Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies

BACHELOR THESIS

Comparison of Factors Determining Czech Exports in Different Aggregations:
Analysis by means of an Adjusted Gravity Model

Author: **Eva Smotlachová**

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

Academic Year: **2013/2014**

Declaration of Authorship

I hereby declare that I compiled this thesis independently, using only the listed resources and literature.

I also declare that this thesis was not used for obtaining another degree.

I grant to Charles University permission to reproduce and distribute copies of this thesis document in whole or in part.

Prague, May 13, 2014

Eva Smotlachová

Acknowledgments

I am grateful especially to my supervisor doc. Ing. Vladimír Benáček CSc. for providing me with valuable advice and useful comments.

Similarly, I would like to express gratitude to Mgr. Božena Bobková and Mgr. Eva Michalíková for their help with technical details of my research.

Lastly, I am grateful to my family and friends for their support and motivation.

Bibliography record

Smotlachová, E. (2014). Comparison of Factors Determining Czech Exports in Different Aggregations: Analysis by means of an Adjusted Gravity Model. *IES FSV, Charles University in Prague. Bachelor thesis, 77 pp.* Supervisor: doc. Ing. Vladimír Benáček CSc.

Abstract

The aim of the thesis is to estimate determinants of Czech export. The key novelty of the research consists in estimating export flows at three different levels of aggregation (total, machinery, automobiles) and a subsequent comparison of results. An augmented gravity model is implemented for the empirical research and estimated with the use of Ordinary least squares (including time or country dummies) and Poisson pseudo maximum likelihood estimators. Consequently, we propose the comparison based on the Poisson estimates due to incontestable weaknesses of the OLS estimator. Besides, we emphasize the importance of identifying the panel data structure, as we have obtained different results under the two structures (time series and cross-sectional). The predicted models reveal a substantial impact of home and partner's GDP, distance and the European Union. Total and machinery export seem to follow same tendencies, whereas the development of Czech automobile export differs and evolves on an idiosyncratic path. Therefore, it does not coincide with a path expected by trade theories based on the assumption of perfect competitive markets.

JEL Classification	F01, F14, C23, C52
Keywords	Czech Republic, Export, Gravity model, Poisson estimation, Machinery, Automobiles
Author's e-mail	eva.smotlachova@gmail.com
Supervisor's e-mail	benacekv@fsv.cuni.cz

Abstrakt

Cílem této práce je odhadnout faktory ovlivňující český export. Hlavní přínos výzkumu spočívá v odhadu pohybů exportu na základě tří různých agregátních skupin (celkový, strojírenský a automobilový export) a v následném porovnání výsledků. Pro vlastní výzkum je použit rozšířený gravitační model, který je odhadován pomocí metody nejmenších čtverců (OLS) s dummy proměnnými a Poissonovské regrese. Následně nabízíme porovnání výsledků získaných pouze z Poissonovského odhadu kvůli nesporným nedostatkům odhadování pomocí OLS. Kromě toho klademe důraz na nezbytnost rozlišení mezi panelovými strukturami dat, protože jsme odhalili rozdílné výsledky pro jednotlivé struktury (tzn. časové a průřezové dělení). Odhadnuté modely odhalily významný efekt domácího a partnerského HDP, geografické vzdálenosti a Evropské Unie. Navíc celkový a strojírenský export se řídí stejnými zákonitostmi, kdežto vývoj českého exportu automobilů se liší a nechová se v souladu se standartními obchodními teoriemi, odvozenými z předpokladu dokonale konkurenčních trhů.

Klíčová slova	Česká republika, Vývoz, Gravitační model, Poissonův odhad, Strojírenství, Automobily

Rozsah práce: 56 stran, 94 459 znaků

CONTENT

LIST OF TABLES

LIST OF FIGURES

BACHELOR THESIS PROPOSAL

Author:	Eva Smotlachová
Supervisor:	doc. Ing. Vladimír Benáček, CSc.
Defense planned:	June 2014
Language of the work:	English

Topic: Comparison of factors determining Czech exports in different aggregations

Key words: export, gravity model, intermediates, Czech Republic

Topic characteristics:

The main goal of my thesis will be to figure out various factors determining volumes of Czech export regarding different commodity classes. We will concentrate on a machinery industry and will proceed systematically from the general case to more detailed products. In order to estimate models of these exports we will use gravity model. First, we will give some background theory of the gravity model and connected risks. At this part, a description of the environment and factors of Czech export will be included as well. The core of the work will consist in estimating 3 different models for which the essential thing will be to collect convenient data. We will search for data for about 60 countries from year 1995 to 2012. After revealing coefficients of the models, we will evaluate results and compare the theoretical values with the real ones. Finally, we will discuss on the differences and try to draw consequences indicating possible opportunities for Czech exporters.

Core bibliography:

AFMAN, E.R. and MAUREL, M. (2010): "Diplomatic relations and trade reorientation in transition countries." In: *The Gravity Model in International Trade*. Cambridge University Press.

BALDWIN R., TAGLIONI, D. (2011): "Gravity Chains: Estimating Bilateral Trade Flows When Parts And Components Trade Is Important." National Bureau of Economic Research, Inc, NBER Working Paper No. 16672

EGGER, P. (2000) “A Note on the Proper Econometric Specification of the Gravity Equation.” *WIFO Working Papers:108*.

HAMILTON, J. D. (1994): “Time Series Analysis.” New Jersey: Princeton University Press.

HARUKA, Y. (2013): “Prospects for Trade in Intermediates and Trade in Services:

What Does the Gravity Model of Bilateral Trade Tell Us?” *OSIPP Discussion Paper: DP-2013-E-002*, Osaka University

LINDERS, G. J. and Henri L. F. de GROOT (2006) “Estimation of the gravity equation in the presence of zero flows.” *Tinbergen Institute Discussion Paper (072)*, March.

SANTOS SILVA, J. M. and S. TENREYRO (2006) “The log of gravity.” *The Review of Economics and Statistics 88 (4)*, pp.641-658.

WOOLDRIDGE, J. M. (2001): “Econometric Analysis of Cross Section and Panel Data.” Volume 1 of MIT Press Books. The MIT Press.

WOOLDRIDGE, J. M. (2002): “Introductory Econometrics”, 3rd Ed. South-Western College Publishing.

_____________ ______________

Author Supervisor

1 INTRODUCTION

The Czech Republic is a small economy whose prosperity was put behind by the Communist regime in the last century. The Czech economy, however, succeeded to profit from its good initial conditions and nowadays, it creates an integral part of the European market. Moreover, the Czech Republic takes a prominent position with respect to international trade, particularly from view of machinery and automobile export. This thesis focuses on revealing determinants of Czech export flows. For that purpose, we employ a standard econometric tool for estimating international trade flows, i.e. a gravity model of trade.

A considerable amount of works in recent period focused on the problematic of the gravity model. A father of this econometric tool is Jan Tinbergen who designed the gravity model of trade on a basis of the Newton's law of universal gravitation (Tinbergen, 1962). Later on, researchers were introducing multiple procedures of implementing the gravity model and ensuring consistency of results. In a response, the most widely accepted tool in recent years is the Poisson pseudo maximum likelihood estimator, which was first suggested by Santos Silva & Tenreyro (2006).

The main contribution of the thesis lies in estimating three models at different aggregation levels. To be precise, we propose a comparative analysis of total, machinery and automobile export functions. This approach is innovative in the way that we reveal different motivations behind export development of more disaggregated commodity categories. To our knowledge, this research represents the first attempt for a comparison of export functions between the main (aggregated) group and its subgroups (machinery, automobiles).

Moreover, we approve the inaccuracy of using Ordinary least squares (OLS) method for estimating trade flows. Its weakness lies in transforming a gravity equation into a log-linearized form, which may cause inconsistency of coefficients due to presence of heteroscedasticity and occurrence of zero flows. Therefore, we use the models estimated by the OLS only as benchmark models for comparison with Poisson estimates.

Last but not least, we emphasize the importance of distinguishing between structures of datasets. In detail, we propose estimations under time series and cross-sectional data structure and show considerable differences between the two approaches. We anticipate the incompleteness of stating results obtained only under one structure (usually time series), as it is applied in most works.

To put our research into a framework, we briefly describe Czech industry in Chapter 2, by underlying the importance of the export sector. Chapter 3 proposes a comprehensive characteristic of the gravity model, including the related literature review. Later on, Chapter 4 concentrates on the actual research, particularly presents the applied methodology and a description of our models. The output of our analysis is summarized in Chapter 5, Chapter 6 and Chapter 7, focusing respectively on the total, machinery and automobile predictions. Finally, Chapter 8 concludes and Appendix covers tables and figures that are not included in the actual text.

2 CZECH INDUSTRY

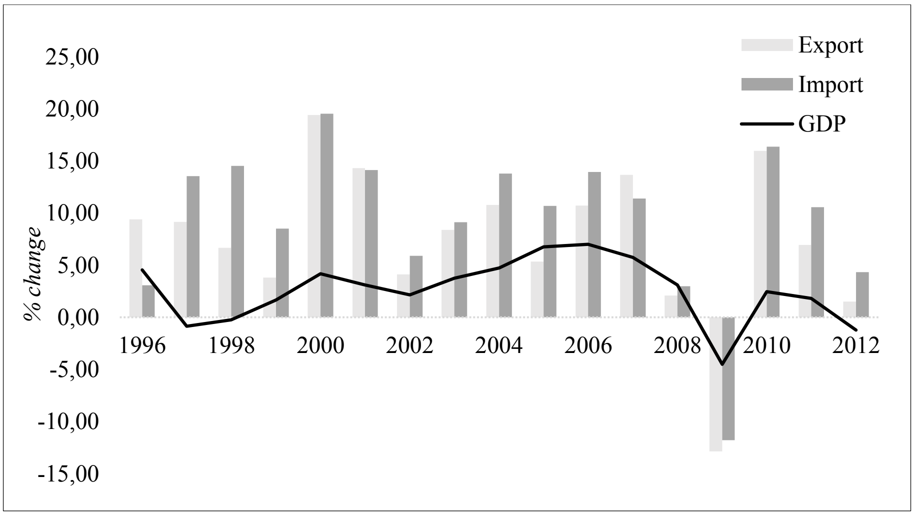
Our research concentrates on determining factors of Czech export for years 1995-2012 and thus, all subsequent statements and analyses are derived with respect to this period. This chapter serves as an introduction to the topic, to understand better the motivation of our study. We proceed as follows. Firstly, we present some basic facts about Czech economy and we continue by outlining a recent development in the export sector. Later on, we specify the export structure and at the end, we introduce a description of Czech automobile industry.

2.1 Basic Characteristics

The Czech Republic is a small open economy, occupying a strategic position in the centre of Europe. The economic growth and especially the growth of export to developed markets was put behind by the Communist era (1948-1989). Though, the Czech Republic succeeded to take advantage of its good initial position after the political changes and today is classified as the second most prosperous country among the post-Communist states (behind Slovenia). Significant positive effects on the Czech development implied the membership in international organisations – the European Union in the first place and then OECD (Organisation for economic cooperation and development) and WTO (World Trade Organisation).

In fact, all these organisations facilitated Czech international trade and encouraged the economic growth. However, as the Czech Republic has gradually become an export-driven economy, it is dependent on the economic performance of its export partners and particularly sensitive to changes in their demand. For this reason, the financial crisis, which hit in Western Europe in 2008, cut down the demand for Czech goods. Consequently, the drop in export affected the GDP growth, as well as the spending propensity of the domestic aggregate demand. Accordingly, as Czech exporters were generally increasing their dependence on imported goods in the production, the level of import lowered as well. Therefore, we can observe in Figure 1 that all three indicators move in the same direction and react similarly to exogenous shocks.

Figure 1: GDP, Export and Import development
(annual percentage change)



Source: IMF, own estimation

Unfortunately, the Czech Republic gradually loses its sovereign position among post-Communist states, as the other countries are converging to the Czech performance. Table 19 in Appendix demonstrates that the Czech Republic's GDP per capita in PPS in 2012 still represents about 80% of the EU28's respective figure but there is no evident progress in the proportion since 1995. Whereas, other post-Communist countries, such as Slovakia or Poland, marked an apparent rise in GDP per capita relative to EU28.

2.2 Export Development

At the beginning of 90s, the Czech Republic was still closely adherent to the East and Czech export was mainly oriented to the former Soviet Union and its political satellites in Central Europe. However, during the 90s, four export promotion institutions were created – CzechInvest, Export Guarantee and Insurance Corporation, Czech Export Bank and CzechTrade. The institutions aimed to promote international trade by increasing competitiveness of Czech firms and facilitating them access to foreign markets (Janda *et al.,* 2013).

Obviously, the crucial step in the export development was the trade diversion from the former Communist block to Western markets. The accession to the European Union in 2004 facilitated not only the intra-European trade but thanks to numerous international agreements of the EU, the Czech Republic gained the possibility to enter markets that were not accessible before. Consequently, the amount of overall export raised significantly.

Nevertheless, the Czech Republic principally concentrates on European countries. Table 20 in Appendix shows that since 1995 over 80% of overall Czech export every year has been transferred to the European Union. To offer an illustration about Czech export partners, Table 1 presents countries, to which the Czech Republic exported the most in 2012.

