Stockholm School of Economics
Department of Economics
Master's Thesis in International Economics

The gravity model revisited

- An evaluation of the predictive ability of the gravity model.

ABSTRACT: The end of the communist era in central and Eastern Europe raised many questions about which new patterns of trade would emerge when these countries would liberalize and open up for trade. To answer this question, among other methods, the gravity model of international trade was used to make predictions about future trade flows. In this thesis we evaluate the predictive ability of the gravity model by comparing the trade flows predicted in the end of the 1980s with the actual trade flows referring to the same time. We find that the gravity model gives very poor results in making predictions about future trade flows. There are several reasons for this; first, the model is specific considering both the countries and time for which it was being calibrated. Second, the method for using the gravity model is ad-hoc, based on conventions and has not undergone a thorough theoretical discussion and in connection with this the proxy variables used by the model are poorly examined. Third, the model is not specified to explain trade re-orientation, yet it is for this purpose it has been used. From this we conclude that before the gravity model is used again for explaining changes in trade patterns it needs a thorough theoretical discussion aimed at correcting these problems. Otherwise, we suggest that the model should not be used at all for this purpose.

Authors: Elin Ryrfeldt (18744)
Olof Sundblad (18921)

Tutor: Professor Örjan Sjöberg

Discussants: Anna Eriksson (18802)
Petra Persson (18736)

Examiner: Richard Friberg

Presentation: January 25 2006
10.15-12.00 in room 328

Acknowledgments

We would like to thank our tutor Örjan Sjöberg for his support during our work with this thesis. We are also grateful to Baldwin and Wang and Winters who provided us with much of their data. We would also thank Mattias Frithiof and Robert Östling for their valuable help.

Table of Contents

Introduction

When the Berlin wall fell and when the communists were thrown out of many countries in central and eastern Europe we witnessed the beginning of a new era. Everyone people, policymakers, workers and company leaders -all wondered how the economic patterns would be altered in the new Europe when the distortions on the economic growth and trade were lifted. In order to find answers to these questions, economists started to make predictions about the future trade patterns of Europe. Among other methods, the gravity equation was used for making these predictions.

Today, ten to fifteen years later, this is history and most of the former Centrally Planned Economies (CPEs) are fully functioning market economies welcomed into the European Community. The former trade distortions are supposed to be lifted. Therefore we ask ourselves, how did it go? Has the actual development followed the development predicted in the studies? What conclusions are possible to draw?

The purpose of this thesis is to evaluate, ex-post, the quality of the predictions generated by the gravity model about the trading potential of central and eastern Europe in the early 1990s.

More specifically, we have chosen to evaluate two models making predictions about the trading potential of eastern Europe. Wang & Winters (1991 and 1994) and Baldwin (1994) who both made their predictions early on in the transition of the eastern European economies. Doing so, we believe that we have two possible things to learn. First, we are interested in evaluating whether the gravity equation is doing a good job explaining trade between countries over time. It is widely known that the model achieves good marks in explaining trade between countries at a given point in time. But can trade flows in the future be predicted using the model? Second, by evaluating the results of the predictive power of the gravity equation, we hope to be able to point out what the problems with the model are, and show where more research is necessary in order for researchers to confidently use the gravity equation for this kind of studies.

Structure

In the following, we start out by providing a background, where we present the gravity equation to the reader. We explain its characteristics and describe the uses to which it has been put. We then show the different attempts made to find some theoretical foundation for the gravity model in economics, building upon the work of Anderson (1979), Bergstrand (1985 and 1989) and Helpman and Krugman (1985). We also show what variables are usually used as proxies, as well as some ways to augment the gravity equation and what the implications of that are. The background section is then finished by a section on the transition of the new EU members where we argue that the former CPEs are now fully transformed into market economies.

In the next section, we present the two models of Baldwin (1994) and Wang & Winters (1991), the predictive ability of which we are testing in this thesis, and we motivate why these models are of certain interest to evaluate. In the data section we present how we have collected the data for the evaluation, the characteristics of our data and what computations have been necessary in order to make the evaluation. What we have done is to use our data to put into the estimated regression models made by Baldwin and Wang & Winters. The results of the predictions have then been compared to the actual trade flow between these particular countries, those particular years.

In the analysis we first present the results of our evaluation, where we are conclude that the models have done a poor job in predicting the development of trade flows to and from Central and Eastern Europe. We therefore ask ourselves why we get these results, and investigate whether the model is too specific, both considering time and calibration or whether there is a poor methodology underlying the use of the model and the specification of proxy variables. We conclude that we have a problem in the trade off between the ability of using a model for general conclusions and estimating it under specific circumstances. The essay is then finished by a section where we open up for more research in order to calibrate the model better.

Background

The Gravity Model

In its purest form, the Newtonian gravity equation, presented in 1687, describes the attractive force (F) between two objects *i* and *j* in the universe as dependent on the masses (M) of the objects, the distance (D) between them and some gravitational constant (G)[1]:

$$F_{ij} = G\frac{M_i M_j}{D_{ij}^2}$$

Among the first writers to use the gravity equation for describing trade among the world's market economies was Jan Tinbergen who in 1962 showed how this concept could be used to describe trade flows. Hans Linnemann (1966) was one of Tinbergen's students in Rotterdam and he continued to develop the concept. The aim of Linnemann's study was to find a quantitative explanation of the trade flows of commodity trade between individual countries.

In Linnemann's experiment he modeled the size of a bilateral trade flow (X) to be dependent of the importer's demand, the exporter's supply and "resistance" to a trade flow, i.e. their costs of doing business. In his equation, he included as explanatory variables, the respective countries gross national product (Y), their population size (N), their graphical distance (D) and potential preferential-trade factors (P). The equation is summarized as follows (Linnemann, 1966: 36):

$$X_{ij} = \delta_0 \frac{Y_i^{\delta_1} Y_j^{\delta_3} P_{ij}^{\delta_6}}{N_i^{\delta_2} N_j^{\delta_4} D_{ij}^{\delta_5}}$$

[1] An overview of the development of the gravity equation in economics is provided by Keith Head (2003).

Linnemann based his study on 6300 bilateral trade flows between eighty countries, deliberately excluding all trade with and between Communist countries. From this he made several conclusions about the size of the trade flows. Among other things, he found that there could be no established effect of per capita income on trade since he found a constant proportionality between trade and each country's national product and he concluded that geographical distance could be used as a proxy for trade-reducing effects of natural barriers to trade. Moreover, population size has a negative effect on trade flows because of a trade reducing effect, i.e. that large countries have better ability to be self-sufficient.

The gravity equation can best be understood as the long-term equilibrium of a market economy's volume of trade in the current context. Hence, as stated by Wang & Winters (1994), eastern Europe was not included in the first estimations of the gravity equation. However, they made the explicit assumption that eastern Europe eventually would get the same trading patterns as the sample of market economies included in the study. All writers making predictions about the trading patterns among the transition economies use this assumption.

Other than for economic trade flows, the gravity model has been used to explain flows of migration, commuting, tourism and commodity shipping (Bergstrand, 1985).

The theoretical foundation of the gravity model

Economists using the gravity equation investigating trade flows motivate the use of their model by referring to studies searching for a theoretical underpinning of the model. Generally, they refer to the work of Anderson (1979), Bergstrand (1985 and 1989) and Helpman and Krugman (1985). Because of this, we present the theoretical foundation suggested by these writers.

Anderson (1979) sets out to explain the theory behind the gravity equation applied to commodities. His theory explains the multiplicative form of the equation, it makes an interpretation of the distance variable, and it assumes that all countries have identical expenditure functions. It does, however, not explain how, or why, income and population affects the trade share of the budget, nor does it explain the linkage between a short-run model with prices and the long-run model that is commonly used. In his theories, Anderson starts out from the Cobb-Douglas expenditure system, with one good for each country, no tariffs and no transportation costs. From this he explains the flow from country *i* to country *j* as proportional to the GDP of the importer and exporter, by applying a constraint on the balance of payments of the countries. The next step is to take into account interregional and international variations in the shares of traded goods in the total expenditure. As a final step, he allows for trade frictions.

Helpman and Krugman (1985:chapter 8) continue to explain the gravity equation with support from the outline made by Anderson (1979). They set out from a case where there is complete specialization in production, either intra-industry or inter-sectoral specialization, increasing returns to scale in the tradable sector and identical homothetic preferences as well as access to the same prices by all customers. In this case, they conclude that each country in the world consumes a fraction of every good that is produced in the world economy and that it, likewise, is exporting a fraction of each good produced in the country. Exports are therefore expressed as a fraction of GDP in a single country, but also of the other countries in the world. Since economies of scale leads to more specialization, the gravity equation will typically perform better the more important increasing returns are in production. This model is a two-sector model where all consumers face the same prices, which simplifies the connection a great deal. The authors express the trade flow from *i* to *j* as,

$$X_{ij} = \beta_j Y_i = \beta_j \beta_i \overline{Y},$$

where $\overline{Y}$ is the world GDP and B_i and B_j are the share of world GDP produced in *i* and *j* respectively. The total volume of trade in the world as,

$$VT_w = \beta_w \left[1 - \sum_{i \in w} \left(\beta_w^i\right)^2\right] Y_w$$

Bergstrand (1985) finds it problematic that prices generally are excluded from the gravity equation. Earlier writers argue that prices only have the function of equating supply and demand.[2] Bergstrand then presents a general equilibrium model of world trade, building on a utility and profit maximizing agent in each country, where each country holds one factor of production. The result is a model that explains the trade flows as a function of all available resources in the country, taking both trade barriers and transportation costs into account. This model, however, may not explain the gravity equation since it treats the bilateral trade flow as endogenous variables instead of exogenous.

Bergstrand simplifies his model by making several assumptions; 1, the trade flow from country *i* to *j* is small relative to the world trade; 2, all countries have identical utility and production functions – in accordance with the Heckscher-Ohlin-Samuelson model of inter industry trade; 3, perfect substitutability in production and consumption, which means that trade flows are differentiated by national origin; 4, perfect commodity arbitrage; 5, zero tariffs; 6, zero transport costs. However, it is a problem that price is not included in the model specification. To solve this issue, Bergstrand specifies a new "augmented" gravity equation that includes price variables. Empirical testing and deviations from the purchasing power parity across nations suggests that the generalized gravity equation is more appropriate than the partial equilibrium approach, i.e. the assumptions 3-6 are refutable. Bergstrand's final model is, however, not the commonly specified gravity equation, since prices are included.

Bergstrand (1989) continues the mission of finding theoretical foundations for the gravity equation. A few years after his first try, Bergstrand makes a new attempt to formulate a gravity model including both income and population. The result is a general equilibrium model of world trade with two factors, multi-industries and multi-countries. Compared to the earlier study from 1985, this explanation also incorporates relative factor endowments and non-homothetic tastes. He starts out by assuming that individuals are maximizing a "nested" Cobb-Douglas-CES-Stone-Geary utility function where demand is dependent on national income, per capita income and import prices. It is shown that the national income elasticity of demand for the manufactured (none-manufactured) good is larger (smaller) than

[2] See for example Linnemann (1966:41) where he assumes "…that for individual countries the 'general world price level' is given…".

one if per capita income rises. Firms are assumed to operate in an environment of simultaneous intra- and inter-industry trade characterized by monopolistic competition. Firms in this market are producing uniquely differentiated products with increasing returns to scale and distributing them to markets under diminishing returns.

