universität wien

DISSERTATION

Titel der Dissertation

„Three Essays on Analyzing Bilateral Trade Flows“

Verfasser

Elif Nuroğlu

Angestrebter akademischer Grad

Doktorin
der Sozial- und Wirtschaftswissenschaften
(Dr. rer. soc. oec.)

Wien, im Mai 2010

Studienkennzahl lt. Studienblatt:	084 140
Dissertationsgebiet lt. Studienblatt:	Volkswirtschaft
Betreuer / Betreuerin:	Univ.-Prof. Dr. Robert M. Kunst

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor Prof. Dr. Robert M. Kunst for his generous support during the whole process of writing this thesis. He was very kind, patient and thorough in his explanations. Without his ideas, comments and support this thesis would not have come into existence.

I would like to thank my second supervisor Prof. Dr. Jesus Crespo-Cuaresma for his valuable comments and creative ideas which provided valuable improvements to all three essays.

Furthermore, I would like to thank

My dean Prof. Dr. Ertuğrul Taçgın at the International University of Sarajevo, for his support and ideas which combined engineering with economics,

My vice Dean Assist. Prof. Dr. Hasan Korkut for his encouragement,

Prof. Dr. Mehmet Can at the International University of Sarajevo for his comments about the first essay and for opening up his library in Sarajevo to me,

My husband Hüseyin Hayri Nuroğlu for the technical and emotional support he provided, for taking care of our little baby son when I had to study, and for making me believe that I would be able to complete the thesis one day,

My father İbrahim Oyuk, mother Safiye Oyuk and sister Ayşe Oyuk for their emotional support and encouragement,

My sister- and mother-in-law Ayşegül and Sema Nuroğlu for their emotional support and my father-in-law Abdurrahman Nuroğlu for his encouragement,

Feyza Çetin, for her encouragement during the completion of the thesis and for taking care of my little baby,

My son A. Mirza Nuroğlu for being a sweet and calm baby and providing me with the motivation to finish my thesis, Mr. Yusuf Ziyaeddin Sula and Mr. Yusuf Kara for their encouragement and support, Ms. Reyyan Bal for her proofreading of all the essays and her kind friendship and support, and Ms. Elizabeth Raab for her proofreading,

All my friends at the International University of Sarajevo,

My friend Iwona Jaszczyszyn for her emotional support and to her husband Wilhelm Hummer for his proofreading of the German abstract.

TABLE OF CONTENTS

INTRODUCTION 1

ESSAY 1: THE IMPACT OF EXCHANGE RATE VOLATILITY ON INTERNATIONAL TRADE USING A MODIFIED GRAVITY MODEL AND A FUZZY APPROACH 7

1- Introduction 7

2- A Modified Gravity Model of Total Trade 11

3- A Fuzzy Approach to Total Trade 14

4- Conclusion 23

5- Appendices 23

7- References 31

ESSAY 2: DIFFERENT SPECIFICATIONS OF THE GRAVITY EQUATION: A THREE-WAY MODEL OR BILATERAL INTERACTION EFFECTS? 35

1. Introduction 35

2. A short survey 37

3. The Data 39

4. A Modified Gravity Model of Bilateral Trade 39

5. Results 41

6. The Forecasting Performance of Main Effects Model and Bilateral Effects Model 44

7. Conclusion 45

8. References 46

ESSAY 3: ESTIMATING BILATERAL TRADE FLOWS AMONG EU15 COUNTRIES USING PANEL DATA ANALYSIS AND NEURAL NETWORKS 51

1- Introduction 51

2- Studies on Neural Networks and Traditional Models 52

3- Neural Network Modeling 56

4- Data 58

5- A Modified Gravity Model of Bilateral Exports and its Application to Panel Data Analysis 59

6- A Modified Gravity Model of Bilateral Exports and its Application to Neural Networks 61

7- A Comparison between Panel Data Analysis and Neural Network Model 63

8- Conclusion 77

9. Appendices 78

10- References 83

CONCLUSIONS 90

ZUSAMMENFASSUNG 93

LIST OF TABLES 96

LIST OF FIGURES 97

LEBENSLAUF 98

INTRODUCTION

This thesis consists of three essays which estimate bilateral trade flows among European countries from 1962 to 2003 by using a modified gravity model of total trade. The basic gravity model explains bilateral trade between two countries with the product of their income and distances between them. According to this model, as the countries become richer, they are expected to trade more. Moreover, if they are further away from other trading partners, they tend to trade less because larger distances impose additional costs to trade such as higher transportation costs, cultural barriers, etc.

In this thesis, the gravity model was extended with the population of exporting and importing countries and bilateral exchange rates between their currencies. Population is included due to our belief that higher population does affect how much people consume, how much they produce, buy (import) and sell (export) as a result. The regression results have proven that population has a significant effect on bilateral trade flows. Furthermore, real exchange rates are also expected to influence bilateral trade flows because they have a direct impact on how much an exporter and importer earns. During periods when there is much exchange rate fluctuation, traders' earnings show some fluctuations as well, and therefore volatility in exchange rates is, in turn, supposed to decrease the amount of bilateral trade.

Although the modified gravity model used in this thesis consists of real GDPs, population numbers and volatility of real exchange rates, in the second essay we include main and bilateral interaction effects and time effects. The main effects model consists of importing- country, exporting-country and time effects. Exporting-country and importing-country effects are used to capture any tendency to export and import while business cycle (time) effects are expected to control for cyclical changes and their effects on bilateral exports. On the other hand, the bilateral interaction effects model uses country-pair or bilateral effects to see any geographical, political, historical or cultural event that could affect bilateral trade between two countries. We refer to the main effects model as the 3-way model and to the bilateral interaction effects model as the 2-way model. According to the results obtained, both the 3-way and the 2-way model explain a larger variation in bilateral trade as compared to the modified gravity model used in this study.

In the literature, the gravity model is used in different versions by inserting more explanatory variables such as common borders, a common language, being in the same trade union, or having free trade agreements. Some of these studies do not take the income of each country separately into the equation, but the product of both countries` incomes. They apply the same to the population as well. In our modified gravity model of total trade, however, we insert income and also population of exporting and importing countries individually into the model. The reasoning behind this is that if both countries` incomes have the same effect on their bilateral trade, then the coefficients of exporter and importer country`s GDPs should be equal. However, the results indicate that exporter and importer country`s GDPs influence bilateral trade by different amounts and this strengthens our expectations.

What is the motivation?

This thesis has two major aims. One is to see the effects of exchange rate volatility on bilateral trade flows. In the international trade literature, there is no agreement as to the effects of exchange rate volatility on trade flows. Some studies suggest a significant negative effect, while others show no association between exchange rate volatility and total trade. When starting this study, our expectation was that exchange rate volatility has a negative impact on bilateral trade. It was supposed that fluctuations in exchange rates make exporters and importers` earnings less predictable, and may lead them to behave more cautiously. As a result, in a non-stable environment they tend to trade less in order not to take more risks.

Our second aim is to compare panel data analysis with different techniques which have become popular in the last decades. Although we think that panel data analysis or statistical methods give the most reliable results and are proven to be one of the best means for analyzing bilateral trade flows, we use fuzzy logic and neural networks as alternative methods to analyze bilateral trade flows and to compare their performance with panel data analysis.

In the first article, after running a regression for the panel data and obtaining significant results, we construct a fuzzy rule and a fuzzy decision table to calculate the effect of exchange rate volatility on bilateral trade. Fuzzy logic has been proposed by Lotfi Zadeh in 1965 as an alternative to probability theory; however, it has been somewhat exploited since its invention. When it was first introduced, there was some debate about using it because until that time problems had been mostly solved by using crisp models. This method encountered resistance especially in the mathematically oriented western world which is based on the Aristotelian either-or approach. However, most of the expressions in the real world cannot be categorized sharply in one or another set. For this kind of vague phenomena, the fuzzy approach may provide better solutions obtained when using crisp models. When employing fuzzy logic in our study to analyze bilateral trade flows, we use the expertise and knowledge about bilateral trade flows which we acquire from statistical methods. Indeed, when there is a sufficient and reliable data set, using panel data analysis may give satisfactory enough results. Our objective in using fuzzy logic is to make a robustness check to see whether fuzzy logic can give similar results to the panel data analysis. Our study proves that when the user has expertise in the topic and can construct a reliable fuzzy decision table and a fuzzy rule set, it is possible to get some approximate results by using fuzzy logic without requiring a data set. As is well known, panel data analysis requires a large data set to give reliable results. However, when there is any problem in obtaining or processing the data, this may impede acquiring good and reliable results. In such cases where there are problems with the data set, we should have an alternative to get first approximate results. For that reason, we made the robustness check for the fuzzy approach, even though we have a very large data set, to see whether we can use it in other cases where there is an insufficient data set. We believe that our results on the fuzzy logic are promising.

In the second article only statistical methods are used. Our modified gravity model of total trade was extended by using exporting and importing country as well as business cycle effects. This model is called the 3-way or main effects model. Another extension is made with bilateral interaction effects and time effects and is called the 2-way model. We first compared the explanatory power of each model and then did a forecasting exercise. The bilateral interaction effects model turns out to be superior to the main effects model and also to the modified gravity model used in this thesis, in estimating and also in forecasting bilateral trade flows. One important characteristic of the bilateral interaction effects model is that it enables us to capture any geographical, cultural, historical or political event which may have an influence on bilateral trade between two countries. However, using bilateral interaction effects hinder us from seeing the individual contribution of some variables in explaining the variance in the dependent variable, such as distance, common languages, common borders, and others.

In the third article, we compare panel data analysis with a neural network model. In panel data analysis, we estimate bilateral exports with the explanatory variables that are income, population, distance and volatility of exchange rates. Then, we construct a neural network where the inputs are income, distance, population and volatility of exchange rates and the output is bilateral exports. Here, the inputs of the neural network are the explanatory variables of the panel model, and the output corresponds to the dependent variable of the panel model.

As a first step, we use the neural network to learn the relationship between the inputs and the outputs from the introduced examples that are called the training set. Once the network has been trained with the training set, it is asked to predict the outputs in the validation set, by using the weights and biases which are determined in the training phase. When the error (MSE between actual outputs and the outputs produced by the neural network) in the validation set starts increasing, the neural network stops training. The weights and biases are chosen at the point where the minimum validation set error is reached. Weights and biases are determined using the back-propagation method, which is one of the differentiating characteristics between neural network models and traditional statistical procedures. The main difference between back-propagation and traditional statistical methods is that the back-propagation algorithm can consecutively consider data records, and readjusts the parameters after each observation in a gradient search manner. By contrast, traditional methods such as maximum likelihood and least squares use an aggregated error across the whole sample in the estimation.

The last step is to measure whether the network has learned the relationship between the inputs and the outputs very well. To do this, the network is asked to produce its own outputs for the given inputs of the test set. If the network has learned the relationship appropriately it should give a low MSE. However, the MSE is not the sole performance measure. Another option for measuring the performance of a neural network is to perform a regression analysis between neural network outputs and the corresponding targets (actual outputs). When the slope of the best linear regression and the correlation coefficient between network outputs and targets are obtained after simulating the network, it is seen that the results are quite satisfactory and the neural network model can explain 97% of the variation in bilateral exports.

Another contribution of the third article is the comparison of the forecasting performance of the panel model and the neural network model. In performing this analysis, the panel model is asked to estimate bilateral exports from 1964 to 1993 and then to forecast bilateral export flows for the years from 1994 to 2003. Then, a neural network is constructed by using the data set from 1964 to 1993 and then the data set from 1994 to 2003 is used to test this neural network. Lastly, the MSE produced by each model are compared. Whereas both models give very close and satisfactory R-squared values, which reflects their in-sample explanatory power, the neural network model appears to be superior in out-of-sample forecasting.

Even though we use three different methods in the entire thesis, the basic model we employ is the extended gravity model. Of course, it is not surprising that different methods offer results which differ in quantity. However, the similarity between the results obtained from the three methods and the comparison with the existing literature indicates that the adjusted gravity model, which serves as an umbrella in the whole thesis, is very successful in explaining bilateral trade flows.

THE IMPACT OF EXCHANGE RATE VOLATILITY ON INTERNATIONAL TRADE USING A MODIFIED GRAVITY MODEL AND A FUZZY APPROACH

Elif Nuroğlu[+], Robert Kunst[++]

(+)International University of Sarajevo

(++)University of Vienna

Abstract: In this paper through the use of a gravity model and cross sectional data for 91 pairs of EU15 countries, a negative effect of exchange rate volatility on bilateral trade across European countries is found for the years from 1964 to 2003. Results illustrating the effects of exchange rate volatility on bilateral trade are obtained both by a modified gravity model and by a fuzzy approach. A remarkable match is observed between the results of these two approaches.

Keywords: *exchange rates, bilateral trade, volatility, gravity model, fuzzy.*

1- Introduction

This study investigates the effects of exchange rate volatility on bilateral trade flows across European countries from 1964 to 2003 by employing panel data analysis and a fuzzy approach. The former approach uses a classical gravity model that is extended by including population of exporting and importing countries and exchange rate volatility as explanatory variables of bilateral trade flows. This model is estimated for a data set of EU15 countries from 1964 to 2003. The second, alternative approach uses fuzzy modeling. The effect of exchange rate volatility on bilateral trade flows is then compared across the two models.

International trade history shows that different exchange rate regimes were preferred at different periods. In the last decades there is a tendency towards purely fixed or purely floating exchange rate regimes. A survey (Fischer, 2001) indicates that most countries have abandoned intermediate exchange rate regimes and instead prefer a purely floating or a purely fixed exchange rate. The percentage of fixed exchange rate regimes increased from 16% in 1991 to 24% in 1999 while percentage of floating exchange rate regimes increased from 23% to 42% during the same period. On the other hand, the number of intermediate regimes declined from 62% in 1991 to 34% in 1999. According to Fischer (2001), this movement from the intermediate regimes is towards currency boards, dollarization or currency unions on the hard peg side, and towards a variety of floating exchange rate regimes on the other side. The main reason suggested for this change is that "soft pegs are crisis-prone and not viable over long periods". Moreover, Bubula and Otker-Robe (2003) provide some support for the proponents of the bipolar view. They find that during 1990–2001, intermediate regimes were more frequently subject to crises as compared with purely fixed and floating ones, while even the latter have not been totally free of pressures.

The choice of exchange rate regime gives a country the freedom to use macroeconomic policies to manipulate the economy and enables it to fight recessions, crises etc. Furthermore, exchange rates influence the level of international trade as well. Therefore, the effects of volatility in exchange rates and of exchange rate regimes on the economy and on international trade have long been studied. There are two sides in the literature. One side claims that exchange rate uncertainty/volatility/variability does not have any impact on trade while there is another side which tries to prove the opposite. Hooper and

Kohlhagen (1978) analyze the impact of exchange rate uncertainty on the volume of the US – German trade between 1965 and 1975 and conclude that there is no statistically significant effect. Gotur (1985) reaches the same conclusion by analyzing the effects of exchange rate volatility on the volume of trade among the US, Germany, France, Japan and the UK. A famous IMF study (1984) summarizes that the large majority of empirical studies could not find a significant relationship between exchange rate variability and the volume of trade either on aggregated or bilateral basis. More recently, this view was supported by Bacchetta and van Wincoop (2000) who find that exchange rate uncertainty, or different exchange rate systems do not have any impact on trade.

On the other hand, Ethier (1983) analyzes the effects of exchange rate uncertainty on the level of trade and finds that uncertainty in future exchange rates reduces level of trade. Cushman (1983) estimates fourteen bilateral trade flows among industrialized countries and finds a significant negative effect of exchange risk on trade. Akhtar and Hilton (1984) establish a significant negative effect of nominal exchange rate uncertainty on bilateral trade between Germany and the US. Kenen and Rodrik (1986) analyze the effects of volatility in real exchange rates on the volume of trade and conclude that volatility depresses volume of trade. De Grauwe and De Bellefroid (1986) employ cross sectional techniques for the European Economic Community countries for 1960-1969 and 1973-1984, and investigate the effects of variability in real exchange rates on trade. They find significant negative effects. Lanea and Milesi-Ferretti (2002) examine the effects of appreciation and depreciation of exchange rates on trade and conclude that in the long run, larger trade surpluses are to be expected with more depreciated real exchange rates. Viane and de Vries (1992) study this issue from a different perspective, by analyzing the effects of exchange rate volatility on exports and imports separately and find that exporters and importers are affected differently by the changes in exchange rates, because they are on opposite sides of the forward market.

In the last decade, artificial intelligence methods such as neural networks and fuzzy logic have been employed in econometric studies especially in time series analysis. Tseng et al. (2001) propose a fuzzy model and apply it to forecast foreign exchange rates. Lee and Wong (2007) use an artificial neural network and fuzzy reasoning to improve the decision making under foreign currency risk and analyze the effect of trading strategy on the changes in exchange rates. They use fuzzy logic because they claim that it is capable to perform text reasoning of macroeconomic news.

Employing statistical methods together with artificial intelligence methods in one study gives the user the opportunity to compare both results. In econometric analysis, a large data set and a strong model is needed to obtain reliable results. However, there may be some cases in which it is difficult to obtain a large data set sufficient to get reliable results, or there may be some missing data which affects the reliability of results. In these cases, combining a fuzzy approach with the expertise in the topic studied can be a good solution to get the first approximate results. For this reason, our study compares the results obtained by panel data analysis and by using fuzzy logic. While we have a large data set and thus it can be argued that fuzzy logic is not really necessary, we propose to use this approach as a robustness check on traditional modeling. Once fuzzy logic proves to be a good alternative, it can also be used in cases of data problems that impair the validity of traditional methods.

The structure of the paper is as follows. Section 2 introduces the modified gravity model of total trade used in this study and reports the regression results obtained. Section 3 shows how the fuzzy approach is used to explain the effects of exchange rate volatility on bilateral trade. Section 4 concludes. Appendices 1 and 2 provide additional information about the fuzzy approach and fuzzy mathematics used.

2- A Modified Gravity Model of Total Trade

According to the Gravity Model, trade flows between two countries depend on their income positively and on the distances between them negatively as shown in Equation 1

$$T_{ij} = C \times \frac{GDP_i \times GDP_j}{D_{ij}} \qquad \text{(Eq. 1)}$$

where c is a constant term, T_{ij} is the value of trade between country i and country j, GDP_i is country i`s income, $GDP_j$ is the country $j$`s income and D_{ij} is the distance between two countries (Krugman and Obstfeld, 2006).

The gravity model says that large economies are expected to spend more on imports and exports; so, the higher the GDP of a country, the higher its total trade. The gravity model can be extended to catch other effects such as population, exchange rates, having a common language and common border or being in the same trade union that promotes bilateral trade.

In this study, the gravity model is extended with additional variables, namely the population of exporting and importing country and exchange rate volatility. Another difference from the original model is that incomes of country i and j are not taken as products with the same coefficient but as separate variables. The same approach applies to the population, where we have different coefficients for each country. The proposed model that is used to capture the effects of exchange rate volatility on bilateral trade is:

$$\begin{aligned} \ln T_{ijt} &= \alpha + \beta_1 \ln D_{ij} + \beta_2 \ln Y_{it} + \beta_3 \ln Y_{jt} + \beta_4 \ln Pop_{it} \qquad \text{(Eq. 2)} \\ &+ \beta_5 \ln Pop_{jt} + \beta_6 VolXR_{ijt} + \varepsilon_{ijt} \end{aligned}$$

where T_{ijt} represents total bilateral trade between country *i* and country *j* during time *t* which is calculated as the sum of exports from country *i* to country *j* and imports from country *j* to country *i*. Exports and imports are measured in nominal terms and then are converted to the volumes by using GDP deflators for each country at time *t*. D_{ij} is the distance between capital cities of country *i* and country *j* that is measured in kilometers. Two basic variables of gravity model are Y_{it} and Y_{jt}, real GDP of country *i* and *j* respectively. Pop_{it} and Pop_{jt} are the populations of country *i* and country *j* in time *t*.

