theory. The second step is performed in R using the FlexMix package developed by (Friedrich Leisch, 2004). We use data at a two-digit ISIC level for 27 European countries as exporters and 171 countries as import partners for the year 2006. Our approach differs from James E. Anderson and Yoto V. Yotov (2012) in that we do not explore the variation in both, importer and exporter fixed effects, but compare different theories for the

[1] However, this is not a trivial exercise and Imai and Tingley (2012) suggest to start with two, which is what we do here as a first step.

variation in the exporter fixed effects. Furthermore, by employing finite mixture modelling, we let the data speak for itself.

Our findings are threefold. Firstly, both the CES MC DSK and the heterogeneous firms model with a log-concave demand system show significant explanatory power for the variation in the exporter-sector fixed effects. Secondly, we find a remarkably clear clustering in alignment to the two models with respect to ISIC sectors, where most observations can be sorted with a very high probability.[2] Our results also suggest that different models might be suitable for different sectors. Thirdly, the finite mixture modelling approach presents itself as a convenient tool to compare economic models with a possibly broad scope of further applications.

The rest of the article is structured as follows. In section 2, we briefly review gravity modelling theory and describe the models that we use for our comparison. In section 3, we show how finite mixture modelling works regarding model comparison. Section 4 describes our empirical approach and the estimation strategy, while section 5 covers our data and the respective sources. In section 6, we present our results, followed by section 7 in which we explore the robustness of our results. Section 8 concludes.

2 Gravity Modelling Theory

This section draws heavily on Head and Mayer (2014), who present definitions for the model class of gravity models as well as various theoretical trade models that result in gravity-type equations. We briefly describe the defintion and the underlying theoretical models we are going to compare. We adopt their notation to facilitate referencing.

2.1 Gravity Definition

Head and Mayer (2014) define structural gravity models as a subset of a wider class of general gravity models. A model fits their general gravity definition, if it leads to a bilateral trade equation of the following multiplicative form:

[2] This in itself is not a new result, since e.g. Christian Broda and David E. Weinstein (2006) estimate elasticities at a very disaggregated level, and theoretical models featuring CES demand mostly model one sector/good with a CES for different varieties.

$$X_{ni} = GS_i M_n \phi_{ni}, \tag{1}$$

where S_i refers to exporter capabilities, M_n to importer characteristics and ϕ_{ni} to bilateral factors influencing trade from exporter i to importer n - trade costs and the corresponding elasticities. G is termed a gravitational constant, though only in a cross-sectional setting. Structural gravity equations fulfill a slightly stricter definition:

$$X_{ni} = \frac{Y_i}{\Omega_i}\frac{X_n}{\Phi_n}\phi_{ni}, \tag{2}$$

where the first fraction on the right-hand side corresponds to the exporter capabilities, S_i, and the second to the importer characteristics, M_n, from the more general definition. Y_i represents the value of production, X_n total expenditures of n on imports from all its trade partners, while Ω_i and Φ_n represent the famous multilateral resistance terms (see (Anderson and Wincoop, 2003)). Head and Mayer (2014) define them in the following way:

$$\Omega_i = \sum_l \frac{\phi_{li} X_l}{\Phi_l} \quad \text{and} \quad \Phi_n = \sum_l \frac{\phi_{nl} Y_l}{\Omega_n}. \tag{3}$$

According to Head and Mayer (2014), this definition relies on two conditions, the importers' allocation of expenditures towards all exporters, and market-clearing for exporters. Firstly, it is crucial that the share of expenditure towards a specific exporter (including the importing country itself) can be expressed in a multiplicatively separable way, yielding an "accessibility-weighted sum of the exporter capabilities" (Head and Mayer, 2014, p. 9) and that these shares sum to one. Secondly, a country's exports to all importers (including the exporting country itself) should be expressible in a similar multiplicatively separable fashion, and be equal to the value of production. These two conditions ensure that a trade model results in a gravity equation according to equation 3. We now turn to the specific gravity variants we want to compare.

2.2 Gravity Flavors

Head and Mayer (2014) summarize seven model setups under the headline of structural gravity.

We refer to their handbook chapter and especially the original articles for detailed derivations. The models can be categorized according to "demand-side" and "supply-side" derivations. Demand side models feature exogenous wages and constant returns to scale or constant markups that "neutralize the supply side of the model"(Head and Mayer, 2014, p. 10). Supply side models make distributional assumptions (either using the Fréchet or the Pareto distribution) that neutralize demand side terms. All these models yield the same importer-specific components X_n/Φ_n. They also lead to the same trade cost term ϕ_{ni}, except for some differences in terms of structural parameters. Since these are not our focal point, we leave this issue for future research. In line with our main interest, we have a closer look at the exporter-specific terms of these models, since this is where the main difference in terms of the structural gravity equations comes from.

On the demand side, Head and Mayer (2014) distinguish a CES National product differentiation ("Anderson-Armington") model (James E. Anderson, 1979), the CES Monopolistic competition ("Dixit-Stiglitz-Krugman") model (one of the earliest derivations is referenced back to Wei (1996), a CES demand with CET production model (Scott L. Baier and Jeffrey H. Bergstrand, 2001), and a Heterogeneous consumer model based on Anderson, De Palma and Thisse (1992). On the supply side, a heterogeneous industry ("Ricardian Comparative Advantage") model (Eaton and Kortum, 2002), and two heterogeneous firms models, one based on Chaney (2008) that features CES monopolistic competition with constant markups, one based on Arkolakis et al. (2012) that features a quite general, log-concave demand system.