Table 1: 10 major export partners for the Czech Republic in 2012

Country	**Code**	**Export** (value in mil. CZK)	**Share of total export**
Germany	DE	966,190	31%
Slovakia	SK	277,491	9%
Poland	PL	187,115	6%
France	FR	155,152	5%
United Kingdom	GB	147,084	5%
Austria	AT	141,646	5%
Russian Federation	RU	118,025	4%
Italy	IT	110,329	4%
Netherlands	NL	99,312	3%
Belgium	BE	73,511	2%
Total		**3,072,598**	

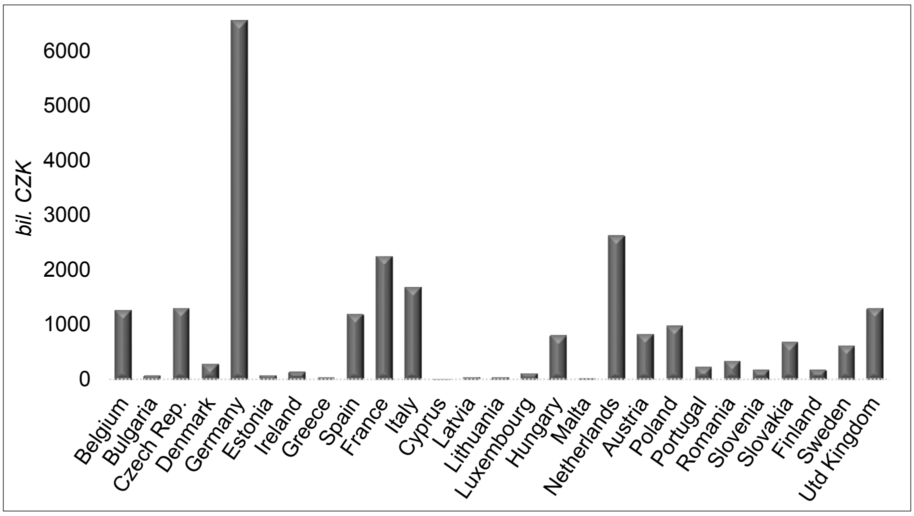
Source: CZSO, own estimation

As one can see, 9 out of the 10 largest export partners are members of the European Union. Regarding the Czech export destinations, we can raise a hypothesis that factors influencing the international trade are geographical distance, economic prosperity, market size and historical ties. Nevertheless, the share of export to Germany is enormous and this dependence can imply serious economic complications (the recent crisis is a good example, where trade destruction and forced trade diversion could play a role). Probably, also for that reason, the Czech government (i.e. its Ministry of Industry and Trade) has the aim to promote export to some advancing countries to diversify the export structure. It is a question how far such initiative could change the territorial export structure, which is one of the objectives of this study.

2.3 Export Articles

Czech economy is based on manufacturing industry, whose share in GDP is the highest among OECD countries. Correspondingly, the main export commodity category, according to 1-digit SITC, is machinery and transport equipment. Table 21 in Appendix reveals that this category accounted for almost 55% of overall Czech export in 2012. The dominance of machinery over other commodity groups is present at the European level as well, where the respective figure amounted to 40% of total export of the EU27.

Figure 2: Intra-EU27 export of machinery in 2012



Source: Eurostat, own estimation

The tradition of Czech manufacturing industry dates back to the 19th century, when Czech region was the most prospering economic part of the Austro-Hungarian Empire. The competitiveness of Czech machinery sector is still remarkable. Figure 2 compares amount of machinery export, traded within the European Union. Leaving alone the enormous amount of German export, the Czech Republic ranks among the biggest exporters, reaching the same level as Belgium or the United Kingdom. This is the reason why we have selected this particular sector as a target of our analysis.

2.4 Automobile Industry

Czech automobile industry plays a crucial role in the manufacturing sector and in the whole economy. There are several reasons for the success of the car production. The optimal location of the Czech Republic offers business opportunities for potential investors, including direct access to the European market. Moreover, high-level technical education, skilled labour force and large car industry capacities contribute to R&D and thus, to the car sector performance (Novosák *et al.*, 2010).

The Czech Republic produces international car brands with high quality. Essentially, there are three main carmakers in the Czech Republic: Škoda Auto (Volkswagen group), TPCA (Toyota & PSA Peugeot Citroën joint venture) and Hyundai Motor Manufacturing Czech. Together, all Czech car factories produced 1,171,747 passenger cars in 2012, which classified the Czech Republic among the 15 largest world car producers (see Table 23 in Appendix) and ensured the supreme position among world car leaders.

Many foreign investors see the potential of Czech environment and they increase the competitiveness of Czech factories by providing investment into the automobile sector. Consequently, the Czech Republic is able to export large share of domestically produced automobiles, stimulating economic performance. Furthermore, Czech industry is fully integrated into the European value chain, which ensures a quite stable position.

Export of automobiles develops in the same way as machinery or total export (Figure 3 in Appendix) but it is more resistant to external shocks. The reason is that the Czech Republic specializes in a production of lower-class passenger cars, which are a demanding article even in the period of economic troubles. The development of automobile export is illustrated in Figure 4 in Appendix, where the tendency of rising proportion over time is obvious.

3 GRAVITY MODEL

Gravity model is a common econometric tool, especially while estimating international trade flows from the point of view of economic geography. In this chapter, we firstly summarize existing work about the gravity model. Afterwards, we outline its initial development and at the end, we present some typical risks and several implemented estimation methods.

3.1 Literature Review

The popularity of gravity models in international trade has been growing steadily in last years. We can observe this trend in an increasing number of recent works dealing with different ways of estimating gravity models. The most common economic outcome, estimated by gravity models, is still the amount of international trade but more and more authors use gravity equations for explaining other types of flows. For example, Frankel & Wei (1993) applied this approach on exchange rate variability, Rose (2000) showed an impact of currency unions and Janda *et al.* (2013) studied the effect of credit support for export.

3.1.1 Initial Development

Tinbergen (1962) enabled this phenomenon, while he first estimated trade flows based on a gravity equation. Tinbergen reported a mathematical formulation as an analogy to the Newton's law of universal gravitation and applied it on estimating bilateral international trade. Despite an immediate reception of the gravity approach, numerous criticisms arose, pointing at an insufficient theoretical foundation of the model.

Accordingly, Anderson (1979), as the first one, came up with an attempt to develop a theoretical foundation for the model. He introduced a term "average barrier", describing resistance to trade of two regions with respect to all their partners. Afterwards, many researchers have been developing this concept by discussing multilateral resistance terms. Regardless of the lack of micro-economic foundation, the majority of authors followed the gravity pattern estimation, seeing in it a simple and effective way of explaining any type of flows.

Besides, some authors have theoretically grounded gravity models using different trade theories, for instance Deardorff (1998) explained the gravity equation through factor endowment theories or Helpman & Krugman (1985) described it with the help of increasing returns to scale. Later, McCallum (1995) underlined the utility of the gravity equation and suggested a new concern of several subsequent works – "border effect". The "border effect" attracted so much attention mainly because of an irrelevant coefficient predicted by the analysis. The estimated parameter had indeed an expected sign but the size of the coefficient was inflated.[1]

Together with a solution to the McCallum's "border puzzle", Anderson & van Wincoop (2003) have come out with the first widely accepted theoretical foundation of the gravity model. The work resumed the ideas outlined already by Anderson (1979), especially the role of the multilateral prices. Authors underlined the importance of including country-specific dummy variables in the gravity equation to be able to estimate consistently the model and to deliver a correct comparative static analysis. Thus, the contribution of this paper lies both in explanation of the inflated border effect and in providing a useful methodology based on a theoretical justification.

Consequently, Baldwin & Taglioni (2006) further developed a concept of the multilateral resistance and generalized it for the use with panel data methods. To clarify, they enabled varying over time by including time-varying dummy variables. It therefore allows for removing also the time series part of bias and thus, for a correct estimation. Above all, the authors introduced a six-step derivation of the gravity equation. The procedure is purely based on the Anderson & van Wincoop's theory but offers a simple and widely implemented way of gravity derivation.

[1] For instance, McCallum obtained a coefficient of 22 for the border effect between U.S. and Canada, i.e. without border the U.S. - Canada trade would have risen by 2,200 per cent with respect to the inter-provincial trade in Canada.

One of the weaknesses behind the theory of Anderson & van Wincoop (2003) was the assumption of symmetric bilateral trade costs. This question was a motivation for the work of Baier & Bergstrand (2009). They augmented the literature dealing with the role of multilateral resistance and suggested a method that yielded exogenous multilateral terms. The authors used a "first-order log-linear Taylor series expansion of the system of nonlinear price equations" and with this approximation method, allowing for asymmetric bilateral trade costs, obtained better results than with the Anderson & van Wincoop's approach.

3.1.2 Estimation Techniques

Due to the vast gravity literature, there are many different procedures for estimating gravity models. One common methodology, followed by majority of researchers in the past, was to estimate a log-linearized version of a gravity equation. However, many worried about the existence of zero-flows that are frequent in international trade and which would cause undefined logarithm. Accordingly, several suggestions have appeared in the literature with a purpose to solve this issue.

One of the approaches, suggested for example by Linnemann (1966), is simply omitting zero-flows from the sample. Another approach is adding a small constant to zero trade flows (e.g. Rose & Spiegel, 2010). These are simple and widely adopted methods, delivering a technical well-defined logarithm. Nevertheless, these methods are only valid, if the zeroes are randomly distributed; otherwise, they would produce a sample selection bias (as it is usually the case).

Heckman (1979) or Helpman *et al.* (2008) introduced a solution in the form of a sample selection correction, or a Heckman estimator. The applied method is called a two-stage estimation technique and was designated to deal with the sample selection bias. The first stage consists of identifying the probability of trade between countries i and j. Afterwards, the second stage estimates the log-linearized gravity equation under the condition from the first stage. In brief, the two-step procedure offers a solution for treating zero trade flows but does not correct for another potential source of biased estimates – the heteroscedasticity.

From that perspective, a more suitable technique seems to be a Poisson pseudo maximum likelihood estimation, promoted especially by Santos Silva & Tenreyro (2006). Authors referred to the Jensen's inequality as a reason for biased estimation of log-linearized equations. In the presence of heteroscedasticity, the log-linearized equations produce biased estimates by definition. For this reason, they abandoned the standard procedure and estimated gravity equation in its original multiplicative form. The problem with zero trade flows disappears and the model corrects for heteroscedasticity. This method has many followers; for instance Westerlund & Wilhelmsson (2009) or Babecká *et al.* (2010).

Due to the incontestable advantages of the Poisson estimator, the method is becoming more and more popular in the recent literature. On the other hand, a considerable number of researchers implemented a fixed or random effects approach for the gravity models (eg. Caporale *et al.*, 2009; Carrère, 2006 or Egger, 2002). The fixed effects method controls for the potential bias caused by omission of invariant factors but the problem is that one can include only variables not collinear with the fixed effects. On the contrary, the random effects approach allows us to include the collinear variables with the fixed effects and this method usually proves to be more efficient but it is consistent only under restrictive assumptions. Therefore, there is a trade-off between the consistency and efficiency of the estimator and hence, many researchers apply a combination of these two approaches for gravity modelling.

No matter what approach they use, a majority of researchers estimate gravity models of final goods and services at aggregated level. To clarify, they use the trading data for final goods and services and determine influencing factors for total volume of trade. Only few studies focus on a trade of intermediate goods and services. We can cite particularly Miroudot *et al.* (2009) or Yane (2013), who also involved disaggregated trade data.

3.2 History

It has passed many years since a Nobel prized economist from year 1969, Jan Tinbergen, came out with an idea of an application of the Newton's law of universal gravitation to economics. The inspiration from physics was not a coincidence, as Tinbergen studied mathematics and physics in the Leiden University in Netherlands.

This Dutch economist formulated an econometric analysis that marked a breakout in international trade research.[2] In his gravity equation, he defined volume of international trade between two countries, as being positively related to economic sizes of the two countries (measured in GDP) and negatively related to distance between them. In fact, it exactly reproduces the law of universal gravitation, which states that force of gravitation between two objects is directly related to product of their masses and indirectly related to square distance between them (Newton, 2002). For illustration, the basic form of the gravity equation is defined as follows (Bergeijk & Brakman, 2010):

$$T_{ij} = \frac{GDP_i^{\alpha} GDP_j^{\beta}}{D_{ij}^{\theta}}$$

where T_{ij} indicates bilateral trade between countries i and j; GDP_i and GDP_j represent economic sizes of two countries, respectively; and D_{ij} stands for the distance separating both countries. The coefficients α, β and θ represent sensitivity of bilateral trade regarding both countries' GDP and the distance between them.

As we mentioned previously, researchers generally estimate the equation in a log-linearized form rather than multiplicative. Correspondingly, after adding a regression constant c and a random error term ε_{ij}, the model would take the following form:

[2] The formulation is included in the Appendix 4 in Tinbergen (1969).

$$\log T_{ij} = c + \alpha \log GDP_i + \beta \log GDP_j + \theta \log D_{ij} + \varepsilon_{ij}$$

The coefficients α, β and θ represent respective elasticities, which are to be estimated.

Even though the gravity model had a remarkable success, the primary development was linked with the lack of micro-economic foundation. The popularity of gravity models laid particularly in its simplicity and significant explanatory power, as they were able to explain many different economic phenomena by indicating various influencing factors. However, the unsatisfactory theoretical foundation implied rather suspicious acceptance of empirical results and therefore undervalued the merit of the gravity model.

The search for a sound theoretical foundation was the aim of many researchers. The influx of suggestions for theoretical explanations of the gravity model has launched Anderson (1979) but his contribution was not adequately appreciated. A study that was quoted the most for justifying the derivation of gravity equations has become Anderson & van Wincoop (2003). Authors extended Anderson (1979)'s findings, overcame their weaknesses and suggested a reasonable theoretical explanation.

The main feature of their work is the implementation of multilateral resistance terms into a model. These multilateral resistance terms capture trade costs across all trading partners and are an influencing factor for every single bilateral trade. In early literature, researchers did not include these country-specific resistance terms, which are correlated with trade costs (covered in the distance variable). Their models consequently suffered from an omitted variable bias and estimated parameters were not consistent (Shepherd, 2013).

As we noted previously, the clear and reasonable suggestions of Anderson & van Wincoop (2003) have implied that their work has become a main reference for subsequent work. Baldwin & Taglioni (2006), for example, resumed ideas of Anderson & van Wincoop (2003) by introducing a simple six-step derivation of gravity equations. Baldwin & Taglioni (2006) just resumed the theoretical derivation; beginning with a supply-demand equation, passing through the CES demand structure and trade costs and finalizing by combining sequentially derived equations. Finally, they obtained the standard formulation of the gravity equation.[3] This simplified procedure has been thereupon widely quoted, especially because of the possibility of implementing it on time series and panel data analyses.

3.3 Risks

The gravity model serves most often as a tool for estimating trade flows. Nevertheless, there are two problematic features associated with trade data: heteroscedasticity (since we are mixing trade data for very large and very small countries) and a substantial number of zero-flows. Some datasets can consist of more than 50% of zero-flows and this number is rising exponentially with the level of disaggregation in data. Similarly, heteroscedasticity is typical for trade datasets because they consist of bilateral trade flows of multiple countries.

The heteroscedasticity itself does not cause estimates to be biased or inconsistent but particular problems arise while logarithmic transformation. Estimation of log-linearized equations by Ordinary Least Squares (OLS) is common in early gravity work that usually neglected the issue of heteroscedasticity and concentrated on cross-section data for a single year. In a response, Santos Silva & Tenreyro (2006) pointed out the appropriateness of this approach and introduced a reasonable explanation for their argumentation. They reminded the Jensen's inequality that implies the following:

$$E(\log y) \neq \log E(y)$$

It means that the expected value of the logarithm of a random variable does not equal to the logarithm of its expected value. This equation is crucial while estimating log-linearized equations by OLS in the presence of heteroscedasticity. To explain, one of the assumptions for consistency of OLS is the statistical independence of the error term ε_{ij} on regressors (*zero conditional mean* condition) (Wooldridge, 2003, pp. 47).

