Balance of payments and fiscal sustainability

Balance of Payments and Fiscal Sustainability

Veronika Akhmadieva

Department of Economics, Mathematics and Statistics
Birkbeck, University of London

This dissertation is submitted for the degree of
Doctor of Philosophy in Economics and Finance

June 2020

I would like to dedicate this thesis to my loving family who helped me in all things great and small and to my supervisors, Professor Ron P. Smith and Professor Sandeep Kapur, who supported me every step of the way and got me from the start to finish line.

Declaration

I hereby declare that this thesis is the result of my own work with the fifth chapter being adopted from a paper written jointly with Professor Ron P. Smith, who kindly gave his permission to use material from the paper for this thesis. The thesis has not been submitted, in whole or in part, for consideration for any other degree or qualification at this university or any other institute of learning and complies with Birkbeck, University of London guidelines on length and format.

Veronika Akhmadieva
June 2020

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Professor Ron P. Smith, who invested enormous amount of time and effort guiding me through my MPhil/PhD program. I thank him for consistently tracking my progress and providing me with everything I need to successfully cross the finish line. It was an honour to have such an exceptional supervisor and a great mentor. I am also grateful for an opportunity to write a joint paper with Professor Ron P. Smith, it was a very valuable learning experience.

I am also extremely grateful to my second supervisor, Professor Sandeep Kapur, for supporting me from day one and for constantly keeping me on track. I cannot thank him enough for everything he has done for my teaching career and studying at Birkbeck.

I would like to thank my family without whom this whole experience would have never been possible. I am very grateful to my loving parents, my first two teachers, who not only encouraged me to pursue my dreams and goals, but also gave me every opportunity to do so, and who supported me every step of the way. Thank you to my brilliant big sister, who left our childhood home first and, by her own example, inspired me to explore all the wonderful opportunities the world has to offer, and who helped me to take my place in it. Thank you to my strong-willed and wise grandmother who, being a teacher herself, inspired me to try myself as one and to explore the academic world. I am also grateful to my grandfather, who was an embodiment of work-life balance, and who always saw a glass as being half full.

I would like to thank Professor Ray Barrell, who was my personal tutor during my Bachelor's studying, and who set me in the right direction and always gave me wise and down-to-earth advices.

Thank you to all the academics at our department and outside Birkbeck, who shared their knowledge with me and provided useful comments on my thesis. I am also grateful to my colleagues who were always there when I needed help, and thanks to whom these years of studying towards PhD among other things, were also fun and exciting. Last but not least, I thank my students, whose questions and in-class discussions inspired me to constantly deepen my knowledge.

I am very grateful to everyone who in any way helped me in writing this thesis. I truly appreciate every contribution, big or small.

Abstract

This thesis focuses on sustainability of the balance of payments as well as government surplus and debt.

Chapter 2 uses the standard model for exports and imports demand proposed by Khan (1974) and used by Haynes and Stone (1983) and many others. The income-trade relationship is first analysed using WTO and IMF data for 15 countries for 1970-2013. We estimate ARDL/ECM models and find evidence of strong short-run and long-run effect of income on trade. Focusing on the UK and US cases, for which longer trade data is available from ONS and BEA, we adopt the theoretical framework proposed in Boyd et al. (2001) and find that the Marshall-Lerner condition holds only for the UK. This raises the question of how the balance of payments adjusts.

Chapter 3 examines adjustment directly using a reduced form autoregressive model of feedback for the balance of payments for a panel of 17 countries for 1870-2016. The data comes from the Jorda-Schularick-Taylor (JST) Macrohistory Database. The theoretical framework is based on the works by Bohn (1998, 2007) and Obstfeld and Rogoff (2002). Pooled estimates (allowing for cross-country fixed effects) for both, the balance of trade and current account models, suggest that both processes are stabilising. This provides strong evidence of the balance of payments sustainability. There appears to be roughly symmetric adjustment by imports and exports. When the data are disaggregated by country and sub-sample, the evidence is somewhat more mixed.

Chapter 4 examines the government surplus (deficit) and debt, which raises similar sustainability issues to the balance of payments. Again a reduced form autoregressive specification and the JST data are used to analyse the sustainability of the public sector surplus and debt. Pooled estimates suggest that public surplus is stabilising on its own lagged values, however unlike the balance of payments, the adjustment is not symmetric but done mainly by revenue. However, when we consider heterogeneous estimates, the feedback from debt is mostly insignificant. It seems that the level of debt a government can maintain mostly depends on the credibility of the borrowing government. In addition, we find that there is stronger pressure to adjust on countries running current account (or public) deficits versus those running surpluses.

Chapter 5 compares stand-alone European countries that have control over their interest rate and exchange rate, which are traditionally seen as main instruments of adjustment, with European Monetary Union members that do not. We investigate whether joining the euro caused structural changes to promote adjustment. We construct counterfactuals using both single equation models and a six equation vector autoregression with foreign exogenous variables. We find no major differences between the two groups, thus, other adjustment mechanisms than structural change must be at play.

There is a conclusion that discusses the contribution that the use of long span heterogeneous panels can make to the analysis of sustainability.

This thesis was written before the pandemic associated with the coronavirus. The pandemic disrupted trade and prompted large increases in government debts and deficits. These events will have important implications for the analysis of balance of payments and fiscal sustainability. This topic, while not covered in this thesis, might be an area for future research.

Table of contents

List of figures

List of tables

Chapter 1

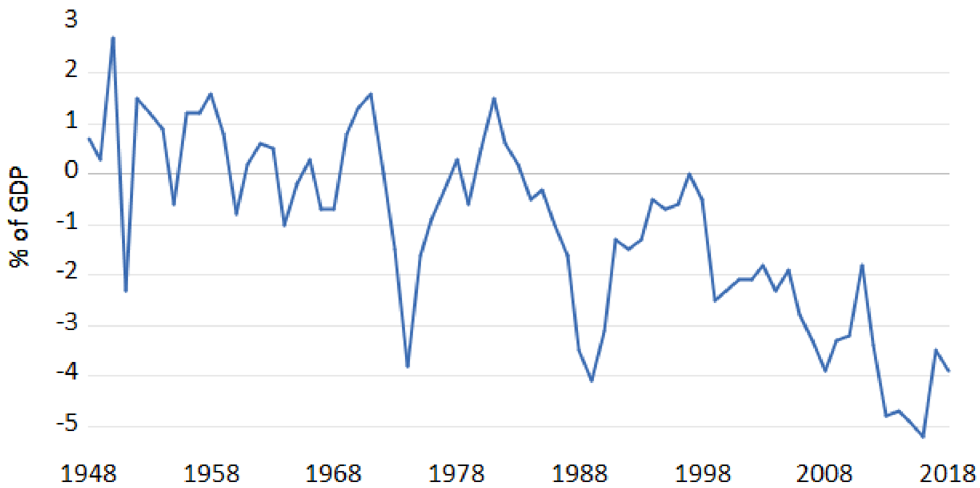
Introduction

It is often argued that certain levels of balance of payments deficits, budget deficits or debt-GDP ratios are unsustainable. For instance, the Maastrich euro convergence criteria put certain restrictions on the European Union members, such as the debt-to-GDP ratio has to be less than 60%, government budget deficit relative to GDP must not exceed 3% as well as being in the exchange rate mechanism under the European Monetary System (EMS). Having certain constraints on how much a nation or a government can borrow proved to be of a vital importance during the 2007-2008 crisis.

As such, the initial motivation for the thesis was to analyse the adjustments associated with the slowdown in the growth of world trade after the 2007-2008 crisis. The period since the crisis has threatened the long-run trend towards globalisation and was characterised by the outbreak of tariff wars and increase in protectionism. In addition, some countries have seen deterioration in their balance of payments positions which are regarded as unsustainable. For instance, Figure 1.1 shows the downward trend in the UK balance of payments in 1948-2018. The UK has had a persistent current account deficit since the 1990s. In fact, it has been reporting the balance of payments deficit every year since 1984, except 1997. The trend deterioration in the balance of payments since the 1970s does not seem sustainable. Similar graphs could be provided for other countries, such as the US. These deficits are matched by increasing surpluses elsewhere.

In standard open economy models balance of payments adjustment is normally thought of as done through the exchange rates. A large balance of payments deficit represents a higher demand for foreign goods and services and causes a real exchange rate depreciation. Meanwhile, government deficit adjustment is usually thought to happen through the changes in the interest rates, such that if governments borrow too much, interest rates for them go up, making further borrowing more expensive. However, do these adjustments actually happen and if so, what are the main adjustment mechanisms? In this thesis we analyse sustainability

Fig. 1.1 UK Balance of Payments: Current Account Balance as Percent of GDP (1948-2018; Annual Data)



Source: Office for National Statistics (2017).

condition and the adjustment processes for the balance of payments and government debt using a long span of data for many countries.

We begin our analysis by searching within the context of what is the standard model, general demand for exports and imports equations proposed by Khan (1974) and used by Haynes and Stone (1983) as well as many others. Within the context of this model, adjustment happens through the real exchange rate and Marshall-Lerner (ML) condition, according to which a depreciation will improve the current account balance as long as the sum of the demand elasticities for exports and imports is greater than one (in absolute value). In the beginning of the second chapter, as in the three chapters that follow it, we provide a literature review on the topic discussed in the chapter. When we estimate our models we check how sensitive the results are by considering different sub-periods and comparing estimates obtained by analysing data from two different data sources. We also check whether the Marshall-Lerner condition holds and compare these findings with ones found in the literature on this subject.

Moving to the empirical part, we first estimate a general model for exports which has a function of the world and domestic incomes, foreign and domestic consumer price indices and exchange rate. We use World Trade Organisation (WTO) data for 15 countries for 1970-2014. Aiming to end up with one common model for all 15 economies, we follow basic steps of the algorithm proposed by Hoover and Perez (1999) and drop variables with insignificant coefficients in order to move from the general unrestricted model to the specific model. However, this analysis produces mixed results, which are sensitive to the model specifications, sample size and estimation period.

Nonetheless, world income effect on exports appears to be significant for most countries in our sample. Hence, we then focus on the income effect and estimate a simple error-correction model (ECM) for exports with world income as the only explanatory variable. The ECM model produces estimates which are more consistent with the economic theory, but the results on the long-run effect of income on trade are mixed.

We also compare our estimates with the paper written by Constantinescu et al. (2015). They follow similar framework, but use International Monetary Fund data. When we try to estimate short-run and long-run income elasticity of trade using two datasets, WTO and IMF, for the same sample and estimation period we find that there seems to be a relationship between IMF and WTO short-run and long-run elasticities, more so in case of short-run ones. However, the actual values of these coefficients seem to vary depending on the dataset.

To check sensitivity of the estimates to the sample period we run structural break analysis and also check whether a common breakpoint might be around the crisis 2007-2008. We find that breaks occur at different years in different countries, but when we split our data in pre-2000 and post-2000 sub-periods, most short-run and long-run coefficients are significant. Hence, for our data 2000 appears to be a suitable year for a common breakpoint.

By this stage it became clear that the longer data could substantially improve to what extent one can reliably estimate the effect of income on trade. While availability of sufficiently long data is an issue for the analysis of the relationship under consideration, fairly long quarterly data are available for countries such as the UK and US (1955Q1-2018Q1 for the UK and 1950Q1-2018Q1 for the US). Hence, we use these data for UK and US to further analyse income-trade relationship and check whether Marshall-Lerner condition holds for these two countries. We compare our estimates with the results from the paper by Hooper et al. (2000), who obtained fairly stable income-trade elasticities estimates for the G7 countries. We find that our short-run and long-run income elasticity estimates are somewhat different from theirs, but the effect of income on trade is once again apparent. We also find that the Marshall-Lerner condition holds in case of the UK, but not for the US.

Overall, there is strong evidence of the significant effect of income on trade when we use sufficiently long data. It is, however, more challenging to establish significance of the income effect on trade over short sub-periods, and it appears that our estimates are sensitive to the model specifications, estimation period and dataset we used. Correlations between variables may have caused part of this sensitivity. We also find that the Marshall-Lerner condition holds for the UK, but not for the US. Hence, it seems the evidence of an Marshall-Lerner type of exchange rate adjustment are mixed.

However, since we believe that balance of payments deficits cannot explode, there must be some sort of mechanism ensuring its sustainability. The initial analysis of the balance of trade components (exports and imports) shows that more structural models do not seem to be working very well, hence we then consider a reduced form autoregressive specification instead. In the third chapter we look at the feedback directly and over long period for many countries to see whether we could improve precision of the short-run estimations by averaging over time and countries. This uses the panel (JST data[1]) of 17 countries over 1870-2016. In addition, to further analyse the adjustment patterns we also focus on three sub-periods, 1870-1914, 1915-1950 and 1951-2016.

The JST dataset has been used for many purposes, but to the best of the author's knowledge, not for the study of solvency. In the main theory of solvency, due to Bohn (2007), the adjustment is typically seen in terms of stock (net debt) rather than flow. In case of the balance of payments the stock is net foreign assets of a country. These data are not available for the long span covered by the JST panel, which is an issue as we have established that there is a big advantage of analysing solvency using a big panel and long span data. Hence, we approach solvency of the balance of payments in terms of autoregressive feedback mechanism using the balance of trade and current account data instead.

We consider a variable y_t with a steady state value y_t^*, there is stabilising adjustment if $\lambda > 0$ in the error-correction/autoregressive model

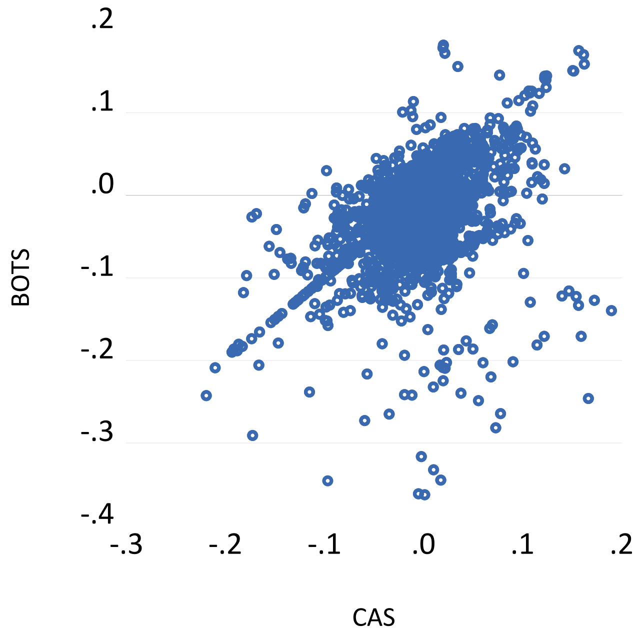
$$\Delta y_t = \lambda (y_t^* - y_{t-1}) + \nu_t.$$

We use two proxies for the balance of payments: the current account and the balance of trade. The current account data are directly available in the JST dataset, while balance of trade is computed as the difference between nominal exports and imports, both are expressed as shares of GDP. It is reasonable to try these two different proxies for the balance of payments, as they seem to be quite different from each other (Figure 1.2) with the correlation coefficient of 0.44, which is relatively low. Not surprisingly, these two proxies for the balance of payments produce quite different results.

Building on the theoretical framework proposed by Bohn (2007), we use autoregressive models for the balance of trade and current account to check if these two processes are stabilising on their own lagged values. If they do, then we can conclude that the necessary

[1]Oscar Jordà, Moritz Schularick and Alan M. Taylor. 2017. "Macrofinancial History and the New Business Cycle Facts." in NBER Macroeconomics Annual 2016, volume 31, edited by Martin Eichenbaum and Jonathan A. Parker. Chicago: University of Chicago Press.

Fig. 1.2 Relationship between the Current Account and the Balance of Trade (Expressed as Shares of GDP; 1870-2016; 17 Countries)



Source: Jordà-Schularick-Taylor Macrohistory Database. (2019).

corrective adjustments are taken in order to prevent the balance of payments deficits from exploding.

We find strong evidence of sustainability for the balance of trade for pooled data for 17 countries for the whole period, 1870-2016. It appears that the balance of trade adjusts by about 10% per year to maintain a sustainable level. The adjustment done by its components, exports and imports, is about symmetric with both adjusting by about 5% per annum.

If the model is estimated for individual countries over the whole period, the sustainability condition is confirmed for more than half of the sample, 10 out of 17 countries. However, it appears that sufficiently long data are required to assess sustainability of the balance of trade as there is less evidence of stabilisation if the model is estimated for individual countries over shorter sub-periods.

When we consider the average adjustment to equilibrium across countries it falls from 35% per year in 1870-1914 to 34% in 1915-1950 to 16% in 1951-2016. But the mean is sensitive to outliers of which there are a number in the first two sub-periods, for instance, in case of Japan 74% of the difference from equilibrium is removed each year when we consider the first sub-period. When the country-specific equations are estimated over the whole period, 1870-2016, the average annual adjustment is 15%. This is similar, though slightly larger than the homogeneous fixed-effect panel estimate of 10%. It is possible that under the Gold Standard there were much stronger pressures to adjust rapidly than under the more recent floating rate period.

The sustainability condition is also confirmed for the current account when the autoregressive model is estimated over 1870-2016. Then the adjustment of the current account components is considered by estimating individual models for exports shares of GDP, imports shares of GDP and for the non balance of trade elements of the current account represented by z_t.

The adjustment of current account seems to be largely by imports (10% per year) and exports (7%), with a smaller portion (4%) of the annual adjustment been done by the other components of the current account. For the pooled data current account adjustment also comes to 21% per year. Heterogeneous estimates some support for the stabilisation of the current account, but less so when we estimate the model over shorter sub-periods. Nonetheless, over the whole period the current account appears to be a stationary process in 13 out of 17 countries and the average adjustment is about 21% per year. Moreover, in 14 out of 17 economies there is evidence of stationarity in at least one of the sub-period. The exceptions are three countries, for one of which there is insufficient number of observations in the first sub-period.

We also consider whether there might be more pressure to adjust on the countries running deficits versus those running surpluses. Hence, we check for the asymmetry in the adjustment towards equilibrium, and based on the estimations for pooled long span data the annual adjustment in countries running deficits is twice higher than in those running surplus (26% versus 13% per year). Hence, we conclude that there does seem to be asymmetric adjustment. There is also evidence of asymmetric adjustment when we consider 17 countries individually and estimate the model over the whole period. Nonetheless, when we consider shorter sub-periods we run into an issue of insufficient number of observations for seven countries for the first sub-period and in general most coefficients appear to be insignificant when we reduce sample size and/or estimation period.

Overall, we conclude that there is strong evidence of sustainability of the balance of trade, current account and, hence, the balance of payments, when we analyse pooled long span data.

Under traditional view, if an economy is running balance of payments deficit, besides exchange rate depreciation and increase in the interest rate, another important part of the adjustment process of the balance of payments back to the equilibrium, is to reduce the government deficit, as it often goes hand in hand with the trade deficit. As government spending increases, part of it is spent on imports, hence, worsening an economy's balance of payments position. Thus, we consider whether there is stabilisation on the fiscal policy side.

Even though we do not have sufficiently long data for our sample for foreign assets to estimate sustainability for the balance of payments in terms of net debt, we do have the stock of debt data for the analysis of the fiscal sustainability. Thus, in the fourth chapter we look at the solvency condition in terms of sustainability of the government debt and surplus and analyse how the government surplus responds to both, the lagged surplus and stock of debt. As in the chapter on the balance of payments sustainability, we are building our theoretical framework on Bohn's (2007) paper. The algebra in terms of theoretical framework for the balance of payments and fiscal sustainability is similar. However, there are two main differences.