$VolXR_{ijt}$ is the volatility of nominal exchange rate between exporter and importer country in year *t* which is calculated as the moving average of standard deviations of the first difference of logarithms of quarterly nominal bilateral exchange rates (Kowalski, 2006).

Results of the Modified Gravity Model

The sample period covers 40 years from 1964 to 2003. Countries included are the EU-15, where Belgium and Luxembourg are taken as one country because of data availability. The sources for the data are World Bank`s World Development Indicators 2005, OECD`s International Trade by Commodity Statistics and IMF`s International Financial Statistics.

The model is estimated using bilateral trade flows across the EU-15 countries from 1964 to 2003. From the data set of 14 countries, 91 bilateral trade flows are obtained during fixed, flexible and Euro periods. Equation 2 is estimated by using bilateral trade volumes and results are shown in Table 1-1.

According to the gravity theory, the income of a country is expected to affect its trade in a positive way. Table 1 - 1 shows that both income terms for country *i* and *j* have the expected positive sign. The difference from previous studies is emphasized by the discrepancy in the two coefficients. The contributions by the income terms of each country to the bilateral trade are quite different. We find that a 1 percent increase in the income of country *i* (exporting country) leads to a 0.09% higher bilateral trade. On the other hand, a 1% increase in the income of country *j* (importing country) results in a 1.1% increase in bilateral trade.

	Coefficient	t-Statistic	Prob.
C	-14.57	-20.38	0.00
Distance	-0.85	-41.18	0.00
Exporter GDP	0.09	2.30	0.02
Importer GDP	1.10	35.64	0.00
Exporter Population	0.67	16.50	0.00
Importer Population	-0.41	-12.11	0.00
Exchange Rate Volatility	-0.21	-3.29	0.00
R-squared	0.85		
Adjusted R-squared	0.84		
AIC	2.04		
Schwarz criterion	2.12		
Hannan-Quinn criter.	2.07		
Number of observations	3601		
Sample (adjusted)	1964-2003		
Periods included	40		
Cross-sections included	91		

Table 1 - 1: Balanced panel estimates with period fixed effects, dependent variable: log of total bilateral trade

Moreover, population has a negative sign for the importing country. The intuition behind is that higher population is assumed to decrease income per capita which may lower the need for imports and also the level of exports. On the other hand, population of the exporting country has a positive effect on bilateral trade suggesting that the higher the population the higher the production and exports as a consequence. Additionally, higher population may increase the need for the imported goods as well.

One of the basic elements of the gravity model is the distance between countries, which is on the denominator of the gravity equation (Equation 1). Since it is on the denominator, it should have a negative sign with the assumption that higher distances tend to decrease international trade by increasing transportation costs and imposing other impediments to trading such as informational and psychological frictions (Huang: 2007). Last but not least, exchange rate volatility has a negative effect on bilateral trade. As the results indicate, when volatility increases by 1%, the percentage change in bilateral trade is -0.21%.

3- A Fuzzy Approach to Total Trade

As shown in section 2, exchange rate volatility leads to fluctuations in the volume of trade. The main objective of this study is to compare the results obtained by panel data analysis with the ones obtained by using fuzzy logic to see how close they are. If fuzzy rules are set appropriately, and if the intuition used is realistic and in accordance with theory, fuzzy reasoning may give very similar results to panel data analysis without requiring a very large data set which is necessary in panel data analysis. For econometric methods, the data set is crucially important. When there is any problem in obtaining or processing the data or in the specification of the model, it is impossible to get reliable results. Moreover, if no sufficient data is available, conventional models cannot give reliable results. For these cases, fuzzy reasoning can be suggested as an alternative to get some approximate results.

In this section, the effects of exchange rate volatility on bilateral trade will be analyzed using fuzzy reasoning. The theory of fuzzy sets has been applied first to engineering fields and then spread to a wide range of areas such as economics, management, artificial intelligence, psychology, linguistics, information retrieval, medicine etc. (Fu and Yao, 1980).

Steps to be taken to apply a fuzzy approach to total trade are:

(i) setting the fuzzy decision table;

(ii) determining the change in total trade following a 1 percent increase in exchange rate volatility.

To start with, it is needed to fuzzify exchange rate volatility and the decrease in bilateral trade. Describing process states by means of linguistic variables and using these variables as inputs is a very important step in fuzzy approach. Table 1-2 shows the partitioning of the universe of exchange rate volatility into three fuzzy sets which are high, medium and low. Table 1-3 shows an analogous partitioning of the universe of total trade. When defining these expressions, membership values are assigned to each state intuitively based on experience (McNeil and Thro, 1994). A fuzzy set is defined solely by its membership function (Zimmermann, 2001). Membership degrees lie between 0 and 1, meaning that if an object completely belongs to the fuzzy set it has a membership value of 1. If an object does not belong to the fuzzy set at all, it has a membership value of 0. Membership degrees of borderline cases lie between 0 and 1. The more an element is characteristic of a fuzzy set, the closer to 1 is its membership degree (Driankov et al., 1996).

According to Table 1 - 2, "high increase in exchange rate volatility" is meant to be a 1% increase in volatility. If the increase is 0.9%, this volatility is considered to be high with a membership value of 0.75. When the volatility increases by 0.8%, the membership value for a high volatility decreases to 0.5. $\tilde{A}_1$ in Table 1-2 is the fuzzy set which describes a high increase in exchange rate volatility. Furthermore, a 0.5% increase in exchange rate volatility is defined as being medium and therefore is assigned a membership value of 1 in $\tilde{A}_2$, which is a fuzzy set that describes a medium increase in exchange rate volatility. Similarly,

$\tilde{A}_3$ in Table 1-2 represents the fuzzy set which describes a low increase in exchange rate volatility. These three fuzzy sets are described with the membership functions as Table 1-2 shows. In fuzzy language, “high”, “medium” and “low” (increase in exchange rate volatility) are called linguistic values. “A” in general is the linguistic variable that represents “exchange rate volatility”.

	0%	**0.1%**	**0.2%**	**0.3%**	**0.4%**	**0.5%**	**0.6%**	**0.7%**	**0.8%**	**0.9%**	**1%**	
high	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.50	0.75	1	$\tilde{A}_1$
medium	0.00	0.00	0.25	0.50	0.75	1.00	0.75	0.50	0.25	0.00	0	$\tilde{A}_2$
low	0.50	0.75	1.00	0.75	0.50	0.25	0.00	0.00	0.00	0.00	0	$\tilde{A}_3$

Table 1 - 2: Increase in Exchange Rate Volatility Partitioning

	0.00%	**0.025%**	**0.05%**	**0.075%**	**0.10%**	**0.125%**	**0.15%**	**0.175%**	**0.20%**	**0.225%**	**0.25%**	
medium	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.50	0.75	1.00	$\tilde{B}_1$
low-medium	0.00	0.00	0.25	0.50	0.75	1.00	0.75	0.50	0.25	0	0	$\tilde{B}_2$
low	0.50	0.75	1.00	0.75	0.50	0.25	0.00	0.00	0.00	0.00	0	$\tilde{B}_3$

Table 1 - 3: Decrease in Total Trade Partitioning

On the other hand, $\tilde{B}_1$, $\tilde{B}_2$ and $\tilde{B}_3$ are the fuzzy sets that describe a “medium”, “low-medium” and “low” decrease in total trade respectively. “B” in general is the linguistic variable which stands for “total trade”.

The benefit of fuzzy set theory should be mentioned here. One number is not necessarily high or low with a 100% certainty. If a value is closer to the target, its membership value is closer to 1. For example, in Table 1-2, a 0.7% increase in exchange rate volatility is categorized as a high increase with a 0.25 membership degree, while it is also possible to count it as a medium increase with a membership value of 0.5. The membership degree to the fuzzy set $\tilde{A}_2$ is higher than $\tilde{A}_1$ because 0.7% is closer to 0.5% than it is to 1%. By contrast, in the crisp sets, variables are categorized in certain classes and they can only belong to one class. If a number belongs to one class, it cannot be member of another.

In this study, triangular membership functions are used due to computational efficiency (Figure 1-1). However, it is also possible to use other functions such as trapezoidal or bell-shaped functions. The triangular membership function

$\Lambda : \mathrm{U} \rightarrow [\,0,\ 1\,]$ is defined by Driankov et al. (1996) as follows[1]:

$$\Lambda(u;\alpha,\beta,\gamma) = \begin{cases} 0 & \text{for } u < \alpha \\ (u-\alpha)/(\beta-\alpha) & \text{for } \alpha \le u \le \beta \\ (\gamma-u)/(\gamma-\beta) & \text{for } \beta \le u \le \gamma \\ 0 & \text{for } u > \gamma \end{cases} \qquad \text{(Eq. 3)}$$

Figures 1-2 and 1-3 show the partitioning of the universe of exchange rate volatility and that of total trade into three fuzzy sets. These figures depict the information given in Tables 1-2 and 1-3 respectively. The linguistic variable “exchange rate volatility” in Figure 1-2 is described via 3 linguistic values which are “high”, “medium” and “low” increase in exchange rate volatility. Similarly, in Figure 1-3 the linguistic variable is “total trade” and linguistic values for it are “medium”, “low-medium” and “low” decrease in total trade.

[1] The third line of this membership function was in error in the mentioned source, therefore it was corrected by the author of this article.

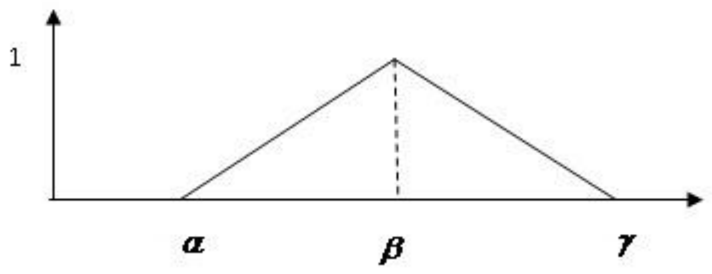


Figure 1 - 1: An example of a triangular function (Source: Driankov et al., 1996)

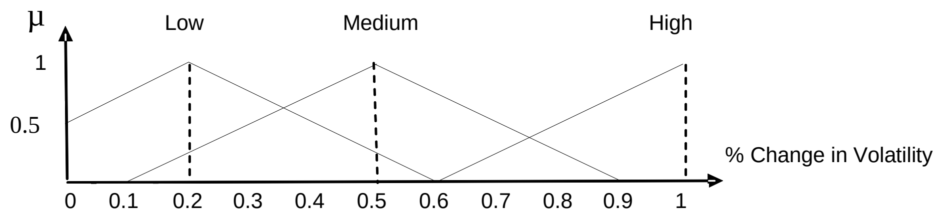


Figure 1 - 2: Linguistic Values for Variable "Exchange Rate Volatility"

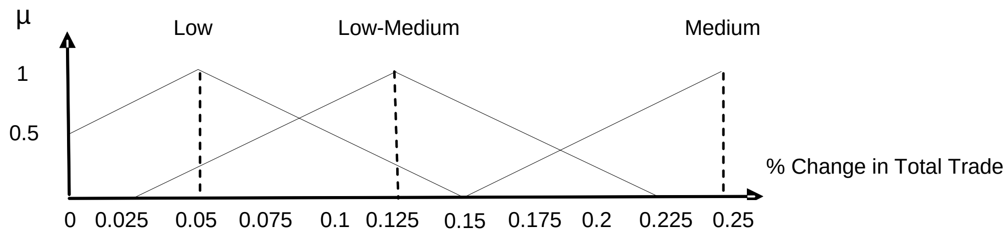


Figure 1 - 3: Linguistic Values for Variable "Total Trade

When dealing with fuzzy sets, the entire knowledge of the system is stored as rules in the knowledge base (Zimmermann, 2001). Thus, the rules play a very important role in fuzzy systems and therefore a considerable effort should be taken when defining the rules. Detailed information on the problem to be solved and experience are necessary to design a reliable fuzzy rule set and to obtain good results. If the designer does not have sufficient prior knowledge about the system or topic, it becomes impossible to develop a reliable fuzzy rule (Aliev et al., 2004). Fuzzy rules are the means that will translate inputs into the actual outputs (McNeil and Thro, 1994).

Under normal circumstances, traders do expect a stable economic environment and also no high volatility in exchange rates because high volatility in exchange rates means high volatility in their revenues as well. For this reason, when exchange rates fluctuate a lot, the impact of this change on total trade will be considerable because people do not expect enormous changes in exchange rate volatility. The fuzzy rule used in this study is constructed by considering that any increase in exchange rate volatility will affect total trade; however, the amount of decrease in total trade will not be exactly by the same percentage but lower.

	FUZZY RULE	
	If increase in exchange rate volatility is high ; Then decrease in total trade is medium	$\tilde{A}_1 \times \tilde{B}_1$
ELSE	If increase in exchange rate volatility is medium ; Then decrease in total trade is low-medium	$\tilde{A}_2 \times \tilde{B}_2$
ELSE	If increase in exchange rate volatility is low ; Then decrease in total trade is low	$\tilde{A}_3 \times \tilde{B}_3$

Table 1 - 4: Fuzzy rule for explaining the effects of increase in exchange rate volatility on bilateral trade

According to the fuzzy rule used (see Table 1- 4), a high increase in exchange rate volatility (1 percent) results in a medium (0.25 percent) decrease in bilateral trade, while a medium (0.5 percent) increase in exchange rate volatility leads to a low-medium (0.125 percent) decrease in bilateral trade. Furthermore, a low increase in exchange rate volatility causes a low decrease in bilateral trade.

After defining fuzzy rules, it is necessary to compute all rule-consequences (Zimmermann, 2001). In the fuzzification process, linguistic variables are described via linguistic values and quantitative values are assigned to these values. Then, possible consequences are defined for each possible input with

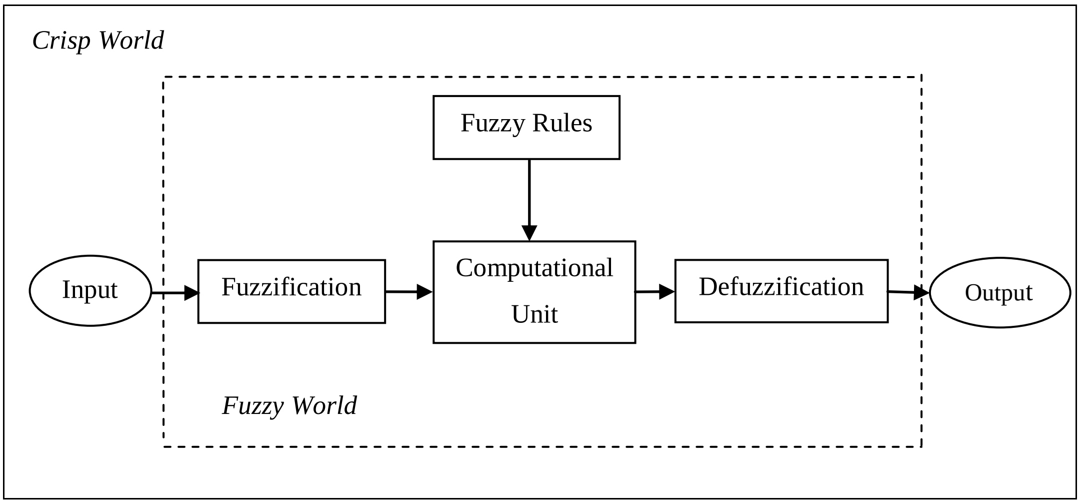


Figure 1 - 4: The process of fuzzification and defuzzification (Reconstructed from Zimmermann (2001) and Driankov et al. (1996)).

"if …., then ….” rules (see Table 1-4), and the consequences are aggregated into a fuzzy set (see Figure 1-4). The last step is defuzzification where one crisp value is generated from the fuzzy output set. The crisp value obtained after defuzzification enables the interpretation of the effect of a “1% increase in exchange rate volatility” on bilateral trade as a percentage value. Figure 1-4 shows how the whole process works.

Given the conclusions obtained by individual fuzzy rules shown in Table 1-4, the overall fuzzy relation ($\tilde{R}$) is calculated by taking the union of all individual effects:

$$\tilde{R} = \bigcup_{i=1}^{3} \tilde{A}_i \times \tilde{B}_i = (\tilde{A}_1 \times \tilde{B}_1) \cup (\tilde{A}_2 \times \tilde{B}_2) \cup (\tilde{A}_3 \times \tilde{B}_3) \qquad \text{(Eq. 4)}$$

where $\tilde{A}_i$ and $\tilde{B}_i$ are fuzzy sets and "x" denotes Cartesian product. The Cartesian product of $\tilde{A}_1$ and $\tilde{B}_1$ shows the impact of a high increase in exchange rate volatility on bilateral trade in a matrix form. Similarly, $\tilde{A}_2 \times \tilde{B}_2$ and $\tilde{A}_3 \times \tilde{B}_3$ depict the effect of a medium and a low increase in exchange rate volatility on bilateral trade respectively, again in a matrix form. The combination of these three individual effects is obtained by applying the union operator to these three matrices and the resultant matrix is $\tilde{R}$. Using this fuzzy relation ($\tilde{R}$) in matrix form, the impact of "1 percent increase in exchange rate volatility" on bilateral trade will be determined (See Appendix A.1.e for fuzzy mathematics used and Appendix 2 for the calculation of $\tilde{R}$).

To determine this effect we need to fuzzify "1% increase in exchange rate volatility". The fuzzy set $\tilde{C}$, called "1% increase in exchange rate volatility", is described by the membership function illustrated in Figure 1-5[2].

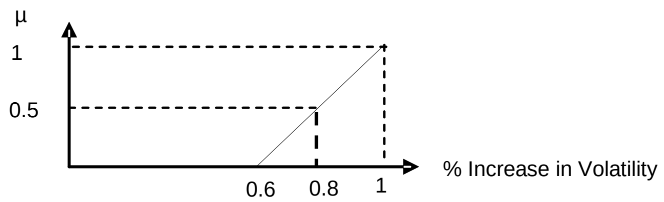


Figure 1 - 5: Membership Function for a "1% Increase in Volatility"

According to this membership function, 1% increase in exchange rate volatility has a membership value of 1 to the fuzzy set $\tilde{C}$. When the increase in exchange rate volatility is nearer to 1%, for example 0.9%, its membership value is 0.75. A

[2] Although it appears that the membership function of "1% increase in exchange rate volatility" corresponds to the fuzzy set which describes "high volatility in exchange rates" in Table 2 and also in Figure 2, it is just a coincidence. Different membership functions could also be used to define this fuzzy set such as $\tilde{C} = [0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0.75\ 1]$ or $\tilde{C} = [0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1]$.

0.8% increase in exchange rate volatility is the member of the fuzzy set of "1% increase in exchange rate volatility" with a degree of 0.5.