The equation developed by Bergstrand suggests that the products traded are luxuries in consumption and capital intensive in production. He argues that this is not unrealistic since the gravity equation is estimated based on trade among major industrialized countries. This model does still differ from the commonly used gravity model due to the inclusion of price variables. The assumption of uniquely differentiated products makes the existence of a world price impossible by definition. There have been several attempts to generalize the gravity model into a multi-industry model. Bergstrand is providing one himself, but concludes that his results are indecisive and that more research is necessary.

Generally, it is held among economists that the gravity model has a weak link between theory and empirics, and that the fitted models are merely descriptive. This makes it hard to interpret the estimated model results. Leamer and Levinsohn (1995) identify Anderson (1979) as the formally most correct theoretical underpinning, but admit that it is too complex to be generally used. They explain the lack of theoretical explanation by two things; that the gravity model does not belong to the same theoretical domain as general international economics – whereas international economics is interested in trade of a country vis-à-vis the world, the gravity model is dealing with questions of total trade between pairs of countries. The other explanation provided by Leamer and Levinsohn is that usage of the gravity model does not discuss the benefits of free trade.

Economics researchers like to argue that they are working with 'science'. They follow the scientific method of stating a theory that may be generalized, then tested and refuted or verified. That is, they have sound statistical underpinning. Economics build upon three fundamental themes. First the assumption that individuals engage in maximizing rational behavior, second that finding equilibrium is an important goal of every theory, and third the importance of efficiency (Lazear, 2000).

In this sense the gravity equation is problematic. The theory in itself was developed in the 1960s and did then constitute a quantitative description of trade patterns between the world's market economies (see Linnemann, 1966). As we have seen, it was not until years later, these theories started to get some sort of theoretical explanation. Some people, particularly within social science think that social science is not supposed to be consisting of theoretical paradigms, rather, it should be an accumulation of empirical findings (Sutton and Staw, 1995). However, we daresay that most economics researchers do not share this view.

What is theory? Researchers do not explicitly agree on whether a model and a theory can be distinguished, they, however, find it easier to agree on what theory is not. Sutton and Staw (1995) have identified five features of academic work that are not theory. These five features are references, data, variables, diagrams and hypotheses. They have noticed that many writers are trying to hide their lack of theory behind a facade of references, and that they do not explain to the reader why these references are relevant or how they help in developing the theory. Data are not theory, data simply describe *which* empirical patterns were observed, whereas theory should explain *why* these patterns were observed or why they are expected to be observed. Moreover, listing interesting variables affecting a certain pattern does not constitute a theoretical argument. A good theory may be visualized by the usage of diagrams, but such devices do not in themselves explain anything. Finally, hypotheses are supposed to serve as a bridge between data and theory, making the logical argument explicit, they state *what* is expected, not *why*.

Is the theory explaining the gravity models a strong theory or not? Our authors seem to be divided. On the one hand, Baldwin (1994) agrees that the gravity model has a poor reputation due to a historical lack of theory that now has been remedied. He thinks the model has sound theoretical foundations and reviews the work of Helpman & Krugman (1985), Bergstrand (1989) and Huang (1993). Wang & Winters (1994:13) on the other hand admit that *"…several authors have tried to provide the model with such a theoretical underpinning…"* They particularly mention Anderson (1979), Bergstrand (1985) and Helpman & Krugman (1985), but none of these authors have been successful in explaining exactly the model specified by Wang & Winters. They end this section by concluding that "[f]or our purposes the important issue is the empirical application of the model to trade flows between countries, and hence we are more concerned with the model's empirical robustness than with its theoretical purity. Nonetheless, theoretical soundness is comforting…" (Wang & Winters, 1994:13) and from this they set out making a brief review of the earlier mentioned theories.

From this it is clear that the authors have differing views on the quality and the need for a strong theory supporting their models. However, neither of the studies actively uses the theory backing up the models. They do not explicitly take the theory into account when they formulate their model. Instead they go on taking the model as given by prior authors having estimated world trade using the gravity equation. Why the different proxy variables included in the model are there is explained as departing from the gravity model as such and not from economics or trade theory. The standard, or convention, for how a gravity study is supposed to be done was thus in many ways set already by Linnemann (1966). None of the studies we have read, working with the gravity model has included a discussion about the method of their study, instead they again refer to what has already been done by other writers.

Specification of the Gravity Equation

Head (2003) provides a discussion on how the gravity equation commonly is estimated and what the included variables are supposed to represent. The economic mass in the equation is usually represented by gross domestic product. Some studies have also occasionally used population. The trade reducing formula is a little bit less straightforward. The most commonly used variable is distance, measured by the "great circle" formula, which calculates the minimum distance along the earth's surface. The great circle distance is a measure taking into account the spherical shape of the earth and the coordinates of points between which the distance is measured. The formula is,

$$\text{distance} = r * \text{acos}\left[\sin(lat1) * \sin(lat2) + \cos(lat1) * \cos(lat2) * \cos(lon2 - lon1)\right]$$

Where r is the radius of the earth in whatever desired unit, which is $r = 3437.74677$ (nautical miles), $r = 6378.7$ (kilometers) or $r = 3963.0$ (statute miles). *lat*1 and *lon*1 are the latitude and longitude of the first point and *lat*2, *lon*2 are the coordinates of the second point.

Other writers, such as Wang & Winters instead use the road and shipping distance in nautical miles. The distance variable is one of the key variables in the gravity equation when it comes to implication for the outcome. The distance variable is used as a proxy to explain many factors influencing trade, among them:

1. Transport costs. The cost of physically shipping the products.
2. Time elapsed during shipment. Rather, the risks included in the shipping that may damage the product in any sense.
3. Synchronization costs. Costs of a Just-In Time system not functioning properly or the cost of having to store goods in a warehouse before they are used.
4. Communication costs. (Im)possibilities of personal contact and transmitting of information.
5. Transaction costs. Costs of searching for new trade opportunities and the establishment of trade partnerships.
6. "Cultural distance". Costs caused by cultural misunderstandings or other cultural clashes in trade.

The typical distance effect is reported to be about -0.6, suggesting that doubling the distance will reduce trade by a half.

There have been several attempts of augmenting the gravity model through inclusion of border effects and preferential trade agreements. Some researches also include variables such as income per capita, because they believe that high-income countries trade more in general. It is moreover popular to include dummy variables, for adjacency (a common border) or other trade creating effects, such as sharing a common language or having had colonial links. It is common for the adjacency variable to lie close to 0.5, in a log-linear gravity relationship implying that trade is 65 percent higher between countries that share a common border. Countries that speak the same language are trading two to three times as much as countries that do not speak the same language.

On the other hand, unlike applications of spatial interaction models in other disciplines like geography, transport or migration studies, intervening opportunities and effects of third countries are seldom explicitly entered into the discussion, meaning that each good is purchased in the nearest place offering it, when exception is made for multi-purpose trips (Olsson, 1965). A discussion like that is likely to have implications for the result of the predicted trade flows as it does so in Olsson's studies about migration. Put differently, there is a measure of endogenity in the model that is not properly accounted for.

The Models

We have chosen to evaluate two models making predictions about the trading potential of Eastern Europe. Wang & Winters (1991 and 1994) and Baldwin (1994) who all have made their predictions early on in the transition of the Eastern European economies. Both these models constitute a small part in larger studies on trade patterns. Wang & Winters wrote their paper as a part of a Centre for Economic Policy Research (CEPR) study on "Eastern Europe's international trade" discussing how EU should treat the newly liberalized Eastern European economies. Baldwin also made his contribution to a CEPR study, "Towards an Integrated Europe", concerning the EU enlargement.

Wang & Winters

When predicting their model, Wang & Winters use data for GDP, Population and Distance. Moreover, they include dummies for adjacency and preferential trade agreements if the countries have any kind of trade agreement. The dummy variables in their model are assumed to take the value 2 if a condition is satisfied and 1 otherwise.

Their regression is specified as follows:

$$\ln X_{ij} = a + \beta_1 \ln Y_i + \beta_2 \ln N_i + \beta_3 \ln Y_j + \beta_4 \ln N_j + \beta_5 \ln D_{ij} + \beta_6 \ln A_{ij} + \sum_k \gamma_k \ln P_{kij}$$

Where X_{ij} is the trade flow from country i to country j measured in thousands of dollars. Y and N is representing the countries respective GDP and population. D is the distance between the countries economic centers. A is a dummy for adjacency and P is a number of preferential trade agreements.

Wang & Winters uses data from 76 countries, 19 industrial and 57 developing market economies and in total this sums up to 7,500 bilateral trade flows, between the years of 1984-1986, averaged over a three year period. They first calculate an extensive model (Table 1) which they soon re-parameterized into a shorter version (Table 2) later used by the authors when they examined the trade flows in Eastern Europe.

Constant	GDP Exporter	GDP Importer	Population Exporter	Population Importer	Distance	Adjacency
-12.49	1.02	1.17	-0.22	-0.38	-0.75	0.78
(32.42)	(42.75)	(58.19)	(8.19)	(15.67)	(22.28)	(3.27)

EEC	EFTA	UK	France	ACP	GSP	aid
0.70	-0.02	1.91	0.73	1.05	0.35	0.89
(2.17)	(0.05)	(4.96)	(1.24)	(5.27)	(2.92)	(4.20)

ECOWAS	SADCC	CACM	AG	LAIA	ASEAN	
0.31	1.25	2.10	0.38	0.96	2.25	
(0.34)	(0.97)	(1.32)	(0.55)	(2.85)	(5.15)	

EEC, EFTA, UK, France, ACP, GSP, aid, ECOWAS, SADCC, CACM, AG, LAIA and ASEAN all correspond to preferential trade agreements in place at the time of Wang & Winters' study. All variables are expressed in their natural logarithm, ln. T-statistics are shown in brackets. Wang & Winters (1991).

Table 1 The first model specified by Wang & Winters.

Constant	GDP per Capita		Population		Distance	Adjacency
	Exporter	Importer	Exporter	Importer		
-12.49	1.17	1.02	0.79	0.8	-0.75	0.78

Wang & Winters (1991)

Table 2 The model used by Wang & Winters.

The model of Wang & Winters is providing a goodness of fit of 0.70. Since we are only interested in the trade between the newly liberalized parts of Eastern Europe and these countries' trade with the world, we have no use for the preferential trade agreements included in Figure 1.

Baldwin

Baldwin follows a very similar procedure to Wang & Winters and his model includes variables for GDP per capita, population and distance. He includes only two dummy variables; adjacency and membership of EEA, which he specifies as the European Union and EFTA. The reason for this is that Baldwin bases his model solely on trade between western Europe, EFTA, the US, Japan, Canada and Turkey, that is 17 countries. In total, his data consisted of 3,390 observations, that is, panel data from the ten-year period between 1979 to 1988 expressed in 1985 international dollar.

Baldwin's regression follows;

$$X_{ij} = \beta_1 + \beta_2 \ln \frac{GDP_e}{Pop_e} + \beta_3 \ln Pop_e + \beta_4 \ln \frac{GDP_i}{Pop_i}_i + \beta_5 \ln Pop_i + \beta_6 \ln D + \beta_7 A + \beta_8 EEA$$

Where X_{ij} is the bilateral trade flow from country *i* to country *j*, GDP and Pop are the GDP and population of the exporter (e) and importer (i) respectively. D is a dummy for distance and A and EEA are dummies for adjacency and membership in EEA.

Constant	GDP per Capita		Population		Distance	Adjacency	EEA
	Exporter	Importer	Exporter	Importer			
-17.5	1.16	1.22	0.77	0.79	-0.88	0.28	0.53
(12)	(13)	(16)	(26)	(25)	(11)	(2)	(3)

The T-statistics are shown in brackets. Baldwin (1994)

Table 3 The gravity model specified by Baldwin.