[3] For a detailed derivation of the six-step procedure, see Bergeijk & Brakman (2010)

The expected value of $\log \varepsilon_{ij}$ depends on both, the expected value of ε_{ij} and its variance. The potential heteroscedasticity in the multiplicative equation then implies that the $\log \varepsilon_{ij}$ depends on regressors and violates the *zero conditional mean* condition. Therefore, resulting estimates are to be biased and inconsistent.

Similarly, zero flows are problematic particularly in case of the logarithmic transformation. The reason is that the logarithm of zero is not defined and thus, some correcting procedures have to be implemented. There are generally three sources of these zeroes and it is almost impossible to distinguish them. Rounding errors, missing observations or the actual lack of trade flows usually stand for zero occurring. The zeroes are not randomly distributed because smaller countries (with lower GDP) are more likely to mark zero flows, whether caused by the measurement error or the actual zero trade flow. Therefore, two standard correcting procedures, i.e. omitting zeroes or adding a small constant, produce sample selection bias and inconsistency of parameters of interest.

A typical suggested solution to the zero flows problem is to truncate the sample, or to implement a variation of the Heckman sample selection model. The method in the first stage explains why the trade takes place and in the second stage estimates the equation on a subsample conditional on trade occurring (Xiong & Chen, 2012). On the other hand, the Heckman model usually does not deal with heteroscedasticity and therefore, as a more appropriate procedure, Santos Silva and Tenreyro (2006) introduced the Poisson pseudo maximum likelihood estimator. [4]

Last but not least, as with any other econometric analysis, there are risks connected to a violation of an assumption, crucial for a specific estimation technique. The most common source of omitted variable bias is country heterogeneity. It is thus necessary to allow for some country-specific factors and choose an appropriate estimation technique based on a specific dataset.

[4] We will discuss the PPML estimation more closely in section 3.5.

3.4 Types of Estimation

The initial Tinbergen's gravity model was gradually augmented by additional variables to increase the explanatory power (e.g. adjacency, numerous policy institution variables or size of population). Thus, gravity models have become the focus of many researchers, who have consequently introduced various estimation methods. In this section, we will discuss more closely only those methods that are relevant for our study.

A common approach in recent works was to estimate gravity models as cross-sections. However, this method could not efficiently describe trade data and very often suffered from omitted variable bias, caused by country heterogeneity (Shepherd, 2013). For parameters to be consistently estimated, one has to add time dimension into a model, which controls for variables with changing values over time. Thus, panel data estimation with both cross-sectional and time series dimension has become a standard tool.

An important point while estimating panel data models is to include unobserved heterogeneity, or fixed effects, which capture all unobserved time-invariant factors, affecting a response variable. The time-varying error, called idiosyncratic, cannot be ignored either (Wooldridge, 2003, pp. 420-421). In the panel, there are constant (distance) and varying factors (GDP) and one has to select a correct estimation technique to control for both types.

A basic procedure, which serves generally as a benchmark estimation, is pooled Ordinary least squares (OLS). Two crucial conditions for consistent estimates are that the unobserved heterogeneity cannot be collinear with any regressors and idiosyncratic errors have to be homoscedastic and serially uncorrelated across time. Especially the requirement for unobserved heterogeneity is usually violated and this method hence produces biased and inconsistent estimates (Wooldridge, 2003, pp. 409-413). Accordingly, the pooled OLS procedure is the simplest method, which is though inappropriate for proper estimation and it is used only for a comparison with more advanced tools.

An alternative approach that allows for an arbitrary correlation of the heterogeneity error and regressors is called fixed effects transformation. A weakness of this approach lies in inability of including time-constant variables into a model and inability of estimating their parameters of interest. For this reason, the random effects procedure offers a solution, which allows for inclusion of variables correlated with fixed effects. A crucial assumption of this approach is that the heterogeneity effect has to be uncorrelated with each explanatory variable in all time periods. The RE method produces more efficient coefficients than with the FE under this condition but the condition for consistency is obviously too restrictive.

The above-mentioned methods are those, employing a log-linearized version of a gravity equation. Consequently, they are likely to produce inconsistent estimates due to the omitted variable and heteroscedasticity bias. Santos Silva & Tenreyro (2006) recommended to estimate gravity equation in its multiplicative form and suggested to use a Poisson pseudo maximum likelihood estimator. The PPMLE is in last years the most widely implemented tool and we apply it for our study as well.

There are several other techniques, which all deliver some improvements, but also drawbacks. For illustration, Martin & Pham (2008) partly dented reputation of the PPMLE, as they underlined its weak performance in the presence of large number of zeroes. Authors suggested a Tobit estimator that is just one type of the Heckman estimator, proposed by Heckman (1979) or Helpman *et al.* (2008). The Heckman two-step estimator can effectively deal with zero trade flows but is heavily dependent on an accurate identification of a restriction criterion for the first stage.

Similarly, Martinez-Zarzoso (2013) disproved partially the results of Santos Silva & Tenreyro (2006) and compared different types of estimators in her work, such as a Gamma pseudo maximum likelihood, a Non-linear least squares or a Feasible generalized least squares. She concluded that the choice of an estimator has to be specific for each dataset because all of the approaches can produce misleading results for some samples. Other methods that several authors implement are a Hausman-

Taylor approach, employing instrumental variables to correct for bias (Egger *et al.*, 2001 or Brun *et al.,* 2005) or a Two-stage method of moments (Xiong & Chen, 2012).

3.5 Poisson Estimation

As we mentioned previously, the Poisson pseudo maximum likelihood estimator is the most popular tool for estimating gravity models in recent literature. Its popularity implied many criticisms but most of them are irrelevant. Moreover, Santos Silva & Tenreyro (2011) disproved the criticisms, confirmed their previous findings from year 2006 and classified the PPMLE as the most appropriate estimator. Many recent authors employed PPMLE for their research (Westerlund & Wilhelmsson, 2009; Babecká Kucharčuková *et al.*, 2010 or Bobková, 2012), prospering from its efficiency and easy interpretation.

The originality of the PPMLE lies in estimating the gravity equation in its multiplicative form and modelling the expected value as an exponential function (as follows), which ensures that y is non-negative.

$$E(y|x_i) = \exp(x_i' \beta), i = 1, \ldots, n$$

Next, we obtain the estimates by maximizing the log-likelihood function, taking the following form (Wooldridge, 2003, pp. 548):

$$\mathcal{L}(\beta) = \sum_{i=1}^{n} \ell_i(\beta) = \sum_{i=1}^{n} \{y_i \boldsymbol{x}_i \boldsymbol{\beta} - \exp(\boldsymbol{x}_i \boldsymbol{\beta})\}$$

An advantage of the Poisson estimation is that for producing consistent estimates, variables do not have to be Poisson distributed. The requirement for equality of the conditional variance and the conditional mean is rarely fulfilled but we can still obtain unbiased results. Nevertheless, one has to adjust standard errors by a correcting technique, such as clustering.

The Poisson estimator is usually used for count data models but can be generalized for non-linear models such as gravity as well. The PPMLE offers us a useful tool for dealing with common zero trade flows and avoiding issues connected with heteroscedasticity in log-linearized equations. Moreover, the interpretation of the estimated coefficients is straightforward and is in line with the interpretation of standard OLS estimates.

4 EMPIRICAL ANALYSIS

The gravity model is a widely accepted tool due to its indisputable advantages that we discussed previously. Given the large number of works implementing gravity models, it may seem that this area of research is saturated, converging potentially to a standard analytical routine. This is far from being the truth, as was pointed out by Baldwin & Taglioni (2006), Shepherd (2013) and by the research, which follows the advances in the paper of Eaton & Kortum (2002) and Helpman *et al.* (2008).

Our work searches to provide an innovative way of employing the gravity model. The majority of existing works study mutual relationships among all countries, participating in international trade. Therefore, obtained results are indeed applicable on all countries but they are average, i.e. one cannot draw implications for a specific country. In contrast, our study concentrates solely on Czech export and we get concrete results, which are directly related to Czech environment.

We will compare three models at different levels of aggregation. We begin with estimating general export function for the Czech Republic and going through machinery industry, we end with modelling factors influencing Czech export of automobiles. To our knowledge, our work offers the first attempt of a similar approach. The most common procedure in existing literature is to estimate total trade at aggregated level, rarely there are works studying flows of intermediates, let alone a comparison of differently aggregated trade flows. Consider please our research as a pioneering venture testing the viability of gravity models for different levels of aggregation.

4.1 Data Characteristics

We have collected data for 50 countries over period from 1995 to 2012, yielding 900 observations for each model. The 50 countries represent trading partners, to which the Czech Republic export the most automobiles at the beginning of the relevant period and the same dataset is employed for all three models.[5] This approach allows us to restrict the number of zero flows to minimum and avoid related difficulties.

One can argue that the sample of 50 countries is not predicative but the total export to these countries represents almost 95% of Czech export worldwide. A possible extension of the sample, and thus a prevention of a sample selection bias is left for further research. In this work, we focus on suggesting reasons for the variation of the amount exported to individual countries and we pay particular attention to differences among the three models. Hence, we leave behind rationalization for the actual existence of international trade.

We have obtained data for Czech export, the dependent variable, through the Czech Statistical Office (2012). To get relevant data for each model, we employ Standard International Trade Classification (SITC) system. Machinery is classified at one digit level with a code 7, and we use a commodity grouping called *Motor cars* with a 3-digit code 781 to obtain values for export flows of automobiles. The export data are expressed in CZK in nominal terms in order to get correct, unbiased results, which will be relevant for Czech exporters.

4.2 Variables Description

In this section, let us introduce the explanatory variables used in our model. We summarize all variables and their properties in Table 2, which is structured according the way of interpreting estimated parameters of interest. The choice of variables follows standard theoretical approaches, for example those of Baldwin & Taglioni (2006) or Babecká *et al.* (2010).

[5] Table 26 in Appendix A covers a complete list of the countries.

Table 2: List of explanatory variables

Variable	Name	Values	Units	Expected sign	Source
GDP PPP	GDP_{jt}		mil. CZK	+	*IMF, own estimation*
GDP PPP (home)	GDP_{it}		mil. CZK	+	*IMF, own estimation*
Distance	$Dist_{ij}$		km	-	*CEPII*
Population	Pop_{jt}		mil.persons	-	*World Bank*
ERDI	$ERDI_{jt}$	US ≅ 1	units	-	*Own estimation*
Tariff freedom	T_{jt}	0-100	per cent	+	*Econ. Freedom dataset, own estimation*
Regulatory trade barriers	Reg_{jt}	0-100	per cent	+	*Econ. Freedom dataset, own estimation*
Education	$Educ_{jt}$	0-100	per cent	+	*World Bank*
Government effectiveness	G_{jt}	0-100	per cent	+	*World Bank*
Political stability	PS_{jt}	0-100	per cent	+	*World Bank*
Business Freedom	Bus_{jt}	0-100	per cent	+	*Heritage Foundation*
Trade Freedom	$Trade_{jt}$	0-100	per cent	+	*Heritage Foundation*
Fiscal Freedom	Fis_{jt}	0-100	per cent	+	*Heritage Foundation*
Government Spending	Gov_{jt}	0-100	per cent	+	*Heritage Foundation*
Monetary Freedom	Mon_{jt}	0-100	per cent	+	*Heritage Foundation*
Investment Freedom	Inv_{jt}	0-100	per cent	+	*Heritage Foundation*
Financial Freedom	Fin_{jt}	0-100	per cent	+	*Heritage Foundation*
Property Rights	Pro_{jt}	0-100	per cent	+	*Heritage Foundation*
Freedom Corruption	Cor_{jt}	0-100	per cent	+	*Heritage Foundation*
EU/Eurozone	EU_{jt}	1/2/3		+	*Own estimation*
Recession dummy (home)	Rec_{it}	0/1	dummy	-	*CZSO, own estimation*
Common Border	Bor_{j}	0/1	dummy	+	*Own estimation*
Austro-Hungary	AH_{j}	0/1	dummy	+	*Own estimation*
Landlockedness	$Land_{j}$	0/1	dummy	-	*Own estimation*

Source: own estimation

The essential explanatory variables in a gravity model are GDPs and a distance variable. The distance is assumed to have a diminishing effect on export, as it increases the transportation costs, and the expected sign of the parameter is thus unambiguously negative. GDP variables represent economic masses that explain the countries' potency to trade internationally. We use GDP at PPP (Purchasing Power Parity) for our model, which is more suitable for international comparison of countries' wealth. We have retrieved data in USD from IMF and converted them into CZK by the annual average USD-CZK exchange rate, computed by OECD.

As the exchange rate does not involve the purchasing power poverty, we have to include ERDI variable to control for this inaccuracy. ERDI (exchange rate deviation index) indicates by how many times the real exchange rate is lower than the PPP rate and is obtained as a ratio of GDP at PPP and GDP in current prices. The ERDI is calculated with respect to some currency (USD in our case) and takes value of one for the USA.

Poorer countries have generally higher values of ERDI and the value converges to one for developed countries. Higher values of ERDI imply weaker currencies and thus, a lower purchasing power in comparison with the USA. Consequently, these countries have lower ability to import goods and services from abroad.

One can argue that we should use rather ERDI with respect to the Czech Republic. To get that, we would employ another conversion constant, which could produce a measurement error if chosen inappropriately and besides, the results would not be significantly different. Therefore, we leave the original variable and expect a negative sign of the estimated parameter.

The group of explanatory variables, which are to be taken in logarithm, is completed with a population variable. The size of population serves as a proxy for the openness of an economy. We expect that larger population have lower tendency to trade internationally. The more people there are in an economy, the more self-sufficient the economy is and thus, there is less need to import goods and services.

Another group of variables are those expressed in per cent with an expected positive sign of estimated parameters. Let us first explore tariff freedom and regulatory trade barriers. Both variables are obtained from Economic Freedom Dataset. Authors of the dataset assign values from 0 to 10 to countries according to their support of international trade. The maximum value of 10 would receive a country imposing no import tariffs or setting no international trade barriers. To facilitate the interpretation of the coefficients, we multiplied the figures by 10 to receive a percentage scale.