The effect of 1 percent increase in exchange rate volatility on bilateral trade can be obtained by applying the compositional rule of inference to the fuzzy set $\tilde{C}$ and fuzzy relation $\tilde{R}$: (see A.1.e. Definition 6 for details about the compositional rule of inference and operator "$\circ$").

$$\tilde{B}` = \tilde{C} \circ \tilde{R}$$

$$\tilde{B}` = [0\ \ 0\ \ 0.25\ \ 0.25\ \ 0.25\ \ 0.25\ \ 0.25\ \ 0.25\ \ 0.5\ \ 0.75\ \ 1\ ]$$

where $\tilde{C} = [0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0.25\ \ 0.5\ \ 0.75\ \ 1\ ]$ as shown in Figure 5.

$\tilde{B}`$ is the fuzzified decrease in bilateral trade where each number is a weight factor between 0 and 1, corresponding to the percentage values between 0 and 1 with an increment 0.025 (see Table 1-3).

The last step requires the defuzzification process, which converts the overall fuzzy conclusion ($\tilde{B}`$) into a real number that represents the decrease in bilateral trade following a 1% increase in exchange rate volatility. There are different defuzzification methods. Here, the centroid method—the center of the output membership function—is employed in the defuzzification process. This method uses a weighted average. From the mathematical point of view, it corresponds to the expected value of probability (Zimmermann, 2001). The centroid method yields that:

$$\%\,Change = \frac{\begin{array}{c}0\times0+0.025\times0+0.05\times0.25+0.075\times0.25+0.1\times0.25+\\0.125\times0.25+0.15\times0.25+0.175\times0.25+0.2\times0.5+0.225\times0.75+0.25\times1\end{array}}{0+0+0.25+0.25+0.25+0.25+0.25+0.25+0.5+0.75+1} = 0.183$$

In words, this means that 1 percent increase in exchange rate volatility leads to 0.18 percent decrease in bilateral trade. It is evident that this result is in accordance with the coefficient of 0.21 that was obtained by using panel data analysis with period fixed effects and is reported in Table1-1.

4- Conclusion

This study explains the effects of exchange rate volatility on bilateral trade between EU15 countries by using cross sectional methods and a fuzzy approach. Considering data for 40 years, panel data analysis with period fixed effects shows a significant negative impact of exchange rate volatility on bilateral trade. A fuzzy approach delivers a very similar result. The key elements of the fuzzy approach are to set fuzzy decision rules describing the event and to assign membership functions to the fuzzy sets intuitively based on experience. It should be emphasized that although the use of econometric methods is essential to obtain reliable results, employing a fuzzy intuitive approach can be useful for estimating first approximate results, especially in the absence of adequate data. Here, the fuzzy approach gives results very close to the ones from traditional panel data analysis, which supports our recommendation to use it more generally as a complement to statistical methods.

5- Appendices

Appendix 1

A.1.a. What is Fuzzy Set Theory?

The difference between conventional dual logic and fuzzy set theory is that in conventional dual logic a statement can be either true or false; in set theory, an element can be either a member of a set or not. However, real situations are very often uncertain. Lack of information, for instance, may cause the future state of the system to be unknown. This type of uncertainty has been handled by statistics and probability theory. Fuzziness can be found in many areas of life such as meteorology, medicine, engineering, manufacturing etc. In daily life, the meaning of words is often vague. When we say “tall man”, “beautiful women”, “successful company” the meaning of a word may change from person to person or from culture to culture. Fuzzy set theory provides a mathematical framework to study vague phenomena precisely. It is defined as a modeling language for fuzzy relations, criteria and situations (Zimmermann, 2001).

In fuzzy set theory, normal sets are called crisp sets to be differentiated from fuzzy sets (Driankov et al., 1996). Let C be a crisp set and F a fuzzy set defined on the universe U. For any element u of U, either $u \in C$ or $u \notin C$. However, in fuzzy set theory it is not necessary that either $u \in F$ or $u \notin F$. In fuzzy set theory, a membership function μ_F assigns a value to every $u \in U$ from the unit interval [0, 1], instead from the two element set {0, 1} as is done in crisp sets. A fuzzy set is defined on the basis of a membership function.

According to Zimmermann (2001), major goals of fuzzy set theory are the modeling of uncertainty and the generalization of classical methods based on dual logic from dichotomous to gradual features. Moreover, it aims to reduce the complexity of data to an acceptable degree by means of linguistic variables. Computational units (see Figure 1-4) process these linguistic expressions, use membership functions of fuzzy sets and finally retranslate the fuzzified result into the words via linguistic approximation which is explained in A.1.c.

A.1.b. Advantages and Disadvantages of Fuzzy Systems

The use of fuzzy logic in various fields has been quite popular due to its advantages (McNeill and Thro, 1994). In fuzzy logic, linguistic variables are used instead of numerical ones, which makes it similar to the way human beings think. The need for fewer values, rules and decisions than conventional models can be counted as another advantage. However, it is hard to develop a model from a fuzzy system. Even though they are easier to design and faster to prototype than conventional methods, fuzzy logic may face cultural bias in some countries who favor mathematically precise or crisp systems. This is why Japanese firms exploited fuzzy systems before the United States. Furthermore, as the complexity of a system increases, it becomes more difficult to specify the correct set of rules and membership functions to describe the behavior of the system appropriately (Aliev et al., 2004).

A.1.c. Linguistic Variables

Zadeh (1975) defines a linguistic variable as a variable whose values are words or sentences in a natural or artificial language. For example, age is a linguistic variable when it is defined as "young, very young, old, not very old" instead of 18, 15, 60 or 40.

The following framework, cited from Driankov et al. (1996), explains the notions of linguistic variable, linguistic value, actual physical domain and semantic function:

$$(X, \mathcal{L}(X), \mathcal{X}, Mx).$$

Here, X represents the symbolic name of a linguistic variable, for example age, temperature, error, weight, etc. In section 3, instead of X we have *"A"* and *"B"* which are the linguistic variables representing "exchange rate volatility" and "total trade" respectively.

$\mathcal{L}(X)$ denotes the set of linguistic values that X can take. Again, in our case

(A) = {high, medium, low}. $\mathcal{L}(X)$ can also be called as the term-set of X or the reference- set of it.

Furthermore, $\mathcal{X}$ is the actual physical domain over which the linguistic variable X can take its quantitative values. In the case of the linguistic variable "exchange rate volatility", $\mathcal{X}$ is the interval [0%, 1%] with 0.1 increments.

Mx is a semantic function which gives a quantitative interpretation of a linguistic value from the interval $\mathcal{X}$ and is defined as

$$Mx: \mathcal{L}(X) \rightarrow \widetilde{\mathcal{L}(X)}$$

where $\widetilde{\mathcal{L}(X)}$ is a denotation for a fuzzy set defined over $\mathcal{X}$. Put differently, *Mx* returns the meaning of a word into the fuzzy terms. Instead of $\widetilde{L(X)}$ it is also possible to use μ_{LX} which is the membership function.

The symbolic translation of natural language in terms of linguistic variables is explained by Driankov et al. (1996) as follows. The symbolic representation of the natural language expression "Error has the property of being negative-big" is written as "E is NB" and called an atomic fuzzy proposition.[3] The interpretation of this atomic representation is defined by the fuzzy set $\widetilde{NB}$ or the membership function μ_{NB} on the normalized physical domain $\varepsilon = [-6{,}6]$ of the physical variable "error",

$$\forall e \in \varepsilon : \widetilde{NB} = \mu_{NB} = \text{"membership function".}$$

where μ_{NB} shows the degree to which a specific quantitative crisp value of the physical variable error, e, belongs to the set $\widetilde{NB}$. For example, the degree of membership of -3.2 to the fuzzy set of negative big is $\mu_{NB}(-3.2) = 0.7$. This degree of membership shows the degree to which the symbolic expression "E is *NB*" is satisfied given the following circumstances: *NB* is interpreted as μ_{NB} and E takes the value -3.2.

A.1.d. Fuzzy If-then Statements

A fuzzy conditional or a fuzzy if-then statement describes the relationship between process state (which contains a description of the process output) and control output variables (which describe the control output that should be produced given the particular process output).

The meaning of the expression

if X is *A*, *then Y* is *B*

[3] The symbol *E* denotes the physical variable "error" and *NB* the particular value "negative big" of error.

is represented as a fuzzy relation defined on $\mathcal{X} \times \Upsilon$ where $\mathcal{X}$ and Υ are the physical domains of the linguistic variables X and Y. The meaning of "X is A" is called the *rule antecedent* and represented by the fuzzy set $\tilde{A} = \int_x \mu_A(x)/x$. The meaning of "$Y$ is B" is called the *rule consequent* and represented by the fuzzy set $\tilde{B} = \int_y \mu_B(y)/y$.[4]

Then, the meaning of the fuzzy conditional is a fuzzy relation μ_R such that

$$\forall\, x \in \mathcal{X} \;\; \forall\, y \in \Upsilon\colon\; \mu_R(x, y) = \mu_A(x) * \mu_B(y),$$

where "*" can be either Cartesian product or any fuzzy implication operator (Driankov et al., 1996).

To give an example, first of the three *if-then* rules used in section 3 is:

if <increase in exchange rate volatility is high> *then* <decrease in total trade is medium>

represented by $\tilde{A}_1 \times \tilde{B}_1$ in Table 1-4. Cartesian product was used to process the relation between the variable "exchange rate volatility" and "total trade".

A.1.e. Fuzzy Set Mathematics

The following definitions except Definition 2 and 6 will be cited from Zimmermann (2001).

Definition 1: If X is a collection of objects denoted generically by x, then a fuzzy set $\tilde{A}$ is a set of ordered pairs: $\tilde{A} = \{(x, \mu_{\tilde{A}}(x) \mid x \in X\}$

where $\mu_{\tilde{A}}$ is called the membership function of x in $\tilde{A}$ that maps X to the membership space M. The range of the membership function is a subset of the nonnegative real numbers whose supremum is finite. However, as a matter of convenience it is assumed that fuzzy sets are normalized to the range [0, 1].

[4] This fuzzy set was in error in the mentioned source, therefore it was corrected by the author of this article.

The membership function is the fundamental part of a fuzzy set. Therefore, operations with fuzzy sets are defined through their membership functions. It is defined by Driankov et al. (1996) as follows:

Definition 2: The membership function μ_F of a fuzzy set F is a function

$$\mu_F : U \rightarrow [0,1].$$

So, each element u from U (universe) has a membership degree $\mu_F(u) \in [0,1]$.

F is completely determined by the set of tuples

$$F = \{(u, \mu_F(u)) | u \epsilon U\}.$$

Definition 3: The membership function $\mu_{\tilde{C}}(x)$ of the intersection $\tilde{C} = \tilde{A} \cap \tilde{B}$ is pointwise defined by

$$\mu_{\tilde{C}}(x) = \min\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\}, x \in X .$$

Definition 4: The membership function $\mu_{\tilde{D}}(x)$ of the union $\tilde{D} = \tilde{A} \cup \tilde{B}$ is pointwise defined by

$$\mu_{\tilde{D}}(x) = \max\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\}, x \in X .$$

Definition 5: The Cartesian product of fuzzy sets is defined as follows: Let $\tilde{A}_1, \tilde{A}_2, ..., \tilde{A}_n$ be fuzzy sets in $X_1, ..., X_n$. The Cartesian product is then a fuzzy set in the product space $X_1 \times X_2 \times ..... \times X_n$ with the membership function

$$\mu_{(\tilde{A}_{1 \times ... \times} \tilde{A}_n)}(x) = \min\{\mu_{\tilde{A}_i}(x_i) | x = (x_1, ..., x_n), x_i \in X_i\}.$$

Definition 6: Compositional Rule of Inference: If $\tilde{R}$ is a fuzzy relation from U to V, and $\tilde{x}$ is a fuzzy subset of U, then the fuzzy subset $\tilde{y}$ of V which is induced by $\tilde{x}$ is given by the composition of $\tilde{R}$ and $\tilde{x}$; that is

$$\tilde{y} = \tilde{x} \circ \tilde{R}$$

in which $\tilde{x}$ plays the role of a unary relation (Zadeh, 1973).

Appendix 2: Calculation of Overall Fuzzy Relation

$$\tilde{R} = \bigcup_{i=1}^{3} \tilde{A}_i \times \tilde{B}_i = (\tilde{A}_1 \times \tilde{B}_1) \cup (\tilde{A}_2 \times \tilde{B}_2) \cup (\tilde{A}_3 \times \tilde{B}_3)$$

if < increase in exchange rate volatility is high> *then* <decrease in total trade is medium>

$\tilde{A}_1 \times \tilde{B}_1$		**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.25**	**0.50**	**0.75**	**1**
(High x medium)	**0**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0
	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0
	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0
	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0
	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0
	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0
	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0
	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.25	0.25	0.25
	0.5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.50	0.50	0.5
	0.75	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.50	0.75	0.75
	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.50	0.75	1

if < increase in exchange rate volatility is medium> *then* <decrease in total trade is low-medium>

$\tilde{A}_2 \times \tilde{B}_2$		**0.00**	**0.00**	**0.00**	**0.25**	**0.50**	**0.75**	**1.00**	**0.75**	**0.50**	**0.25**	**0**
(Medium x low-medium)	**0**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0
	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0
	0.25	0.00	0.00	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.00	0
	0.5	0.00	0.00	0.25	0.25	0.50	0.50	0.50	0.50	0.50	0.00	0
	0.75	0.00	0.00	0.25	0.25	0.50	0.75	0.75	0.75	0.50	0.00	0

1	0.00	0.00	0.25	0.25	0.50	0.75	1.00	0.75	0.50	0.00	0
0.75	0.00	0.00	0.25	0.25	0.50	0.75	0.75	0.75	0.50	0.00	0
0.5	0.00	0.00	0.25	0.25	0.50	0.50	0.50	0.50	0.50	0.00	0
0.25	0.00	0.00	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.00	0
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0
0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0

if <increase in exchange rate volatility is low> *then* <decrease in total trade is low>

$\tilde{A}_3 \times \tilde{B}_3$		**0.50**	**0.75**	**1.00**	**0.75**	**0.50**	**0.25**	**0.00**	**0.00**	**0.00**	**0.00**	**0**
(Low x low)	**0.5**	0.50	0.50	0.50	0.50	0.50	0.25	0.00	0.00	0.00	0.00	0
	0.75	0.50	0.75	0.75	0.75	0.50	0.25	0.00	0.00	0.00	0.00	0
	1	0.50	0.75	1.00	0.75	0.50	0.25	0.00	0.00	0.00	0,00	0
	0.75	0.50	0.75	0.75	0.75	0.50	0.25	0.00	0.00	0.00	0.00	0
	0.5	0.50	0.50	0.50	0.50	0.50	0.25	0.00	0.00	0.00	0.00	0
	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.00	0.00	0.00	0.00	0
	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0
	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0
	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0
	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0
	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0

$\tilde{R}$	0.5	0.5	0.5	0.5	0.5	0.25	0.00	0.00	0.00	0.00	0.00
	0.5	0.75	0.75	0.75	0.5	0.25	0.00	0.00	0.00	0.00	0.00
	0.5	0.75	1	0.75	0.5	0.25	0.25	0.25	0.25	0.00	0.00
	0.5	0.75	0.75	0.75	0.5	0.5	0.5	0.5	0.25	0.00	0.00
	0.5	0.5	0.5	0.5	0.75	0.75	0.75	0.5	0.25	0.00	0.00

0.25	0.25	0.25	0.5	0.75	1	0.75	0.5	0.25	0.00	0.00
0.00	0.00	0.25	0.5	0.75	0.75	0.75	0.5	0.25	0.00	0.00
0.00	0.00	0.25	0.5	0.5	0.5	0.5	0.5	0.25	0.25	0.25
0.00	0.00	0.25	0.25	0.25	0.25	0.25	0.25	0.5	0.5	0.5
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.5	0.75	0.75
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	0.5	0.75	1

7- References

Akhtar M.A. and R.S. Hilton (1984) “Effects of Exchange Rate Uncertainty on German and US Trade”. Federal Reserve Bank of NewYork Quarterly Review 9(1): 7-16.

Aliev Rafik A., Bijan Fazlollahi and Rashad R. Aliev (2004) Soft Computing and its Applications in Business and Economics. Springer Verlag: Berlin Heidelberg.

Baccheta Philippe and Eric van Wincoop (2000) “Does Exchange Rate Stability Increase Trade and Welfare?”, American Economic Review 90(5): 1093-1108.

Bubula Andrea and Inci Otker-Robe (2003) “Are Pegged and Intermediate Exchange Rate Regimes More Crisis Prone?”. IMF Working Paper.

Cushman David O. (1983) “The Effects of Real Exchange Rate Risk on International Trade”. Journal of International Economics 15(1): 45-63.

De Grauwe Paul (1987) “International Trade and Economic Growth in the European Monetary System”. European Economic Review 31: 389-398.

Driankov Dimiter, Hans Hellendoorn and Michael Reinfrank (1996) An Introduction to Fuzzy Control. Springer Verlag: Berlin-Heidelberg-NewYork.

Ethier Wilfred (1973) “International Trade and the Forward Exchange Market”. American Economic Review 63(3): 494-503.

Fischer Stanley (2001) “Distinguished Lecture on Economics in Government: Exchange Rate Regimes: Is the Bipolar View Correct?”. Journal of Economic Perspectives 15(2): 3-24.

Fu King-Sun and James T. P. Yao (1980) “Application of Fuzzy Sets in Earthquake Engineering”. Working Paper Purdue University: CE-STR-80-11.

Hooper Peter and Steven W.Kohlhagen (1978) “The effects of Exchange Rate Uncertainty on the Prices and Volume of International Trade”. Journal of International Economics 8(4): 483-511.

Huang Rocco R. (2007) “Distance and Trade: Disentangling Unfamiliarity Effects and Transport Cost Effects”. European Economic Review 51: 161-181.

Gotur Padma (1985) “Effects of Exchange Rate Volatility on Trade”. IMF Staff Papers 32: 475- 512.

International Monetary Fund (1984) “Exchange Rate Volatility and World Trade: A Study by the Research Department of the IMF”. Occasional Paper 28.

Kenen Peter B. and Dani Rodrik (1986) “Measuring and Analyzing the Effects of Short-term Volatility in Real Exchange Rates”. The Review of Economics and Statistics 68: 311-315.

Kowalski Przemyslaw (2006) “The Impact of the Economic and Monetary Union in the EU on International Trade- A Reinvestigation of the Exchange Rate Volatility Channel”. PhD Thesis, University of Sussex.

Krugman Paul R. and Maurice Obstfeld (2006) International Economics: Theory and Policy. 7th Edition, Pearson Addison Wesley: London.

Lane Philip.R. and Gian Maria Milesi-Ferretti (2002) “External Wealth, the Trade Balance and the Real Exchange Rate”. European Economic Review 46: 1049-1071.

Lee Vincent C.S. and Hsiao Tshung Wong (2007) “A Multivariate Neuro-fuzzy System for Foreign Currency Risk Management Decision Making”. Neurocomputing 70: 942-951.

McNeill F. Martin and Ellen Thro (1994) Fuzzy Logic: A Practical Approach. Academic Press Limited: London.

Tseng Fang-Mei, Gwo-Hshiung Tzeng, Hsiao-Cheng Yu and Benjamin J.C. Yuan (2001) “Fuzzy ARIMA Model for Forecasting the Foreign Exchange Market”. Fuzzy Sets and Systems 118: 9-19.