The model's R^2-value of Baldwin's models is 0.99, an extremely high value.

Distance data

One of the larger differences between the models is how they treat the distance variable. Baldwin expresses the distance in kilometres as a straight line between the capitals of the countries.

Wang & Winters on the other hand use a much more complicated measure and start out with the shortest navigable distance, in nautical miles, between the countries main ports thereafter adding the road distance from the ports to the country's main economic centre. For continental Europe, however, only road distance is used. They used *Reed's Marine Distance Tables* and *Philip's Distance Tables*.

Data sources

Baldwin and Wang & Winters have to some extent used different sources for their data collections. Wang & Winters received the data on bilateral trade flows, measured from the side of the importer, from the IMF's Direction of Trade Statistics Yearbook (DOTSY). Baldwin used the UN's COMTRADE database and measured data from the exporter's side.

Wang & Winters have collected their data on GDP and population from the World Banks' World Development Indicators. Baldwin on the other hand used the Summers and Heston database for population and trade. Moreover, he collected his data on Eastern European GDP figures from PlanEcon.

How to test the models

In order to test how good the models are for making predictions about future trade with the new EU member countries it has been necessary for us to decide on a point in time where such an evaluation is possible. As we will argue later, we believe that the restrictions that were put on trade due to the central planning systems now are lifted. We may therefore treat these countries as being market economies.

We evaluate the models on a continuous basis starting from 1993 and until 2003, the year before the accession of these countries to EU. For reasons explained later, we have in many cases been forced to present our studies with data from 1993 to 2003. The reason for choosing to do the evaluation up to 2003, the year before the EU-entry is that we do not wish the actual enlargement to influence our results.

We have also narrowed down our analysis to exclude the estimated trade before 1993. The reason for this is that data are very imprecise before this year. Part of the problem is due to the fact that only three of the new EU members report trade before this year, namely Czechoslovakia, Hungary and Poland. The end of 1992 was also the year when Czechoslovakia split up into the Czech Republic and the Slovak Republic. Since we do not wish to control for factors like this, we have simply chosen to evaluate the models between the years of 1993 to 2003. Another reason for excluding the first years of the 1990s is that among the eastern European countries, only Romania reports frequent trade with the new EU, however, the countries from the new EU reports trade to the eastern European countries. We could have corrected for this, but the more corrections we make, the more uncertainty we bring into the model.

We have collected data on bilateral trade flows, GDP and population for the 33 countries included in our study. Where we received the data from and how the data are treated is explained in detail in the next section on data.

We have calculated the predicted trade flows using the models presented by Baldwin and Wang & Winters and compared the results with the actual trade flows from the same time. The results have later been presented using the four groups presented below and visualized using graphs in the analysis. If the gravity estimations have made good forecasts of the future trade development, we expect to see a declining trade among the former centrally planned economies and increasing trade among the western economies and the former CPEs. We also expect the discrepancy between actual trade and potential trade to decline with time as the historical trade distortions disappear for the CPEs. We have support for this hypothesis from, among others, Jakab et. al. (2001).

Is the transition over?

When Baldwin and Wang & Winters wrote their articles, they all had an idea about what the trade patterns of Eastern Europe would be like, once the restrictions imposed by communism were lifted. Their general idea was that trade under communism was distorted, and that once communism were gone, trade would follow some predictable normal pattern.

Communism resulted in distortions on trade in two distinct ways. First, the CPEs economic performance were far below the performance of the Western European countries, resulting in lower GDP per capita in the East. Second, the CPEs were, as implied by their name, centrally planned, and so was their trade. They followed a system with trade monopolies, and used a system of fixed exchange rate "conversion rates". They were also trying to become self-sufficient within the so-called Eastern bloc, by producing everything themselves.

As such, the former CPEs not only have to bother about being transformed to a market economy, but also to fight a large income catch up. The predictions about their future trade focused mostly on how the trade patterns would look like when the CPEs have reached the same level of income as the countries in the western parts in Europe. In order to evaluate the predictions we need to examine the CPEs and see whether they have actually transformed and thereafter evaluate the predictions themselves. The countries became members of the European Union in 2004 and we therefore look at which of the countries had completed their transition from a planned economy to a market economy by 2003, the year their EU entry was decided upon.

We are aware of the fact that all definitions about how far the transformation of the CPEs is from being fully transformed are somewhat arbitrary. What a country needs to do to transform itself is not clear-cut, and there is no country in the world that is absolutely free from trade restrictions, like restrictive laws or taxes created to hinder trade. Many countries in western Europe have regulated prices on some sectors of their economies; others have state ownership in large export companies.

In order to see how the new EU members have done with finishing their transition process by 2003, we have chosen to use the Transition Report (2003) published by the European Bank of Reconstruction and Development (EBRD). EBRD was established in 1991 and has from the start aimed at supporting the ex-Soviet block countries to nurture a private sector in a democratic environment. 60 countries together with the two inter-governmental organisations, the European Community and the European Investment Bank, own it. Also the former CPEs are members themselves. The method used in supporting the transition is foreign direct investments. The transition report from the EBRD evaluates the transition using a number of indicators, constructed to measure the reformed economy and the development of democracy and rule of law.

The transition of the economy is measured in nine areas, covering the four elements of market economy, trade, infrastructure and enterprises, and financial institutions. The indicators are:

1. large scale privatization
2. small scale privatization
3. governance and enterprise restructuring
4. price liberalization
5. trade and foreign exchange systems
6. competition policy
7. banking reforms and interest rate liberalization
8. security market and none-bank financial institutions
9. infrastructure

The countries are graded on a scale from 1 to 4+, where 1 means that virtually no reform has been undertaken and 4+ means that the country has the same standard as an industrialized market economy. Unfortunately there are no CPEs received 4 on all measures, so we had to use some other criteria. The new EU members which were the countries of most interest to us, however, got 4+ in the fifth category, meaning that they are members of WTO and have removed most barriers to trade. Moreover, they had 4+ in the second criteria, implying that there is no state ownership in small-scale businesses. To sum up, on none of the criteria they had less than 3. Using this evaluation, we are fully confident that Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovak Republic and Slovenia are transformed market economies.

If measuring the transition process concerning the economy is hard, measuring the level of democracy in a country is even harder. The EBRD has several indicators for how democracy and rule of law are developing in the CPEs. We will, however, not go into detail with these, but rather simply conclude that our eight countries are democratic enough to make the accession into the European Union, why we assume that they can be characterized as functioning market economies.

Which countries and why

Primarily we have been interested in the countries whose transition period is over and who have turned into full market economies. Hence, we have chosen to focus our study on the countries that are done with their transition. This has been described in more detail earlier and the countries are: Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovak Republic and Slovenia

Our logic for testing the predictive power of the gravity equation among different countries was that we wanted to group them into four distinct categories. First, we have the free trade area of EU, to which our former centrally planned economies now belong. Then we also wished to pair them with other industrial countries as well as some of the old centrally planned economies with which our eight countries used to "trade" in the old planning system.

For industrial countries other than EU, we have chosen to look at the part of OECD situated outside Europe. Excluded are thus Iceland, Norway, Switzerland and Turkey. We had to exclude Mexico as well due to significantly low or completely missing trade flows. We would simply not have data to compare with for most of our evaluation years.

As for Eastern Europe, we have included only five European ex-communist countries. We excluded the ex-Yugoslavian countries, Albania and Moldova because of the persistent turmoil in these countries that are distorting trade in unpredictable ways.

Overall, we ended up with a sample of 34 countries grouped into four categories: new EU, EU, OECD and Eastern Europe. The 34 countries quickly became 33, since the models we evaluate, as well as our data sources, treat Belgium and Luxembourg as one single entity.

New EU	EU-15	OECD	Eastern Europe
Czech Republic	Austria	Australia	Belarus
Estonia	Bel/Lux	Canada	Bulgaria
Hungary	Denmark	Japan	Romania
Latvia	Finland	Korea	Russia
Lithuania	France	New Zealand	Ukraine
Poland	Germany	USA	
Slovak Republic	Greece		
Slovenia	Holland		
	Ireland		
	Italy		
	Portugal		
	Spain		
	Sweden		
	U.K.		

Table 4 The countries included in our evaluation study.

Data

In our evaluations of the studies, it has been important for us to get the same quality of the data as the writers themselves used, in order to control for errors due to differences in measurement techniques. Hence, our data on GDP and population are collected from the World Bank's statistical (on-line) database *World Development Indicators* (WDI). Data on exchange rates and GDP-deflators are taken from IMF's *International Financial Statistics Yearbook* (IFS). Finally, our bilateral trade flow data are collected from IMF's *Direction of Trade Statistics Yearbook* (DOTSY).

Distance

We have tried to use distance data that are as similar to the one used by the authors. We have requested and received some of the distance data sets. This includes the data set from **Wang & Winters**. It has been quite some detective work to understand the data and it contains a lot of errors and very few explanations, which is why we have chosen to use proxies for some of the distances. We use air distance between the European capitals instead of the road distances in nautical miles used by Wang & Winters. We have used the nautical distances between the two closest ports provided by Wang & Winters for distances when estimating trade between countries in Europe and countries on other continents. It should be noted that for Slovakia we use the same distance as for the Czech Republic, Slovenia uses Greece as a proxy and Estonia, Latvia and Lithuania is estimated by Finland's distance to countries outside Europe. This is because these countries did not exist when Wang & Winters wrote their paper. Since we only made these corrections for overseas trade, we believe it is of minor importance for our results.

We also received the data sets from **Baldwin**, who contrary to Wang & Winters uses the straight-line distance between the countries' capitals expressed in kilometres according to the great circle formula. Baldwin's dataset contained data from all European countries in our data set, but we have had to complete the data with distances to the non-European countries (i.e. Australia, Canada, New Zealand, Japan, Korea and the United States). The distances for these countries we have obtained from an internet program designed to calculate the distance according to the great circle formula (DBCC and GCDW).

Current vs. constant prices

Since our data set contains data for a long period, between the years of 1993 to 2003 and since some of our data are expressed in current dollars and other parts of our data are expressed in constant dollar prices, it is important to remember the differences between current and constant prices. Current prices are an expression of the value of output today, expressed in prices of today. Constant prices on the other hand, is the value of output produced today, expressed in the prices of a certain point in time, i.e. the base year. In analyzing changes in GDP, expressing the data in constant prices is valuable, because they control for inflation. The best way of rebasing the data from one year to another is to use the GDP deflator. The reason for this being that a Consumer Price Index takes the output as given, whereas the GDP deflator allows for changes in the composition of output.

In order to be able to evaluate the chosen estimations, we found it necessary to rebase all our data to the same year as the base-year of the regression analyses in the studies. Baldwin, and Wang & Winters uses data expressed in 1985 year dollar value. We believe that a rebasing of the data is necessary because the models are not designed to account for inflation and general growth, but merely to estimate the trade between given nations at given volumes of GDP or population.

The data computations were conducted as follows: data on GDP and GDP per capita are expressed in constant USD prices for base year 2000. These data are merely rebased using the United States GDP deflator with base year 2000 = 100 to the desired year by means of a very simple formula:

$$GDP_{1985} = GDP_{2000} \text{ x } (1 - \text{GDP deflator}).$$

For example in 1998, Poland had a GDP of 153,908,000,000 in constant 2000 USD and the GDP deflator (base year = 2000) for the United States for 1985 is 67,713. This means that Poland's GDP from 1998 was 49,692,275,960 expressed in 1985 value.