Here, we briefly present the assumptions and structural terms for the two models which we compare, the CES MC DSK and Heterogeneous firms models. In order to enhance readiblity, we present the aggregate versions, while later on we are going to use sectoral versions for our estimations.

Wei (1996) presents a version of the gravity model which is slightly modified in Head and Mayer (2014). It uses the DSK monopolistic competition framework (e.g. Paul R. Krugman (1979)), where in each country there are N_i competing firms, each supplying one variety of a good on international markets. Utility is modelled by a constant elasticity of substitution function with the elasticiy being the same between all varieties in the world and denoted by σ. Solving the model for the gravity terms yields $S_i = N_i w_i^{1-\sigma}$ characterizing the exporter attributes with w_i being wages, $M_n = X_n/\Phi_n$ as the importer-specific terms as in the definition of structural gravity, and $\phi_{ni}^{1-\sigma}$ trade costs between exporter i and importer n. These terms imply that both, the wage and the trade cost elasticity are given by $1 - \sigma$.

We now turn to the supply-side derivation in our comparison approach. We chose Arkolakis et al. (2012)'s version of a heterogeneous firms model since it yields gravity terms that are easy to interpret. They use a general log-concave demand system which is essentially neutralized in the resulting gravity equation due to distributional assumptions. On the supply side, their model features monopolistic competition a la Marc J. Melitz (2003) where firms competing in markets draw firm-specific productivities from a Pareto distribution.[3] The resulting gravity terms are then $M_n = {}^{X_n}/_{\Phi_n}$ as in both models above, $S_i = N_i \bar{\alpha}_i^{-\theta} w_i^{-\theta}$, where N_i and w_i again denote the number of firms (essentially varieties) and wages respectively, and $\bar{\alpha}_i$ reflects the upper support of the production cost distribution from which the firms draw.[4] Here, the trade cost, wage and productivity elasticities are all given by $-\theta$, which represents heterogeneity on the supply side.

To sum up, our comparison features two theoretical derivations of a structural gravity model that feature the same importer-specific terms and essentially the same bilateral terms. They are different with respect to the exporter-specific terms, which is what we are going to exploit here. They also differ in terms of parameter elasticities, especially trade cost elasticities. However, both θ and $1 - \sigma$ are inverse measures for heterogeneity. Heterogeneity of consumer tastes increases in ${}^1/_{\sigma - 1}$, while heterogeneity among firms[5] - essentially productivity differences - increases in ${}^1/_{\theta}$.

Having laid out which theoretical models we are going to compare, we now turn to briefly describing how we use finite mixture modelling in order to perform the comparison.

3 Finite Mixture Modelling for competing theories

Imai and Tingley (2012) propose using finite mixture modelling to jointly estimate models implied by competing theories. In our case, both models that we describe in section 2 attempt to explain trade flows from exporter i to importer n. The main difference between these models is the specification of the exporter-specific terms. Here, the models essentially compete.

Using Imai and Tingley (2012)'s notation, let $f_m(y|x, \theta_m)$ represents the m-th statistical model

[3] On a side note, by including a choke price, this model is also able to generate zero-trade flows.

[4] In the originial article by Arkolakis et al. (2012), the $\bar{\alpha}_i^{-\theta}$ is actually b_i^{θ}, the lower bound of the productivity distribution, which the productivity variable we use later on should capture more closely.

[5] Or industries in case of the Eaton and Kortum (2002) model.

of the set of M models that compete, with y being the outcome variable, x a set of explanatory variables and θ_m a vector of model parameters. To relate this to our situation at hand, y represents the exporter-specific shifts in trade flows and the x the variables that are only exporter-specific in each model.

The joint observed-data likelihood L_{obs} of the competing models may then be expressed as

$$L_{\text{obs}}(\Theta, \Pi | y, x) = \prod_{i=1}^{N} \left[\sum_{m=1}^{M} \pi_m f_m(y|x, \theta_m) \right] \quad . \tag{4}$$

$\Theta = \{\theta_1, \ldots, \theta_M\}$ denotes the set of model parameters for all competing models, and $\Pi = \{\pi_1, \ldots, \pi_M\}$ is the set of proportions each model m contributes to the joint model, satisfying the conditions $\sum_{m=1}^{M} \pi_m = 1$ and $\pi_m > 0,\ m = 1, \ldots, M$. In practice π_m is not known and has to be estimated. This can be achieved by applying the EM-Algorithm proposed by Dempster, Laird and Rubin (1977). This algorithm can be applied to maximum likelihood estimation methods, where parts of the likelihood are not observed, and therefore the maximization of the likelihood is not straightforward. In the present case, the π_m are assumed to be missing. Basically, instead of maximizing the likelihood directly, the following iterative procedure is applied (Dempster, Laird and Rubin, 1977):

E-step Estimate the expected likelihood given the estimated parameters of the likelihood of the preceding step.

M-step Maximize the likelihood given the estimated missing data.