Viane Jean-Marie and Casper G. de Vries (1992) “International Trade and Exchange Rate Volatility”. European Economic Review 36(6): 1311-21.

Zadeh, L. A. (1973) “Outline of a New Approach to the Analysis of Complex Systems and Decision Processes”. IEEE Trans. On Systems, Man and Cybernetics 3(1):28-44.

Zadeh L.A. (1975) “The Concept of a Linguistic Variable and its Application to Approximate Reasoning – I”. Information Sciences 8(3): 199-249.

Zimmermann Hans-Jürgen (2001) Fuzzy Set Theory and its Applications. 4th Edition, Kluwer Academic Publishers: USA.

DIFFERENT SPECIFICATIONS OF THE GRAVITY EQUATION: A THREE-WAY MODEL OR BILATERAL INTERACTION EFFECTS?

Elif Nuroğlu[+], Robert Kunst[++]

[+] International University of Sarajevo

[++] University of Vienna

Abstract: In this study we investigate bilateral trade flows across EU15 countries from 1962 to 2003 by using two different specifications of the gravity model. We augment the basic gravity model with (i) exporting-country, importing country and time effects (3-way model) and (ii) country pair and time effects (2-way model), compare the results of the 3-way model with those of the 2-way model, show the differences between them and claim the superiority of the latter to the former due to its higher explanatory and predictability power.

Keywords: *gravity model, main effects, bilateral interaction effects.*

1. Introduction

This study explains bilateral trade flows among European countries from 1962 to 2003 using different specifications of the gravity model, one is with exporting-country, importing country and time effects and the other is country pair and time effects. This sheds some light on deciding which model to use when we have some specific purposes such as to see how fertility affects bilateral trade or to learn about the influences of bilateral variables such as the distance, common border, trade agreement etc. on bilateral trade.

We start with the most basic gravity model that includes only GDPs of exporter and importer countries and distances between them. Then we insert population and real bilateral exchange rates into the equation and see that variation in bilateral trade is explained better. Our objective is to specify the model so that the most relevant effects are included; therefore, finally by including exporting-country, importing-country and time effects (3-way model) on the one hand and by including country pair effects together with time effects on the other hand (2-way model) we show that our new models can explain a much larger part of the variation in bilateral exports. In the last step, we compare out-of-sample forecasting performances of the 3-way and the 2-way model and claim the superiority of one to the other.

Following Egger and Pfaffermayr (2003) and Harris and Matyas (1998) we apply a main effects three-way model and a bilateral interaction effects model to a panel of EU15 countries for the years between 1962 and 2003. We show how the results are affected by the use of different specifications concerning the existence of exporting and importing-country and country-pair effects. We find that the 2-way model explains a larger variation in bilateral exports across Europe. However, the 3-way model may be preferable if one wants to see the effects of bilateral variables such as distance, common language, common border etc. on bilateral trade separately.

The plan of the paper is as follows: section 2 gives a short survey, section 3 describes the data and section 4 explains about the augmented gravity model. Section 5 summarizes our findings, section 6 compares the forecasting performance of the main effects model with the bilateral effects model and section 7 concludes.

2. A short survey

Using the gravity model to explain and estimate bilateral trade flows has become very popular in the international trade literature. The basic gravity model which is developed by Tinbergen in the 1960s says that bilateral trade between two countries depends on their economic size positively and distances between them negatively (Tinbergen 1962: 263-264). This model suggests that higher income tends to boost trade by leading to more production, higher exports and also higher demand for imports (Dell`Ariccia: 1999; De Grauwe and De Bellefroid: 1986; Cushman: 1983, Balogun: 2007; Clark et al.: 2004; Glick and Rose: 2002; Matyas: 1997; Rose et al.: 2000). Furthermore, larger distances between countries tend to decrease bilateral trade (Clark et al.: 2004; Glick and Rose: 2002; Rose et al.:2000) by imposing higher transport costs and some other difficulties to trade such as informational and psychological frictions (Huang: 2007). As is well known, transport costs are an important barrier to trade and therefore they tend to reduce international trade (Jacquemin and Sapir: 1988; Neven and Röller: 1991).

The gravity model can be extended by including the populations of exporting and importing countries to see how the population affects bilateral trade. Matyas (1997), among others, finds that population has a positive effect on trade and increase the level of specialization by creating gains from specialization. Other authors, such as Bergstrand (1989) and Dell`Ariccia (1999), find negative population coefficients, suggesting that imports and exports are capital intensive.

In addition to the population, real bilateral exchange rates are also important variables in explaining bilateral trade flows. International trade theory suggests that an appreciation in a country makes its goods more expensive abroad and therefore decreases exports. However, at the same time an appreciation makes imports into that country cheaper and has a tendency to increase imports. If the decrease in exports exceeds the increase in imports for a country, the result will be a decrease in total trade.

Some studies which analyze the effects of exchange rate uncertainty and/or volatility on international trade find significant negative effects (Lane and Milesi-Ferretti: 2002; Dell`Ariccia: 1999; De Grauwe: 1987; De Grauwe and De Bellefroid: 1986; Kennen and Rodrik: 1986; Thursby and Thursby: 1985; Thursby and Thursby: 1987; Akhtar and Hilton: 1984; Cushman: 1983; Ethier: 1973; Kowalski: 2006; Wei: 1999; Chowdhury: 1993; De Grauwe: 1988; Rose et al.:2000). One explanation for this negative effect is offered by Frank and Bernanke (2007) who suggest that uncertainty in exchange rates under flexible exchange rate systems makes exporters` profits less predictable, therefore it may make people more reluctant to export and reduce total trade. However, there is another side in the literature which claims that there is no significant effect of exchange rate uncertainty and/or volatility on the volume of trade. Some of these studies argue that even if there is some small significant effect of exchange rates on trade, this effect is neither stable nor consistent (Hooper and Kohlhagen: 1978; Gotur: 1985; Bacchetta and van Wincoop: 2000). One reason that these authors could not find any significant effect may be their focus on short-run measures. Moreover, while Perée and Steinherr (1989) find that exchange rate uncertainty has negative effects on the volume of trade among industrial countries, their conclusion is noteworthy: The effects of uncertainty can be expected to vary from country to country depending on the structural characteristics of the country. Therefore, it is not surprising that different studies have different results.

More recently, Clark et al. (2004) find a negative association between exchange rate volatility and trade in certain country groupings. However, when they analyze the time of the increase in volatility and decrease in trade, they see that the decrease in trade may not be attributed only to the increase in exchange rate volatility. At crises, for instance, even if volatility in exchange rates increases, the fall in domestic demand is a much more important factor which decreases imports. When they allow for time-varying fixed effects they do not find a negative association between exchange rate volatility and trade.

3. The Data

The data used in this study is obtained from IMF`s International Financial Statistics, World Bank`s World Development Indicators 2005, and OECD`s International Trade by Commodity Statistics. The sample period covers 42 years from 1962 to 2003. Countries included are the EU-15 where Belgium and Luxembourg are taken as one country because of data availability.

Our model uses bilateral trade flows which cover fixed and flexible exchange rate periods as well as the Euro era.

4. A Modified Gravity Model of Bilateral Trade

Our study analyzes trade flows across Europe from 1962 to 2003 by using cross sectional data. We expand the basic gravity model which explains bilateral trade between two countries only with their GDPs and distances to a more general model. We form a model which includes the population of importing and exporting countries, real exchange rates between exporter and importer's currency, exporting country, importing country, country pair and time effects. Exporting-country and importing-country effects are supposed to capture the tendency to export and import while business cycle effects control for cyclical changes and their effects on bilateral exports. Country-pair or bilateral effects are used to see any geographical, political, historical or cultural event that influence bilateral trade between countries.

In the first step of this study, we insert exporting-country, importing-country, business cycle and country-pair effects into the gravity model. In the second step, we compare the 3-way model with the 2-way model. Main and bilateral effects models have been used by Matyas (1997), Harris and Matyas (1998) and Egger and Pfaffermayr (2003) previously. In contrast to previous studies, we do not include main and bilateral effects in one model, but form two different models, one with main effects and the other with bilateral effects, to see which one explains the variation in bilateral trade better and which one serves what purpose.

The augmented model used here is:

$$\ln Exp_{ijt} = \beta_0 + \beta_1 \ln D_{ij} + \beta_2 \ln Y_{it} + \beta_3 \ln Y_{jt} + \beta_4 \ln Pop_{it} + \beta_5 \ln Pop_{jt} + \beta_6 d\ln(XR_{ijt}) + \alpha_i + \lambda_j + \gamma_t + \delta_{ij} + \varepsilon_{ijt},$$

where *i*=exporter, *j*= importer, and

- Exp_{ijt} represents nominal exports from country *i* to country *j*,
- D_{ij} is the distance between country *i* and country *j* measured in kilometres,
- Y_{it} is exporting country's real GDP,
- Y_{jt} is importing country's real GDP,
- Pop_{it} is exporter country's population in year *t*,
- Pop_{jt} is importer country's population in year *t*,
- XR_{ijt} is the real exchange rate between exporter and importer country in time *t*,
- α_i is exporter country effects,
- λ_j is importer country effects,
- γ_t is business cycle effects,
- δ_{ij} is country pair or exporter-by-importer effects,
- ε_{ijt} is the error term.

α_i and λ_j may differ depending on the countries` tendency to export and import. Exporter country dummy α_i equals 1 whenever country *i* is exporting and 0 otherwise. Similarly, λ_j equals 1 whenever country *j* is importing and 0 otherwise.

Time or business cycle effect dummy γ_t equals 1 in the current year and 0 otherwise. Moreover, country pair dummy δ_{ij} equals 1 whenever country *i* is exporting to country *j* and 0 otherwise.

5. Results

Table 2-1 compares the results of the main effects three-way model (Model 1), the bilateral effects model with time effects (Model 2), the main effects three-way model without distance (Model 3), simple OLS results (Model 4) and the very basic gravity model (Model 5). The bilateral effects model with time effects (Model 2) appears to be superior to all others, and also to the main effects three-way model (Model 1) as indicated by adjusted R-squared and Akaike`s "an information criterion" (AIC).

The exporter and importer GDP coefficients do not differ so much in the first three models, while these coefficients become slightly higher in the simple OLS model. In the first model, the exporter GDP's coefficient of 1.48 means that a 1% increase in exporter country's GDP increases its bilateral trade by 1.48%. In the second model, the exporter GDP's coefficient of 1.49 implies that a 1% increase in exporter country's GDP increases its bilateral trade by 1.49%. Positive GDP coefficients in all models prove that higher income tends to boost bilateral trade for the exporter as well as the importer country. Moreover, in the first three models exporter population has a high positive coefficient which shows that the effect of a 1% increase in exporter's population on its bilateral trade is about 4% increase in its exports. The positive effect of exporter country's population on bilateral trade suggests that higher population increases the level of specialization, leads to higher production and more trade. On the other hand, importer country's population has a negative impact on bilateral trade in all models.

When comparing exporter population coefficients in the simple OLS model with the first three models, we see that exporter population has a negative effect on bilateral trade in the simple OLS model. In the literature, the effect of population on bilateral trade is interpreted in a different way in the simple OLS models than other models. Cross sectional OLS models, in most cases, suggest a negative population coefficient which refers to the long run effect of population on bilateral trade, where higher population tends to decrease income per capita and therefore causes production and exports to decrease. Additionally, lower income per capita is expected to decrease the demand for imports as well. In the first three models in this study, population coefficients show the short run effect of population on bilateral trade. In the short-run, higher population is good for the country as it means more labor and more products to export; therefore exporter population has a positive coefficient in the first three models.

Another variable of interest in our study is the real exchange rate which has very close coefficients in the first three models, meaning that a 1% appreciation in the exporter country leads to a 0.33% decrease in bilateral trade.

Finally, we have the results of the most basic gravity model which explain bilateral exports only with the distance and income of exporter and importer countries. Including population and real exchange rate variables increases the explanatory power of the model from 80% to 89%. Furthermore, including exporter, importer and time effects (Model 1) increases adjusted R-squared to 95%. Ultimately, Model 2 with time and bilateral effects can explain the variation better than all other models (98%). The Wald tests indicate that exporter, importer, time and also bilateral effects are all highly significant and contribute well to the models.

It appears that the bilateral and time effects model (2-way model) is superior to the 3-way model (Model 1) and all other models (Models 3 to 5) in this study. The only disadvantage of relying on a two-way model is that the two-way model

	Model 1: Main effects (exporter, importer and time)		Model 2: Time and bilateral effects		Model 3: Model 1 without distance		Model 4: Simple OLS results		Model 5: The basic gravity model	
	Coefficient	**t-Statistic**	**Coefficient**	**t-Statistic**	**Coefficient**	**t-Statistic**	**Coefficient**	**t-Statistic**	**Coefficient**	**t-Statistic**
C	-101.60	-13.47	-1.20	-24.22	-107.79	-11.18	-43.78	-87.09	-35.67	-56.51
Distance	-0.96	-48.44					-0.72	-35.71	-0.81	-30.53
Exporter GDP	1.48	11.51	1.49	18.07	1.48	8.97	2.32	73.17	1.11	77.59
Importer GDP	1.15	12.53	1.15	19.58	1.15	9.83	1.37	51.31	0.98	68.26
Exporter Population	4.02	10.60	4.01	16.49	4.04	8.32	-1.41	-42.57		
Importer Population	-0.52	-1.59	-0.49	-2.35	-0.54	-1.30	-0.64	-21.15		
Real Exchange Rate	-0.33	-2.24	-0.33	-3.49	-0.32	-1.70	-0.60	-2.93		
R^2	0.95		0.98		0.92		0.89		0.80	
Adjusted R^2	0.95		0.98		0.92		0.89		0.80	
AIC	1.47		0.60		1.96		2.23		2.84	
Schwarz criterion	1.59		0.82		2.08		2.24		2.84	
Number of Observations	3760		3760		3760		3760		3887	
		p-value		**p-value**		**p-value**				
Wald-tests:										
Exporter effect: F-stat.	143.22	0.00			139.70	0.00				
Importer effect: F-stat.	51.31	0.00			72.93	0.00				
Time effect: F-stat.	13.84	0.00	33.66	0.00	8.47	0.00				
Bilateral interaction: F-stat.			235.89	0.00						

Table 2 - 1: Balanced panel estimates, dependent variable: log of nominal bilateral exports

makes us unable to include some observable bilateral variables such as distance, common border, common language etc. into the model separately. However, higher R-squared and a much lower AIC of the two-way model prove that even if the two-way model does not show the effects of bilateral variables individually, it is really powerful in explaining bilateral trade flows with aggregate bilateral effects.

6. The Forecasting Performance of Main Effects Model and Bilateral Effects Model

So far R-squared and AIC have been used to compare the main effects model with the bilateral effects model. In this part, the forecasting performance of both models is compared. The whole sample is cut into in- and out-of-samples randomly and forecast errors are calculated for each model as shown in Table 2-2.

The forecast results indicate that the bilateral effects model (Model 2) gives lower forecast errors than the main effects model (Model 1) in all samples. This shows that in addition to its higher explanatory power, Model 2 offers better forecasts with lower errors for each sample used. These results strengthen the view that the bilateral effects model (2-way model) is superior to the main effects model (3-way model).

	Forecast Sample	2003	2002-2003	1997-2003	1996-2000	1990-1995	1995-1998	1986-1990	1971-1974
Model 1 (Main effects model)	Number of observations	55	146	601	455	429	364	455	364
	Root Mean Squared Error	0.53	0.51	0.52	0.50	0.51	0.45	0.51	0.56
	Mean Squared Error	0.28	0.26	0.27	0.25	0.26	0.20	0.26	0.31
Model 2 (Bilateral Effects Model)	Number of observations	55	146	601	455	429	364	455	364
	Root Mean Squared Error	0.44	0.42	0.44	0.42	0.43	0.33	0.37	0.36
	Mean Squared Error	0.20	0.17	0.19	0.18	0.18	0.11	0.14	0.13

Table 2 - 2: Forecast Errors produced by Model 1 and Model 2

7. Conclusion

In this study different specifications of the gravity model are used to show the differences between a three-way and two-way model, which were integrated by Egger and Pfaffermayr (2003). Instead of combining main and bilateral effects in one model, we analyze bilateral trade among the EU-15 countries with the help of two different models to see which model is more helpful for which purpose. It is found that the adjusted R-squared is higher and the AIC is much lower in the bilateral effects model (two-way model) which shows that the model with time and bilateral effects explains the variation in bilateral exports between EU15 countries better than all other models used in this study, including main effects 3-way model. Moreover, when some forecasts for randomly chosen samples are performed, the two-way model is found to offer lower forecast errors than the three-way model. The results of this study prove that the two-way model is superior to the three-way model due to its higher explanatory and predictability power.

8. References

Akhtar M.A. and R.Spence Hilton (1984) “Effects of Exchange Rate Uncertainty on German and US Trade”. Federal Reserve Bank of New York Quarterly Review 9(1): 7-16.

Bacchetta Philippe and Eric Van Wincoop (2000) “Does Exchange Rate Stability Increase Trade and Welfare?” American Economic Review 90(5): 1093-1109.

Balogun Emmanuel Dele (2007) “Effects of Exchange Rate Policy on Bilateral Export Trade of WAMZ Countries”. MPRA Paper: 6234.

Bergstrand Jeffrey H. (1989) “The Generalized Gravity Equation, Monopolistic Competition, and the Factor-proportions Theory in International Trade”. The Review of Economics and Statistiscs 71(1): 143-153.

Chowdhury Abdur R. (1993) “Does Exchange Rate Volatility Depress Trade Flows? Evidence from Error Correction Models”. The Review of Economics and Statistics 75(4): 700-706.

Clark Peter, Natalia Tamirisa, Shang-Jin Wei, Azim Sadikov and Li Zeng (2004) “Exchange Rate Volatility and Trade Flows - Some New Evidence”. International Monetary Fund.

Cushman David O. (1983) “The Effects of Real Exchange Rate Risk on International Trade”. Journal of International Economics 15: 45-63.

De Grauwe Paul (1987) “International Trade and Economic Growth in the European Monetary System”. European Economic Review 31: 389-398.

De Grauwe Paul (1988) “Exchange Rate Variability and the Slowdown in the Growth of International Trade”, IMF Staff Papers 35(1): 63-84.

De Grauwe, Paul and De Bellefroid (1986) “Long Run Exchange Rate Variability and International Trade”. Chapter 8 in S. Arndt and J.D. Richardson (eds), Real Financial Linkages Among Open Economies. The MIT Press: London.

Dell`Ariccia Giovanni (1999) “Exchange Rate Fluctuations and Trade Flows: Evidence from the European Union”. IMF Staff Papers 46(3): 315-334.

Egger Peter and Michael Pfaffermayr (2003) “The Proper Panel Econometric Specification of the Gravity Equation: A three-way Model with Bilateral Interaction Effects”. Empirical Economics 28: 571-580.

Ethier Wilfred (1973) "International Trade and the Forward Exchange Market". American Economic Review 63(3): 494-503.

Frank Robert H. and Ben S. Bernanke (2007) Principles of Economics. Mc Graw Hill Irwin.

Glick Reuven and Andrew K. Rose (2002) "Does a Currency Union Affect Trade? The Time-series Evidence". European Economic Review 46: 1125 – 1151.

Gotur Padma (1985) "Effects of Exchange Rate Volatility on Trade". IMF Staff Papers 32: 475-512.

