CHARLES UNIVERSITY IN PRAGUE

FACULTY OF SOCIAL SCIENCES

INSTITUTE OF ECONOMIC STUDIES

DIPLOMA THESIS

Various Estimation Techniques of the Gravity Model of Trade

Author: **bc. Lucie Davidová**

Supervisor: **doc. Ing. Vladimír Benáček, CSc.**

Academic year: **2014/2015**

Declaration of Authorship

The author hereby declares that she compiled this thesis independently, using only the listed resources and literature.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, January 5, 2015

Lucie Davidová

Poděkování

Na tomto místě bych ráda poděkovala vedoucímu své práce, panu docentu Benáčkovi, za řadu cenných rad a připomínek, za vstřícný osobní přístup a za velmi dobrou spolupráci při publikaci společného výzkumu. Také děkuji všem, kteří si mou práci přečetli, a zvláště pak těm, kteří se se mnou podělili o své názory na ni.

Bibliographic Record

DAVIDOVÁ, Lucie. *Various Estimation Techniques of the Gravity Model of Trade*. Prague, 2015. Diploma Thesis (Mgr.) Charles University in Prague, Faculty of Social Sciences, Institute of Economic Studies. Supervisor: doc. Ing. Vladimír Benáček, CSc.

Abstract

This diploma thesis deals with alternative estimation possibilities of the gravity model in trade. We provide the reader with a synthetic methodological overview of the technical problems with the estimation of gravity equations. Consequently, we test for the heterogeneity of datasets used in gravity models of trade which leads us to a conclusion that behavioural patterns of exporters and importers built in the datasets are very complicated and a single generalized specification of gravity equation can lead to bias in estimates and/or to similarly generalized conclusions that hide important robust idiosyncrasies in behavior present in some subsamples of economic agents. Both the theory of estimation techniques and dataset heterogeneity are applied in the empirical part estimating Austria's export function.

JEL classification: C23, C26, C38, C52, F12, F14, F17

Keywords: gravity model, international trade, Austria, fixed effects, Poisson, Mundlak, Hausman-Taylor, cluster analysis

Author's email: davidova.lucka@gmail.com

Supervisor's email: benacekv@fsv.cuni.cz

Abstrakt

Tato diplomová práce se věnuje alternativním možnostem odhadu gravitačního modelu obchodu. Nabízíme souhrnný metodologický přehled technických problémům, kterým ekonomové čelí při odhadování gravitačních modelů. Následně testujeme heterogenitu dat se závěrem, že charakter chování vývozců a dovozců je nesmírně komplexní. To způsobuje, že příliš zobecněné specifikace gravitačních modelů mohou vést k vychýleným výsledkům a případně i k příliš obecným závěrům, které nezohledňují specifičnosti chování v jednotlivých pod-skupinách ekonomických agentů. Jak teorie o technikách odhadu, tak heterogenitu datasetu následně aplikujeme v empirické části při odhadu vývozní funkce Rakouska.

Klíčová slova: gravitační model, mezinárodní obchod, Rakousko, fixní efekty, Poisson, Mundlak, Hausman-Taylor, shluková analýza

Rozsah práce: 49 stran, 100 461 znaků

CONTENTS

LIST OF TABLES

INTRODUCTION

The aim of this thesis is to enrich the previous research done in my bachelor diploma thesis (Davidová, 2012), the abbreviated version of which was published in the IES Working Papers (Davidová & Benáček, 2013). The core framework to be scrutinized is the gravity model of bilateral trade flows. By such model we try to reveal the driving forces shaping the short-term and long-term patterns of international trade flows. In doing so we must first specify the trade function. This is done by aligning the empirical tests with economic theories. Consequently, we are able to calculate the coefficients of the applied model, which take most often the form of elasticities with respect to specified estimators (determinants). As an outcome, we are able to estimate also the trade potential and proceed with predictions in the behavior of exporters or importers.

Literature published globally on the models of international trade typically estimates the complete panel dataset, yielding a general trade function. This approach takes advantage of typically huge number of observations which makes the results seemingly very reliable (significant). It is, however, obvious that the huge trade datasets grouping most varied countries and products cannot avoid the risk of putting together observations whose data are subject to incompatible statistical reporting and cultural (behavioral) factors that are too heterogeneous for being explained by one common specification. This was unveiled by Egger & Průša (2014), Baltagi & Egger (2014), Egger & Nigai (2014) and Fidrmuc (2009) among others. Also Paulus et al. (2014) and Bobková (2014) at IES FSV UK tested the heterogeneity and the poolability of data in gravity models. The problem gets even more confusing once the estimation techniques assign different weights to observed heterogeneity in behavior in subpopulations of data.

Our paper therefore deals with this issue by simply easing the assumption that the foreign trade is ruled by a single deterministic equation that would be homogenous across all countries and time. Instead, we build our model on a contrary assumption that the more specific trade function we are looking for, the more exact results we obtain. This is why we constrain our trade function to one trade direction (export) and a single home-country (Austria) only, as we believe that these data characteristics in fact impact the shape of the trade function as well.

Whenever we estimate a panel data regression, we implicitly assume that the coefficients are homogenous for all countries and that the relationship with the trade flow is exactly linear. This is in fact a far-reaching arbitrary assumption as each home and partner country is represented by a number of real exporters and importers, who are heterogeneous – they trade in different products, and their countries have different endowments, productivities and policies. All these imply that the risk of "omitted variables" can end up in estimation bias, up to the spurious regression syndrome (Egger & Nigai, 2014). The decision-makers are actually a very diverse group of business units of different sizes, different preferences and of course of different aims. Therefore, what is usually revealed are just some kind of "average" behavioral properties with respect to the independent variables.

It therefore makes sense to search for more specific export functions, valid for a particular sup-population of the sample in particular. For example, we could reasonably expect the Austrian export function for European countries would significantly differ from the one valid for poor African partner-countries. Trade functions heterogeneity was empirically confirmed by e.g. Bobková (2014), Sarafidis & Weber (2014), Chang-Ching & Ng (2012), Sarafidis & Weber (2009) or Kapetanios (2006). This underlying assumption leads us to testing hypothesis of data heterogeneity, i.e. that our dataset contains several meaningful clusters of countries, whose estimation of mutual trade is subject to different theoretical underpinnings and thus their functions should be treated separately. As a result, we obtain different elasticity coefficients as well as different set of significant variables. Our complete dataset consists of about 211 countries, over the period 1995-2011.

Similarly, the trade determinants coefficients are typically assumed to be constant over time. Also this assumption must not necessarily be true, as the trade behavior could be subject to time evolvement and could also react to certain global events like e.g. natural disasters or economic crises. We therefore also account for the option what if trade determinants have been evolving in time, as scrutinized by e.g. Bleaney & Neaves (2013), Yotov (2012), Chaney (2013, June), Berthelon & Freund (2008), Disdier & Head (2008), Coe et al. (2007) or Brun et al. (2005). The most glowing example of this is definitely the distance (representing transport costs), of which importance must have fallen in the past decades. Not only the price affordability of high-distance transport, but also the accessibility to this service and its impediments to speed has can be presumed to have been increasing. As will be argued in the body of this work, this phenomenon is not always confirmed by the empirical studies and we will try to find some of its evidence in our data using a bit different approach.

According to our previous experience with estimations, we found out that the estimated functions and their outcomes actually significantly differ depending on the estimation technique employed. The problem is that models of gravity, spanning across most diverse countries of the world, are highly extensive in their coverage of factors and their role is to capture very complicated behavioral patterns of trade decision-makers that are not fully consistent with the econometric assumptions of BLUE (best linear unbiased estimators). A generally recommended and very commonly employed estimator is fixed effects model (Head & Mayer, 2013), of which main advantage lies in its ability to account for unobserved individual effects of each trade partner. Unfortunately, this is done on the basis of time demeaning which results in the elimination of time-invariant variables and the inability of the model to obtain estimation of those coefficients. Distance, common border, language, landlockness, or colonial link are typical examples of eliminated variables. Their impact is, however, of particular interest as well and this is why we compare standard fixed effects with alternatives like Poisson PML estimator, Mundlak model and Hausman Taylor estimator – that are able to give results for time-invariant variables as well.

As was already pointed out by numerous research papers, various estimators also differ in their sensitivity to assumption violations, missing observations or number of observations, leading to different results. These data features can adversely impact the reliability of the estimations or affect the significance and magnitude of the coefficients, which makes some estimation results invalid (Goméz & Milgram, 2009). We are thus going to assess the applicability of these various alternative estimators on our dataset, based on data characteristics and the reliability of the estimation obtained. By this experiment we would like to point out that we cannot fully rely on just one estimate and draw the quantitative conclusions precisely. In other words, the elasticities and trade potentials can never be known as exactly as they are numerically estimated, including their confidence intervals.

1 GRAVITY MODELS IN TRADE

As suggested by the name, gravity equations model bilateral trade flows depending on the size and distance effect, analogically to the Newton's gravity formula. The models have become an extremely popular tool for applied trade analysis as it provides us with quite intuitive results. The outcomes actually seem to lay down certain stylized facts about the determinants of bilateral trade (Head & Mayer, 2013).

There are several reasons, why gravity modeling of trade has become that widespread in the past years. Firstly, international trade flows are a key element in all manner of economic relationships and decision making, there is thus a large demand for assessing the normal or potential trade flows. Secondly, the data needed for such analysis are easily accessible and thirdly, there exist already a number of respected papers using gravity models, which establishes a set of standard practices (Baldwin & Taglioni, 2006).

The gravity models are not being used for analysis of trade determinants solely. They appear to have rather several useful attributes – gravity model is a popular tool for analysis of trade liberalization (e.g. Egger & Pfaffermayr, 2004), for discussion of so-called home bias (McCallum, 1995) or the effects of currency unions on trade (Rose, 2000). Further applications estimate gravity models for trade in services (Kimura & Lee, 2006), intra-firm trade (e.g. Egger & Pfaffermayr, 2005 Dec), and FDI (Egger & Pfaffermayr, 2004). For a more extensive description of gravity models' development, see e.g. Davidová (2012).

1.1 Model Specification

In contrary to the previously mentioned studies, we are going to employ the gravity model in a bit different way. Usually, the data set contains N countries trading with each other over a period of length T. This yields all together $\binom{N}{2} \cdot T$ observations – each for one trading pair and each year. As Baldwin & Taglioni (2006) notice, you have actually four numbers for each bilateral trade flow – two export records and two import records. In theory, the export number should correspond to its respective import number at the export partner. This is, however, not always true due to different methodology of measurement.

The authors state that there is an old tradition of using import data only on the grounds that nations spend more on measuring imports than exports in order to avoid tariff fraud. This is, nonetheless, not true anymore for the EU, as trade data is gathered from the VAT statistics since 1993. What do the authors in the literature usually do is that they average all the four bilateral trade numbers. Baldwin & Taglioni (2006) explain why this average should be geometric (i.e. sum of the logs) instead of arithmetic (log of the sums), even though most authors do the latter. In case of perfect data balance (export = import at each pair), the averaging actually does not matter. This is usually not the case, though. E.g. the dataset of Baldwin & Taglioni (2006), crafted to that of Micco et al. (2003), contains only 6% of country pairs of which the trade balance is close to zero.

Actually such an averaging of the export/import data eliminates very important information that is the exported amount per country, this distorts the model and the researchers are then able to obtain only a very general trade function. As we would like to avoid similar generalization, we are going to employ a one-way model, by which we mean that only one country is selected as a home country and we measure its export diversification among its trade partners. In this way we arrive to a concrete export function of one specific country only, which is, as we believe, much more concrete, further applicable and reliable estimate. However, even after such a contraction, we will see that there is still left sufficient space for the test of the heterogeneity of behavioral patterns.

One should keep in mind that the export and import decision makings are rather different and that these processes might be subject to different trade determinants. This was confirmed by my previous analysis of import and export functions of Austria, where different variables play a significant role.[1] Therefore, we do not regard averaging import and export data as a good idea and we will treat them rather separately.

[1] See Davidová (2012), Davidová (2013), and Davidová& Benáček (2013) for more details.

Our model is thus going to take the following log-log form:

$$\log(X_{it}) = \beta_0 + \beta_1 \cdot \log(GDP_{it}) + \beta_2 \cdot \log(GDP\_AT_t) + \beta_3 \cdot \log(D_i) + \varepsilon_{it}$$

with i being an index for partner country and t a time index. Note that the depending variable has only two dimensions. This of course diminishes the dataset but on the other hand, it makes the equation applicable for the home country, i.e. Austria, in particular.

1.2 Log-linear Gravity Model under Heteroskedasticity

As pointed out by Santos Silva & Tenreyro (2006), the estimates of the log-linear gravity equation are biased in case the data suffer from heteroskedasticity in error terms. Santos Silva & Tenreyro (2006, p. 653) have remarked the essential point that: "*the log-linearization of the empirical model in the presence of heteroskedasticity leads to inconsistent estimates because the expected value of the logarithm of a random variable depends on higher-order moments of its distribution*". If we have a look at the gravity equation in terms of expected values, we come to the following:

$$\mathbb{E}\left[\log\left(X_{ij}\right)\right] = \mathbb{E}[\log(\beta_0)] + \cdots + \beta_3 \mathbb{E}\left[\log\left(D_{ij}\right)\right] + \mathbb{E}\left[\log\left(\mu_{ij}\right)\right]$$

The conditional distribution of X_{ij} is altered due to Jensen's inequality:

$$\mathbb{E}\left[\log(\mu_{ij})\ \right] \neq \log\left[\mathbb{E}\left(\mu_{ij}\right)\right].$$

Hence, the variance of our estimated coefficients would be biased and our estimates inefficient. Moreover, in the presence of heteroskedasticity (as usual in trade data) the coefficients' estimations would be biased as well.