Next, an education variable could have a significant predicative power. In our case, the variable represents the ratio of tertiary education graduates to total population at the typical age of graduation. Education should be positively related to the amount exported, as educated inhabitants recognize the incontestable economic advantages of international trade. Similarly, government effectiveness and political stability should take a positive sign of parameter.

Other 9 variables are various institutional factors, characterising each importing country. There are higher for developed countries as the economic growth is generally evolved from a good institutional base. The expected sign for all coefficients is positive. However, it can happen that some variables will be omitted from the estimation due to multicollinearity.

There are also several discrete variables in our model. The EU variable, indicating the membership in the European Union or Eurozone. We divided countries into 3 groups: non EU-members, EU-members and Eurozone members and assigned them values from 1 to 3, respectively. We suppose that a membership in a European community increases the financial stability (the further integration, the better) of international transactions and thus, the expected sign is positive.

The first dummy variable, recession, stands for a recession in home country, i.e. the Czech Republic. By definition, a country is in recession if, in at least two consecutive quarters, the economic growth is negative. We naturally expect it to influence negatively the response variable. As well as the landlockeness, which is a dummy, taking value one in case of a landlocked country. Although the Czech Republic has no access to the sea, shipping is one of the most common means of transport facilitating international trade.

Finally, both border and Austro-Hungary dummies are positively related to Czech export. A common border lowers transportation costs connected with trade and decreases the delivery period. Former members in the Austro-Hungarian Empire share part of the historical development and thus, maintain closer relationships among each other.

4.3 Methodology

We apply a panel data analysis for the research. Panel data denote that various characteristics of the 50 countries are observed across time and in opportunity costs in the competition between countries (i.e. potential partners). Therefore, panels are more informative than simple time series or cross-sections because they have two dimensions, implying more observations. They allow us to control for individual heterogeneity and consequently, estimated coefficients are less likely to be biased due to an omitted variable.

To estimate impacts of various determinants of Czech export, we employ an augmented gravity model of trade, whose properties were described in Chapter 3. For the actual analysis, the choice of estimator plays a crucial role. We have decided to follow recent researchers in applying the Poisson pseudo maximum likelihood estimator, as it is assumed to provide the most consistent estimates.

However, for a comparison of estimated parameters, we employ simple OLS regression with time or country dummies as a benchmark model. The log-linearized form of the gravity equation is typically used for the OLS estimation. Hence, as mentioned early, in the presence of heteroscedasticity, the OLS coefficients are likely to be inconsistent as trade data generally fail to fulfil the very restrictive assumption on properties of error term.[6] Another source of biased coefficients is the omission of zero trade flows because of the undefined logarithm of zero.

The PPMLE proposes a viable alternative that should correct for both types of biases. Accordingly, the obtained results differ significantly under these two estimations, suggesting the inconsistency of OLS estimates on top of the different optimisation procedures in the estimation routines. To control for heteroscedasticity, we cluster standard errors.[7] It means that we allow for correlation of standard errors within a group, i.e. country. Clustering actually replaces standard robust command.

Different results can be attained by the fixed or random effects approach. The fixed effects model is theoretically more appropriate for our analysis because our sample of countries is not randomly chosen. Moreover, we confirm our hypothesis by running a Hausman test, which has rejected the null hypothesis in favour of fixed effects approach in most cases.

Our estimations are hence based on fixed effects method. However, using standard fixed effects method, we would be forced to drop all invariant variables from our model (e.g. the distance), which is not our intention. We use instead Least Squares Dummy Variables (LSDV) procedure that allows us to keep crucial variables in the model, even though they are constant over time (entities).

The LSDV method lies in estimating the gravity equation by OLS with a set of country (time) dummies. In other words, we would include 49 country dummies (one is omitted due to multicollinearity) or 17 time dummies into the equation. We want to keep some invariant variables in the model (especially the distance, which is the generic variable of gravity models). Therefore, we have to omit additional country dummies because they can be correlated with some other “sluggish” variables.

[6] The issue was discussed in section 3.3.

[7] The presence of heteroscedasticity was discovered by a Breusch-Pagan test.

One of the uncommon aspects of our analysis is estimating the gravity model under different data structures. It means, once as time series and then as cross-sections. We estimate the equation as 50 time series with 18 observations per group and consequently, as 18 cross-section with 50 observations per group. That obviously suggests a radically different concept.

This distinction certainly plays a significant role because each approach follows a different logic and accordingly, the interpretation of coefficients differs. For time series dimension, we observe the development of countries' behavioural characteristics over time. Whereas for cross-sections dimension, we look at the characteristics of decision-making as varying factors across entities in a given year. Therefore, the estimated coefficients can differ significantly and we focus on interpreting these differences.

4.4 Model Specification

While applying the OLS procedure (i.e. the LSDV), we use the log-linearized form of the gravity equation for the regression. As mentioned previously, variables that are not expressed in per cent or are not dummies, are taken in logarithm in order to get elasticities and to be able to simply interpret their estimates. The initial model is designed as follows:[8]

$$\log X_{ijt} = \beta_0 + \beta_1 \log GDP_{jt} + \beta_2 \log GDP_{it} + \beta_3 \log D_{ij} + \beta_4 \log ERDI_{jt} + \beta_5 \log P_{jt} + \beta_6 T_{jt} + \beta_7 RT_{jt} + \beta_8 Educ_{jt} + \beta_9 G_{jt} + \beta_{10} PS_{jt} + \beta_{11} Bus_{jt} + \beta_{12} Trade_{jt} + \beta_{13} Fis_{jt} + \beta_{14} Gov_{jt} + \beta_{15} Mon_{jt} + \beta_{16} Inv_{jt} + \beta_{17} Fin_{jt} + \beta_{18} Pro_{jt} + \beta_{19} Cor_{jt} + \beta_{20} EU_{jt} + \beta_{21} Rec_{it} + \beta_{22} Bor_j + \beta_{23} AH + \beta_{24} Land + \log e_{ijt}$$

$$j = 1, \dots, 50$$
$$t = 1, \dots, 18$$

[8] Country or time dummies are not stated in the equation.

The respective model designated for the Poisson regression takes the following form:

$$X_{ijt} = \exp(\beta_0 + \beta_1 \log GDP_{jt} + \beta_2 \log GDP_{it} + \beta_3 \log D_{ij} + \beta_4 \log ERDI_{jt} +$$
$$\beta_5 \log P_{jt} + \beta_6 T_{jt} + \beta_7 RT_{jt} + \beta_8 Educ_{jt} + \beta_9 G_{jt} + \beta_{10} PS_{jt} + \beta_{11} Bus_{jt} +$$
$$\beta_{12} Trade_{jt} + \beta_{13} Fis_{jt} + \beta_{14} Gov_{jt} + \beta_{15} Mon_{jt} + \beta_{16} Inv_{jt} + \beta_{17} Fin_{jt} +$$
$$\beta_{18} Pro_{jt} + \beta_{19} Cor_{jt} + \beta_{20} EU_{jt} + \beta_{21} Rec_{it} + \beta_{22} Bor_j + \beta_{23} AH + \beta_{24} Land + e_{ijt})$$
$$j = 1, \ldots, 50$$
$$t = 1, \ldots, 18$$

With such a large number of variables, the results can be biased. For that reason, we implement VIF test after the OLS regression to reveal potential multicollinearity issues. In case of a high value, we have two possibilities how to proceed.[9] If the problematic variable is an important one (e.g. GDP), we explore a correlation matrix to discover which variables it is highly correlated with. Consequently, we drop these variables from the model (if they are not crucial for us). On the other hand, if the variable with a high value of VIF is for example a country dummy or other less important determinant, we simply omit it to ensure the relevance of our analysis.

We have to pay particular attention while dealing with institutional variables. Some of them are subject to low variability (i.e. inertia) in time, which can lead to problems with multicollinearity and can influence estimates of other variables. As a result, we had to keep only four institutional indices. Similarly, as there are many country (time) dummies, some of them may influence estimates of essential determinants. It is hence preferable for us to reveal the problematic dummies and drop them from the model.

The set of explanatory variables is thus substantially reduced. The explanatory power slightly lowers but the estimates are more statistically significant and we are able to draw meaningful conclusions from the results. The corrected model for the OLS analysis takes the following form:

[9] A value above 10 is a sign of a multicollinearity issue.

$$\log X_{ijt} = \beta_0 + \beta_1 \log GDP_{jt} + \beta_2 \log GDP_{it} + \beta_3 \log D_{ij} + \beta_4 \log ERDI_{jt} + \beta_5 RT_{jt} + \beta_6 Educ + \beta_7 Bus_{jt} + \beta_8 Gov_{jt} + \beta_9 Inv_{jt} + \beta_{10} EU_{jt} + \beta_{11} Rec_{it} + \log e_{ijt}$$

$$j = 1, \ldots, 50$$
$$t = 1, \ldots, 18$$

Analogically, the model for the Poisson regression is designed as follows:

$$X_{ijt} = \exp(\beta_0 + \beta_1 \log GDP_{jt} + \beta_2 \log GDP_{it} + \beta_3 \log D_{ij} + \beta_4 \log ERDI_{jt} + \beta_5 RT_{jt} + \beta_6 Educ + \beta_7 Bus_{jt} + \beta_8 Gov_{jt} + \beta_9 Inv_{jt} + \beta_{10} EU_{jt} + \beta_{11} Rec_{it} + e_{ijt})$$

$$j = 1, \ldots, 50$$
$$t = 1, \ldots, 18$$

Last but not least, for our analysis it is crucial to apply the same procedure on all three models (overall, machinery and automobile) in order to propose a relevant comparison of predicted parameters. We use the same set of explanatory variables for all regressions, even if it would be preferable in some cases to include a different set. Therefore, we have to observe carefully the development of the correlation matrix to reveal possible collinearity issues and rationalize some unexpected coefficients.

5 TOTAL EXPORT MODEL

In this section, we present results from our econometric analysis of total Czech export. We indicate predicted coefficients and propose an argumentation about their sign and significance. Our aim is to stress differences between the two estimation techniques (the LSDV and the PPML) in order to reveal potential bias in estimates based on log-linearized form of the gravity equation.

Moreover, we confirm suggestions of Hyžíková (2012), for example, who stated that the distinction between two structures of a dataset is essential, as the estimates are not equal. Besides, the interpretation of the coefficients differs as well because each structure examines data in a different way. Therefore, we run regressions under both types and for each of them, the predicted coefficients are obtained by the OLS and the Poisson regression, in order to cover all possibilities.

Nevertheless, we have to underline a severe weakness of our analysis that lies in a limited number of countries included in the sample. To be precise, we have created a model, which is subject to a sample selection bias because the countries were not chosen randomly. Therefore, we have to consider this weakness while interpreting results and we cannot draw general implications. In fact, we are particularly interested in respective figures of the three models under different specifications. Hence, we anticipate that the subsequent interpretation of results is relevant only for our country set.

For clarity, we divide this section into two parts. Firstly, we study the time series structure of the dataset and interpret results obtained by both estimation methods. In the second part, the cross-sectional data structure is examined and the two sets of estimates are compared.

5.1 Time Series Data Structure

Our dataset is now structured as 50 time series. Thus, the estimation considers particularly the development of indices over time. Hence, with the time series structure, we examine a decision making process of Czech exporters rather from a long-term perspective. Alternatively, the analysis allows for a possibility of exporters to adjust accordingly to changes in partners' countries.

As mentioned previously, after running the Hausman test, we have concluded for the fixed effects estimation due to the inability of our data to fulfil the strict conditions, necessary for obtaining unbiased estimates by the random effects approach. Therefore, we add country dummies into the model for estimating by OLS and PPML.

The first step is to estimate the benchmark model based on the log-linearized equation. Table 3 presents the OLS coefficients together with their significance and confidence interval.[10] Firstly, we can say that the model is specified quite well as the R-squared equals to 0.9172. That means that the explanatory variables account for almost 92% of observed variation in Czech export in the data.

All explanatory variables, except regulatory trade barriers, have expected signs. Due to the clustering of standard errors, correcting for heteroscedasticity, the p-value is generally quite high. Nevertheless, crucial variables (the GDPs and the distance) are significant at 0.01 - 0.05 levels and influence substantially the response variable.

Regarding the exact values, one per cent increase in partner's GDP would cause Czech export to increase by 0.65 per cent, assuming all other factors fixed. Home GDP has a similar positive impact on the response variable. However, the predicted distance elasticity even exceeds unity and implies a strong negative effect of geographical

[10] For this and subsequent tables holds that country or time dummies are not stated in the table, values are rounded to four decimal places and variables significant at 0.1 level are marked in bold.

distance on Czech export. Thus, these three variables are, as expected, highly statistically and economically significant.

Table 3: Total export: time series data structure results – LSDV estimation

Linear regression

Number of observations = 900
R-squared = 0.9172
Root MSE = 0.68276

(Std. Err. adjusted for 50 clusters in Country)

$log(X_{ijt})$	Coefficient	Robust Std. Err.	p-value	[95% confidence interval]	
$log(GDP_{jt})$	**.6531**	**.0858**	**0.000**	**.4807**	**.8255**
$log(GDP_{it})$	**.5238**	**.2137**	**0.018**	**.0943**	**.9533**
$log(Dist_{ij})$	**-1.1339**	**.1694**	**0.000**	**-1.4744**	**-.7935**
$log(ERDI_{jt})$	-.4232	.2749	0.130	-.9752	.1288
Reg_{jt}	**-.0185**	**.0074**	**0.016**	**-.0334**	**-.0036**
$Educ_{jt}$	**.0138**	**.0062**	**0.032**	**.0013**	**.0263**
Bus_{jt}	.0099	.0074	0.186	-.0050	.0248
Gov_{jt}	.0043	.0038	0.265	-.0034	.0120
Inv_{jt}	.0068	.0064	0.290	-.0060	.0196
EU_{jt}	**.5028**	**.1413**	**0.001**	**.2189**	**.7867**
Rec_{it}	-.0049	.0373	0.897	-.0798	.0701

Source: own estimation

Another significant indicator is the membership in the EU or Eurozone, whose positive impact on Czech export is by far not negligible. It confirms the general evidence that the European Union facilitates international trade. In fact, the EU variable could cause the surprising coefficient on regulatory trade barriers. Their relative correlation of almost 0.6 shows a moderate relationship, implying that the EU and regulatory trade barriers indicator vary in a similar way and thus, part of the variation of one can be hidden in a predicted coefficient of another. For confirmation, we have run a simple regression with only four explanatory variables and the predicted coefficient of regulatory trade barriers had an expected positive sign but its economic impact was still negligible.