Harris Mark N. and Laszlo Matyas (1998) "The Econometrics of Gravity Models". Melbourne Institute Working Paper 5/98.

Hooper Peter and Stewen W. Kohlhagen (1978) "The Effect of Exchange Rate Uncertainty on the Prices and Volume of International Trade". Journal of International Economics 8: 483-511.

Huang Rocco R. (2007) "Distance and Trade: Disentangling Unfamiliarity Effects and Transport Cost Effects". European Economic Review 51: 161-181.

Jacquemin Alexis and Andrè Sapir (1988) "International Trade and Integration of the European Community". European Economic Review 32: 1439-1449.

Kennen Peter B. and Dani Rodrik (1986) “Measuring and Analyzing the Effects of Short-term Volatility in Real Exchange Rates”. The Review of Economics and Statistics 68: 311-315.

Kowalski Przemyslaw (2006) “The Impact of the Economic and Monetary Union in the EU on International Trade – A Reinvestigation of the Exchange Rate Volatility Channel”. PhD Thesis submitted at the University of Sussex.

Lane Philip.R. and Gian Maria Milesi-Ferretti (2002) “External Wealth, the Trade Balance and the Real Exchange Rate”. European Economic Review 46: 1049-1071.

Matyas Làszlò (1997) “Proper Econometric Specification of the Gravity Model”. World Economy 20(3): 363-368.

Neven Damien J. and Lars-Hendrik Röller (1991) “European Integration and Trade Flows”. European Economic Review 35 (6): 1295-1309.

Perée Eric and Alfred Steinherr (1989) “Exchange Rate Uncertainty and Foreign Trade”. European Economic Review 33: 1241-64.

Rose Andrew K., Ben Lockwood and Danny Quah (2000) “One Money, One Market: The Effect of Common Currencies on Trade”. Economic Policy 15(30): 7-45.

Tinbergen Jan (1962) Shaping the World Economy: Suggestions for an International Economic Policy. New York: Twentieth Century Fund.

Thursby Marie and Jerry Thursby (1985) “The Uncertainty Effects of Floating Exchange Rates: Empirical Evidence on International Trade Flows”. in Arndt S. W., R. J. Sweeney, and T. D. Willett, Exchange Rates, Trade and the US Economy. Ballinger Publishing Company: Cambridge, MA.

Thursby Jerry G. and Marie C. Thursby (1987) “Bilateral Trade Flows, the Linder Hyphotesis and Exchange Risk”. The Review of Economics and Statistics 69(3): 488-495.

Wei Shang-Jin (1999) “Currency Hedging and Goods Trade”. European Economic Review 43: 1371- 1394.

ESTIMATING BILATERAL TRADE FLOWS AMONG EU15 COUNTRIES USING PANEL DATA ANALYSIS AND NEURAL NETWORKS

Elif Nuroglu

International University of Sarajevo

Abstract: The objective of this paper is to investigate bilateral trade flows among European countries from 1964 to 2003 and their determinants by using panel data analysis and neural networks. In this study, the gravity model of bilateral trade is extended with population and volatility of exchange rates and this modified gravity model is used in the above mentioned models. Lastly, the results as well as the forecasting performance of both models are compared. It is demonstrated that neural networks are superior to traditional panel data analysis in learning and explaining the relationship between inputs and outputs. Moreover, it is shown that the out-of-sample forecasting performance of our neural networks dominates linear panel models.

Keywords: *gravity model, panel data, neural networks.*

1- Introduction

This study investigates bilateral trade flows among EU15 countries from 1964 to 2003 by using panel data analysis and neural networks. The basic concept for both approaches is the gravity model of bilateral trade which was developed by Tinbergen

(Tinbergen, 1962: 263-264). This model explains bilateral trade flows between two countries via the product of their incomes and distances between them. Here, we consider an extended version that also takes the population of both countries and exchange rate volatility into account. This augmented gravity model is estimated on a panel data set, with real bilateral exports as the dependent variable and the income and population of both countries, distances between them and the volatility of exchange rates as influence factors. We contrast this linear panel model with a neural network, in which income and population of both countries, distance and exchange rate volatility are the inputs of the network (influence factors in panel model) and bilateral exports (the dependent variable in panel model) is the output. We divide the data set of 3586 observations into training, validation and test sets; and we train the network. Finally, we compare the performances of panel model and neural network, and discuss the advantages of each model.

The structure of the paper is as follows. Section 2 compares neural networks to traditional econometric and statistical methods by giving examples from conducted studies. Section 3 gives an overview about neural networks and discusses advantages and disadvantages of them. Section 4 and 5 introduces the data set used and our modified gravity model with its application to panel data analysis respectively. Section 6 shows how the modified gravity model is used to build a neural network. Section 7 discusses the results of panel data analysis and neural network model, and compares the forecasting performance of both models. Finally, section 8 concludes. Detailed information about neural networks can be found in Appendix 1.

2- Studies on Neural Networks and Traditional Models

The comparative advantage of neural networks over traditional econometric models is that they can approximate any function arbitrarily well if the networks are allowed to contain a sufficiently large number of hidden units (Hornik, Stinchcombe, and White 1989; 1990). White (1990) states that, as the size of the training set increases and the network acquires more experience, the probability of network approximation error approaches zero. However, in finite samples the asymptotic zero errors do not necessarily mean perfect forecasts.

When working with neural networks, there is no need to make any assumptions about the properties of the distribution of the data (West et al., 1997). In regression models the relationship between dependent and independent variables should be specified with a mathematical function based on past experience or hypotheses. Such specifications are not necessary for neural networks because the network learns complex relationships between inputs (independent variables) and output (dependent variable) through hidden layers (see Appendix 1 for more information about the hidden layer) (Gronholdt and Martensen, 2005). Instead of presupposing some statistical properties of the underlying population, neural nets with at least one hidden layer use the data set to understand the internal relationship between inputs and outputs. They have the ability to ignore the data that is not significant and to focus on the data which is most influential (Shachmurove, 2002). Moreover, the flexibility of neural networks in forming both linear and nonlinear relationships can be counted as an important factor in their superior performance compared to other methods.

In a sample of stock returns on the S&P 500 index from 1960 to 1992, Qi (1999) finds that the nonlinear neural network model not only fits the data better than the linear model, but it also provides fairly accurate out-of-sample forecasts. The recursive neural model used in this study has smaller RMSE, MAE, and MAPE and higher Pearson correlation and percentage of correct signs than the linear model in the whole out-of-sample forecast period. Kuan and Liu (1995) show that nonlinearity in exchange rates may be utilized to improve point and sign forecasts of exchange rates using the data for five exchange rates against the US dollar from 1980 to 1985. Hutchinson et al. (1994) find that linear models exhibit considerably weaker performance than the neural network models where S&P 500 futures options data was used from 1987 to 1991.

Hill and O`Connor (1996) compare time series forecasts based on neural networks with forecasts from traditional statistical time series methods including exponential smoothing, Box-Jenkins and a judgment-based method. According to their results, the neural network model outperforms traditional statistical and human judgment methods when forecasting quarterly and monthly data which is used in the forecasting competition of Makridakis et al. (1982), whereas the results are comparable on annual data. Moreover, the neural network model has almost always lower variance of forecasts than those of the traditional models.

Kuo and Reitsch (1996) test the accuracy of forecasts produced by time series, multiple regression and neural network models by using two data sets first of which consists of 56 months measuring 14 variables and the second of which is time series data where the dependent variable is monthly dollar sale volumes of a tuxedo rental firm. Their results prove that neural networks tend to do better than conventional methods in all cases. They find neural networks especially valuable where inputs are highly correlated, some data is missing or the systems are non-linear.

West et al. (1997) show that neural networks make superior predictions concerning consumer decision processes. Based on a sample of 800 people, they conclude that, when modeling consumer judgment and decision making, neural network models produce significantly better results than traditional statistical methods. The suggested reason for this improvement is their ability to capture nonlinear relationships without making an assumption of a parametric relationship between product attributes, perceptions and behavior. They demonstrate that neural networks are highly promising for improving model predictions in nonlinear decision contexts as well as linear decision contexts. They show that the neural network model outperforms statistical methods in terms of explained variance and out-of-sample predictive accuracy. In addition to these, when predicting consumer choice in nonlinear and linear settings neural networks are again superior to traditional statistical methods.

A very recent study carried out by Giovanis (2008) examines the effects of some factors on greenhouse effects for the fifteen countries of the European Union by using annual data from 1990 to 2004. The factors included are concerning not only gases, but also some economic variables, such as gross domestic product, consumption etc. In this study, the forecasting performance of panel regression analysis is compared with that of neural network modeling and it is found that forecasting performance of neural networks is much better than traditional econometric methods.

The results of these studies are quite promising for the neural networks, nevertheless most of them are specific to a data set. Therefore, it is not possible to claim that neural networks will always perform better than traditional models. Besides, most studies carried out so far compare the traditional econometric models and neural networks; however, they are not necessarily substitutes. It is possible to combine neural networks with regression analysis to generate a much stronger forecasting tool (Kabundi, 2004).

There is another side in the literature claiming that positive results for any new model or approach are always more interesting than negative ones; therefore, studies which show the superiority of neural networks on traditional models tend to be published more (Chatfield, 1995). Chatfield (1993) reexamines the study carried out by Refenes et al. (1993) and finds that the way they compare neural network forecasts with classical smoothing techniques is unfair. He points out that there are some studies (Hoptroff, 1993) which even do not compare neural networks with any alternatives but only mention about successful applications of neural networks and see the black box character of neural networks as an advantage because people with little knowledge or expertise can also make reasonable forecasts. Faraway and Chatfield (1998) think that this is especially dangerous, because without expertise unreasonable results can be obtained and wrong conclusions can be drawn. They show that it is only possible to construct a good neural network model for time series data by combining traditional modeling skills with the knowledge of time series analysis, and at the same time by knowing the problems that are possible to face with in fitting neural network models. The architecture of the network, activation functions and appropriate starting values for the weights should be carefully chosen when working with neural networks.

3- Neural Network Modeling

Neural network models are mathematical imitation of the neurophysical structure and decision making way of the human brain. Although they are closely related to generalized linear models from the statistical point of view; artificial neural networks are nonlinear and use estimation procedures like feed-forward and back-propagation (for details see Appendix 1), while traditional statistical models use least squares or maximum likelihood (West et al., 1997). Artificial neural networks have different names such as connectionist networks, parallel distributed networks or neuromorphic systems.

A neural network is defined as a nonlinear regression function which characterizes the relationship between the dependent variable (*t*, target, output) and n-vector of explanatory variables (*p*, inputs). Instead of forming a nonlinear function, many basic nonlinear functions are combined via a multilayer structure which is shown in Figure 3-9 in Appendix 1 (Kuan and Liu, 1995).

Neural networks are proven to be useful tools when the relation between variables is not known but some examples of inputs and outputs already exist and there is some evidence on a functional relationship between inputs and outputs. To estimate the output, first, some examples of inputs and corresponding outputs (targets) are introduced to the network for training purposes and the network is allowed to generalize the relation between inputs and corresponding outputs. Their only way to learn the relationship between inputs and outputs is to use these introduced examples. Therefore, data collection and deciding about the training set is extremely important in neural network modeling. If a neural network is not provided with a rich set of examples (training set) that show the relation between inputs and outputs from as different aspects as possible, obtaining reliable results may not be possible.

Neural networks have many advantages. One advantage of them is that they can handle incomplete or fuzzy information and are suitable for the cases where generalization or inference is required (Lodewyck and Deng, 1993). Even though the data is partly missing, neural networks are still able to give good results. However, when the performance of the network decreases because of missing data, then it can be concluded that the missing data was important to the network and since the network does not have it, its performance is low. On the contrary, traditional methods have more difficulties in working with missing data. Missing data may even directly result in insignificant results under traditional methods. Another advantage of neural networks is that they have tolerance for error. Even when some neurons have deteriorated and are not able to work, the network can still produce results. However, depending on the importance or the position of that neuron, the performance of the neural network may be lower. Öztemel (2003) claims that artificial neural networks are the best and most powerful means to process missing, unusual or uncertain data.

On the other hand, artificial neural networks have some disadvantages as well. First, there are no certain rules that help the user how to construct a network, how to decide about the learning rate, the number of nodes and hidden layers etc. (Details about neural networks terminology can be found in Appendix 1). It is very important for the user to have some experience in working with neural networks for different problems so that the user knows which activation function to use, which learning rate to set, which topology to use etc. Second, neural networks work only with numerical data. If the data is not numerical, the user has to convert it. The reliable solution to the problem may be impeded when the data is not converted successfully into numerical terms. A third disadvantage is that there are no concrete rules to determine when to stop training the network. When the error is reduced to an acceptable level the user can stop training. However, it cannot be concluded that the network has produced optimum results, but it has produced good results. Last but not least, once a solution to a problem is offered by a neural network the user cannot explain why and how this solution is produced.

Neural networks are spread to a wide range of areas such as finance, medicine, engineering, biology, psychology, statistics, mathematics, business, insurance, and computer science (West et al., 1997). They have become quite popular among economists, mathematicians and statisticians since they do not require any assumptions about population distribution (Shachmurove, 2002).

4- Data

The data used in this study is obtained from IMF`s International Financial Statistics, World Bank`s World Development Indicators 2005, and OECD`s International Trade by Commodity Statistics. The sample period covers 40 years from 1964 to 2003. Countries included are the EU-15 where Belgium and Luxembourg are taken as one country because of data availability. The model is estimated using bilateral trade flows among EU15 countries from 1964 to 2003. For these 15 countries, 91 bilateral trade flows are obtained which cover fixed and flexible exchange rate periods as well as Euro period. The number of total data points analyzed is 3586.

Nominal exports in the data set are converted into export volumes by using GDP deflators. Volatility of exchange rates is calculated as the moving average of standard deviations of the first difference of logarithms (i.e. percentage changes) of quarterly nominal bilateral exchange rates (Kowalski, 2006). $Vol(xr_{ijt})$ is the 5-year ("t-4,...,t") average of standard deviations from the average quarter-on-quarter percentage change in bilateral nominal exchange rate calculated over the last 4 quarters, given by the following formula:

$$Volxr_{ijt} = \frac{1}{20}\sum_{q}^{q-19}\delta_q , \qquad \text{Eq. 1}$$

where q is the last quarter in year t and

$$\delta_q = \sqrt{\sum_{q}^{q-3}\frac{1}{3}\left(de_q - \sum_{q}^{q-3}\frac{1}{4}de_q \right)^2} . \qquad \text{Eq. 2}$$

δ_q is a standard deviation from the average quarter-on-quarter percentage change in bilateral nominal exchange rate calculated over the last 4 quarters where $de_q = e_q - e_{q-1}$ and e_q is a logarithm of bilateral exchange rate at the end of quarter q.

5- A Modified Gravity Model of Bilateral Exports and its Application to Panel Data Analysis

The gravity model is extensively used in international trade literature to analyze international trade flows. The original gravity model explains bilateral trade between two countries with their incomes and distances between them. According to this model, the product of income of both countries affects bilateral trade positively while distance imposes an obstacle to trade. This original gravity model has been extended later on by including population, exchange rates, common language, common borders, foreign currency reserves etc. to better explain the variation in bilateral trade. We insert population of both countries and volatility of exchange rates into the equation and do not take the product of incomes, but take each variable separately into the equation.

The modified gravity model of bilateral trade used is given by:

$$\ln Exp_{ijt} = \beta_0 + \beta_1 \ln D_{ij} + \beta_2 \ln Y_{it} + \beta_3 \ln Y_{jt} + \beta_4 \ln Pop_{it} + \beta_5 \ln Pop_{jt} + \beta_6 vol(xr_{ijt}) + \varepsilon_{ijt}$$

Eq. 3

where *i*=exporter, *j*= importer, and

- Exp_{ijt} represents the volume of exports from country *i* to country *j* in year *t*,
- D_{ij} is the distance between country *i* and country *j* measured in kilometers,
- Y_{it} is the exporting country`s real GDP in year *t*,
- Y_{jt} is the importing country`s real GDP in year *t*,
- Pop_{it} is exporter country`s population in year *t*,
- Pop_{jt} is importer country`s population in year *t*,
- $Vol(xr_{ijt})$ is the volatility of nominal exchange rate between exporter and importer country in year *t*,
- ε_{ijt} is the error term.

Some authors find that as distance becomes larger, bilateral trade between countries tends to decrease (Clark et al.: 2004; Glick and Rose: 2002; Rose et al.: 2000). Furthermore, higher income in the exporting country will have a positive effect on bilateral trade by leading to more production and higher exports (Dell`Ariccia: 1999;

De Grauwe and De Bellefroid: 1986; Cushman: 1983, Balogun: 2007; Clark et al.: 2004; Glick and Rose: 2002; Matyas: 1997; Rose et al.: 2000). For a similar reason, higher income tends to increase the level of imports as well, by forcing countries to import more to be able to consume and produce more.

On the other hand, population has an ambiguous effect on bilateral trade, which is positive for some countries and negative for others (Matyas: 1997; Bergstrand: 1989; Dell`Ariccia: 1999). Moreover, the impact of population on trade may also change depending on the length of the estimation period (short-term vs. long-term). The last variable of interest is the volatility of exchange rates. The expected effect of volatility of exchange rates on bilateral trade is negative, because when the economic environment is not stable, prices are very changeable depending on the fluctuations in exchange rates. As a result, profits of exporters become less predictable and this may reduce bilateral trade (Frank and Bernanke, 2007: 889).

6- A Modified Gravity Model of Bilateral Exports and its Application to Neural Networks

An essential decision for most approximation techniques and particularly for neural networks is about the type and complexity of the model. Different approaches and various network architectures can be used depending on the type of the approximation function used to solve the problem. Even if the user decides on one type of network architecture based on prior knowledge, the question about the appropriate complexity of the architecture remains (Hutchinson et al., 1994). Furthermore, some issues should be clarified very carefully when constructing a neural network such as the number of hidden layers and the number of neurons in each hidden layer etc., because when the model is not constructed properly neural nets can give inferior results (Chatfield, 1997). Additionally, explanatory variables are also of a vital importance. It is a very well known fact that, the success of any model, whether traditional or neural network, relies on the explanatory variables (Church and Curram, 1996).

In this study, bilateral export flows will be analyzed through a panel model and a neural network model. In both models, exactly the same data set and the same explanatory and dependent variables will be used. Equation 3 will be the base for the panel model as well as the neural network model. In panel model, distance, real GDP of the exporting (country *i*) and importing country (country *j*), population of the exporting and importing country and the volatility of exchange rates are regressed on bilateral exports to explain the variation in them. On the other hand, the inputs of the neural network are also the same as the explanatory variables of the panel model (distance, real GDP of the exporting (country *i*) and importing country (country *j*), population of the exporting and importing country and volatility of exchange rates). Neural network model is expected to learn the relationship between these 6 inputs and bilateral exports (the output) through neurons in each layer. The hidden layer which processes and sends the information received from the input layer to the output layer consists of 25 neurons. The output layer has one output neuron namely bilateral exports from country *i* to country *j*. The structure of the neural network model is shown in Figure 3-1 and Figure 3-2. Our network has tan-sigmoid transfer function in the hidden layer and a linear transfer function in the output layer. The role of these transfer functions is explained in Appendix 1 in detail.