Firstly, from the economics side in the balance of payments case the net foreign assets have to add to zero across countries. Therefore, there are lenders (countries who have positive net foreign assets) and borrowers (those with negative net foreign assets). In the public sector case, in general, governments only have debt, they are mainly net borrowers not net lenders. Secondly, balance of payments deficit can be financed through exchange rate adjustments or changing reserves and borrowing. As for the government deficit, it can be financed through issuing new debt (there may be interest rate adjustment) and printing money. Otherwise, basic theory for our analysis of the solvency condition for the balance of payments and government debt is similar.

In his paper Bohn (2007) emphasises the importance of an error-correction response of surplus to the stock of debt. He argues that if surplus-GDP ratio is a positive linear function of debt-GDP ratio, then one can conclude that a government takes adjusting actions in response to augmentation of debt, so that the debt-GDP ratio remains bounded. Hence, fiscal policy can be considered sustainable. We estimate an augmented error-correction model for surplus, which allows for the feedbacks from not only lagged debt-GDP, but also lagged surplus. We then split the main equation for surplus into two components, revenue and expenditure, to see which of these two components is doing most of the adjustment. We use the same dataset (JST data) as for the analysis of the balance of payments solvency.

When we look at the pooled results for surplus, revenue and expenditure, in each case there is a stabilising feedback on both, the lagged debt-GDP and lagged surplus. However, when we estimate these equations for individual countries, there does not seem to be a significant feedback coming from the lagged debt-GDP. Nonetheless, the heterogeneous estimates suggest that large deficits do tend to adjust themselves, and the adjustment is mainly done through revenue, not expenditure.

In addition, there does not seem to be a natural equilibrium level for the stock of debt, which suggests that the debt-GDP ratio does not have steady state value. In fact, this is consistent with the debt-GDP ratios being very high in some countries and very low in others. For instance, according to the IMF database in 2018 the debt-GDP ratio was about 100% in Belgium and 238% in Japan, while only about 14% in Russia and 22% in Luxembourg (IMF World Economic Outlook, 2018). Thus, we conclude that different countries can have very different debt-GDP ratios, and there is no tendency to converge to some common or standard one. In contrast, a feedback process from the surplus suggests that large deficits do tend to adjust themselves.

We also try to include key macroeconomic variables, namely GDP growth, long-term interest rate and inflation, in the model for surplus. They are likely to be correlated with the debt-GDP and, therefore, their inclusion might alter the feedback on debt-GDP ratio. We find that the coefficient of debt-GDP ratio is significant for the pooled panel data with fixed country effects, but only for two economies when we estimate the model for each country individually.

Hence, while there is a significant feedback coming from debt for the pooled data, the heterogeneous estimates are mixed and overall we cannot know whether the best estimate for the effect of the lagged debt-GDP on the dependent variable is its coefficient or zero. Hence, it seems that on the individual level, in most cases there is a feedback coming from the lagged surplus, but not the lagged debt-GDP, and the adjustment of surplus to its previous values is mainly done through revenue.

Moreover, pooled data estimates provide evidence of the asymmetric adjustment when we consider the whole period. As in case with the current account deficit, there seems to be more pressure to adjust on governments running deficits than on those running surpluses.

We also use a basic national accounting identity that links balance of payments and government surplus together and check how the two adjust to each other. Our estimates suggest little cross-surplus adjustment even when the models are estimated over the whole period, 1870-2016.

In traditional economic theory, in case of balance of payments deficit, governments tend to increase interest rates in order to reduce income. This would be done to deflate the domestic economy in order to reduce demand for imports and bring the balance of payments back to the equilibrium. However, the extent of these adjustment policies depends on the degree of control an economy has over setting its monetary target. This can be quite limited in case a country belongs to a monetary union, such as European Monetary Union (EMU), that does not allow its members to freely adjust exchange rate or short interest rates. The two solvency conditions that we look at, balance of payments as well as government deficit and debt, are central constraints in the Maastricht criteria. Hence, a case of an economy belonging to the monetary union such as EMU is of a particular interest when studying open economy adjustment processes.

Thus, the focal question of the fifth chapter is suppose you remove the nominal interest rate and exchange rate adjustments, what else adjusts and what else changes? The fifth chapter is an expanded version of a paper written jointly with Professor Ron P. Smith[2]. We focus on analysing whether being a member of a monetary union affects country's ability to maintain stable level of inflation, exchange and interest rates. In addition, we check whether results change if we consider different structural breaks.

For this analysis we use data from the GVAR toolbox for 1979Q3-2016Q4, Mohaddes and Raissi (2018), and estimate vector autoregressive models with various specifications for 12 European countries, 8 of which joined the EMU and 4 did not. For each country we use a set of individual variables and their global equivalents. A three-year transition period from national currencies to euro and single monetary policy began in 1999. Hence, the model is estimated up to 1998Q4 and these estimates then used to forecast GDP over the following 72 Quarters.

The results are sensitive to the specification choices, and the confidence intervals around counterfactuals are too large to allow for any specific conclusions about distinctions between EMU and non-member countries. The main structural breaks are found to be in the interest and exchange rate equations, most likely due to a clear institutional change in their determination following the formation of the EMU. However, even after we try treating short interest rate and exchange rate as exogenous, our estimates and actuals remain to be quite different for most countries.

We also estimate Taylor Rule for each of the countries in the sample with short- and long-term interest rates, but again find no major difference between estimates for EMU countries and for non-member economies. In addition, it appears that the actual and predicted estimates are closer in the model with long-term interest rates than in the one with short-term interest rates.

Since joining the EMU did not seem to have caused any structural changes to take account of, there were inevitable tensions to break out within the union and those tensions were reflected in the divergence between the northern and southern members. Hence is the fact that all these years after the formation of the EMU, its members still did not adjust on both sides, such that Germany is still running large surplus and Greece − a large deficit.

[2] A note of gratitude to Professor Ron P. Smith with whom the "The Macroeconomic Impact of the Euro" (2019) paper was written. I appreciate the opportunity to use this paper for the fifth chapter of this thesis.

This thesis makes a number of contributions. We find numerous evidence of strong effect of income on trade in the short and in the long run. However, the actual magnitude of income elasticities of exports and imports vary depending on the dynamic specifications, estimation period and choice of the dataset. It appears that endogeneity and other issues make empirical analysis of the income-trade relationship complicated. When it comes to the analysis of the sustainability of the balance of payments and public finances, using a long span data and many countries seems to produce more precise estimates and indicate that there is on average a stabilising effect for current account, balance of trade and government surplus. We find that there seems to be more pressure to adjust on countries running current account deficit versus those running surplus. Meanwhile, it seems that the government surplus is stabilising on its own lagged values and the sustainability is achieved mainly through the adjustment of revenue, not expenditure. However, it is harder to establish significant feedback when considering each country individually or focusing on shorter estimation periods, because of various confounding variables coming in and different shocks that provide noise around the data. Nevertheless, when we can average over long period and many countries, the stabilising effect becomes apparent. In addition, we find no major differences between stand-alone countries who have control over setting their interest and exchange rates versus EMU members, who do not have such autonomy, suggesting that becoming part of this monetary union at least did not make those countries worse off. We conclude that while interest rates and exchange rates may be important in the adjustment process, there must be other forces at play, but what they are is an open question.

The outline of this thesis is as follows, the second chapter focuses on the income-trade relationship, the third one is dedicated to the balance of payments solvency, the fourth one analyses the sustainability of the government debt and surplus (deficit) and the fifth chapter considers open economy's adjustment processes in case when control over interest and exchange rates are removed. There follows a short conclusion which discusses limitations of the analysis and areas for future research.

Chapter 2

Trade Elasticities

2.1 Introduction

This chapter aims to examine the changing patterns of international trade-income relationship over the last four-five decades. First this chapter estimates short and long-run income elasticity of trade, using an autoregressive distributed lag (ARDL) model and an error correction model (ECM). Then it proceeds by estimating import and export demand elasticities for the United Kingdom and the United States, consequently checking whether Marshall-Lerner condition holds when trade elasticities are estimated over sufficiently long period of time. This analysis of the relationship between national trade flows, world income and the real effective exchange rate is performed using autoregressive distributed lag (ARDL) and vector error-correction models (VECM) and is based on Boyd et al. (2001) and Hooper et al. (2000) papers. Finally, we also check for parameter stability.

To begin with, the literature provides three main rationales for trade: comparative advantage that arises from differences in production technologies (the Ricardian theory), differences in factor endowments (the Heckscher-Ohlin theory) and increasing returns to scale (the New trade theory). There are many factors that might affect trade. For instance, the geographic location of a country may be one of them. The gravity equation, by analogy with Newton's law of gravity, encompasses the relationship between volume of trade, economic size and distance between countries (Tinbergen, 1966). According to the gravity model, countries located close to each other trade more, as transport costs are lower. Some other factors, such size of an economy, common language and similar culture, hence consumer preferences, may also enhance trade between closely located countries. However, studies on the gravity model are largely cross-sectional, meaning use short panel data (Anderson and Van Wincoop, 2003; Disdier and Head, 2008), while an analysis of the changing trade-income relationship requires long span data. Hence, the gravity model related literature is not directly relevant to the empirical analysis presented in this chapter. Finally, some other researchers suggested that possible factors that affect growth are the increase in vertical specialisation (Yi, 2003) and changes in energy prices (Bridgman, 2008).

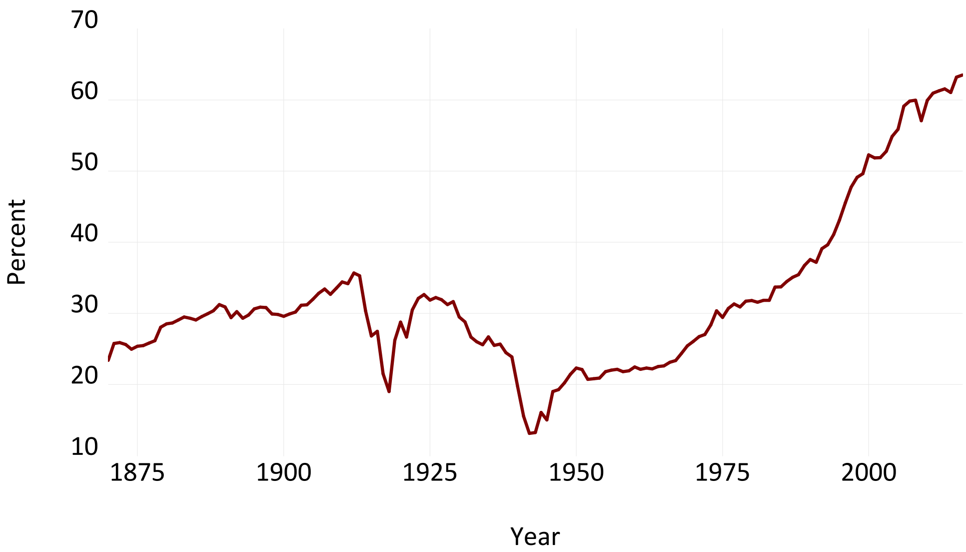
Nonetheless, the one relationship that seemed to be obvious for decades has been the one between trade and world income (Constantinescu et al., 2015). However, a slower recovery of trade growth, relative to the global economy after the financial crisis 2007-2008, called for a more in-depth analysis of the trade-income relationship. This chapter studies the relationship between world income and global trade flows and attempts to answer the questions, such as is there a relationship between trade and income in the short and/or in the long run; was there a structural break in this relationship and if so, when; are the results consistent if data used are from different sources (for instance, from well-known large databases, such as the International Monetary Fund (IMF) and World Trade Organisation (WTO))?

According to the January 2018 Global Economic Prospects (the World Bank, 2018), the world income growth has been higher than expected and was projected to reach 3.1% in 2018. While it appears the global economy has mostly recovered after the financial crisis 2007-2008, trade growth has been fluctuating ever since, struggling to recover to its pre-crisis rate (Caliendo et al., 2015). For instance, trade in global manufactured goods has been increasing since October 2016 and reached 5.2% in September the following year only to fall to 4.9% next month. Hence, it is hard to predict how the volume of global trade will change in the future and what are the factors that determine the patterns of trade growth in the long run.

Taking the UK as an example and looking at the change in the UK trade as percentage of the national GDP since 1870 (Figure 2.1), the increasing dependence of the UK economy on trade since 1960s, comparing to the first half of the 20th century, is obvious. The UK trade has been mostly increasing since 1870, when it accounted for about a quarter of the national GDP, to 2016, when trade nearly reached 70% of the UK income. Nonetheless, there were a few significant drops in the UK trade growth, most notably around the First and Second World Wars with a temporary recovery during the interwar period in the 1920s. This emphasises dependence of the trade growth not only on economic growth, but on political stability in the world, which is, however, difficult to measure.

Looking at the UK exports and imports separately (Figure 2.2), from the 1870s till late 1910s the UK exported more than it bought from abroad. The trade patterns changed following the World War I. Imports decreased from 16% of GDP to barely half of that during late 1910s-early 1920s, while exports dropped from around 22% of national income to approximately 8%. There was a short period of recovery during the interwar years, when the trade growth picked up the pace just to drop to the historical low during the Second World

Fig. 2.1 The UK Trade as % of UK GDP (1870-2016; Annual Data)



Source: Bank of England (2018).

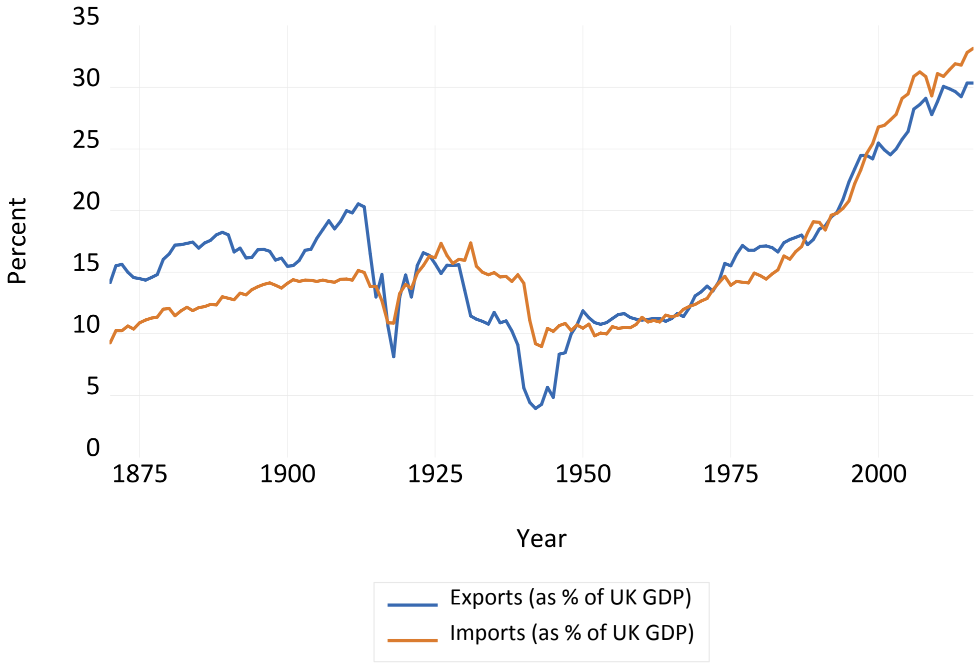
War, with imports decreasing to about 8%, and exports falling to approximately 4% of the national GDP. After those turbulent times the UK mostly run balance of trade deficit, and even though, the trade growth seemed to have been recovering after the global financial crisis 2007-2008, in 2016 the UK still continued to import slightly more than it exported.

Besides economic conditions, export and import volumes depend on the relative prices, which also determine demand for the home and foreign goods. Classic economic theory suggests that depreciation of the home currency should improve the trade balance, as exports become relatively cheaper than imports, and demand for exports increases. The condition for this to happen is known as the Marshall-Lerner (ML) condition.

To check whether the Marshall-Lerner condition holds, the analysis in this chapter uses time-series data for the UK exports (for period 1960Q1-2017Q1) and imports (1955Q1-2018Q1) as well as for the US exports (1960Q1-2017Q1) and imports (1950Q1-2018Q1). Income and price elasticities are estimated using an ARDL model. Moreover, assuming that the variables under consideration might be co-determined, and, hence, endogenous, VAR models are also estimated, followed by the cointegration analysis to estimate long-run effect of income and prices on the trade flows.

Even though this chapter focuses on the imports and exports demand elasticities, that are considered to be determined by the income and relative prices (Boyd et al., 2001), there are many other factors that have influenced the trade growth over time. They include trade

Fig. 2.2 The UK Exports and Imports as % of UK GDP (1870-2016; Annual Data)



Source: Bank of England (2018).

liberalisation, globalisation and technological progress, but not all these factors can be easily estimated numerically. Overall, even when we focused only on the relationship between trade flows, income and prices, it proved to be a challenge to empirically analyse this relationship in the long run due to complex nature of the trade growth and noise in the data.

The outline of this chapter is as follows, section 2.2 reviews the literature and provides some theoretical background on the changing income-trade relationship. In section 2.3 we discuss model specifications and methodology. Section 2.4 discussed the data used in this chapter. General-to-specific modelling (ARDL model) results are presented in section 2.5 and specific-to-general modelling (ECM) estimates are in section 2.6. Section 2.7 covers structural stability analysis and section 2.8 compares results obtained using two different datasets, WTO and IMF. Sections 2.9 and 2.10 review two specific cases we focus on with former section being dedicated to the UK case and the latter one being devoted to the US case. Section 2.11 contains some concluding comments.

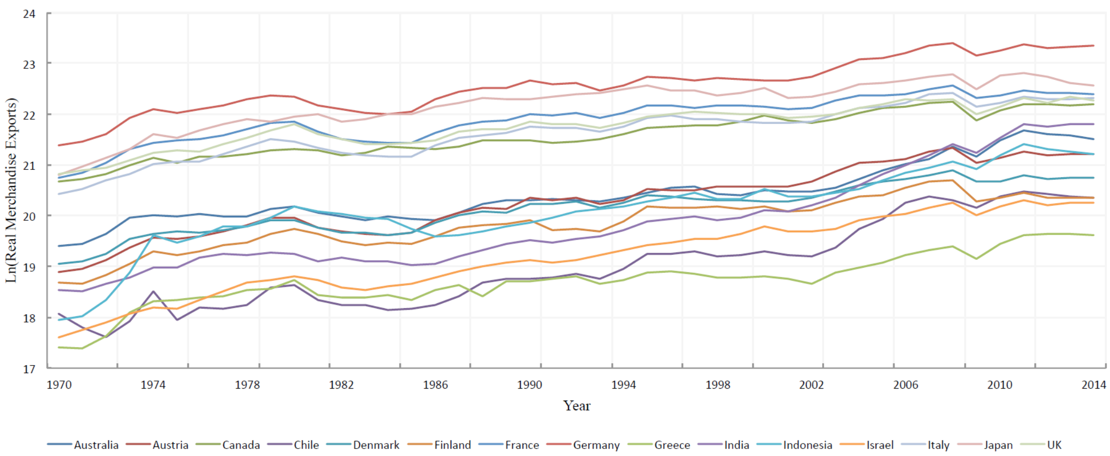
2.2 Literature review

2.2.1 The Global Trade Growth: Overview

From the early 1980s until the financial crisis 2007-2008, trade grew about twice as fast as GDP (Caliendo et al., 2015). The trade growth dropped below GDP level to its historic low of the Great Recession 2.2% in 2010. If trade continued to expand in accordance with its historical trend, it would have been about 20% above its actual level in 2014, 2.8% (Constantinescu et al., 2015).