Santos Silva & Tenreyro (2010) actually suggest an alternative estimation, called the Poisson Maximum Likelihood Estimator (PMLE). This method was employed among others by Santos Silva & Tenreyro (2006), Westerlund & Wilhelmsson (2011), Martinez-Zarzoso et al. (2007), Babecká Kucharčuková et al. (2012) and Martin & Pham (2008). Poisson estimation is not only able to account for the bias caused by the logarithmic form of the gravity equation in case of heteroskedasticity in the error term. It, moreover, solves for zero trade flows as well. Santos Silva & Tenreyro (2006) tested this estimator against other methods and found its performance satisfactory even in the presence of measurement errors in the explained variable.

1.3 Panel Data

As Egger (2002) pointed out, the framework in 1990s was in most cases cross-sectional analysis – e.g. Wang & Winters (1991), Hamilton & Winters (1992), Brulhart & Kelly (1999) or Nilsson (2000). Newly, the authors tend to make use of panel econometric methods, as e.g. Baldwin (1994), Gros & Gonciarz (1996), Mátyás (1997), Egger (2000), Wall (2000), or Rose & van Wincoop (2001) did. In the one-way panel estimation there is a quite substantial and often neglected issue by the econometricians, that the interpretation of coefficients must be either related to cross-sections or to time-series. The simultaneous two-way estimation (with dummies for time and space dimensions present simultaneously) is appropriate only if the tests of poolability in time and space do not preclude it.

In the applied part, we are also going to take advantage of panel data structure, as common among researchers The advantage of panel data over simple cross sections or simple time series is that it is more informative (data encompasses more variability, less collinearity) and thus allows for more degrees of freedom, which in turn makes the estimates more efficient. Moreover, longitudinal data allow for controlling for individual unobserved heterogeneity, which is a very relevant question at trade data (Brüderl, 2005).

Actually, Goméz & Milgram (2009) warn that the results from cross-sectional data may vary substantially depending on the group selected, which leads to an estimation bias. Also Serlenga & Shin (2007) point out that cross-section estimation is misspecified since it is not able to deal with bilateral heterogeneity, which is extremely likely to be present in bilateral trade flows. In this regard, a panel-based approach is desired as heterogeneity issues can be controlled for by including individual effect dummies.

Máthyás (1997) also claims that the correct econometric specification of the gravity model should contain time, exporter and importer effects specified as fixed and unobservable. Egger & Pfaffermayr (2003) further demonstrate that this triple-way model can reduce to two-way model, including time and country effects only. This paper is going to employ panel data and add both time and country effects to the equation, as in line with the above-stated theory.

1.4 Model Non-Stationarity

Zwinkels & Beugeldijk (2010) point out that whereas usage of panel data enables to track development over time and increases efficiency of estimates, one also needs to take into account issues arising from time-series econometrics. Whenever we regress two trending time series (or time series integrated of order one, e.g. a random walk) on each other, spurious regression can arise. This is a problem – even though the variables are not subject to any causal relationship, the estimates appear to be highly significant and the whole model suggests great fit with a very high R^2 (Wooldridge, 2012). Of course, spurious regression is something we want to avoid in any empirical analysis, since it completely distorts the results. Unfortunately, some of the variables usually included in the gravity models tend to be non-stationary.

In particular, export data and both domestic and partners' GDPs are variables typically integrated of order one. Zwinkels & Beugeldijk (2010) also argue that these variables tend to be co-integrated. There is, therefore, space for such a spurious relationship (Engle & Granger, 1987; Hayashi, 2000).Although the spurious correlation problem is less important in panels than in time series analysis, as the fixed effects estimator for non-stationary data is asymptotically normal (see Kao & Chiang, 2000), the results of standard panel unit root tests are still biased (Fidrmuc, 2009).

Further, when standard estimators of gravity models are applied, they usually do not consider any possible endogeneity between trade and output, i.e. their long-run relationship. However, as demonstrated by Fidrmuc (2009), also the simple fixed effects model performs relatively well in comparison to panel cointegration techniques, like fully modified OLS suggested by Pedroni (1996, 2001) or dynamic OLS proposed by Kao & Chiang (2000). These methods deal with non-stationarity of the variables and with the long-run relationship between trade and output.

Nonetheless, Pedroni's panel cointegration tests confirm a good performance of gravity models under the presence of the cross-sectional correlation. It seems that fixed and time effects and the estimation of heterogeneous slopes in the long-run gravity models help to deal with cross-sectional correlation typical for gravity models (Fidrmuc, 2009). Since panel cointegration techniques[2] are very close to those obtained from fixed effects model, that is predominantly used in the recent literature, the possible bias due to the non-stationarity of gravity models is rather small. Hence, we do not consider trending variables to distort our results systematically. Actually, the fixed effect model performs well and deals with data non-stationarity, cross-correlation and endogeneity in a very similar manner than the specialized estimation techniques.

In addition, Zwinkels & Beugeldijk (2010) also admit that macro-economic variables like GDP and trade are trending relatively slowly. Therefore, they state, the problem of non-stationarity is less of an issue in short time series. This is in fact the case of our panel data, of which cross-sectional dimension is much larger that the time-series one. Admittedly, non-stationarity could be an issue in the sense that later periods would overly influence the estimation procedure, having higher values and thus higher weight. This could lead to biased and inefficient estimates. Dividing the sample into sub-periods, as suggested by Zwinkels & Beugeldijk (2010), solves for this problem well as it shortens the time-periods even further. Data sub-sampling is presented in the second part of the empirical section.

1.5 Distance Puzzle

It is no surprise that the world has become more trade-open and that the markets are getting more integrated. As stated in the World Bank's 1995 World Development Report, "*This globalizing trend has been driven by breakthroughs in transportation, communications and industrial technology, and above all by the opening of national markets to international trade*" (p. 8). At the same time, the gravity literature more than often fails to confirm this phenomenon by the empirical results. There is no consistent

[2] Like DOLS and FMOLS

empirical support for the globalization and for diminishing effects of distance on bilateral trade (Yotov, 2012).

Disdier & Head (2008) scrutinize trends in variation of 1,467 distance estimates from 103 papers in their meta-analysis study; they still could not find any significantly diminishing effect of distance to bilateral trade flows. They even conclude that "*the estimated negative impact of distance on trade rose around the middle of the century and has remained persistently high since then*". This result is now commonly known in literature as *the distance puzzle* and the authors are not able to get rid of it even after controlling for many important differences in samples and methods. Actually, Coe et al. (2007, p. 3) wittily add that "*globalization is everywhere but in estimated gravity models*".

Actually, despite the apparent technical progress in the transport industry, distance appears to be more rather than less important as a trade determinant. This finding was also supported by Berthelon & Freund (2008), Brun et al. (2005) or Leamer & Levinsohn (1995). Interestingly, this might be connected to log-linear specification of the model which has difficulties with zero or near-zero observations. For example, Coe et al. (2007) find that when dependent variable is not in logarithmic form, distance effects are indeed declining.

Yotov (2012) proposes a simple solution to distance puzzle and shows that the diminishing effect of distance is actually present in the data for both emerging and developed countries. The author claims that existing studies measure international trade costs relative to other international trade costs. Therefore, it is no surprise that the effects are stable over time assuming that the effects of globalizations are spread evenly among the different trading partners.

On contrary, when Yotov (2012) measured the effects of distance relative to the corresponding effects within national markets, the negative impact of distance on international trade has decreased over time. Specifically, the effect diminished significantly over the period 1965-2005 by 37% according to Poisson PMLE and by 28% according to simple OLS. This decrease was found at both poor and rich countries, even though the globalization effect is more marked at the poor ones.

Also Bleaney & Neaves (2013) have contributed to the debate that transport costs have declined over time and thus distance effects on trade should decrease as well. These authors investigate countries' openness to international trade (that is the ratio of exports plus imports to GDP) and have come to logical conclusion that trade is inversely dependent on geographical remoteness, land area and lack of access to sea. As these variables are most probably negatively correlated with transport costs and, in addition, effects of remoteness and land area have declined over time, the conclusion is that the effect of transport costs on trade has declined as the transport costs themselves declined. In other words, countries with high values of proxies for transport costs (remoteness, land area, sea access) have experienced an increase in openness relative to those with low values, ceteris paribus. The authors follow standard country fixed effects regression to investigate time trends and the evolution of the transport cost coefficients over time. This paper is quite similar to Guttmann & Richards (2006), who, nonetheless, do not deal with the time-dimension of trade openness or time variation in cross-section effects.

However, Chaney (2013, Aug) has come up with a different approach that the impact of distance on trade may be immune to changes in the technology for trading goods, in the types of goods traded, in the political barriers to trade or in the set of countries involved. Quite on contrary, he claims that distance effect on trade has been stable over time and should be close to the elasticity of -1. Chaney's explanation is based on the emergence of a stable network of input-output linkages between firms which is in his model presented by two assumptions that were in fact confirmed empirically: (i) the impact of distance on trade depends on the shape of the distribution of firm sizes. Chaney's dataset suggested that firm sizes could be well approximated by Zipf's law. (ii) Another assumption is made that larger firms tend to export over longer distances on average, as over time, a firm acquires more suppliers and customers, which tend to be further away. These two factors in turn imply distance elasticity of trade close to -1, which is in line with commonly achieved empirical results. Still, one may argue that also the distribution of firm sizes has shifted towards the larger ones, which results in change in distance elasticity.

One of the aims of this analysis is actually also to observe the change of the respective trade determinants over different time periods and across different country groups. As our model is designed in a simpler way and we offer 17 observations for each partner country, the year-by-year regressions actually make sense so that we could reveal the possible presence of (missing) distance puzzle in our data.

1.6 Alternative Models of International Trade

Besides the gravity model in trade, that links export and import to those traditional determinants GDP and distance, there have emerged many alternatives how to model exports of a country. This is usually done on the basis of augmenting the gravity model by additional independent variables or by modification of the dependent one. This is, nonetheless, still usually called gravity model of trade, as the trade-GDP-distance relationship is kept. On the other hand there have been a number of alternative approaches that disregard those traditionally used variables and model the export/import in the basis of a different idea. Let us just mention several of such experiments in order to emphasize that the gravity model might still disregard some deterministic behavior of international trade, as actually any other model does as well.

One example of such alternative approach is Chaney (2013, Aug), who has come up with idea of liquidity constrained exporters. He models international trade while taking into account entry costs that firms must pay in order to access foreign markets. In case firms lack sufficient liquidity, they are unable to export, as they cannot access financial markets and cover entry costs. As a result, more productive and wealthy firms that inherit large amount of liquidity are more likely to export. Chaney therefore claims that total amount of liquidity and its distribution matters for the behavior of exporters and may thus modify the pattern of aggregate exports. Actually, access to liquidity creates artificial links between different markets and generates export amplification. As soon as a firm starts exporting to a new foreign market, it is able to generate additional liquidity from this export. This may allow the firm to enter more foreign markets in the future.

Another example is an innovation of the gravity model presented by Anderson & van Wincoop (2003). These authors employed a parametric normalization to derive their empirical gravity equation, determining preferences conditional on observables. Holding preferences and the nominal unit of measure constant, their gravity equation deviates from the original trade equation by an endogenous factor. Anderson and van Wincoop derivate the nominal trade equation subject to the unit expenditure function that indicates price index as a function of prices and trade costs obtaining the following form:

$$x_{ij} = \frac{y_i y_j}{y^W} \cdot \left(\frac{t_{ij}}{P_i P_j}\right)^{(1-\sigma)}$$

Where x_{ij} represents the nominal trade flow from region i to region j, P_i and P_j represent the regions price index and y_i, y_j and y_W represent income of region i, j and the world, respectively. Distance and border-related transportation costs are further represented by the factor t_{ij} and σ stands for the substitution parameter of CES preference function.

2 ESTIMATION METHODS

Currently, there exists no econometric estimator that would strictly dominate all the others. Each method brings some advantages and disadvantages: some of them solve the heteroskedasticity or the zero problems but are too costly, whereas other simpler methods are not useful in the presence of those two characteristics, or do not take into account the multilateral dimension of trade. For that reason, it becomes a frequent practice in the literature to include several estimation methods using the same database, in order to check which one performs better (Goméz & Milgram, 2009; Head & Mayer, 2013). We also employ such approach in this thesis.

The estimation and consequent interpretation of gravity equations for bilateral trade involves a careful consideration of the theoretical underpinnings since it has become clear that naive approaches to estimation lead to biased and frequently misinterpreted results (Head & Mayer, 2013). As the authors correctly argue, there are more theory-consistent estimation methods and sole reliance on only one of them is not advisable. One should employ a toolkit approach instead to establish robustness. The estimation via different methods has actually become just a first step before a deeper analysis of the implications of the results.

The estimation techniques of gravity models were plenteously discussed over the past several years and we list and describe several commonly used estimation methods applicable to panel data. We are going to take advantage of panel data structure, as common among researchers, see e.g. Wall (2000), Egger (2000), or Rose & van Wincoop (2001). The drawbacks of simple cross-sectional or time series analysis were explained in the previous chapter.

Generally, the pooled OLS, the fixed effects model and the random effects model have been widely used in many previous studies, in various contexts. As pointed out by Serlenga & Shin (2007), the assumption that unobserved individual effects are uncorrelated with all the regressors is convincingly rejected in many of them, which makes the fixed effects estimation the preferable one, in order to avoid potential bias, as in e.g. Cheng & Wall (2005).

2.1 Pooled OLS

Gómez & Milgram (2009) state, that multiplicative gravity model was traditionally linearized and estimated using OLS techniques. This can be done under the assumption that the variance of the error is constant across observations (homoskedasticity). Pooled OLS estimation method applicable to a pooled dataset is based on the simple OLS:[3]

$$y_{it} = \alpha + \beta_1 x_{it1} + \beta_2 x_{it2} + \cdots + \beta_k x_{itk} + \nu_{it}, \quad t = 1, \ldots, T$$

$$\nu_{it} = a_i + u_{it}$$

The pooled OLS is a biased and inconsistent estimator, whenever a_i and $\mathbb{X}_{it}$ are correlated. This is true even under the assumption that the composite error ν_{it} is uncorrelated with $\mathbb{X}_{it}$. The resulting bias in pooled OLS is sometimes called heterogeneity bias, as it is caused by omitting a time-constant variable (Baltagi, 2008), which is attributed to a reliance of pooled OLS solely on a between comparison (Brüderl, 2005). Also Monte-Carlo simulations by Head & Mayer (2013) showed that pooled OLS is a poor estimator under the structural gravity data generating process. Their estimates of explanatory variables were biased towards zero and this method was not robust to non-complete sample.