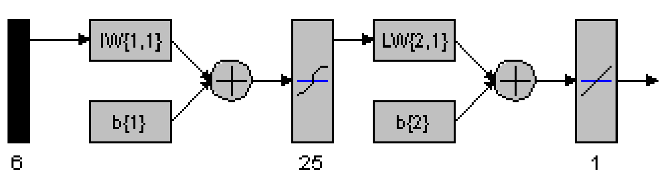


Figure 3 - 1: The multilayer network with 6 inputs, 25 hidden neurons and 1 output

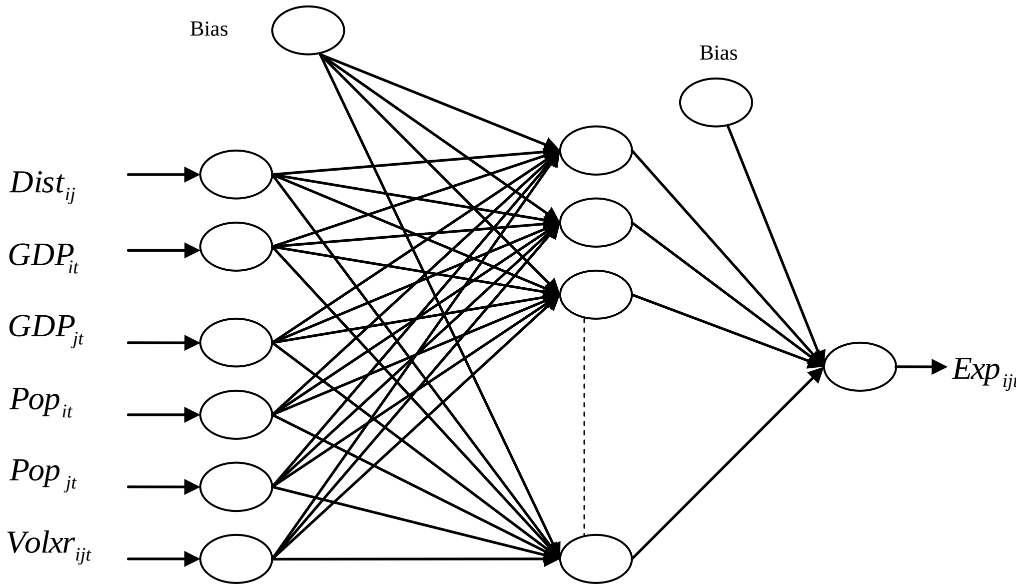


Figure 3 - 2 : The topology of the panel network

7- A Comparison between Panel Data Analysis and Neural Network Model

7.1 Results of Panel Data Analysis

The data set used in this study consists of bilateral export flows, GDPs, population, volatility of exchange rates and distances among the EU-15 countries from 1964 to 2003. For each country pair we have 40 years of data. Our initial attempt is to investigate how trade flows across European countries can be explained by income, population, distance and also volatility of exchange rates. To this aim, we apply panel data analysis on the basis of equation 3.

	Coefficient	Std. Error	t-Statistic	Prob.
C	-18.97	0.75	-25.21	0.00
Distance	-0.84	0.02	-38.50	0.00
Exporter GDP	0.53	0.04	12.20	0.00
Importer GDP	0.90	0.03	27.84	0.00
Exporter Population	0.27	0.04	6.39	0.00
Importer Population	-0.17	0.04	-4.80	0.00
Volatility of Exchange Rate	-0.26	0.07	-3.91	0.00
R-squared	0.84			
Adjusted R-squared	0.84			
AIC	2.14			
Schwarz criterion	2.22			
Number of observations	3586			
Sample Period	1964-2003			

Table 3 - 1: Panel least squares with period fixed effects, dependent variable: log of real bilateral exports

Table 3-1 shows the results of panel data analysis that are in consistency with the international trade theory. The results indicate that as distance becomes larger, bilateral trade between countries tends to decrease. Furthermore, higher income in the exporting country has a positive effect on bilateral trade by leading to more production and higher exports. As Table 3-1 shows, when income of exporting country increases by 1%, its exports increase by 0.53%. For a very similar reason, higher income tends to increase the level of imports as well. According to Table 3-1, a 1% increase in the importing country's real GDP increases its imports by 0.90%.

Additionally, population of the exporting country has a positive effect on bilateral exports. This shows that higher population will create opportunities for specialization which will boost production and exports from that country. The last variable of interest is the volatility of exchange rates. Our results indicate that volatility of exchange rates has a negative effect on real bilateral exports. For all variables that are used to explain the variance in bilateral exports, coefficients are highly significant.

7.2 Results of the Neural Network Model

This section deals mainly with the feed-forward network used to analyze bilateral exports across European countries and their determinants.

Before constructing and training the network, the inputs (distance, GDP and population of exporting and importing country, volatility of exchange rates) and targets (bilateral exports) are normalized so that the mean is zero and variance unity. In some situations, the input vector has a large dimension, but the correlation among components of the vectors is quite high. In this case, it is useful and necessary to reduce the dimension of input vectors. An effective method for performing this operation is principal component analysis (Demuth et al., 2002). When we perform principal components analysis, we see no redundancy in our data set, because the size of input vectors does not reduce.

The objective of model selection is to construct a model with acceptable levels of model bias and variance. Therefore, it is necessary to divide the data set into three subsets: training, validation and test sets. The training set is used to determine the network parameters: weights and biases. Weights show how strongly a signal from one node affects the other node, while bias influences the strength of the effects of inputs on the output. So, the net effect of an explanatory variable or input on the output is calculated as the product of the input and weight, including the impact of the bias (see Appendix 1 for more details). The outcome of the validation set is predicted by using these weights and biases calculated using the examples given in the training set. Then, the network architecture –a combination of weights, biases and the number of hidden neurons- that gives the smallest validation set error is chosen and the network`s performance is evaluated by using the test set (Hung et al., 2002).

The network is trained using early stopping which is a method employed for improving generalization. In this method, while the network is being trained the error on the validation set is observed simultaneously. The training and validation set error normally decreases during the initial stage of training. However, the error on the validation set starts increasing when the network begins to overfit the data. When the validation set error increases for a pre-determined number of iterations, the training must be stopped, and the weights as well as biases which result in minimum validation set error are accepted (Demuth et al., 2002).

After dividing the data set into three subsets, a feed forward network with two layers is created as shown in Figure 3-1 and 3-2. The number of neurons in the hidden layer is selected experimentally. As stated in Hung et al. (2002), too many hidden layers may cause overfitting of the data resulting in a low model bias and a high model variance. It is found that the network with 25 hidden neurons gives quite satisfactory results.

The Levenberg-Marquardt training function is used to train the network. This is a numerical optimization technique and is especially adapted to the minimization of the error. This algorithm is known as the fastest technique for training moderate-sized feed forward neural networks (Demuth et al., 2002). One advantage of this training method is its fast convergence about minimum and its good prediction performance (Koker, 2007). Kisi (2004) provides a comprehensive explanation for this algorithm.

Training results of the network are shown in Figure 3-3. As the figure shows, when the number of epochs increases the errors of all three sets decline. At the beginning of training, the decrease in squared error is very sharp, but then it decreases at a lesser pace. In our neural network, training is stopped after 48 epochs, because at that point

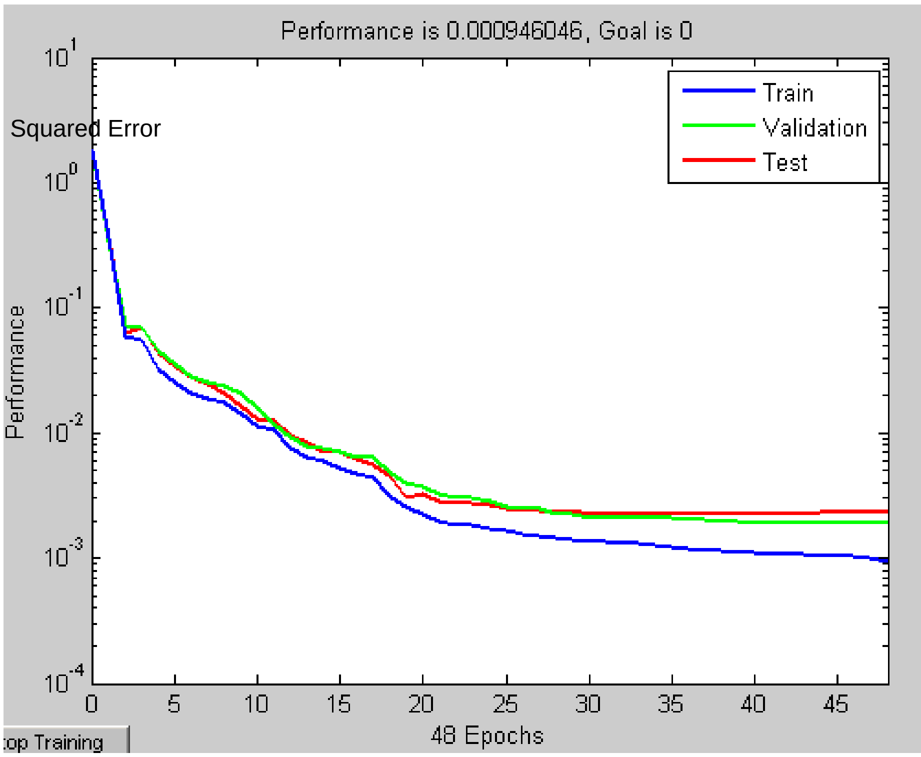


Figure 3 - 3: Training, validation and test set errors

the error for validation set starts increasing. If the model is constructed successfully, the test set and validation set error should show similar characteristics. Figure 3-3 shows that they both follow the same pattern which proves that our model is reliable. The performance of the model is measured with the mean squared error which reduces to 0.00095 after 48 epochs.

The errors on the training, validation and test sets are the first means to obtain some information about the performance of a trained network. However, the network response needs to be measured in more detail. One choice is to carry out a regression analysis between the network response which are outputs produced by the network and the corresponding targets that are actual outputs in the data set (Demuth et al., 2002). Therefore, the next step is to simulate the trained network. Since the targets were normalized before constructing the network so that the mean was 0 and the standard deviation was 1, the network outputs are needed to be transformed back into the original units.

In a linear regression analysis between the network outputs and the corresponding targets, there are two parameters to interpret:

$m = 1.0001$

$r = 0.9871$

The first parameter m stands for the slope of the best linear regression between targets (actual outputs) and network outputs. In the case of a perfect fit, which means network outputs exactly equal to the targets, the slope should be 1. The second statistic r is the correlation coefficient (R-value) between network outputs and targets. It shows how well the variation in network output is explicated by the targets. If this number is equal to 1, then there is perfect correlation between targets and network outputs (Li and Liu, 2005). In our results, the numbers are very close to 1, which implies a very good fit. To compare neural network model with panel least squares we need R-squared for the neural network model which is the square of the correlation between targets and network outputs (r, R-value). R-squared calculated is 0.97. According to this value of R-squared, it is possible to conclude that the neural network model can explain 97% of the variation in bilateral exports with the given 6 inputs.

In Figure 3-4, the network outputs are plotted against the targets. The best linear fit is displayed by a dashed line. The perfect fit, which requires network outputs equal to targets, is shown by the solid line. Here, it is very difficult to differentiate the best linear fit line from the perfect fit line because the fit is very good.

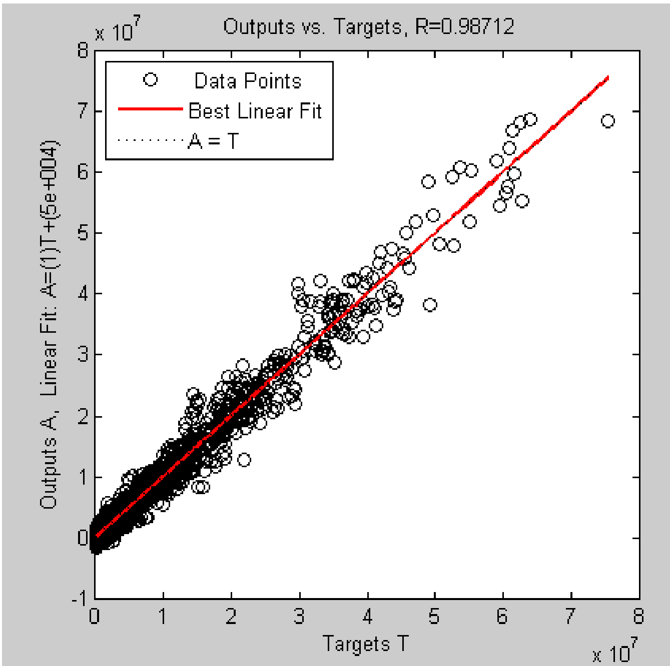


Figure 3 - 4: Outputs produced by the neural network versus targets

7.3. Forecasting Performance of the Neural Network and of the Panel Model

In this section, the performance of the panel model will be compared with that of neural network model in out-of-sample forecasting. To perform this, a panel least squares model is estimated with period fixed effects using the data from 1964 to 1993 and this model is used to forecast bilateral export volumes from 1994 to 2003. The estimation output and the results of panel model forecasting are shown in Table 3-2 and 3-3 respectively. A neural network model calculates MSE based on standardized data, while the panel model errors are based on non-standardized data. To make a fair

	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.33	0.05	-6.20	0.00
Distance	-0.30	0.01	-30.22	0.00
Exporter GDP	0.22	0.03	6.84	0.00
Importer GDP	0.68	0.03	27.11	0.00
Exporter Population	0.30	0.03	9.52	0.00
Importer Population	-0.18	0.02	-7.65	0.00
Volatility of Exchange	-1.07	0.18	-6.02	0.00
R-squared	0.82			
Adjusted R-squared	0.82			
Akaike info criterion	1.12			
Schwarz criterion	1.20			
Number of observations	2704			
Sample Period	1964- 1993			

Table 3 - 2: Panel Least Squares Model with Period Fixed Effects, Dependent Variable: Log of Real Bilateral Exports

	Log(export volume)
Forecast sample	1994-2003
Included observations	903
Root Mean Squared Error	2.26
Mean Squared Error	5.10
Mean Absolute Error	1.44
Mean Absolute Percentage Error	687.91

Table 3 - 3: Forecast Results of Panel Least Squares Model with Period Fixed Effects

comparison, the whole dataset was standardized before running the regression and the forecasts were obtained. The results shown in Tables 3-2 and 3-3 indicate that MSE of panel model is 5.1.

On the other hand, the data set that will be used in the neural network model is divided into two parts. The first part consists of the years from 1964 to 1993 and is used to construct and train the neural network. The second part which includes the data from 1994 to 2003 is used to test the constructed neural network in terms of mean squared error and R-squared, and to make a comparison between neural network and panel data forecasting.

After constructing and training the network, mean squared errors are obtained for the three subsets of the first part of the data set (years from 1964 to 1993). MSE is calculated as the average squared difference between normalized network outputs and targets. Table 3-4 shows the mean squared error and R-value for training, validation and test sets.

Since 25% of the whole sample is employed as validation set and the other 25% as test set, 1341 data points out of 2683 are used for training the network. The

	Number of Samples (Total: 2683)	**MSE**	**R-value**
Training	1341	1.25697e-3	0.99
Validation	671	1.95853e-3	0.98
Test	671	1.98984e-3	0.98

Table 3 - 4: Results of Training

regression R-value shown in Table 3-4 measures the correlation between unnormalized network outputs and targets for each subset. It is seen that for each subset we have satisfactorily high R- values. Figures 3-5, 3-6 and 3-7 plot these R-

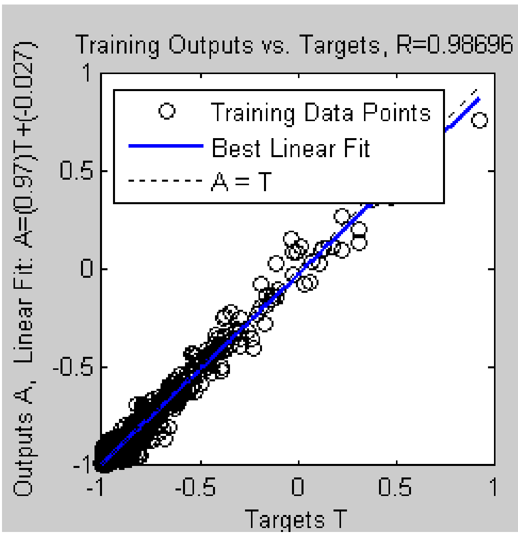
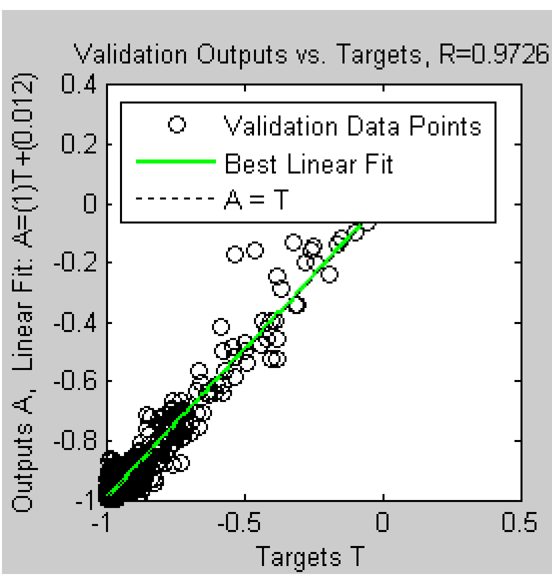


Figure 3 - 5: Training Set Outputs versus Targets



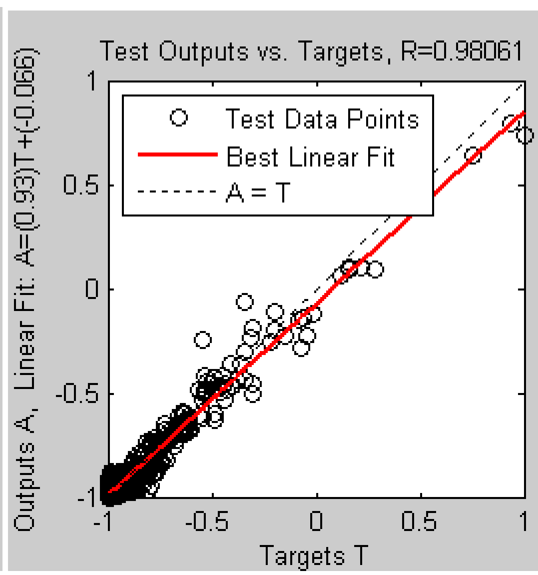


Figure 3 - 6: Validation Set Outputs versus Targets **Figure 3 - 7:** Test Set Outputs versus Targets

values for training, validation and test set respectively. When the network is trained with a different number of hidden neurons, for example with 10 hidden neurons, the results do not change enormously but very slightly.

After constructing the neural network by using the data from 1964 to 1993, we use the rest of the data set, which is from 1994 to 2003, to test the network's performance. If the network has learned the relationship between inputs and outputs appropriately, it is expected to perform well when a new data set is introduced to it. The performance of a neural network is measured as follows: The network determines the weights and biases by using the training set chosen from the observation points between the years 1964-1993. Then, the neural network is asked to produce its own outputs that are bilateral export volumes for the years from 1994 to 2003, for given input values for the same period (GDP and population of exporting and importing country, distance and volatility of exchange rates). Then, actual/observed values of bilateral export volumes from 1994 to 2003 will be compared with the ones that neural network produces using its own weights and biases. MSE is computed by comparing the network`s outputs with observed values of bilateral exports from 1994 to 2003.