To get an idea as to how trade flows and global income changed during 1970-2014, a few graphs were plotted. Looking at total merchandise exports for 15 countries (Figure 2.3) that are used in the sample for this chapter, the trade flows from these countries have been growing relatively slowly from 1970 and declined around 1980, when the Oil Crisis (1979) and a few other economic disturbances led to the 1980s recession.

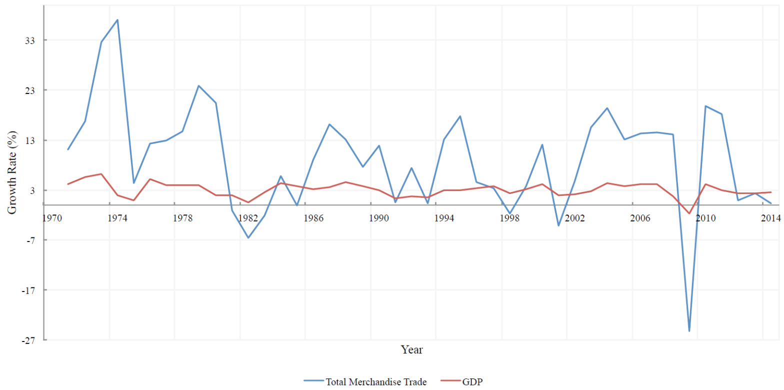
Fig. 2.3 Total Merchandise Trade (Real Exports; 1970-2014)



Source: WTO (2017).

The only exceptions were Japan and Australia which experienced a relatively sharp growth during this time. Since it was the US, the world's largest importer, that suffered from a deep recession between 1980 and 1982, it affected European exporters more than Australia and Japan, which traded less with the US. However, there were a few other growth spans in merchandise exports, experienced by the majority of countries in the next two decades. The exports were rising dramatically, beginning around 1985 and until 1992, the year of the European Currency Crisis, and then from 2003 to 2007, when the global financial crisis struck. It appears that starting from around 2012, trade flows began to flatten off. Will trade growth return to its pre-crisis level remains to be seen.

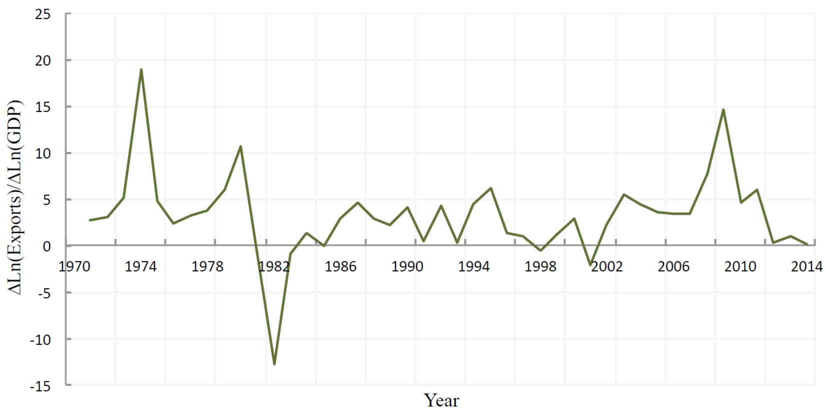
Fig. 2.4 Total Merchandise Trade and GDP (World; 1970-2014)



Source: WTO (2017).

Considering the relationship between world GDP and trade growth rates (Figure 2.4), one can see that on average trade grew faster than GDP in 1970-2014. However, trade growth fell dramatically during the financial crisis 2007-2008, much more so than world output growth. Overall, the trade-GDP ratio (Figure 2.5) is very volatile, reaching nearly 18% around 1974 and falling to disappointing -13% less than a decade later. Nonetheless, changes in trade growth seem to be affected by the economic conditions in the world.

Fig. 2.5 Income Elasticity of Trade



Source: WTO (2017).

For instance, the negative effect of the financial crisis 2007-2008 persisted for about five years, with the demand for imports being 19% and GDP being 4.5% below the predicted level (in the absence of crisis). There were many factors behind the dramatic collapse of trade growth following the crisis, some were cyclical, while others were structural in their nature. For instance, a weak demand, an example of a cyclical factor, was thought to be one of the main reasons for the global trade slowdown and accounted for up to 90% of the contraction (Constantinescu et al., 2015). As for the structural factors, one example is a changing relationship between global trade and income. For instance, according to Constantinescu et al. (2015) trade became less sensitive to changes in GDP with long-run income elasticity of trade falling from 2.2 in 1986-2000 to about 1.3 in 2001-2013. Hence, it appears that on one side, the decrease in income growth might have contributed to the trade slowdown, but on the other hand, trade might have become less responsive to changes in GDP.

There are many possible reasons for decline in the income elasticity of trade following the crisis 2007-2008. One example is the change in the composition of aggregate demand. The income elasticity of trade is a weighted average of import elasticity of individual aggregate demand components, thus the change in composition of the aggregate demand shifts the overall income elasticity of trade (Anderton and Tewolde, 2011; Bems et al., 2013). Thus, the weak recovery in the business investment, which is considered to be the most import-intensive component of the domestic demand, contributed to a decrease in the income elasticity of trade after the crisis. Meanwhile, financial institutions deleveraged the pressure by cutting back on credit growth to boost liquid assets and impaired credit channels caused trade finance to become costlier (Chor and Manova, 2012).

To summarise, the income elasticity of trade appears to have changed at the beginning of the 21st century. Whether there is still strong effect of income on trade in the short and/or in the long run is among income-trade relationship related questions empirically addressed in this chapter.

2.2.2 Modelling Trade Growth: Theoretical Background

Nearly two hundred years ago, David Ricardo developed his theory of comparative advantage, and since then the Ricardian model was considered one of the most useful tools in explaining what generated trade growth. David Ricardo stated that the differences in production technologies explain why trade is beneficial to two parties even if one of them has the absolute advantage in all production (Krugman, 2014). A famous paper by Dornbusch et al. (1977) offers a version of the Ricardian model in which free trade equilibrium is characterised by efficient international specialisation with an underlying assumption that goods should be produced in a country where labour is cheaper and exported to another country.

In comparison, Eli Heckscher and Bertil Ohlin argued that rationales for trade come from the differences in factor endowments, such as land, labour and capital (Wood et al., 2009). According to the Heckscher-Ohlin theory, a country exports goods that intensively use its abundant factors and imports goods that are intensive in its relatively scarce factors. While the Heckscher-Ohlin model may be intuitively compelling, this theory received little empirical support. In 1953 Wassily W. Leontief studied imports and exports of the US, which is essentially a capital-abundant country. Nonetheless, he found that the US imports more capital-intensive goods than it exports (Leontief, 1953). This surprising empirical finding is known as Leontief's paradox. Many other empirical studies on the Heckscher-Ohlin model also rejected this theory in favour of models that allow for technological differences (Baldwin, 2008; Bowen et al., 1986).

In the 1970s, new trade theorists proposed an alternative explanation for international specialisation and trade. They analysed the role of increasing returns from bilateral trade and claimed that due to economies of scale, trade could increase for a variety of goods, keeping the average cost of production relatively low. A Nobel-prize recipient Paul Krugman (1979) developed a model based on the economies of large-scale production, which allowed for technological differences among producers. He explained international trade growth through increasing returns, monopolistic competition and factor mobility, which caused a process of agglomeration of certain types of labour in one region or the other. Furthermore, an industry may tend to cluster in one location because of trade impediments, for instance tariff barriers or transportation costs. The obvious advantages of trade, according to this theory, include an increased productivity, a reduction of costs and a wide variety of goods. However, a monopoly setting means first mover advantage creates entry barriers and the potential issue of price discrimination.

Nevertheless, most traditional trade theories are based on assumptions that do not fully account for the characteristics of real world trading patterns, which are affected by many factors, such as "border effects" or simultaneous export and import of the goods of the same product category (Krugman, 2014). Endogeneity issue is another major problem when analysing trade flows, as not only output determines demand for exports and imports (Khan and Khanum, 1997; Khan, 1974), but also trade is a source of economic growth, hence, the two variables are co-determined (Constantinescu et al., 2015; Disdier and Head, 2008). Moreover, often data on trade are not readily available and comes at a highly disaggregate level. It is then left to a researcher to aggregate the information to an industry or country level, which gives a rise to a problem of measuring structural heterogeneity Goldberg and Pavcnik (2016). Nonetheless, addressing endogeneity and heterogeneity issues remains to be one of the major problems of modelling global trade (Bussière et al., 2009).

From the empirical perspective, three main approaches of modelling trade can be distinguished. The gravity model, introduced by Jan Tinbergen (1966), is among the most successful empirical tools used to explain trade flows in the world economy. In its simplest form, the gravity model states that the volume of trade between any two countries is determined by the economic size of these countries and the distance between them. Transport costs and geography have a crucial effect on trade flows among countries and give rise to various trade agreements, such as NAFTA to enhance trade between closely located economies (Krugman, 2014). Empirical validity of the gravity equation received strong support from many economists, such as Disdier and Head (2008), Anderson and Van Wincoop (2003) and Eaton and Kortum (2002). The empirical works on the gravity equation often use short panel data for a wide range of country pairs, where estimates are sometimes obtained from as little as one year of data (Silva and Tenreyro, 2006).

Nonetheless, for a long time the gravity model of trade lacked theoretical foundation. The economic size has hardly any effect on trade patterns in prominent economic theories like the Ricardian or Heckscher-Ohlin models of international trade. Armington (1969) was the first to provide a theoretical basis for the gravity equation by employing the product differentiation by country of origin, the preference structure commonly known as Armington assumption. In addition, some new trade theorists like Helpman (1987) attempted to provide some foundation for the gravity equation by focusing on product differentiation among firms and claimed that the gravity model provides support for the monopolistic competition explanation of trade. Moreover, by obtaining database of 1467 estimates from 103 papers, Disdier and Head (2008) found that, on average, with each 10% increase in distance, bilateral trade falls by about 9%. However, despite strong empirical support for the gravity model, there is still some ambiguity regarding theoretical foundation for the effect of economic size on trade and even fewer theoretical models can explain the role of distance in a current age (Chaney, 2011).

Other macroeconomists focus on major empirical puzzles in international economics that were summarized by Obstfeld and Rogoff (2002). Using a two country stylized model, Obstfeld and Rogoff proposed that trade frictions, such as transport costs, tariff and non-tariff barriers lie behind international macroeconomics puzzles. These puzzles include home bias in trade and equity portfolios, Feldstein-Horioka puzzle (correlation between domestic investment and national saving rates), international consumption correlations, the purchasing power parity puzzle and the exchange rate disconnect puzzle (Obstfeld and Rogoff, 2002). Obstfeld and Rogoff proposition received a substantial quantitative support. Reyes-Heroles et al. (2017), using a multicountry dynamic model of international trade (Eaton et al., 2011), concluded that falling trade costs account for a 69% increase in net trade imbalances in the period 1970 to 2007. Meanwhile, Kortum et al. (2016) argued that eliminating trade barriers resolves a number of these macro puzzles.

While trade frictions may indeed have a significant effect on trade imbalances, the collapse of the trade growth following the financial crisis 2007-2008, and its exceptionally sluggish, comparing to GDP, recovery showed that there must be other factors involved

(Eichengreen, 2010). Analysis of the changing income-trade relationship and of the price elasticity of exports and imports should help to shed some light on the changes in the trade growth, as both, income and prices have direct effect on demand for exports and imports (Khan and Khanum, 1997; Khan, 1974). Hence, while the gravity model and six major empirical puzzles in international economics are of great interest, this chapter focuses on the effect of income and relative prices on trade flows, as the need for further research on this topic is apparent.

2.2.3 Measuring Trade

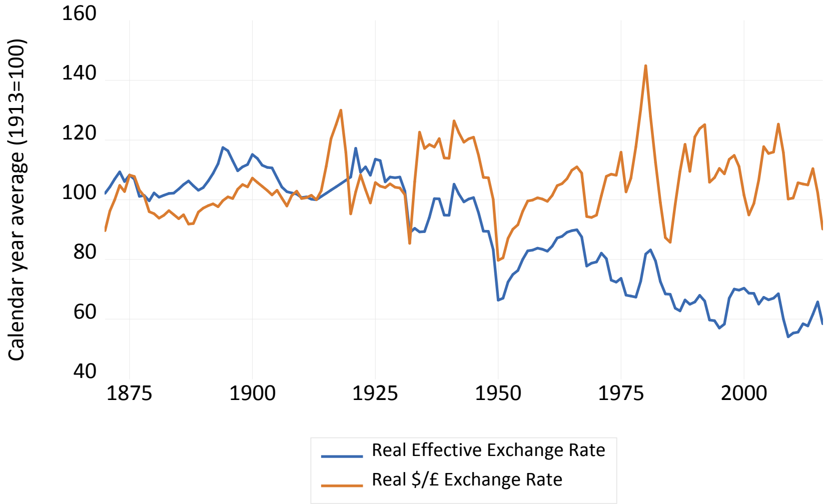
Measuring openness to trade has proved to be a challenge from the empirical perspective. One possible approach is to use a measure of trade policies, such as average tariff rate, as an indicator of openness. However, not all trade barriers are easily quantifiable (Pritchett, 1996). Moreover, trade policies might be correlated with other domestic policies that affect national income and visa versa, hence, giving rise to a potential endogeneity issue.

There are two ways of dealing with the endogenity issue, one is instrumental variables and the other is using vector autoregressive (VAR) models, which allow for all variables to be co-determined. Dollar and Kraay (2003) were among economists that measured trade using instrumental variables. They chose lagged values of trade as a fraction of GDP as instruments and found a significant effect of trade on growth, but did not focus on the effect of income on imports and exports. Geographic characteristics, for instance, size of an economy and distance between countries were also tried as instruments for measuring trade by some researchers, such as Frankel and Romer (1999), but those results were replicated with various success. Since, it is not clear what would make a reliable instrumental variable for the estimation of the income-trade relationship, an alternative empirical approach, the use of VAR models, has been adopted for this chapter.

Moving to the price elasticity of imports and exports, this relationship received a substantial attention from the economists in the field. A Marshall-Lerner condition is a well-known proposition that suggests if the sum of export (ε_x) and import (ε_m) price elasticities (in absolute value) is greater than unity ($\varepsilon_x + \varepsilon_m > 1$), then depreciation of home currency will improve this country's external balance by increasing demand for its exports and reducing demand for imports, as domestic goods and services become relatively cheaper comparing to the foreign ones (Krugman, 2014).

The dollar to pound exchange rate has been very jumpy ever since 1870s (Figure 2.6). However, the pound real effective exchange rate has been mostly decreasing ever since late 1920s. This suggests that pound has been depreciating against other currencies, comparing to US dollar that has been relatively stable since about 1925.

Fig. 2.6 Sterling Exchange Rates

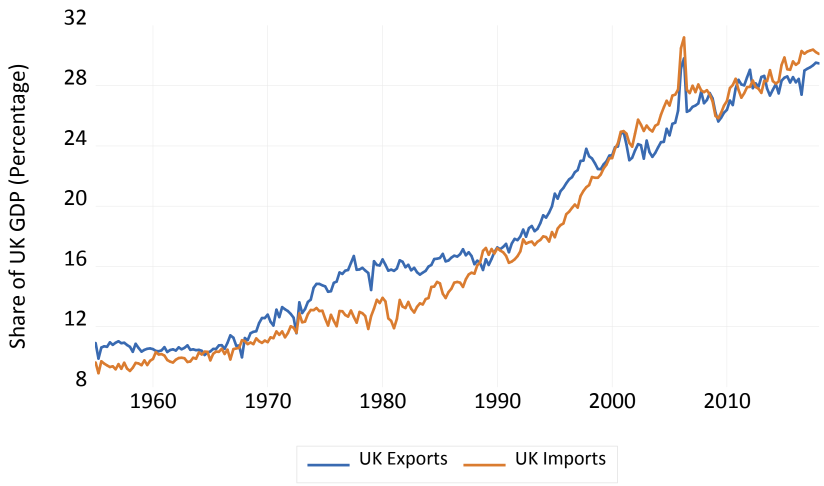


Source: Bank of England (2018).

If the Marshall-Lerner condition held, depreciation would have caused UK exports to become relatively cheaper for the foreign consumers, which should have resulted in trade balance surplus. Nonetheless, the UK has been running trade deficit ever since the 1990s. Trade deficit averaged to 1453.27 million of pound sterling in 1955-2018, reaching an all time low of 6.2 billions of pounds in September of 2016 (Trading Economics, 2018). Explanation for this may lie in the Balassa-Samuelson effect (Balassa, 1964; Samuelson, 1964), which encompasses the idea, that the real exchange rate is determined by the relative productivity growth, and this depreciation of pound relative to the world may not have been the consequence of the differential productivity growth, but rather an attempt to stimulate the exports by decreasing relative exports prices. However, looking at the real UK exports and imports of goods and services (Figure 2.7), expressed as shares of the UK GDP, it appears that, on average, the UK was running balance of payments deficit, even though during that period the real sterling exchange rate continued to depreciate against other countries.

Empirical results on the effect of the exchange rate on trade flows are mixed. In the classic paper by Houthakker and Magee (1969), it was suggested that the price elasticities estimates were not precise enough to confirm or refute the Marshall-Lerner condition. Other papers provided some empirical support for the Marshall-Lerner condition, concluding that the depreciation should in fact reduce trade deficit (Arize, 1986; Caporale et al., 2015b; Hasan and Khan, 1994). For instance, Arize (1986) estimated the relationship between trade and relative prices for nine developing countries for 1960-1982 and found that the

Fig. 2.7 The UK Exports and Imports as Shares of UK GDP



Source: Bank of England (2018)

absolute sum of the long-run export and import demand elasticities is greater than one for majority of countries in his sample, implying that the Marshall-Lerner condition holds. He also emphasised that there was a strong connection between exchange rate policy and trade balance.

The empirical papers on trade elasticities, especially the early ones, are mostly focused either on one type of product (Altinay, 2007; Sousa, 2014) or on certain importers (Burke and Liao, 2015; Soderbery, 2015). Only a few studies attempted to estimate the effect of change in relative price on trade flows for a large sample of countries or for various products. Among them is a paper by Kee et al. (2008), who based their empirical analysis on GDP function approach described by Kohli (1991). Kee et al. (2008) estimated import demand elasticity for 4,900 products for a sample of 117 countries, focusing on the consequences of the trade liberalisation and the effect of tariffs on income. They found that import demand for homogeneous products was more price-elastic than for differentiated goods and estimated that import demand elasticities for the countries in their sample averaged to -1.67. However, their study covered a relatively short period of 1988-2001.

For comparison, Imbs et al. (2010) used a multi-country demand system with constant elasticity of substitution preferences nested in the model and found that import elasticities, constrained to sectoral homogeneity, ranged between 0.5 and 2.7 while export demand elasticities ranged from 0.9 to 2.25 (in absolute values). They also emphasised that unconstrained elasticities, the ones that vary depending on the product specialisation and trade across sectors, tended to be higher for specialised, open, developing countries.

Building on Kee et al. (2008) modelling approach, Ghodsi et al. (2016) estimated import demand elasticity for 167 importing countries using data from the Commodity Trade Statistics Database (COMTRADE) on 5124 products for 1996-2014. They found that South Asia and North America have highest import demand elasticities while Europe, especially Eastern European countries, is associated with quite inelastic import demand. Hence, distribution of import demand elasticities seem to vary depending on geographic location among other factors. As for product category, demand for luxury goods found to be the most inelastic for majority of countries in the sample, while the demand for mineral products, live animals, animal and vegetable fats was the most elastic.