The last statistically significant variable is the indicator of education. The coefficient is positive but it does not seem to play an important role while deciding where to export. To be precise, one per cent increase of the ratio of tertiary education graduates to total population in a partner's country increases Czech export to this country by 0.014 per cent. The other institutions are insignificant and predicted estimates are low.

Next, we run the Poisson estimation that we assume to be a more correct model. Again, standard errors are clustered in countries to allow for intragroup correlation. The obtained results can be seen in Table 4. As well as OLS, PPMLE shows high R-squared; it reaches even a slightly higher value than in the LSDV model.

Table 4: Total export: time series data structure results – Poisson estimation

Poisson regression Number of observations = 900

R-squared = 0.9342

Log pseudolikelihood = -2.721e+09

(Std. Err. adjusted for 50 clusters in Country)

X_{ijt}	Coefficient	Robust Std. Err.	p-value	[95% confidence interval]	
$log(GDP_{jt})$	**.2308**	**.1204**	**0.055**	**-.0051**	**.4667**
$log(GDP_{it})$	**1.0226**	**.3824**	**0.007**	**.2732**	**1.7721**
$log(Dist_{ij})$	**-.7688**	**.1472**	**0.000**	**-1.0574**	**-.4803**
$log(ERDI_{jt})$	-.3589	.3706	0.333	-1.0853	.3675
Reg_{jt}	**-.0198**	**.0072**	**0.006**	**-.0340**	**-.0057**
$Educ_{jt}$	**.0110**	**.0051**	**0.033**	**.0008**	**.0210**
Bus_{jt}	**.0191**	**.0068**	**0.005**	**.0059**	**.0324**
Gov_{jt}	**.0096**	**.0031**	**0.002**	**.0035**	**.0156**
Inv_{jt}	**.0065**	**.0033**	**0.052**	**-.0001**	**.0130**
EU_{jt}	**.3049**	**.0874**	**0.000**	**.1337**	**.4761**
Rec_{it}	.0007	.0341	0.984	-.0662	.0675

Source: own estimation

Comparing results with the OLS coefficients, the Poisson estimation delivers generally more significant estimates. An apparent change can be observed in values of our crucial variables. More specifically, partner's GDP has now lower influence, i.e. the elasticity amounts only to 0.23. On the contrary, the elasticity of home GDP becomes superior to unity and thus, has a substantial impact on Czech export. Moreover, the distance variable has still a significant influence but the value is lower in absolute terms.

Concerning the institutional variables, they are generally significant at 0.05 significance level but their impact is still negligible. The European Union influences now the response variable by a lower rate. Finally, the two remaining variables – the ERDI and the recession – are still highly insignificant and we keep them just for a comparison with other models.

Based on the analysis, we can conclude that the choice of an estimator does matter and due to the considerable differences, we suspect the predicted OLS coefficients to be biased because of the violation of crucial assumptions for consistency.

5.2 Cross-Sectional Data Structure

In this part, we propose estimation results from view of a short-term perspective. In other words, exporters' decisions are observed in each year, we do not take into account a possibility of adjustment to changes over time. As can be seen in Table 5, R-squared value is naturally lower than in time-series structure because the number of additional time dummies is substantially smaller than country dummies, added previously.

From the significance point of view, the essential change has experienced home GDP. More precisely, it has become highly insignificant, which is understandable, as its value does not change within a group (year). The other two crucial variables – the partner's GDP and the distance – are statistically significant at 0.01 level and their influence on export is considerable. The only two factors that are more significant now are the education and the government spending. However, the real economic impact of both variables is negligible and in case of the government spending, the sign is surprisingly slightly negative.

The predicted parameter of the EU is lower than the relative figure in the time series data structure. Again, the explanation can be the inability of exporters to adjust to changes on a yearly basis. To be precise, Czech export are by 24.6% higher to a member of the European Union or the Eurozone with respect to non-members.[11]

Table 5: Total export: cross-sectional data structure results – LSDV estimation

Linear regression				Number of observations = 900
				R-squared = 0.8209
				Root MSE = .99003
				(Std. Err. adjusted for 18 clusters in Year)

$log(X_{ijt})$	Coefficient	Robust Std. Err.	p-value	[95% confidence interval]	
$\boldsymbol{log(GDP_{jt})}$	**.7314**	**.0266**	**0.000**	**.6752**	**.7877**
$log(GDP_{it})$	.3124	.485	0.528	-.7108	1.3357
$\boldsymbol{log(Dist_{ij})}$	**-.9230**	**.0340**	**0.000**	**-.9947**	**-.8513**
$log(ERDI_{jt})$	-.0677	.1436	0.643	-.3706	.2352
Reg_{jt}	-.0042	.0031	0.193	-.0108	.0024
$\boldsymbol{Educ_{jt}}$	**.0172**	**.0012**	**0.000**	**.0146**	**.0198**
Bus_{jt}	.0052	.0047	0.282	-.0047	.0151
$\boldsymbol{Gov_{jt}}$	**-.0076**	**.0037**	**0.055**	**-.0155**	**.0002**
Inv_{jt}	.0032	.0031	0.311	-.0033	.0098
$\boldsymbol{EU_{jt}}$	**.2239**	**.0483**	**0.000**	**.1220**	**.3259**
Rec_{it}	-.0124	.1739	0.944	-.3793	.3546

Source: own estimation

Regarding the predicted parameters in Table 6, obtained by the Poisson estimation, we can say that there are no such differences, with respect to OLS, as in the time series approach. Again, the R-squared value exceeds the relative figure in OLS model, implying the results to be more convincing.

On the contrary, while comparing to the Poisson model under the time series data structure, the differences are considerable. Focusing on a significance criterion, two variables deserve a particular attention – the home GDP and the education. Education

[11] We use a correction $\%\Delta y = 100 \cdot [\exp(\hat{\beta} \cdot \Delta x) - 1]$ suggested by Wooldridge (2003, pp. 184).

has now not only a negligible economic impact but is also highly statistically insignificant. It implies that the level of education in partners' countries can influence decisions of Czech exporters only in long-term perspective, as they cannot adjust to the partner's ratio of tertiary graduates to total population in a given year.

Table 6: Total export: cross-sectional data structure results – Poisson estimation

Poisson regression Number of observations = 900

R-squared = 0.8659

Log pseudolikelihood = -5.551e+09

(Std. Err. adjusted for 18 clusters in Year)

X_{ijt}	Coefficient	Robust Std. Err.	p-value	[95% confidence interval]	
$log(GDP_{jt})$	**.7781**	**.0126**	**0.000**	**.7533**	**.8028**
$log(GDP_{it})$	.1946	.9265	0.834	-1.6213	2.0104
$log(Dist_{ij})$	**-1.2884**	**.0361**	**0.000**	**-1.3590**	**-1.2177**
$log(ERDI_{jt})$	.2037	.1377	0.139	-.0662	.4736
Reg_{jt}	-.0061	.0064	0.341	-.0187	.0065
$Educ_{jt}$	.0019	.0023	0.406	-.0026	.0064
Bus_{jt}	**.0290**	**.0039**	**0.000**	**.0214**	**.0365**
Gov_{jt}	**.0084**	**.0015**	**0.000**	**.0055**	**.0113**
Inv_{jt}	**.0094**	**.0037**	**0.012**	**.0020**	**.0167**
EU_{jt}	**.1988**	**.0331**	**0.000**	**.1339**	**.2637**
Rec_{it}	.1082	.2763	0.696	-.4334	.6497

Source: own estimation

In a conclusion, we have approved that the way of structuring of a dataset plays an important role, as the predicted parameters and related significance differ substantially. Therefore, we confirm our hypothesis that the different attitude towards the time series and cross-sectional structure has not a negligible impact and should be definitely taken into consideration.

6 MACHINERY EXPORT MODEL

For estimating the export function of machinery, we follow the same procedure as for estimating total export determinants. Thus, we predict coefficients of variables under different structures, applying two estimation methods. Further, we suggest an interpretation and economic rationalization of results.

Accordingly, the first part of the chapter deals with the time series structure of the dataset. The second part studies the cross-sectional structure and finally, the last part introduces a comparative analysis based on predicted coefficients from the estimated total and machinery export functions. We anticipate that the first two parts focus particularly on differences between results based on the two estimators and under the different data structures. Consequently, the final part deals with the differences between the total and machinery export estimations.

6.1 Time Series Data Structure

To begin, we examine the relevance of the model. R-squared attains a high value and thus, the model delivers convincing coefficients that are mostly significant. Similarly, we operate with robust standard errors, clustered in countries, which controls for heteroscedasticity in data. Consequently, we assume our model to be well specified and coefficients to be predicative. Table 7 presents parameters estimated by the OLS regression.

Both GDPs and distance variables are highly significant and their economic impact is considerable. For concreteness, one per cent rise in partner's GDP increases imports of Czech machinery goods by 0.66 per cent. Home GDP and distance have even higher influence, while their elasticity exceeds unity. The signs of predicted parameters coincide with our expectations and their magnitude affirms the essential role of these three determinants.

Table 7: Machinery: time series data structure results – LSDV estimation

Linear regression

Number of observations = 900

R-squared = 0.8964

Root MSE = .79916

(Std. Err. adjusted for 50 clusters in Country)

$log(X_{ijt})$	Coefficient	Robust Std. Err.	p-value	[95% confidence interval]	
$log(GDP_{jt})$	**.6621**	**.0787**	**0.000**	**.5039**	**.8203**
$log(GDP_{it})$	**1.0086**	**.2884**	**0.001**	**.4289**	**1.5882**
$log(Dist_{ij})$	**-1.0190**	**.1601**	**0.000**	**-1.3408**	**-.6971**
$log(ERDI_{jt})$	**-.8219**	**.2815**	**0.005**	**-1.3875**	**-.25632**
Reg_{jt}	**-.0292**	**.0082**	**0.001**	**-.0457**	**-.0128**
$Educ_{jt}$	**.0121**	**.0059**	**0.047**	**.0002**	**.0240**
Bus_{jt}	.0104	.0077	0.182	-.0051	.0258
Gov_{jt}	**.0071**	**.0039**	**0.076**	**-.0008**	**.0149**
Inv_{jt}	.0050	.0062	0.424	-.0075	.0175
EU_{jt}	**.6728**	**.1322**	**0.000**	**.4071**	**.9384**
Rec_{it}	-.0110	.0439	0.804	-.0991	.0772

Source: own estimation

Actually, the model detects other two variables that can cope (in influence) with the primary ones. These are the ERDI and the European Union. The exchange rate deviation index has high explanatory power as its estimate equals to -0.82. It means that one per cent decrease in the index implies an increase by 0.82 of Czech machinery imports.

We have to admit that a part of the high estimate can be the result of a not negligible correlation of ERDI with other variables, e.g. the regulatory trade barriers or the EU. Nevertheless, we still cannot ignore the considerable influence on Czech machinery export. Moreover, the sign of the coefficient conforms to our expectations and implies that a weaker currency (a higher value of ERDI) corresponds to a lower purchasing power in the world market.

Table 8: Machinery: time series data structure results – Poisson estimation

Poisson regression Number of observations = 900

R-squared = 0.9291

Log pseudolikelihood = -1.545e+09

(Std. Err. adjusted for 50 clusters in Country)

X_{ijt}	Coefficient	Robust Std. Err.	p-value	[95% confidence interval]	
$\boldsymbol{log(GDP_{jt})}$	**.2962**	**.1063**	**0.005**	**.0879**	**.5045**
$\boldsymbol{log(GDP_{it})}$	**1.6981**	**.3883**	**0.000**	**.9370**	**2.4591**
$\boldsymbol{log(Dist_{ij})}$	**-.5754**	**.1527**	**0.000**	**-.8748**	**-.2761**
$\boldsymbol{log(ERDI_{jt})}$	**-.8832**	**.3748**	**0.018**	**-1.6178**	**-.1487**
$\boldsymbol{Reg_{jt}}$	**-.0317**	**.0098**	**0.001**	**-.0511**	**-.0124**
$Educ_{jt}$	.0081	.0050	0.105	-.0017	.0179
$\boldsymbol{Bus_{jt}}$	**.0189**	**.0071**	**0.008**	**.0050**	**.0327**
$\boldsymbol{Gov_{jt}}$	**.0091**	**.0025**	**0.000**	**.0042**	**.0140**
$\boldsymbol{Inv_{jt}}$	**.0061**	**.0035**	**0.080**	**-.0007**	**.0128**
$\boldsymbol{EU_{jt}}$	**.4631**	**.1135**	**0.000**	**.2406**	**.6856**
Rec_{it}	-.0370	.0396	0.350	-.1147	.0406

Source: own estimation

The other variables, even if statistically significant, have negligible impact on the Czech machinery export. Therefore, we can conclude that from a long-term perspective, the destination of the Czech machinery export is mostly identified based on few determinants. Namely, the influencing factors are partner's and home GDP, distance, ERDI and EU membership.

Let us now move on the Poisson pseudo maximum likelihood estimation, whose output is transferred into Table 8. The R-squared value is superior to the LSDV one and the change in predicted estimates is not negligible. Above all, we observe a substantial drop in the impact of partner's GDP. It implies a positive bias in the LSDV model and reveals a lower real impact of the partner's GDP on the Czech machinery export.

On the other hand, home GDP seems to play a more important role for determining an amount of export. It signifies that in good years, the productivity is high and the proportion of export increases accordingly.[12] This theory is particularly valid for the most profitable sector in the Czech Republic, i.e. machinery.

The estimate on distance was probably negatively biased in the LSDV model, as the respective figure in the Poisson model is less negative. Similarly, the EU estimate shows a different value, attributing a lower influence of the EU membership. On the contrary, the PPMLE confirms the substantial negative influence of the ERDI, as the elasticity exceeds 0.8 in absolute value and it is still significant at 0.05 level.

As well as the OLS, the Poisson estimator predicts negligible influence of other variables. Therefore, they belong to our model because they are mostly statistically significant and can have some aggregate effect but we do not consider their influence individually.

6.2 Cross-Sectional Data Structure

In this part, we propose an estimation of machinery export under the cross-sectional structure. Data are arranged into 18 groups (years) with 50 observations (countries) for each year. This approach allows us to reveal which country characteristics are important to determine the level of Czech machinery export in a given year. Table 9 presents the output of the OLS regression.

The first point to retain is the logical insignificance of home GDP (it does not vary within a group). Similarly, the exchange rate deviation index is no more a predicative variable but still reveals a negative impact on machinery export. Another key difference with regard to the time series model is the drop in influence of the distance and the EU variable.