The only missing information in order to have a fully specified likelihood in our case is the set Π. Following Leisch (2004), in the E-step the $\pi_m,\ m = 1, \ldots, M$, are estimated by taking the mean posterior probabilites $p_{m,i}$ of unit i resulting from theory m over all units $i = 1, \ldots, N$. These posterior probabilites $p_{m,i}$ are given by,

$$p_{m,i} = \frac{\pi_m f_m(y_i | x_i, \theta_m)}{\sum_{j=1}^{M} \pi_j f_j(y_i | x_i, \theta_j)} \tag{5}$$

In the M-step the classical maximum likelihood estimates, weighted by the $p_{m,i}$ obtained from the E-step, are computed for each model separately. These two steps are repeated until convergence. For a deeper discussion on the computational implementation we refer to Leisch (2004).

From this joint model we are mainly interested in the $p_{m,i}$ and π_m. As Imai and Tingley (2012, p. 222) state the π_m can be interpreted as the "overall performance of theory m", wheras $p_{m,i}$ "measures the consistency between a specific observation and a particular theory".

4 Econometric Model and Estimation Strategy

From section 2 we have two theoretical models, each yielding a structural gravity equation, also for sectoral data. As already mentioned, we now emply sectoral versions of these models. According to Head and Mayer (2014), all structural gravity models can be consistently estimated using fixed effects for the exporter-sector- and importer-sector-specific characteristics and various bilateral trade cost variables. What we propose is to make use of the exporter-sector-specific fixed effects which represent shifts in the trade shares for each exporter-sector. The theoretical models suggest different explanations for these shifts. To evaluate the performance of the two theoretical models in explaining the exporter-sector-specific shifts in trade shares, we proceed in two steps. In a first step, we estimate a gravity model as described above, obtaining estimates for the exporter-sector fixed effects. In a second step, these fixed effects can be used as the dependent variable and the variables suggested by each of the three theories as the explanatory variables in order to examine the models explanatory power.

Thus, our first-step equation reads:

$$X_{ink} = exp[\beta_0 + FE_{ik} + FE_{nk} + \beta_1 lnDIST_{in} + \beta_2 CONTIG_{in} + \beta_3 RTA_{in}] + \varepsilon_{ink}, \tag{6}$$

where X_{ink} denotes exports from i to n in sector k, FE_{ik} denotes fixed effects for each exporter-sector ik, FE_{nk} denotes fixed effects for each importer-sector nk, $DIST$ refers to bilateral distance between countries i and n, $CONTIG$ is a dummy variable indicating if countries i and n share a common border, RTA is a dummy that captures a free trade agreement between i and n, and ε_{ink} denotes an error term. Following J. M. C. Santos Silva and Silvana Tenreyro (2006), we estimate equation 6 using the Poisson Pseudo Maximum Likelihood (PPML) estimator in order to account for heteroskedasticy and the substantial share of zero-trade flows.[6] Afterwards, we use the estimated exporter-sector fixed effects and regress these on variables (as described in section 2) which, according to the different theoretical models, should determine the exporter-specific effects.

In essence, we have the following equations:

$$FE_{ik} = \begin{cases} N_{ik} w_{ik}^{1-\sigma} & \text{CES MC DSK,} \\ N_{ik} \bar{\alpha}_{ik}^{-\theta} w_{ik}^{-\theta} & \text{Log-concave heterogeneous firms,} \end{cases} \tag{7}$$

where, as already described in section 2, N_{ik} is the number of firms, w_{ik} denotes wages, and $\bar{\alpha}_{ik}$ denotes the upper support of the production cost distribution, all in country i and sector k respectively.

Employing the finite mixture modelling approach described in section 3, we estimate log-linearized versions of both models simultaneously while, at the same time, calculating probabilities that an observation can be attributed to one of the two models. As a benchmark, both models were also estimated separately. We were agnostic about the clustering that we might find. However, we expect to find certain structures, since it is unlikely that one of the models turns out to be of a "one-size-fits-all" type.

In the next section, we describe the data used for the estimations and their respective sources.

5 Data

For our first-step gravity estimation, we use bilateral trade data measured in billion Euros at the two-digit ISIC Revision 3.1 level sourced from the UN comtrade database[7] for 27 member states of the European Union and 171 import partners and 22 manufacturing sectors for the year 2006. Our trade cost variables all stem from the full gravity dataset provided by CEPII.[8] We calculate

[6]Using Monte-Carlo simulations, Head and Mayer (2014) compare the performance of several estimation methods for this setup. Although the PPML method is not perfect, it performed rather well. The only other method with a comparable performance (the performance of the estimators varied according to the nature of the data-generating process, i.e. how zero-trade flows come about) was a Tobit estimator suggested by Jonathan Eaton and Samuel Kortum (2001). As robustness checks, we include this estimator later on as well as LSDV-type regressions.

[7]accessed through the World Integrated Trade Solution sytem; see `https://wits.worldbank.org/`.

[8]accessible at `http://www.cepii.fr/CEPII/en/bdd_modele/presentation.asp?id=8`.

internal trade using production and trade data downloaded from Eurostat's Prodcom database.[9] In total, we have 415793 observations for our first-step gravity regression.