Moreover, in the case that$a_i$is uncorrelated with all right-hand-sided variables in each time period, the pooled OLS standard errors and test statistics are generally invalid. They in fact ignore the substantial serial correlation in the composite error (Wooldridge, 2012). Also Egger (2002) argues that convenient OLS estimates are very likely to result in inconsistent estimates, as the most important dimensions of variation are importer and exporter effects. He adds that a conclusion should not be based on simple OLS estimates, as consistent estimation is a must.

Pooled OLS should be regarded as a simple benchmark which more complex models can be compared to or based on. This method actually disregards the panel data structure as it compounds the observations from the cross sections and time series all together. Pooled OLS can thus remedy the problem of unobserved heterogeneity across neither the countries, nor the years. Still, many early influential papers (e.g. Rose, 2000

[3] See for example Wooldridge (2012) and Baltagi (2008) for further explanation.

or McCallum, 1995) employ pooled OLS specification, disregarding the fact that this estimator ignores heterogeneity among countries. This neglect can lead to distorted estimates, as shown by e.g. Cheng & Wall (2005). The pooled OLS estimation method is not employed in our empirical part as it does not provide us with reliable results. It is, nonetheless, a start-up line for other estimators, e.g. the fixed and random effects models.

2.2 Fixed Effects

Theory consistent estimation of gravity model should always account for the multilateral resistance terms that are a key feature of general and structural gravity. Historically, the very first approach was to proxy multilateral resistance terms with remoteness terms. As soon as this approach appeared too weak, researchers switched to more structural approaches. One of such is also fixed effects estimation that accounts for individual unobserved fixed effects at each country. Harrigan (1996) appears to be the first one who employed fixed effects in his model.

The fixed effects model uses a transformation to remove the unobserved effect a_i from the equation so that it could be estimated by simple OLS. This is done in the manner that the averages are subtracted from the actual values and as the fixed effect a_i equals to its average over time, we get rid of it. So, the final equation has the following form:

$$y_{it} - \bar{y}_i = \beta_1(x_{it1} - \bar{x}_{1i}) + \cdots + \beta_k(x_{itk} - \bar{x}_{ik}) + (u_{it} - \bar{u}_i)$$

$$i = 1, \dots, N$$

$$t = 1, \dots, T$$

The β–coefficients are estimated by pooled OLS. The key assumption in this model is that the fixed effects model allows for arbitrary correlation between a_i and the explanatory variables in any time period $Corr(a_i; \mathbb{X}_i) \neq 0$. Hence, we are essentially assuming that our subjects of interest and their variances are identical.

However, the FE estimator does not work for constant variables across time, as these are ruled out by the averaging within transformation. This is the case for e.g. the distance variable that is actually one of the core regressors of the whole model. There still exists a modification of this estimator that allows for including also time-invariant variables called Fixed-effects Vector Decomposition introduced by Plümper & Troeger (2007).

Another pitfall of the fixed effects model is that whenever the number of geographical entities N gets large, it leads to a great loss of degrees of freedom. Moreover, inference made by the fixed effects estimator is more sensitive to non-normality, heteroskedasticity, and serial correlation in the idiosyncratic errors (Wooldridge, 2012). Still, the fixed effects model is very commonly used in the empirical work, as it is able to account for individual effects, sometimes called “resistance terms”. Head & Mayer (2013) emphasize that estimating gravity equations with fixed effects, as is now common practice and recommended by major empirical trade economists, does not involve strong structural assumptions on the underlying model. As long as the precise modeling structure yields and equation in multiplicative form, using fixed effects will yield consistent estimates of the fixed effect components.

There is also another additional advantage of using country fixed effects over other estimation methods. Head & Mayer (2013) correctly claim that there can be systematic tendencies of a country to export large amounts relative to its GDP and other observed determinants. They suggest the Netherlands and Belgium as an example, as much of Europe’s trade flows through Rotterdam and Antwerp. Sure that in principle the production location should be used as the exporting country and the consumption location as the importing country. However, this becomes vexed in reality, when use of warehouses and other reporting issues makes this difficult so there is reason to expect that trade flows to and from such countries are over-sated. Fixed effects can control for this, since they will account for any unobservable that contributes to shift the overall level of exports or imports of a country.

2.3 Random Effects

Unlike to the fixed effects model, the unobserved fixed effects are sometimes not likely to be correlated with the $\mathbb{X}_i$ independent variables matrix. Quite the opposite, the effects might be assumed to be distributed randomly across the units of interest. In such case, the fixed effects model is not a right choice as it yields inefficient estimates. That is why there exists also a random effects model that works under contradictory assumptions compared to the fixed effects model.

Contrary to the fixed effects model, the random effects one is based on quasi-demeaning – the transformation here subtracts only λ-fraction of the averages. By this we come to the following:

$$y_{it} - \lambda\bar{y}_i = \beta_0(1-\lambda) + \beta_1(x_{it1} - \lambda\bar{x}_{i1}) + \cdots + \beta_k(x_{itk} - \lambda\bar{x}_{ik}) + \nu_{it} - \lambda\bar{\nu}_i$$

$$t = 1, \dots, T$$

$$\text{where } \ \nu_{it} - \lambda\bar{\nu}_i = (a_i + u_{it}) - \lambda(a_i + \bar{u}_i) = a_i(1-\lambda) + (u_{it} - \lambda\bar{u}_i)$$

The quasi-demeaned equation is again estimated by pooled OLS. The drawback of the random effects estimator is that the parameter λ is never known and it can only be estimated. Therefore, we are never able to arrive to unbiased results. The value of the estimated transformation parameter $\hat{\lambda}$ actually indicates, whether the estimates are likely to be closer to the pooled OLS or the fixed effects estimates. It is evident, that for $\lambda = 1$ the RE estimator is identical to the FE one. On the contrary, as $\hat{\lambda} \to 0$, the RE estimates are be closer to the pooled OLS estimates.

The crucial assumption for the random effects model is the orthogonality of the individual effects and the regressors: $Corr\left(a_i, x_{itj}\right) = 0 \ \forall \, t, j$. This means that the unobserved heterogeneous component a_i is supposed to be randomly distributed with given mean and variance among the observed countries. In such case the fixed effects estimator would be inefficient (Wooldridge, 2012; Baltagi, 2008).

It is very important that we obtain white noise residuals as a consequence of a consistent and efficient estimator. Such residuals do not have any more systematic variation. The application of random effects approach is actually rather problematic due to the likelihood of its inconsistency as subject to correlation between some of the explanatory variables with the unobserved individual effects. Pure random effect model assumes absolutely no correlation of the individual effects with any of the regressors.

2.4 Hausman Test

Due to the contrary assumptions at FE and RE models, the researcher should decide for only of these methods. The fixed effects model assumes that individual groups have different intercept in the regression equation, while random effects hypothesize individual group have different disturbance. The complicated part is to select the one correct model – including either fixed or random effects as, strictly speaking, if one model is right, the other ones must be wrong.

Park (2010) suggests a simple scheme based on which one should decide for fixed or random effects model (or pooled OLS if data is poolable, i.e. no unobservable effect is actually present). Most importantly, you should always perform two basic tests that reveal whether the dataset contains fixed and/or random effects. F test (or Wald test) checks for presence of fixed effects, whereas Breusch-Pagan LM test can be used for random effects testing.

Technically, the tests should be run as follows: by the F test it is not meant the overall goodness-of-fit test reported by many softwares together with the estimate result. In order to test fixed effect, we fit the least squares dummy variable model with simple OLS and then test for the joint significance of the dummy variables for the groups. On the other hand, Breusch-Pagan LM test for random effects can be conducted by command .xttest0 in STATA.

Table 1: Model selection scheme

	Fixed effect (F test of Wald test)	**Random effect (Breusch-Pagan test)**	**Selected model**
H_O rejected? (type of effect present)	NO (no fixed effect)	NO (no random effect)	**→Pooled OLS** (poolable data)
	YES (fixed effect present)	NO (no random effect)	**→Fixed effects model**
	NO (no fixed effect)	YES (random effect present)	**→Random effects model**
	YES (fixed effect present)	YES (random effect present)	→ Choose FE or RE by the **Hausman test**

Source: Park (2010)

The choice of the model theoretically depends on the results of the stated tests. Still, panel data often leads to significant presence of both types of effects. In this case, we could use a model including both fixed and random effect. In theory, however, we are not allowed to do this as this is contradictory conceptually. FE and RE models have in fact contradictory assumptions about the correlation between the effects and independent variables. Park (2010) admits that it is possible to fit a model with e.g. a fixed group effect and random time effect (or vice versa) using both least squares dummy variable (LSDV) model and a random effect model. This possibility is, however, least recommended largely due to the loss of parsimony and degrees of freedom.

Still, distinguishing between fixed and random effects affects the final interpretation of the results. Actually, we associate different estimators with short-term and long-term time horizons, when comparing the results (Egger, 2002). Whereas fixed effects (and consistent random effects) model estimates reflect short run parameters, between model estimates are closer to long-run parameters. For more details, see Pirotte (1999).

The Hausman test states the following hypotheses:

H_0: $Cov(a_i; x_{it}) = 0 \quad \forall t$

H_A: $Cov(a_i; x_{it}) \neq 0$ for at least some t

Under the null hypothesis, both FE and RE estimators are consistent, RE estimator is, however, more asymptotically efficient. Under the alternative, FE estimator is still consistent, whereas the RE estimator is not. Hence, if we are able to reject the H_0, FE is the preferable one and vice versa.

Interestingly, the results of the Hausman test seem to be rather sensitive to the set of explanatory variables or country sample. The point is that the two compared models do not necessarily include the same set of variables. Therefore, the results of the Hausman test have to be interpreted carefully (Fidrmuc, 2009). As either fixed or random effects model is necessarily misspecified, we conduct the Hausman test also in this thesis in order to prefer one of the models.

Among others, Bleaney & Neaves (2013) also conducted a Hausman test in order to prefer either fixed or random effects model. In their case, random effects were always rejected in favor of fixed effects.

2.5 Poisson Estimator

The Poisson estimation technique is a log-linear pseudo maximum likelihood estimator, very popular one recently. Santos Silva & Tenreyro (2006), Martinez-Zarzoso et al. (2007), Westerlund & Wilhelmsson (2011), Martin & Pham (2008), and Babecká Kucharčuková et al. (2012) employed it, among others. Poisson estimation is able to account for the bias caused by the logarithmic form of the gravity equation in case of heteroskedasticity in the error term. Moreover, it solves for zero trade flows between two countries as well. Santos Silva & Tenreyro (2006) tested this estimator against other advanced methods and found its performance satisfactory even in the presence of measurement errors in the explained variable. Gourieroux et al. (1984) claimed that Poisson-PML estimation procedure is fairly easy to implement and robust to misspecifications. Another advantage of the Poisson estimator is that it allows for a continuous dependent variable.

Technically speaking, Poisson regression models are generalized linear models assuming the response variable y_{it} has a Poisson distribution, thus applicable to count data. The basic model takes the following form, assuring y_{it} being nonnegative:

$$y_{it} = \exp(\alpha + \beta_1 x_{it1} + \beta_2 x_{it2} + \cdots + \beta_k x_{itk}) + \nu_{it}$$

Let us switch to a matrix representation of the mean of the predicted Poisson distribution:

$$\mathbb{E}(Y|X) = \exp(\beta X)$$

β can be estimated by maximum likelihood method under specific conditions listed below. Generally speaking, maximum likelihood means that the estimator cannot be explicitly expressed by mathematical formula and must be found numerically. From Poisson distribution's probability mass function, we come to a formula which needs the probability to be maximized:

$$L(\beta|X,Y) = p(y_1,\dots,y_t|x_1,\dots,x_t;\beta) = \prod_{t=1}^{T} \frac{\exp(y_t\beta x_\mathrm{i})\cdot\exp[-\exp(\beta x_\mathrm{i})]}{y!}$$

This expression can be simplified by application of logarithm, which is a non-decreasing transformation that does not change the maximum. We thus come to log-likelihood function to be maximized instead:

$$l(\beta|X,Y) = \log[L(\theta|X,Y)] = \sum_{i=1}^{n}[y_i\beta x_i - \exp(\beta x_i) - \log(y!)]$$

Derivation of this equation with respect to β has no closed-form solution and is typically handled by a convex optimization. See e.g. Cameron & Trivedi (2013) and Winkelmann (2003) for more details on the Poisson regression for count data.

However, as first noted by Gourieroux et al. (1984), the data need not to be necessarily Poisson. Moreover, y_t does not have to be an integer at all for the estimator based on the Poisson likelihood function to be consistent. Actually, as pointed out by Silva & Tenreyro (2006), all that is needed for this PPMLE to be consistent is the correct specification of the conditional mean, i.e.:

$$\mathbb{E}(y_\mathrm{i}|x) = \exp(x_\mathrm{i}\beta)$$

Standard econometric packages thus allow for y_t to be non-integer. The dependent variable is usually inserted in linear form, while the right-hand-side variables come in logarithms, wherever elasticity desirable (typically distance and GDPs). Furthermore, the Poisson estimator does not take full account of the heteroskedasticity in the model. It actually calculates the covariance matrix, standard errors of the estimates and confidence intervals by a robust covariance matrix estimator developed by Eicker and White (Eicker, 1963; White, 1980).

As claimed by Babecká Kucharčuková et al. (2012), pseudo-maximum likelihood estimation techniques such as Poisson regression allow us to correct for biases resulting both from heteroskedastic errors and from missing trade between country pairs. Even though the Poisson estimator corrects for possible biases resulting from heteroskedastic error terms in log-linear specifications, it does not eliminate the need for correction of standard errors due to presence of heteroskedasticity. In compliance with customs, we shall employ Huber-White’s method.