When the test set which has 903 observations - from 1994 to 2003- is introduced to the neural network, it yields the results summarized in Table 3-5.

	Number of Samples	**MSE**	**R-value**
Test Set	903	2.11193e-2	0.9176

Table 3 - 5: The Test Results of the Neural Network Model

R-squared for the test set is calculated as the square of the R-value in Table 3-5 and it is 0.84. A comparison between R-squared produced by neural networks (0.84) and panel model (0.82) reveals that neural network modeling offers slightly higher explanatory power. Figure 3-8 plots test set outputs against test set targets.

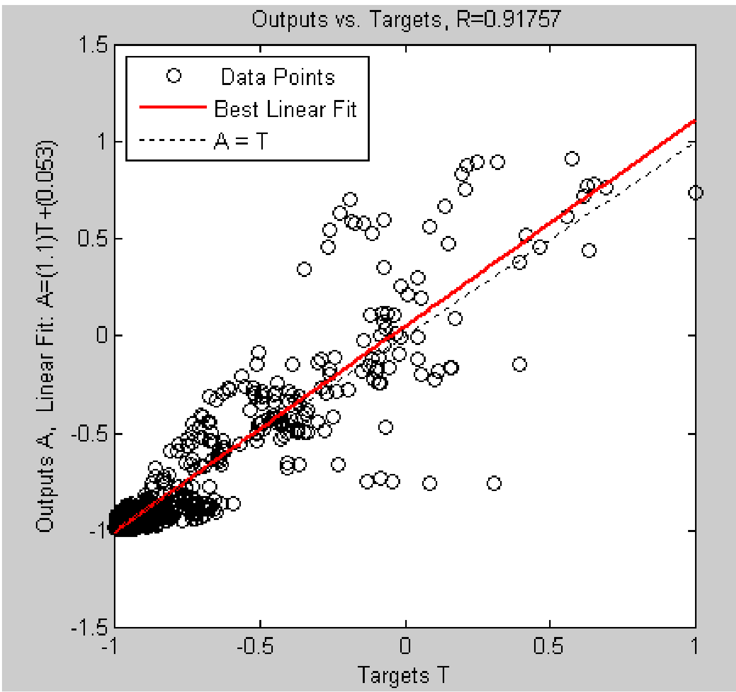


Figure 3 - 8: Test Set Outputs versus Targets

When the MSE of neural network model -which is 0.0211- is compared to the MSE produced by panel data forecasting -which is 5.1- it is seen that neural networks offer a much lower MSE. The reason why there is so much difference between neural network and panel model forecasts is that the regression model uses logged data set while inputs given into the neural network are non-logged. Therefore, MSE of panel model is recomputed for logged data in terms of non-logged data. It is done as follows: First, forecasts are computed in logs and then are transformed using the exponential into the forecasts of the variable without logs. Lastly, MSE is computed based on these predictions, which is 2.97. Our analysis reveals that there is still a large discrepancy between the MSE produced by the neural network model (0.0211) and panel model (2.97) which suggests that neural networks lead to much better out-of-sample forecasting performance than the panel model.

7.4. Panel Model Forecast Errors with Different Out-of-Sample Periods

Since our model is a static model, it is not really necessary to choose the end of the sample period as the out-of-sample period. To extend the comparison made, different periods have been chosen as the forecast sample, and the average of forecasting errors were calculated. Since we have 40 years of data and one-fourth of it is used to test the neural network; panel model uses 30 years of the data set for estimation and 10 years of it as the forecast period. Table 3-6 indicates the forecast errors which are obtained by randomly cutting the whole sample into in- and out-of-samples and performing the out-of-sample predictions on this division by using panel model.

It is seen that when the end of sample is chosen as the forecast sample, a very high MSE is obtained in panel data analysis (2.97). When the forecast sample is changed, the errors decrease enormously almost in all cases; however, the difference between errors produced by neural networks (0.02) and panel data analysis (on average 0.32) remains still high which shows that neural networks lead to high improvements in prediction.

Estimation Sample	Prediction Sample	Number of observations	RMSE	MSE
1964-1980 and 1991-2003	1981-1990	910	0.34	0.12
1964-1970 and 1981-2003	1971-1980	910	0.46	0.22
1964-1983 and 1994-2003	1984-1993	884	0.33	0.11
1964-1973 and 1984-2003	1974-1983	910	0.42	0.17
1974-2003	1964-1973	910	0.55	0.31
1964-1990 and 2001-2003	1991-2000	884	0.37	0.14
1964-1985 and 1996-2003	1986-1995	884	0.33	0.11
1964-1988 and 1999-2003	1989-1998	884	0.35	0.13
1964-1978 and 1989-2003	1979-1988	910	0.36	0.13
1964-1972 and 1983-2003	1973-1982	910	0.43	0.19
1964-1974 and 1985-2003	1975-1984	910	0.40	0.16
1964-1966 and 1977-2003	1967-1976	910	0.53	0.28
1964-1966 and 1971-1997	1967-1970 and 1998-2003	903	1.29	1.65
1964-1966 and 1972-1998	1967-1971 and 1999-2003	903	1.06	1.13
1964-1975, 1981-1992 and 1998-2003	1976-1980 and 1993-1997	910	0.38	0.15
1964-1980, 1986-1997 and 2003-2003	1981-1985 and 1998-2002	910	0.37	0.14
			Average:	0.32

Table 3 - 6: Forecast Errors of Panel Model for Different Samples

8- Conclusion

This study compares the results produced by panel data analysis and neural networks. Both models give satisfactory results which show that our modified gravity model of bilateral trade can well explain the variation in bilateral exports among European countries from 1964 to 2003. Balanced panel estimates have the advantage of explaining the individual effect of each independent variable on bilateral exports and showing whether this effect is significant or not. R-squared given by panel data analysis is 84%, which shows that 84% of the variation in bilateral exports can be explained by distance, GDP and population of exporting and importing countries, and the volatility of exchange rates. Then, we construct a neural network by using the same independent and dependent variables. We find that 97% of the variation in bilateral exports can be explained with these variables using a neural network model. Neural networks seem superior to traditional panel data analysis in explaining bilateral exports when R-squared is used as a criterion.

When we make out-of-sample forecasting by employing a panel model with period fixed effects, and use the period that is forecasted in panel model as the test set of our neural network, we see that neural networks produce a much lower MSE (0.02) which makes them superior to the panel model with an MSE of 2.97.

However, neural networks lack the ability of showing the individual effect of each influence factor on bilateral exports. Once we construct the neural network, it learns the relationship between bilateral exports and their determinants internally through hidden layers by using the examples introduced to the network, although it does not explain what kind of a relationship it is exactly and how the network can explain it. This does not mean that the results produced by the network are not reliable, because we have some tests that give a signal whether the model and the network constructed is appropriate to draw some conclusions such as mean squared error and regression analysis results. One of the main relative benefits of the neural network model is nonlinearity, as it uses sigmoid functions instead of linear functions as building blocks. This partly explains its success in our exercise.

9. Appendices

Appendix 1

A Neuron Structure

A simple neuron consists of six elements which are shortly described below:

1- Inputs (p): Information from the environment or from other neurons that come into the system to be processed. For example, inputs of our neural network model are distance, GDP and population of exporting and importing country, and volatility of exchange rates as shown in Figure 3-2.

2- Weights (w): Inputs have an effect on the output depending on their importance. Weights show how strongly a signal from one node influences another node by entering into the activation equation of it (Anderson, 1999). Weights are especially important for neural networks because the intelligence of a neural network depends on how true its weights are. The knowledge of a network is stored in the connection weights. A neural network is used to compute a function by forcing input units to produce given outputs (Kryzanowski et al., 1993). By modifying the strength of the connections among neurons, which are called weights, the network learns the relationship between inputs and outputs, and by using these weights it can predict the value of the dependent variable for other given independent variables (Anderson, 1999).

3- Bias (b): The product of inputs and weights is affected by a bias and the effects of the inputs may be reduced through the bias. The bias has a constant input of 1.

4- Sum function: This function calculates the net input into the neuron with the following formula:

$$NET = \sum_{i}^{n} p_i w_i + b_i \qquad \text{Eq. 1}$$

where *i* is the input neuron, *n* is the total number of neurons, *p* is the input and, *w* is the weight and *b* is the bias.

5- Transfer function (Activation function): Takes the net input into the neuron and produces output *a*. Three of the most commonly used transfer functions are hard limit, pure linear and log-sigmoid functions. In multilayer networks mostly the log-sigmoid function is used because it is differentiable.

6- Output (a): The value produced by the activation function. This output is either sent to the environment or to another neuron as an input.

Learning Rule

A learning rule is defined by Demuth et al. (2002) as a process for adjusting the weights and biases of a network. At the beginning of this process, a set of examples, called the training set, is introduced to the learning rule, such as

$$[p_1, t_1], [p_2, t_2], \ldots.. [p_n, t_n]$$

where *p* is the input, *t* is the accompanying target output and *n* is the total number of observations in the training set. When the inputs are provided to the network, the network is expected to produce its own outputs to be compared to the targets. To do this, the weights and biases of the network are determined using the training set and modified by using the learning rule to move network outputs closer to the targets. At the end, the outputs produced by the network and real outputs are compared to see whether the network has learned the relationship appropriately.

The objective is to reduce the error $e=t-a$, where a is output produced by the network for a given input and t is the target (actual/observed output). The performance of a network is measured by MSE which is equal to squared error (e). If there is no difference between network outputs and targets, no learning takes place (Demuth et al., 2002; Rumelhart, 1986).

Multi Layer Networks

When the relation between inputs and outputs is not linear, a single layer network cannot find a solution. In this case, we need more developed models of networks. Multilayer networks are good solutions to the nonlinear cases.

As shown in Figure 3-9, a feed forward multi-layer network has three layers:

Input layer: This layer receives inputs from the environment and sends them to the hidden layer without processing. Depending on the nature of the problem, there may be more than one input. Each process element (neuron) must have at least one input and one output.

Hidden layer: Information received from the input layer is processed in the hidden layer and sent to the output layer. Multi-layer networks may have more than one hidden layer and more than one neuron in each layer. Hidden layers are expected to improve the networks ability to model complex data, including allowing piecewise approximation (Hill and O'Connor, 1996). The number of hidden layers and the number of nodes in each hidden layer can be selected randomly; however, when making this decision it should be taken into account that too many nodes in the hidden layer may result in a neural network that only memorizes the input data and produces outputs without having any ability to generalize (Shachmurove, 2002).

Output layer: In this layer, information received from the hidden layer is processed and a network output to each given input is produced. Multilayer networks use supervised learning method which is explained under "learning rule" in the previous page.

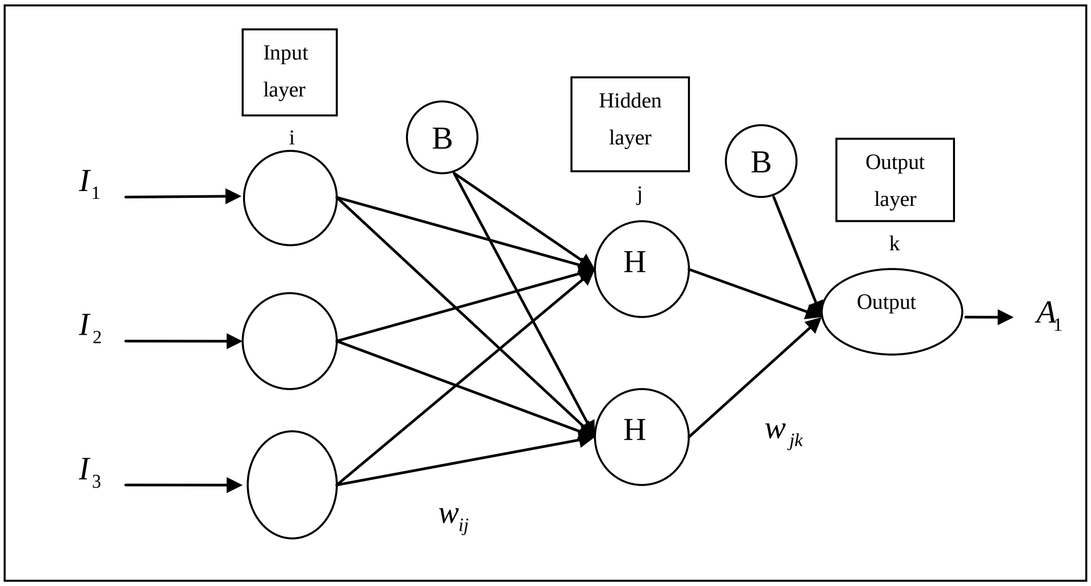


Figure 3 - 9: A Multi Layer Feed Forward Network Structure (Source: Köker, 2007)

Training a Neural Network

Neural networks can be trained for function approximation, pattern association or pattern classification (Demuth et al., 2002). Training in an MLP network requires two steps, forward and backward operation (Köker, 2007). During the first phase – forward pass-, the network is introduced with inputs and propagated forward to produce its own outputs using current weights. Then, outputs of the neural network and targets are compared, and error is calculated for each output. At the second stage, called backward pass, the weights and biases of the network are repeatedly changed to minimize the network performance function that is measured by the mean squared error, the average squared error between network outputs and the targets (Demuth et al., 2002; Kröse and van der Smagt, 1996). As the number of training cases increases, the network is expected to generate more reliable results, and have a better and “more educated” generalization capability (Lodewyck and Deng, 1993).

The back-propagation estimation algorithm used by neural networks is one of the differentiating characteristics between neural network models and traditional statistical procedures. As mentioned above, the back-propagation algorithm is a gradient search technique where the objective function is to minimize the squared error between an observed output (target) and the computed output of the network using given input values. A primary difference between the back-propagation and traditional statistical methods is that the back-propagation algorithm can sequentially consider data records, readjusts the parameters after each observation in a gradient search manner; whereas estimation procedures such as maximum likelihood and least squares use an aggregated error across the entire sample in the estimation (West et al., 1997).

Appendix 2

MATLAB Code for feed-forward multilayer neural network model training and simulation

```
fixunknowns(data);

t=ans(:,2);

p=ans(:,3:8);

p=p';

t=t';

[pn,pp1] = mapstd(p);

[ptrans,pp2] = processpca(pn,0.001);

[tn,tp] = mapstd(t);
```

```
[R,Q] = size(ptrans); iitst = 2:4:Q;

iival = 3:4:Q; iitr = [1:4:Q 4:4:Q];

vv.P = ptrans(:,iival); vv.T = tn(:,iival);

vt.P = ptrans(:,iitst); vt.T = tn(:,iitst);

ptr = ptrans(:,iitr); ttr = tn(:,iitr);

net = newff(minmax(ptr),[25 1],{'tansig' 'purelin'},'trainlm');

net.trainParam.show = NaN;

[net,tr]=train(net,ptr,ttr,[],[],vv,vt);

plot(tr.epoch,tr.perf,'r',tr.epoch,tr.vperf,':g',tr.epoch,tr.tperf,'
-.b');

legend('Training','Validation','Test',-1);

ylabel('Squared Error')

an = sim(net,ptrans);

a = mapstd('reverse',an,tp);

[m, r] = postreg(a,t);
```

10- References

Anderson Philip (1999) "Complexity Theory and Organization Science". Organization Science 10(3): 216-232. Special Issue: Application of Complexity Theory to Organization Science.

Balogun Emmanuel Dele (2007) "Effects of Exchange Rate Policy on Bilateral Export Trade of WAMZ Countries". MPRA Paper: 6234.

Bergstrand Jeffrey H. (1989) "The Generalized Gravity Equation, Monopolistic Competition, and the Factor-proportions Theory in International Trade". The Review of Economics and Statistiscs 71(1): 143-153.

Chatfield Chris (1993) "Neural networks: Forecasting Breakthrough or Passing Fad?". International Journal of Forecasting 9: 1-3.

________ (1995) "Positive or Negative?". International Journal of Forecasting 11: 501-502.

________ (1997) "Forecasting in the 1990s". The Statistician 46(4): 461-473.

Church Keith B. and Stephen P. Curram (1996) "Forecasting Consumers' Expenditure: A Comparison between Econometric and Neural Network Models". International Journal of Forecasting 12: 255-267.

Clark Peter, Natalia Tamirisa, Shang-Jin Wei, Azim Sadikov and Li Zeng (2004) "Exchange Rate Volatility and Trade Flows - Some New Evidence". International Monetary Fund.

Cushman David O. (1983) "The Effects of Real Exchange Rate Risk on International Trade". Journal of International Economics 15: 45-63.

De Grauwe Paul and De Bellefroid (1986) “Long Run Exchange Rate Variability and International Trade”. Chapter 8 in S. Arndt and J.D. Richardson (eds), Real Financial Linkages Among Open Economies. The MIT Press: London.

Dell`Ariccia Giovanni (1999) “Exchange Rate Fluctuations and Trade Flows: Evidence from the European Union”. IMF Staff Papers 46(3): 315-334.

Demuth, H., M. Beale and M. Hagan (2002) Neural Network Toolbox User’s Guide for Use with MATLAB. Version 4. The Mathworks, Natick, MA.
http://www.mathworks.com/support/product/NN/ Accessed Dec. 15, 2009.

Faraway Julian and Chris Chatfield (1998) “Time Series Forecasting with Neural Networks: A Comparative Study Using the Airline Data”. Applied Statistics 47 (2): 231-250.

Frank Robert H. and Ben S. Bernanke (2007) Principles of Economics. Mc Graw Hill Irwin.

Giovanis Eleftherios (2008) “A Panel Data Analysis for the Greenhouse Effects in Fifteen Countries of European Union”. MPRA Paper: 10321, Online at http://mpra.ub.uni-muenchen.de/10321/.

Glick Reuven and Andrew K. Rose (2002) “Does a Currency Union Affect Trade? The Time-series Evidence”. European Economic Review 46: 1125 – 1151.

Grønholdt Lars and Anne Martensen (2005) “Analysing Customer Satisfaction Data: A Comparison of Regression and Artificial Neural Networks”. International Journal of Market Research 47(2).

Hill Tim, Marcus O'Connor and William Remus (1996) “Neural Network Models for Time Series Forecasts”. Management Science 42(7): 1082-1092.

Hornik Kurt, Maxwell Stinchcombe and Halbert White (1990) “Universal Approximation of an Unknown Mapping and its Derivatives Using Multilayer Feedforward Networks”. Neural Networks 3(5): 551-560.

Hornik Kurt, M. Stinchcombe and Halbert White (1989) “Multilayer Feedforward Networks are Universal Approximators”. Neural Networks 2: 359-368.

Hung M. S., M. Shanker and M. Y. Hu (2002) “Estimating Breast Cancer Risks Using Neural Networks”. The Journal of the Operational Research Society 53(2): 222-231. Part Special Issue: The Process of OR.

Hutchinson James M., Andrew W. Lo and Tomaso Poggio (1994) “A Nonparametric Approach to Pricing and Hedging Derivative Securities via Learning Networks”. The Journal of Finance 49(3) :851-889. Papers and Proceedings Fifty-Fourth Annual Meeting of the American Finance Association: Boston, Massachusetts, January 3-5.

Kabundi Alain N. (2004) “Using Artificial Neural Networks to Forecast Economic Variables”. CTSA Conference Proceedings, USA.