For the second-step estimation, we obtain more than 500 exporter-sector fixed effects[10] from the first step. For the explanatory variables, we use data sourced from Eurostat's Structural Business Statistics (SBS) database[11], which provides data on business demographics at the NACE Revision 1.1 two-digit level until 2008.[12] Incidentally, the two-digit versions of ISIC 3.1 and NACE 1.1 match so that we did not have to work with concordance tables. From the SBS database, we obtained data on the number of firms as well as wages measured in million Euros in each country and sector. Since we are not able to directly obtain information on the upper support of the production cost distribution in each country and sector, we proxy the theoretical variable using wage-adjusted labor productivity. Since not every variable is available for every year, we tried to fill missing values in 2006 by averaging values for the years 2004 to 2008. This leaves 318 observations for the second-step regression.

6 Estimation results

Table 1: Gravity Results for Bilateral Trade

	Distance	Contiguity	RTA	cons
estimate	-1.175542***	0.2166882***	0.6688995***	-0.8288285**
s.e.	(0.0161062)	(0.0162882)	(0.0587304)	(0.3636775)
415793 observations	Signif. codes: 0<'***'<0.001<'**'<0.01<'*'<0.05<'°'<0.1<' '<1			

Table 1 presents the results obtained from our first-step gravity model estimation.[13] All three trade cost variables have the expected sign and are of a reasonable magnitude. Based on an earlier analysis by Anne-Célia Disdier and Keith Head (2008), Head and Mayer (2014) perform a meta analysis of more than 2500 gravity estimates from 159 articles. For structural gravity models, they find a median distance coefficient of -1.14 (mean -1.1), a median contiguity coefficient of 0.52 (mean 0.66), and a median RTA coefficient of 0.28 (mean 0.36). This implies we are almost in the center for distance, a bit on the high end for the RTA dummy, and a bit on the low end

[9] see `http://epp.eurostat.ec.europa.eu/portal/page/portal/prodcom/data/database`.

[10] The dummy for Sweden in ISIC sector 36 is dropped as the base category

[11] see `http://epp.eurostat.ec.europa.eu/portal/page/portal/european_business/introduction`.

[12] Afterwards, the data is provided in the NACE Revision 2 classification.

[13] Note that the PPML estimator automatically calculates robust standard errors.

for the common border dummy. Thus, having obtained results that are very much in line with the structural gravity literature gives us confidence that we also have reasonable results for our exporter-sector fixed effects.

Our results from the second-step regression are shown in table 2.

Table 2: Results from the competing models under the finite mixture approach and seperate estimation.

	Joint Finite Mixture Model		Seperate Estimation	
	Het. Firms	CES MC DSK	Het. Firms	CES MC DSK
(Intercept)	-11.58***	-1.43*	-5.86**	-3.40***
	(2.54)	(0.53)	(2.06)	(0.37)
Wages	0.82***	0.83***	0.79***	0.80***
	(0.10)	(0.08)	(0.07)	(0.07)
Firms number	0.21°	-0.15	0.02	-0.02
	(0.11)	(0.10)	(0.08)	(0.08)
Productivity	1.02*		0.44	
	(0.44)		(0.37)	
Adj. R-squared:			0.4431	0.4423
F-statistic:			85.08 (DF 3, 314)	126.7 (DF 2, 315)

318 observations Signif. codes: 0<'***'<0.001<'**'<0.01<'*'<0.05<'°'<0.1<' '<1

The first two columns display the estimates from the mixture estimation, while columns 3 and 4 display the estimates from separate estimations. As can be seen from the R^2 of the separate estimations, both models have significant explanatory power, which in our view suggests a reasonable justification as a theoretical basis for gravity equations. The R^2 are not as high as those found by Anderson and Yotov (2012). However, in contrast to our approach, they explore the theoretical model specified in James E. Anderson and Eric van Wincoop (2004). The interpretation of the coefficients is less clear-cut. Wages are highly significant and the coefficients do not vary across models, both in the mixture as well as the separate estimations. The number of firms does not have explanatory power with respect to the CES MC DSK model in neither estimation. However, through our mixture estimation we find a positive and significant influence in the heterogeneous firms model, at least at the 10%-level. Through the mixture estimation, we also obtain a positive and significant effect for productivity. We will explore these issues and their implications for the σ and θ heterogeneity parameters in more detail at a later stage of the project.

We now turn to the model performance evaluation. As explained in section 3, the model performance can be evaluated by examining the share of observations determined to be consistent

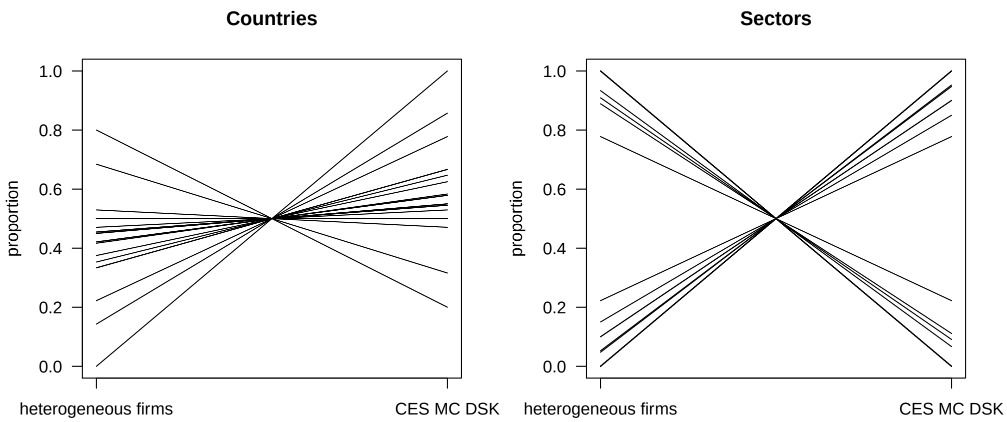

Figure 1: Clustering of observations according to underlying model, by country (left panel) and by sector (right panel).