[12] Obviously, this causality may act backwards as well, as the increasing export can boost economic performance.

Table 9: Machinery: cross-sectional data structure results – LSDV estimation

Linear regression

Number of observations = 900
R-squared = 0.7974
Root MSE = 1.1016

(Std. Err. adjusted for 18 clusters in Year)

$log(X_{ijt})$	Coefficient	Robust Std. Err.	p-value	[95% confidence interval]	
$\boldsymbol{log(GDP_{jt})}$	**.7824**	**.0317**	**0.000**	**.7155**	**.8493**
$log(GDP_{it})$	.3662	.6548	0.583	-1.0154	1.748
$\boldsymbol{log(Dist_{ij})}$	**-.8271**	**.0459**	**0.000**	**-.9240**	**-.7303**
$log(ERDI_{jt})$	-.2332	.1692	0.186	-.5901	.1237
Reg_{jt}	-.0058	.0039	0.156	-.0141	.0025
$\boldsymbol{Educ_{jt}}$	**.0168**	**.0017**	**0.000**	**.0133**	**.0203**
$\boldsymbol{Bus_{jt}}$	**.0096**	**.0046**	**0.055**	**-.0002**	**.0194**
Gov_{jt}	-.0063	.0039	0.123	-.0145	.0019
Inv_{jt}	.0006	.0034	0.864	-.0065	.0077
$\boldsymbol{EU_{jt}}$	**.3530**	**.0500**	**0.000**	**.2475**	**.4585**
Rec_{it}	-.0175	.2267	0.939	-.4958	.4608

Source: own estimation

Next, let us examine the Poisson regression. Table 10 reveals again a higher level of R-squared with regard to the LSDV model, suggesting that data fits better the Poisson model. Besides, predicted coefficients do not differ a lot. We underline the coefficient on distance because the lower value in the Poisson model implies a possible positive bias in the OLS estimation.

We consider the change in the estimate on distance an important feature because, compared to time series, it moves in another direction for the Poisson estimation then in case of the LSDV. To explain, in case of the LSDV, the coefficient on distance amounts to -1.02 under the time series structure, whereas the relative figure for cross-sections is only -0.827. On the other hand, the Poisson estimation delivers a coefficient of -.575 for the time series approach and -1.079 for the cross-sectional one. Therefore, we suspect the OLS coefficients to be biased, but surprisingly in an opposite direction under the two structures.

Table 10: Machinery: cross-sectional data structure results – Poisson estimation

Poisson regression Number of observations = 900

R-squared = 0.8683

Log pseudolikelihood = -2.870e+09

(Std. Err. adjusted for 18 clusters in Year)

X_{ijt}	Coefficient	Robust Std. Err.	p-value	[95% confidence interval]	
$\boldsymbol{log(GDP_{jt})}$	**.7974**	**.0125**	**0.000**	**.7729**	**.8220**
$log(GDP_{it})$	.7565	1.0497	0.471	-1.3009	2.8138
$\boldsymbol{log(Dist_{ij})}$	**-1.0787**	**.0360**	**0.000**	**-1.1494**	**-1.0081**
$log(ERDI_{jt})$	-.1943	.1477	0.188	-.4837	.0951
Reg_{jt}	-.0114	.0073	0.118	-.0258	.0029
$Educ_{jt}$	-.0001	.0020	0.967	-.0040	.0038
$\boldsymbol{Bus_{jt}}$	**.0307**	**.0043**	**0.000**	**.0224**	**.0391**
$\boldsymbol{Gov_{jt}}$	**.0090**	**.0018**	**0.000**	**.0054**	**.0125**
$\boldsymbol{Inv_{jt}}$	**.0082**	**.0043**	**0.056**	**-.0002**	**.0167**
$\boldsymbol{EU_{jt}}$	**.3078**	**.0373**	**0.000**	**.2348**	**.3809**
Rec_{it}	.1019	.3156	0.747	-.5166	.7203

Source: own estimation

As we assume the Poisson estimates to be correct, we should support them with a meaningful explanation for such a development. We know that distance does not change over time and thus, it would be omitted from the time series structured model using standard fixed effects method. We have bypassed this mandatory exclusion by implementing the LSDV approach. However, one needs to remember that models structured as time series ascribe lower importance to variables constant over time. Therefore, either models deliver the time-invariant variables insignificant or their explanatory power is reduced. That holds for the cross-sectional structure with variables constant across entities as well (e.g. home GDP).

We use another criterion for a comparison of time series and cross-sectional structures – a confidence interval. As majority of researchers do not distinguish between these two structures, we would suppose that the respective confidence intervals for a given variable would at least partly coincide. That is often not true in our case. For instance, we are 95 per cent sure that the estimate on partner's GDP predicted under the time series structure does not equal to that under the cross-sectional structure. It is probably even more than 95 per cent because there is a considerable gap between the upper limit of the time series confidence interval (0.505) and the lower limit of the cross-sectional one (0.773).

There are some coefficients, whose confidence intervals overlap, such as the ERDI. It can eventually happen that one confidence interval constitutes a subset of another, e.g. the home GDP. In these cases, there is a chance of equality of coefficients. Nevertheless, none of the confidence intervals is equal to its counterpart from the other structure and thus, we underline again the importance of distinguishing between the structures.

6.3 Comparative Analysis

This part proposes a summation of the obtained results and particularly, the comparison of estimated functions from view of overall and machinery export. We use the Poisson estimates predicted under different structures for that purpose. Table 11 summarizes results for time series structured data and Table 12 points out coefficients with their significance level, acquired from cross-sectional structured models.[13]

Comparing the two differently aggregated models, the R-squared reaches almost the same values in both structures of datasets. It corresponds to the fact that Czech export consist mostly of machinery products and thus, the determinants in both models should behave in a similar way. Our analysis confirms that hypothesis, as the predicted parameters and their significance levels more or less coincide.

Let us begin with the time series structure. As one can see, there are no considerable differences between the models. A variable that behaves in a different way is the home GDP. In both cases, its elasticity exceeds unity but evidently, the level of home GDP

[13] Estimates are rounded to three decimal places. The last column reveals a similarity of predicted coefficients and their significance between the models.

has a larger impact on machinery export. It is a common knowledge that mostly developed economies focus on machinery production. Therefore, it is not a surprise that the Czech economic performance represents a crucial determinant for the Czech machinery production and consequently for machinery export.

Table 11: Comparison of total and machinery export functions – time series structure

X_{ijt}	Total export	Machinery export	Similarity
$log(GDP_{jt})$	.231*	.296***	✗
$log(GDP_{it})$	1.023***	1.698***	✓
$log(Dist_{ij})$	-.769***	-.575***	✓
$log(ERDI_{jt})$	-.359	-.883**	✗
Reg_{jt}	-.020***	-.032***	✓
$Educ_{jt}$	.011**	.008	✗
Bus_{jt}	.020***	.019***	✓
Gov_{jt}	.010***	.009***	✓
Inv_{jt}	.006**	.006*	✓
EU_{jt}	.305***	.463***	✓
Rec_{it}	.001	-.037	✓
R-squared	0.9342	0.9291	

Source: own estimation
*Note: *** p<0.01, ** p<0.05, * p<0.1*

Another substantial change can be observed in the ERDI estimate and particularly in its significance. For total export, the exchange rate deviation index does not seem to play any important role. On the contrary, for the export function of machinery, the index becomes not only statistically significant, but its economic impact is by far not negligible. It implies that the variation of countries' purchasing power plays a bigger role in the international machinery trade market.

Similarly, the EU variable attains a higher value in the machinery model. In other words, Czech exporters tend to benefit from the European trade facilitation more on the machinery level than on the overall one. On the other hand, the distance influences machinery export by a lower rate. It suggests that machinery products are more specific and thus, trade partners cannot be chosen simply on a distance basis.

Moving on the cross-sectional structure, we can see that both models follow the same tendencies with respect to the time series structure. For instance, lower R-squared and less significant variables. Similarly, the differences in estimates between the models are comparable to the time series structure. For illustration, distance has again a lower impact on machinery export and the EU plays a more important role in the machinery export function. Therefore, we can conclude that both structures agree on the predicted differences between the total and machinery export function, even though the actual estimates do not reach the same values.

Table 12: Comparison of total and machinery export functions – cross-sectional structure

X_{ijt}	Total export	Machinery export	Similarity
$log(GDP_{jt})$	.778***	.797***	✓
$log(GDP_{it})$	.195	.756	✓
$log(Dist_{ij})$	-1.29***	-1.079***	✓
$log(ERDI_{jt})$	.204	-.194	✓
Reg_{jt}	-.006	-.011	✓
$Educ_{jt}$	.002	-.000	✓
Bus_{jt}	.029***	.031***	✓
Gov_{jt}	.008***	.009***	✓
Inv_{jt}	.009**	.008*	✓
EU_{jt}	.199***	.308***	✓
Rec_{it}	.108	.102	✓
R-squared	0.8659	0.8683	

Source: own estimation
*Note: *** p<0.01, ** p<0.05, * p<0.1*

7 AUTOMOBILE EXPORT MODEL

Finally, we run regressions for automobiles, which reveals the most surprising evidence. The first two parts of this chapter introduce results obtained with the use of two estimators and under two data specifications. In the final part, we emphasize essential differences among models covering three levels of aggregation – overall, machinery and automobiles.

7.1 Time Series Data Structure

We first examine the OLS estimation, whose output is presented in Table 13. We can see that R-squared is relatively high but explanatory variables are generally insignificant. It implies a limited predicative power of the variables. Moreover, the model omits zero flows and the potency of the estimation is thus not fulfilled (only 854 out of 900 possible observations). The model is then not convincing but definitely worth investigating.

The most surprising outcome represents partner’s GDP. It contradicts to our expectations. Partner’s GDP is expected to be an essential determinant, positively related to export. In our case, not only the variable is insignificant, but also its predicted impact is negative. Therefore, we suppose Czech export of automobiles to be influenced by other unobservable factors that can impair traditional trade theories.

On the other hand, home GDP seems to take a prominent position among determinants of automobile export. Home GDP is statistically significant at 0.05 level and its elasticity notably exceeds unity. To be precise, one per cent increase in home GDP boosts export of automobiles by 1.59 per cent.

Table 13: Automobiles: time series data structure results – LSDV estimation

Linear regression — Number of observations = 854; R-squared = 0.7188; Root MSE = 1.3582

(Std. Err. adjusted for 50 clusters in Country)

$\log(X_{ijt})$	Coefficient	Robust Std. Err.	p-value	[95% confidence interval]	
$log(GDP_{jt})$	-.1574	.2607	0.549	-.6812	.3664
$\boldsymbol{log(GDP_{it})}$	**1.5922**	**.6585**	**0.019**	**.2689**	**2.9156**
$\boldsymbol{log(Dist_{ij})}$	**-.9576**	**.2067**	**0.000**	**-1.3731**	**-.5422**
$log(ERDI_{jt})$	-.0374	.4594	0.935	-.9607	.8858
Reg_{jt}	-.0232	.0231	0.319	-.0697	.0232
$Educ_{jt}$	.0065	.0118	0.586	-.0172	.0301
Bus_{jt}	.0143	.0132	0.284	-.0122	.0408
Gov_{jt}	.0078	.0054	0.151	-.0030	.0186
Inv_{jt}	-.0080	.0057	0.161	-.0194	.0033
$\boldsymbol{EU_{jt}}$	**1.5452**	**.3896**	**0.000**	**.7622**	**2.3281**
Rec_{it}	-.0239	.1072	0.824	-.2393	.1915

Source: own estimation

A similar impact on export has the European variable. Its coefficient reaches almost the same value as that on home GDP and it is significant even at 0.01 level. Distance significantly influences automobile export as well, but naturally in the inverse direction. Other variables reveal a low significance and neither their estimates seem to be influential.

Let us now examine the Poisson estimation (Table 14). The model seems to be a more appropriate one as the R-squared reaches a higher value, more coefficients are significant and particularly, it counts with all possible observations. The two additional significant variables, with respect to the LSDV model, are the government spending and the investment freedom. Therefore, also institutional variables play some role for Czech exporters of automobiles.

Concerning the signs of coefficients, they coincide to the OLS ones. However, their values and significance differ. For instance, the estimate on home GDP is even higher that the relative predicted figure in the OLS estimation, suggesting a negative bias in the OLS estimate. According to the Poisson results, an increase in home GDP causes almost a double increase in Czech export of automobiles.

The distance and the EU variables worth investigating as well. Both have a lower impact on automobile export comparing to the OLS estimates. It implies inconsistency of the OLS coefficients, confirming an inconvenience of the log-linearization of the gravity equation.

Table 14: Automobiles: time series data structure results – Poisson estimation

Poisson regression — Number of observations = 900

R-squared = 0.8209

Log pseudolikelihood = -5.616e+08

(Std. Err. adjusted for 50 clusters in Country)

X_{ijt}	Coefficient	Robust Std. Err.	p-value	[95% confidence interval]	
$log(GDP_{jt})$	-.0382	.1506	0.800	-.3335	.2570
$\boldsymbol{log(GDP_{it})}$	**1.9306**	**.5450**	**0.000**	**.8624**	**2.9988**
$\boldsymbol{log(Dist_{ij})}$	**-.3409**	**.1739**	**0.050**	**-.6818**	**.0000**
$log(ERDI_{jt})$	-.5003	.4679	0.285	-1.4173	.4167
Reg_{jt}	-.0121	.0226	0.594	-.0564	.0323
$Educ_{jt}$	.0132	.0096	0.168	-.0055	.0320
$\boldsymbol{Bus_{jt}}$	**.0229**	**.0077**	**0.003**	**.0078**	**.0380**
$\boldsymbol{Gov_{jt}}$	**.0123**	**.0042**	**0.004**	**.0040**	**.0206**
Inv_{jt}	-.0020	.0045	0.660	-.0109	.0069
$\boldsymbol{EU_{jt}}$	**.6753**	**.2594**	**0.009**	**.1668**	**1.1837**
Rec_{it}	.0904	.0819	0.269	-.0700	.2509

Source: own estimation

7.2 Cross-Sectional Data Structure

In this part, we study predicted models with data structured as cross-sections. After examining the OLS output in Table 15, a key point to retain is a much lower value of R-squared. It implies that in a given year there are other unobservable factors, influencing decision making of Czech automobile exporters.

