Academic year 2018-2019

Erasmus Mundus Joint Master's Degree
in Economics of Globalization
and European Integration

The international Balassa-Samuelson effect in Central and Eastern European countries: Is its importance decreasing?

Master dissertation

Student	Carl-Philip Dörge
Home institution	Vysoká škola ekonomická v Praze
Supervisor	Vilém Semerák
Submission date	September 2019

DECLARATION OF AUTHORSHIP

I, Carl-Philip Dörge, hereby declare that the dissertation thesis “The international Balassa-Samuelson effect in Central and Eastern European countries: Is its importance decreasing? “ was written by myself, and that all presented results are my own, unless stated otherwise. The literature sources are listed in the list of references.

Udine, September 30th 2019

ACKNOWLEDGEMENT

I would like to express my sincerest gratitude to the supervisor of my thesis, Vilém Semerák, Ph.D. , for his professional guidance, endless patience and very valuable recommendations through out the whole process of writing the thesis.

Also, I would like to thank my family and my friends for their support.

Abstract

This dissertation examines the external Balassa-Samuelson effect in Central and Eastern European countries. Eleven of the twelve Central and Eastern European are investigated empirically for the period from 1995 to 2017 with annual data. For that period the external Balassa-Samuelson effect is estimated using a Vector Error Correction Model. Empirical findings agree with empirical literature revised.

List of content

Dissertation

Introduction

The Balassa-Samuelson hypothesis proposes that countries experiencing relatively high productivity growth will also experience relatively high appreciation of their exchange rates as growth of productivity in tradable sector leads to growth of prices in non-tradable one. The Balassa-Samuelson effect consists of two interlinked effects. Firstly, the internal effect explains a change in relative price in the non-tradable sector with a change in the relative productivity of sectors. This effect is also referred to as Baumol–Bowen effect. Secondly, the international effect explains a change in the real exchange rate with a change in relative productivity of countries which is the Balassa-Samuelson effect. The second effect is sometimes also referred to as external Balassa-Samuelson effect. This dissertation will focus on the external effect.

The basic explanation of the Balassa-Samuelson effect is a model with two countries, with two sectors, and with labour as factor of production. The two sectors are distinguished by their goods which are either tradeable or not. There is a sector producing tradeable goods and a sector producing non-tradeable goods. One of the most important assumptions is that labour moves freely between the two sectors within a country. Free labour mobility implies relative homogeneity of labour as well as absence of other barriers which could be administrative to the mobility of labour. Furthermore, free labour mobility implies that wages in both sectors must be equal within a country. One more important assumption is that the marginal productivity in the non-tradeable sector is equal for both countries and constant. A third assumption is that trade restrictions are non-existent or at least minimal so that the price of the tradable good is equal for both countries. Whilst the third assumption was strictly assumed when the model was first developed, it has been shown that the dynamic effect persists when this is not the case to the full extent. As the wage in a sector is closely related to the sectorial marginal labour productivity and the price of the good in that sector an increase in productivity in the tradable good sector leads to higher wages in that sector because the price of the tradable goods will stay constant as they are dictated by the international market. Free labour mobility implicates that wages in the non-tradeable sector rise subsequently. Which leads to a rise of the price of the non-tradeable good. Hence, the overall price level of a country with a higher increase in productivity in the tradeable sector increases more than the overall price level of a country with a lower increase in productivity in that sector. Naturally, this creates pressure on the inflation to increase for the country with the higher growth in productivity and pressure on the real exchange rate to appreciate.

Genuine interest in the Balassa-Samuelson effect as a research field exists because the Balassa-Samuelson effect is an approach to deviations of the law of one price, it helps to determine the exchange rate, and explains price level variations between countries with different levels of development. Within Europe a difference in productivity development is apparent in form of the catching up process of the newer members. Hence it is reasonable to ask if the Balassa-Samuelson effect plays a role. As such, this dissertation aims to examine empiric evidence for Central and Eastern European Countries. The Central and Eastern European countries are Albania, Bulgaria, Croatia, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovenia, and Slovenia. This set of countries is particularly interesting as there is a debate about the magnitude of the Balassa-Samuelson effect that affects them. Considering the impact of the Balassa-Samuelson effect on inflation, its occurrence strongly affects fiscal policies of all Central and Eastern European Countries and monetary policies of the countries that are not within the European monetary union to achieve the Maastricht criterion. As policies geared towards reaching an inflation rate that is lower than their ideal inflation rate, are likely to affect their tradeable sector negatively.

The name Balassa-Samuelson effect is the most commonly used name after the authors of the most renown underlying papers. The two papers are "The Purchasing Power Parity Doctrine: A Reappraisal" by Bela Balassa (1964) and "Theoretical Notes on Trade Problems," by Paul Samuelson (1964). As the effect was discovered independently and simultaneously it is not named after one economist. In fact, the effect had been described earlier in 1911 by David Ricardo (Ricardo 1911) and the first fully worked out model was produced by Roy Harrod in 1933 (Harrod 1933). As there were even more economists who had worked on the topic, in the nineties Samuelson himself suggested that Ricardo-Viner-Harrod-Balassa-Samuelson-Penn-Bhagwati effect as a name. But due to convenience the names Balassa-Samuelson effect and Harrod-Balassa-Samuelson effect are predominantly used. Whereby the name Harrod-Balassa-Samuelson effect is a more modern occurrence.

This dissertation aims to examine the external Balassa-Samuelson effect in Central and Eastern European Countries for as many years and countries as possible. However, it is wise to follow most empirical research and start the period observed no earlier than 1995 due to the strong effects in the transition period. Annual data is available for more countries over the full time period than quarterly data. Annual data has further advantages like the avoidance of seasonal adjustment issues with unit roots and the reduced number of lags in the Augmented Dickey-Fuller Test (MacKInnon 1996, p. 613). However, for an equal time period annual data naturally provides less observations reducing the power of test like unit root tests to an extent and is more likely to fail to capture macroeconomic effects that manifest themselves in the short run.

This dissertation is structured as follows. Chapter 2 is the literature review. The literature review aims to grant an overview of the research in this field and to highlight empirical findings of studies that focus on Central and Eastern European Countries. Chapter 3 provides the deduction of the theoretical and empirical model. Chapter 4 contains data, the construction of the variables needed and a basic data description. Chapter 5 is about the methodological approach of the empirical estimation. Chapter 6 provides the results of the empirical analysis. Chapter 7 concludes.

Literature review

There are three more recent surveys of empirical evidence on the Balassa-Samuelson effect. The oldest of the three is "Productivity Bias Hypothesis and the Purchasing Power Parity: A Review Article" by Bahmani-Oskooee, Mohsen & Nasir from 2005. A year later Egert, Halpern and McDonald published "Equilibrium Exchange Rates in Transition Economies: Taking Stock of the Issues". Also, in 2006, Josip Tica and Ivo Družić published their paper called “The Harrod-Balassa-Samuelson Effect: A Survey of Empirical Evidence”. Which is a rather complete summary of empirical work up to that point including 58 empirical papers. Their paper not only analyses the empirical results it also describes the evolution of the model over time and the development in the methodological approach incorporating the most important empiric issues and their solutions. Also, its findings reflect the recently increasing popularity of the Balassa-Samuelson effect better (Tica 2006, p.5).

Tica and Družić found almost exclusively proof for the validity of the Balassa-Samuelson effect as: “A huge majority of papers have resulted in statistically significant coefficients and theoretically predicted signs for the majority of analysed countries. “(Tica 2006, p.12). Specifically, cross-section analysis has shown strong support for the Balassa-Samuelson effect as only two papers that use this approach were not supporting the model (Tica2006, p.12). 49 out of 58 empirical papers surveyed support the Balassa-Samuelson effect with all their findings. Furthermore, six papers of the nine papers that did not support the Balassa-Samuelson effect fully had findings that partially did not support whereas the remaining three papers completely dismissed the hypothesis (Tica 2006, p. 14).

The empirical model used in research started out as a rather simple ordinary least squares model that analyses the log-linear relationship between the levels of price and productivity within countries whereby the demand side was not properly included in the analysis (Tica 2006, p.6). Until the early nineties of the 20th century, the ordinary least squares dominated the research in this area. Naturally, it progressed as more econometric methods were developed and more extensive data became available. In 1996, the Johansen cointegration method was used to analyse the Balassa-Samuelson effect for the first time by Bahmani-Oskooee and Hyun-Jae Rhee in their paper “"A Time-Series Approach to Test the Productivity Bias Hypothesis in Purchasing Power Parity". This approach then became the most often used method (Tica2006, p.9). Nonetheless, there are many econometric methods applied. The Robert Engle and Clive Granger's (1987) cointegration technique, and the autoregressive distributed lag (ARDL) technique are more prominent among them (Tica 2006, p.9). What most more modern approaches have in common is a sectorial approach (Tica 2006, p.14). The productivity proxy variable is the main explanatory variable in the econometrical model. There has been an ongoing debate in the literature of this field about which proxy variable is the correct one as there are plenty of good reasons for both main options. Researchers are split between total factor productivity and average labour productivity (Tica 2006, p.12). Independent of the conclusion drawn about which proxy is better in theory, most researchers go with average labour productivity due to scarcity of data (Tica 2006, p.6).

While in 1964 Samuelson and Balassa wrote an original, name-giving paper each independently of each other, Samuelson only wrote a theoretical work consisting of notes and thoughts whereas Balassa did an empirical analysis as well as theoretical work (Tica 2006, p.6). In this first test of the hypothesis Balassa and Samuelson proposed, Balassa used an ordinary least squares approach for a cross-country analysis between nine countries for the year 1955. Balassa found “a positive correlation between the growth of manufacturing productivity and the ratio of the GNP deflator to the price index for manufactured goods. (Balassa 1964, pp.594)”. Which is evidence for his propositions.