with each model. Overall, we find 142 out of 318 observations consistent with the heterogeneous firms model, and 176 observations consistent with the CES MC DSK model. On itself, these results suggest that both models come with a certain validity and explanatory power. A more interesting pattern is revealed by looking at the observations that are consistent with each model in more detail.

Figure 1 shows the proportions of observations that are in line with either the heterogeneous firms model or the CES MC DSK model. In the left panel, each line represents the observations for a specific exporting country. Here, we do not find a clear structure since for most countries, observations are almost evenly split between the two models. However, the results in the right panel are striking. Here, each line represents the observations for a specific sector. The structure of these lines shows that observations for most sectors can be almost exclusively attributed to either one of the models.[14] For example, observations in sector 19 (”Manufacture of rubber and plastic products”) and 20 (”Manufacture of other non-metallic mineral products”) are determined to be exclusively consistent with the heterogeneous firms model. In contrast, sectors 17 (”Manufacture of coke, refined petroleum products and nuclear fuel”) and 21 (”Manufacture of basic metals”) completely align with the CES MC DSK model. This suggests that there might be specific characteristics to the products in these sectors that are much more in line with specifications of one of the two models. To support our results, figure 2 (in the appendix) displays a rootogram, showing that

[14]Tables 6 and 7 in the appendix show the number of observations for each country and sector that are attributed to either model.

most observations are also aligned to either one of the models with a relatively high probability.[15] In our view, these results indicate that both models have their value in terms of structural gravity and that there is no "one-size-fits-all" model, which is to be expected when considering the assumptions underlying each model. Also, this is not a really new result, since e.g. Broda and Weinstein (2006) estimate elasticities at a very disaggregated level, and theoretical models featuring CES demand mostly model one sector/good with a CES for different varieties. However, our results also show that certain sectors seem to be much more in line with one model than the other. Firstly, this underlines the results of Lankhuizen, de Graaff and de Groot (2012), who obtain a pretty clear sectoral clustering in terms of distance. Secondly, the cluster patterns might actually hold value for both, theoretical and empirical future research. We hesitate to draw strong conclusions since we are at an early stage and have only compared two models for now, but if we find similar patterns for other models, too, we would have a possible test to indicate which sectoral patterns are best explained by one of several demand and supply systems. Also, predetermined sectoral clustering could be incorporated in future estimations. Summing up, what we think our results show is that it is a worthwhile exercise to have a closer look at theories underlying gravity equations, and that finite mixture modelling could be a convenient tool for further applications of economic model comparisons.

7 Robustness Analysis

There are several issues, both technical and data- and model-wise, concerning our estimation procedure that warrant robustness checks. Firstly, since we have quite a high share of zeroes in our dataset, we cannot solely rely on the PPML estimator since its performance relies on the underlying data generating process. Therefore, we also estimate the first-step gravity equation using the Tobit estimator suggested by Eaton and Kortum (2001), as well as the a Poisson Quasi Likelihood estimator suggested by Kevin E. Staub and Rainer Winkelmann (2013) and applied by Sören Prehn and Bernhard Brümmer (2011), and compare the results to evaluate the respective performances. We also perform a parametric bootstrap and take the correlation from the obtained

[15] These results are not as clear-cut as those obtained by Imai and Tingley (2012) when re-analyzing Michael J. Hiscox (2002). Hiscox explores if voting patterns on trade policy in the US Senate and House are in line with either a Heckscher-Ohlin or a Ricardo-Viner model. In this case, factor mobility is a clear theory-predicting variable, which we do not have for our analysis.

variance-covariance matrices into account to adjust our standard errors. Furthermore, our dependent variable in the second step, consists of estimated sector-fixed effects. Here, too, we employ a parametric bootstrap to account for the possible ambiguity.

Secondly, for now our dataset is only of a cross-sectional nature. We find that its use is appropriate since, to date, all gravity models are static. However, since gravity models are also widely estimated with panel data, we also want to test our model comparison approach with trade data from several years. However, due to computational issues, we have to reduce the number of import partners in this context. We also hope to go on a lower sector-level using UN Indstat data to explore a possible aggregation bias.

Apart from these issues, we plan to cover a lot more models in the future, also including more countries and years when we will have obtained the necessary data. If possible, we will also apply this methodology to more disaggregated data.

8 Concluding remarks

In this paper, we employ finite mixture modelling to compare different theories leading to structural gravity equations.

References

Anderson, James E. 1979. "A Theoretical Foundation for the Gravity Equation." *American Economic Review*, 69(1): 106–116.

Anderson, James E., and Eric Van Wincoop. 2003. "Gravity with Gravitas: A Solution to the Border Puzzle." *American Economic Review*, 93(1): 170–192.