Poisson pseudo-maximum likelihood regressions exhibit interesting and beneficial properties that can be very useful for the estimation of trade gravity equations. Specifically, the estimation of gravity with Poisson estimator while exporter and importer effects are included is consistent with a more structural approach (as in Anderson & van Wincoop, 2003) that imposes further restrictions on exporter and importer multilateral resistance terms (Fally, 2012). Moreover, it can be shown that whenever exporter and importer fixed effects are used in the Poisson regression, fitted output perfectly matches the observed output, which is a unique property to Poisson-PML (Fally, 2012). To sum up Fally's findings, we can say that multilateral resistance indices can be neglected, whenever Poisson estimator with fixed effects is employed.

2.6 Mundlak Model

Mundlak model was first proposed by Mundlak (1978). This model is based on random effects regression model, where group-means of variables that vary within groups are added to the set of independent variables. This technique was proposed as a way to relax the assumption in the random effects estimator that the observed variables are uncorrelated with the unobserved variables. Additionally, the degree of statistical significance of the estimated coefficients on the group means can be used to test whether such assumption holds for individual regressors (Perales, 2013). See also Chapter 10 in Wooldridge (2012) and Chapter 11 in Greene (2011) for further details.[4]

Mundlak's work (1978) is a follow-up to the discussion on fixed and random effects model. The author scrutinizes the decomposition of error term:

$$v_{it} = m_i + s_t + u_{it}$$

Where m_i and s_t are systematic components (or effects) associated with ith economic unit (country) and tth period (year). Mundlak raises a question which method of

[4]As there exists no in-built command for this model in Stata, we employ a module programmed by Perales (2013) downloadable from IDEAS at the Research Division of the Federal Reserve Bank of St. Louis.

estimation is correct to use. He argues that whenever random effects model is assumed, the consequences of the correlation, which may exist between the effects and the explanatory variables, remain completely neglected. Such correlation leads to a biased estimator. Mundlak (1978) shows that it can always be assumed the effects are random and that we can view the FE inference as conditional on the effects. Moreover, it is argued that when the effects are not correlated with the explanatory variables, the within and between estimators are the same. Therefore, any weighted combination of thereof will be the same. Mundlak assumes the following basic equation to be estimated:

$$Y = X\beta + Z\alpha + u$$

Where Z is a matrix of dummies and α is a vector of effects, for which he assumes orthogonality to the error:

$$\mathbb{E}(u'X) = \mathbb{E}(u'Z\alpha) = 0$$

In order to take an explicit relationship between the X's and the effects, Mundlak introduces the following auxiliary regression:

$$\alpha_i = X_{it}\pi + w_{it}$$

Parameter $\pi = 0$ only if the explanatory variables are uncorrelated with the effects. The auxiliary regression is subsequently averaged over t for a given i:

$$\alpha_i = X_i\pi + w_i$$

The Mundlak model was employed in an interesting way by Egger & Pfaffermayr (2005, Jan) who use it as an approximation of a general dynamic autoregressive distributed lag model for short and fat panels (i.e. a few time periods compared to many countries in our context). The authors explain that while the long run effects are mainly captured by between estimates, i.e. cross-sectional change in the data, the within estimates (i.e. fixed effects) represent short-run effects. Egger and Pfaffermayr demonstrate that disregarding the dynamic process by omitting the lagged endogenous variable results in an approximation error and in autocorrelated residuals.

The Mundlak model provides in fact both long run and short run parameter estimates. Egger & Pfaffermayr (2005, Jan) show that in the absence of a lagged dependent variable the Mundlak model is a perfect representation of a model with lagged exogenous variables, and the underspecified lag dynamics is fully compensated by the inclusion of the group mean as a control.

2.7 Hausman-Taylor Estimator

Although fixed effects estimation method has been widely accepted as theoretically correct, it has this drawback that time-invariant variables (such as distance or dummies of common border, language or colonial history) cannot be included due to multicollinearity with the intercept. Similarly to the Mundlak model, also Hausman-Taylor estimator deals with this weakness. Hausman-Taylor is an efficient instrumental variables estimation technique and it was introduced by its authors in 1981. This method allows us to obtain consistent estimation also of the coefficients of time-invariant regressors, similarly to previously introduced Mundlak model. It deals with the case when the individual effects are correlated with some of the time invariant variables and some of the $X's$. Valid instruments are given by the other time invariant and time varying variables in the equation (Krishnakumar, 2004).

The model is estimated in traditional form:

$$Y_{it} = \beta X_{it} + \gamma Z_i + \varepsilon_{it}$$

$$\text{with } \varepsilon_{it} = \alpha_i + u_{it}$$

Again, the error term ε is composed of ith individual effect α that might be correlated with variables X as well as Z and is assumed to be time-invariant random variable, distributed independently across individuals.Second, u is a zero mean idiosyncratic random disturbance uncorrelated across cross section units and over time periods. u is also assumed to have zero mean and constant variance, conditional on X and Z. We have reason to believe that:

$$\mathbb{E}(\varepsilon_{it}|X_{it}, Z_i) = \mathbb{E}(\alpha_i|X_{it}, Z_i) \neq 0$$

Time invariant regressors such as distance, common language and common borders dummies are now included in time invariant Z. This matrix has each column filled by blocks of T identical entries (e.g. distance in kilometres). γ is vector of coefficients to be estimated. These coefficients are impossible to estimate by a within estimator, i.e.

fixed effects model, where constant variables are wiped out by demeaning. In addition, Hausman &Taylor (1981) also argue that fixed effects estimate of remaining β coefficients is not fully efficient, since it ignores variation across individuals in the sample.

Another possible approach in the simultaneous equation model is to find instruments for those variables in X and Z which are potentially correlated with α. It is however usually difficult to find a suitable and theoretically correct instrument. An instrumental variable needs to be correlated with individual specific variable but not with unobserved individual effects. In any case, employing an external instrumental variables means ignorance of the time-invariant characteristic of the latent effect variable α.

Instead, Hausman-Taylor method uses those variables from X-matrix that are uncorrelated with α in order to: (i) produce unbiased estimates of the β's by deviations from individual means, and (ii) provide valid instruments for the columns of Z that are correlated with α by using individual means. Because the only component of the disturbance which is correlated with an explanatory variable is time-invariant, any vector orthogonal to a time-invariant vector can be used as an instrument. In other words, Hausman-Taylor estimator actually generates its own instrumental variables from time-variant ones. Still, one needs to be quite careful in choosing among the columns of X for those variables which are uncorrelated with α. Authors emphasize that such non-correlation can be tested under certain circumstances, so that the method does not rely on a priori assumptions

Mathematically expressed, the estimator is yielded when simple OLS is applied to:

$$P_A\widehat{\Omega}^{-1/2}Y_{it} = P_A\widehat{\Omega}^{-1/2}X_{it}\beta + P_A\widehat{\Omega}^{-1/2}Z_i\gamma + P_A\widehat{\Omega}^{-1/2}\varepsilon_{it}$$

This is just a transformed version of the model equation by $P_A\widehat{\Omega}^{-1/2}$ where

(i) P_A is the orthogonal projection onto column space of matrix $[X_{1it} \vdots Z_{1i}]$[5]

(ii) Ω is a disturbance covariance matrix $\Omega \equiv cov(\varepsilon_{it}|X_{it}, Z_i)$ and $\widehat{\Omega}$ any consistent estimate thereof.

[5]For more technical background on theoretical derivation of the estimator see original paper of Hausman & Taylor (1981).

Hausman & Taylor (1981) claim that their method does not assume a specification of the components of α and may be less sensitive to potential lack of knowledge about the unobservable individual-specific effect. To summarize, Hausman & Taylor have developed a consistent and efficient estimator that on the hand side accounts for unobservable individual effects, and on the other hand provides an estimate also of time-invariant variables.

For example, Brun et al. (2005) applied Hausman-Taylor estimator. They used infrastructure and population as instruments for standard trade-barrier function such as distance, common border and common language dummies, assuming these are not correlated with individual effects. Next, Egger (2002) takes advantage of this estimation method and compares it with fixed and random effects, among others. Egger also complaints that time-invariant variables cannot be estimated by fixed effects and makes use of the several dimensions of panel data. He tries to overcome the correlation of α bilateral effect with some explanatory variables without any variables from outside the model. Egger checks the model appropriateness on the basis of a Hausman and Taylor test for over-identifying restrictions.

Also Serlenga & Shin (2007) employed the Hausman-Taylor estimation generalized to heterogeneous panels with time-specific factors. They attempted to deal with exporter/importer heterogeneity which they consider to be likely present in bilateral trade flows. Serlenga and Shin came to an interesting finding that once the correlation between the common language dummy and unobserved individual effect is accommodated by the Hausman-Taylor estimation, there is evidence that the effects of geographical distance proxy variables (i.e. distance and common border dummy) might be mutually compensated, whereas the role of cultural affinities approximated by common language dummy becomes more significant.

To summarize, Hausman-Taylor method is preferable one in case it is consistent. This is testable by the over-identification test. If this test implies inconsistence, fixed effects model appears as the only valid alternative (Egger, 2002). In addition, simulation results show that the Hausman-Taylor model with perfect-knowledge about the underlying data structure (instrument orthogonality) has on average smaller bias than fixed effects. However, simple non-IV rival estimators performs equally well or even better compared in case of imperfect-knowledge and instruments chosen by statistical criteria only (Mitze, 2010).

3 DATA AND METHODOLOGY

Previous chapters provided a critical assessment of decision-making criteria of an econometrician who is challenged by the estimation of a gravity model based on panel data. Our next step is to illustrate the previous on an extensive empirical model of Austrian exports. Our dataset contains various macroeconomic, geographical and institutional variables in 211 countries over the period 1995-2011. There are of course several countries that had to be omitted from the analysis, mainly due to poor data availability. Still, the set of partner countries is very large compared to other gravity model studies, where usually only certain group of countries is included (e.g. CEE, EU or OECD member states that can be expected to be more homogenous than the full set of countries). Our extensive country group can be justified by the fact that one of the aims of this thesis is to estimate an overall export function of a particular country (Austria in our case) and to compare it to export functions estimated for smaller clusters of the partner countries (like rich EU countries, distant poor countries, etc.). As will be revealed in Chapter 4, the export determinants differ significantly in terms of both sizes and statistical significances of the respective variables across the clusters.

Instead of possible 3,587 observations, the reduced data set yields 3,396 complete observations due to some missing observations. The zero or rather missing export observations do not present a significant part of our sample (namely 5%). They are thus being treated in the most conventional way – i.e. omitting the observation in case of its incompleteness. We are aware that this solution is not optimal in the case when the missing observations account for a more substantial proportion of the sample due to possibility of selection bias. In our case, however, omitting of incomplete observations leads to reasonable outcomes that are in line with both common-sense expectations and with usual gravity model results presented in the literature.[6]

[6]For comparison of different zero and missing observations treatments please see Davidová (2012).

3.1 Variables Description

The core idea of the gravity model is to model the bilateral trade on the distance and size of the two trading countries, analogically to the Newton's law of universal gravitation. In our model, the dependent variable is thus the exported value of Austria, whereas the main independent variables are GDPs of the two countries and their mutual distance.

However, as argued in many gravity model studies, such model would be too simplistic, as it does not account for other barriers to trade except for the distance. Therefore, our model is augmented by other distance-measuring variables (landlockness, contiguity, and trade unions memberships), institutional indices (government effectiveness index), economic variables (common currency and recession dummy), and last but not least by common language and common colonial history dummies. The complete list of our variables is presented in Table 2below. For detailed variables description and the methodology, see Davidová (2012).

Table 2: Variables description

		Values	Unit	Source
1.	log(export) as dependent variable		EUR	Eurostat
2.	log(GDP partner)		mil. EUR	IMF
3.	log(GDP Austria)		mil. EUR	IMF
4.	log(distance)		km	CPII
5.	common language	0/1	dummy	CPII
6.	common border	0/1	dummy	CPII
7.	common political history (colony)	0/1	dummy	CPII
8.	landlockness (no direct sea access)	0/1	dummy	CPII
9.	recession	0/1	dummy	own estimation
10.	trade barriers (EU, EFTA, WTO etc.)	0/1	dummy	own estimation
11.	government effectiveness	0-100	per cent	World Bank
12.	common currency (euro)	0/1	dummy	own estimation

Source: Davidová (2012)

3.2 Model Specification

In this thesis, a log-log version of the gravity model is employed in order to obtain some coefficients as elasticities (GDP and distance). The model is designed as follows:

$$\log(export_{jt}) = \beta_0 + \beta_1 \log(GDP\ partner_{jt}) + \beta_2 \log(GDP\ AT_t) + \beta_3 \log(distance_j) + \beta_5 language_j + \beta_6 contiguity_j + \beta_7 colony_j + \beta_8 landlockness_j + \beta_9 recession_t + \beta_{10} euro_{jt} + \beta_{11} barriers_{jt} + \beta_{12} gov\_effectiveness_{jt}$$

$j = 1, \dots, 211$ stands for the partner country

$t = 1995, \dots, 2011$ represents time (years)

Note that some variables do not have t index, since they are time-invariant (distance, language, contiguity, colony and landlockness). We have to omit these in the fixed effect estimation due to multicollinearity with the intercept.

3.3 Variables Hypotheses

Let us now have a look at the expected relationship between selected variable, according to previous published results or economic theory.

It is rather intuitive that exports rise proportionately with the economic size of the destination. In gravity modeling, we use GDP as a proxy of economic size, so the relationship is expected to be analogous. An interesting experiment is presented by Head & Mayer (2013) who observe relationship of Japanese exports and imports to each EU member state. They do this by simply plotting export and import value against GDP in 2006 and they found out a very good fit (85% and 75%, respectively). Moreover, the elasticity is very close to 1 (1.001 and 1.03, respectively). Actually the same experiment was repeated for years 2000-2009 and the average elasticity value was 0.98. As Japan is very distant from the EU and it does not share language, currency, border or colonial history, this result is expected to be purified from other influence. Also the meta-analysis by Head & Mayer (2013) supported this finding when the average coefficient at GDP of origin (i.e. the exporter) was 0.98 over 700 regressions included.