The first one to formulate a model fully and apply it within a General Equilibrium framework was Rogoff in 1992 in the paper “Traded Goods Consumption Smoothing and the Random Walk Behaviour of the Real Exchange Rate and he did indeed find that the base assumption of the Balassa-Samuelson effect holds (Rogoff 1992, p. 32.). More specifically, the newly formulated model was applied to the dollar/yen exchange rate using quarterly data between 1975 and 1990 in which the exchange rate was floating (Rogoff 1992, p.19) and found that unexpected shocks to productivity affect the exchange rate, though stronger in the earlier time of the data set than in the later years despite the shortness of data (Rogoff 1992, p. 32). Also, important to note is that Rogoff incorporated the demand side of the economy.

A very good example for this enablement is another key paper in this field which is "International Evidence on Tradables and Non-tradeables Inflation" by Jose De Gregorio, Alberto Giovannini and Holger Wolf from 1994. When applying OECD data for 14 countries including 20 sectors for the years between 1975 and 1990 the three researchers found strong evidence for the Balassa-Samuelson effect (De Gregorio, p.20) and that in the long run the price level of an economy does not remain unaffected by demand side effects (De Gregorio, p.19). Furthermore, the researchers introduced a ten percent threshold for tradability by sector (De Gregorio, p.7). However, it is noteworthy that this measure of tradability is not in common practice in later papers due to the scarcity of data (Tica 2006, p.14).

Another paper that influenced research in this field greatly is “The Balassa-Samuelson General Equilibrium Model: A General Equilibrium Appraisal” by Patrick Asea and Enrique Mendoza. This paper was published in 1994. Besides introducing the Balassa-Samuelson effect successfully into a dynamic setting (Asea, p.25) Asea and Mendoza show that “the ratio of sector output per capita level and not the aggregate level of output determines the relative price of non-tradeables” (Tica 2006, p.9). Even though this broadens the way for more accurate research in theory, the common scarcity of adequate data makes it hard for researchers to use this.

In 2004 Paul Bergin, Reuven Glick, and Alan Taylor published a paper called “Productivity, tradability, and the long-run price puzzle”. This is a key paper in the field as it contrary to other papers before who took varying approaches to classifying tradability if not following De Gregorio et all´s ten percent threshold they did not treat tradability as exogenous (Bergin, p.2). More specifically, sector productivity is disaggregated to explain the correlation between productivity and tradability (Bergin, pp.5). Treating tradability as endogenous also allows for the model to have a better to fit to the in reality constantly changing degrees of tradability of sectors and goods. The empirical example used in this paper is an illustration of their proposal as a possible improvement to the Balassa-Samuelson model. But the paper does not focus on how achievable incorporation of their approach is with regards to data. However, their proposal is incorporating a new concept from more genuine trade research into the Balassa-Samuelson effect.

A very recent paper called “The Balassa-Samuelson effect reversed: new evidence from OECD countries” by the researchers Gubler and Sax applies three annual data sets to a sample consisting of 18 major Organisation for Economic Co-operation and Development countries for a long time period from the seventies to the financial crisis. The researchers test their panel data sets for cointegration and estimate the relationship with the dynamic ordinary least square approach. Overall, their findings do not support the Balassa-Samuelson effect (Gubler 2019, p. 12).

Naturally, most empiric papers researching the Balassa-Samuelson effect in Central and Eastern European countries were written when interest and relevance for that subject and for those countries peaked. Interest peaked when the Central and Eastern European countries were considered for the ascension into the European Union. Naturally, follow-up studies became increasingly infrequent over time.

A good example of a paper published during the peak time of interest is “The Balassa–Samuelson effect in Central and Eastern Europe: myth or reality?” by Egert, Drine, Lommatzsch, and Rault. It was published in 2003. Nine of the Central and Eastern European countries are being investigated in the paper. These nine countries are Croatia, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia, and Slovenia. The time considered is from 1995 to 2000 and the researchers argue that data from before 1995 should not be included to avoid the effects of the turbulent first years of transitions (Egert 2003, p. 558). After the necessary unit root tests, the Pedroni panel cointegration test is used to find the estimates. They find supportive evidence for the Balassa-Samuelson effect. Another paper from Egert from 2002 researches the Balassa-Samuelson effect for the Czech Republic, Hungary, Poland, Slovakia and Slovenia during the transition period. Using data from 1991 to 2000 for three countries and data from 1993 to 2001 for two countries Egert finds when applying the Johansen cointegration approach that the Balassa-Samuelson effect is present in all five countries but to different extents. In 2006 Egert also published "Balassa-Samuelson Meets South Eastern Europe, the CIS and Turkey: A Close Encounter of the Third Kind?". In this paper six countries are examined for the Balassa-Samuelson effect. Three of them, namely Bulgaria, Romania, and Croatia, are of interest here. Estimating data for the years 1993 to 2003 with a dynamic ordinary least square approach, Egert finds that only Croatia is affected strongly by the Balassa-Samuelson effect. The most recent paper by Egert about the Balassa-Samuelson effect is "Catching-up and inflation in Europe: Balassa-Samuelson, Engel's Law and other culprits" from 2011. It uses panel estimation and more recent data than the previous papers of the author on most Central and Eastern European countries. It finds the Balassa-Samuelson effect to be smaller in magnitude compared to the previous papers. However, it is not focused on Central and Eastern European countries.

Another paper out of that time period with focus on Central and Eastern European countries is "The Balassa–Samuelson Effect in Central Europe: A Disaggregated Analysis" from Mihaljek and Klau published in 2004. Croatia, the Czech Republic, Hungary, Poland, Slovakia and Slovenia are being examined from different points of the early 90s depending on availability of data until 2001 using ordinary least square estimation. The authors find that the Balassa-Samuelson effect is evident. In 2008 Mihaljek and Klau publish another paper with similar approach called "Catching-Up and Inflation in Transition Economies: The Balassa-Samuelson Effect Revisited" (Mihaljek, D., & Klau, M. (2008). Catching-up and inflation in transition economies: the Balassa-Samuelson effect revisited.). In this second paper the two researchers mainly improve data. They do so by expanding the countries considered as the only Central and Eastern European country not considered is Albania, by applying newer more reliable and more extensive data, and by increasing the time span considered. Whilst finding that the Balassa-Samuelson effect is present in all countries, they also find that its magnitude and direction are not unanimous.

In a paper called "Estimating the impact of the Balassa-Samuelson effect in Central and Eastern European Countries: A Revised Analysis of Panel Data Cointegration Tests" from 2011 Miletic applies panel estimation techniques for nine Central and Eastern European countries and find smaller magnitudes for the Balassa-Samuelson effect than in their previous papers.

The paper "Harrod-Balassa-Samuelson effect in selected countries of central and eastern Europe" from 2004 by a group of scholars composed of Baszkiewicz, Kowalski, Rawdanowicz, and Wozniak examines the Balassa-Samuelson effect in Bulgaria, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia, and Slovenia between 1995 and 2003. The scholars find that the Balassa-Samuelson effect is affecting price level within the countries for the whole sample of countries but that the effects on exchange rate are ambiguous.

A more recent paper is "Estimating the impact of the Balassa-Samuelson effect in Central and Eastern European Countries: A Revised Analysis of Panel Data Cointegration Tests" by Miletić which was published in 2011. The paper focuses on the Balassa-Samuelson effect for nine Central and Eastern European countries. The nine countries are Bulgaria, Croatia, Czech Republic, Hungary, Macedonia, Poland, Romania, Slovenia, and Slovakia. In this paper, data for a time frame from the early nineties to 2010 is used. Miletić finds that the Balassa-Samuelson effect plays a role. But that role is not the most important one for most countries.

The researcher Konopczak published a paper called "The Balassa-Samuelson Effect and the Channels of its Absorption in the Central and Eastern European Countries" in 2013. The paper examines the Czech Republic, Hungary, Poland, and Slovakia with respect to the years 1995 to 2010. The researcher finds that all countries are affected by the Balassa-Samuelson effect but not in the same magnitude.

The paper "Harrod, Balassa, and Samuelson (re)visit Eastern Europe" from Sonora and Tica, which was published in 2014, covers eleven out of twelve Central and Eastern European countries. The country that is not included is Albania. The time span considered is from 2000 to 2009 for all eleven countries and 1995 to 2009 for six of them. With panel estimation technique and an error correction model the researchers find that the Balassa-Samuelson effect is evident but does not have a big impact.

Another paper from 2014 which is called "Wage spillovers across sectors in Eastern Europe" by D'Adamo is of interest despite it being different in its approach compared to most papers accorded directly to researching the Balassa-Samuelson effect. The approach is to examine which sector is driving in the wage setting process. Which, when compliant with Balassa-Samuelson framework, should be the sector that produces tradable goods. For the time between 2000 and 2011 the paper examines the wage setting process of the countries Bulgaria, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovak Republic, and Slovenia. The researcher finds that the tradable goods sector is not the wage-driving sector for all countries and that labour mobility is far from perfect. Hence, a bias is implied in conventional models of estimating the Balassa-Samuelson effect as they assume perfect labour mobility.

In 2018 the paper "Balassa-Samuelson Effect in 10 CEE Countries: Does Productivity Influence their Price Convergence to EU15" from Adámek was published. The ten Central and Eastern European countries included in the sample are Bulgaria, the Czech Republic, Estonia, Croatia, Hungary, Latvia, Lithuania, Poland, Slovakia and Slovenia and the years included are 1995 to 2017. The model used here is adopted by the papers results for the validity of two main assumptions made in the theoretical framework and estimated with ordinary least squares with heteroscedasticity consistent estimator. The findings are that the Balassa-Samuelson effect plays a role in all countries considered but is the strongest in Poland, Slovenia, and Slovakia.

Table one provides a quick overview over selected empirical studies estimating the Balassa-Samuelson effect. To sum up the findings of reviewing the literature, no two papers use the same model specification or, the exact same proxies for all variables. Especially the proxy variable for productivity varies from paper to paper. Nor is there a golden rule for the estimation method established. Though the common trend appears to be panel estimation in recent years. Nonetheless, the results of the empirical evidence for the Balassa-Samuelson effect shift from mostly very supportive towards still supporting the Balassa-Samuelson but finding its effect to be smaller in more recent years (Miletić 2011, p. 476). For example, the papers of Egert and the papers of Mihaljek and Klau reflect this development nicely. Also, it is noteworthy that Albania is not included in any of the papers mentioned above indicating the poor availability of data for this country.