When estimating the effects of income or prices on trade in the long run, one possible technique to use is a cointegration approach. For instance, Chinn and Johnston (1996) built their empirical analysis on the dynamic model proposed by Rogoff (1992). It is based on Cobb-Douglas production functions (supply side) and a utility function of a representative consumer that smooths expected marginal utility over time. Chinn and Johnston (1996) claimed that it was possible to form a cointegrating relationship among the real exchange rate, traded and nontraded productivity levels across sectors as well as government spending, when analysing panel data, but not for an individual exchange rate. They noted that cointegrating relationship may also include terms-of-trade and income per capita.

Afzal and Ahmad (2004) used cointegration analysis to estimate trade elasticities and check the Marshall-Lerner condition. They found that depreciation did not improve the trade balance, suggesting that decrease in home currency value may trigger some other forces, that combat positive effect of depreciation on the trade balance. For instance, expansionary monetary policy that may be adopted following depreciation, in order to prevent rise in unemployment level and fall in the living standard, at the same time may lead to the balance of payments continuing to deteriorate (Afzal and Ahmad, 2004). Although, it worth to note that they focused on the case of Pakistan and, hence, we shall generalise from these findings with caution.

However, there are papers that claimed that depreciation generally has small effect on imports and even less so, on exports, especially in developing countries, where import and export elasticities tend to be lower than in developed, open economies (Edwards, 1986; Naqvi, 1982; Upadhyaya and Upadhyay, 1999). For instance, Naqvi (1982) argued that a certain way to boost exports is through provision of subsidies and increasing production in home country, rather than through change in home-to-foreign exchange rate (Naqvi et al, 1983).

Hence, it is not clear whether there is a significant effect of exchange rate and relative prices on trade flows, and whether one can consider them exogenous. Boyd et al. (2001)

used a cointegrating vector autoregressive distributed lag (VARDL) model to estimate the relationship among output, balance of trade and the real exchange rate. One of the strengths of this approach is that this model allows to start analysis with an assumption that all variables under consideration might be endogenous. Using reduced-form equation for the balance of trade, Boyd et al. (2001) estimated price elasticity of trade without identifying the 'deep' parameters, which otherwise would be required (Boyd et al., 2001; Lee and Chinn, 1998; Rose, 1991). They also used generalised impulse response functions to investigate the response of domestic and foreign output, balance of trade and the real exchange rate to shocks. That allowed them to check whether there was empirical evidence of J-curve (which encompasses the idea that depreciation causes initial worsening of trade balance, followed by the increase in exports, as the volume effect outweighs the value effect, and eventually demand for cheaper domestic products increases). A cointegration approach was used to properly handle non-stationary time-series data, such as global income and international trade flows, that normally exhibit trend. Besides providing evidence of J-curve effects, another important finding of this paper was conclusion that income and real exchange rate could be treated as weakly exogenous when used to model balance of payment. Therefore, they are suitable as explanatory variables when modelling trade-income relationship.

As in Boyd et al. (2001), Hooper et al. (2000) also used conventional import and export demand equations that assume that trade flows are determined by the real national and global incomes and relative prices. They explained that in large open economies, it might be that not only income affects trade, but also the other way around, in which case output and prices can be considered endogenous. Hence, Johansen's (1988) maximum-likelihood cointegration method can be used to estimate long-run elasticities, as it recognises simultaneity among income, trade and prices. As for the effects of relative prices and income on imports and exports in the short run, Hooper et al. (2000) used the ECM version of ARDL model to estimate these elasticities. They estimated short-run and long-run trade elasticities for G-7 economies, namely Canada, France, Germany, Italy, Japan, the UK and US. They found that for the majority of countries in their sample, with the exception of Germany and France, the Marshall-Lerner condition was holding. Moreover, according to their estimates, in the short run, national economic development was transmitted largely through changes in income rather than prices. They also added that short-run fluctuations might have been influenced by factors, such as delivery lags, dock strikes or transitory changes in trade policy. However, further empirical research would be required to estimate the effect of these factors on trade. In this chapter we focus on the effects of income on trade and check whether the Marshall-Lerner condition holds for the UK and US for which longer quarterly data is available.

2.3 Model Specifications and Methodology

2.3.1 Income Elasticity of Trade

There is no consensus in the literature as to what is the right way to measure trade. For instance, some researchers (Baldwin and Taglioni, 2006) argued that trade should be measured as the sum of the logarithms of exports and imports, while the use of the logarithm of the average trade flows can potentially bias the results. Nonetheless, following Schnatz et al. (2006), who analysed both definitions from the empirical perspective and did not find significant difference between them, this chapter assumes that the results are not sensitive to the use of one definition or the other.

Our empirical approach is based on Khan (1974), who first developed the model that is based on an assumption that import and prices adjust to their equilibrium values without delay, and that prices are exogenous. Relative prices and global income determine exports, while imports are determined by the relative prices and national income. In both cases one would expect that increase in relative unit value would decrease exports and imports, while increase in world output or domestic income will boost exports or imports, respectively. Rise in domestic price of foreign currency (nominal exchange rate), however, will have different effect in each case, increasing exports and deflating imports. The disequilibrium model of imports demand for a country $i's$ can be written as follows:

$$lnM_{it} = \alpha_0 + \alpha_1 lnPM_{it} + \alpha_2 lnY_{it} + \mu_{it}, \tag{2.1}$$

where M_{it} is the real imports of a country i, Y_{it} is the real gross national product of a country i, PM_{it} are the relative imports prices and μ_{it} is the error term.

The relative imports prices are measured as follows:

$$lnPM_{it} = ln\left(\frac{PML_{it}}{PD_{it}}\right), \tag{2.2}$$

where PML_{it} is the imports price index denominated in local currency and PD_{it} is the domestic GDP deflator of country i in local currency.

Meanwhile, demand for a country $i's$ exports can be formulated as follows:

$$lnX_{it} = \beta_0 + \beta_1 lnPX_{it} + \beta_2 lnY_t^* + \nu_{it}, \tag{2.3}$$

where X_{it} is real exports of a country i, Y_t^* is the real world income, PX_{it} are the relative exports prices and ν_{it} is the error term.

The relative exports prices are measured as the logarithm of the ratio between the export price index of the ith country and the US gross domestic product (GDP) deflator, both expressed in US dollars:

$$lnPX_{it} = ln\left(\frac{PXL_{it} \times E_{it}}{PD_{USt}}\right), \tag{2.4}$$

where PXL_{it} is an exports price index denominated in local currency, E_{it} is the price of the ith foreign currency in terms of US dollars and PD_{USt} is the foreign GDP deflator in US dollars.

These conventional demand equations for imports, (2.1), and exports, (2.3), are based on several assumptions. The domestic and foreign products are considered to be imperfect substitutes, the price homogeneity holds and trade elasticities with respect to income and relative prices are assumed to be constant over time.

Following this framework and Griliches (1967), this chapter proposes a version of the disequilibrium model of trade that includes lagged variables to capture long-term effects, following the assumption that the adjustment is only partly achieved within a period t. The lagged variables allow for the investigation of the dynamic of the link between trade flows and other macroeconomic variables.

When empirically estimating the models that include macroeconomic time series, such as GDP, prices and exchange rate, the cointegration analysis proved to be a useful approach. Two or more variables are said to be cointegrated, if they share a common trend, meaning they might temporarily diverge from each other, but there is still a relationship between them in the long run (Granger, 1986). A combination of two or more cointegrated non-stationary variables has linear properties and is stationary; hence, it can be estimated by OLS. The long-run relationship between cointegrated variables can be estimated using various tests, such as the Wald test (Constantinescu et al., 2015), the Engle and Granger (1987) residual based test or the maximum likelihood estimation test suggested by Johansen and Juselius (1990). In this chapter we check the significance of the long-run effect of income on trade using the Wald test.

2.3.2 The Marshall-Lerner Condition

To check whether the Marshall-Lerner condition holds we again use exports, (2.3), and imports, (2.1), demand equations, introduced by Khan (1974). In (2.1) and (2.3) the prices are considered to be exogenous. However, as was mentioned above, income and trade are likely to be co-determined, hence, there is a possibility that prices might not in fact be exogeneous either. Therefore, the analysis of the relationship between income, relative prices and trade flows starts with an estimation of VAR models for exports and imports. This allows modelling of macroeconomic data without imposing a restriction on variables to be exogenously determined and does not require a strong economic theory to be based on.

Assuming the distributed lag of order p, in order to analyse the relationship between income, prices and trade flows, $VAR(p+1)$ is estimated for the m $\times$ 1 vector $\boldsymbol{w}_{it}$:

$$\boldsymbol{w}_{it} = \boldsymbol{a} + \sum_{i=1}^{p+1} \boldsymbol{\Gamma} \boldsymbol{w}_{it-1} + \boldsymbol{v}_{it}, \tag{2.5}$$

where $\boldsymbol{w}_{it}$ is either the exports vector $(lnX_{it},\ lnY_t^*,\ lnPX_{it})$ or the imports vector $(lnM_{it},$ $lnY_{it},\ lnPM_{it})$, $\boldsymbol{a}$ is a vector of intercepts, Γ is a matrix of coefficients and $\boldsymbol{v}_{it}$ is a vector of error terms. In the exports vector: X_{it} is the exports of a country i in constant prices; Y_t^* is the real world income; PX_{it} denotes the relative exports prices ($PX_{it} = \left(\frac{PXL_{it} \times E_{it}}{PD_{USt}}\right)$, where PXL_{it} is an exports price index denominated in local currency, E_{it} is the price of the ith foreign currency in terms of US dollars and PD_{USt} is the foreign GDP deflator in US dollars). In the imports vector: M_{it} is the imports of a country i in constant prices; Y_{it} is the real gross national product of a country i; PM_{it} denotes the relative imports prices ($PM_{it} = \left(\frac{PML_{it}}{PD_{it}}\right)$, where PML_{it} is the imports price index denominated in local currency and PD_{it} is the domestic GDP deflator in local currency).

In principle one could estimate a large system in $(lnX_{it}, lnM_{it}, lnY_{it}, lnY_t^*, lnPX_{it}, lnPM_{it})$. However, VAR models are very densely parameterised in comparison to the dimension of the time-series data. Thus, to avoid estimation problem, VAR models are estimated with the lowest possible number of parameters. In particular, we will work with two smaller systems, one in $(lnX_{it},\ lnY_t^*,\ lnPX_{it})$ and one in $(lnM_{it},\ lnY_{it},\ lnPM_{it})$.

This model, (2.5), can be written in the VECM form as:

$$\Delta \boldsymbol{w}_{it} = \boldsymbol{\mu} + \boldsymbol{\alpha\beta} \boldsymbol{w}_{it-1} + \sum_{i=1}^{p} \boldsymbol{\Gamma} \Delta \boldsymbol{w}_{it-1} + \boldsymbol{v}_{it}, \tag{2.6}$$

where $\boldsymbol{w}_{it}$ is either the export vector $(lnX_{it},\ lnY_t^*,\ lnPX_{it})$ or the import vector $(lnM_{it},$ $lnY_{it},\ lnPM_{it})$, $\boldsymbol{\mu}$ includes unrestricted deterministic elements, $\boldsymbol{\beta}$ are cointegrating vectors ($\beta = (\beta_{1i}, \beta_{2i}, \beta_{3i})$ characterises the ith long-run relation among $(lnX_{it},\ lnY_t^*,\ lnPX_{it})$ for exports and $(lnM_{it},\ lnY_{it},\ lnPM_{it})$ for imports, correspondingly), the elements of $\boldsymbol{\alpha}$ measure the speed of adjustment coefficients, often referred to as feedback or loading coefficients, Γ is a matrix of coefficients and $\boldsymbol{v}_{it}$ is a vector of error terms.

Hooper et al. (2000) calculated income-trade elasticities following a theoretical framework and estimation techniques similar to the ones presented in this chapter. Hence, the results of the analysis presented here are compared to the estimates from Hooper et al. (2000).

The model (2.5) is also sensitive to the number of lags included in the system. To choose the appropriate lag length for the VAR model we use Akaike (AIC) and Bayesian (BIC) Information Criteria. In addition, as Toda and Yamamoto (1995) emphasised, if the variables in the VAR are integrated of order one, it might be necessary to include more lags, beyond the number suggested by the chosen selection criteria, as the usual tests may no longer produce reliable estimates. Thus, the optimal number of lags is chosen not only depending on model selection criteria, but also on whether the coefficients have correct signs, meaning signs that are consistent with economic theory.

This is followed by the Granger Causality test to check whether knowing current value of one of a variable helps to predict any other variables in the model. Furthermore, whereas Granger-causality only reflects predictability from previous periods, the generalised impulse response functions (GIRFs) reflect the within quarter correlations between the residuals of the equations in the VAR models. In addition, correlation coefficients are obtained from the correlation matrix.

Then, the order of integration of the variables is determined using Augmented Dickey-Fuller (ADF) unit root test. If the variables are integrated of the same order, we proceed with the Johansen's cointegration test (1988) in order to check for presence of the long-run relationship among the variables. Cointegration approach is appropriate for this analysis, since judging by the graphs (Figures 2.8 and 2.9 for the UK; Figures 2.10 and 2.11 for the US), there exists at least one equilibrium relationship that cancels out a common stochastic trend that variables under consideration have in common. If the presence of cointegrating vector(s) is established, then following Patterson (2000) it is common to normalise one coefficient in each cointegrating vector by setting it equal to one.

For example, considering the VECM, (2.6), for exports. If there is one cointegrating vector, which we normalise on exports as the dependent variable, and if the feedback coefficients on the world income and export prices are insignificant, we can treat them as exogenous and rewrite the export equation in the reduced form of ARDL model, the ECM, as in (2.7) below:

$$\Delta lnX_{it} = \alpha_x + \beta_x \Delta lnY_t^* + \eta_x \Delta lnPX_{it} + \rho_x lnX_{it-1} + \mu_x lnY_{t-1}^* + \nu_x lnPX_{it-1} + \varepsilon_{it}^x, \qquad (2.7)$$

where X_{it} is the quantity of exports of a country i; Y_t^* is the real world income; PX_{it} are the relative exports prices ($PX_{it} = \left(\frac{PXL_{it} \times E_{it}}{PD_{USt}}\right)$, where PXL_{it} is an exports price index denominated in local currency, E_{it} is the price of the ith foreign currency in terms of US dollars and PD_{USt} is the foreign GDP deflator in US dollars).

As for coefficients, β_x and η_x are short-run income and price elasticities of exports, respectively, and ε_{it}^x is the error term.

The significance of the long-run income and price effects on trade flows are measured using the Wald test.

Similarly with imports, if there is one cointegrating relation between the variables under consideration, we proceed with normalising on imports. If the loading coefficients on the national income and imports relative prices are insignificant, we conclude that these variables can be treated as exogenous and proceed by estimating the following ECM:

$$\Delta lnM_{it} = \alpha_m + \beta_m \Delta lnY_{it} + \eta_m \Delta lnPM_{it} + \rho_m lnM_{it-1} + \mu_m lnY_{it-1} + \nu_m lnPM_{it-1} + \varepsilon_{it}^m, \quad (2.8)$$

where M_{it} is the quantity of imports of a country i; Y_{it} is the real gross national product of a country i; PM_{it} are the relative imports prices ($PM_{it} = \left(\frac{PML_{it}}{PD_{it}}\right)$, where PML_{it} is the imports price index denominated in local currency and PD_{it} is the domestic GDP deflator in local currency).

As for the coefficients, β_m and η_m are short-run income and price elasticities of imports, correspondingly, and ε_{it}^m is the error term.

Finally, to check whether income elasticity of trade has changed following the financial crisis 2007-2008, the model bellow is estimated for exports:

$$\begin{aligned}\Delta lnX_{it} = \alpha_{0x} + \beta_{0x}\Delta lnY_t^* + \gamma_{0x}\Delta lnPX_{it} + \alpha_{1x}X_{it-1} + \beta_{1x}lnY_{t-1}^* + \\ + \gamma_{1x}lnPX_{it-1} + \delta_{0x}D\Delta lnY_t^* + \delta_{1x}DlnY_{t-1}^* + \varepsilon_{it}^x\end{aligned} \quad (2.9)$$

and for imports:

$$\begin{aligned}\Delta lnM_{it} = \alpha_{0m} + \beta_{0m}\Delta lnY_{it} + \gamma_{0m}\Delta lnPM_{it} + \alpha_{1m}lnM_{it-1} + \beta_{1m}lnY_{it-1} + \\ + \gamma_{1m}lnPM_{it-1} + \delta_{0m}D\Delta lnY_{it} + \delta_{1m}DlnY_{it-1} + \varepsilon_{it}^m,\end{aligned} \quad (2.10)$$

where, D takes a value of one for 2009Q1 onwards and equals zero for the period before 2009Q1, β_{0x} and β_{0m} measure short-run income elasticities of exports and imports from 2009Q1 onwards, respectively, and β_{1x} and β_{1m} measure long-run effect of income on exports and imports during that time, correspondingly, and ε_{it}^x is the error term.

To summarise, starting from the assumption, that variables of interest might be endogenous, the unrestricted VAR model is estimated. Lag length is chosen using AIC and BIC. This is followed by checking for the Granger causality among income, prices and trade flows.

Then the model is checked on the presence of any cointegrating vectors, and if such exist(s), the reduced form of VAR model, a VECM, is specified. The VECM coefficients are then used to determine whether income and prices are exogenous in exports and imports equations. If they are, the analysis then proceeds with an ARDL model to calculate short and long-run coefficients. The analysis is concluded by running a parameter stability check.

2.4 Data and Variables

2.4.1 Income Elasticity of Trade: Sample of 15 Countries

The cross-sectional studies and time-series papers on trade use different sets of variables. The studies on a gravity model of trade are characterised by using variables that do not vary over time. Those are normally dummy variables, which are included in the models to account for the distance effects, for the differences among countries due to their economic development (developed or transition countries) as well as for the differences in industry and product-level data of each country in a sample (Disdier and Head, 2008; Hanson et al., 2015).

Other variables, such as gross domestic product, relative prices and the exchange rate are often used in various time-series and cross-sectional papers on trade, including the papers on gravity model (Binh et al., 2011; Bussière et al., 2009; Schnatz et al., 2006), on trade growth (Baldwin, 2008; Krugman, 1979; Yi, 2003) and on major empirical puzzles in international economics (Dumas, 1992; Obstfeld and Rogoff, 2002).

However, due to limited availability of trade data, most papers on international trade focus on imports and/or exports of manufactured goods and services and normally use short time-series data for a large number of countries. However, an analysis of a changing trade-income relationship calls for long time-series data and there are a few sources these data can be obtained from, such as World Trade Organisation (WTO) and International Monetary Fond (IMF) databases.

For a preliminary analysis of the relationship between trade flows and income we use trade data from the two databases mentioned above, WTO and IMF. Besides the data for imports and exports, we also use time series for GDP, price indices and exchange rate. The estimation period for preliminary research using WTO data is 1970-2014. When we move to the comparison of the WTO and IMF datasets, the estimation period is shortened to 1980-2013, due to limited availability of IMF trade data for individual countries.

The real exchange rate is measured as a product of the US dollar exchange rate, E_{it}, and the ratio of domestic and foreign consumer prices (CPI_{it} and CPI_{USt}, respectively). The non time-varying dummy variables mentioned above are captured by the intercept. The required data are collected for the world as a whole and for a sample of 15 countries for a more in-depth analysis of the trade-income relationship. The sample includes Australia, Austria, Canada, Chile, Denmark, Finland, France, Germany, Greece, India, Indonesia, Israel, Italy, Japan and the United Kingdom. The choice of countries was determined by the data availability. The sample includes the countries for which the longest time series with fewest missing observations were available. Overall, this sample provides a good spread of different types of countries, allowing for analysis of possible country-specific patterns.