The data on bilateral trade flows were a bit more complicated to rebase, and included several more steps on the road. The trade data were collected in current USD prices, therefore we had to use three steps to rebase them:

1. The data was first exchanged back to the national currency using the exchange rates from the IFS. Most countries report their data to DOTSY in their national currency (DOTSY 2004:ix) these figures are then converted to dollars using the exchange rates published in IFS country pages, namely the series rf or rh. These exchange rates refers to period averages of market exchange rate or official exchange rate of the reporting country (IFS 1998:ix).
2. After this conversion, we deflated the data using the particular country's GDP deflator to the desired year, β*, and
3. After this exchanged the values back into dollars, using the exchange rate for that year also reported in IFS.

For every bilateral trade flow we have had the possibility of using two different deflators when rebasing the data. When we wish to express the trade flow between Sweden and Czech Republic from 2000 in 1985-constant dollars we may chose to either use the Swedish or the Czech GDP deflator. Note that the value in 1985 is dependent on the differing GDP deflators between these years and that the result may be significantly different depending on our choice of deflator. Thus, in all cases when we are concerned with trade between one old market economy and our eight countries we have chosen to use the deflator of the old market economy. The reason for this is that these deflators have been more stable over time than the deflators of the transition economies.

When two of the new EU-members have traded with each other, and both measures are equally unstable, we have consequently used the GD deflator of the importer in the data rebasing. Finally, when one of our new EU members was trading with any of the eastern European countries in our sample we have chosen the GDP deflator of the new EU member, simply because there are no reported GDP deflators for Russia and these are only available during a very short period for the other countries.

Another problem with rebasing trade between the former centrally planned economies to years before 1992-93 is that some of the countries simply did not exist before these years. In these cases, we have followed the method suggested by Baldwin, where we have rebased the values as far as our data allows, and thereafter converted our result to US dollars with the exchange rate applicable at that point in time. After this, we have used the US GDP-deflator for the last years.

Here we can see the equation we used for transforming the trade flow in current dollars to the value of dollars in the base year.

$$TradeBY_{ij\alpha} = Trade_{ij\alpha}\left(\frac{Exchrate_{J*\alpha}}{Exchrate_{j*\beta*}}\right)\left(\frac{deflator_{j*\beta*}}{deflator_{j*\alpha}}\right)\left(\frac{deflator_{US\beta}}{deflator_{US\beta*}}\right)$$

$TradeBY_{ij\alpha}$ is the value of the trade flow from country *i* to country *j*, for year α, measured in base year dollar.

$Trade_{ij\alpha}$ is the value of the trade flow from country *i* to country *j*, for year α, measured in current dollar.

$Exchrate_{J*\alpha}$ is the exchange rate, measured as national currency per US dollar, for the GDP-deflator country, j*, in year α

$Exchrate_{j*\beta*}$ is the exchange rate, measured as national currency per US dollar, for the deflator country, j*, in the year β*, note that this often the same as the base year β.

$deflator_{j*\beta*}$ is the GDP deflator for country deflating country for year β* The deflator county is normally the importing country but this also depends on what data is available for the different countries, as described above.

$deflator_{j*\alpha}$ is the GDP deflator for country deflating country for year α.

$deflator_{US\beta}$ is the US GDP deflator for the base year.

$deflator_{US\beta*}$ is the US GDP deflator for the year β*.

Note that in most cases the terms of $deflator_{j*\beta*}$ and $deflator_{US\beta*}$ are excluded.

Import vs. export data

All data in DOTSY appear twice in the sense that Sweden reports its export to the Czech Republic and the Czech Republic reports its imports from Sweden. These numbers are not necessarily equal. Import data are generally said to be more reliable since importers are assumed to have better knowledge of what is actually brought to the country than vice versa (Baldwin, 1994:85). We have, however, chosen to use the same data as the writers, implying that in the case of Wang & Winters we have used data collected from the importer and in the case of Baldwin we have used data reported from the exporter. This might result in some differences in results between the two models.

The Euro-zone

The Euro area, consisting of Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal and Spain decided on introducing the Euro as their currency in 1998 (Council Regulation 2866/91). Hence, from January 1, 1999 all Euro member states have a new currency. We have converted the Euro back to their former currencies using the conversion rate provided to us in the Council Regulation from above.

Belgium and Luxembourg

Our data sources, as well as the writers of the previous studies report Belgium and Luxembourg as one economic entity until 1996. For this reason we have decided to treat them as one entity throughout our data set, simply by adding their GDP and trade data. As capital when calculating the distances, we have chosen Brussels.

Analysis

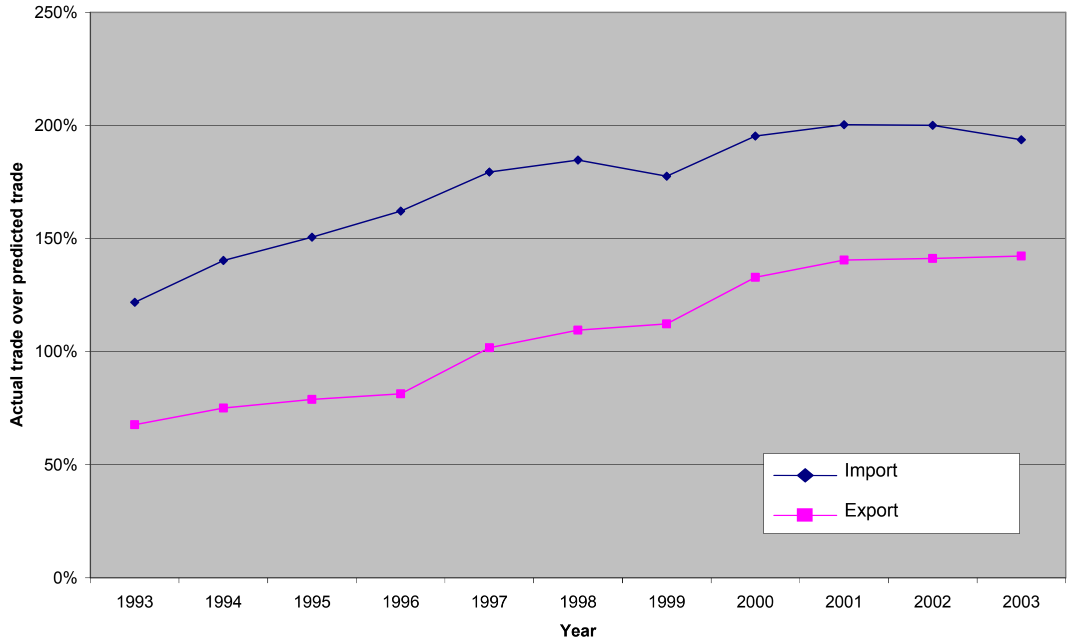
In order to analyze the question, we have calculated the predicted trade flows of the former centrally planned economies that have now made their accession into the European Union. We have chosen to name these countries 'the new EU'. We have made our calculations for every year between 1993 to 2003, a period of eleven years in which the selected countries have gone through a fundamental transformation from being command economies to today being accepted as market economies.

In this section the results of the analysis is presented and an explanation of the results is offered. To sum up, there are also suggestions for what may be done in order to overcome the shortcomings the model is clearly demonstrating. But first things first, let us start with looking at our results.

The model predictions

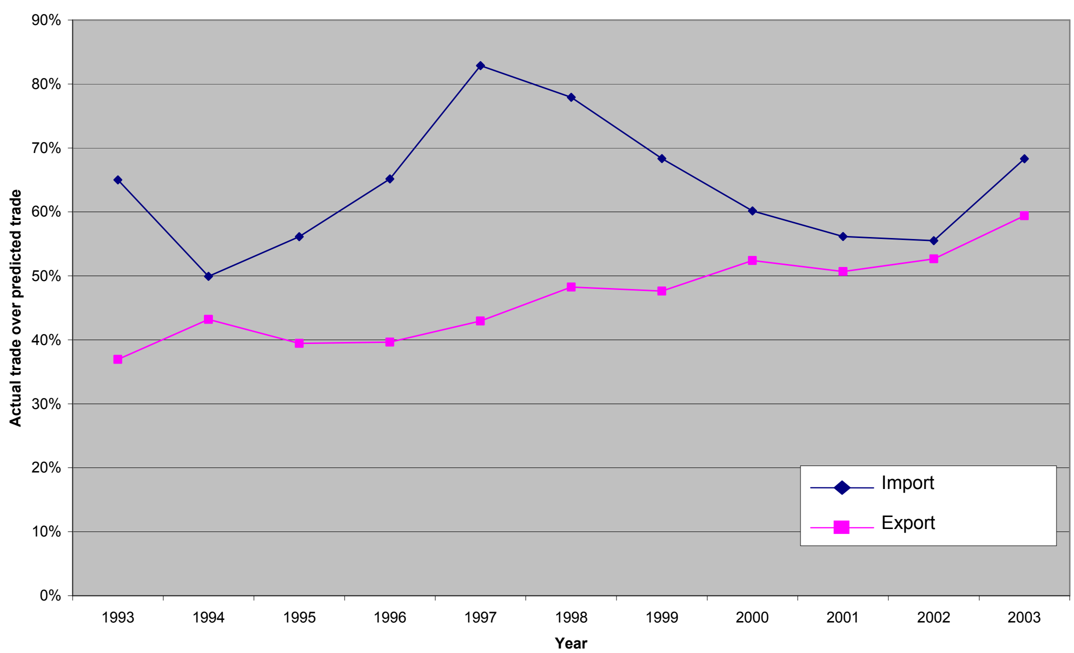
We have evaluated the trade flows using two statistically based methods. First, we calculated the predicted trade flows and compared the result with the actual trade flows of the same year to see whether the trade flows are converging or not. Second, we regressed new models based on recent trade flows to see whether the coefficients resemble the coefficients in the old studies.

In order to see whether or not the trends are going in the assumed direction we plugged in data on GDP, population and so forth, and then calculated each predicted trade flow between the years of 1993 to 2003 based on the coefficients from the models of Baldwin and Wang & Winters. The values received was then compared to the corresponding actual trade bilateral trade flows. In order to make the results easy to comprehend we have chosen to present them in graphs showing the values of actual trade between each bilateral trade flow in relation to its predicted trade flow. In this way, it may clearly be shown whether the model is over or underestimating the observed trade flows. The graphs from this analysis are shown on the next four pages (Figure 5 to 13). Lines with diamonds are representing imports and lines with squares are representing exports. Note that the scales are not the same in all graphs.