There are three main exporting car companies in the Czech Republic with one occupying a prominent position – Škoda Auto. The Škoda takes part in the German concern Volkswagen Group, which has a substantial influence on Czech carmakers. German investors can have a different intention with the development of Škoda's export in order to meet requirements of other subsidiary companies or to coordinate their actions. Therefore, the Škoda reacts to the situation in its mother company, which is obviously not predictable.

As well as under the time series structure, the most influencing factor represents home GDP. In the cross-sectional case, its influence seems to be even more convincing regarding the other variables. Partner's GDP stays insignificant but the ERDI variable becomes an essential determinant of Czech automobile export. Aside from a high significant level, the estimate of the ERDI is considerable. One per cent increase in a partner's index implies 0.78 per cent decrease in export of automobile to that country.

Table 15: Automobiles: cross-sectional data structure results – LSDV estimation

Linear regression — Number of observations = 854; R-squared = 0.4722; Root MSE = 1.8325

(Std. Err. adjusted for 18 clusters in Year)

$log(X_{ijt})$	Coefficient	Robust Std. Err.	p-value	[95% confidence interval]	
$log(GDP_{jt})$	.0392	.0246	0.130	-.0128	.0911
$\boldsymbol{log(GDP_{it})}$	**1.7674**	**.5311**	**0.004**	**.6469**	**2.8880**
$\boldsymbol{log(Dist_{ij})}$	**-.6666**	**.0830**	**0.000**	**-.8417**	**-.4914**
$\boldsymbol{log(ERDI_{jt})}$	**-.7780**	**.2156**	**0.002**	**-1.2329**	**-.3231**
Reg_{jt}	-.0043	.0058	0.470	-.0165	.0080
$Educ_{jt}$	.0069	.0041	0.105	-.0016	.0155
Bus_{jt}	-.0079	.0071	0.276	-.0228	.0069
$\boldsymbol{Gov_{jt}}$	**-.0120**	**.0057**	**0.051**	**-.0240**	**.0001**
$\boldsymbol{Inv_{jt}}$	**.0114**	**.0034**	**0.004**	**.0042**	**.0186**
$\boldsymbol{EU_{jt}}$	**.8492**	**.0769**	**0.000**	**.6867**	**1.0114**
Rec_{it}	-.0606	.1657	0.719	-.4103	.2891

Source: own estimation

The impact of the EU membership is mitigated under the cross-sectional structure, as its elasticity does not reach unity. The same holds for the distance estimate. Other coefficients do not reveal any important differences with respect to the time series structure. Therefore, we can conclude that their impact is negligible.

Apart from the relatively low R-squared, Table 16 shows us generally low p-values of the estimated Poisson coefficients. It suggests that in the Poisson estimation, the individual variables play some role. However, regarding the low estimates of institutional indices, we do not consider them individually.

On the contrary, the effect of home GDP is huge. To be specific, if home GDP would double, the export of automobiles would almost triple, holding other factors fixed. The automobile sector is definitely seriously dependent on the actual Czech economic performance. The causality can be obviously reverse, when automobile export represents a great impulse to the economy.

Table 16: Automobiles: cross-sectional data structure results – Poisson estimation

Poisson regression — Number of observations = 900; R-squared = 0.5500

Log pseudolikelihood = -1.411e+09

(Std. Err. adjusted for 18 clusters in Year)

X_{ijt}	Coefficient	Robust Std. Err.	p-value	[95% confidence interval]	
$\log(GDP_{jt})$	**-.1806**	**.0210**	**0.000**	**-.2219**	**-.1394**
$\log(GDP_{it})$	**2.8311**	**.5882**	**0.000**	**1.6782**	**3.9839**
$\log(Dist_{ij})$	**-.5114**	**.0207**	**0.000**	**-.5520**	**-.4708**
$\log(ERDI_{jt})$	**-1.1286**	**.1843**	**0.000**	**-1.4898**	**-.7674**
Reg_{jt}	**-.0174**	**.0046**	**0.000**	**-.0264**	**-.0083**
$Educ_{jt}$	.0007	.0014	0.604	-.0020	.0035
Bus_{jt}	**.0206**	**.0057**	**0.000**	**.0094**	**.0317**
Gov_{jt}	**-.0036**	**.0015**	**0.016**	**-.0065**	**-.0007**
Inv_{jt}	.0048	.0029	0.101	-.0009	.0106
EU_{jt}	**.5414**	**.0737**	**0.000**	**.3970**	**.6858**
Rec_{it}	.0903	.1560	0.563	-.2155	.3961

Source: own estimation

The exchange rate deviation index seems to be a crucial determinant of Czech exporters. The negative elasticity even exceeds unity, which attributes to a considerable influential power of the variation in foreign currency. Moreover, we have to emphasize that the variable is much more significant while deciding in a given year (cross-sections) than in the long-term perspective.

The most striking evidence is the significant negative estimate of the partner's GDP. It confirms our hypothesis that the Czech automobile export is controlled from abroad and is subject to numerous unobservable factors. Therefore, it can happen that crucial determinants reveal an unexpected evidence.

The distance variable still coincides to our expectations and reveals a negative impact. Its impact is even higher than under the time series structure. Correspondingly, the facilitation of the international trade by the European Union boosts the automobile trade, as the coefficient on the EU variable is positive and significant.

7.3 Comparative Analysis

This section presents a summation of our existing findings by comparing models covering export data at three different levels of aggregation. We distinguish once more between the two structures of our datasets. Table 17 presents the predicted results under the time series structure and Table 18 covers the cross-sectional results.

R-squared value serves as a goodness of fit measure that detects the potential inaccuracy of a model. Correspondingly, only the automobile export model, particularly with cross-sectional data, is prone to deliver less convincing results. As mentioned previously, there are many unobservable determinants of Czech automobile export, which implies a lower explanatory power of our model.

Let us first investigate the three models under the time series data structure. A common feature is the negligible impact of institutional indices with business freedom to be the most influential one. The membership in the European Union or Eurozone has a considerable positive impact for all models. However, the crucial determinant in our analysis is home GDP. It designs the strong relationship between the economic prosperity and the amount of export, no matter in which direction the causality holds. A compelling evidence is a rising tendency in the impact of the EU and home GDP with the level of disaggregation.

On the other hand, we cannot reveal any common feature in the partner's GDP. It proves to be the most significant for machinery export, less significant for total export and for export of automobiles; not only it is insignificant but reveals a negative value. Considering the distance variable, Poisson estimation predicts the higher impact, the higher level of aggregation is concerned. In other words, distance influences the most considerably total Czech export.

From the long-term perspective, purchasing power of export partners shows to be a relevant determinant of Czech export only in the machinery sector. Nevertheless, all three models deliver the ERDI estimates with an expected negative sign and in the machinery case, the ERDI's impact is substantial.

Table 17: Comparison of total, machinery and automobile export functions - time series structure

X_{ijt}	Total export	Machinery export	Automobile export	Similarity
$log(GDP_{jt})$	.231*	.296***	-.038	✗
$log(GDP_{it})$	1.023***	1.698***	1.931***	✓
$log(Dist_{ij})$	-.769***	-.575***	-.341*	✓
$log(ERDI_{jt})$	-.359	-.883**	-.500	✗
Reg_{jt}	-.020***	-.032***	-.012	✗
$Educ_{jt}$	.011**	.008	.013	✓
Bus_{jt}	.020***	.019***	.023***	✓
Gov_{jt}	.010***	.009***	.012***	✓
Inv_{jt}	.006**	.006*	-.002	✓
EU_{jt}	.305***	.463***	.675***	✓
Rec_{it}	.001	-.037	.090	✓
R-squared	0.9342	0.9291	0.8209	

Source: own estimation
*Note: *** p<0.01, ** p<0.05, * p<0.1*

The predicted coefficients under the cross-sectional structure (Table 18) in the automobile model seem to be more significant but we suspect the overall model of inaccuracy due to the lower R-squared. Consequently, the estimates differ considerably.

First thing to retain is the variation of the estimate on partner's GDP. The values in the total and machinery export's estimation coincide and they are statistically significant. On the other hand, the respective figure for automobile export implies a negative impact of partner's GDP. Another surprising point is the substantial positive effect of home GDP on automobile export, whereas the other two models reveal the variable as insignificant.

The distance variable follows the same tendency as in time series, i.e. the impact of geographical distance is lower on export at more disaggregated level. On the other hand, the ERDI does not seem to play any important role for total and machinery export but it is a crucial determinant of Czech automobile export in a given year. Finally, the positive influence of the European Union escalates with the disaggregation level.

Table 18: Comparison of total, machinery and automobile export functions - cross-sectional structure

X_{ijt}	Total export	Machinery export	Automobile export	Similarity
$log(GDP_{jt})$	.778***	.797***	-.181***	✗
$log(GDP_{it})$	.195	.756	2.831***	✗
$log(Dist_{ij})$	-1.29***	-1.079***	-.511***	✓
$log(ERDI_{jt})$	.204	-.194	-1.129***	✗
Reg_{jt}	-.006	-.011	-.017***	✗
$Educ_{jt}$	.002	-.000	.001	✓
Bus_{jt}	.029***	.031***	.021***	✓
Gov_{jt}	.008***	.009***	-.004**	✓
Inv_{jt}	.009**	.008*	.005	✓
EU_{jt}	.199***	.308***	.541***	✓
Rec_{it}	.108	.102	.090	✓
R-squared	0.8659	0.8683	0.5500	

Source: own estimation
*Note: *** p<0.01, ** p<0.05, * p<0.1*

Regarding the other less influential variables, the business freedom confirms its position as the most significant and stable variable across all models. The recession dummy does not seem to play any meaningful role for amount of export. However, most of its estimates are slightly positive, illustrating the focus of the Czech Republic on production of low-cost automobiles that are demanding articles also during a crisis.

To summarize, the predicted total and machinery export functions more or less coincide. It is due to the high proportion of the machinery goods in total export. However, results from the estimation of automobile export differ considerably but it may not be true that export of automobiles follow different tendencies. The Czech main car exporters are controlled from abroad and their development can be thus driven by other factors that are not observable for us. Consequently, the results based on the estimation of the automobile export function may not be convincing.

8 CONCLUSION

The main interest of the thesis lies in examining determinants of Czech export. The Czech Republic is an export-oriented country and the export sector significantly contributes to the economic performance. By employing illustrative figures, we approved the machinery to be the fundamental industry and automobiles taking a supreme position among machinery export products.

The research thus focuses on estimating export functions based on total, machinery and automobiles export flows data. For that purpose, we used a widely implemented tool for estimating international trade – an augmented gravity model. The models were then predicted by the Ordinary least squares (LSDV method) and Poisson pseudo maximum likelihood estimators. After examining both approaches, we are in accordance with the existing literature in preferring the Poisson method due to the possible inconsistency of the OLS estimation.

We also approved the importance of distinguishing between data structures. The estimations were performed under the time series and the cross-sectional structure and the obtained results differed, particularly for the essential determinants. It implies that the development of Czech export does not vary in the same way from the short-term and long-term perspective.

Our main finding suggests the similarity of determinants for Czech total and machinery export but a considerable deviance of export of automobiles. The predicted coefficients in total and machinery models are in accordance with our expectations, whereas the export of automobiles shows to be influenced by some unobservable factors that impair standard trade theories. The Czech automobile sector is based on three multinational oligopolies that behave according to their endogenous trade strategy. Therefore, the predicted determinants of Czech export of automobiles may be biased as they do not take into account factors influencing development of related foreign companies.

We have concluded for main determinants of Czech export to be GDP of an export partner, home GDP, distance and membership in the European Union or the Eurozone. Home GDP and the EU variable are more influential for more disaggregated products. On the contrary, the distance plays a more important role towards a higher level of aggregation.

Our study enlarged the gravity literature by an innovative way of employing the augmented gravity model. We also contributed to the analytical research of the Czech economy and revealed some surprising evidence about development of Czech export. For further research, we suggest concentrating on an enlargement of the country set and bypassing a potential sample selection issue. Besides, it would be meaningful to proceed further towards a higher level of disaggregation to reveal possible similar tendencies.

BIBLIOGRAPHY

Anderson, J. E. (1979). A Theoretical Foundation for the Gravity Equation. *American Economic Review 69(1)*, 106-116.

Anderson, J., & van Wincoop, E. (2003). Gravity with Gravitas: A Solution to the Border Puzzle. *NBER Working Paper 8079.*

Babecká Kucharčuková, O., Babecký, J., & Raiser, M. (2010). A Gravity Approach to Modelling International Trade in South-Estern Europe and the Commonwealth of Independent States: The Role of Geography, Policy and Institutions. *CNB Working Paper 4*.

Baier, S. L., & Bergstrand, J. H. (2009). Bonus vetus OLS: A simple method for approximating international trade-cost effects using the gravity equation. *Journal of International Economics 77(1)*, 77-85.

Baldwin, R., & Taglioni, D. (2006). Gravity for Dummies and Dummies for Gravity. *NBER Working Paper 12516.*

Bergeijk, P. A., & Brakman, S. (2010). *The Gravity Model in International Trade: Advances and Applications.* Cambridge University Press.

Bobková, B. (2012). Gravity model estimation using panel data - is logarithmic transformation advisable? *IES FSV, Charles University, Prague*.

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

Caporale, G. M., Rault, C., Sova, A., & Sova, R. (2009). On the bilateral trade effects of free trade agreements between the EU-15 and the CEEC-4 countries. *Review of World Economies 145(2)*, 189-206.

Carrère, C. (2006). Revisiting the effects of regional trade agreements on trade flows with proper specification of the gravity model. *European Economic Review 50*, 223-247.

Deardorff, A. (1998). Determinants of Bilateral Trade: Does Gravit Work in Neoclassical World. In J. A. Frankel, *The Regionalization of the World Economy*, 7-32. University of Chicago Press.

Eaton, J., & Kortum, S. (2002). Technology, Geography, and Trade. *Econometrica 70(5)*, 1741-1779.

Egger, P. (2002). An econometric view on the estimation of gravity models and the calculation of trade potentials. *The World Economy 25(2)*, 297-312.

Egger, P., & Pfaffermayr, M. (2001). Distance, Trade, and FDI: A Hausman-Taylor SUR Approach. *WIFO Working Paper 164*.

Frankel, J., & Wei, S. (1993). Trade Blocs and Currency Blocs. *NBER Working Paper 4335*.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica 47*, 153-161.

Helpman, E., & Krugman, P. R. (1985). *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy.* Cambridge: The Massachusetts Institute of Technology.

Helpman, E., Melitz, M. J., & Rubinstein, Y. (2008). Estimating Trade Flows: Trading Partners and Trading Volumes. *Quarterly Journal of Economics 123*, 441-487.