Theoretical and empirical model

Theoretical model

The Balassa-Samuelson hypothesis is mainly that the real exchange rate is increased by an increase in productivity level. As mentioned in the previous section the first fully formulated model was introduced in 1992 by Rogoff (Tica 2006, p.6). The basic model of the for underlying framework consists of a two country and two sector model where the goods produced in one sector are tradable (T) and the goods produced in the other sector are not tradable (N). The assumption this is based on is that all goods are clearly distinguishable by international tradability. Additionally, the price of the traded good cannot be affected by the two countries as it is determined by international trade and is therefore constant and equal for the two countries. The second, important assumption for the model is that labour (L) is fully mobile in between sectors of each country but there is no labour mobility between countries and that capital (K) is perfectly mobile between countries. In the following notation * is used to indicate the foreign country.

When inserting the general price (P) level equations of both countries which are

$P = P_T^{\alpha} \times P_N^{1-\alpha}$ (equation 1)

and $P^* = P^{*\alpha}_T \times P^{*1-\alpha}_N$ (equation 2)

into the definition for the real exchange rate (RER) which is

$RER = \frac{P}{\varepsilon * P^*}$ (equation 3)

where ε is the nominal exchange rate we get

$RER = \frac{P_T^{\alpha} \times P_N^{1-\alpha}}{P_T^{*\alpha} \times P_N^{*1-\alpha}}$.

Now as it is assumed before, the prices for the traded goods in both countries are equal and for the simplification equal to one. This gives

$RER = \frac{P_N^{1-\alpha}}{P_N^{*1-\alpha}}$ (equation 4).

To consider the supply side of the economy the basic production function for the home country

$Y = Y_T + Y_N;\ Y_T = A_T F(K_T, L_T);\ Y_N = A_N F(K_N, L_N)$ (equation 5)

and the basic production function for the foreign country

$Y^* = Y_T^* + Y_N^*;\ Y_T^* = A_T^* G(K_T^*, L_T^*);\ Y_N^* = A_N^* F(K_N^*, L_N^*)$ (equation 6)

and the assumption of perfect labour mobility within countries which means that real wages are equal in both sectors are used in a profit maximization. This maximization demonstrates that that the marginal product of capital equals to the world interest rate (factor price of K), and the marginal product of labour equals to the nominal wage (factor price of L) in two sectors (Dumitru and Jianu 2009, pp. 885). Hence,

$W_T = W_N = W$ and $W_T^* = W_N^* = W^*$.

This leads to

$W = P_T \times A_T = P_N \times A_N; since\ P_T = 1; P_N = A_T/A_N$ (equation 7)

for the home country and

$W^* = P_T^* \times A_T^* = P_N^* \times A_N^*; since\ P_T^* = 1; P_N^* = \frac{A_T^*}{A_N^*}$ (equation 8)

for the foreign country. Inserting equations 7 and 8 into equation 4 applying natural logarithms gives the deterministic Balassa-Samuelson model

$\ln(RER) = (1 - \alpha)\beta((\ln A_T - lnA_T^*) - (\ln A_N - lnA_N^*))$ (equation 9).

The ratio β stands for the share of labour in tradable and non-tradable sectors. By assuming that those shares are even β, naturally becomes 1. Obstfeld and Rogoff show in 1996 that this ratio comes very close to one even when they work with losing that assumption (Obstfeld and Rogoff 1996, pp. 208). Equation 9 can be further simplified by assuming that productivity in non-tradable sectors are equal and not changing for both countries. On a side note, this simplification also helps deal with the scarcity of data as less data is required. The simplification of equation 9 gives

$\ln(RER) = (1 - \alpha)((\ln A_T - \ln A_T^*))$ (equation 10).

This is a basic theoretical model for the Balassa-Samuelson effect where (1-α) reflects the Balassa-Samuelson effect. A positive coefficient for (1-α) confirms the Balassa-Samuelson effect.

Empirical model

The empirical model is derived from the basic theoretical model (equation 10) by writing it as stochastic model with constant and error term. Also, the sector classification can be dropped as only one sector is considered in the basic theoretical model. This gives

$\ln(RER_{c,t}) = c + \beta_1 \times ln\left(\frac{A_{c,t}}{A_t^*}\right) + u_{c,t}$ (equation 11).

This model is expanded with control variables to improve the explanatory power of it and to reduce the risk of the omitted variable bias. The expanded model with control variable (CV) is

$\ln(RER_{c,t}) = c + \beta_1 \times ln\left(\frac{A_{c,t}}{A_t^*}\right) + \sum_{j=1}^{J} \beta_j \times CV_{j,t} + u_{i,t}$ (equation 12).

The four control variables included are terms of trade, gross domestic product per capita, employment rate, and public debt as share of the gross domestic product. Terms of trade and gross domestic product per capita have been implemented in the research on the Balassa-Samuelson effect before. Terms of trade as an explaining variable was introduced by De Gregorio et al. (1994) in their before-mentioned paper (De Gregorio 1994) and often applied after (Tica 2006, p.10) because it adds supply side effects into the model due "to the home bias in consumption preferences" (Gubler 2019, p. 5). Gross domestic product per capita has been used before by, for example, Drine and Rault (Drine 2005). Employment rate and public debt as share of the gross domestic product are not applied often in research on the Balassa-Samuelson effect but have been shown to have effects on the exchange rate in other fields of economic research.

Variables, data set, and data description

Variables

The selection of data for the variables is an attempt to balance the advantage of having large numbers of observations in the estimations, the goal of covering all Central and Eastern European Countries, and the scarcity of data. As evident from the previous chapter, there are six variables in our model consisting of the dependent variable, the explanatory variable of interest, and four control variables. These consist of a variable for the exchange rate, a variable for the relative productivity, a variable for the terms of trade, a variable for gross domestic product per capita, employment rate, and public debt as share of the gross domestic product.

The variable for the exchange rate is constructed from real effective exchange rate index data from a dataset provided by Bruegel. The real effective exchange rate is the real exchange between a country and a basket of its trading partners (Ho 2012, p.2). Whilst in the theoretical model a country trades with one country, in reality each country trades with multiple partners. Hence the real effective exchange rate reflects the real exchange rate with respect to a reference country from the theoretical two country model because the real effective exchange rate of the country consists of the exchange rate of a country with regards to all its trade. Thus, the real effective exchange rate is also a better approximation of the empirics to the theoretical framework for this sample than a singular real exchange rate of a country with a reference country even if the reference country is the main trading partner.

The proxy variable for the productivity variable is more difficult. The before-mentioned, theoretical discussion is negligible in this case because data is scarce, especially for sectorial data. The ratio of average labour productivity in the tradable sectors of a country to the average labour productivity in the tradeable sectors of the whole of the European Union is the proxy variable for relative productivity. Sectorial average labour productivity is calculated by dividing output by hours worked with regards to sectors and the average labour productivity of the European Union is unweighted. The selection of which sector is tradeable is taken from the recent selection from Sonora and Tica (2014) who examine a comparable period and sample of countries and apply the methodology suggested in an earlier mentioned paper by De Gregorio (1994). The industry sector, but without the construction sector, and the agriculture sector are taken as tradeable sectors. All data used for this variable is extracted from the Eurostat database.

Terms of trade is the ratio of the price level of exports to the price level of imports. All data for the two price levels in constant prices has been taken from the Penn World Table version 9.1. So, the variable for terms of trade has been constructed from scratch and is not a proxy in contrast to the explained variable and the productivity variable.

The gross domestic product per capita is taken from the World Development Indicators Database. No adaptations are necessary here as the World Development Index contains the gross domestic product per capita in constant prices.

The employment rate is ratio of the number of persons working to the number of persons in the labour force. The variable is constructed from the unemployment rate per country from World Development Index.

The public debt as share of gross domestic product consists of government consolidated gross debt in percent of the gross domestic product. This data is taken from Eurostat.

Data set

For eleven of the twelve Central and Eastern European Countries it was possible to find matching data for the six variables. Data for Albania is limited. Including Albania would have drastically reduced the amount of years covered and would have ruled out the proxy variable for the productivity variable. Hence, Albania is excluded from the analysis. As noted in the previous chapter, Albania is not investigated in any of the studies considered in this dissertation due to the poor availability of data. Even for the remaining countries the data lacks observations. For Bulgaria, the observations for the first two years of public debt are not available. For Estonia and Poland, the observations for hours worked by sector are not available from 1995 to 1999. Naturally, this implies that the proxy variable for productivity cannot be built for the first five years for the two countries.

As the observations of the six variables vary over time and for entities which are the countries, all six variables combined create a panel data set. It is common practice to describe panel data sets as short or long. Whereby "a short panel has many entities but few time periods, while a long panel has many time periods but few entities" (Park 2011, p. 3). This panel data set is more on the short end of the spectrum of short versus long classification for panel data sets. But as numbers of entities in this case countries and the length of the time are generally relatively small this classification does not fully apply. Furthermore, the panel data set is unbalanced as is evident from the lack of the twelve observations. On the one hand, unbalanced panel data sets are no problem for vector error correction models. Thus, the full panel data set can be applied for that part of the analysis. On the other hand, for panel unit root testing and for panel cointegration balanced panel data sets are required. Therefore, two subsets of the data which are both balanced are created. One subset contains data for all countries and the years for which observations are available for all eleven countries. That period is 2000 to 2017. The second subset contains data for all years and countries for which observations are available for all years. The balanced panels are referred to as Panel 2 and Panel 3. That restricted group of countries is Croatia, the Czech Republic, Hungary, Latvia, Lithuania, Romania, Slovenia, and Slovenia. Most importantly, as the countries included in do not change over time and the time measurement is constant and regular the panel data sets are fixed.

Data description

To illustrate the time series that are to be examined Graph 1 and Graph 2 are added in the Appendix A.