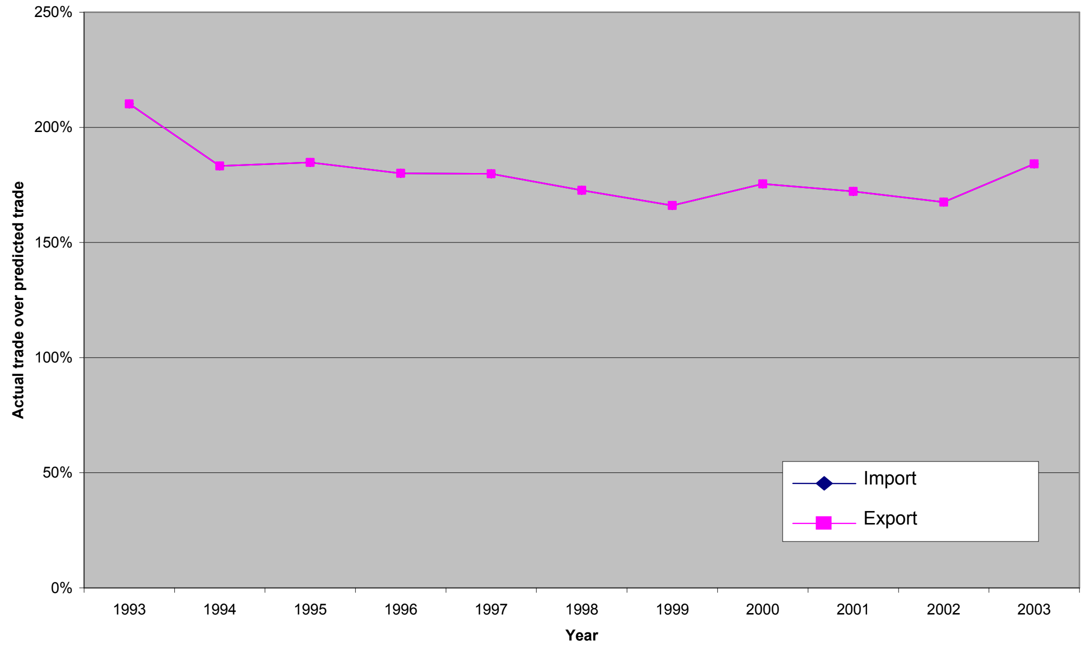
Actual over predicted trade. Own calculations based on Baldwin (1994). Diamonds represent imports and squares represent exports.

Figure 5 Trade between New EU and EU according to Baldwin



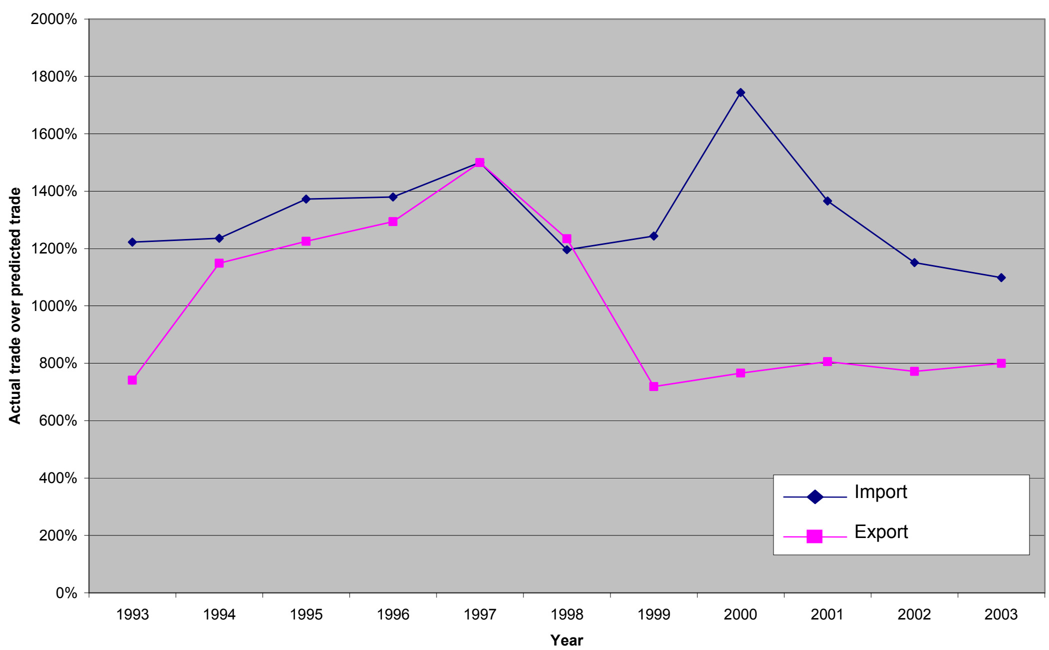
Actual over predicted trade. Own calculations based on Baldwin (1994). Diamonds represent imports and squares represent exports.

Figure 6 Trade between New EU and OECD according to Baldwin



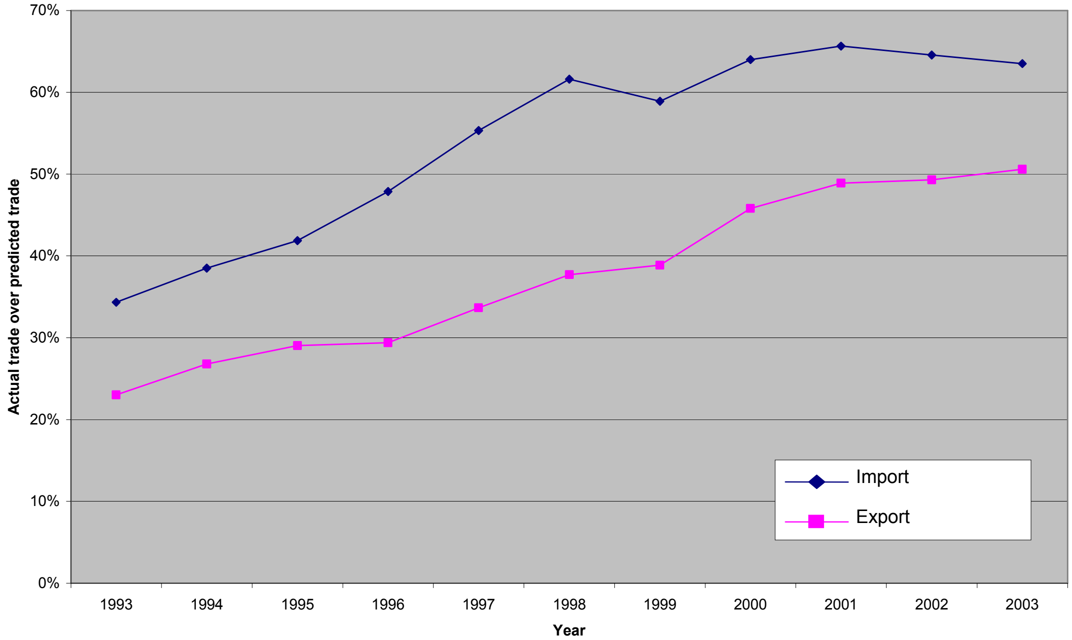
Actual over predicted trade. Own calculations based on Baldwin (1994). Diamonds represent import and squares represent exports.

Figure 7 Trade amongst the new EU members.



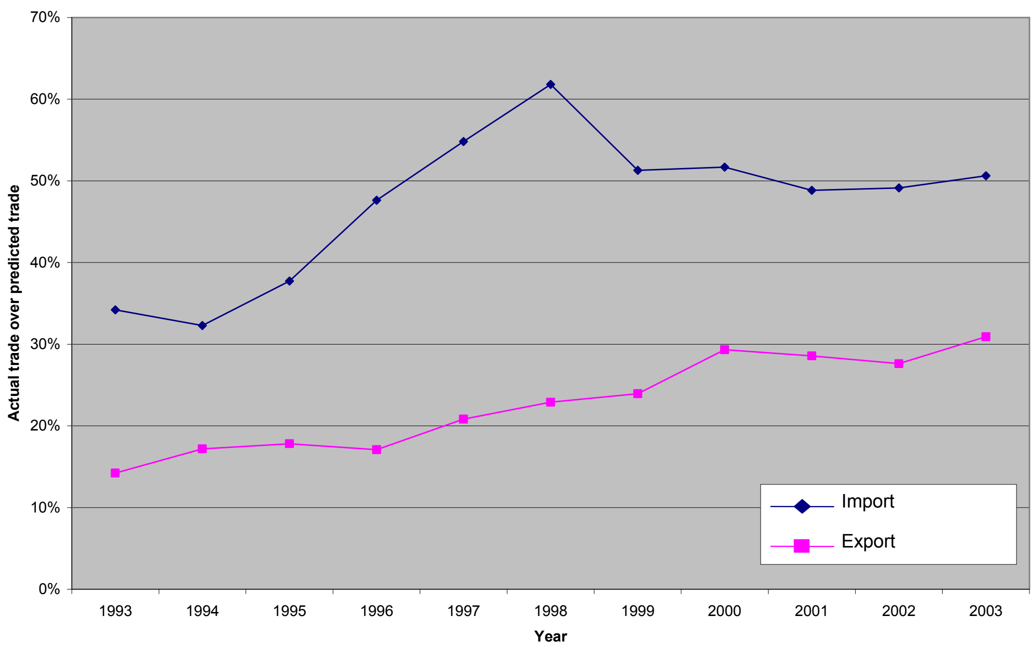
Actual over predicted trade. Own calculations based on Baldwin (1994). Diamonds represent import and squares represent exports.

Figure 8 Trade between EU and East Europe according to Baldwin.



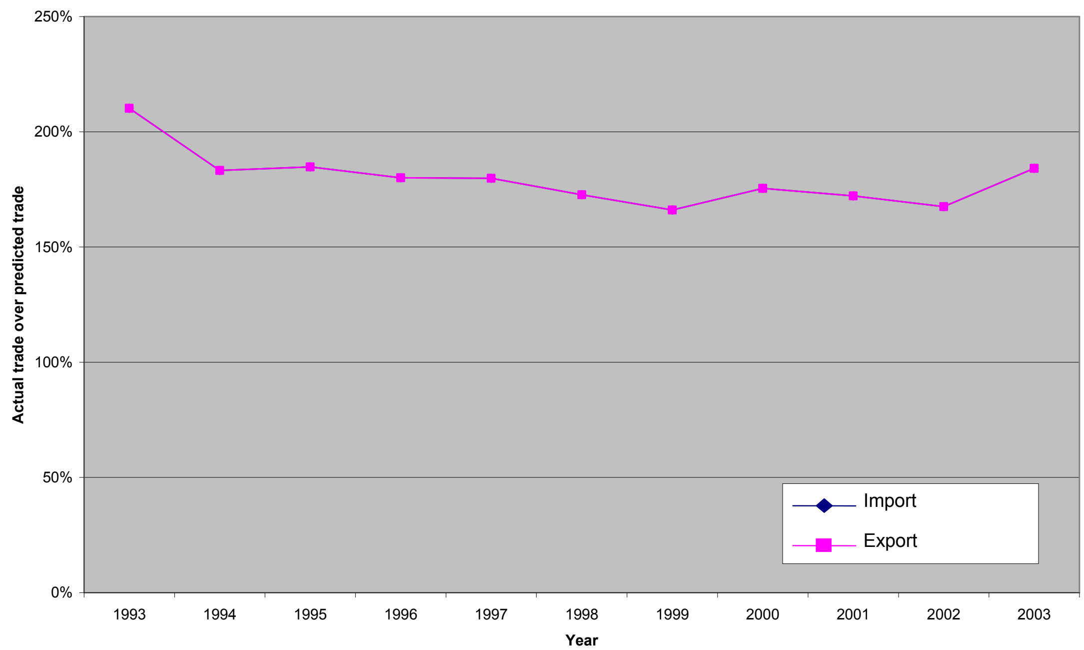
Actual over predicted trade. Own calculations based on Wang & Winters (1991). Diamonds represent import and squares represent exports.

Figure 9 Trade between New EU and EU according to Wang & Winters.