Anderson, James E., and Eric van Wincoop. 2004. "Trade Costs." *Journal of Economic Literature*, 42(3): 691–751.

Anderson, James E., and Yoto V. Yotov. 2012. "Gold Standard Gravity." National Bureau of Economic Research, Inc NBER Working Papers 17835.

Anderson, Simon P, Andre De Palma, and Jacques Franois Thisse. 1992. *Discrete choice theory of product differentiation.* MIT press.

Arkolakis, Costas, Arnaud Costinot, Dave Donaldson, and Andrés Rodríguez-Clare. 2012. "The Elusive Pro-Competitive Effects of Trade." *Unpublished, MIT.*

Baier, Scott L., and Jeffrey H. Bergstrand. 2001. "The growth of world trade: tariffs, transport costs, and income similarity." *Journal of International Economics*, 53(1): 1–27.

Baier, Scott L., Jeffrey H. Bergstrand, and Michael Feng. 2014. "Economic Integration Agreements and the Margins of International Trade." *Journal of International Economics*, forthcoming(0).

Broda, Christian, and David E. Weinstein. 2006. "Globalization and the Gains From Variety." *The Quarterly Journal of Economics*, 121(2): 541–585.

Chaney, Thomas. 2008. "Distorted Gravity: The Intensive and Extensive Margins of International Trade." *American Economic Review*, 98(4): 1707–1721.

Dempster, A. P., N. M. Laird, and D. B. Rubin. 1977. "Maximum Likelihood from Incomplete Data via the EM Algorithm." *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1): 1–38.

Disdier, Anne-Célia, and Keith Head. 2008. "The Puzzling Persistence of the Distance Effect on Bilateral Trade." *Review of Economics & Statistics*, 90(1): 37–48.

Eaton, Jonathan, and Samuel Kortum. 2001. "Trade in capital goods." *European Economic Review*, 45(7): 1195–1235.

Eaton, Jonathan, and Samuel Kortum. 2002. "Technology, Geography, and Trade." *Econometrica*, 70(5): 1741–1779.

Egger, Peter, and Mario Larch. 2011. "An assessment of the Europe agreements' effects on bilateral trade, GDP, and welfare." *European Economic Review*, 55(2): 263–279.

Glick, Reuven, and Andrew K. Rose. 2002. "Does a currency union affect trade? The time-series evidence." *European Economic Review*, 46(6): 1125–1151.

Head, Keith, and Thierry Mayer. 2014. "Chapter 3 - Gravity Equations: Workhorse,Toolkit, and Cookbook." In *Handbook of International Economics*. Vol. Volume 4, , ed. Kenneth Rogoff Elhanan Helpman and Gita Gopinath, 131–195. Elsevier.

Head, Keith, Thierry Mayer, and John Ries. 2010. "The erosion of colonial trade linkages after independence." *Journal of International Economics*, 81(1): 1–14.

Hiscox, Michael J. 2002. "Commerce, Coalitions, and Factor Mobility: Evidence from Congressional Votes on Trade Legislation." *American Political Science Review*, 96(03): 593–608.

Imai, Kosuke, and Dustin Tingley. 2012. "A Statistical Method for Empirical Testing of Competing Theories." *American Journal of Political Science*, 56(1): 218–236.

Krugman, Paul R. 1979. "Increasing returns, monopolistic competition, and international trade." *Journal of International Economics*, 9(4): 469–479.

Lankhuizen, Maureen B.M., Thomas de Graaff, and Henri L.F. de Groot. 2012. "Product Heterogeneity, Intangible Barriers and Distance Decay: The Effect of Multiple Dimensions of Distance on Trade across Different Product Categories." Tinbergen Institute Tinbergen Institute Discussion Paper 12-065/3, Amsterdam and Rotterdam:Tinbergen Institute.

Leisch, Friedrich. 2004. "FlexMix: A General Framework for Finite Mixture Models and Latent Class Regression in R." *Journal of Statistical Software*, 11(8): 1–18.

Melitz, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica*, 71(6): 1695–1725.

Melitz, Marc J., and Gianmarco I. P. Ottaviano. 2008. "Market Size, Trade, and Productivity." *The Review of Economic Studies*, 75(1): 295–316.

Novy, Dennis. 2013. "International trade without CES: Estimating translog gravity." *Journal of International Economics*, 89(2): 271–282.

Prehn, Sören, and Bernhard Brümmer. 2011. "Estimation issues in disaggregate gravity trade models." Georg-August University of Gttingen, Department of Agricultural Economics and Rural Development (DARE) DARE Discussion Papers 1107.

Silva, J. M. C. Santos, and Silvana Tenreyro. 2006. "The Log of Gravity." *Review of Economics and Statistics*, 88(4): 641–658.

Staub, Kevin E., and Rainer Winkelmann. 2013. "Consistent estimation of zero-inflated count models." *Health Economics*, 22(6): 673–686.

Wei, Shang-Jin. 1996. "Intra-National versus International Trade: How Stubborn are Nations in Global Integration?" National Bureau of Economic Research Working Paper 5531.