Graph 1 shows the development of the real effective exchange rate for all countries over time. The development for all countries over the full period is an appreciation. Nine out of eleven countries peak in 2008 or 2009 which is caused by the lagged effect of the financial crisis. As the Central and Eastern European countries are mainly exporting countries an appreciation followed by a strong deprecation as response to the financial crisis is to be expected. The countries that show the highest value in 2008 are Croatia, the Czech Republic, Hungary, and Poland. The countries that reach their highest value in 2009 are Bulgaria, Latvia, Lithuania, Slovakia, and Slovenia. One exception is Estonia which has its highest value in 2017. Another exception is Romania where the real effective exchange rate index reaches the highest point one to two years earlier than most countries in 2007. All countries show overall appreciation up to 2009 compared to 1995. Almost all countries with exception of Estonia depreciate slightly after 2009. But the depreciation post 2009 is small compared to the appreciation before 2009. Estonia which peaks in 2017 does appreciate from 2009 to 2017 but compared to its appreciation before 2009 it is marginal. To sum up, most countries appreciate from 1995 to 2009 and then stagnate or depreciate slightly where the depreciation is nowhere near the magnitude of the earlier appreciation.

Graph 2 shows the development of the relative average productivity in the tradable sector for all countries over time. All countries display an overall increase of relative average productivity in the tradable sector for the full period. The countries that have high relative average labour productivity of the tradeable sector, especially towards the end of observed time period, are the Czech Republic, Slovenia, and Slovakia. Countries with low relative average labour productivity of the tradeable sector are Bulgaria and Romania. Magnitude of change in relative average labour productivity of the tradeable sector of all countries are displayed in Table One. Bulgaria and Lithuania experience the highest increase in percent while Croatia and Slovenia experience the lowest increase. However, it is noteworthy that the occurrence of a higher increase up to 2009 compared to the development after 2009 seems to be similar. Croatia, the Czech Republic, Latvia, Poland, Romania, Slovakia, and Slovenia display a strong increase in relative average productivity of the tradeable sector and all seven countries display slow down significantly after, some even show an overall decrease in relative average labour productivity in the tradeable sector. Estonia and Hungary display an only moderate reduction of their increase in relative average labour productivity in the tradeable sector after 2009 compared to the rate of increase before 2009. Bulgaria seems to increase faster after 2009.

The development of the two variables seem to be similar for Croatia, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia, and Slovenia. As there is a strong increase until around 2009 for nine countries in both variables and afterwards the increase slows down significantly, sometimes even turns into a decrease. For Bulgaria this occurrence is not quite so evident. The real effective exchange rate develops in a pattern similar to the nine countries mentioned before. But Bulgaria`s relative average labour productivity in the tradeable sector seems to be constant in its rate of increase. Overall, there seem to be similarities in the development of the two variables with regards to countries over time.

Methodological approach

The Balassa-Samuelson effect in the model used in this dissertation implies a long-term relationship between the explanatory variable and the proxy variable for productivity in the tradeable sector. The commonly used approach is to examine if there is a long-term relationship and then estimate it. To test if there is a long-term relationship between time series cointegration tests are the appropriate tool. As the cointegration tests require the time series considered to be of the same order of integration, it is necessary to test for unit root. Once cointegration is established, the model can be regressed using an error correction model (Hackl 2004, p.278). All estimations are made using R. Packages used are referred to in the list of references.

Panel unit root

Testing for stationarity of time series is important for analysing the long-term relationship between time series. A stationary time series` statistical properties such as mean, variance, and autocorrelation are all constant over time. Hence, a non-stationary time series has inconsistent properties over time. This implies the risk of spurious regression when two non-stationary time series are linearly combined. Therefore, time series are tested for stationarity before analysing them further. This kind of test is referred to as Unit Root Test. If a time series has a unit root that means that the time series is non-stationary. And if a time series does not have a unit root, it means that the time series is stationary. There are various methods for transforming a non-stationary time series into a stationary time series. For example, taking the first difference of observations is such a technique. This one is particularly important as it is employed to determine the order of integration. If a time series is stationary, it is integrated of the order 0. If a time series is non-stationary and its first difference is stationary, it is integrated of the order 1. If a time series is non-stationary and turns stationary upon applying its second difference, it is integrated of the order 2. This can be continued for additional orders. However, in practice orders of integration bigger larger than 2 hardly ever occur. To find an equal for order of integration for time series, is a necessary condition for testing them for cointegration.

Tests for Unit Root in panel data are called Panel Unit Root tests. Two Panel Unit Root tests are used in this paper. One Panel Unit Root test was developed by Levin et al. (2003) and the other test is from Im et al. (2003). Both tests are based on the Augmented-Dickey-Fuller test (Im 2003, p. 52).

Levin et al. recommend their test for sample sizes of moderate magnitude (Levin 2002, p.2). They determine moderate sample size as "10 and 250 individuals, with 25–250 time series observations per individual. (Levin et al. 2002, p. 2). Levin et. al examine the panel with regards to the null hypothesis that each time series contains a unit root (Levin et. al 2002, p. 5).

Im et al. state that their test is more powerful within small samples with regards to N (Im 2003, p.44). In the case of Balassa-Samuelson effect research N are countries. Im et al. test for the null hypothesis as stated:

"$H_0: \beta_i = 0$ for all I,

against the alternatives,

$H_o: \beta_i < 0, i = 1,2, \dots, N_1, \ \beta_i = 0, i = N_1 + 1, N_1 + 2, \dots, N.$ " (Im 2003, p.55).

This alternative hypothesis means that the panel is stationary for at least one time series. Hence, it can be argued that the higher power for small samples is due to the loser hypothesis.

For both Panel Unit Root tests there are different specifications. Noteworthy out of these is that an intercept is included in the testing but no trend variable as no complex trend is expected. The Schwarz information criterion is used for choosing the number of lags. For each variable both tests are applied to increase the robustness of the result.

Engle-Granger cointegration

Under the assumption that both panels considered are integrated of the same order a cointegration test is performed. The approach suggested by Engle and Granger in 1987 is such a test (Engle and Granger 1987). Their two-step approach examines the residuals of a linear model which includes the times series of interest. More precisely, the test checks if the residuals of a linear combination of the two series are stationary or not. If the residuals are stationary the time series are cointegrated (Hackl 2004, p.278). The null hypothesis tested is that the residuals are non-stationary. Hence, a rejection of the null hypothesis implies that the variables in the linear model of the first step are cointegrated.

Also noteworthy is that the econometricians Engle and Granger suggest that an error correction model is the most convenient way to estimate cointegration (Engle and Granger 1987, p.260). As there has been development in the econometrical methods it might not necessarily be the most convenient way anymore. Yet, it does still imply that an error correction model is a plausible approach.

Vector error correction model

Long-term relationship models can also be expressed with an error correction model (Hackl 2004, p. 278). Estimating a part of the error correction model with fixed effects for the countries is a possibility to find the country specific impact of relative average labour productivity of the tradeable sector on the real effective exchange rate. Mathematical proof that the fixed effect regression reflects the Balassa-Samuelson effect in the model derived earlier (equation 12) starts with an equation of the general notation of ADL (1,1)-models. The equation is defined as

$Y_t = \alpha + \varphi Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + u_t$ (equation 13)

by Hackl (Hackl 2004, p 278). When applying notation methods used one gets

$reer_{ct} = i + \alpha_{1c} * reer_{c,t-1} + \alpha_{2c} * prod_{c,t-1} + \sum_{j=1}^{J} \beta_{jc} * X_{jct} + \varepsilon_{ct}$ (equation 14)

Note that notations are simplified and slightly altered for clarity by

$reer_{ct} = \ln(RER_{c,t})$; $prod_{c,t} = ln\left(\frac{A_{c,t}}{A_t^*}\right)$; $\sum_{j=1}^{J} \beta_{jc} * X_{jct} = \sum_{j=1}^{J} \beta_j \times CV_{j,t}$ where X also includes fixed effects. i is the constant.

Expanding with $+reer_{c,t-1} - reer_{c,t-1}$ and $-X_{j,c,t} + X_{j,c,t}$ leads to

$reer_{ct} + reer_{c,t-1} - reer_{c,t-1} = i + \alpha_{1c} * reer_{c,t-1} + \alpha_{2c} * prod_{c,t-1} + \sum_{j=1}^{J} \beta_{jc} * (X_{jct} + X_{j,c,t-1} - X_{j,c,t-1}) + \varepsilon_{ct}$ (equation 15).

First differences give

$\Delta reer_{ct} = i + (\alpha_{1c} - 1) * reer_{c,t-1} + \alpha_{2c} * prod_{c,t-1} + \sum_{j=1}^{J} \beta_{jc} * (\Delta X_{jct} + X_{j,c,t-1}) + \varepsilon_{ct}$ (equation 16).

Introducing brackets give

$\Delta reer_{ct} = i + [(\alpha_{1c} - 1) * reer_{c,t-1} + \alpha_{2c} * prod_{c,t-1} + \sum_{j=1}^{J} \beta_{jc} * X_{j,c,t-1}] + \sum_{j=1}^{J} \beta_{jc} * \Delta X_{jct} + \varepsilon_{ct}$ (equation 17).

The term $\alpha_{1,c} - 1$ is put out of the brackets to get

$\Delta reer_{ct} = i + (\alpha_{1c} - 1) * [reer_{c,t-1} + \frac{\alpha_{2c}}{(\alpha_{1c}-1)} * prod_{c,t-1} + \sum_{j=1}^{J} \frac{\beta_{jc}}{(\alpha_{1c}-1)} * X_{j,c,t-1}] + \sum_{j=1}^{J} \beta_{jc} * \Delta X_{jct} + \varepsilon_{ct}$ (equation 18).

$\Delta reer_{ct} = i + -(1 - \alpha_{1c}) * [reer_{c,t-1} - \frac{\alpha_{2c}}{(1-\alpha_{1c})} * prod_{c,t-1} - \sum_{j=1}^{J} \frac{\beta_{jc}}{(1-\alpha_{1c})} * X_{j,c,t-1}] + \sum_{j=1}^{J} \beta_{jc} * \Delta X_{jct} + \varepsilon_{ct}$ (equation 19).