Total merchandise exports and imports from a country i to the rest of the world come from the WTO and are expressed in US dollars and current prices. Since these data reflect not only changes in trade volumes, but also the effect of inflation, in the general-to-specific (section 2.5) and specific-to-general (section 2.6) modelling, which focus on the exports demand equation and income elasticity of exports in particular, real exports are used. Real exports are calculated by deflating total merchandise exports in current US dollars, ($NX_{it} \times E_{it}$), by the United States consumer price index, CPI_{USt}, in order to enable the quantity comparison between different time periods.

One of the reasons the real exchange rate was included in the general model (section 2.5) is to control for the effects of movements in a country's currency relative to the US dollar (Graham et al., 2004; Micco et al., 2003). Moreover, some researchers claim that the exchange rate and exports relative price are powerful determinants of a country's exports (Warner and Kreinin, 1983). Therefore, these variables are included in the model in order to check their significance in relation to trade flows.

However, it is not clear whether the right approach to measuring trade would be to express exports in local currency or in US dollars. Taking nominal exports in local currency, NX_{it}, might not be the best approach as nominal values do not account for the effect of inflation. There is an option to divide exports in local currency by a country i's consumer price index, CPI_{it}, to get real exports in local currency, $\left(\frac{NX_{it}}{CPI_{it}}\right)$. However, using real exports in local currency, $\left(\frac{NX_{it}}{CPI_{it}}\right)$, does not allow for a proper comparison of export flows among the countries that have different currencies. Another option is to express exports of the country i in real US terms, and there are two possible ways to do that. Local currency exports, NX_{it}, multiplied by the dollar exchange rate, E_{it}, and divided by the US price index, CPI_{USt}, gives real exports of the country i expressed in the US consumer goods, $\left(\frac{NX_{it} \times E_{it}}{CPI_{USt}}\right)$. If instead of the US price index, CPI_{USt}, we divide nominal exports in US dollars, $(NX_{it} \times E_{it})$, by the US GDP deflator, PD_{USt}, that would give us real exports of the country i expressed in terms of the components of the US GDP, $\left(\frac{NX_{it} \times E_{it}}{PD_{USt}}\right)$.

The appropriate dependent variable could be a domestic real measure of exports, $\left(D = \frac{NX_{it}}{CPI_{it}}\right)$, or one of the possible world (dollar) measures of trade, for instance $\left(W = \frac{NX_{it}*E_{it}}{CPI_{USt}}\right)$. It is not clear which measure is the ideal choice, but domestic measure of exports can be expressed as a product of world measure of exports and real exchange rate:

$$D = \frac{NX_{it}}{CPI_{it}} = \frac{\frac{NX_{it}*E_{it}}{CPI_{USt}} * CPI_{USt}}{E_{it} * CPI_{it}} = \frac{W * CPI_{USt}}{E_{it} * CPI_{it}}. \tag{2.11}$$

Hence,

$$D = W * \left(\frac{CPI_{USt}}{E_{it} * CPI_{it}}\right). \tag{2.12}$$

Therefore, domestic and world measures of exports are related through the real exchange rate. Thus, the general model (section 2.5) used in this chapter nests both (domestic and world measures of exports) to account for both possible cases. One possibility is that the correct dependent variable is the domestic measure. The other case is that the correct dependent variable is the world measure and the real exchange rate is allowed to vary independently.

For the estimation of the income elasticity of trade for 15 countries the data for the nominal dollar exchange rate, E_{it}, and consumer price indices, CPI_{it} and CPI_{USt}, are obtained from the Federal Reserve Economic Data database (FED, 2017). The real domestic and world GDP data (Y_{it} and Y_t^*, respectively) come from the World Bank (World Bank, 2017), and is expressed in US constant prices.

One of the aims of this chapter is to check whether income-trade analysis performed using data from different data sources produces similar results. Hence, trade data from two different databases, the World Trade Organisation and International Monetary Fund, are collected, and imports (not exports) data are used for this part of the analysis to enable comparison to Constantinescu et al. (2015), who performed the analysis similar to the one presented in this chapter, and they used imports. As was mentioned, from the WTO database trade data come as total merchandised exports and imports, expressed in US dollars at current prices. Hence, real imports are calculated by deflating total merchandise imports in current US dollars ($NM_{it} \times E_{it}$) by the United States consumer price index (CPI_{USt}). From the IMF database we use the volume of total imports of goods and services, expressed in constant US dollars. Hence, IMF data come in constant US dollars and the WTO trade data are converted from current to constant US dollars. We then compare estimates obtained using data from these two different databases, the WTO and IMF.

The initial analysis of the income-trade relationship using annual data for a sample of 15 countries showed that use of longer time series can be a strong advantage for the analysis of the relationships considered in this chapter. Longer quarterly data are available for the UK and US, but not for many other countries. Hence, due to data availability limitations, we use different data for the analysis of the Marshall-Lerner condition (quarterly data for the UK and US) than for the estimation of the income elasticity of trade (annual data for 15 countries). The data we use for the analysis of the Marshall-Lerner condition for the UK and US are discussed below.

2.4.2 The Marshall-Lerner Condition: The UK and US Cases

For the analysis of the trade-income elasticities the longer the data the better. After preliminary research the data availability became an apparent issue, as for most countries the data for imports and exports are annual and only available for a few decades. Nonetheless, there are exceptions, such as the UK and US, for which there is quarterly data that go back to 1950s (1955Q1 for the UK and 1950Q1 for the US) and is available to 2018Q1 (on the moment of writing this chapter). Hence, in order to further analyse the relationship between trade, imports and exports prices as well as income we now focus on the cases of the UK and US.

Gross domestic product is often used to measure income. However, quarterly data for the world GDP are not available for sufficiently long period. Therefore, when checking the Marshall-Lerner condition for the UK and US, we follow Hooper et al. (2000) and use combined GDP of the G7 countries (Canada, France, Germany, Italy, Japan, the UK and US), which accounts for roughly 30% of world GDP, as a proxy for the world income. This also makes our estimates more comparable to the ones in Hooper et al. (2000). The GDP data are collected for the G7 countries in current values, local currency. The US data come from the Bureau of Economic Analysis (BEA), the French one are from National Institute for Statistics and Economic Studies, the UK and Canadian GDP data are from the Office for National Statistics (ONS) and for Germany, Japan and Italy are from Quarterly National Accounts. The exchange rates of local currency per US dollar data come from OECD Main Economic Indicators publications, except for Canada these data come from IMF (and are in US dollars per Canadian dollar exchange rate). The GDP data are converted to US dollars using nominal exchange rates and then divided by the US GDP deflator. Finally, the G7 members' incomes are added up. Since some data, such as quarterly GDP for most countries in G7 group, are not available prior 1960Q1, the final G7 GDP series covers 1960Q1-2017Q1. The resulting time series is a G7 real GDP in US dollars, which is used as a proxy for world income when estimating the UK and US exports demand equations.

The UK: Data (Quarterly, 1955Q1-2018Q1)

Starting from the UK case, the Table 2.1 presents a summary of the data that were used to analyse UK trade.

Table 2.1 The UK Quarterly Data

Variable	Notation	Units	Year [1]	Timespan	Source
Exports	X_{UK}	MM £; constant prices	2016	1955:1-2018:1	ONS
UK Exports Defl.	PXL_{UK}	Constant price value index	2012	1955:1-2018:1	ONS
Exch. rate (£/$)	E_{UK}	Price index	-	1950:1-2018:1	ONS
US GDP Deflator	PD_{US}	Price index	2009	1950:1-2018:1	BEA
Imports	M_{UK}	MM £; constant prices	2016	1955:1-2018:1	ONS
UK Real GDP	Y_{UK}	MM £; constant prices	2016	1955:1-2018:1	ONS
UK Imports Defl.	PML_{UK}	Constant price value index	2012	1955:1-2018:1	ONS
UK GDP Defl.	PD_{UK}	Price index	2016	1955:1-2018:1	ONS

[1] Reference Year.

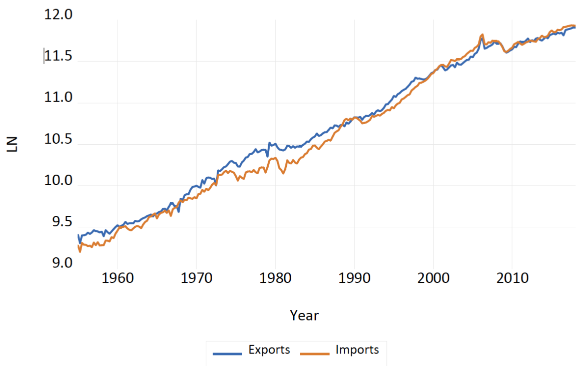
Most of the quarterly data come from the ONS and cover 1955Q1-2018Q2. The exceptions are the exchange rate and the US data, that cover slightly longer period with the US data being from the Bureau of Economic Analysis. Exports, imports and GDP are in millions of pounds, constant prices. Other variables in the table come in index form.

During 1955-2018 the UK exports and imports moved closely together (Figure 2.8), except for a period in the 1970-1990, when constant price export flows were generally higher than constant price imports. Both time series have an increasing trend, however, there was a drop in trade volumes around 1967-1968, time of the UK Devaluation of Sterling, and in the beginning of 1980s, years of the UK recession as well as in late 2000s, when the global financial crisis 2007-2008 happened.

Second part of the Figure 2.8 presents the UK real GDP expressed in percentage change form. This series is very jumpy, with major growth rate drops, around 2%, in the second quarter of 1958, the Middle East Crisis, then in 1973, the oil price shock, in early 1980s during the recession and then during the financial crisis 2007-2008. The UK is a large open economy that relies heavily on international trade, therefore, all those global economic disturbances triggered sharp falls in the British GDP growth rate.

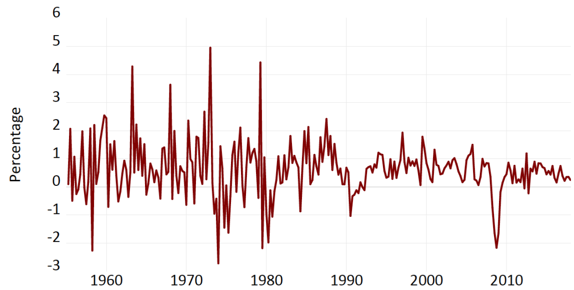
Moving to the prices (Figure 2.9 Part 1), UK relative exports prices are more volatile than import ones, with significant jumps around 1970, 1980, 2002, 2008 and the most recent drop in 2014-2015. For instance, in the late 1970s, following the end of the post-war economic boom, raw material prices were rising and the UK real effective exchange rate as well as the pound to dollar real exchange rate, were fluctuating dramatically. Imported goods and

Fig. 2.8 The UK Trade and Income (Quarterly Data, 1955-2018).



Source: ONS (2017).

Part 1. Exports and Imports.

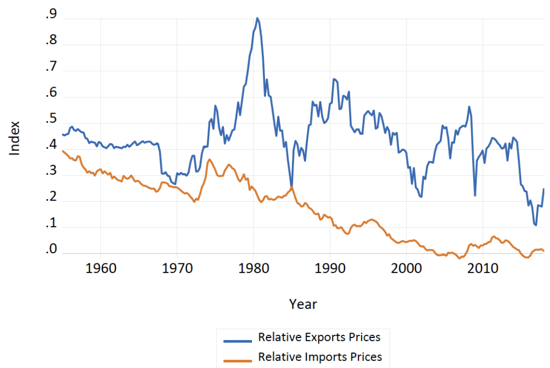


Source: ONS (2017).

Part 2. The UK Real GDP.

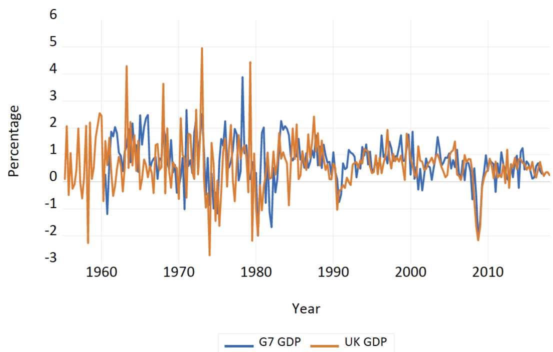
services were valued higher than the domestic ones, which led to the balance of payments problems in the UK. This period was also characterised by the political unrest within the country, including issues in the Northern Ireland and strikes caused by the increasing power of the trade unions. Overall, it seems that exports prices adjust to changes in economic conditions with a slight delay comparing to imports prices, and changes are bigger in magnitude. The UK relative imports prices have been mostly declining since mid-1970s, with a few minor increases during the timespan covered by the estimation period.

Fig. 2.9 The UK Trade Relative Prices and Income (Quarterly Data, 1955-2018).



Source: ONS (2017) and BEA (2018).

Part 1. Relative Exports and Imports Prices.



ONS (2017) and BEA (2018).

Part 2. The G7 and UK Real GDP.

In turn, the G7 and UK incomes (Figure 2.9 Part 2) series have a few especially pronounced spikes, around 1963, the European Economic Community (EEC) Crisis, then around 1971, International Monetary Crisis, followed by early 1980s, time of recession in the UK, then another spike around 1992, remembered by the Black Wednesday (16 September 1992), and a more recent decline was around the crisis 2007-2008.

The UK trade and price series appear to be nonstationary, with increasing trend. The unit root analysis confirms that all variables under consideration are I(1), and the long-run relationship among them can be estimated using not only ARDL model, but a cointegration approach, in order to compare estimates produced by the two techniques.

In addition, looking at the descriptive statistics summary (Table 2.2), the UK imports and exports series have similar coefficients of central tendency measures (around 10.60) as well as maximum (about 11.92) and minimum (around 9.25) values. Both series have a relatively low standard deviation of 0.82 with few outliers in the data. However, the UK exports seem to have more values below the average, while the exact opposite is true for the UK imports.

Table 2.2 Descriptive Statistics: The UK Quarterly Data (Summary; Logarithms)

	Mean	**Med.**[1]	**Max.**[2]	**Min.**[3]	**Std. Dev.**[4]	**Skew.**[5]	**Kurt.**[6]	**Jarque-Bera**	**Prob.**[7]	**Obs.**[8]
X_{UK}	10.69	10.65	11.91	9.31	0.80	-0.08	1.73	17.21	0.00	253
M_{UK}	10.63	10.54	11.93	9.20	0.84	0.03	1.69	18.24	0.00	253
Y_{UK}	12.45	12.44	13.13	11.62	0.45	-0.16	1.83	15.46	0.00	253
Y^*	15.92	15.97	16.65	14.95	0.50	-0.23	1.84	14.88	0.00	229
PX_{UK}	0.44	0.43	0.90	0.11	0.12	0.67	5.11	65.75	0.00	253
PM_{UK}	0.17	0.19	0.39	-0.02	0.12	0.00	1.59	20.91	0.00	253

All variables are in logarithmic form;
[1] Median; [2] Maximum; [3] Minimum; [4] Standard Deviation; [5] Skewness; [6] Kurtosis; [7] Probability; [8] Observations.

The G7 and UK GDPs data have a low standard deviation of around 0.50, with a slight negative skew and thin tails, as suggested by the skewness (around -0.20) and kurtosis (about 1.84) coefficients. Relative imports prices exhibit a clear downward trend. Both price indices have a very low standard deviation of 0.12, lowest when compared to other variables under consideration. The exports prices data have more extreme values, with fatter tails than normally distributed data have, while the opposite is true for the imports prices.

When the variables are graphed in the level form, none of them appear to be normally distributed. In the analysis the logarithms of the variables are used to allow for a more straightforward interpretation of the coefficients (as elasticities).

The US Data (Quarterly, 1950Q1-2018Q1)

To estimate income and price elasticities of the US exports and imports, the quarterly data were collected from the BEA. The US quarterly data are summarised in the Table 2.3. The data cover five-year longer period, 1950Q1 to 2018Q1, than the UK one (1955Q1 to 2018Q1). Imports and exports of goods and services as well as the US income and G7 GDP are all expressed in millions of constant US dollars. The US imports, exports and GDP deflators are in index form. All variables are transformed in logarithmic form, as in the UK case.

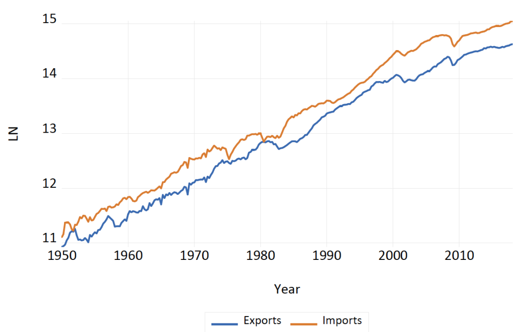
Table 2.3 The US Quarterly Data (1950-2018)

Variable	Notation	Units	Year[1]	Timespan	Source
Exports	X_{US}	MM \$; constant prices	2009	1950:1-2018:1	BEA
US Exports Defl.	PXL_{US}	Constant price value index	2009	1950:1-2018:1	BEA
Imports	M_{US}	MM \$; constant prices	2009	1950:1-2018:1	BEA
US Real GDP	Y_{US}	MM \$; constant prices	2009	1950:1-2018:1	BEA
US Imports Defl.	PML_{US}	Constant price value index	2009	1950:1-2018:1	BEA
US GDP Defl.	PD_{US}	Price index	2009	1950:1-2018:1	BEA

[1] Reference Year.

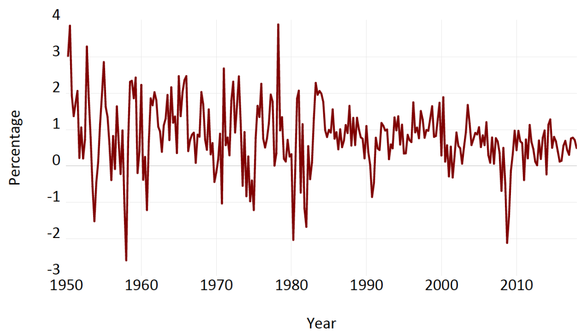
When the logarithms of the US exports and imports plotted on the same graph (Figure 2.10), the trade flows are comparable to those of the UK, exhibiting similar increasing trade with a few minor decreases around early 1980s, time of the major recession in the US and UK, and around the financial crisis 2007-2008, which originated in the US. The US GDP is plotted as percentage change in the Figure 2.10 Part 2 and has similar jumps around 1980s and sharp fall around 2007-2008, as the UK GDP series in the Figure 2.10 Part 2. However, the maximum the US GDP ever reached over this period is lower (less than 4%) than the UK GDP (5% increase around year 1973).

Fig. 2.10 The US Trade and Income (Quarterly Data, 1950-2018)



Source: BEA (2018).

Part 1. The US Exports and Imports.

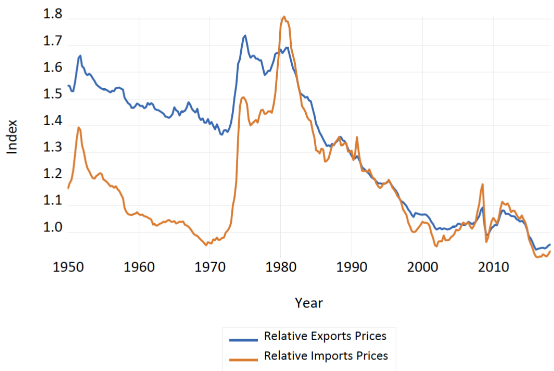


Source: BEA (2018).

Part 2. The US Real GDP.