Appendix

Table 3: Exporters included in first-step gravity estimation

1	AUT	Austria
2	BEL	Belgium
3	BGR	Bulgaria
4	CYP	Cyprus
5	CZE	Czech Republic
6	DEU	Germany
7	DNK	Denmark
8	ESP	Spain
9	EST	Estonia
10	FIN	Finland
11	FRA	France
12	GBR	United Kingdom
13	GRC	Greece
14	HUN	Hungary
15	IRL	Ireland
16	ITA	Italy
17	LTU	Lithuania
18	LUX	Luxembourg
19	LVA	Latvia
20	MLT	Malta
21	NLD	Netherlands
22	POL	Poland
23	PRT	Portugal
24	ROM	Romania
25	SVK	Slovak Republic
26	SVN	Slovenia
27	SWE	Sweden

Table 4: Importers included in first-step gravity estimation

1	ABW	Aruba
2	AGO	Angola
3	ALB	Albania
4	AND	Andorra
5	ARE	United Arab Emirates
6	ARG	Argentina
7	ARM	Armenia
8	ATG	Antigua and Barbuda
9	AUS	Australia
10	AUT	Austria
11	AZE	Azerbaijan
12	BDI	Burundi
13	BEL	Belgium
14	BEN	Benin
15	BFA	Burkina Faso
16	BGD	Bangladesh
17	BGR	Bulgaria
18	BHR	Bahrain
19	BHS	Bahamas, The
20	BIH	Bosnia and Herzegovina
21	BLR	Belarus
22	BLZ	Belize
23	BMU	Bermuda
24	BOL	Bolivia
25	BRA	Brazil
26	BRB	Barbados
27	BRN	Brunei
28	BTN	Bhutan
29	BWA	Botswana
30	CAF	Central African Republic
31	CAN	Canada
32	CHE	Switzerland
33	CHL	Chile
34	CHN	China
35	CIV	Cote d'Ivoire
36	CMR	Cameroon
37	COG	Congo, Rep.
38	COL	Colombia
39	CRI	Costa Rica
40	CUB	Cuba
41	CYP	Cyprus
42	CZE	Czech Republic
43	DEU	Germany
44	DJI	Djibouti
45	DMA	Dominica
46	DNK	Denmark

47	DOM	Dominican Republic
48	DZA	Algeria
49	ECU	Ecuador
50	EGY	Egypt, Arab Rep.
51	ERI	Eritrea
52	ESP	Spain
53	EST	Estonia
54	ETH	Ethiopia(excludes Eritrea)
55	FIN	Finland
56	FJI	Fiji
57	FRA	France
58	GAB	Gabon
59	GBR	United Kingdom
60	GEO	Georgia
61	GHA	Ghana
62	GIN	Guinea
63	GMB	Gambia, The
64	GNB	Guinea-Bissau
65	GRC	Greece
66	GRD	Grenada
67	GRL	Greenland
68	GTM	Guatemala
69	HKG	Hong Kong, China
70	HND	Honduras
71	HRV	Croatia
72	HTI	Haiti
73	HUN	Hungary
74	IDN	Indonesia
75	IND	India
76	IRL	Ireland
77	IRN	Iran, Islamic Rep.
78	IRQ	Iraq
79	ISL	Iceland
80	ISR	Israel
81	ITA	Italy
82	JAM	Jamaica
83	JOR	Jordan
84	JPN	Japan
85	KAZ	Kazakhstan
86	KEN	Kenya
87	KGZ	Kyrgyz Republic
88	KHM	Cambodia
89	KOR	Korea, Rep.
90	KWT	Kuwait
91	LAO	Lao PDR
92	LBN	Lebanon
93	LBR	Liberia
94	LBY	Libya

95	LKA	Sri Lanka
96	LSO	Lesotho
97	LTU	Lithuania
98	LUX	Luxembourg
99	LVA	Latvia
100	MAC	Macao
101	MAR	Morocco
102	MDA	Moldova
103	MDG	Madagascar
104	MDV	Maldives
105	MEX	Mexico
106	MKD	Macedonia, FYR
107	MLI	Mali
108	MLT	Malta
109	MMR	Myanmar
110	MNG	Mongolia
111	MOZ	Mozambique
112	MRT	Mauritania
113	MUS	Mauritius
114	MWI	Malawi
115	MYS	Malaysia
116	NAM	Namibia
117	NER	Niger
118	NGA	Nigeria
119	NIC	Nicaragua
120	NLD	Netherlands
121	NOR	Norway
122	NPL	Nepal
123	NZL	New Zealand
124	OMN	Oman
125	PAK	Pakistan
126	PAN	Panama
127	PER	Peru
128	PHL	Philippines
129	PNG	Papua New Guinea
130	POL	Poland
131	PRT	Portugal
132	PRY	Paraguay
133	QAT	Qatar
134	ROM	Romania
135	RUS	Russian Federation
136	RWA	Rwanda
137	SAU	Saudi Arabia
138	SDN	Fm Sudan
139	SEN	Senegal
140	SGP	Singapore
141	SLE	Sierra Leone
142	SLV	El Salvador