Equation 18 is the error correction model where the term in square brackets

$[reer_{c,t-1} - \mu_{1c} * prod_{c,t-1} - \sum_{j=1}^{J} \pi_{jc} * X_{j,c,t-1}]$ (simplified version from equation 20)

is the long-term relationship between exchange rate and productivity. To get

$\Delta reer_{ct} = i - (1 - \alpha_{1c}) * [reer_{c,t-1} - \mu_{1c} * prod_{c,t-1} - \sum_{j=1}^{J} \pi_{jc} * X_{j,c,t-1}] + \sum_{j=1}^{J} \beta_{jc} * \Delta X_{jct} + \varepsilon_{ct}$ (equation 20)

the deviation term is simplified, and some coefficients are redefined. Using the implication of the long-term relationship that (t-1) can be changed to t for an explicit notation. This gives

$reer_{ct} = 1 + \mu_{1c} * prod_{ct} + \sum_{j=1}^{J} \pi_{jc} * X_{jct} + \omega_{ct}$ (equation 20).

Instead of estimating the whole model, it is more straightforward to apply a fixed effects regression for the term in the brackets pointed out above because the long-term relationship reflected by the term is the Balassa-Samuelson effect. As the variables of interest are in natural logarithm values, the country specific coefficient can be in interpreted as elasticities, more specifically percentual change of the dependent variable per one percent change in the explanatory variable.

Results

Panel Unit Root

The results of the Panel Unit Root estimations are presented in table three and table four. As the panel data set is unbalanced and balanced panels are a requirement for the analysis software it is necessary to create two panel data sets. Panel 2 consists of observations for all elven countries and a reduced period and Panel 3 covers the whole period for eight countries. First, one example variable is interpreted step by step to illustrate the read of the results with regards to integration order. Second, the implications for the rest of the analysis are deducted.

The variable real effective exchange rate rejects the null hypothesis for the Levin et al. And the Im et al. test at the 0.05 significance level. The rejection of the null hypothesis implies for the Levin test that not each time series has a unit root. Which means that is stationary. In case of the Im et al. the rejection of the null hypothesis also means that the panel is stationary at 0.05 significance level. The only result that slightly deviates from this outcome is the variable for relative average labour productivity from panel when tested with the Im et al. The slight deviation is that in this case the rejection of the null hypothesis is at the 0.10 significance level. This means that all variables should be stationary. However, the discrepancy between the two Panel Unit Root tests for one out of four panels tested suggest a weak robustness of the results. This indication can be explained with the small sample with regards to observations over time. The period covered is limited by two factors. Firstly, data availability before 1995 becomes increasingly poor for more and more countries. To illustrate, the division into two panels is already caused by a lack of data for three countries before 2000. Secondly, the transition time period before 1995 causes different strong effects and is hence avoided by most of the relevant research. So that the number of observations is limited to 23 for Panel two and 18 to Panel three. More importantly, not capturing short process effects adequately is an inherent disadvantage of annual data which reduces their explanatory power. As shown technically by Asimakopoulos et al. annual data provides less explanatory value compared to quarterly data (Asimakopoulos 2013, p.2). This finding is also supported by Chambers how argues that the aggregation of high frequency data to into low frequency data causes a loss of information (Chambers 2016, p.400). This supports the indication of lacking reliability of the result of the Panel Unit Root test. The implication for the following analysis is that the condition for testing for cointegration cannot reliably be confirmed nor denied.

Engle-Granger approach

The theoretical assumptions are not strictly met. Yet due to the weak explanatory power of panel unit root tests with annual data the Augmented Dickey-Fuller test on the residuals is conducted anyway to test for the potential possibility of cointegration. To do so, it is assumed that the panels are cointegrated.

Under that assumption the result of the Engle Granger approach is -12.60112 for the test statistic where critical values are -2.58 for 0.01 level, -1.95 for 0.05 level, and -1.62 for 0.1 level. Therefore, the null hypothesis can be rejected confidently at the 0.01 significance level. If both panels were integrated of the same order this would imply that there is cointegration. But the same reliability issue ensues as the Augmented Dickey-Fuller test is based on a unit root test as well. Though, the high level of confidence provided suggests that a result for the data when integration was not merely assumed would support cointegration.

Fixed effects regression

The Error Correction approach is based on cointegration. As the results of the test for the condition are not reliable the uncertainty can be eased by considering a previous paper. The paper that is perhaps best comparable with regards to sample selection and period is Sonora (2014). Researching the exact same sample of countries for the shorter period that quarterly data is accessible for. Their evidence finds cointegration between exchange rate and productivity (Sonora 2014, p.10). Despite the shorter period considered the use of quarterly data provides explanatory power. Hence, leaning on their findings cointegration between exchange rate and productivity is assumed.

The country specific results of regressing the Balassa-Samuelson coefficient are presented in Table five. The log-log linear combination in the estimation means that the measured effect is an increase in percent of the dependent variable to a one percent increase in the explanatory variable. Five countries do not show an effect that is different to zero at the 0.05 significance level. These countries are Bulgaria, Estonia, Lithuania, Slovenia, and Romania. The highest measured effect is 0.0856% for Poland. The lowest effect that was not dismissed as statistically equal to zero is 0.0072% for Croatia. Noteworthy, is that the lowest value found is hence zero or very close to zero as the five countries show no effect. To sum up, the Balassa-Samuelson effect is found in six out of eleven countries investigated.

Whilst contrary to some previous papers no reverse effects are found, the distribution of few countries with relatively high effect and a majority of Central and Eastern European countries experiencing low to no Balassa-Samuelson effect is not contrary to the trend in literature of decreasing magnitude and quantity of Balassa-Samuelson effects.

Conclusion

Whilst the results for the international Balassa-Samuelson effect could not be determined with the level of certainty that is strived for one modern economics, the findings themselves are well in line with the empirical findings in the field.

The difficulty of the field is well described by Mihaljek who states that “Obtaining precise and reliable estimates of these effects is much more difficult than, for instance, obtaining estimates of potential GDP. In particular, measurement errors and room for discretion in transforming the data and applying even the simplest estimating procedures are not negligible. Issues of equal treatment would inevitably arise if one sought to standardise these estimating procedures in practice. Therefore, one would be hard pressed to recommend, in good confidence, an operationalisation of the concept of the Balassa-Samuelson effect for the assessment of the Maastricht inflation criterion.” MIhaljek 2008, p.17). Additionally, the scarcity of data which has been discussed in this dissertation is troubeling. The key concept applied in this dissertation to overcome this issue was introducing a high number of control variables and extending the period and sample of countries as far as possible. Yet the explanatory power provided by the data was not sufficient to provide reliability to the extend obtainable in economic research. Hence, an important conclusion is that quarterly data is preferable in this field.

Appendix A: List of references

Adámek, E. (2018). Balassa-Samuelson Effect in 10 CEE Countries: Does Productivity Influence their Price Convergence to EU15?. on European Integration 2018, 41.

Asea, P. K., & Mendoza, E. G. (1994). The Balassa-Samuelson Model: A General-Equilibrium Appraisal. Review of International Economics, 2(3), 244-267.

Asimakopoulos, S., Paredes, J., & Warmedinger, T. (2013). Forecasting fiscal time series using mixed frequency data.

Bahmani-Oskooee, M., & Rhee, H. J. (1996). Time-Series Support for Balassa's Productivity-Bias Hypothesis: Evidence from Korea. Review of International Economics, 4(3), 364-370.

Bahmani-Oskooee, M., & Nasir, A. B. M. (2005). Productivity bias hypothesis and the purchasing power parity: a review article. Journal of Economic Surveys, 19(4), 671-696.

Balassa, B. (1964). The purchasing-power parity doctrine: a reappraisal. Journal of political Economy, 72(6), 584-596.

Bergin, P. R., Glick, R., & Taylor, A. M. (2006). Productivity, tradability, and the long-run price puzzle. Journal of Monetary Economics, 53(8), 2041-2066.

Bergstrand, J. H. (1991). Structural determinants of real exchange rates and national price levels: Some empirical evidence. The American Economic Review, 81(1), 325-334.

Blaszkiewicz, M., Kowalski, P. A., Rawdanowicz, L., & Wozniak, P. (2004). Harrod-Balassa-Samuelson effect in selected countries of central and eastern europe. CASE Network Reports, (57).

Bruegel, 2019, Real effective exchange rates for 178 countries: a new database, Retrieved from https://bruegel.org/publications/datasets/real-effective-exchange-rates-for-178-countries-a-new-database/ , 18.09.2019

Canzoneri, M., Cumby, R., Diba, B., & Eudey, G. (2002). Productivity trends in Europe: implications for real exchange rates, real interest rates, and inflation. Review of international economics, 10(3), 497-516.

Chambers, M. J. (2016). The estimation of continuous time models with mixed frequency data. Journal of econometrics, 193(2), 390-404.

Croissant Y, Millo G (2018). _Panel Data Econometrics with R: the plm package_. Wiley.

D'Adamo, G. (2014). Wage spillovers across sectors in Eastern Europe. Empirical Economics, 47(2), 523-552.

Darvas, Zsolt (2012a) 'Real effective exchange rates for 178 countries: a new database', Working Paper 2012/06, Bruegel, 15 March 2012

Darvas, Zsolt (2012b) 'Compositional effects on productivity, labour cost and export adjustment', Policy Contribution 2012/11, Bruegel, 22 June 2012

De Gregorio, J., Giovannini, A., & Wolf, H. C. (1994). International evidence on tradables and nontradables inflation. European Economic Review, 38(6), 1225-1244.

Drine, I., & Rault*, C. (2005). Can the Balassa-Samuelson theory explain long-run real exchange rate movements in OECD countries?. Applied Financial Economics, 15(8), 519-530.

Dumitru, I., & Jianu, I. (2009). The Balassa–Samuelson effect in Romania–The role of regulated prices. European Journal of Operational Research, 194(3), 873-887.