Balwin & Taglioni (2006) argue that it matters a lot which GDP measure is included in the model – whether real or nominal. They explain that the gravity equation is in fact an expenditure equation and it is therefore logical to relate the value of bilateral export to the value of importing nation’s expenditure measured in nominal numeraire converted to common international currency (e.g. USD or EUR) by current exchange rate. Using real GDP, as became common practice in gravity models, treats the gravity equation as if it was demand equation for the whole country. Still, some authors use the real GDP.[7] We do not follow that methodology in this thesis though by setting the GDP at current purchasing power standards. As explained in detail in Davidová (2012), measuring GDP in purchasing power corresponds to the country’s purchasing power at home, which also considers the tradeoffs between spending on imports and spending on domestic alternatives. So, indirectly, we included the opportunity costs of domestic agents in their expenditure. This specification, according to theory, must be complemented by capturing the exchange rate as a factor enhancing (via undervaluation) or abating (via overvaluation) the income from exports. Thus in our models the variables of GDP in exporter and importer countries were complemented by the variable of ERDI, which serves to straighten distortion between the exchange rate and the real purchasing power parities.

The expected effect of distance, trade barriers and recession is expected to be negative, while common language, border, currency or colonial history should enhance the exports. In addition, we also include governmental effectiveness index measuring the level of institutional development and effectiveness. Proper institutional background is expected to further support the international trade, as indicated by a study from Wu et al. (2011). These authors found that countries with better governance environment (called as rule-based countries) trade more than relation-based or family-based countries whose rule of law are weaker. The results of our estimation are discussed in Chapter 4.

[7] That is nominal GDP in national currency deflated by a national price index and then converted to US dollars at current exchange rate.

3.4 Clustering

In order to re-estimate our sample country group-wise, we need to split Austria's partner countries into groups. This is done on the basis of clustering. Similar approach was employed by e.g. Bobková (2014), Sarafidis & Weber (2014), Chang-Ching & Ng (2012), Sarafidis & Weber (2009), or Kapetanios (2006). Cluster analysis is able to re-group the sample into several sub-groups based on their common characteristics. To group the countries into clusters, we perform k-means clustering, which was originally introduced by MacQueen (1967). It is a popular algorithm allowing to partition N observations into k clusters. K-means clustering requires a fixed number of groups in which the countries are supposed to be divided. Based on combination of couple of stopping rules and the number of countries in each cluster, nine was chosen as an optimal number of the clusters.

We employ several important variables when running the k-means clustering procedure – partners' GDP, country distance from Austria, and government effectiveness. Austria's GDP and recession were obviously omitted due to zero variance among countries. Also trade barriers, common language, colonial link and common border dummies were not used in the cluster analysis. This is due to typically very low number of countries in the whole sample sharing language/political history/border with Austria. When these dummy variables were employed, it typically lead to overemphasis on these variables and very uneven split of countries into the clusters.

Since k-means clustering cannot take account for panel data, we transformed our panel into a cross-section. Moreover, we aim to overcome the problem of countries switching from one to other cluster in time. This was done on the basis of simple averaging of variables for each given country over the 17 years, which yields the typical characteristic of each partner country, based on the size of its economy (GDP), distance from Austria and its level of institutional development (government effectiveness index). The resulting clusters are described and discussed more in detail in the following chapter in section 4.2.

4 RESULTS

This chapter presents commented numerical results. We apply various estimation techniques that were discussed above to Austrian panel dataset of exports to 211 countries within 17 years (1995-2011). Firstly, we estimate the model under the time-series data structure using the whole dataset (3,396 observations). This enables us to compare the performance of different estimators on our particular dataset. Based on the data characteristics and assumption violations, some estimators are considered as unreliable, while others as preferable. The most suitable estimator appears to be the Hausman-Taylor estimator due to its ability to deal with data issues in the most comprehensive and reliable way.

Hausman-Taylor is thus employed in the consequent cluster analysis. In the first clustering experiment, we split the 211 partner countries into several sub-groups according to their trade characteristics using k-means clustering technique. Consequently, we re-estimate the model for each cluster separately in order to compare the results between the groups. We find significant difference in resulting export functions both in significance and magnitude of the coefficients. In the second clustering experiment, we split the 17 years period into several sup-periods in order to track the development of the trade determinants in time. We find significant impact of the economic crisis on Austrian export function as well as other time trends.

4.1 Various Estimation Methods

As already indicated in the theoretical part of the thesis, we are going to estimate the full model by the following four estimators: fixed or random effects, Poisson pseudo maximum likelihood (PPML) estimator, Mundlak model based on random effects and Hausman-Taylor estimator. These models were chosen due to their good performance despite data imperfection and assumptions violation. In addition, all except for the fixed effects allow for time-invariant variables, while still accounting for individual effect of each partner country. All four models have been extensively employed in recent gravity model literature, as already noted earlier in Chapter 2.

Table 3: Wald test for fixed effects (based on LSDV model)

Joint significance of LSDV dummy variables:
H_0: state_id_$i =$ 0 $\forall i$ i.e. no FE present
H_A: state_id_$i \neq$ 0 for at least one i i.e. FE present

F(205, 3179) = 28.12
Prob > F = 0.0000

$\rightarrow$ H_0 strongly rejected, i.e. FE present

Source: author's calculation

First, we had to decide between fixed and random effects model. The theoretical foundations of this dilemma were described in detail in section 2.4, referring mostly to Park (2010). Let us now apply the procedure in practice. As the initial step, we tested for the presence of fixed effects so that we estimated the Least Squares Dummy Variable (LSDV) regression by simple OLS. Consequently we applied the Wald test to check for joint significance of all dummy variables present in the model. We have obtained a very low p-value indicating strong joint significance and thus we conclude that fixed effects are apparently present (see Table 3).

Table 4: Breusch and Pagan Lagrangian multiplier test for random effects

log_export[state_id,t] = Xb + u[state_id] + e[state_id,t]

Estimated results:

	Var	sd = sqrt(Var)
log_export	11.860	3.444
e	0.686	0.828
u	1.490	1.221

Test: Var(u) = 0 (i.e. no random effects)
chi2(1) = 6547.3
Prob > chi2 = 0.000

$\rightarrow$ H_0 strongly rejected, i.e. RE present

Source: author's calculation

As the second step, we control for presence of random effects. This was done on the basis of Breusch and Pagan Lagrangian multiplier test for random effects (see Table 4). P-value very close to zero implies that we strongly reject the null hypothesis of no significant difference across units (i.e. no panel effect), which means that random effects are in fact present. This result in combination with previous test for fixed effects indicates that both types of effects are present in the data and that Hausman test needs to be applied.

Table 5: Hausman test

	-------------------	Coefficients	-------------------	
	(b)	(B)	(b-B)	sqrt(diag(V_b-V_B))
	fixed	random	difference	S.E.
log (GDP partner)	0.548	0.810	-0.262	0.032
log (GDP Austria)	1.515	1.107	0.407	0.048
recession	0.054	0.046	0.008	.
trade barriers	-0.397	-0.274	-0.123	0.043
gov. effectiveness	0.000	0.001	-0.001	0.001
euro	-0.330	-0.210	-0.119	0.057

Fixed (b) = consistent under H_0 and H_A; obtained from xtreg
Random (B) = inconsistent under H_A, efficient under H_0; obtained from xtreg

Test: H_0: difference in coefficients not systematic

$\chi^2(6)$ = (b-B)'[(V_b-V_B)^(-1)](b-B)
= 78.13
Prob > χ^2 = 0.000

→ H_0 strongly rejected, i.e. RE estimator is inconsistent

Source: author's calculation

According to Hausman test, the null of zero covariance was strongly rejected and thus the conclusion is that fixed effects are the correct model to apply over random effects (see Table 8). We conclude that random effects would be inconsistent and thus prefer a consistent estimator based on fixed effects instead. The preference towards FE as compared to the RE is in line meta-analysis of Head & Mayer (2013) and other recent literature.

When testing for assumptions, strong heteroskedasticity was detected in the data, as expected.[8] Therefore, we employed Huber-White's robust standard errors, whenever applicable. In fact, also non-normality of the data seems to be present according to the relevant tests.[9] We cannot easily solve for this imperfection and we will thus rely on the law of large numbers that ensures that the estimators will be consistent anyway. In fact, slight serial correlation is present as well, which indicates that some of the variance in export remained unexplained by the included variables. Nonetheless, the results of this test are not that convincing[10] and in combination with high explanatory values of the models we will consider this problem as minor.

Table 6: Complete sample – comparison of different estimators

log(export)	**Fixed Effects**	**Poisson PML**	**Mundlak**	**Hausman-Taylor**
Observations	3396	3396	3396	3315
F/Wald statistic	91.18	27,744	2,483	1,832
Prob > F	0.0000	0.0000	0.0000	0.0000
R^2	0.72	0.97	0.85	.
log(GDP partner)	0.55***	0.78***	0.55***	0.58***
log(GDP Austria)	1.51***	1.40***	1.51***	1.22***
log(distance)	(omitted)	-0.86***	-1.19***	-0.93***
language	(omitted)	0.39***	0.14	0.46
contiguous	(omitted)	0.70***	0.45**	1.18*
colony	(omitted)	0.23***	0.51	0.45
landlockness	(omitted)	-0.21***	-0.13	-0.38
recession	0.053	-0.08	0.05	-0.03
trade barriers	-0.40***	0.29***	-0.39**	-0.17
gov. effectiveness	-0.0004	0.002**	0.008	0.002
euro	-0.33***	0.09**	-0.33**	0.09**
constant	-7.2***	-0.33	15.87*	-0.13

Source: author's estimation

Table 6 presents an overview of our numerical results. Each column represents one estimator: fixed effects model, Poisson PML estimator, Mundlak model and Hausman-

[8]See results of White's test in the Appendix, Table 15.

[9]See results of tests for normal data in the Appendix, Table 16.

[10] See results of autocorrelation test in the Appendix, Table 17.

Taylor estimator. All four estimators were applied to the full data set, estimating a model of all 12 independent variables including constant. In fixed effects, distance, language, common border, colonial history and landlockness had to be omitted due to multicollinearity with the intercept. This is probably the reason for relatively lower $R^2 = 72\%$, compared to other methods. The dependent variable is logarithmized exported value (with exception of PPML, where the logarithm is not employed due to the log-nature of the estimator itself). All estimations are based on 3,396 observations except for Hausman-Taylor that has 3,315 observations. This is because we also included lagged export variable in order to avoid the endogenous issues like bidirectional causality between GDPs and trade. In Hausman-Taylor model, we treat export as dependent; GDPs and lagged export as time-variant endogenous; recession, trade barriers, government effectiveness and euro are considered to be exogenous time-variant; while distance, language, contiguity, colony and landlockness are treated as exogenous time-invariant in the model.

Table 6 indicates that all four models actually perform sufficiently well, pointing to similar behavioral characteristics for the core economic variables. All models have rather high joint F/Wald statistics indicating strong joint significance of the variables. This is further confirmed by very high goodness of fit measured as R^2. Poisson PML obviously outperforms its peers in terms of significance and fit, which, however, does not necessarily mean that the model is the most correct one. Actually quite on contrary – all FE, Mundlak and Hausman-Taylor signal the significance in a similar way, while Poisson indicates almost all variables as strongly significant which makes it more suspicious and less reliable.

The fact that all four estimators give in fact pretty similar results indicate that none of them is subject to a severe bias. The partners' GDP is strongly significant in all models, yielding elasticity of approximately 0.6-0.8. This number is in line with meta-analysis of Hear & Mayer (see Table 7) who find the median coefficient at GDP of destination to be 0.85 in all gravity analyses involved and 0.67 among structural gravity models. Also Austrian GDP appears to be a very significant driver of exports with elasticity of 1.2 to 1.5. This finding is again in line with general tendency that domestic GDP plays more important role than the foreign one, as indicated by Head & Mayer (2013). Still, the numerical result is rather too high in comparison with the existing research. One reason of that might be the discussed non-stationarity of GDP, which would be eliminated by the consequent cluster analysis. In case the GDP coefficients would be lower in time sub-periods, we could conclude that non-stationarity plays a role here.

Table 7: Head & Mayer meta-analysis results

	All Gravity				**Structural Gravity**			
Estimates:	median	mean	s.d.	#	median	mean	s.d.	#
Origin GDP	0.97	0.98	0.42	700	0.86	0.74	0.45	31
Destination GDP	0.85	0.84	0.28	671	0.67	0.58	0.41	29
Distance	-0.89	-0.93	0.40	1,835	-1.14	-1.10	0.41	328
Contiguity	0.49	0.53	0.57	1,066	0.52	0.66	0.65	266
Common language	0.49	0.54	0.44	680	0.33	0.39	0.29	205
Colonial link	0.91	0.92	0.61	147	0.84	0.75	0.49	60
RTA/FTA	0.47	0.59	0.5	257	0.28	0.36	0.42	108
EU	0.23	0.14	0.56	329	0.19	0.16	0.50	26
CUSA/NAFTA	0.39	0.43	0.67	94	0.53	0.76	0.64	17
Common currency	0.87	0.79	0.48	104	0.98	0.86	0.39	37
Home	1.93	1.96	1.28	279	1.55	1.90	1.68	71

Note: The number of estimates is 2,508, obtained from 159 papers. Structural gravity refers here to some use of country fixed effects or ratio-type method.

Source: Head & Mayer (2013)

The coefficient of distance is again strongly significant yielding almost negative unity elasticity across the three models (0.9-1.2), which is pretty in line with the theory outlined earlier. Also Head & Mayer find that median distance impact on trade in the literature is -0.9 (among 1,835 studies). Negative impact of distance was definitely expected due to increasing transport costs. It will be, however, interesting to observe the development of this coefficient in Austria across time and also across country-groups.

Contiguity is another consistently significant variable that has positive impact on Austrian exports, while trade barriers more or less consistently decrease the trade in line with economic theory. Language, colonial history and landlockness are significant only in the Poisson PML model, while playing no important role in Mundlak or Hausman-Taylor. This confirms our hypothesis that Poisson estimator in out paper tends to overestimate the significance of variables. Recession, government effectiveness and euro currency have ambiguous impact on trade according to our results.