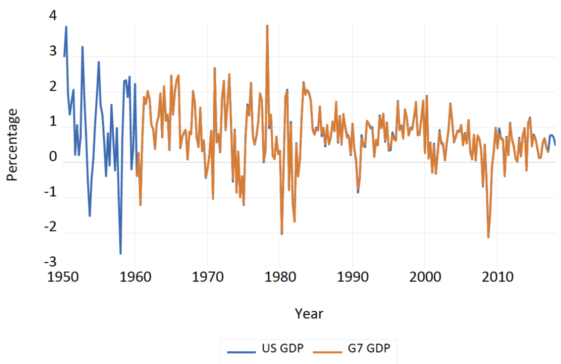
The US relative exports and imports prices (Figure 2.11 Part 1) are quite different prior early 1970s, comparing to those of the UK (Figure 2.9 Part 1). Relative imports prices increased sharply in 1972-1975, almost reaching the level of the relative exports prices, and in the subsequent decade relative imports prices were above the exports ones, but only for a few years. Starting from early 1980s the US relative imports and exports prices changed similarly, while the UK exports prices remained much higher then the British imports relative prices. In addition, the US GDP is almost identical to the G7 income, which reflects the fact that the US is the largest economy in the world and, hence, its income represents the largest share of the G7 series, comparing to other countries in the group.

Fig. 2.11 The US Trade and Income (Quarterly Data, 1950-2018)



Source: BEA (2018).

Part 1. The Relative Exports and Imports Prices.



ONS (2017) and BEA (2018).

Part 2. The G7 and US Real GDP.

Furthermore, the descriptive analysis (Table 2.4) shows that the US imports were on average slightly higher (13.25) than exports (12.92). However, other descriptive statistics coefficients for these series are similar. The US exports and imports data are more spread out (standard deviation is about 1.15), comparing to those of the UK (about 0.82). Both US trade series are slightly negatively skewed (skewness coefficients are within -0.09 to -0.04 range), and their distributions have slimmer tails than normal (kurtosis coefficients are about 1.70).

Table 2.4 Descriptive Statistics: The US Quarterly Data (Summary; Logarithms)

	Mean	Med.[1]	Max.[2]	Min.[3]	Std. Dev.[4]	Skew.[5]	Kurt.[6]	Jarque-Bera	Prob.[7]	Obs.[8]
X_{US}	12.92	12.84	14.63	10.93	1.14	-0.09	1.69	19.78	0.00	273
M_{US}	13.25	13.22	15.05	11.11	1.17	-0.04	1.72	18.76	0.00	273
Y_{US}	15.76	15.78	16.67	14.55	0.62	-0.22	1.80	18.59	0.00	273
Y^*	15.92	15.97	16.65	14.95	0.50	-0.23	1.84	14.88	0.00	229
PX_{US}	1.33	1.39	1.74	0.93	0.23	-0.15	1.62	22.61	0.00	273
PM_{US}	1.17	1.09	1.81	0.91	0.20	1.12	3.83	64.90	0.00	273

All variables are in logarithmic form;
[1] Median; [2] Maximum; [3] Minimum; [4] Standard Deviation; [5] Skewness; [6] Kurtosis; [7] Probability; [8] Observations.

As was clear from the Figure 2.11 Part 2 for the US and G7 incomes, two variables are very similar. The two series have mean values of about 15.80, with relatively low standard deviation of 0.55, slight negative skew and peaked distributions. Moving to the remaining two series, the US relative exports prices are on average slightly higher than imports ones (1.33, comparing to 1.17) with a similar standard deviation of about 0.20. However, while exports price series is slightly negatively skewed, imports price series has a positive skewness coefficient of 1.12 and a much higher kurtosis of 3.83 (comparing to 1.62 for exports price series). In addition, as in case with the UK data, the Jarque-Bera coefficients suggest that none of the US variables have normal distribution.

2.5 General-to-Specific Modelling (Exports; 15 Countries)

Since the determinants of trade growth are not well studied, the chapter first uses general-to-specific modelling approach, which is well-known for excellent model selection abilities (Campos et al., 2005). Moreover, diagnostic tests using a general model are more likely to be valid than in a potentially unspecified restricted model.

The ideal outcome of using general-to-specific method would be a simplified (restricted) model that adequately characterises the empirical evidence within a theoretical framework. To move from the general to specific model, the paper follows basic steps of the Hoover and Perez's algorithm (Hoover and Perez, 1999), starting with an assumption that the general statistical model is appropriate, since it allows a certain flexibility in accessing what drives the trade growth. Then, a variable that satisfies the simplification selection criteria (its coefficient is statistically insignificant) is eliminated and the simplified model is checked to remain congruent. This procedure continues until only variables that have a statistically significant effect on the dependent variable remain.

In addition, most economic time-series, such as logarithm of world income, are trended. To account for that, the trend variable is included in the general model and checked for significance. Building on the demand equation for exports, (2.3), and accounting for the fact that different (domestic and world) measures of exports might be suitable, we estimate the following general ARDL model for 1970-2014:

$$ARDL(1,1): lnX_{it} = \alpha_0 + \beta_0 lnN_{it} + \beta_1 lnN_{it-1} + \alpha_1 lnX_{it-1} + \delta t + u_{it}, \tag{2.13}$$

where $N_{it} = Y_{it}, Y_t^*, E_{it}, CPI_{it}, CPI_{USt}$.

The dependent variable, X_{it}, is real total merchandise exports (world measure) from country i to the rest of the world, meaning total merchandise exports in the US dollar current prices, deflated by the US CPI; α_0 is a constant; Y_{it} and Y_t^* are GDP of country i and the world at market prices (2010 = 1.0), respectively, CPI_{it} and CPI_{USt} are CPI for a country i and the US CPI, correspondingly (2010 = 1.0), E_{it} is a US dollar to national currency spot exchange rate for a country i, δt is a time trend and u_{it} is an error term. The dynamic behaviour of the demand for exports is captured by the time trend (δt) and lagged variables of the model (X_{it-1} and N_{it-1}).

In (2.3) the relative exports prices are measured as $PX_{it} = \left(\frac{PXL_{it} \times E_{it}}{PD_{USt}}\right)$, where PXL_{it} is an exports price index denominated in local currency, E_{it} is the price of the ith foreign currency in terms of US dollars and PD_{USt} is the foreign GDP deflator in US dollars. In contrast, in (2.13) we include variables $E_{it}, CPI_{it}, CPI_{USt}$ on the right-hand side instead and check which ones are actually significant in hope that it would help to identify most suitable functional form of the demand for exports equation.

The ARDL(1,1) model, (2.13), can also be estimated in the error-correction form (ECM),

$$ECM : \Delta lnX_{it} = a_0 + b_0 \Delta lnN_{it} + b_1 lnN_{it-1} + a_1 lnX_{it-1} + \delta t + u_{it}, \qquad (2.14)$$

where Δ represents the change in a variable and all other variables are the same as in (2.13).

First, the unrestricted ARDL model for each country is used with either lagged variables or first differences (Δ), depending on which parameterised version of a general ARDL model fits the data for a country i better.

Starting from the unrestricted model, the variables that have no statistically significant effect on total merchandise exports are dropped one by one (starting from the least significant one) until only statistically significant variables are left in the model (the results are presented in an appendix A.2 Tables A.1 and A.2). However, the regression results for the unrestricted model are mostly inconsistent with the economic theory. For instance, the results suggest that the coefficient of the world GDP for the majority of countries in the sample is insignificant, meaning world income has no significant effect on the exports of most of the countries under consideration. In addition, some coefficients seem to have "wrong" signs, such as, for Australia, Austria and a few other countries for which exchange rate variable is statistically significant, it seems to have a positive relationship with the dependent variable. This is opposite to what one would expect as it suggests that as the US dollar to local currency exchange rate increases, the exports from country i to the rest of the world increase as well. However, the international economics theory (Krugman, 2014) suggests that if the local currency appreciates against the foreign currency, then exports from this country become relatively more expensive and, as a result, demand for them decreases.

To summarise, there are three main problems one runs into when applying general-to-specific modelling on multiple countries with short time series. Firstly, it would be ideal to have a common model for all countries under consideration, but using this approach we did not get it. Secondly, there is a question of how general a model should be to begin with? A model might be overparameterised and poorly estimated. Thirdly, a chosen (restricted) model is difficult to interpret. Since the general-to-specific modelling does not produce sensible results, the alternative approach is to start with a minimum number of explanatory variables and expand from there. Therefore, next the chapter outlines specific-to-general modelling approach to analyse income-trade relationship.

2.6 Specific-to-General Modelling (Exports; 15 Countries)

This procedure begins with a simple model which estimates the relationship between change in real exports of a country *i* (to the rest of the world) and the global GDP. The price variables (domestic and US consumer price indices) have been dropped as well as domestic income, leaving only world income as an explanatory variable. The trend variable is not included, as it is not significant in the general model for the vast majority of countries in the sample. Thus, the general ECM, (2.14), is used, but with only one independent variable:

$$N_t = Y_t^*,$$

where Y_t^* is a world GDP at market prices (2010 = 1.0).

These types of demand equations for exports and imports are fairly intuitive and have been widely used in the empirical trade literature (Constantinescu et al., 2015; Freund, 2009; Irwin, 2002). Hence, adopting (2.14), the following model is estimated:

$$\Delta lnX_{it} = \alpha_0 + \beta_0 \Delta lnY_t^* + \beta_1 lnY_{t-1}^* + \alpha_1 lnX_{it-1} + u_{it}, \tag{2.15}$$

where a dependent variable ΔX_{it} is a change in real total merchandise exports (world measure) from country *i* to the rest of the world, meaning total merchandise exports in US dollar current prices, deflated by the US GDP deflator; α_0 is a constant; ΔY_t^* is a change in world GDP at market prices (2010 = 1.0); X_{it-1} and Y_{t-1}^* are the lagged variables and u_{it} is an error term.

We estimate this model using OLS and running diagnostic tests to ensure that results are consistent with a theoretical framework. The regression analysis using WTO data for the same sample of 15 countries for 1970-2014 (as in general-to-specific modelling) produces results that are summarised in the Table 2.5.

The output is consistent with the economic theory, for instance the signs of the coefficients are the ones we would expect. The effect of change in world GDP, the short-run income elasticity of trade, is significant for all countries, with the exception of Indonesia. A lagged

Table 2.5 Equation (2.15): Restricted Model (Dependent Variable: lnX_{it}; 1970-2014)

	Australia	**Austria**	**Canada**	**Chile**	**Denmark**	**Finland**	**France**
α_0	-2.55	-6.74*	-4.44***	-16.17***	-2.83	-2.33	-0.81
	(1.71)	(3.61)	(1.49)	(4.97)	(1.93)	(2.26)	(1.34)
β_0	5.33***	3.93***	4.40***	5.58***	2.81***	5.01***	3.48***
	(1.01)	(1.03)	(0.66)	(1.91)	(0.93)	(1.16)	(0.91)
α_1	-0.03	-0.22*	-0.32***	-0.29***	-0.21**	-0.15	-0.17*
	(0.06)	(0.11)	(0.11)	(0.10)	(0.10)	(0.10)	(0.08)
β_1	0.10	0.35*	0.36***	0.69***	0.22*	0.17	0.14
	(0.09)	(0.18)	0.12	(0.22)	(0.12)	(0.13)	(0.09)

	Germany	**Greece**	**India**	**Indonesia**	**Israel**	**Italy**	**Japan**	**UK**
α_0	-3.08	-3.45	-4.57*	-3.95	-6.41*	-2.76	-0.15	-1.36
	(1.95)	(2.27)	(2.52)	(3.12)	(3.63)	(1.97)	(1.49)	(1.33)
β_0	4.24***	3.78**	4.06***	2.65	4.61***	3.09***	3.68***	3.06***
	(1.03)	(1.48)	(1.01)	(1.83)	(0.68)	(0.97)	(1.00)	(0.96)
α_1	-0.18*	-0.20**	-0.03	-0.19**	-0.17*	-0.20**	-0.16**	-0.21**
	(0.10)	(0.10)	(0.04)	(0.07)	(0.09)	(0.10)	(0.08)	(0.11)
β_1	0.22*	0.23*	0.16	0.25*	0.31*	0.23*	0.12	0.19*
	(0.13)	(0.12)	(0.11)	(0.14)	(0.17)	(0.12)	(0.09)	(0.11)

Source: WTO (2017).
Note: Standard errors in parentheses; *** indicates a significance level of 1%, ** of 5%, and * of 10%.

real exports coefficient, which measures the speed of adjustment to the equilibrium level in one period, is significant for 12 countries, while the lagged world GDP has significant effect on change in the real exports in case of 10 countries, but for most only at 10% level.

To check if there is in fact a relationship between trade and income in the long run, further testing is required. First, F-test (Table 2.6) is used to estimate the joint effect of lagged total exports and world income on the dependent variable. Since there is some ambiguity about the distribution of the data under consideration, following Pesaran et al. (2001) the non-standard critical values are obtained to check the joint significance of the explanatory variables. These non-standard critical values are 4.94 for I(0) data and 5.73 for I(1) data. Even using standard critical values there is not a lot of evidence of a jointly significant effect of the lagged level terms, and using the non-standard critical values for I(1) variables − none at all.

The test for the long-run effect of world income on exports is based on a non-linear function of the parameters, hence, we then use a Wald test to check whether there is in fact a long-run relationship between income and exports. It is worth noting that the Wald test is quite sensitive to how the expression for the null hypothesis is written. For the purpose of checking whether there is a long-run world income effect on the real exports, we write the null hypothesis for the test as: $-\beta_1/\alpha_1 = 0$ (where, from (2.15), α_1 is the lagged real exports

Table 2.6 Equation (2.15): Existence of the Long-Run Relationship (Redundant Variables)

	F-statistic	Probability
Australia	1.38	0.26
Austria	2.04	0.14
Canada	4.75	0.01
Chile	5.22	0.01
Denmark	2.20	0.12
Finland	1.21	0.31
France	2.59	0.09
Germany	1.66	0.20
Greece	2.12	0.13
India	3.89	0.03
Indonesia	3.82	0.03
Israel	1.86	0.17
Italy	2.40	0.10
Japan	4.21	0.02
UK	2.15	0.13

Source: WTO (2017). Note: null hypothesis: lnX_{t-1} and lnY^*_{t-1} are jointly insignificant.

and β_1 is the lagged world income). The coefficients for the income elasticity of exports in the short and in the long run are summarised in the Table 2.7.

The Wald test results suggest that world GDP has a significant effect on trade growth in the long run for all countries except two, Australia and India. The average income elasticity of trade is about 4.08 in the short run and 1.26 in the long run, and it is unclear why trade seems to be more responsive to changes in world income in the short run than in the long run. Nonetheless, this paradoxical result is consistent with other empirical papers on income-trade relationship (Constantinescu et al., 2015). In addition, the speed of adjustment coefficients (an appendix, Table A.3) are negative and all, except three, are significant. The results imply a sluggish adjustment of exports, about 0.21 on average.

Hence, the two test procedures, F-test and Wald test produce somewhat conflicting results. This may happen in small samples or when there are structural breaks in the data. Since, our sample is not necessarily small, it might be helpful to continue our analysis by checking for the structural instability.

To sum up, there seems to be a strong relationship between world income and exports in the short run. However, when it comes to estimating the significance of this relationship in the long run, the F-test and Wald test procedures conflict. This does not allow for a conclusion without doubt as to whether there is a long-run effect of world income on the real exports or not. Since one possible explanation for these conflicting results is the structural

Table 2.7 Equation (2.15): Short-Run and Long-Run Income Elasticities

	Short-run		**Long-run**	
	Coefficient	SE	Coefficient	SE
Australia	5.33***	1.01	3.04	3.63
Austria	3.93***	1.03	1.63***	0.17
Canada	4.40***	0.66	1.12***	0.07
Chile	5.58***	1.91	2.40***	0.28
Denmark	2.81***	0.93	1.07***	0.16
Finland	5.01***	1.16	1.10***	0.28
France	3.48***	0.91	0.84***	0.22
Germany	4.24***	1.03	1.26***	0.20
Greece	3.78**	1.48	1.14***	0.25
India	4.06***	1.01	6.14	6.54
Indonesia	2.65	1.83	1.32***	0.37
Israel	4.61***	0.68	1.82***	0.15
Italy	3.09***	0.97	1.12***	0.18
Japan	3.68***	1.00	0.73**	0.29
UK	3.06***	0.96	0.89***	0.16
Average	4.08		1.26	

Source: WTO (2017). Note: *** indicates a significance level of 1%, ** of 5%, and * of 10%. SE stands for standard error.

instability, we estimate the global income-exports relationship further to see if we can obtain more consistent results when we allow for structural breaks.

2.7 Structural Stability Analysis (Exports; 15 Countries)

One possible reason for some inconsistency in the regression results described above is the presence of structural breaks. There are many tests for structural stability available. We use Bai-Perron, Quandt-Andrews and Chow tests to check whether they produce similar results. The results, summarised in the Table 2.8, truly bring out the sensitivity of our estimates to the choice of test.

Using the Bai-Perron method of testing for structural stability, it is found that the data for seven countries, such as Austria, Denmark, France, Germany, Greece, India and Italy, do not have any significant structural breaks over the estimated period. However, for Austria, Denmark and France the Chow test provides weak evidence of one breakpoint that occurs at year 1986, but the test statistics is only significant at a 10% level for Austria and Denmark (and 5% level for France). The Quandt-Andrews test confirms this breakpoint for France

Table 2.8 Equation (2.15): Structural Instability Analysis

	Bai-Perron Test[1]				Q-A Test[2]			Chow Breakpoint Test		
	N.[3]	Years	F-stat.[4]	Critical value	Year	LR F-stat.	Prob.[5]	Year	F-stat.	Prob.[5]
Australia	1	2008	8.14	18.11	2008***	6.03	0.0021	2008***	6.03	0.00
Austria	0		9.82	16.19	1986	2.45	0.3957	1986*	0.06	0.06
Canada	3	1977	17.90	19.64	1978**	4.76	0.0169	1977***	6.06	0.00
		1993						1978***	4.76	0.00
		2008						1993***	4.69	0.00
								2008***	4.29	0.01
Chile	1	1981	7.50	18.11	1981***	11.71	0.0000	1981***	11.71	0.00
Denmark	0		8.76	16.19	1986	2.19	0.5131	1986*	2.19	0.09
Finland	1	1981	10.23	18.11	1981**	4.38	0.0306	1981***	4.38	0.01
France	0		15.40	16.19	1986*	3.85	0.0676	1986**	3.85	0.01
Germany	0		7.98	16.19	1986	2.00	0.6074	1986	2.00	0.12
Greese	0		11.16	16.19	2002	2.62	0.3324	2002*	2.62	0.05
India	0		14.98	16.19	1978*	3.74	0.0786	1978**	3.74	0.01
Indonesia	1	1977	17.34	18.11	1978***	8.04	0.0001	1977***	10.31	0.00
								1978***	8.04	0.00
Israel	1	1981	14.18	18.11	1981**	4.48	0.0262	1981***	4.48	0.00
Italy	0		15.15	16.19	1980*	3.79	0.0739	1980**	3.79	0.01
Japan	1	1996	17.48	18.11	1996***	5.37	0.0063	1996***	5.37	0.00
UK	2	1981	5.88	18.93	1981***	7.11	0.0003	1981***	7.11	0.00
		1991						1991*	2.40	0.07

Source: WTO (2017).
Note: [1] Bai-Perron multiple breakpoint tests; [2] Quandt-Andrews unknown breakpoint test; [3] - number of breakpoints; [4] - scaled F-statistics; [5] - probability; *** indicates a significance level of 1%, ** of 5%, and * of 10%.

only (at 10% significance level). Furthermore, the Chow test suggests three breakpoints for Canada, while Quandt-Andrews test identifies 1978 as the most likely place for a breakpoint. In case of the UK, the Chow test once again confirms breakpoints in 1981 and 1991, as suggested by the Bai-Perron test, while the Quandt-Andrews test suggests that the breakpoint most likely occurred in 1981. Australia, Chile, Finland, Indonesia, Israel and Japan data happen to have one structural break. All three tests confirm a break in 2008 in Australia, in 1981 in Chile, Finland and Israel, in 1996 in Japan and around 1977 in Indonesia.