Égert, B. (2011). Catching-up and inflation in Europe: Balassa–Samuelson, Engel's Law and other culprits. Economic Systems, 35(2), 208-229.

Égert, B. (2002). Estimating the impact of the Balassa–Samuelson effect on inflation and the real exchange rate during the transition. Economic Systems, 26(1), 1-16.

Égert, B. (2005). Balassa-Samuelson meets South Eastern Europe, the CIS and Turkey: A close encounter of the third kind?.

Égert, Balázs, László Halpern, and Ronald MacDonald. "Equilibrium exchange rates in transition economies: taking stock of the issues." Journal of Economic surveys 20.2 (2006): 257-324.

Égert, B., Drine, I., Lommatzsch, K., & Rault, C. (2003). The Balassa–Samuelson effect in Central and Eastern Europe: myth or reality?. Journal of comparative Economics, 31(3), 552-572.

Engle, R. F., & Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. Econometrica: journal of the Econometric Society, 251-276.

Eurostat.,2018, “sts_inlb_a”, Retrieved from https://ec.europa.eu/eurostat/data/database, 18.09.2019

Eurostat.,2018, “sts_colb_a”, Retrieved from https://ec.europa.eu/eurostat/data/database, 18.09.2019

Eurostat.,2018, “gov_10dd_edpt1”, Retrieved from https://ec.europa.eu/eurostat/data/database, 26.09.2019

Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer (2015), "The Next Generation of the Penn World Table" American Economic Review, 105(10), 3150-3182, available for download at www.ggdc.net/pwt

Gaure. lfe: Linear group fixed effects. R package version 2.8-3, 2019

Gubler, M., & Sax, C. (2019). The Balassa-Samuelson effect reversed: new evidence from OECD countries. Swiss Journal of Economics and Statistics, 155(1), 3.

Hackl, P. (2008). Einführung in die Ökonometrie (Vol. 7118). Pearson Deutschland GmbH.

Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2019). dplyr: A Grammar of Data Manipulation. R package version 0.8.3. https://CRAN.R-project.org/package=dplyr

Harrod, R. F. (1933). International Economics (Cambridge: Nisbet & Cambridge University Press).

Ho, L. S. (2012). Globalization, exports, and effective exchange rate indices. Journal of International Money and Finance, 31(5), 996-1007.

Im, K. S., Pesaran, M. H., & Shin, Y. (2003). Testing for unit roots in heterogeneous panels. Journal of econometrics, 115(1), 53-74.

Konopczak, K. (2013). The Balassa-Samuelson effect and the channels of its absorption in the Central and Eastern European Countries. National Bank of Poland Working Paper, (163).

Levin, A., Lin, C. F., & Chu, C. S. J. (2002). Unit root tests in panel data: asymptotic and finite-sample properties. Journal of econometrics, 108(1), 1-24.

MacKinnon, J. G. (1996). Numerical distribution functions for unit root and cointegration tests. Journal of applied econometrics, 11(6), 601-618.

Mihaljek, D., & Klau, M. (2004). The Balassa–Samuelson effect in central Europe: a disaggregated analysis. Comparative Economic Studies, 46(1), 63-94.

Mihaljek, D., & Klau, M. (2008). Catching-up and inflation in transition economies: the Balassa-Samuelson effect revisited.

Miletić, M. (2012). Estimating the impact of the Balassa-Samuelson effect in central and eastern European countries: A revised analysis of panel data cointegration tests. Panoeconomicus, 59(4), 475-499.

Park, H. M. (2011). Practical guides to panel data modeling: a step-by-step analysis using stata. Public Management and Policy Analysis Program, Graduate School of International Relations, International University of Japan, 1-52.

Pfaff, B. (2008) Analysis of Integrated and Cointegrated Time Series with R. Second Edition. Springer, New York. ISBN 0-387-27960-1

Ricardo, D. (1911). On the Principles of Political Economy and Taxation, introduced by Michael P. Fogart, Ed., JM Dent & Sons LDT, London.

Rogoff, K. (1992). Traded goods consumption smoothing and the random walk behavior of the real exchange rate (No. w4119). National bureau of economic research.

Obstfeld, M., & Rogoff, K. (2000). The six major puzzles in international macroeconomics: is there a common cause?. NBER macroeconomics annual, 15, 339-390.

Samuelson, P. A. (1964). Theoretical notes on trade problems. The review of economics and statistics, 145-154.

Sonora, R. J., & Tica, J. (2014). Harrod, Balassa, and Samuelson (re) visit Eastern Europe. Cogent Economics & Finance, 2(1), 920557.

Tica, J., & Družić, I. (2006). The Harrod-Balassa-Samuelson effect: a survey of empirical evidence. EFZG working paper series, (07), 1-38.

World Bank, World Development Indicators. (2019). SL.UEM.TOTL.ZS. https://databank.worldbank.org,24.09.2019

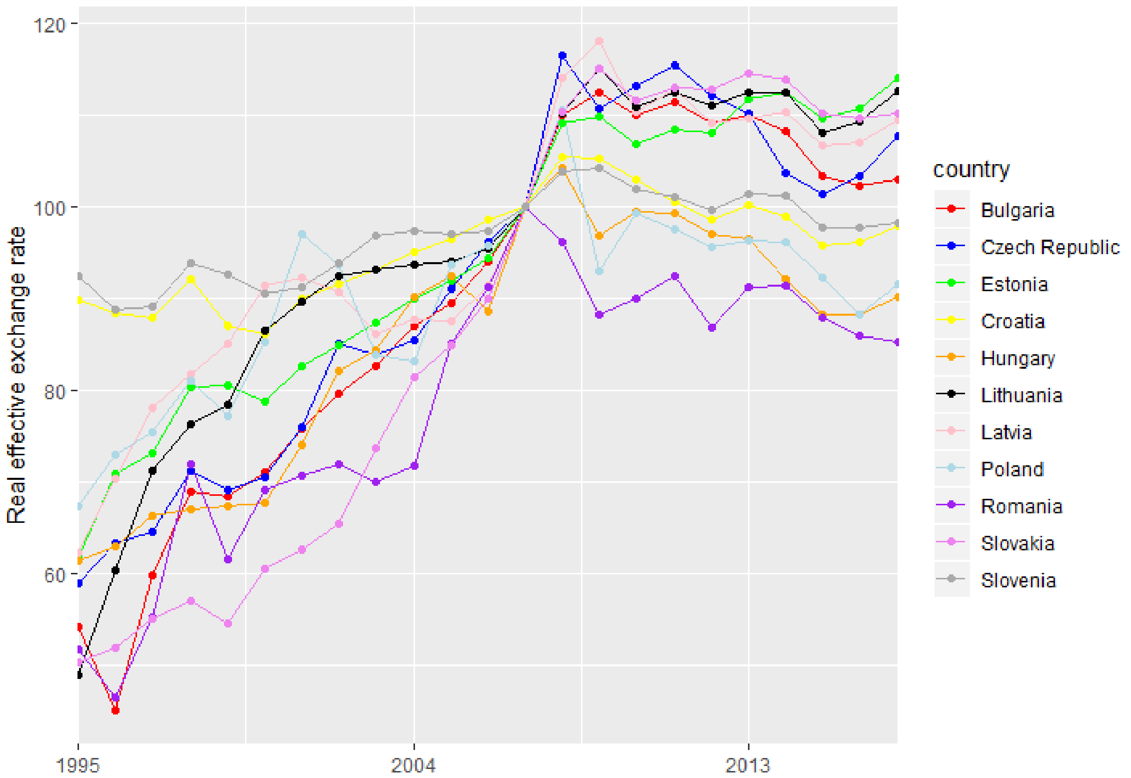
World Bank, World Development Indicators. (2019). NY.GDP.PCAP.KD. https://databank.worldbank.org,24.09.2019

Zeileis A (2019). _pwt9: Penn World Table (Version 9.x)_. R package version 9.1-0, <URL:

https://CRAN.R-project.org/package=pwt9>.

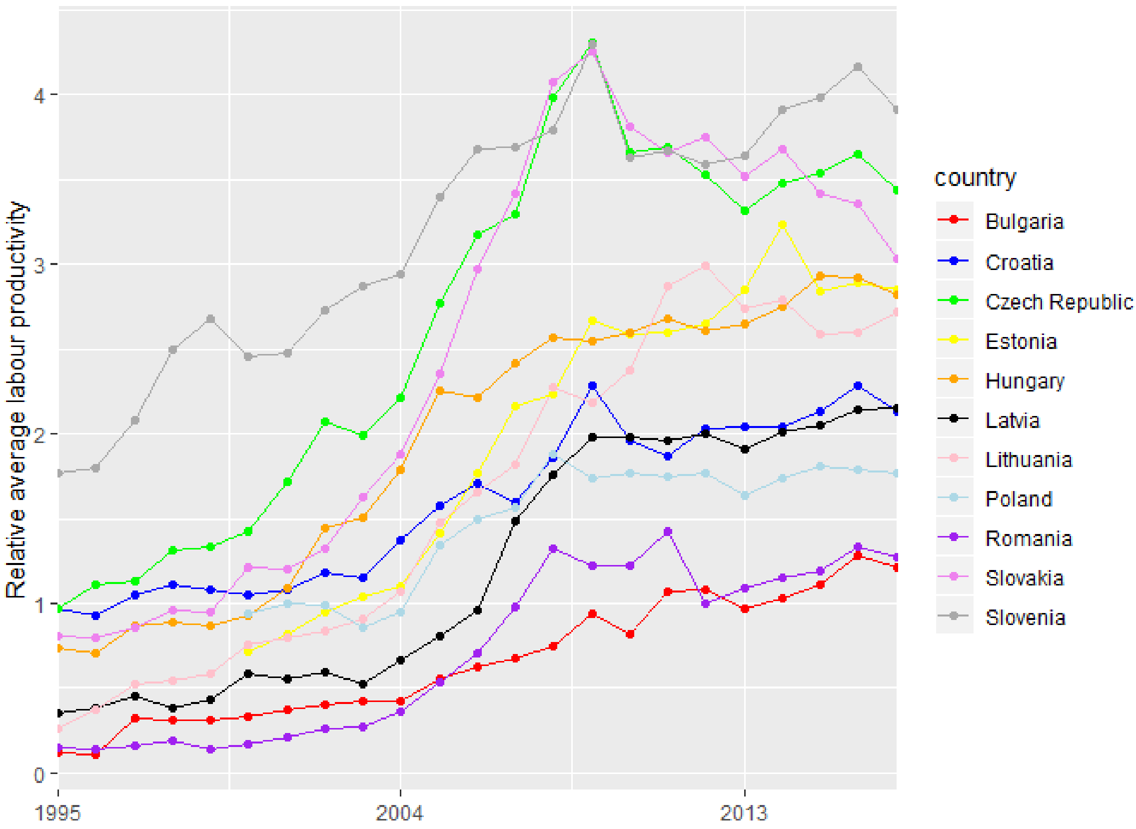
Appendix B: Graphs

Graph 1: Real effective exchange rate - Annual Data



Source: Bruegel data, own plot

Graph 2: Relative average labour productivity in the tradable sector



Source: Eurostat, own plot

Appendix C: Tables

Table one: Literature review of empirical studies with focus on Central and Eastern European countries