Let us now compare the four selected estimation method and assess their performance on our data sample, so that we can proceed to data clustering only with one estimator in use. Fixed effects in fact perform well – the magnitude and significance of the coefficients is in line with other estimators. Also, even though this estimator is rather sensitive to data imperfections resulting in heteroskedasticity, non-normality and serial correlation in the residuals, FE model does not lay down strong assumptions on the underlying model and yields consistent estimates. This is why Head & Mayer (2013) recommend FE model and also why it became common practice among researchers. Its advantage is that FE model accounts for unobserved individual country-specific effects. The penalty for that is, however, the inability to estimate time-invariant variables including distance or contiguity. This is the major drawback and the reason why FE will not be our preferred estimator.

Poisson PML is believed to account for bias caused by logarithmic form in case heteroskedasticity is present (Santos Silva & Tenreyro, 2006) and the model is also pretty robust to misspecifications (Gourieraux et al., 1984). However, it does not seem to correspond to our data well, which is indicated by clearly overvalued significance level and this is why it would not be wise to rely on this estimator in the consequent analysis.

Mundlak estimator is in fact based on the random effects with added group-means in order to relax the orthogonality assumption. It is able to decompose errors to country-specific and time-specific and also provides both short-run and long-run estimates. In other words, this estimator in fact combines the estimation of data under the time-series data structure with cross-sectional data structure,[11] which makes the results rather clumsy to interpret. Otherwise, the estimator performs pretty well in terms of all fit, significance and magnitude of the coefficients.

Finally, we presented Hausman-Taylor estimator that also enables to account for country-specific unobserved effects, while yielding the consistent and efficient estimates for the time-invariant variables as well. Its great benefit is that it takes the advantage of instrumental variable technique and also accounts for bidirectional

[11] See Davidová & Benáček (2013) for further details on this topic.

relationship between the international trade and countries' GDPs. In addition, Hausman-Taylor estimator appears to be the most sober one in terms of assessing the significance which makes it the most trustworthy for risk-averse policy makers. From those three reasons we have decided to apply Hausman-Taylor estimator also in the consequent analysis.

4.2 Clusters by Countries

This section presents empirical results of re-estimation of the whole sample by groups of countries. The underlying hypothesis is that it is unrealistic to assume that exporters' behavior is homogenous. In fact, different characteristics of the partner country also modify the exporting function as various determinants gain or lose their significance. Following k-mean clustering methodology (further described in the methodology section 3.4), we have come to nine clusters, each consisting of 1-58 countries. Detailed list of countries in each cluster is presented in the Appendix in Table 18 and Table 19. Let us only summarize the general characteristics of each cluster with the help of Table 8 below that shows mean values of selected variables for each cluster together with number of countries included. In fact, our clusters turned out to be intuitively reasonable groups of countries that share many economic, political or geographical characteristics.

Table 8: Country clusters mean characteristics

Cluster	N	Export (EUR m)	Partner GDP (EUR bn)	Distance (km)	Barriers (dummy)	Gov. eff. (%)	Euro (dummy)
1	1	4,110	9,181	8,124	1.0	91	0.00
2	5	7,340	2,122	3,980	0.4	83	0.40
3	11	1,150	252	2,016	0.3	83	0.50
4	25	518	40	1,858	0.6	80	0.46
5	9	1,520	680	6,857	0.8	71	0.22
6	50	41	35	8,891	1.0	61	0.02
7	20	5	4	15,181	1.0	42	0.00
8	58	8	10	6,155	1.0	21	0.00
9	32	121	26	2,600	1.0	38	0.10

Note: Export, trade barriers and euro dummy were not included as determinants for clustering. They are presented for illustration purposes only.

Source: author's calculation

Cluster 1 consists of the USA only and we thus do not perform any separate estimation for this cluster. Cluster 2 is composed of 5 very big and rich economies with developed institutional background that are on average not that distant from Austria. 40% of the sample shares euro with Austria and trade barriers do not pose significant obstacle in trade with these countries which all together causes very significant trade flow. They are France, Germany, China, Japan and UK. Third cluster consists of 11 countries of which 9 are European.[12] These are rather rich country group geographically the second closest to Austria with advanced government effectiveness, lower trade barriers and either common or closely related currencies. This cluster includes among others Belgium, Denmark, Netherlands, Sweden, Norway or Switzerland.

Cluster 4 is composed of mostly medium-developed European countries as e.g. Croatia, Czech Rep., Hungary, Lithuania, Slovakia or Slovenia, including also some non-European, quite closely located, developed and mid-class rich economies. This cluster is geographically the most adjacent to Austria. Cluster 5 has only 9 countries that tend to be geographically distant from Austria, rich economies with relatively high barriers to the EU trade as e.g. Australia, Canada, Brazil, India, Mexico or Russia. Cluster 6 consists of 50 small and very distant economies with high level of barriers to trade and relatively above-average governmental effectiveness, as compared to the rest of developing world. Typical examples of countries included are from Latin America (Argentina, Bolivia, Belize, Colombia, Costa Rica, Chile, Peru, and Uruguay) or from South-East Asia (Hong Kong, Singapore, Philippines, Thailand or Malaysia).

Countries included in Cluster 7 tend to be very tiny economies that are the most geographically distant from Austria, with no preferential trade agreements and rather low level of institutional efficiency, which all in combination implies very low imports from Austria. This cluster includes e.g. Fiji, Falkland Islands, Marshall Islands, Micronesia, Vanuatu or Wallis. Cluster 8 is the most numerous one containing 58 countries that share the following common characteristics: very low GDP, very distant and extremely low level of governmental efficiency. These are mostly either African (Afghanistan, Congo, Djibouti, Ethiopia, Somalia or Yemen) or Asian countries

[12]Including Turkey

(Bangladesh, Laos, Mongolia, N. Korea or Nepal). Finally, Cluster 9 represents 32 quite poor but not too distant countries from Austria with low government effectiveness and no trade liberalization agreements, as e.g. Albania, Armenia, Bosnia, Bulgaria, Egypt, Jordan, Morocco or Ukraine.

Numerical results of our regressions are summarized in Table 9 below. Some dummy variables had to be omitted in several clusters due to multicollinearity (they did not vary across the sub-sample). Number of observations for Cluster 2 is rather low ($N = 85$), which, however, did not negatively impact the significance of the variables. Other regressions tend to have 150-950 observations, which offer sufficient degrees of freedom. Generally speaking, partner's GDP is the main determinant of exports across all clusters, yielding very significant results with coefficient varying from 0.3 to 0.6. Clusters 3 and 5 (i.e. rich, developed countries with often times trade agreements) show the lowest impact of GDP to the exports. Austrian GDP seems to play a very significant role as well – most strongly statistically significant coefficient with magnitude of 1.0-2.3 shows again that domestic GDP has greater power to determine the exports than the partners' GDPs. Only Cluster 7 (very poor and distant countries with almost no trade with Austria) does not indicate any significant relationship.

Table 9: Sample clustered by countries – estimation results

Country cluster	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
log (GDP partner)	0.62***	0.30***	0.52***	0.32***	0.49***	0.59**	0.59***	0.45***
log (GDP Austria)	1.45***	1.33***	1.6***	2.17***	2.31***	0.62	0.97***	1.23***
log (distance)	-7.47***	-0.39*	-1.14**	-1.22***	9.38	12.47	2.19**	-1.08***
language	-1.87***	0.06	-0.24	.	.	.	.	.
contiguous	.	1.44*	1.01	0.25	.	.	.	.
colony	.	.	0.43	.	.	.	.	0.35
landlockness	.	.	0.01	.	-1.79*	.	0.25	0.44
recession	-0.11*	-0.05	-0.17*	-0.03	-0.02	-0.14	-0.06	0.15
trade barriers	12.84***	-0.12	-0.21*	.	.	.	.	-0.06
gov. effectiveness	0.01**	0.02***	0.01*	0.02***	-0.004	-0.01	0.0002	0.01***
euro	-1.98***	0.44	-0.2	-0.70	-0.93**	.	.	0.02
constant	47.9***	1.71	-0.07	-1.31	-104.7**	-121.6	-24.8***	2.77
# observations	85	185	416	153	812	172	950	525

*Note: Significance levels at 1%, 5%, and 10% denoted by ***, ** and *, respectively.*

Source: author's estimation

If we have a look at the distance variable, we can observe quite ambiguous results. This may be caused by the time-invariance of distance and thus lower number of observations per sample. Mostly we can see around-unity negative impact on exports (Clusters 4, 5, 9 – i.e. standard trade partners of Austria) with exception of Cluster 2, where the coefficient is -7.5, which can be attributed to only 5 countries and thus 5 observations in the sample. On the contrary, Cluster 3 (rich European countries) seems to be less sensitive to the distance as transport distance is less important than GDPs and contiguity. Clusters 6 and 7 did not reveal any significant impact of distance – these are all very distant countries and as the distance vary across the subsamples it does not in fact affect the transport costs that significantly. That is why landlockness and possibly common currency play more important role among these countries. Shockingly, distance has positive impact on trade in Cluster 8 (very poor and rather distant African and Asian countries). Most probably, distance does not play any significant role within Cluster 8, as those countries are far enough anyway so that their relative differences in distance are not large and their GDP is the main determinant of their imports from Austria. In reality, these countries do not compete among themselves in attracting imports and the explanatory power of the gravity equation is thus reduced.

Recession seems to negatively impact only Clusters 2 and 4 (i.e. big economies and poor European countries) which would support the hypothesis that developing countries are not that sensitive to financial crises in terms of trade. Government effectiveness is a mostly significant variable except for very poor and distant countries in Clusters 6, 7 and 8, where transport costs (not very well approximated by distance in sub-samples) and landlockness together with local GDP (i.e. purchase power) play the most important role. As a bottom line to this section, it should be stressed that estimates by pools (clusters) cannot be taken as alternatives replacing full-scale estimates. In fact, they are mere complements pointing to the extent of heterogeneity in data and the weaker robustness of estimates of the full sample. Results in Table 9 offer just a partial view within the given cluster. They abstract from mutual competition of Austrian exports between clusters.

4.3 Clusters by Years

In the final part of the empirical chapter let us present the results of data clustering by very short time periods (mostly 3 years). The aim of this experiment is to twofold: (i) this time-wise re-estimation enables us to assess how the importance of the trade determinants has been developing in time; (ii) shorter time dimension significantly eliminates the non-stationary results, as discussed in section 1.4, which makes the results more reliable with respect to spurious regression. We regrouped the data in six clusters of 3 years.[13] Such number of clusters enables us to get rid of non-stationarity and observe time development of the coefficients as well as keep the number of observations per one regression reasonable. The numerical results are presented in Table 10 below with stars indicating the significance.[14] We employed Hausman-Taylor estimator analogically to the two previous sections, as already justified earlier.

Table 10: Sample clustered by triple-years – estimation results

Sub-period	1995-1996	1997-1999	2000-2002	2003-2005	2006-2008	2009-2011
log (GDP partner)	0.84***	0.69***	1.03***	0.63***	0.96***	0.09
log (GDP Austria)	-6.66***	0.82	1.08	2.13***	6.36***	1.26
log (distance)	-0.93***	-0.84***	-0.63***	-1.58***	-0.64**	-1.21**
language	0.05	0.35	0.44	0.7	0.59	0.51
contiguous	1.10	1.58	2.02**	1.65	1.73*	2.62
colony	0.55	0.98	1.64*	-0.03	1.25	-0.14
landlockness	-0.24	-0.27	-0.04	-0.9*	-0.45	-0.78
trade barriers	0.12	-0.34	0.71	-0.05	0.48	-1.13
gov. effectiveness	0.003	0.01**	0.01***	0.01	0.002	0.02***
euro	0.43	-0.13	1.06**	0.04	1.35*	0.01
constant	94.8***	6.4	0.11	0.40	-64.7***	13.2
# observations	375	564	581	589	608	598

*Note: significance levels at 1%, 5%, and 10% denoted by ***, ** and *, respectively.*

Source: *author's estimation*

[13] The first cluster has two years only due to the length of the dataset.

[14] Recession variable was dropped due to lack of observations.

According to the results in Table 10, the main determinants of Austrian exports in 1995-2011 were both partners' GDP and the distance. Common border, currency, landlockness and government effectiveness play marginal role due to the limited significance and/or low magnitude of the coefficients. These results confirm that exporting companies take only the core determinants into account in the short term. In addition, the significance of both arguably non-stationary variables (GDPs) indicates that the previous significant results were not distorted by spurious regression phenomenon, since they play a significant role in very short term period as well.

We can observe that export partner's GDP has had a stable and important impact on the export, ranging in years 1995-2008 between 0.7-1.0, which is in line with our previous findings when estimating the whole dataset. We are not able to identify any clear trend in the importance of partner's GDP except for the crisis years 2009-2011, when it completely lost its both significance and magnitude. This would indicate that the drop in exports is not distributed in the same way as the impact on the crisis on partners' economies.

Domestic GDP is larger in magnitude than the partners' GDP, which is in line with our previous findings as well as gravity research in general. However, the magnitude of the coefficients is extremely volatile, reaching even negative effect in years 1995-1996. This great uncertainty is associated with the extremely low number of observations in each sub period (only 2 or 3 per one regression), as opposed to partner's GDP, where we have different observation for each country.

Distance is a variable of great interest as it contributes to the debate about distance puzzle existence. Our results reveal great significance of this variable in each sub-period with the coefficient ranging from -1.6 to -0.6. Over years 1995-2008, the magnitude of the coefficient has been gradually decreasing with the exception in 2003-2005, when it suddenly jumped up to its maximum, for which we do not see any reason. Later on, there was another sudden jump in the crisis period 2009-2011, which could be attributed to larger sensitivity of exports to high transport costs during the crisis years.

CONCLUSION

Gravity models in trade traditionally focus on estimation of a general trade function. However, as we already indicated in Davidová & Benáček (2013), we believe that this approach could be misleading since different country groups and different periods of time are associated with different trading patterns. Therefore, we tried to reveal the driving forces shaping the short-term and long-term patterns of international trade flows by complementing these with the geographic sampling. Before doing so, however, we concentrated on the role of estimation techniques in the models of trade gravity as such by comparing their different estimators and assessing their reliability.