Various historical events might have caused sharp changes in real exports flows in different countries (a brief historical overview is presented in an appendix A.1). For instance, a drop in real exports of Japan falls close to the Asian financial crisis of 1997, while the most recent financial crisis 2007-2008 appears to have strongly affected only Australia and Canada in our sample. Nonetheless, structural breaks mostly occur at different years in different countries, thus it is difficult to generalise from these results.

The results obtained are based on the exports data from the WTO, but long time-series data on trade flows are also available from the IMF database, which is of a particular interest since one of the recent papers on trade, Constantinescu et al. (2015), established a clear short-run and long-run income-trade relationship using this data source. Therefore, next we compare the results obtained using the WTO and IMF datasets as well as attempt to identify the most likely place for a common breakpoint in the data.

2.8 IMF and WTO Datasets: Comparison (Imports; World and 15 Countries Sample)

In estimating income elasticity of trade, Constantinescu et al. (2015)[1] used IMF WEO data, supplementing the post-1980 series of the publicly available dataset with the data for the period 1970-1979 (which are accessible only from within the IMF). Using the final data received from Dr. Constantinescu directly (presented in an appendix, Table A.4), we successfully replicate their results (an appendix, Table A.5) and then compare these estimates to the results previously described in this chapter.

To begin with, Constantinescu et al. (2015) uses ECM

$$ECM : \Delta lnM_{it} = \alpha + \beta \Delta lnY_{it} + \delta lnY_{it-1} + \gamma lnM_{it-1} + \varepsilon_{it}, \tag{2.16}$$

where M_{it} is the volume of imports of goods and services, Y_{it} is the GDP growth at the purchasing power parity (PPP) rates, both variables being in the index form (year 2000 = 100), and ε_{it} is an error term.

The estimated income elasticity of trade comes to 2.82 for the short-run and 1.7 for the long run. They continue their analysis by splitting the data in three approximately equal time-periods and estimating the following model:

$$\begin{aligned}\Delta lnM_{it} = \alpha_1 * DV_1 + \beta_1 \Delta lnY_{it} * DV_1 + \gamma_1 lnM_{it-1} * DV_1 + \delta_1 * lnY_{it-1} * DV_1 + \\ + \alpha_2 * DV_2 + \beta_2 \Delta lnY_{it} * DV_2 + \gamma_2 lnM_{it-1} * DV_2 + \delta_2 * lnY_{it-1} * DV_2 + \\ + \alpha_3 * DV_3 + \beta_3 \Delta lnY_{it} * DV_3 + \gamma_3 lnM_{it-1} * DV_3 + \delta_3 * lnY_{it-1} * DV_3 + \varepsilon_{it},\end{aligned} \tag{2.17}$$

where DV_1 is a dummy for 1970-1985 period, DV_2 is a dummy for 1986-2000 period and DV_3 is a dummy for 2001-2013 period.

Comparing their estimates with our results (for 15 countries, using the WTO exports data), our short-run elasticity averages to 4.08, which is much higher than 2.82, corresponding estimate from the Constantinescu et al. (2015). However, our average long-run elasticity, 1.26, is more comparable to theirs, 1.70, especially taking into account that their long-run income

[1] I would like to thank Dr. Ileana Cristina Constantinescu for sending me the dataset she used in her joint paper with Dr. Mattoo and Dr. Ruta "Global Trade Slowdown: Cyclical or Structural?" so that I could replicate their results and compare my estimates with theirs.

elasticity of trade is lower, 1.31, for periods 1970-1985 and 2001-2003. An average speed of adjustment for the sample is 0.21, which was considered sluggish before, but appears to be higher than 0.12 from the Constantinescu et al. (2015) study. Therefore, it might be within norm for the speed of adjustment to be relatively low when estimating the income-trade relationship. In addition, as in case with the WTO estimations, a trend variable appears to be insignificant when estimating the income-trade relationship using IMF data.

One difference between the WTO data we used for 15 countries and Constantinescu et al. (2015) dataset is that they used imports in their estimations, not exports. In order to run a close comparison between estimations obtained using the WTO versus the IMF, we collect the WTO data for imports for the world and for our sample of 15 countries. Since the IMF data for individual countries are not available for 1970-1979, in order to allow for direct comparison of the WTO and IMF estimates for 15 countries, the estimation period is shortened to 1980-2013. Thus, splitting it in three sub-periods with breaks in 1985 and in 2000 would be cutting the first period too short. If the data are broken approximately in half, then the breakpoint is around 2000.

The IMF and WTO estimates suggest different breakpoints for different countries (the results are presented in an appendix, Table A.7). However, the Chow test confirms a break in 2000 for seven countries for the WTO data and for five economies and the world for the IMF data. Overall, the year 2000 might be if not ideal then at least a suitable choice for a common breakpoint. Hence, we try to estimate income-trade relationship using the WTO and IMF data for the whole period (1980-2013) and then for two sub-periods, pre-2000 and post-2000.

2.8.1 IMF Dataset (1980-2013 and 2 Sub-Periods)

We use IMF data for 1980-2013 to estimate income elasticity of imports for the world, for a whole sample (i.e. averages, calculations description is in an appendix, A.4.5) and for 15 countries individually. We use (2.16):

$$\Delta lnM_{it} = \alpha + \beta \Delta lnY_{it} + \delta lnY_{it-1} + \gamma lnM_{it-1} + \varepsilon_{it},$$

where M_{it} is the real imports of goods and services, Y_{it} is the real GDP and ε_{it} is an error term.

The results for this analysis are summarised in the Table 2.9. The long-run income elasticity coefficient is insignificant for the world in 1980-2013, but for 14 countries in the sample, the exception is Greece, this coefficient is significant and averages to 2.35. The short-run elasticity coefficient is significant for all countries and for the world. On average, short-run elasticity of income is about 2.27 for the sample and 3.19 for the whole world.

Table 2.9 Income Elasticity of Imports: IMF Data (1980-2013; Equation 2.16)

	Short-Run Elasticity (β)		**Long-Run Elasticity ($-\delta/\gamma$)**		**Whole Period (1980-2013)**	
	Pre-2000	Post-2000	Pre-2000	Post-2000	Short-run	Long-run
World	2.06***	3.43***	2.42***	1.29***	3.19***	1.32
	(0.41)	(0.23)	(0.47)	(0.07)	(0.31)	(0.80)
15 countries	2.93***	2.93***	3.24*	2.37***	2.91***	2.50***
	(0.45)	(0.24)	(1.71)	(0.07)	(0.23)	(0.12)
Australia	2.28***	4.72**	1.96***	2.68***	2.51***	2.16***
	(0.68)	(2.02)	(0.11)	(0.15)	(0.67)	(0.09)
Austria	2.19**	2.98***	2.07***	1.93***	2.75***	2.10***
	(0.99)	(0.26)	(0.13)	(0.18)	(0.39)	(0.05)
Canada	2.31***	3.58***	2.57***	2.13***	2.65***	2.02***
	(0.37)	(0.34)	(0.14)	(0.17)	(0.57)	(0.16)
Chile	2.64***	3.63***	0.15	2.11***	2.60***	2.07***
	(0.25)	(0.66)	(17.14)	(0.09)	(0.24)	(0.42)
Denmark	0.60	3.54	3.46***	9.06***	1.93**	3.88***
	(1.26)	(0.98)	(0.56)	(1.63)	(0.82)	(0.49)
Finland	1.98***	1.85***	2.22***	2.14***	2.00***	2.19***
	(0.27)	(0.25)	(0.30)	(0.11)	(0.17)	(0.08)
France	2.81***	2.90***	2.48***	2.42***	2.92***	2.62***
	(0.42)	(0.26)	(0.19)	(0.12)	(0.26)	(0.07)
Germany	1.88***	1.42**	1.24*	4.96***	1.94***	3.87***
	(0.36)	(0.49)	(0.61)	(1.31)	(0.25)	(1.19)
Greece	1.54***	2.18***	3.93***	1.23***	1.73***	1.33
	(0.46)	(0.45)	(0.24)	(0.30)	(0.32)	(1.09)
India	2.70**	2.69**	1.24***	1.44***	2.56***	1.38***
	(0.93)	(0.90)	(0.24)	(0.19)	(0.79)	(0.07)
Indonesia	1.07	5.91	1.12***	1.24***	1.18**	1.12***
	(0.66)	(5.28)	(0.17)	(0.26)	(0.51)	(0.08)
Israel	1.82***	2.08***	1.39***	0.63***	2.19***	1.06***
	(0.58)	(0.45)	(0.08)	(0.08)	(0.38)	(0.12)
Italy	2.85***	2.78***	2.54***	-3.37	2.85***	3.09***
	(1.19)	(0.41)	(0.17)	(73.38)	(0.38)	(0.42)
Japan	1.91*	2.71***	2.94**	2.90***	2.33***	3.40***
	(0.96)	(0.34)	(1.26)	(0.29)	(0.47)	(0.83)
United Kingdom	1.70***	2.29***	1.83***	1.73***	1.89***	1.88***
	(0.33)	(0.37)	(0.07)	(0.09)	(0.20)	(0.03)
Average	2.20	2.76	2.21	2.62	2.27	2.35

Source: IMF World Economic Outlook April 2014. Note: Standard errors in parentheses; *** indicates a significance level of 1%, ** of 5%, and * of 10%. 15 countries - aggregated (calculated using weights).

After breaking the data in two sub-periods the majority of coefficients remain significant. The short and long-run income elasticities for the whole sample in 1980-2000 come to 2.20 and 2.21, respectively, and for the world to 2.06 and 2.42, correspondingly. However, over the post-2000 period, the short-run and long-run elasticities for the sample are about 2.76 and 2.62, respectively, and for the world are 3.43 and 1.29, correspondingly. For the aggregate of 15 countries, there is a clear indication of income-trade relationship in the short and in the long run during 1980-2013. Furthermore, when the data are split in pre-2000 and post-2000 periods, the short-run and long-run income elasticity coefficients are still significant in 1980-1999 (2.93 and 3.24, correspondingly) and in 2000-2013 (2.93 and 2.37). In addition, long-run elasticity coefficient for post-2000 is lower than for pre-2000, which is consistent with the results from Constantinescu et al. (2015), and suggests that trade flows might be becoming less responsive to changes in income.

2.8.2 WTO Dataset (1980-2013 and 2 Sub-Periods)

For comparison we use the WTO data for the world and sample of 15 countries for 1980-2013 to estimate (2.16):

$$\Delta lnM_{it} = \alpha + \beta \Delta lnY_{it} + \delta lnY_{it-1} + \gamma lnM_{it-1} + \varepsilon_{it},$$

where M_{it} is the real imports of goods and services, Y_{it} is the real GDP, and ε_{it} is an error term.

The results are summarised in the Table 2.10. The long-run (2.18) and short-run (4.80) income elasticity coefficients for the world are higher than the ones calculated using IMF data (1.32 and 3.19, respectively). Similarly, for the aggregate of 15 countries: the long-run coefficient (3.01) is higher than using the IMF data (2.50) for 1980-2013, while the short-run income elasticity comes to 4.87, comparing to 2.91 using the IMF. For individual countries, long-run income elasticity is significant for 13 out of 15 countries and averages to 1.85, which is close to 1.70 from the Constantinescu et al. (2015). The short-run income elasticity coefficient is significant for all countries, except Greece, and averages to 3.25, which is higher than 2.82 from the Constantinescu et al. (2015) paper and 2.27, using the IMF data for 1980-2013.

When the data are split in two sub-periods, short-run and long-run elasticity coefficients are significant for the whole world (except the world short-run coefficient for 1980-2000) and for most countries in the sample. The average long-run income elasticity for 15 countries comes to 1.48 in 1980-2000, which is lower than 2.21, using the IMF data, and to 3.22 in 2001-2013, which is higher than 2.62, using IMF data. The short-run income elasticity for the aggregate of 15 countries is again higher for both periods, using the WTO data. The coefficients for income elasticity in pre-2000 and post-2000 periods come to 3.92 and 4.96, respectively, using the WTO data, comparing to 2.93 for both periods using the IMF data.

To sum up, there is significant effect of income on imports both, in the short and in the long run. However, the findings are sensitive to the data used in the analysis, and the estimations do not match very closely.

Table 2.10 Income Elasticity of Imports: WTO Data (1980-2013; Equation 2.16)

	Short-Run Elasticity (β)		**Long-Run Elasticity ($-\delta/\gamma$)**		**Whole Period (1980-2013)**	
	Pre-2000	Post-2000	Pre-2000	Post-2000	Short-run	Long-run
World	2.39	5.57***	1.55**	2.57***	4.80***	2.18***
	(1.56)	(0.99)	(0.56)	(0.25)	(0.91)	(0.70)
15 countries	3.92**	4.96***	1.83*	3.63***	4.87***	3.01**
	(1.68)	(1.16)	(0.88)	(0.43)	(0.97)	(1.39)
Australia	2.84**	9.57**	0.87***	3.14***	3.23***	1.73***
	(0.99)	(3.02)	(0.26)	(0.23)	(1.12)	(0.43)
Austria	4.42**	4.83***	1.64***	3.66***	3.81***	2.18***
	(2.00)	(1.29)	(0.56)	(0.51)	(1.05)	(0.29)
Canada	2.32***	4.96***	1.50***	2.59***	2.87***	1.46***
	(0.44)	(0.72)	(0.16)	(0.36)	(0.44)	(0.10)
Chile	3.70***	7.27***	1.33	2.61***	3.86***	2.35*
	(0.51)	(1.84)	(1.39)	(0.53)	(0.52)	(1.17)
Denmark	0.68	3.96***	1.60***	5.44***	2.07**	2.10***
	(1.46)	(0.83)	(0.44)	(0.53)	(0.90)	(0.58)
Finland	2.60***	4.46***	0.92	3.92***	3.04***	1.85***
	(0.82)	(0.49)	(0.54)	(0.33)	(0.58)	(0.51)
France	3.80**	5.24***	1.50**	4.06***	3.76***	2.04***
	(1.74)	(1.38)	(0.59)	(0.61)	(1.13)	(0.45)
Germany	4.16***	3.32***	1.74***	3.71	3.47***	2.54***
	(1.32)	(1.12)	(0.46)	(2.68)	(0.77)	(0.49)
Greece	0.41	0.88	2.97***	2.27**	0.94	2.13***
	(1.20)	(1.00)	(0.55)	(0.91)	(0.56)	(0.23)
India	3.25**	2.69	0.57**	2.36***	2.93**	1.81***
	(1.14)	(2.17)	(0.24)	(0.26)	(1.17)	(0.47)
Indonesia	2.05**	22.32***	0.90	1.30**	2.51***	1.39***
	(0.74)	(4.85)	(0.54)	(0.49)	(0.70)	(0.34)
Israel	3.03***	2.85***	0.91***	1.11***	3.05***	0.94***
	(0.58)	(0.89)	(0.11)	(0.15)	(0.50)	(0.05)
Italy	7.90***	5.08***	3.73	31.80	5.65***	-8.79
	(1.49)	(1.24)	(4.22)	(87.64)	(0.98)	(25.79)
Japan	1.77	4.64***	1.25	6.64***	3.18***	4.20
	(1.24)	(1.29)	(0.76)	(1.55)	(0.93)	(3.25)
United Kingdom	0.89	3.91***	1.63***	2.71***	2.05**	1.58***
	(1.18)	(0.95)	(0.33)	(0.41)	(0.79)	(0.24)
Average	3.64	6.34	1.48	3.22	3.25	1.85

Source: WTO (2017). Note: Standard errors in parentheses; *** indicates a significance level of 1%, ** of 5%, and * of 10%. 15 countries - aggregated.

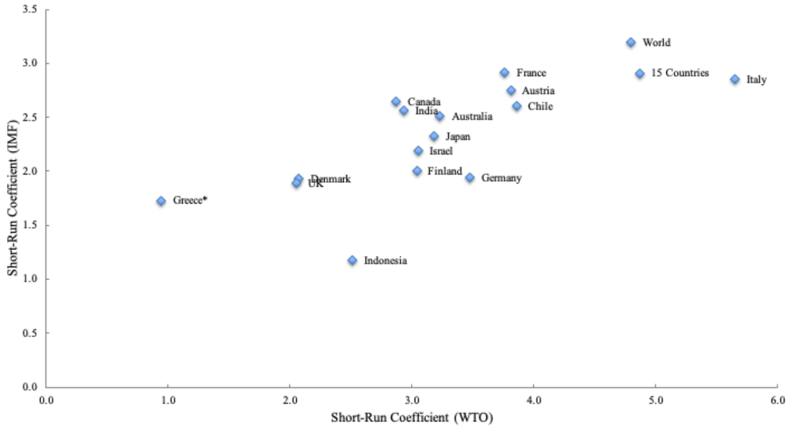
2.8.3 Relationship Between The Estimates

Another way of looking at how similar the estimates obtained using two datasets are, is by using the scatter diagrams for the short-run (Figure 2.12 Part 1) and long-run (Figure 2.12 Part 2) elasticity coefficients for the whole period.

It appears that in many cases the coefficients calculated using IMF and WTO data are comparable with each other. For instance, the diagram for the short-run coefficients (Figure 2.12 Part 1) suggests there is a moderately strong positive relationship between the IMF and WTO coefficients. For some countries the coefficients calculated using the IMF and WTO data are very similar, such as in case of Canada the short-run coefficient equals to 2.07 using

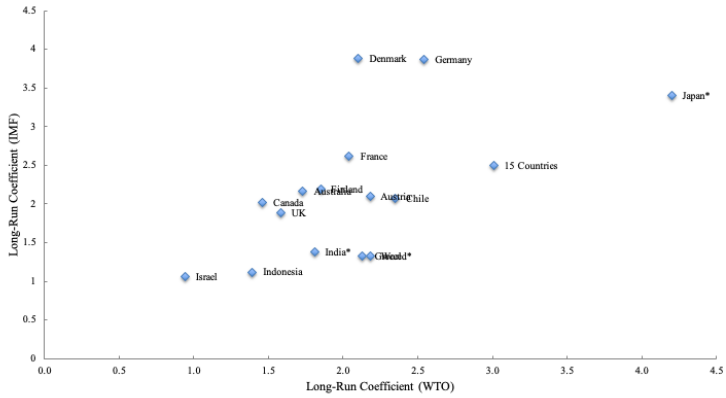
Fig. 2.12 Relationship Between The Estimates (Short-Run and Long-Run Coefficients; 1980-2013).

Part 1. Short-Run Coefficients.



Source: WTO (2017) and IMF (2014). Note: * indicates countries for which one or both coefficients (WTO or/and IMF) are insignificant.

Part 2. Long-Run Coefficients.



Source: WTO (2017) and IMF (2014).
Note: * indicates countries for which one or both coefficients (WTO or/and IMF) are insignificant. Coefficient for Italy is omitted as it is negative and insignificant.

the WTO data and to 1.93 using the IMF data, respectively. However, for other countries the discrepancy between the two coefficients (IMF and WTO) is noticeably larger. For instance, in case of Italy the WTO short-run coefficient is almost twice as large (5.65) as the IMF one (2.85).