Authors (year of publication)	Period	Central Eastern and European Countries considered	Methodology	Empiric Result
Egert (2002)	1991 to 2001	Czech Republic, Hungary, Poland, Slovakia and Slovenia	Johansen Cointegration	Some countries no BS-effect, weak results confirmation for BS-effeect
Egert et al. (2003)	1995 to 2000	Croatia, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia, and Slovenia.	Panel Cointegration	BS-effect is no bigger than other factors influencing the exchange rate
Baszkiewicz et al. (2004)	1995 to 2003	Bulgaria, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia, and Slovenia	Pane FMOLS, Panel PMGE	Below 3 percentage points, Potentil for negative effects in the future
Mihaljek and Klau (2004)	~1993 to 2001	Croatia, the Czech Republic, Hungary, Poland, Slovakia and Slovenia	Ordinary least square estimation	5.72 percentage points
Egert (2006)	1993 to 2003	Bulgaria, Romania, and Croatia	Dynamic ordinary least squares	0.8 percentage points Croatia, 1.9 Romania, Bulgaria might be slightly negative
Mihaljek and Klau (2008)	~1997 to 2008	CEEC without Albania	ADF	1.2 average BS-effect in percentage points
MIletic (2011)	~1993 to 2010	Bulgaria, Croatia, Czech Republic, Hungary, Macedonia, Poland, Romania, Slovenia, and Slovakia	Johansen und Pedroni panel cointegration	1 average BS-effect in percentage points
Konopczak (2013)	1995 to 2010	Czech Republic, Hungary, Poland, and Slovakia	Panel, Cointegration	2.13 average BS-effect in percent points
Sonora and Tica (2014)	2000 to 2009	CEEC without Albania	Panel, VECM	BS-effect a third of Impact of Terms of trade
Adamek (2018)	1995 to 2015	Bulgaria, the Czech Republic, Estonia, Croatia, Hungary, Latvia, Lithuania, Poland, Slovakia and Slovenia	ADF, OLS/HCE	0.39 average Balassa-Samuelson effect

Source: Literature, references are in Appendix A

Table two Changes in relative average labour productivity in the tradeable sector

Country	Overall change in percent	Average annual change in percent
Bulgaria	231.52	11.59
Croatia	78.40	3.56
Czech Republic	127.02	5.78
Estonia	296.12	8.10
Hungary	135.14	6.14
Latvia	182.29	8.29
Lithuania	234.58	10.66
Poland	87.39	3.69
Romania	217.01	9.86
Slovakia	138.88	6.04
Slovenia	79.74	3.62

Source: own calculations

Table three: Results Panel Unit Root test for panel 2

Panel Unit Root test applied	Real effective exchange rate	Relative average labour productivity of the tradeable sector
Levin et al. (2002)	-4.5736 (2.397e-06)	-5.5651 (1.31e-08)
Im et al. (2003)	-3.4096 (0.0003252)	-3.525 (0.0002117)

Notes: (1) results provided: Test-statistic (according p-value), (2) all results reject the null hypothesis at 5% level of significance; Source: Own estimations

Table four: Results Panel Unit Root test for panel 3

Panel Unit Root test applied	Real effective exchange rate	Relative average labour productivity of the tradeable sector
Levin et al. (2002)	-6.0432 (7.554e-10)	-4.4146 (5.059e-06)
Im et al. (2003)	-3.4585 (0.0002716)	-1.4743 (0.07019)*

Notes: (1) results provided: Test-statistic (according p-value), (2) all results reject the null hypothesis at 5% level of significance but one which * denotes. This result rejects the null hypothesis at 10% level of significance; Source: Own estimations

Table five: Elasiticities of effects and theirs significance estimated in the error corrected model with fixed effects

Countries	Elasticities	p-Value of Wald Test on significance
Bulgaria	0.007548987*	1.464840e-01
Czech Republic	0.024071718	3.179730e-03
Croatia	0.007215218	4.901704e-02
Estonia	0.001730290*	5.156825e-01
Hungary	0.019004023	2.325852e-05
Latvia	0.072802408	2.617104e-03
Lithuania	0.002959828*	4.562749e-01
Poland	0.085604718	1.810329e-02
Slovakia	0.049528669	2.833296e-05
Slovenia	0.034361634*	1.114810e-01
Romania	-0.000561543*	5.841308e-01

Notes: (1) all results are significant at 0.05 confidence, (2) notation * implies statistically meaningfully different from zero at .05 level; (3) Wald test H_0: effect is zero; Source: own estimations