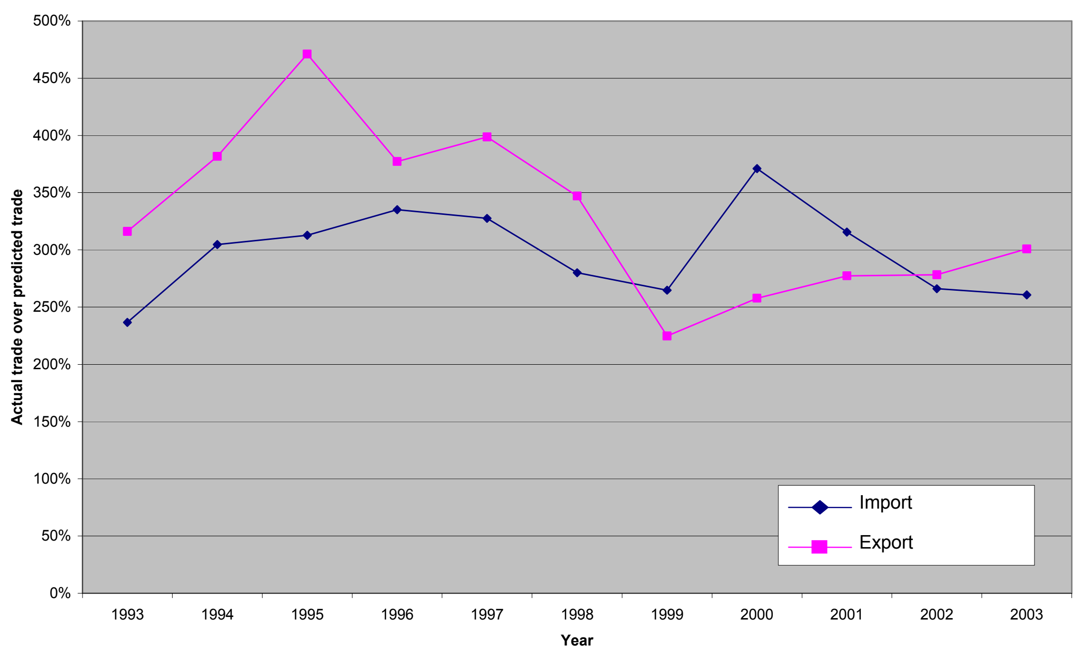
Actual over predicted trade. Own calculations based on Wang & Winters (1991). Diamonds represent imports and squares represent exports.

Figure 10 Trade between New EU and OECD according to Wang & Winters.



Actual over predicted trade. Own calculations based on Wang & Winters (1991). Diamonds represent import and squares represent exports.

Figure 11 Trade between amongst the New EU countries according to Wang & Winters.



Actual over predicted trade. Own calculations based on Wang & Winters (1991). Diamonds represent imports and squares represent exports.

Figure 12 Trade between EU and East Europe according to Wang & Winters.

If the models are making a good job in explaining the trade reallocation of the new EU members, we would expect trade to decrease among the new EU members themselves and to decrease trade with Eastern Europe. On the other hand, we expect a trend of increasing trade with the European Union and the OECD. As may be seen in the graphs, the gap between actual trade and potential trade is not closing significantly on the block level, and when the analysis is conducted on the country level, the results are even more unclear.

The models are systematic in the sense that they underestimate trade with Central and Eastern Europe and that they overestimate trade with EU and OECD. However, there are no clear trends visible in any of the cases, since the trends are not converging towards 100 (where actual trade and potential trade are equally large) in the graphs above. Moreover, for Eastern Europe, the model is making huge underestimations of the trade flows.

It is important to look at the trends when doing this analysis and not to look at the single point estimates. It is obvious that looking at the results from a single point in time may give results that are very specific to that particular year. For example, Gros and Gonciarz (1996) evaluate the results of Baldwin's study by looking at actual and predicted trade in 1992, concluding that the trade of Central and Eastern Europe does not differ considerably from the trade patterns of similar Western European countries. Baldwin (1997) replies that 1992 was the year when the Central and East European countries reached their lowest level of intra-region post-communist trade. Because reasons like this, we have instead decided to look at the trends of trade.

The method of looking at the trend instead of point estimates also yields support from Jakab et. al. (2001) who show that several studies using the gravity equation to estimate trade flows are having prediction intervals at 95% level for the point estimates of ±350%, which makes just about any conclusion about the level of potential trade possible. Instead of using the point estimates, Jakab et. al. define a measure of "the average speed of convergence" as the average growth rate of potential trade divided by the average growth rate of actual trade for the period of their analysis. This method is hypothesized to show convergence towards EU and OECD if the growth rate of potential trade is lower than that of actual trade, negative values, and divergence from the Central and East European countries, positive values. The method used by Jakab et. al. merely quantifies the information presented in the graphs above. When looking at the speed of convergence, presented in

Figures 13 and 14, it is clear that there is convergence towards EU and OECD, however, there is also a reported convergence towards the Central and East European countries, which we would not expect to see.

	New EU	EU	OECD	EE	Total
		Imports			
Czech Republic	447.8	-69.4	-58.2	602.5	-50.0
Estonia	-53.3	-60.1	-70.1	-26.3	-56.2
Hungary	-47.7	-54.8	-59.1	-97.2	-58.1
Latvia	-29.9	-67.6	-65.9	-323.5*	-36.2
Lithuania	-46.5	-66.5	-71.4	-35.1	-55.6
Poland	-53.7	-39.0	-16.7	-53.0	-38.4
Slovak Republic	215.2	-63.4	13.6	230.8	-7.8
Slovenia	-7.0	-5.6	-5.2	-39.2	-5.0
		Exports			
Czech Republic	128.4	-73.3	-67.6	-17.3	-60.7
Estonia	-13.5	-71.8	-76.4	27.6	-65.8
Hungary	-46.8	-64.4	-65.0	-31.3	-60.3
Latvia	75.8	-28.9	-74.1	-251.3*	-17.3
Lithuania	-7.6	-52.9	-76.9	-74.5	-47.9
Poland	-54.0	-45.3	-25.6	-59.5	-45.6
Slovak Republic	1661.1	-70.7	-80.7	-30.5	-43.2
Slovenia	-57.1	-14.6	-16.2	-59.2	-23.3

Speed = Average growth rate of potential /average growth rate of actual x 100 – 100. Own calculations based on the results from Wang & Winters' model. Measure is taken from Jakab et. al. (2001) * Means that the sign is not correct, see the text for explanation.

Figure 13 Average Speed of Convergence between 1993 and 2003 according to Wang & Winters.

	New EU	EU	OECD	EE	Total
			Imports		
Czech Republic	1123.3	-64.8	-64.8	-1084.3*	-44.5
Estonia	-44.8	-56.3	-71.4	-71.3	-54.3
Hungary	-29.4	-8.9	-46.3	0.5	-3.4
Latvia	-12.0	-53.9	21.1	-32.2	-33.5
Lithuania	-59.1	-66.6	-52.1	-47.1	-59.2
Poland	-38.5	-32.7	143.6	-54.4	-28.4
Slovak Republic	296.2	-59.7	-74.8	30.8	-12.3
Slovenia	-56.8	-13.0	-26.8	12.5	-16.9
			Exports		
Czech Republic	101.9	-70.2	-60.4	2274.0	-56.6
Estonia	11.9	-62.0	-62.9	39.6	-49.6
Hungary	-47.6	-67.5	-44.7	-89.8	-64.7
Latvia	83.5	-45.0	-73.2	-158.8*	35.5
Lithuania	-5.1	-48.5	-70.1	-91.3	-44.8
Poland	-57.6	-41.3	-21.6	-54.9	-42.9
Slovak Republic	-1223.1*	-68.6	-79.5	555.2	-22.7
Slovenia	-17.5	20.9	21.7	-38.1	13.1

Speed = Average growth rate of potential /average growth rate of actual x 100 – 100. Own calculations based on the results from Wang & Winters' model. Measure is taken from Jakab et. al. (2001) * Means that the sign is not correct, see the text for explanation.

Figure 14 Average Speed of Convergence between 1993 and 2003 according to Baldwin.

In some of the cases where we have negative values in Figure 13 and 14 above, this is caused by a negative development of actual trade rather than the fact that potential trade grows faster than actual trade. The measures where this phenomenon is occurring are marked with * in the tables above. These numbers are hence inconclusive.

What these figures tell us, is that the new EU members have not reached their potential levels of trade, with any region, but that trade is converging towards the potential levels. From these results, we feel confident when concluding that the models are not generating meaningful results in predicting future trade flows. Moreover, analyzing this in relation to the results presented in the graphs above, as we look at EU we see that trade is not simply converging towards the predicted of trade (100% in the graphs), but continues to grow after having reached the predicted level.

There may of course be explanations for why our countries deviate from the hypothesized path. For example, we know that the Czech and Slovak Republics, since their split in 1992 have reported very high levels of bilateral trade, because of the location of their industries. Explanations like this are also identified by for example Holwmann and Zukowska-Gagelmann (1998), who examine differences in both the scope and speed of economic reform, differences in transaction costs and differences in demand and supply structures of the former CMEA members. Other ways to address this problem are given by Hamilton and Winters (1992) who brake down trade on industry level. However, an underlying assumption when making models like this is to generate general conclusions about trade patterns. When we break down the results and the analysis on a specific level we loose this ability of drawing general conclusions. Like Lazear (2000) notes, the strength of economics is also its greatest weakness: it is rigorous and analytic. To make an analysis possible we are forced to narrow down the focus of the research. Narrowing down the study allow us to receive concrete answers, but prevent us from analysing the wider aspects of the problem.

In order to see whether Baldwin and Wang & Winters are alone in getting results like this, we have conducted the very same type of analysis on a study made by Frisell (1996). Frisell based his study on data from 14 industrialized countries (12 European countries, Japan and the United States), a total of 182 trade flows, all from 1992 and made predictions for trade flows in the Baltic area. In all areas, Frisell's model generates equally poor results as do Baldwin and Wang & Winters: trends are not moving in the hypothesized directions.

To sum up, we have shown that the gravity model does not provide the hypothesized results when predicting the development of trade between the new EU-members and the rest of the world. The results are rather indecisive. More interesting is, however, why we are getting these results.

Explaining the problems

One possible response to our statement that the model did not yield the results expected is that the new EU members are not yet transformed. This would, however, be a stark criticism. First, none of our countries is receiving lower grades than 3 from recent EBRD Transition Reports and both EU and WTO have welcomed them as being market economies. Second, even if the countries would not show the characteristics of a market economy, we would at least expect to see a convergence of potential trade towards actual trade, which we are not observing. Moreover, Brenton and Gros (1997) conducted a study about the definition of transition and when it could be said that the transition to a market economy is complete. They based their study on, at the time, recent economic data, focusing on international trade, finding that it is not possible to distinguish the more advanced countries in Central and Eastern Europe from any advanced market-oriented economy.

A lack of theory?

As we have already discussed in the background section, it is a problem that researchers working with the gravity equation are unable to theoretically derive the model used. As we have already mentioned, prices are excluded from the model, making it unfit for evaluating trade reorientation. The more countries included in the model, the more home production is exchanged for imports, rather than redirecting the present trade flows. For example, if we would include China in the country sample, the model would not predict countries to re-direct their trade to the cheap labor in China, rather it would assume that the countries should exchange home production with imports from China. If prices were included in the model, it might have allowed for trade re-orientation due to changes in relative prices, and a completely different picture could have emerged.

The model is too specific

Other possible explanation for the poor predictive power of the gravity model could be that the model is specific, either specific to the countries it was regressed from, or, specific over time and thereby useless for predicting the future. In order to test for whether the model is specific to the countries in the regression, we have calibrated the model based on the trade flows of the years between 1993 to 2003 and compared this with the results of Wang & Winters' predictions.

The regressions are computed, based on our sample data, but still exclude all trade flows to the Eastern Europe. We have expressed all data in the dollar value of 1985, in order to make the results directly comparable to the results of Wang & Winters. The reason for not doing the same with Baldwin's data is that he is not particularly clear on how he did his regression study as he corrected for reported serial correlation. We ran new regressions for the years of 1994, 1997, 2000 and 2003 in order to see whether the coefficients are moving towards the coefficients in the study from 1991.