Hyžíková, M. (2012). Ekonometrický test vlivu přijetí eura na české a německé exporty. Analýza nákladů a výnosů plynoucích z členství v eurozóně. *Prague: Charles University, Institute of Economic Studies*.

Janda, K., Michalíková, E., & Skuhrovec, J. (2013). Credit Support for Export: Robust Evidence from the Czech Republic. *World Economy 36(12)*, 1588-1610.

Linnemann, H. (1966). An Econometric Study of International Trade Flows.

Martin, W., & Pham, C. (2008). Esimating the Gravity Model When Zero Trade Flows Are Frequent. *Economic series: MPRA Paper 9453*.

Martinez-Zarzoso, I. (2013). The Log of Gravity Revisited. *Applied Economics 45(3)*, 311-327.

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

Miroudot, S., Lanz, R., & Ragoussis, A. (2009). Trade in intermediate goods and services. *OECD Trade Policy Working paper 93*.

Newton, I. (2002). Book 1, Section 1. In I. Newton, & D. R. Wilkins (Ed.), *The Mathematical Principles of Natural Philosophy* (Motte, Trans.).

Novosák, J., Damborský, M., & Linhartová, I. (2010). Internacionalizace automobilového průmyslu v České republice. *Czech regional studies, 1*.

Rose, A. K. (2000, April). One Money, One Market: The Effect of Common Currencies on Trade. *Economic Policy*, 7-33.

Rose, A., & Spiegel, M. (2010). International environmental arrangements and international commerce. In P. A. van Bergeijk, & B. Steven, *The Gravity Model International Trade*, 255-277. Cambridge: Cambridge University Press.

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

Santos Silva, J., & Tenreyro, S. (2011). Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator. *Economics Letters 112(2)*, 220-222.

Shepherd, B. (2013). *The Gravity Model of International Trade: A User Guide.* Bangkok: United Nations .

Tinbergen, J. (1962). Shaping the World Economy; Suggestions for an International Economic Policy.

Westerlund, J., & Wilhelmsson, F. (2009). Estimating the gravity model without gravity using panel data. *Applied Economics 43(6)*, 641-649.

Wooldridge, J. M. (2003). *Introductory Econometrics: A Modern Approach, 2e.* Michigan State University.

Xiong, B., & Chen, S. (2012). Estimating Gravity Equationtions Models in the Presence of Sample Selection and Heteroscadasticity. *Presented at the AAEA Annual Meeting*.

Yane, H. (2013). Prospects for Trade in Intermediates and Trade in Services: What Does the Gravity Model of Bilateral Trade Tell Us? *OSIPP Discussion Paper 2.*

DATA SOURCES

CEPII – Centre d'Etudes Prospectives et d'Informations Internationales

http://www.cepii.fr/anglaisgraph/bdd/distances.htm

CZSO - Czech Statistical Office (2012)

http://apl.czso.cz/pll/stazo/STAZO.STAZO?jazyk=EN&prvni=N

http://www.czso.cz/csu/redakce.nsf/i/recese_deprese

Economic Freedom Dataset (2013), published in *Economic Freedom of the World: 2013 Annual Report. Frazer Institute.* Gwartney, J; Lawson, R.; & Hall, J.

http://www.freetheworld.com/datasets_efw.html

Eurostat

http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/themes

Heritage Foundation

http://www.heritage.org/Index/

IMF – International Monetary Fund (2013)

http://www.imf.org/external/pubs/ft/weo/2013/02/weodata/index.aspx

OECD – Organisation for Economic Co-operation and Development

http://stats.oecd.org/

OICA – Organisation Internationale des Constructeurs d'Automobiles

http://www.oica.net/category/production-statistics/

World Bank

http://databank.worldbank.org/data/views/variableselection/selectvariables.aspx?source=worldwide-governance-indicators

http://data.worldbank.org/indicator

APPENDIX: TABLES AND FIGURES

Table 19: Comparison of GDP per capita in PPS

(values with respect to EU28 = 100)

Country	1995	2000	2005	2008	2009	2010	2011	2012
Belgium	128	126	120	116	118	119	118	119
Bulgaria	32	28	37	43	44	44	46	47
Czech Republic	**76**	**71**	**79**	**81**	**82**	**80**	**80**	**79**
Denmark	131	131	123	124	123	128	125	125
Estonia	36	45	61	69	64	63	68	70
Finland	107	117	114	119	114	113	114	113
France	115	115	110	107	109	108	108	108
Croatia	46	50	57	63	62	58	61	61
Ireland	103	131	144	131	128	128	130	130
Italy	121	117	105	104	104	101	100	98
Cyprus	87	87	93	99	100	96	94	92
Lithuania	35	40	54	64	58	61	66	70
Latvia	31	36	49	58	54	54	58	62
Luxembourg	221	244	253	263	252	262	265	262
Hungary	51	54	63	64	65	65	65	65
Malta	89	87	80	81	84	87	87	86
Germany	128	117	116	116	115	118	121	122
Netherlands	123	134	130	134	132	131	130	128
Poland	43	48	51	56	60	62	64	66
Portugal	77	81	79	78	80	80	78	75
Austria	134	132	125	124	125	127	129	130
Romania	33	26	35	47	47	47	47	49
Greece	75	84	90	92	94	87	79	75
Slovakia	47	50	60	72	73	73	73	75
Slovenia	74	80	87	91	86	83	83	82
United Kingdom	115	120	124	114	112	112	110	110
Spain	91	97	102	103	103	99	96	95
Sweden	125	127	121	123	120	123	126	128

Source: Eurostat, own estimation

Table 20: Czech export to the European Union (27 members)

Year	**Export** (mil. CZK)	Share of total export
1995	470 878	83,2%
1996	494 433	82,2%
1997	587 067	82,8%
1998	709 624	85,1%
1999	796 312	87,6%
2000	965 382	86,1%
2001	1 100 085	86,7%
2002	1 076 670	85,8%
2003	1 196 871	87,3%
2004	1 500 796	87,1%
2005	1 597 541	85,5%
2006	1 837 052	85,7%
2007	2 113 649	85,3%
2008	2 107 915	85,2%
2009	1 811 957	84,7%
2010	2 126 339	84,0%
2011	2 388 968	83,0%
2012	2 486 806	80,9%

Source: CZSO, own estimation

Table 21: Structure of Czech export based on 1-digit SITC in 2012

Commodity	Code	Value in mil. CZK	Percentage share
Food and live animals	0	108 057	3.5%
Beverages and tobacco	1	19 873	0.6%
Crude materials, inedible, except fuels	2	86 439	2.8%
Mineral fuels, lubricants and related materials	3	118 705	3.8%
Animal and vegetable oils, fats and waxes	4	9 177	0.3%
Chemicals and related products, n.e.s.	5	189 474	6.2%
Manufactured goods classified chiefly by material	6	532 504	17.3%
Machinery and transport equipment	7	1 663 416	54.1%
Miscellaneous manufactured articles	8	340 363	11.1%
Commodities and transactions n.e.c. in the SITC	9	4 591	0.1%

Source: CZSO, own estimation

Table 22: Czech export according to 3-digit SITC

Commodity	Code	Value in mil. CZK
Motor cars	781	295 665
Parts and accessories of the motor vehicles	784	202 533
Automatic data processing machines and units thereof	752	200 817
Telecommunications equipment, n.e.s., and parts	764	95 066
Electr. appararatus for switching electrical circuits	772	83 292
Manufactures of base metal, n.e.s.	699	80 322
Electrical machinery and apparatus, n.e.s.	778	76 857
Equipment for distributing electricity, n.e.s.	773	56 350
Heating a. cooling equipment a. parts thereof, n.e.s.	741	53 207
Baby carriages, toys, games and sporting goods	894	51 496
Furniture and parts thereof	821	47 120
Rubber tyres, tyre treads or flaps a. inner tubes	625	46 281
Monitors and projectors; reception apparatus for television	761	46 140
Electric current	351	44 797
Rotating electric plant and parts thereof, n.e.s.	716	40 316
Iron and steel bars, rods, angles, shapes a. sections	676	39 588
Articles, n.e.s. of plastics	893	38 820
Cathode valves and tubes; diodes; integrated circuits	776	36 694
Pumps (excl. for liquids); parts thereof	743	35 526
Pumps for liquids; liquid elevators	742	31 938

Source: CZSO, own estimation

Table 23: World comparison of production of passenger cars

Rank	Country	Number of produced passenger cars
1.	China	15,523,658
2.	Japan	8,554,219
3.	Germany	5,388,456
4.	South Korea	4,167,089
5.	USA	4,105,853
6.	India	3,285,496
7.	Brazil	2,623,704
8.	Russia	1,968,789
9.	Mexico	1,810,007
10.	France	1,682,814
11.	Spain	1,539,680
12.	United Kingdom	1,464,906
13.	**Czech Republic**	**1,171,774**
14.	Canada	1,040,298
15.	Thailand	945,100
16.	Slovakia	900,000
17.	Iran	871,997
18.	Indonesia	743,501
19.	Turkey	576,660
20.	Poland	540,000
21.	Malaysia	509,621
22.	Belgium	507,204
23.	Argentina	497,376
24.	Others	463,990
25.	Italy	396,817
26.	Romania	326,556
27.	Taiwan	278,043
28.	South Africa	274,873
29.	Hungary	215,440
30.	Australia	178,480
31.	Sweden	162,814
32.	Uzbekistan	144,980
33.	Slovenia	126,836
34.	Austria	123,602
35.	Portugal	115,735
36.	Ukraine	69,687
37.	Egypt	36,880
38.	Netherlands	28,000
39.	Serbia	10,227
40.	Finland	2,900

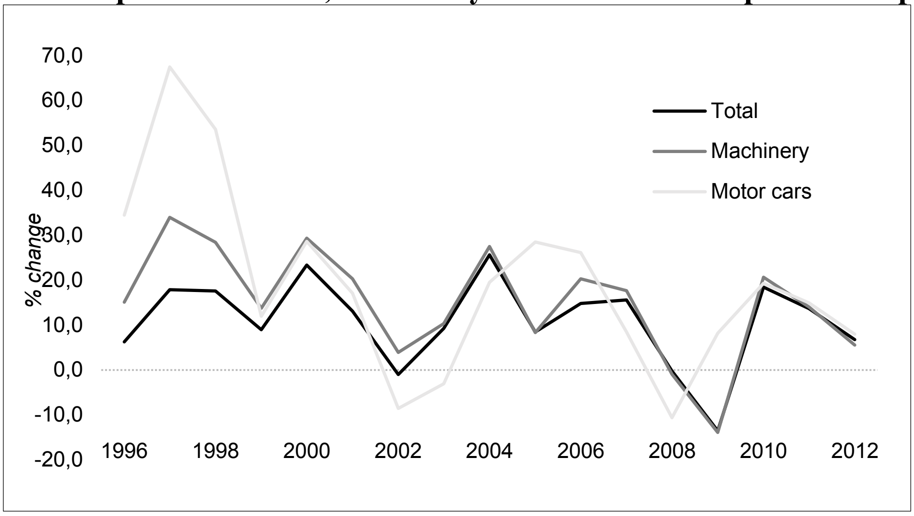
Source: OICA, own estimation

Table 24: List of countries

No.	Code	Country	No.	Code	Country
1	AR	ARGENTINA	26	IS	ICELAND
2	AT	AUSTRIA	27	IT	ITALY
3	BB	BARBADOS	28	JO	JORDAN
4	BE	BELGIUM	29	KZ	KAZAKHSTAN
5	BG	BULGARIA	30	LB	LEBANON
6	BY	BELARUS	31	LT	LITHUANIA
7	CL	CHILE	32	LU	LUXEMBOURG
8	CN	CHINA PR	33	LV	LATVIA
9	CO	COLOMBIA	34	MT	MALTA
10	DE	GERMANY	35	NL	NETHERLANDS
11	DK	DENMARK	36	NO	NORWAY
12	DO	DOMINICAN REPUBLIC	37	PL	POLAND
13	EC	ECUADOR	38	PT	PORTUGAL
14	EE	ESTONIA	39	PY	PARAGUAY
15	EG	EGYPT	40	RO	ROMANIA
16	ES	SPAIN	41	RU	RUSSIA
17	FI	FINLAND	42	SE	SWEDEN
18	FR	FRANCE	43	SI	SLOVENIA
19	GB	UNITED KINGDOM	44	SK	SLOVAKIA
20	GR	GREECE	45	SY	SYRIA
21	HR	CROATIA	46	TM	TURKMENISTAN
22	HU	HUNGARY	47	TR	TURKEY
23	CH	SWITZERLAND	48	UA	UKRAINE
24	IE	IRELAND	49	US	UNITED STATES
25	IL	ISRAEL	50	YE	YEMEN

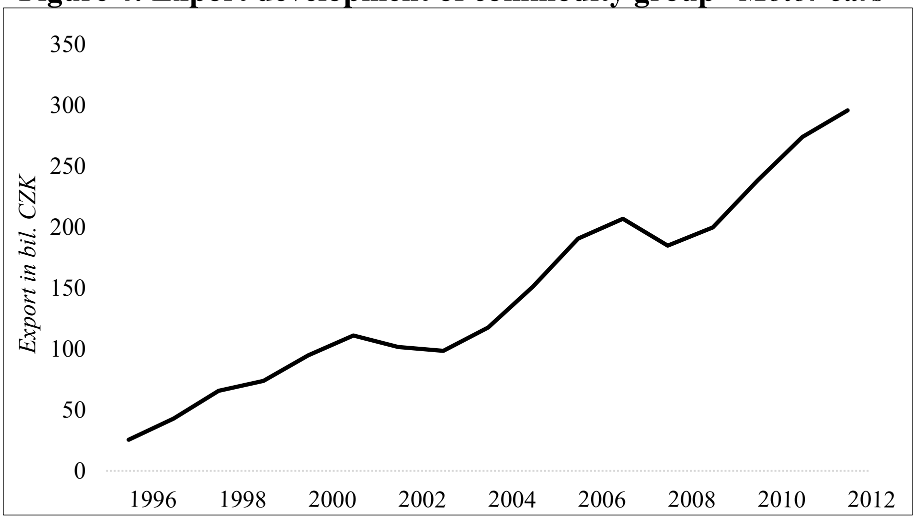
Source: own estimation

Figure 3: Comparison of total, machinery and automobile export developments



Source: CZSO, own estimation

Figure 4: Export development of commodity group "*Motor cars*"



Source: CZSO, own estimation