Fixed effects model belongs to generally recommended estimation techniques for gravity models in international trade, mainly for its ability to include unobservable individual effects that are broadly believed to exist in foreign trade data. However, the main drawback of this technique is its inability to give results on time-invariant variables, as for example distance, common language or common border dummies, which are in fact variables of interest as well. That is why we tried to compare the fixed effect model to alternative estimation techniques that also account for unobserved individual effects but are still able to give results on variables constant in time at the same time. As a result, all three models included (Poisson PML, Mundlak model and Hausman-Taylor) yield very consistent results with very good data fit, implying that none of them is subject to a severe bias. All models indicate results that are in line with general practice in gravity models research as well – i.e. both domestic and partner's GDPs and distance are the major determinants. In addition, common border and trade barriers were found to play an important role as well.

Based on data imperfection and the ability of those three alternative estimators, we assessed their reliability, choosing out preferred one to be employed in the consequent cluster analysis as the only tool. The main advantage of Poisson PML estimator is that it accounts for bias caused by logarithmic form under presence of heteroskedasticity. Unfortunately, its pure version does not include fixed effects and moreover, it does not seem to correspond to our data well due to clearly overvalued significance. Mundlak model is based on the random effects, relaxing the orthogonality assumption and decomposing errors to country-specific and time-specific and also providing both short-run and long-run estimates. The model thus combines the estimation of data under the time-series data structure with cross-sectional data structure, which makes the results very hard to interpret. Hausman-Taylor was chosen as our preferred estimator since it not only accounts for country-specific unobserved effects, but also provides consistent and efficient estimates of the time-invariant variables. It employs instrumental variable technique and is thus able to account for bidirectional relationship between the international trade and countries' GDPs.

In order to present even more specific results, we clustered our partner countries into nine groups with similar economic, geographical and institutional characteristics. Our results indicate that the export function does in fact vary significantly across these country groups. Generally speaking, our numerical results confirmed our expectations that the more advanced countries, the more sophisticated variables come into play. In particular, partner's GDP was revealed as less significant among more developed and richer countries, while contiguity of institutional background appears to be important. Also rich European countries are not sensitive to the distance as opposed to further located Austria's trade partners. Similarly, rich and developed western countries appeared to be much more sensitive to economic crisis starting in 2009, as opposed to poor developing countries typically in Latin America, Asia and Africa.

In our second clustering experiment, we divided the sample into 6 sub-periods, in order to reveal changes in export function over time and dynamic evolvement of the variables (and also to suppress possible negative impact of data non-stationarity). In fact, we did not experience any loss of significance indicating that slight data non-stationarity did not distort our previous results either. Partners' GDP has had constant significant impact in 1995-2008 and lost its significance in the crisis period 2009-2011 indicating that vulnerability of exports to crisis is not distributed in the same way as the vulnerability of GDP. Distance variable is very significant over the whole observed period, experiencing a sudden jump in 2009, which could be attributed to larger sensitivity of exports to high transport costs during the crisis years.

There are two main contributions of this research to the theory and practice of gravity models of trade: (i) Providing a synthetic methodological overview of the technical problems with the estimation of gravity equations, which was revealed by amalgamating the innovative findings of a vast list of literature in the last 10 years; (ii) Testing for the heterogeneity of data sets used in gravity models of trade leading to a conclusion that behavioral patterns of exporters and importers built in the datasets are very complicated and a single generalized specification of gravity equation can lead to bias in estimates and/or to similarly generalized conclusions that hide important robust idiosyncrasies in behavior present in some subsamples of economic agents.

REFERENCES

Anderson, J. E., & van Wincoop, E. (2003). Gravity with Gravitas: A Solution to the Border Puzzle. *American Economic Review, vol. 93(1)*, pp. 170-192.

Babecká Kucharčuková, O., Babecký, J., & Raiser, M. (2012). Gravity Approach for Modelling International Trade in South-Eastern Europe and the Commonwealth of Independent States: The Role of Geography, Policy and Institutions. *Open Economies Review, vol. 23(2)*, pp. 277-301.

Baldwin, R. (1994). *Towards an Integrated Europe.* London: Centre for Economic Policy Research.

Baldwin, R., & Taglioni, D. (2006). Gravity for Dummies and Dummies for Gravity. *NBER Working Paper 12516, National Bureau of Economic Research*.

Baltagi, B. H. (2008). *Econometric Analysis of Panel Data.* Wiley, 4th Edition.

Baltagi, B. H., & Egger, P. (2014). Estimation of Structural Gravity Quantile Regression Models. *Forthcoming Working Paper*.

Berthelon, M., & Freund, C. L. (2008). On the Conservation of Distance in International Trade. *Journal of International Economics, vol. 75(2)*, pp. 310-320.

Bleaney, M., & Neaves, A. S. (2013). Declining Distance Effects in International Trade: Some Country-Level Evidence. *World Economy, vol. 36*, pp. 1029–1040.

Bobková, B. (2014). On Estimation of Gravity Equation - Cluster Analysis. *Working Papers IES 12/2014. Charles University in Prague*.

Brüderl, J. (2005). *Panel Data Analysis.* University of Mannheim.

Brulhart, M., & Kelly, M. J. (1999). Ireland's Trading Potential with Central and Eastern European Countries: A Gravity Study. *Economic and Social Review, vol. 30*, pp. 159-174.

Brun, J.-F., Carrere, C., Guillaumont, P., & Melo, J. (2005). Has Distance Died? Evidence from a Panel Gravity Model. *World Bank Economic Review, vol. 19(1)*, pp. 99-120.

Cameron, A. C., & Trivedi, P. K. (2013). *Regression Analysis of Count Data.* Cambridge University Press, 2nd edition.

Coe, D. T., Bhavnani, R., & Tamirisa, N. T. (2007). The Missing Globalization Puzzle: Evidence of the Declining Importance of Distance. *IMF Working Papers 02/171*.

Davidová, L. (2012). Determinants of Austrian International Trade: Econometric Analysis Based on the Gravity Model. *IES FSV UK, Bachelor Diploma Thesis. Thesis supervisor: doc. Ing. Benáček CSc.*

Davidová, L. (2013). Ekonometrická analýza hlavních vlivů na rakouské dovozy. *University of Economics in Prague, Bachelor diploma thesis. Thesis supervisor: Ing. Sedláček, Ph.D.*

Davidová, L., & Benáček, V. (2013). Determinants of Austrian International Trade: Gravity Model Analysis. *Working Papers IES 05/2014. Charles University in Prague*.

Disdier, A. C., & Head, K. (2008). The Puzzling Persistence of the Distance Effect on Bilateral Trade. *The Review of Economics and Statistics, vol. 90(2)*, pp. 37-48.

Egger, P. (2000). A Note on the Proper Econometric Specification of the Gravity Equation. *Economic Letters, vol. 66*, pp. 25-31.

Egger, P. (2002). An Econometric View on the Estimation of Gravity Models and the Calculation of Trade Potentials. *The World Economy, vol. 25*, pp. 297-312.

Egger, P., & Nigai, S. (2014). Structural Gravity with Dummies Only: Constrained ANOVA-type Estimation of Gravity Panel Data Models. *CEPR Discussion Paper*.

Egger, P., & Pfaffermayr, M. (2003). The Proper Panel Econometric Specification of the Gravity Equation: A Three-Way Model with Bilateral Interaction Effects. *Empirical Economics, vol. 28*.

Egger, P., & Pfaffermayr, M. (2004). Distance, Trade and FDI: a Hausman–Taylor SUR Approach. *Journal of Applied Econometrics, vol. 19*, pp. 227–246.

Egger, P., & Pfaffermayr, M. (2005, January). Estimating Long and Short Run Effects in Static Panel Models. *Econometric Reviews, vol. 23(3)*, pp. 199-214.

Egger, P., & Pfaffermayr, M. (2005, December). The Determinants of Intrafirm Trade: In Search for Export–Import Magnification Effects. *Review of World Economy, vol. 141(4)*, pp. 648–669.

Egger, P., & Průša, J. (2014). The Determinants of Trade Costs: A Random Coefficients Approach. *CESifo Working Paper Series No. 4859*.

Eicker, B. (1963). Limit Theorems for Regression with Unequal and Dependant Errors. *Annals of Mathematical Statistics, vol. 34*, pp. 447–456.

Engle, R. F., & Granger, C. (1987). Co-Integration and Error Correction, Representation, Estimation and Testing. *Econometrica, vol. 55(2)*, pp. 251–276.

Fally, T. (2012). Structural Gravity and Fixed Effects. *Working paper, University of Colorado-Boulder*.

Fidrmuc, J. (2009). Gravity Models in Integrated Panels. *Empirical Economics, vol. 37(2)*, pp. 435-446.

Goméz, E. H., & Milgram, J. B. (2009). Are Estimation Techniques Neutral to Estimate Gravity Equations? An Application to the Impact of EMU on Third Countries' Exports. *Universidad de Granada*.

Gourieroux, C., Monfort, A., & Trognon, A. (1984). Pseudo Maximum Likelihood Methods: Applications to Poisson Models. *Econometrica, vol. 52*, pp. 701–720.

Greene, W. H. (2011). *Econometric Analysis.* Prentice Hall, 7th ediiton.

Gros, D., & Gonciarz, A. (1996). A Note on the Trade Potential of Central and Eastern Europe. *European Journal of Political Economy, vol. 12*, pp. 709-721.

Hamilton, C. B., & Winters, A. L. (1992). Opening up International Trade with Eastern Europe. *Economic Policy, vol. 14*, pp. 77-116.

Hausman, J. A., & Taylor, W. E. (1981). Panel Data and Unobservable Individual Effect. *Econometrica, vol. 49*, pp. 1377-1398.

Hayashi, F. (2000). *Econometrics.* Princeton University Press.

Head, K., & Mayer, T. (2013). Gravity Equations: Workhorse, Toolkit, and Cookbook. *CEPII Working Paper 2013-27*.

Chaney, T. (2013, June). Liquidity Constrained Exporters. *NBER Working Paper 19170, Cambridge*.

Chaney, T. (2013, August). The Gravity Equiation in International Trade: An Explanation. *NBER Working Paper 19285, Cambridge*.

Chang-Ching, L., & Ng, S. (2012). Estimation of Panel Data Models with Parameter Heterogeneity when Group Membership is Unknown. *Journal of Econometric Methods, vol. 1(1)*.

Cheng, J. H., & Wall, H. J. (2005). Controlling for Heterogeneity in Gravity Models of Trade and Integration. *Federal Reserve Bank of St. Louis Review, vol. 87*, pp. 49-63.

Kao, C., & Chiang, M. H. (2000). On the Estimation and Inference of Cointegrated Regression in Panel Data. *Advanced Econometrics, vol 15.*, pp. 179-222.

Kapetanios, G. (2006). Cluster Analysis of Panel Data Sets Using Non-Standard Optimisation of Information Criteria. *Journal of Economic Dynamics and Control, vol. 30(8).*

Kimura, F., & Lee, H. H. (2006). The Gravity Equation in International Trade in Services. *Review of World Economy, vol. 142*, pp. 92-121.

Krishnakumar, J. (2004). Time Invariant Variables and Panel Data Models: A Generalised Frisch-Waugh Theorem and its Implications. *Cahiers du département d'économétrie, University of Geneva*.

Leamer, E., & Levinsohn, J. (1995). International Trade Theory: The Evidence. *Handbook of International Economics, vol. 3*, pp. 1339-1394.

MacQueen, J. (1967). Some Methods for Classification and Analysis of Multivariate Observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, Statistics*, pp. 281-297.

Martin, W., & Pham, C. S. (2008). Estimating the Gravity Model when Zero Trade Flows are Frequent. *Economics Series; MPRA Paper No. 9453*.

Martínez-Zarzoso, I., Felicitas Nowak-Lehmann, D., & Vollmer, S. (2007). The Log of Gravity Revisited. *Center for European, Governance and Economic Development Research. Discussion Paper 64*.

Máthyás, L. (1997). Proper Econometric Specification of Gravity Model. *The World Economy, vol 20.*, pp. 363-368.

McCallum, J. (1995). National Borders Matter: Canada–U.S. Regional Trade Patterns. *American Economic Review, vol. 85*, pp. 615–623.

Micco, A., Stein, E., & Ordonez, G. (2003). The Currency Union Effect on Trade: Early Evidence from EMU. *Economic Policy*, pp. 316-356.

Mitze, T. (2010). Estimating Gravity Models of International Trade with Correlated Time-Fixed Regressors: To IV or not IV? *Munich Personal RePEc Archive Paper No. 23540*.

Mundlak, Y. (1978). On the Pooling of Time Series and Cross Section Data. *Econometrica, vol. 46(1)*, pp. 69-85.

Nilsson, L. (2000). Trade Intergartion and the EU Economic Membership Criteria. *European Journal of Political Economy, vol. 16*, pp. 807-827.

Park, H. M. (2010). Practical Guides to Panel Data Analysis. *Working Paper, International University of Japan*.

Paulus, M., Michalíková, E., & Benáček, V. (2014). German International Trade: Interpreting Export Flows According to the Gravity Model. *Working Papers IES 2014/19, Charles University Prague*.

Pedroni, P. (1996). Fully Modified OLS for Heterogeneous Cointegrated Panels and the Case of Purchasing Power Parity. *Working paper No. 96–020, Department of Economics, Indiana University*.

Pedroni, P. (2001). Purchasing Power Parity Tests in Cointegrated Panels. *The Review of Economics and Statistics, MIT Press, vol. 83(4)*, pp. 727-731.

Perales, F. (2013). Mundlak: Stata module to estimate random-effects regressions adding group-means of independent variables to the model. *IDEAS at the Research Division of the Federal Reserve Bank of St. Louis*.

Pirotte, A. (1999). Convergence of the Static Estimation Toward the Long Run Effects of Dynamic Panel Data Models. *Economic Letters, vol. 63*, pp. 151-158.