143	SOM	Somalia
144	SUR	Suriname
145	SVK	Slovak Republic
146	SVN	Slovenia
147	SWE	Sweden
148	SWZ	Swaziland
149	SYR	Syrian Arab Republic
150	TCA	Turks and Caicos Isl.
151	TCD	Chad
152	TGO	Togo
153	THA	Thailand
154	TJK	Tajikistan
155	TKM	Turkmenistan
156	TTO	Trinidad and Tobago
157	TUN	Tunisia
158	TUR	Turkey
159	TZA	Tanzania
160	UGA	Uganda
161	UKR	Ukraine
162	URY	Uruguay
163	USA	United States
164	UZB	Uzbekistan
165	VEN	Venezuela
166	VNM	Vietnam
167	YEM	Yemen
168	ZAF	South Africa
169	ZAR	Congo, Dem. Rep.
170	ZMB	Zambia
171	ZWE	Zimbabwe

Table 5: Sectors included in first-step gravity estimation

Code	Description
15	Manufacture of pulp, paper and paper products
16	Publishing, printing and reproduction of recorded media
17	Manufacture of coke, refined petroleum products and nuclear fuel
18	Manufacture of chemicals and chemical products
19	Manufacture of rubber and plastic products
20	Manufacture of other non-metallic mineral products
21	Manufacture of basic metals
22	Manufacture of fabricated metal products, except machinery and equipment
23	Manufacture of machinery and equipment n.e.c.
24	Manufacture of office machinery and computers
25	Manufacture of electrical machinery and apparatus n.e.c.
26	Manufacture of radio, television and communication equipment and apparatus
27	Manufacture of medical, precision and optical instruments, watches and clocks
28	Manufacture of motor vehicles, trailers and semi-trailers
29	Manufacture of other transport equipment
30	Manufacture of furniture; manufacturing n.e.c.
31	Recycling
32	Electricity, gas, steam and hot water supply
33	Collection, purification and distribution of water
34	Construction
35	Sale, maintenance and repair of motor vehicles and motorcycles; retail sale of automotive fuel
36	Wholesale trade and commission trade, except of motor vehicles and motorcycles

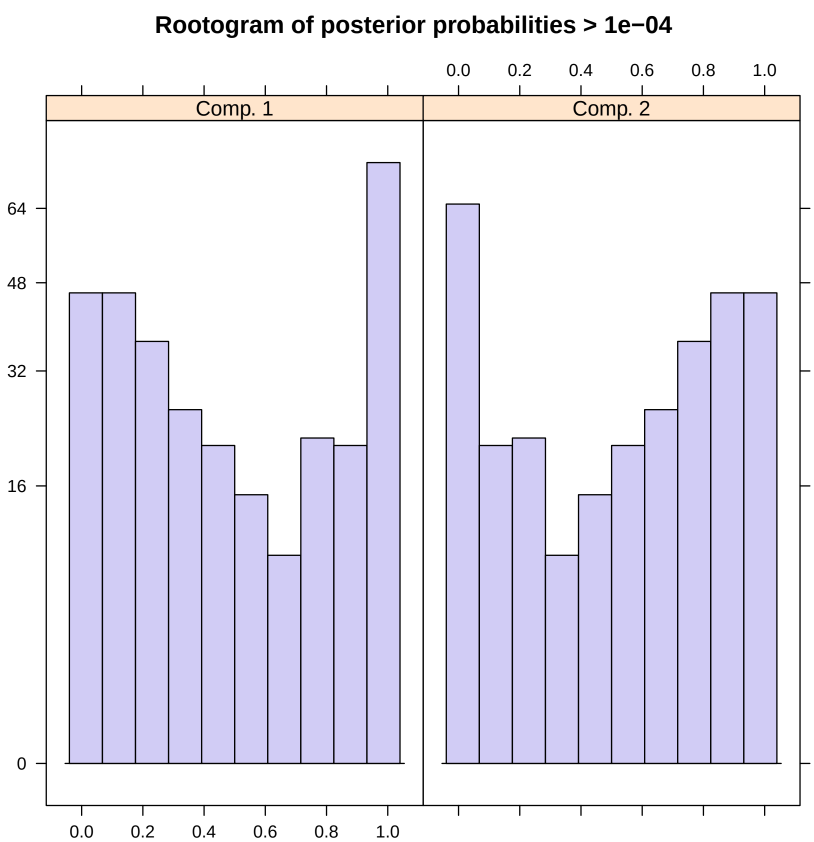
Rootogram of posterior probabilities > 1e−04
Comp. 1
Comp. 2
0
16
32
48
64
0.0
0.2
0.4
0.6
0.8
1.0

Figure 2:

Table 6: Number of observations in line with each model by country.

	Het. Firms	CES MC DSK
AUT	9	11
BGR	2	7
CYP	8	2
CZE	8	9
DEU	10	12
DNK	9	8
ESP	10	12
EST	6	11
FIN	9	11
FRA	10	10
GBR	11	11
GRC	7	7
HUN	8	11
LTU	2	12
LVA	2	4
MLT	0	5
NOR	13	6
PRT	2	4
ROM	3	5
SVK	5	7
SWE	8	11

Table 7: Number of observations in line with each model by sector.

	Het. Firms	CES MC DSK
15	0	9
16	0	4
17	0	13
18	10	1
19	11	0
20	10	0
21	0	13
22	1	9
23	9	0
24	3	17
25	4	14
26	16	0
27	14	1
28	14	4
29	1	20
30	0	13
31	2	18
32	0	20
33	1	18
34	16	2
35	16	0
36	14	0