Appendix D: R-script

```
###data prep

## rpro, tot,emp,gdpc,reer from excel

library(readxl)

allvaribalesbuttot <- read_excel("data/allvaribalesbuttot.xlsx",)

library(dplyr)

datr <- read_excel("data/allvaribalesbuttot.xlsx")

#change two varibale to numeric

datr$hwin<-as.numeric(datr$hwin)

datr$hwag<-as.numeric(datr$hwag)

#transformations

datr$prova<- datr$vaagricp * datr$vaagripi + datr$vaindcp * datr$vaindpi # value added constant prices per country

datr$provaeu<- datr$vaagrieucp * datr$vaagrieuppi + datr$vaindeucp * datr$vaindeupi # same for eu variable

datr$emp<-100-datr$unemp

datr$pro<- datr$prova /(datr$hwin + datr$hwag) # average labeour productivity varibale of countries

datr$proeu<- datr$provaeu /(datr$hwineu + datr$hwagrieu) # average labour productivity varibale eu

datr$rpro <- datr$pro / datr$proeu #relative productivty varibale

datr <- select(datr, rpro,year,country,reer,pd,emp,gdpc) # dataframe with pd, emp, gdpcrpro, year, country, reer 1995-2017
11 countries

## TOT from pwt

#install.packages("pwt9")

library(pwt9)

data("pwt9.1")

# 182 countries, reduce to CEEC for now,

data<-subset(pwt9.1, country== "Bulgaria"|country== "Croatia"|country== "Czech Republic"| country== "Estonia"|
country== "Hungary"| country== "Latvia"| country== "Lithuania"| country== "Poland"| country== "Romania"|
country== "Slovakia"| country== "Slovenia")

#picking varibales needed price of import and export varibal,country, year,

data <- select(data, pl_m, pl_x, year, country,)

# restrict years to 1995-2017

data<-data[data$year>1994 & data$year<2018, ]

#create terms of trade variable

data$tot <- (data$pl_x / data$pl_m) * 100
```

```
#lose price o(dat, tot, year, country)f import and export as not needed anyumore

data <- select(data, tot,year,country,) # includes year country tot 1995 to 2017 11 countires

#join data and datr to have df with with all varibaales and year and country

dat<- merge(data,datr)

#write.table(backup, file= "dat.csv",sep="," )

#backup and cleaning enivroment

backup<-dat

rm(datr,data,allvaribalesbuttot,pwt9.1)


####data description

##data for change table

#change year

source("/PanelLag.R") ## source code found at the end of script

#dat<-backup

dat <- dat[order(dat$country, dat$year), ]

dat$rpro_log <- log(dat$rpro)

dat$rpro_log_lag <- panel_lag(x = dat$rpro_log, lag = 1, id = dat$country, time = dat$year)

dat$rpro_growth <- dat$rpro_log - dat$rpro_log_lag

dat1<-dat

rm(dat)

dat1<-na.omit(dat1)

aggregate(dat1$rpro_growth, by=list(dat1$country), mean)

rm(dat1)

#change whole period

dat<-backup

dat <- dat[order(dat$country, dat$year), ]

dat$rpro_log <- log(dat$rpro)

dat$rpro_log_lag <- panel_lag(x = dat$rpro_log, lag = 22, id = dat$country, time = dat$year)

dat$rpro_growth <- dat$rpro_log - dat$rpro_log_lag

dat$aepro<-dat$rpro_growth*100

rm(dat)

##plots
```

```
dat<-backup

library(ggplot2)

#plot reer

ggplot(dat, aes(x = year, y = reer, color = country, shape = country)) +

  geom_line() +

  geom_point() +

  scale_color_manual(values = c("red",
"blue","green","yellow","orange","black","pink","lightblue","purple","violet","darkgrey")) +

  scale_shape_manual(values = c(16, 16,16,16,16,16,16,16,16,16,16)) +

  ggtitle("Graph 1: Real effective excahnge rate - Annual Data") +

  xlab("Year") +

  ylab("Real effective excahnge rate") +

  scale_x_continuous(limits=c(1995, 2017), breaks = seq(1995, 2017,9), expand = c(0, 0))

#plot rpro

ggplot(dat, aes(x = year, y = rpro, color = country, shape = country)) +

  geom_line() +

  geom_point() +

  scale_color_manual(values = c("red",
"blue","green","yellow","orange","black","pink","lightblue","purple","violet","darkgrey")) +

  scale_shape_manual(values = c(16, 16,16,16,16,16,16,16,16,16,16)) +

  ggtitle("Graph 2: Relative average labour productivity in the tradable sector - Annual Data") +

  xlab("Year") +

  ylab("Relative average labour productivity") +

  scale_x_continuous(limits=c(1995, 2017), breaks = seq(1995, 2017,9), expand = c(0, 0))

rm(dat)

###### panel unit root

#setup of balanced panels

dat<- backup

#add differnecing

dat$reer<-log(dat$reer)

dat$rpro<-log(dat$rpro)

dat<-dat[dat$year>1999 & dat$year<2018,]

#panel1 all countries resticted years)
```

```
library(plm)

##Panel 2 all countries resticted years

#prep for unit root

rm(pat)

pat<-dat

str(pat)

pat$country<-as.integer(factor(pat$country))

#transfomr into paneldataframe #object <- as.data.frame(split(object, id))

pat<- pdata.frame(pat, index = c("country","year"))

#Panel 1

#reer

ram= data.frame(split(pat$reer,pat$country))

purtest(ram, pmax = 4, exo = "intercept", test = "levinlin", lags="SIC",)

purtest(ram, pmax = 4, exo = "intercept", test = "ips", lags="SIC")

rm(ram)

#rpro

ram= data.frame(split(pat$rpro,pat$country))

purtest(ram, pmax = 4, exo = "intercept", test = "levinlin", lags="SIC",)

purtest(ram, pmax = 4, exo = "intercept", test = "ips", lags="SIC")

rm(ram,dat,pat)

#Pnael 3

dat<- backup

#add differnecing

dat$reer<-log(dat$reer)

dat$rpro<-log(dat$rpro)

dat<-subset(dat, country== "Croatia"|country==  "Czech Republic"| country==  "Hungary"| country==  "Latvia"| country==
"Lithuania"| country==  "Romania"| country==  "Slovakia"| country==  "Slovenia")

#panel1 all countries resticted years)

library(plm)

##Panel 3 all year less countires

#prep for unit root

rm(pat)

pat<-dat
```

```
str(pat)

pat$country<-as.integer(factor(pat$country))

#transfomr into paneldataframe #object <- as.data.frame(split(object, id))

pat<- pdata.frame(pat, index = c("country","year"))

#Panel 1

#reer

ram= data.frame(split(pat$reer,pat$country))

purtest(ram, pmax = 4, exo = "intercept", test = "levinlin", lags="SIC",)

purtest(ram, pmax = 4, exo = "intercept", test = "ips", lags="SIC")

rm(ram)

#rpro

ram= data.frame(split(pat$rpro,pat$country))

purtest(ram, pmax = 4, exo = "intercept", test = "levinlin", lags="SIC")

purtest(ram, pmax = 4, exo = "intercept", test = "ips", lags="SIC")

rm(ram,dat,pat)


###traditoinal panel analysis

#ftest for random effect, estimate random ef

dat<- backup

dat$reer<-log(dat$reer)

dat$rpro<-log(dat$rpro)

#require(foreign)

library(plm)

############VECM fixed effects und Waldtest fuer controlvariable

# Schritt 1: Packages:

library("lfe")

library("stargazer")

library("xtable")

# Schritt 2: Variablen-Typ korrigieren:

str(dat)

dat$year <- as.numeric(dat$year)

dat$country <- as.character(dat$country)
```

```
dat$reer <- as.numeric(dat$reer)

dat$rpro <- as.numeric(dat$rpro)

dat$tot <- as.numeric(dat$tot)

dat$pd <- as.numeric(dat$pd)

dat$emp <- as.numeric(dat$emp)

dat$gdpc <- as.numeric(dat$gdpc)

str(dat)

# Step 3: FE-Modelle:

## Step 3.1: Fixed Effects:

country_fe <- factor(dat$country)

year_fe <- factor(dat$year)

## Schritt 3.2: FE-Modell:

p <- felm(log(reer) ~ log(rpro)*factor(country)+ tot*factor(country)+ pd*factor(country)+ emp*factor(country)+
gdpc*factor(country) | year_fe, data = dat)

summary(p, robust = TRUE)

#Engle grange

res_reg_longrun=resid(p)

#step2

library(urca)

y=ur.df(res_reg_longrun, type="none",selectlags="AIC")

summary(y)

y@teststat

y@cval

# Ho: residuals are nonstationary

# if residuals are stationary, series are cointegrated long run relationship

#vecm works

ram(y)

##estimation long term effect.

#p[["coefficients"]]

stargazer(p, se = coef(summary(p, robust = TRUE))[, 2])

##waldtest

#wald<- waldtest(p, ~ `tot` - 0 | `emp` - 0 | `gdpc` - 0 | `pd` - 0, type = "robust")

#wald[1]
```

```
#wald[2]

#wald[3]

#wald[4]

#wald[5]

#wald[6]

##VECM

# step 4: Effekte:

# step 4.1: tables with effectsn:

effect_bl <- coef(p)[which(names(coef(p)) == "log(rpro)")]

effect_cz <- effect_bl + coef(p)[which(names(coef(p)) == "log(rpro):factor(country)Czech Republic")]

effect_cr <- effect_bl + coef(p)[which(names(coef(p)) == "log(rpro):factor(country)Croatia")]

effect_es <- effect_bl + coef(p)[which(names(coef(p)) == "log(rpro):factor(country)Estonia")]

effect_hu <- effect_bl + coef(p)[which(names(coef(p)) == "log(rpro):factor(country)Hungary")]

effect_la <- effect_bl + coef(p)[which(names(coef(p)) == "log(rpro):factor(country)Latvia")]

effect_li <- effect_bl + coef(p)[which(names(coef(p)) == "log(rpro):factor(country)Lithuania")]

effect_pl <- effect_bl + coef(p)[which(names(coef(p)) == "log(rpro):factor(country)Poland")]

effect_sa <- effect_bl + coef(p)[which(names(coef(p)) == "log(rpro):factor(country)Slovakia")]

effect_se <- effect_bl + coef(p)[which(names(coef(p)) == "log(rpro):factor(country)Slovenia")]

effect_rm <- effect_bl + coef(p)[which(names(coef(p)) == "log(rpro):factor(country)Romania")]

effect_table <- rbind(effect_bl, effect_cz , effect_cr, effect_es,
effect_hu,effect_la,effect_li,effect_pl,effect_sa,effect_se,effect_rm)

effect_table <- data.frame(effect_table)

rownames(effect_table) <- c("Bulgaria", "Czech_Republic", "Croatia", "Estonia", "Hungary" , "Latvia", "Lithuania", "Poland",
"Slovakia", "Slovenia", "Romania")

colnames(effect_table) <- c("Elasticities of Countries")

xtable(effect_table)

# step 4.2: Table p-Values Wald Tests:

t101 <- waldtest(p, ~ `log(rpro)` - 0, type = "robust") # Baseline Bulgarien

t102 <- waldtest(p, ~ `log(rpro)` + `log(rpro):factor(country)Czech Republic` - 0, type = "robust") # Tschechien

t103 <- waldtest(p, ~ `log(rpro)` + `log(rpro):factor(country)Croatia` - 0, type = "robust") # Kroatien

t104 <- waldtest(p, ~ `log(rpro)` + `log(rpro):factor(country)Estonia` - 0, type = "robust") # Estland

t105 <- waldtest(p, ~ `log(rpro)` + `log(rpro):factor(country)Hungary` - 0, type = "robust") # Ungarn

t106 <- waldtest(p, ~ `log(rpro)` + `log(rpro):factor(country)Latvia` - 0, type = "robust") # Lettland
```

```
t107 <- waldtest(p, ~ `log(rpro)` + `log(rpro):factor(country)Lithuania` - 0, type = "robust") # Litauen

t108 <- waldtest(p, ~ `log(rpro)` + `log(rpro):factor(country)Poland` - 0, type = "robust") # Polen

t109 <- waldtest(p, ~ `log(rpro)` + `log(rpro):factor(country)Slovakia` - 0, type = "robust") # Slovakei

t110 <- waldtest(p, ~ `log(rpro)` + `log(rpro):factor(country)Slovenia` - 0, type = "robust") # Slovenien

t111 <- waldtest(p, ~ `log(rpro)` + `log(rpro):factor(country)Romania` - 0, type = "robust") # Rumänien

pvalues <- rbind(t101[1], t102[1], t103[1], t104[1], t105[1], t106[1], t107[1], t108[1], t109[1], t110[1], t111[1])

pvalues <- data.frame(pvalues)

pvalues

rm(t101,t102,t103,t104,t105,t106,t107,t108,t109,t110,t111)#housecleaning

#merge both add names

effect_table <- data.frame(effect_table, pvalues)

colnames(effect_table) <- c("Elasticities of Countries", "p-Value of Wald Test on Significance")

xtable(effect_table)

#vorhinvergessn

rm(effect_bl,effect_cr,effect_cz,effect_es,effect_hu,effect_la,effect_li,effect_pl,effect_rm,effect_sa,effect_se)

#panel_lag source code #thx for the help on this one philipp

panel_lag <- function (x, lag, id, time, other) {

  if(missing(other)) {other <- rep(1, length(x))}

  if(missing(x) | missing(id) | missing(time) | missing(lag)) {stop('All parameters have to be provided. Otherwise calculation is not possible!')}

  if (length(id) != length(x) || length(id) != length(time) || length(id) != length(other)) {stop ("Inputs not same length")}


  X <- data.frame(x, id, time, other)

  colnames(X) <- c("x", "id", "time", "other")

  X <- X[order(X$other, X$id, X$time), ]


  x <- X$x

  id <- X$id

  time <- X$time

   x_lag <- x[1:(length(x) - lag)]

  x_lag[id[1:(length(id) - lag)] != id[(1 + lag):length(id)] ] <- NA

  x_lag[time[1:(length(id) - lag)] + lag != time[(1 + lag):length(id)] ] <- NA

  return(c(rep(NA, lag), x_lag))  }
```

```
#citations of packages

citation("dplyr")

citation("lfe")

citation("plm")

citation("pwt9")

citation("urca")
```