Kisi Ozgur (2004) “Multi-layer Perceptrons with Levenberg-Marquardt Training Algorithm for Suspended Sediment Concentration Prediction and Estimation”. Hydrological Sciences–Journal–des Sciences Hydrologiques 49(6): 1025-1040.

Kowalski Przemyslaw (2006) “The Impact of the Economic and Monetary Union in the EU on International Trade- A Reinvestigation of the Exchange Rate Volatility Channel”. PhD Thesis submitted at the University of Sussex.

Köker Rasit, Necat Altinkök and Adem Demir (2007) “Neural Network Based Prediction of Mechanical Properties of Particulate Reinforced Metal Matrix Composites Using Various Training Algorithms”. Materials and Design 28: 616-627.

Kröse Ben and Patrick van der Smagt (1996) An Introduction to Neural Networks. The University of Amsterdam.

Kryzanowski Lawrence, Michael Galler and David W. Wright (1993) “Using Artificial Neural Networks to Pick Stocks”. Financial Analysts Journal 49(4): 21-27.

Kuan Chung-Ming and Tung Liu (1995) “Forecasting Exchange Rates Using Feedforward and Recurrent Neural Networks”. Journal of Applied Econometrics 10(4): 347-364.

Kuo C. and A. Reitsch (1995-96) “Neural Networks vs. Conventional Methods of Forecasting”. Journal of Business Forecasting Methods and Systems 14(4): 17-22.

Li Shouju and Yingxi Liu (2005) “Parameter Identification Procedure in Groundwater Hydrology Artificial Neural Network”. in D. S. Huang, X. P. Zhang and G. B. Huang (Eds): Advances in Intelligent Computing, ICIC Part II, LNCS 3645.

Lodewyck Ronald W. and Pi-Sheng Deng (1993) “Experimentation with a Back-propagation Neural Network, An Application to Planning End User System Development”. Information & Management 24: 1-8.

Makridakis, S., A. Anderson, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen and R. Winkler (1982) “The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition”. Journal of Forecasting 1: 111-153.

MATLAB@ Version 7.4.0.287, The MathWorks, Inc. R2007a.

Matyas Làszlò (1997) “Proper Econometric Specification of the Gravity Model”. World Economy 20(3): 363-368.

Öztemel Ercan (2003) Yapay Sinir Aglari. Papatya Yayincilik: Istanbul.

Qi Min (1999) “Nonlinear Predictability of Stock Returns Using Financial and Economic Variables”. Journal of Business and Economic Statistics 17(4): 419-429.

Rose Andrew K., Ben Lockwood and Danny Quah (2000) “One Money, One Market: The Effect of Common Currencies on Trade”. Economic Policy 15(30): 7-45.

Rumelhart, D. E. and J. L. McClelland (1986) “Parallel Distributed Processing: Explorations in the Microstructure of Cognition”. The MIT Press.

Shachmurove Y. (2002) “Applying Artificial Neural Networks to Business, Economics and Finance”. CARESS Working Papers, University of Pennsylvania.

Tinbergen Jan (1962) Shaping the World Economy: Suggestions for an International Economic Policy. Twentieth Century Fund: New York.

West Patricia M., Patrick L. Brockett and Linda L. Golden (1997) “Comparative Analysis of Neural Networks and Statistical Methods for Predicting Consumer Choice”. Marketing Science 16(4): 370-391.

White Halbert (1990) “Connectionist Nonparametric Regression: Multilayer Feedforward Networks Can Learn Arbitrary Mappings”. Neural Networks 3: 535-549.

CONCLUSIONS

This thesis analyzes bilateral trade flows across EU15 countries from 1962 to 2003 by employing a modified gravity model of total trade to see the effects of exchange rate volatility on bilateral trade. Throughout the thesis, an extended version of the gravity model is used where real GDP and population of exporting and importing countries, distances between them and volatility of exchange rates are the influence factors and bilateral trade flows is the dependent variable.

In the first essay, the effects of exchange rate volatility on bilateral trade between EU15 countries is investigated by using cross sectional methods and a fuzzy approach with the data set from 1964 to 2003. Panel data analysis with period fixed effects shows a significant negative impact of exchange rate volatility on bilateral trade. Then, the same analysis is conducted by using a fuzzy approach and it is seen that the fuzzy approach delivers a very similar result to the panel model. The objective of using a panel data analysis and a fuzzy approach in the same study is to make a robustness check for the fuzzy approach. Even though the use of statistical methods is essential to see the effect of each influence factor on bilateral trade individually, and to see whether this effect is significant or not; it is obvious that when the data set is not large enough, or some data is missing or not reliable, we cannot use panel data analysis and cannot draw any conclusions by using it. In such cases, we suggest the fuzzy approach as a complement to traditional statistical methods to estimate first approximate results.

In the second essay, two different specifications of the extended gravity model of total trade are used. The first specification, called main effects or the 3-way model, is an extension of the modified gravity model with exporter, importer and time effects. The second model is called bilateral interaction effects or the 2-way model. By using these two models and also three more models to serve as a comparison, we analyze the variation in bilateral exports from 1962 to 2003 among EU15 countries.

Moreover, we investigate by how much does the explanatory power of the extended gravity model of total trade increase when main effects and bilateral interaction effects models are used. Lastly, we compare the explanatory power and forecasting performance of the 3-way and 2-way models. It is seen that bilateral interaction effects or the 2-way model outperforms the 3-way model and all other models in estimating bilateral export flows. Furthermore, the 2-way model shows a better performance in out-of-sample forecasting than the 3-way model.

In the third essay, the variation in bilateral exports among European countries from 1964 to 2003 is analyzed via a panel data analysis and a neural network model. When the explanatory powers of both models are compared, it appears that neural networks can explain a larger variation in bilateral exports (97%) compared to the panel data analysis (84%). Although their explanatory power is lower in this exercise, balanced panel estimates have the advantage of explaining the individual effect of each independent variable on bilateral exports and showing whether this effect is significant or not. However, the neural network model does not show the individual effect of each influence factor on bilateral exports, it only explains whether there is a relationship between given inputs (explanatory variables) and the output in total, and shows how much of the variation in bilateral exports can be explained by given inputs. When we compare out-of-sample forecasting performances of panel model and neural networks, we see that neural networks produce a much lower MSE which makes them superior to the panel model. One of the main relative benefits of the neural network model is nonlinearity, as it uses sigmoid functions instead of linear functions as building blocks. This partly explains its success in our study. Another advantage of neural networks is that they make no a priori assumptions about the population distribution and the relationship between explanatory variables and the dependent variable.

Although the results given by panel data analysis, fuzzy approach and neural network model sometimes differ in quantity, the extended gravity model used throughout the thesis proves to be a very good model in explaining bilateral trade flows across European countries. The differences in the results across models can be attributed to different ways of calculating the relationship between explanatory variables and the dependent variable. Furthermore, this thesis serves as a study which combines three different methods for analyzing the same problem, which is the variation in bilateral trade flows among European countries and the effects of exchange rate volatility on bilateral trade flows. It shows that fuzzy logic and neural networks can be employed in economics sometimes as an alternative and yet sometimes as a complement to the statistical methods, and may lead to satisfactory results which are in accordance with the international trade literature.

ZUSAMMENFASSUNG

Diese Arbeit besteht aus drei Teilen (Kapiteln), die die bilateralen Handelsströme zwischen den europäischen Ländern von 1962 bis 2003 mit einem modifizierten Gravitationsmodell schätzen und erklären. Das grundlegende Gravitationsmodell erklärt die bilateralen Handelsströme zwischen zwei Ländern mit ihrem Einkommen und der Entfernung zwischen ihnen. Nach diesem Modell treiben die Länder voraussichtlich mehr Handel, wenn sie reicher werden. Außerdem neigen sie dazu, weniger zu handeln, wenn sie weiter entfernt vom anderen Handelspartnern sind, weil größere Entfernungen zu zusätzlichen Kosten für den Handel (wie höhere Transportkosten, kulturelle Barrieren usw.) führen.

In dieser Arbeit wurde das Gravitationsmodell mit der Bevölkerungszahl der exportierenden und importierenden Länder und dem bilateralen Wechselkurs zwischen ihren Währungen erweitert. Unsere Resultate zeigen, dass die Bevölkerung eine erhebliche Auswirkung auf die bilateralen Handelsströme hat. Außerdem beeinflussen die realen Wechselkurse voraussichtlich auch die bilateralen Handelsströme. In Zeiten höherer Wechselkursschwankungen zeigen die Einnahmen der Händler auch einige Schwankungen. Deswegen ist zu erwarten, dass die Volatilität der Wechselkurse die Höhe der bilateralen Handelsströme verringert. Im zweiten Kapitel haben wir Haupteffekte-und bilaterale Interaktionseffekte in das modifizierte Gravitationsmodell inkludiert. Dieses Modell besteht dann aus Exporteur-, Importeur- und Zeiteffekten. Exporteur- und Importeur-Effekte werden verwendet, um die Tendenz zum Exportieren und Importieren zu erfassen, während Zeiteffekte die zyklischen Veränderungen (Konjunkturzyklen) und deren Auswirkungen auf bilaterale Exporte kontrollieren. Das bilaterale Interdependenz-Modell versucht, die Auswirkungen der geographischen, politischen, historischen oder kulturellen Ereignisse zu erkennen, die den bilateralen Handel zwischen den beiden Ländern beeinträchtigen könnten. Beide Modelle erklären die Variation der

bilateralen Handelsströme besser als das modifizierte Gravitationsmodell, welches in der gesamten Arbeit als Grundmodell dient.

Diese Arbeit hat zwei Hauptziele. Eines ist, die Auswirkungen von Wechselkursschwankungen auf die bilateralen Handelsströme zu erkennen. In der internationalen Fachliteratur, einige Studien finden eine signifikante negative Wirkung, während andere keinen Zusammenhang zwischen der Volatilität des Wechselkurses und den gesamten Handelsströmen aufzeigen. Beim Start dieser Studie war unsere Erwartung, dass die Wechselkurs-Volatilität sich negativ auf die bilateralen Handelsströme auswirkt. Es wurde angenommen, dass die Schwankungen bei den Wechselkursen die Erträge der Exporteure und Importeure weniger vorhersehbar machen, und sie dazu verleiten könnten, sich vorsichtiger zu verhalten. Als Folge neigen sie in einem instabilen Umfeld dazu, weniger Handel zu treiben, um Risiko zu vermeiden.

Unser zweites Ziel ist die Paneldatenanalyse mit Fuzzy-Logik und neuralen Netzen zu vergleichen. Obwohl wir denken, dass die Paneldatenanalyse und statistische Methoden die zuverlässigsten Ergebnisse liefern und nachweislich die besten Mittel für die Analyse der bilateralen Handelsströme sind, verwenden wir Fuzzy-Logik und neuronale Netze als alternative Methode, um die bilateralen Handelsströme zu analysieren und ihre Performanz mit der Paneldatenanalyse zu vergleichen.

Im ersten Artikel konstruieren wir eine Fuzzy-Regel und eine Fuzzy-Entscheidungs-Tabelle, um die Auswirkungen der Wechselkursschwankungen auf den bilateralen Handel zu berechnen, nach einer Regression für Panel-Daten, in der wir aussagekräftige Ergebnisse erhalten. Wir prüfen die Robustheit für den Fuzzy-Ansatz, obwohl wir eine sehr große Datenmenge haben, um zu sehen, ob wir die Fuzzy-Logik in anderen Fällen verwenden können, wo es unzureichende Daten gibt. Wir glauben, dass unsere Ergebnisse mit der Fuzzy-Logik vielversprechend sind. Im zweiten Artikel werden nur statistische Methoden verwendet. Unser modifiziertes Gravitationsmodell wurde durch die Einführung der exportierenden und importierenden Landes sowie Konjunktureffekte erweitert. Dieses Modell wird 3-Wege- oder Haupteffekt-Modell genannt. Eine weitere Erweiterung wurde mit bilateraler Interaktion und Konjunktureffekten durchgeführt und wird 2-Wege-Modell genannt. Wir vergleichen die Aussagekraft und Vorhersageperformanz der beiden Modelle. Das bilaterale Interaktions-Modell erweist sich sowohl bei der Aussagekraft als auch bei der Vorhersage der bilateralen Handelsströme als besser als das Haupteffektmodell. Im dritten Artikel vergleichen wir die Paneldaten-Analyse mit einem neuronalen Netzwerk-Modell. Während die beiden Modelle zu sehr ähnlichen und zufriedenstellenden R^2-Werten führen, scheint das neuronale Netz-Modell überlegen bei der out-of-sample-Prognose.

Auch wenn wir drei verschiedene Methoden in der gesamten Arbeit verwenden, ist das modifizierte Gravitationsmodell das Grundmodell, das wir anwenden. Natürlich ist es nicht verwunderlich, dass verschiedene Methoden zu numerisch unterschiedlichen Ergebnissen führen. Doch die Ähnlichkeit zwischen den Ergebnissen aus den drei Methoden und der Vergleich mit der vorhandenen Literatur zeigt, dass das modifizierte Gravitationsmodell sehr erfolgreich bei der Erklärung der bilateralen Handelsströme ist. Die Unterschiede in den Ergebnissen aus den drei Methoden können auf die verschiedenen Berechnungsmethoden zurückgeführt werden. Diese Arbeit dient als eine Studie, die drei verschiedene Methoden für die Analyse des gleichen Problems, nämlich die Variation der bilateralen Handelsströme zwischen den europäischen Ländern und die Auswirkungen von Wechselkursschwankungen auf die bilateralen Handelsströme, zusammenführt. Diese Arbeit zeigt, dass Fuzzy-Logik und neuronale Netze in der Volkswirtschaftslehre manchmal als Alternative und manchmal als eine Ergänzung verwendet werden können, und auch zu befriedigenden Ergebnissen führen, die im Einklang mit der internationalen Literatur sind.

LIST OF TABLES

Table 1 - 1: Balanced panel estimates with period fixed effects, dependent variable: log of total bilateral trade 13
Table 1 - 2: Increase in Exchange Rate Volatility Partitioning 16
Table 1 - 3: Decrease in Total Trade Partitioning……….…………………………16
Table 1 - 4: Fuzzy rule for explaining the effects of increase in exchange rate volatility on bilateral trade 19
Table 2 - 1: Balanced panel estimates, dependent variable: log of nominal bilateral exports 43
Table 2 - 2: Forecast Errors produced by Model 1 and Model 2 45
Table 3 - 1: Panel least squares with period fixed effects, dependent variable: log of real bilateral exports 64
Table 3 - 2: Panel Least Squares Model with Period Fixed Effects, Dependent Variable: Log of Real Bilateral Exports 70
Table 3 - 3: Forecast Results of Panel Least Squares Model with Period Fixed Effects 70
Table 3 - 4: Results of Training 71
Table 3 - 5: The Test Results of the Neural Network Model 73
Table 3 - 6: Forecast Errors of Panel Model for Different Samples 76

LIST OF FIGURES

Figure 1 - 1: An example of a triangular function 18
Figure 1 - 2: Linguistic Values for Variable “Exchange Rate Volatility” 18
Figure 1 - 3: Linguistic Values for Variable “Total Trade 18
Figure 1 - 4: The process of fuzzification and defuzzification 20
Figure 3 - 1: The multilayer network with 6 inputs, 25 hidden neurons and 1 output 62
Figure 3 - 2 : The topology of the panel network 63
Figure 3 - 3: Training, validation and test set errors.. 67
Figure 3 - 4: Outputs produced by the neural network versus targets 69
Figure 3 - 5: Training Set Outputs versus Targets... 72
Figure 3 - 6: Validation Set Outputs versus Targets ………………………………….. 72
Figure 3 - 7: Test Set Outputs versus Targets……………………………………… 72
Figure 3 - 8: Test Set Outputs versus Targets.. 74
Figure 3 - 9: A Multi Layer Feed Forward Network Structure 81

LEBENSLAUF

Mag. Elif NUROGLU

(vor dem Heirat Elif OYUK)

Geb. 25 November 1982

International University of Sarajevo

Faculty of Economics and Business Administration, Sarajevo-BiH

Address: Hrasnicka Cesta, 71000 Sarajevo - BiH

Tel: +38733957407

E-mail: enuroglu@ius.edu.ba

Ausbildung (Universität, Gymnasium)

März 2005-	Doktoratstudium Universität Wien, Volkswirtschaftslehre Dissertation: Three Essays on Analyzing Bilateral Trade Flows
September 2001- November 2004	Wirtschaftuniversität Wien, Betriebswirtschaftslehre Schwerpunkt: Produktionsmanagement und Public Management
Dezember 1999- Februar 2001	Bilkent Universität, Türkei Internationale Beziehungen
Bis Juni 1999	Imam Hatip Gymnasium, Türkei

	Schwerpunkt: Arabisch und naturwissenschaftliche Fächer

Berufsausbildung

September 2007-	**Lektorin,** Internationale Universität Sarajevo, BiH
Februar 2007- September 2007	**Assistentin,** Internationale Universität Sarajevo, BiH
April 2005 – June 2005	**English Lektorin**, Verein WONDER Wien
Oktober – Jänner 2004	**Tutorin** für „Beschaffung, Logistik und Produktion“ Institut Produktionsmanagement - WU -Wien
August 2004	Memorial Hospital, Istanbul **Praktikum in den Einkauf und Logistik- sowie Management Abteilungen**
Juni 2004	Wiener Allgemeines Krankenhaus, Wien **Praktikum in der Qualitätsmanagement Abteilung**

Konferenzen

- Huseyin Hayri Nuroglu, Elif Nuroglu, *Activity Based Costing leads to Strategic Decision Making*, 4th International Strategic Management Conference Proceedings, Sarajevo/BIH, 19-21 June 2008.

- Elif Nuroglu, Huseyin Hayri Nuroglu, *Relationship Between Strategic Management and Total Quality Management How Can TQM be Implemented In The Service Sector? Case Study: Vienna General Hospital*, 4th International Strategic Management Conference Proceedings, Sarajevo/BIH, 19-21 June 2008.

- Huseyin Hayri Nuroglu, Elif Nuroglu, *An Integration of Game Theory into the Operations Research,* ECCO XXI Operations Research Conference Proceedings, Dubrovnik/Croatia, 29-31 May 2008.

- Elif Nuroglu, *"The Effect of EU Membership on Bilateral Trade, Case Study: Greece as a Balkan Country"*, East West Relations: Past, Present, Future, Co-organized by Ankara Center for Thought and Research(ADAM) and International University of Sarajevo, Sarajevo/BiH, 14-15 February 2008

- E. Oyuk, R. Kunst, E. Tacgin,*"The Impact of Exchange Rates on International Trade in Europe from 1960s till 2000 Using a Modified Gravity Model and Fuzzy Approach",* SOR '07 The 9th International Symposium on Operations Research Proceedings, Nova Gorica, SLOVENIA, September 26–28, 2007

Editierte Bücher

Gozen Ramazan, Mehmet Bulut and Elif Nuroglu (eds.), *The East-West Relations: Turkish and Bosnian Perspectives* (Sarajevo: International University of Sarajevo, 2008)

Sonstiges

Sprachen: - gute Deutsch-Kenntnisse

September 2001- Deutsch Abschlussprüfung

- sehr gute Englisch-Kenntnisse, auch Business English

September 2000 Cambridge Certificate for IELTS Exam

Februar 2005 TOEFL Certificate

- Muttersprache Türkisch

Computer Kenntnisse: SAP, Arena, Adonis, Eviews, Matlab

Hobbys: Reisen, Eislaufen, Tennis