Plümper, T., & Troeger, V. (2007). Efficient Estimation of Time-Invariant and Rarely Changing Variables in Finite Sample Panel Analyses with Unit Fixed Effects. *Political Analyses, vol. 15*, pp. 124-139.

Rose, A. (2000). One Money, One Market: Estimating The Effect of Common Currencies on Trade. *Economic Policy*, pp. 7-45.

Rose, A., & van Wincoop, E. (2001). National Money as a Barrier to International Trade: The Real Case for Currency Union. *American Economic Review, vol. 91(2)*, 386-390.

Santos Silva, J. M., & Tenreyro, S. (2006). The Log of Gravity. *The Review of Economics and Statistics, vol. 88(4)*, pp. 641-658.

Santos Silva, J. M., & Tenreyro, S. (2010). On the Existence of the Maximum Likelihood Estimates in Poisson Regression. *Economics Letters, vol. 107(2)*, pp. 310–312.

Sarafidis, V., & Weber, N. (2009). To Pool or Not to Pool: A Partially Heterogeneous Framework. *MPRA paper no. 36155*.

Sarafidis, V., & Weber, N. (2014). A Partially Heterogeneous Framework for Analyzing Panel Data. *Oxford Bulletin of Economics and Statistics. doi: 10.1111/obes.12062*.

Serlenga, L., & Shin, Y. (2007). Gravity Models of Intra-EU Trade: Application of the CCEP-HT Estimation in Heterogeneous Panels with Unobserved Common Time-specific Factors. *Journal of Applied Econometrics, vol. 22*, pp. 361-381.

Wall, H. J. (2000). NAFTA and the Geography of North American Trade. *Working Paper 2000-17B, Federal Reserve Bank of St. Louis*.

Wang, Z. K., & Winters, A. L. (1991). The Trading Potential of Eastern Europe. *CEPR Discussion Paper No. 610, London*.

Westerlund, J., & Wilhelmsson, F. (2011). Estimating the Gravity Model without Gravity Using Panel Data. *Applied Economics, vol. 43(6)*, pp. 641-649.

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica, vol. 48*, pp. 817–838.

Winkelmann, R. (2003). *Econometric Analysis of Count Data.* Berlin: Springer-Verlag, 4th edition.

Wooldridge, J. M. (2012). *Introductory Econometrics: A Modern Approach.* Cengage Learning, 5th Edition.

Yotov, Y. V. (2012). A Simple Solution to the Distance Puzzle in International Trade. *Economics Letters, vol. 117(3)*, pp. 794-798.

Zwinkels, R. C., & Beugelsdijk, S. (2010). Gravity Equations; Workhorse or Trojan Horse in Explaining Trade and FDI Patterns Across Time and Space? *International Business Review, vol. 19(1)*, pp. 102-114.

APPENDIX

A.1 Detailed Estimation Results on Complete Sample

Table 11: Fixed effects model

Observations	=	3396		R-squared:	within	=	0.236
Number of groups	=	211			between	=	0.775
F(6, 210)	=	91.2			overall	=	0.724
Prob > F	=	0.000			Corr (u_i, xb)	=	0.673

log (export)	**coefficient**	**std. err.**	**p-value**	**[95% confidence interval]**	
log (GDP partner)	0.548	0.075	0.000	0.400	0.696
log (GDP Austria)	1.515	0.203	0.000	1.114	1.915
log (distance)	(omitted)				
language	(omitted)				
contiguous	(omitted)				
colony	(omitted)				
landlockness	(omitted)				
recession	0.054	0.052	0.300	-0.048	0.155
trade barriers	-0.397	0.087	0.000	-0.568	-0.226
gov. effectiveness	0.000	0.003	0.891	-0.006	0.005
euro	-0.330	0.114	0.004	-0.554	-0.105
constant	-7.222	2.250	0.002	-11.658	-2.787

Source: author's estimation

Table 12: Poisson PML estimator

Observations	=	3396	Pseuo R-squared	=	0.973
Log pseudolikelihood	=	-9.2e+10	Wald chi2(11)	=	27745
			Prob > chi2	=	0.000

log (export)	**coefficient**	**std. err.**	**p-value**	**[95% confidence interval]**	
log (GDP partner)	0.810	0.051	0.000	0.709	0.911
log (GDP Austria)	1.107	0.182	0.000	0.751	1.463
log (distance)	-1.314	0.116	0.000	-1.542	-1.086
language	0.270	0.184	0.141	-0.089	0.630
contiguous	0.761	0.246	0.002	0.278	1.244
colony	0.340	0.254	0.180	-0.157	0.838
landlockness	-0.383	0.244	0.117	-0.861	0.096
recession	0.046	0.052	0.374	-0.055	0.147
trade barriers	-0.274	0.072	0.000	-0.416	-0.132
gov. effectiveness	0.001	0.003	0.847	-0.005	0.006
euro	-0.210	0.093	0.024	-0.393	-0.027
constant	6.179	2.273	0.007	1.723	10.635

Source: author's estimation

Table 13: Mundlak model

Observations	= 3396		R-squared:	within	= 0.236
Number of groups	= 211			between	= 0.884
Wald chi2(11)	= 2483.1			overall	= 0.850
Prob > chi2	= 0.000			rho	= 0.685

log (export)	**coefficient**	**std. err.**	**p-value**	**[95% confidence**	**interval]**
log (GDP partner)	0.548	0.042	0.000	0.466	0.630
log (GDP Austria)	1.515	0.110	0.000	1.300	1.729
log (distance)	-1.194	0.138	0.000	-1.465	-0.923
language	0.138	0.660	0.835	-1.156	1.432
contiguous	0.452	0.599	0.450	-0.722	1.626
colony	0.510	0.675	0.450	-0.814	1.834
landlockness	-0.127	0.250	0.611	-0.617	0.363
recession	0.054	0.062	0.385	-0.067	0.175
trade barriers	-0.397	0.166	0.017	-0.723	-0.071
gov. effectiveness	0.000	0.002	0.796	-0.003	0.003
euro	-0.330	0.185	0.075	-0.692	0.033
constant	15.870	50.837	0.755	-83.803	115.54

Source: author's estimation

Table 14: Hausman-Taylor estimator

Number of observations	= 3315	Wald chi2(11)	= 1832.7
Number of groups	= 209	Prob > chi2	= 0.000
		rho	= 0.754

log (export)	**coefficient**	**std. err.**	**p-value**	**[95% confidence**	**interval]**
Time-variant exogenous					
recession	-0.027	0.058	0.644	-0.140	0.086
trade barriers	-0.173	0.154	0.261	-0.475	0.129
gov. effectiveness	0.002	0.001	0.236	-0.001	0.005
euro	-0.134	0.172	0.436	-0.470	0.202
Time-variant endogenous					
log (GDP partner)	0.585	0.037	0.000	0.513	0.656
log (GDP Austria)	1.224	0.101	0.000	1.027	1.421
log (export(n-1))	0.160	0.010	0.000	0.142	0.179
Time-invariant exogenous					
log (distance)	-0.934	0.173	0.000	-1.272	-0.595
language	0.459	0.708	0.517	-0.929	1.846
contiguous	1.181	0.681	0.083	-0.153	2.515
colony	0.450	0.741	0.544	-1.003	1.903
landlockness	-0.376	0.262	0.151	-0.888	0.137
Constant	0.863	1.815	0.634	-2.694	4.421

Source: author's estimation

A.2 Tests for Assumption Violations

Table 15: White's test for heteroskedasticity

H_0: homoskedasticity
H_A: unrestricted heteroskedasticity

$\chi^2(65)$	=	530.47
Prob $>\chi^2$	=	0.0000

Cameron & Trivedi's decomposition of IM-test:

Source	χ^2	df	p-value
Heteroskedasticity	530.47	65	0.000
Skewness	48.83	11	0.000
Kurtosis	17.81	1	0.000
Total	**597.12**	**77**	**0.000**

→ H_0 strongly rejected, i.e. heteroskedasticity detected

Source: author's calculation

Table 16: Skewness/Kurtosis and Shapiro-Wilk W normality tests

H_0: residuals normally distributed
H_A: residuals non-normally distributed

Skewness/Kurtosis tests for Normality

				----------- joint -----------	
Variable	Observations	Pr(Skewness)	Pr(Kurtosis)	adjχ^2	Prob>chi2
residuals	3315	0.000	0.000	.	0.000

Shapiro-Wilk W test for normal data

Variable	Observations	W	V	z	Prob>z
residuals	3315	0.893	201.0	13.7	0.000

→ H_0 strongly rejected, i.e. residuals non-normally distributed

Source: author's calculation

Table 17: Serial autocorrelation test

Observations	3208	R-squared	0.0082
F(14, 3193)	1.89	adjusted R-squared	0.0039
Prob > F	0.023	Root MSE	0.706

log (export)	**coefficient**	**std. err.**	**p-value**	**[95% confidence interval]**	
res_1	-0.060	0.018	0.001	-0.094	-0.025
res_2	-0.010	0.017	0.575	-0.044	0.024
res_3	-0.063	0.017	0.000	-0.096	-0.029
log (GDP partner)	-0.004	0.006	0.481	-0.016	0.007
log (GDP Austria)	0.047	0.077	0.537	-0.103	0.198
log (distance)	0.007	0.020	0.732	-0.032	0.045
language	0.003	0.096	0.974	-0.184	0.191
contiguous	0.011	0.086	0.900	-0.157	0.179
colony	0.001	0.094	0.990	-0.184	0.186
landlockness	-0.002	0.036	0.957	-0.072	0.068
recession	0.002	0.053	0.974	-0.103	0.107
trade barriers	-0.014	0.069	0.837	-0.150	0.121
gov. effectiveness	0.000	0.001	0.989	-0.001	0.001
euro	-0.005	0.067	0.938	-0.136	0.125
constant	-0.577	0.949	0.543	-2.438	1.283

Source: author's estimation

A.3 Cluster Analysis

Table 18: Clusters of countries (1/2)

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 6	
USA	FRANCE	BELGIUM	ANDORRA	ST KITTS	MACAO
	GERMANY	DENMARK	BAHRAIN	ANGUILLA	MALAYSIA
	CHINA	GREECE	CROATIA	ANTIGUA & B.	MALDIVES
	JAPAN	NETHERLANDS	CYPRUS	ARGENTINA	MAURITIUS
	UK	NORWAY	CZECH REP.	BAHAMAS	MOZAMBIQUE
		POLAND	ESTONIA	BARBADOS	NAMIBIA
		SAUDI ARABIA	FAROE ISL.	BELIZE	NL. ANTILLES
		SWEDEN	FINLAND	BERMUDA	PALAU
		SWITZERLAND	GREENLAND	BOLIVIA	PANAMA
		TAIWAN	HUNGARY	BOTSWANA	PERU
		TURKEY	ICELAND	BRUNEI	PHILIPPINES
			IRELAND	CAYMAN ISL.	SAINT HELENA
			ISRAEL	COLOMBIA	SEYCHELLES
			LATVIA	COSTA RICA	SINGAPORE
			LIECHTENSTEIN	CUBA	SOUTH AFRICA
			LITHUANIA	DOMINICA	SRI LANKA
			LUXEMB.	EL SALVADOR	ST LUCIA
			MALTA	GRENADA	ST VINCENT
			OMAN	GUATEMALA	SURINAME
			PORTUGAL	GUYANA	THAILAND
			QATAR	HONG KONG	TRINIDAD & T.
			SLOVAKIA	CHILE	TURKS AND C. ISL.
			SLOVENIA	INDONESIA	URUGUAY
			TUNISIA	JAMAICA	VIET-NAM
			UAE	LESOTHO	VIRGIN ISL.

Source: author's estimation

Table 19: Clusters of countries (2/2)

Cluster 5	Cluster 7	Cluster 8		Cluster 9
AUSTRALIA	KIRIBATI	AFGHANISTAN	LIBERIA	ALBANIA
BRAZIL	TOKELAU	ANGOLA	MADAGASCAR	ALGERIA
CANADA	TUVALU	ARUBA	MALAWI	ARMENIA
INDIA	COOK ISLAND	BANGLADESH	MALI	AZERBAIJAN
ITALY	FALKLAND ISL.	BHUTAN	MONGOLIA	BELARUS
MEXICO	FIJI	BURKINA FASO	MYANMAR	BENIN
RUSSIA	FR. POLYNESIA	BURUNDI	N. KOREA	BOSNIA
S. KOREA	MARSHALL ISL.	CAMBODIA	NEPAL	BULGARIA
SPAIN	MICRONESIA	CAMEROON	NICARAGUA	CAPE VERDE
	NAURU	CENTRAL AFR. R.	NIGER	EGYPT
	NEW CALEDONIA	COMOROS	NIGERIA	GAMBIA
	NEW ZEALAND	CONGO	PARAGUAY	GEORGIA
	NIUE	CONGO, DEM. REP.	RWANDA	GHANA
	N. MARSHAL ISL.	COTE D'IVOIRE	SAINT PIERRE	GIBRALTAR
	PAPUA NEW G.	DJIBOUTI	SAO TOME	GUINEA-BISSAU
	SAMOA	DOMINICAN REP.	SIERRA LEONE	IRAN
	SOLOMON I.	ECUADOR	SOMALIA	JORDAN
	TONGA	EQ. GUINEA	SUDAN	KAZAKHSTAN
	VANUATU	ERITREA	SWAZILAND	KUWAIT
	WALLIS	ETHIOPIA	TAJIKISTAN	LEBANON
		GABON	TANZANIA	LIBYA
		GUINEA	TOGO	MACEDONIA
		HAITI	TURKMENISTAN	MAURITANIA
		HONDURAS	UGANDA	MOLDOVA
		CHAD	UZBEKISTAN	MOROCCO
		IRAQ	VENEZUELA	PAKISTAN
		KENYA	YEMEN	ROMANIA
		KYRGYZSTAN	ZAMBIA	SAN MARINO
		LAOS	ZIMBABWE	SENEGAL
				SERBIA
				SYRIA
				UKRAINE

Source: author's estimation