Year	Constant	GDP per Capita Exporter	GDP per Capita Importer	Population Exporter	Population Importer	Distance	Adj	(Adj R^2)
1994	-8.60	0.92	0.89*	0.81*	0.77*	-1.03	1.35	(0.77)
1997	-8.26	0.87	0.75	1.03	0.78*	-1.02	1.18*	(0.78)
2000	-7.77	0.97*	0.81	0.85*	0.80*	-1.06	1.03*	(0.79)
2003	-6.45	0.85	0.75	0.87	0.80*	-1.02	1.02*	(0.80)
WW	-12.49	1.17	1.02	0.79	0.8	-0.75	0.78	(0.70)

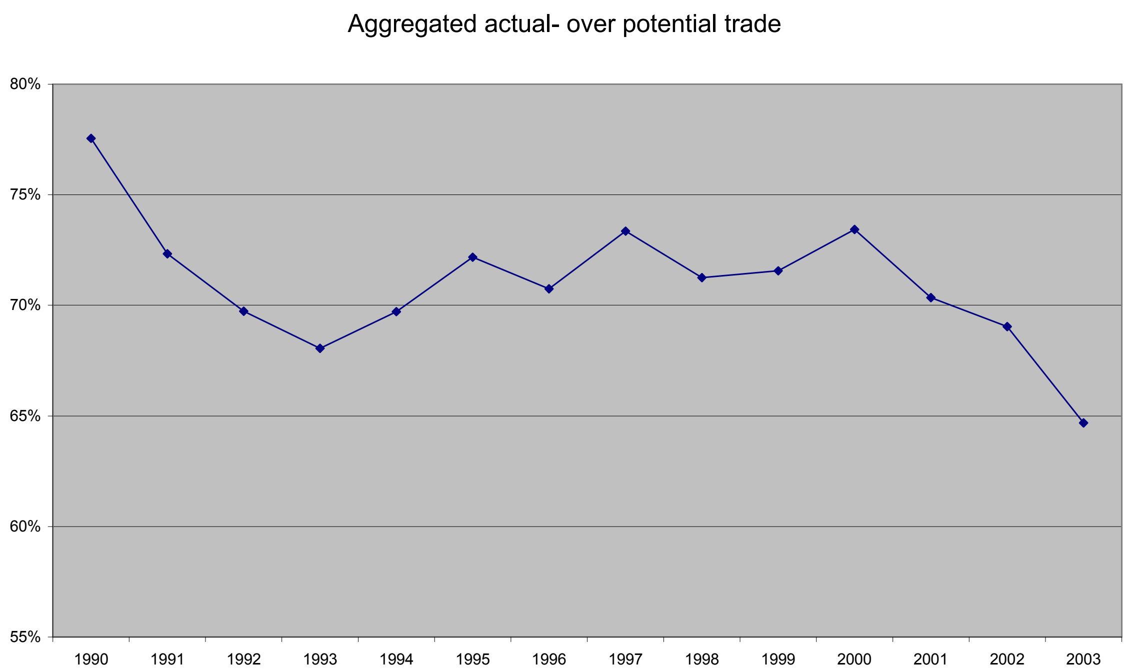
New coefficients for Wang & Winters' study based on our own calculations. * the coefficient cannot be distinguished from Wang & Winters' coefficient on a 95% confidence level.

Figure 15 New coefficients for Wang & Winters' study.

It soon becomes clear, however, that even though the coefficients from Figure 15 all are significant and show the expected signs they do not resemble the coefficients in the study conducted by Wang & Winters. When looking at a 95% confidence interval for the coefficients, we can see that some of our newly regressed variables may be the same variables as in Wang and Winters' study. These variables are marked with * in the table above. However, the corresponding confidence intervals for the Adjacency variable are ranging between 140 and 230%, and hence, just about any value of the coefficients would end up within this range. The confidence interval for the importer's population, on the other hand, is ending up around 25%, indicating that this coefficient might be stable over time. Overall, however, it is clear that our coefficients do not follow any trend of convergence towards the coefficients in the study of Wang & Winters. Therefore, we conclude that the model is not generating stable results for other sample countries than the ones used in the original regression. Note also that the model's goodness of fit increases with every new regressed model.

Is the gravity model also specific for a given time? If so, there may be several reasons for this. One is that trade costs may decrease over time with more efficient means of transportation, making the distance less influential in all sorts of spatial interaction (Mikkonen and Luoma 1999). This statement is very intuitive, but it should be noted that there is a current debate about whether trade costs really are decreasing with distance. It is clear that the costs for transporting an item one kilometre has fallen, but researchers have started to question whether trade costs as such are falling (Durantion and Storper, 2005). In our evaluation these new findings does not constitute any major changes to our analysis, since the distance variable serves as a proxy for transportation costs and not trade costs as such.

In order to test whether the results are stable over time we have taken data between the years of 1990 to 2003, from the same twenty countries Baldwin regressed his model with, and calculated the potential trade between these countries using Baldwin's coefficients and compared these results with the actual trade flows of the same years. The results of this evaluation, actual over predicted trade, are presented in the diagram below.



Own calculations of the trading potential of Baldwin's original sample countries. Presented as a fraction of actual over predicted trade.

Figure 16 Test of the time stability of Baldwin's model.

The results show that actual trade is, on average, developing slower than predicted trade during 1990 to 2003. Even when looking at the development of single countries we see a clear trend of an increased gap between the actual and the predicted trade, the further away from 1989, the last year in Baldwin's sample, we get.

One explanation to this phenomenon is that it is caused by the estimated coefficients for the GDP per capita of the trading countries. Since these coefficients are larger than 1, a small growth in GDP per capita gives an exponential effect on the predicted trade flow[3]. To illustrate this phenomenon, think of two trading countries, where each have a population of 5 million inhabitants and a distance of 150 km between the countries' capitals, and let their GDP grow. Ceteris paribus, their trade will grow exponentially with GDP according to the graph presented below.

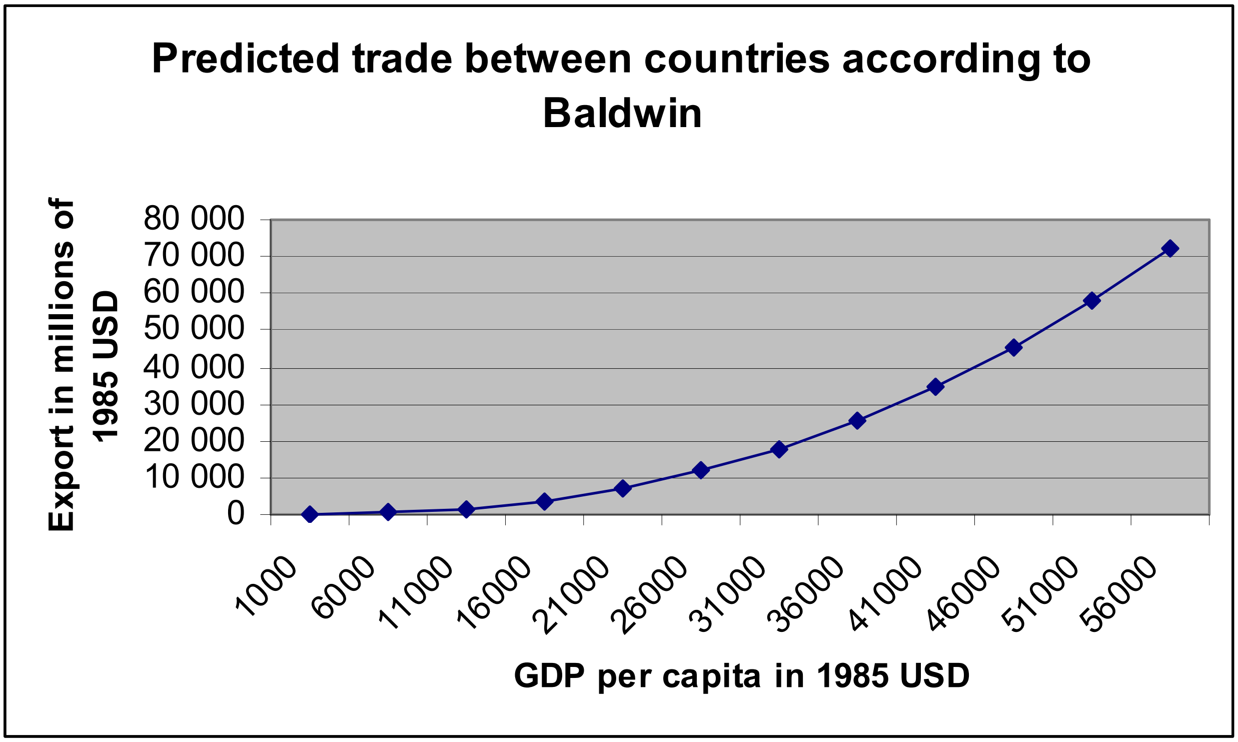


Figure 17 Illustration of the impact of exponential coefficients.

The actual trade flow may, however, not grow at the same speed. Aggregated GDP for the twenty countries in Baldwin's study grew from 12.800 billion to 17.200 US-dollars (1985 values). This in it self may explain the widening gap between the actual and the predicted trade from the model. This fact may also explain why both Baldwin and Wang & Winters get very poor results for trade between the new EU members and Eastern Europe, these

[3] Incidentally, a similar effect was identified in applications of the gravity model already in the 1960s (e.g. Claesson 1964 and 1969), suggesting that Helpman and Krugman (1985) instead are on the right track.

countries having relatively low GDP per capita compared to the countries for which the model was calibrated. From this we conclude that the gravity model as such cannot be used for making predictions of trade flows over periods where the included countries' GDP or GDP per capita has grown substantially.

Ad-hoc method

A third reason for the possible poor predictive ability of the model is that, in spite of the fact that the theory behind the model might be consistent with other trade theory and thoroughly developed over the years, the method for how the model is used is not. Much of the theoretical decisions facing researchers using the gravity equation are not discussed. Neither Baldwin nor Wang & Winters provide any discussion about the methods applied to their study, they are simply taken as given by prior researchers. In this sense, much of the standard for how the gravity model is used today was set already by Linemann in the 1960s. Already in 1970, Olsson conclude that a successful explanation is a prerequisite for successful predictions and that a successful explanation must follow from general law statements. Hence, he notes that "the best insurance against specification error is a good a priori theory" (Olsson, 1970:225).

Breuss and Egger (1999) were the first authors to question the use of cross-section data when regressing a gravity model. When evaluating old studies of the trading potential of Central and Eastern Europe they find that the forecast intervals are around 350%, making just about any conclusions about absolute trade levels extremely unsure. Earlier writers, like Baldwin and Wang & Winters in our studies, do not even report the standard deviation of their estimated regressions. This makes it impossible for others to evaluate the performance of their studies. Moreover, we can clearly question the statistical support of the projections. We have already pointed at this in the section above, and this is why we chose to present the model results with speed of convergence, as suggested by Jakab et. al (2001). Breuss and Egger (1999) hence conclude that cross-section gravity equations provides extremely uncertain results for understanding the trading potential of eastern Europe. However, they do not suggest any other more appropriate way of conducting analyses using the gravity model.

Mátyás (2001) has taken a closer look at the use of regional dummy variables to test the hypothesis that trading blocs have a significant role in explaining trade flows. He suggests that all gravity models used in this area are econometrically misspecified. Instead of the ordinary dummy variables, Mátyás is proposing that there is a need to analyze both the local and target specific parameters that are representing influence on the trade flow beyond the effects explained by GDP or distance and so forth. A large target specific effect indicates openness of the target country, and local country effects are employed to show how efficient one country is in exporting relative to its size and relative to the other countries in the sample. He concludes that if both these effects are large for most countries in a trading bloc then, and only then, may this be interpreted as a trading block effect. How the ordinary dummy variables may be interpreted remains unclear.

Does this force us to conclude that the gravity equation as such is useless for answering these questions? Maybe, if we are interested in trade re-orientation; thanks to the way the model is specified it is only designed to explain trade flows as such, using a logic saying that trade increases with increased GDP. Entering more countries into the model will result in more trade but does not say anything about trade-creation or trade-destruction.

Poor proxies

Along with this discussion of poor method comes the issue of a lacking discussion about the adequacy of the proxy variables used in the studies. There are, however, as many ways of specifying the model as there are authors of papers using the gravity model.

Remember that Linnemann (1966:34) stated that the trade flow from the exporter to the importer depends on:

(i) factors indicating total potential supply of the exporting country
(ii) factors indicating total potential demand of the importing country
(iii) factors representing the resistance to a trade flow between the two countries

This has commonly been specified in terms of gross national product, population size, a set of potential preferential-trade factors, and distance between the two countries. All authors we have come across use some form of GDP of the importer as well as exporter to capture the demand of the importer and supply of the exporter. In what form GDP should be used is however debated. Both Wang and Winters use real GDP per capita, but for instance Christie (2002) uses PPP adjusted GDP. Frisell (1996) however, chose to only include GDP in his study, because he claimed that there is generally a strong positive multicollinearity between population size and GDP. This is, however, questionable, just think about two of the largest countries in the world, China and India, in which this relationship is far from clear.

The most debated area is however how to measure trade resistance. The most commonly used way to do this is by including a distance variable and various numbers of dummy variables that are supposed to capture trade preferences such as common border, common market or speaking the same language. The distance variable in turn comes in many forms, air distances, road distances, distances between ports, distances between capitals, and distances in time. Christie (2002) also tried with using time of shipment instead of distance for the third factor. This did not yield any major differences in the results.

There is no theoretical foundation to help us determine what variables to use; they are all more or less arbitrary chosen. The choice of variables may however severely change the prediction power of the gravity model. There is also a trade off. More dummy variables can be included to improve the goodness of fit when the gravity equation is calibrated but that will lower the ability to make any generalizations of the regressed model. There is also a risk of dummy variables capturing a pattern that exists when the model is calibrated rather than something that is consistent over time. An example of this can for instance be to include a CIS dummy; such a dummy will have a huge effect on trade because of high trade flows between Russia, Ukraine and Belarus. This is of course due to their common Soviet heritage but there is no reason why this relationship would exist over time; on the other hand, size and proximity would suggest sizable trade under free trade circumstances, at least when specialisation have begun to yield complementarities at the expense of similarities in the structure of production created by the Soviet system. The distance variable is as mentioned a quite arbitrary proxy for measuring trade costs. Also, the value of the distance coefficient varies greatly between estimates of different authors. The distance variable as a proxy for trade costs might work for a given period of time but as mentioned earlier, decreasing costs of transportation should make the distance variable unstable over time.

To get an idea of how to specify the gravity model correctly one would need to run extensive robustness tests with all kind of variables to see which of them are consistent over time and between different groups of countries. It is a bit strange that no one has done this given the extensive use of the model as well as the fact that Linnemann first wrote about the gravity equation in the 1960s.

Conclusions

This thesis has aimed at ex-post evaluating the predictions made about trade with and within central and Eastern Europe after the fall of communism. Two different models have been chosen as the bases for the analysis, one written by Wand and Winters (1991) and one written by Baldwin (1994). Both studies use the gravity model framework in calibrating their models for certain samples of countries and years in order to predict the trade flows in central and Eastern Europe once the ex-communist countries had transformed into market economies.

First we concluded that the trade flows were not converging towards the flows predicted by Baldwin or Wang & Winters. New regressions suggested that the coefficients of the model are specific both for the time and countries for which it was being calibrated. However, more robustness tests would be needed in order to analyze if it is possible to calibrate coefficients that are consistent over time.

There are problems with the gravity model such as the general lack of theory underpinning it. This is in turn reinforced by the fact that prices have to be included to make the gravity model work in theory. However, this has never been done in reality. Indeed, the model is not designed to account for trade re-orientation, and this is a possible problem in the longer run, when trade policies and different relative prices makes it desirable for countries to move production and imports to other parts of the world. However, prices are not everything when deciding where to put production, there are also matters of strategic positioning from a logistic point of view that needs to be taken into account.

Finally there are several problems associated with the explanatory variables included in the gravity model, this is especially true for using the distance variable as a proxy of trade costs. Since the model used differs quite a lot from the theoretical model, their proxy variables are not explained, neither explicitly nor implicitly more research in the light of Mátyas (2000) is necessary before the model may confidently be used. Many small corrections in the data are likely to end up resulting in substantial mistakes not accounted for.

The gravity model is often used for explaining trade patterns and it seems as if it has a surprisingly large predictive power. Many economists have therefore used the gravity model when making predictions about future trade patterns. This thesis has, however, showed that these predictions are very uncertain and that using the gravity model for making predictions is associated with many theoretical and practical problems.

It is obviously not within the reach of our thesis to solve the problems surrounding the usage of the gravity model for predicting international trade flows. Rather, it is important to remember these shortcomings when designing policy instruments based on studies using the gravity equation. If these problems not are possible to solve, researchers might need to consider not using the model.

References

James E. Anderson (1979) "A theoretical foundation for the gravity equation", *The American Economic Review*, Vol. 69, No. 1, pp. 106-116.

Richard E. Baldwin (1997), "Comment on Groz and Gonciarz", *European Journal of Political Economy*, Vol 13:1, pp. 187-188.

Richard E. Baldwin (1993), *The Potential for Trade Between the Countries of EFTA and Central and Eastern Europe* EFTA Occasional Paper, 44. Genève: European Free Trade Association.

Richard E. Baldwin (1994), *Towards an Integrated Europe* London: CEPR.

Jeffery H. Bergstrand (1985) "The gravity equation in international trade: some microeconomic foundations and empirical evidence", *The Review of Economics and Statistics*, Vol. 67, Iss. 3, pp. 474-481.

Jeffery H. Bergstrand (1989) "The generalized gravity equation, monopolistic competition, and the factor-proportions theory in international trade", *The Review of Economics and Statistics*, Vol. 71, Iss. 1, pp. 143-153.

Peter A Brenton and Daniel Gros (1997), "Trade reorientation and recovery in transition economies", *Oxford Review of Economic Policy*, Vol 13:2, pp. 65-76.

Fritz Breuss and Peter Egger (1999), "How reliable are estimations of east-west trade potentials based on cross-section gravity analyses?", *Empirica*, Vol. 26, pp. 81-94.

Edward Christie (2002), *Potential Trade in Southeast Europe: a Gravity Model Approach*, *WIIW* Working Papers, No 21, March 2002.

EBDR (2003) *The transition report*, European Bank for Reconstruction and Development : 2003.

Lars Frisell (1996), "Trade in the Baltic States – the gravity model revisited"; Unpublished Master's Thesis Stockholm School of Economics.

Durantion and Storper (forthcoming), "Rising trade costs" Draft paper 2005

Daniel Gros and Andrzej Gonciarz (1996), "A note on the trade potential of Central and Eastern Europe", *European Journal of Political Economy*, Vol. 12:4, pp. 709-721.

Daniel Gros and Andrzej Gonciarz (1997), "Rejoinder", *European Journal of Political Economy*, Vol. 13:1, pp. 195-196.

R Scott Hacker and Henrik Einarsson, "The pattern, pull, and potential of Baltic Sea trade", *Annals of Regional Science*, Vol 37:1 (2003), pp. 15-29.

Carl B. Hamilton and Alan L. Winters (1993), "Opening up inter-national trade with Eastern Europe", *Economic Policy*, Vol. 7:1, pp. 78-116.

Keith Head (2003) "Gravity for beginners", http://pacific.commerce.ubc.ca/keith/. 2005-10-01.

E. Helpman and P.R. Krugman (1985), *Market Structure and Foreign Trade. Increasing Returns, Imperfect Competition, and the International Economy*, The MIT Press, Cambridge, MA/London.

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

Robert Holzmann and Katarzyna Zukowska-Gagelmann (1998), "Trade adjustment in Eastern Europe during transition: the determined, sophisticated and proximate", *Journal of Public Policy*, Vol. 18:1, pp. 29-52.

Zoltán M. Jakab, Mihály A. Kovács and András Oszlay (2001), "How far has trade integration advanced?: an analysis of the actual and potential trade of three Central and Eastern European countries", *Journal of Comparative Economics*, Vol. 29:2, pp. 276-292.

Edward P. Lazear (2000) "Economic imperialism", *Quarterly Journal of Economics*, Vol. 115, Iss. 1, pp. 99-146.

Edward E. Leamer and James Levinsohn (1995) "International trade theory: The evidence" pp. 1339-94 in G.M. Grossman and K. Rogoff (eds.), *Handbook of International Economics, Vol. III*, North Holland: Amsterdam.

Hans Linnemann, (1966) *An Econometric Study of International Trade Flows*, North-Holland Publishing Company, Amsterdam.

László Mátyás, "Proper econometric specification of the gravity model", *The World Economy*, Vol. 20:3 (1997), pp. 363-368.

John McCallum (1995) "National borders matter: Canada-U.S. regional trade patterns", *The American Economic Review*, Vol 85, Iss 3, pp. 615-623.

Kauko Mikkonen and Martti Luoma (1999) "The parameters of the gravity model are changing, how and why?" *Journal of transport geography*, Vol 7, pp. 277-283.

Gunnar Olsson (1965) "Distance and Human Interaction. A migration study", *Geografiska Annaler. Series B, Human Geography*, Vol. 47, No. 1, pp. 3-43.

Gunnar Olsson (1970) "Explanation, prediction and meaning variance: an assessment of distance interaction models", *Economic Geography*, Vol 46, Supplement: Procedings. International Geographical Union. Commission on Quantitative Methods, pp. 223-233

Robert I. Sutton and Barry M. Staw (1995) "What theory is not", *Administrative Science Quarterly*, Vol. 40, Iss. 3, pp. 371-384.

Zhen K. Wang and Alan L. Winters (1994), *Eastern Europe's International Trade* Manchester: Manchester University Press.

Zhen K. Wang and Alan L. Winters (1991), *The Trading Potential of Eastern Europe.* CEPR Discussion Paper, 610. London: CEPR.

Data sources:

Direction of trade statistics. (2004 and 1997) Yearbook/ prep. By the General Statistics Division of the Bureau of Statistics of the International Monetary Fund. Washington, D.C: IMF. (DOTSY)

Distance Between Capital Cities http://dss.ucsd.edu/~kgledits/capdist.html, accessed 2005-09-06. (DBCC)

European Council (1998) "Council Regulation (EC) No 2866/98 of 31 December 1998 on the conversion rates between the euro and the currencies of the Member States adopting the euro". *Official Journal* L 359, 31/12/1998 P, 0001-0002.

Great Circle distance between Capital Cities of the World at http://www.wcrl.ars.usda.gov/cec/java/capitals.htm, accessed 2005-09-06. (GCDW)

International Financial Statitics, www.imfstatistics.org, accessed 2005-09-26, (IFS)

World Development Statistics, http://devdata.worldbank.org/dataonline/ accessed 2005-09-14. (WDI).