Moving to the second scatter plot (Figure 2.12 Part 2), the relationship between the IMF and WTO long-run coefficients is weaker than in the short-run case. There are also countries, such as Japan, for which WTO coefficient is insignificant and on the scatter plot it appears to be the outlier.

Nonetheless, in general there is a positive relationship between the IMF and WTO short-run and long-run coefficients when considering the whole period.

2.9 The Marshall-Lerner Condition: The UK Case

We now check whether the Marshall-Lerner condition holds for two countries in our sample, the UK and US, for which longer quarterly data are available.

2.9.1 The UK Exports

VAR

We begin this analysis by first focusing on the exports demand equation (2.3) for the UK:

$$lnX_{it} = \beta_0 + \beta_1 lnPX_{it} + \beta_2 lnY_t^* + v_{it}.$$

Starting from the unrestricted VAR (2.5):

$$\boldsymbol{w}_{it} = \boldsymbol{a} + \sum_{i=1}^{p+1} \boldsymbol{\Gamma w}_{it-1} + \boldsymbol{v}_{it},$$

where $\boldsymbol{w}_{it}$ is the exports vector $(lnX_{it}, lnY_t^*, lnPX_{it})$.

The lag length of two is chosen using AIC selection criteria. The Granger Causality test for VAR(2) model suggests that past values of the G7 income and UK exports prices seem to predict current level of British exports, and there are no other Granger-causality links among the variables.

Analysis of the generalised impulse response functions suggests that all own effects are significant. A shock to the G7 GDP has a significant positive effect on exports, while a shock to exports does so on prices, although the effect might not be substantial.

The residuals correlation matrix for three variables suggests that the correlation between the G7 GDP and exports is 0.114, while that of relative prices with GDP (-0.003) and with exports (-0.025) are much lower and negative.

The UK exports, relative prices and G7 GDP time series are integrated of order one, using the ADF test (the lag length is chosen using BIC). Therefore, we run the Johansen's test for cointegration and proceed with a model with an intercept and without a trend in the cointegrating equation.

VECM

Next we estimate (2.6):

$$\boldsymbol{\Delta w}_{it} = \boldsymbol{\mu} + \boldsymbol{\alpha\beta w}_{it-1} + \sum_{i=1}^{p} \boldsymbol{\Gamma \Delta w}_{it-1} + \boldsymbol{v}_{it},$$

where $\boldsymbol{w}_{it}$ is the exports vector $(lnX_{it}, lnY_t^*, lnPX_{it})$.

The VAR(2) corresponds to the VECM(1). In VECM(1) long-run income elasticity of exports is 1.38, while the price one is -0.39 (Table 2.11). The loading coefficient for exports, -0.10, measures the extent to which deviations from the long-run relationship feedback on export prices. The speed of adjustment for the G7 income is -0.01 (on the edge of significance), and it is -0.02 and insignificant for export prices, meaning both variables can be treated as weakly exogenous and, thus, used in the ARDL model. Furthermore, the adjustment coefficients for exports and prices have the correct sign, while that for GDP has the wrong sign, but is barely significant at 5% level.

Table 2.11 Exports (The UK)

	Estimation Summary		**Hooper et al. (2000)**	
	ARDL/ECM	**VECM**	**ARDL/ECM**	**VECM**
Estimation Period	1960Q2-2017Q1	1960Q4-2017Q1	1977Q1-1994Q4	
Income S-R El-ty[1]	0.38	-	1.09	-
	(0.26)		(0.65)	
Income L-R El-ty[2]	1.43***	1.38***	-	1.11
	(0.03)	(0.04)		(0.32)
Price S-R El-ty	-0.04	-	-0.24	-
	(0.05)		(0.10)	
Price L-R El-ty	-0.22*	-0.39***	-	-1.55
	(0.12)	(0.15)		(0.51)
Speed of Adj.[3]	-0.14***	-0.10 ***	-0.03	-0.02
	(0.03)	(0.03)	(0.01)	(0.02)
R-Squared	0.12	0.17	0.74	-
SER	0.03	0.03	1.18	-
Coin. Vectors[4]	-	1	-	1
Lag Length	-	1	-	4

Note: *** indicates a significance level of 1%, ** of 5%, and * of 10%; [1]S-R El-ty - short-run elasticity; [2]L-R El-ty - long-run elasticity; [3] speed of adjustment; [4]cointegrating vectors.

ARDL/ECM

The ARDL model is estimated in the ECM form as in (2.7),

$$\Delta lnX_{it} = \alpha_x + \beta_x \Delta lnY_t^* + \eta_x \Delta lnPX_{it} + \rho_x lnX_{it-1} + \mu_x lnY_{t-1}^* + \nu_x lnPX_{it-1} + \varepsilon_{it}^x,$$

because it simplifies interpretation of coefficients. World income and relative prices do not have a statistically significant effect on the UK exports in the short run. There might be various explanations for this. For instance, British exporters might be tied by the contract-related liabilities and not being able to quickly adjust export volumes to a new desirable level in accordance to change in prices or economic conditions in the world.

The loading coefficient for export demand equation is 0.14, while in Hooper et al. (2000) it is 0.03 (3% per quarter), hence, our model suggests a faster quarterly adjustment of exports to equilibrium. All the lagged coefficients are significant at 5% level in exports model. The long-run income elasticity of exports is 1.43, and the relative prices also seem to affect exports over time, with the price elasticity of exports coming to -0.22.

Structural Stability

Moving to the structural stability analysis using ECM, the Sequential Bai-Perron test identified breakpoints in 1969Q2, 1977Q4 and 1989Q3. The first date is also confirmed by the Q-A and Chow tests. Splitting data in three sub-periods (Table 2.12) does not improve significance of the independent variables. From the first sub-period to the next two, the short-run price elasticity was falling to zero, and the long-run income elasticity was increasing, from 0.71 in 1960Q2-1969Q2 to 1.69 in 1989Q4-2017Q1. Furthermore, after estimating (2.9) one can conclude that the relationship between the G7 GDP and UK exports did not substantially change following the financial crisis 2007-2008.

Table 2.12 Structural Stability Analysis: Exports (The UK). Dependent Variable is ΔlnX_t

	1960Q2-69Q2	**1969Q3-89Q3**	**1989Q4-2017Q1**	**Whole Period**
ΔlnY_t^*	0.45	0.19	1.47***	0.38
$\Delta lnPX_t$	-0.49	0.06	-0.02	-0.04
lnM_{t-1}	-0.81***	-0.37***	-0.30***	-0.14***
lnY_{t-1}^*	0.57***	0.44***	0.50***	0.21***
$lnPX_{t-1}$	-0.65***	0.01	0.03	-0.03*
LRGDP	0.71***	1.16***	1.69***	1.43***
LRPX	-0.81***	0.02	0.10	-0.22*
BIC	-4.09	-3.62	-4.42	-3.99
SER	0.03	0.04	0.02	0.03

Note: *** indicates a significance level of 1%, ** of 5%, and * of 10%. LRGDP and LRRP - long-run income and price elasticities, respectively. BIC - Bayesian information criterion; SER - standard error of regression.

Some discrepancies between the results presented in this chapter and the estimates from Hooper et al. (2000) might be due to different estimation periods or difference in the way the

G7 GDP was measured. Since the G7 income data were not readily available, we calculated G7 GDP by summing up G7 members' GDP series expressed in constant US dollars, while Hooper et al. (2000) agglomerated their data using bilateral export shares from Direction of Trade report by IMF (1995). Nonetheless, findings presented in this section confirm that there is a strong effect of world income and relative prices on UK exports in the long run.

2.9.2 The UK Imports

VAR

We now focus on the import demand equation (2.1):

$$lnM_{it} = \alpha_0 + \alpha_1 lnPM_{it} + \alpha_2 lnY_{it} + \mu_{it}.$$

Starting with an assumption that all variables are co-determined, the unrestricted VAR model, (2.5), is estimated for the UK for 1955Q1-2018Q1:

$$\boldsymbol{w}_{it} = \boldsymbol{a} + \sum_{i=1}^{p+1} \boldsymbol{\Gamma} \boldsymbol{w}_{it-1} + \boldsymbol{\nu}_{it},$$

where $\boldsymbol{w}_{it}$ is the imports vector $(lnM_{it}, lnY_{it}, lnPM_{it})$.

BIC and Hannan–Quinn information criterion suggest lag length of two as optimal, while AIC and final prediction error criterion suggest three lags instead. The VAR(3) model seems to fit data slightly better, thus we proceed with three lags. In the VAR(3) model all variables are Granger-causal for each other, except import prices do not predict GDP.

The GIRFs show that all own effects are significant, and shocks to imports have significant positive effect on the GDP and visa versa. The correlation between GDP and imports is 0.343 while that for relative prices and imports (0.016) or income (0.122) are much lower.

Variables are confirmed to be all integrated of order one and according to the AIC, there is one cointegrating vector when long-run relationship among the variables is modelled with an intercept and a linear trend.

VECM

We then estimate the VECM(2), as in (2.6),

$$\Delta \boldsymbol{w}_{it} = \boldsymbol{\mu} + \boldsymbol{\alpha}\boldsymbol{\beta} \boldsymbol{w}_{it-1} + \sum_{i=1}^{p} \boldsymbol{\Gamma} \Delta \boldsymbol{w}_{it-1} + \boldsymbol{\nu}_{it},$$

where $\boldsymbol{w}_{it}$ is the imports vector $(lnM_{it}, lnY_{it}, lnPM_{it})$.

In the VECM(2) with a trend, the change in the GDP lagged two periods has a very significant effect on imports. Both, the UK GDP and import prices, have significant effect on the UK imports in the long run. Income elasticity of imports is 0.73 (Table 2.13), lower than 2.21 in Hooper et al. (2000), and price elasticity is -1.82, while in their paper it is -0.58 and insignificant. The trend coefficient is positive and significant, although very small (0.004). The speed of adjustments are -0.10 for imports, -0.02 for GDP and -0.02 for imports prices. Both, the GDP and prices, may be considered weakly exogenous as loading coefficients for both are close to the edge of significance. Moreover, the income and prices feedback coefficients are quite small, about 2% per quarter, thus, we will treat both variables, the income and prices, as weakly exogenous and proceed by estimating the ARDL model. In addition, similar to the UK exports, the adjustment coefficients for imports and prices have the correct sign, for the GDP it has the wrong sign and is insignificant.

Table 2.13 Imports (The UK)

	Estimation Summary		**Hooper et al. (2000)**	
	ARDL/ECM	**VECM**	**ARDL/ECM**	**VECM**
Estimation Period	1955Q2-2018Q1	1955Q4-2018Q1	1956Q2-1994Q4	
Income S-R El-ty[1]	1.10***	-	1.01***	-
	(0.19)		(0.30)	
Income L-R El-ty[2]	1.65***	0.73**	-	2.21***
	(0.11)	(0.35)		(0.82)
Price S-R El-ty	-0.09	-	X*	-
	(0.17)			
Price L-R El-ty	-0.95**	-1.82***	-	-0.58
	(0.40)	(0.39)		(1.50)
Trend	-	-2.004**	-	-
	-	(0.002)	-	-
Speed of Adj.[3]	-0.11***	-0.10***	-0.02**	0.01
	(0.03)	(0.02)	(0.01)	(0.01)
R-Squared	0.19	0.18	0.23	-
SER	0.03	0.03	2.91	-
Coin. Vectors[4]	-	1	-	1
Lag Length	-	2	-	5

Note: *** indicates a significance level of 1%, ** of 5%, and * of 10%; [1]S-R El-ty - short-run elasticity; [2]L-R El-ty - long-run elasticity; [3] speed of adjustment; [4]cointegrating vectors; *estimation is missing in the paper by Hooper et al. (2000).

ARDL/ECM

By estimating the ECM, (2.8),

$$\Delta lnM_{it} = \alpha_m + \beta_m \Delta lnY_{it} + \eta_m \Delta lnPM_{it} + \rho_m lnM_{it-1} + \mu_m lnY_{it-1} + \nu_m lnPM_{it-1} + \varepsilon_{it}^m,$$

we find that short-run income coefficient is 1.10, close to 1.0 estimated in Hooper et al. (2000). The long-run income elasticity coefficient is smaller for our sample (1.65 comparing to theirs of 2.2), however, their sample covers significantly shorter time period. As for the price elasticities, in Hooper et al. (2000) both were insignificant for the UK. Our estimate of the short-run price elasticity is also insignificant, but the long-run one is significant, negative and close to unity (-0.95). Our estimate of the speed of adjustment is much higher than in Hooper et al. (2000) analysis, 11% per quarter (comparing to 2%), and it is significant in both cases.

Structural Stability

The Sequential Bai-Perron and Chow tests suggest that the likely places for structural breaks are 1981Q3 and 2006Q3. However, breaking data in 1981Q3 and 2006Q3 (Table 2.14) does not affect significance of the coefficients.

Table 2.14 Structural Stability Analysis: Imports (The UK). Dependent Variable is ΔlnM_t

	1955Q2-81Q3	**1981Q4-2006Q3**	**2006Q4-18Q1**	**Whole Period**
ΔlnY_t	1.09***	1.12***	2.11***	1.10***
$\Delta lnPM_t$	0.26	0.32	0.00	-0.09
lnM_{t-1}	-0.61***	-0.28***	-0.55***	-0.11***
lnY_{t-1}	0.98***	0.63***	0.97***	0.19***
$lnPM_{t-1}$	0.26***	0.04	0.09	-0.11**
LRGDP	1.60***	2.22***	1.78***	1.65***
LRRP	0.43***	0.13	0.17	-0.95**
BIC	-3.92	-4.48	-5.29	-4.22
SER	0.03	0.02	0.01	0.03

Note: *** indicates a significance level of 1%, ** of 5%, and * of 10%. LRGDP and LRRP - long-run income and price elasticities, respectively. BIC - Bayesian information criterion; SER - standard error of regression.

Nonetheless, estimates for the sub-periods suggest that short-run income elasticity was increasing over this time, while price one was mostly fluctuating around zero. The long-run price elasticity coefficient seems to be significant only in the first sub-period, 1955Q2-1981Q3, and when estimated over the whole period. Moreover, when dummy variables are created for 1981Q3 and 2006Q3, the coefficients of these dummy variables are significant and their addition somewhat improves the significance of other coefficients. After estimating (2.10) we conclude that, as in case with the UK exports, income elasticity of the UK imports did not seem to have changed following the financial crisis 2007-2008.

Overall, looking at the UK exports and imports price elasticities, their sum is greater than unity (in absolute values). Hence, we can conclude that the Marshall-Lerner condition holds for the UK data, meaning depreciation of pound is likely to improve UK's external balance by making its goods and serviced more competitive and, as a result, increasing demand for them.

2.10 The Marshall-Lerner Condition: The US Case

We repeat the analysis described above for the UK, now using the US data. The estimation procedure is the same, but we arrive to a different conclusion as the Marshall-Lerner condition does not seem to hold for the US data.

2.10.1 The US Exports

VAR

The price and income elasticities of the US exports (2.3) are estimated for 1960Q2-2017Q1. We estimate the unrestricted VAR model (2.5), assuming that all variables might be endogenous and find that the lag length of three is optimal.

In the VAR(3) model all three variables are Granger-causal for each other, with an exception of the G7 GDP not being Granger-causal for the US exports prices. In addition, the GIRFs show that all own effects are significant. Furthermore, shocks to the G7 GDP and prices have positive effect on exports, although the latter is barely significant. Moreover, exports prices respond positively to the shocks to the US exports and G7 income, however, both effects are on the edge of significance. In addition, there is positive correlation between all variables, strongest is between exports and G7 income (0.206).

All variables are integrated of order one, and the AIC suggests that there is one long-run relation among the variables when their relationship modelled with intercept and no linear trend.

VECM

Next we estimate a VECM as in (2.6). According to the VECM(2) coefficients (Table 2.15), all adjustment in the long run is done by the G7 income, with prices being exogenous. The long-run income elasticity of exports comes to 2.38, while the price one is insignificant.

Table 2.15 Exports (The US)

	Estimation Summary		Hooper et al. (2000)	
	ARDL/ECM	**VECM**	**ARDL/ECM**	**VECM**
Estimation Period	1960Q2-2017Q1	1960Q4-2017Q1	1976Q3-1994Q4	
Income S-R El-ty[1]	0.70**	-	1.83**	-
	(0.28)		(0.48)	
Income L-R El-ty[2]	1.85***	2.38***	-	0.83**
	(0.10)	(0.18)		(0.19)
Price S-R El-ty	0.19	-	-0.53**	-
	(0.15)		(0.23)	
Price L-R El-ty	-0.10	-0.50	-	-1.47**
	(0.20)	(0.39)		(0.24)
Speed of Adj.[3]	-0.90***	-0.02**	-0.05	-0.06
	(0.02)	(0.01)	(0.05)	(0.05)
R-Squared	0.10	0.25	0.49	-
SER	0.03	0.03	1.79	-
Coin. Vectors[4]	-	1	-	1
Lag Length	-	2	-	2

Note: *** indicates a significance level of 1%, ** of 5%, and * of 10%; [1]S-R El-ty - short-run elasticity; [2]L-R El-ty - long-run elasticity; [3] speed of adjustment; [4]cointegrating vectors.

The speed of adjustment in the VECM(2) model is quite low (-0.02), comparing to the models discussed previously, suggesting a rather slow quarterly adjustment of the US exports in response to changes in world income and prices. The loading coefficient for prices has the wrong sign and suggests that prices are exogenous, while income seems to respond to the disequilibrium (2.65), but this coefficient is on the edge of significance. Hence, an ARDL model is of interest for estimating long-run income and price elasticities for the US exports. In addition, we find that the US export equation is exceptionally sensitive to the lag order selection. For instance, in the VECM(1) model the long-run price elasticity coefficient becomes significant and has the right sign (-1.27).

ARDL/ECM

According to the ECM (2.7) estimates, the G7 income has significant effect on the US exports in the short run (elasticity equals to 0.70) and in the long run (elasticity is 1.85). However, relative prices do not seem to have strong effect on the US exports either in the short run, or in the long run. The speed of adjustment coefficient is significant, but very high (-0.90), suggesting a 90% quarterly adjustment, which is much higher than the corresponding value for the UK (11%, Table 2.11).

Structural Stability

Hooper et al. (2000) included dummy variables for the US model for 1977Q4 and 1978Q2 to account for oil-price shocks. The Chow test confirms these two breakpoints for our data, however, breaking data in either of these two dates or using dummy variables for them does not improve the model.

Table 2.16 Structural Stability Analysis: Exports (The US). Dependent Variable is ΔlnX_t

	1960Q2-69Q4	**1970Q1-79Q3**	**1979Q4-1989Q2**	**1989Q3-2017Q1**	**Whole Period**
ΔlnY_t^*	-0.85	1.30***	0.14	2.10***	0.70**
$\Delta lnPX_t$	-0.93	0.96***	0.86***	0.60***	0.19
lnX_{t-1}	-1.15***	-1.01***	-0.25***	-0.05**	-0.09***
lnY_{t-1}^*	1.24***	1.38***	0.79***	0.15***	0.16***
$lnPM_{t-1}$	-0.88**	0.53***	0.33***	0.16***	-0.01
LRGDP	1.08***	1.36***	3.17***	3.09***	1.85***
LRRP	-0.77**	0.52***	1.32***	3.50**	-0.10
BIC	-3.22	-4.10	-5.11	-5.38	-3.83
SER	0.04	0.03	0.02	0.01	0.03

Note: *** indicates a significance level of 1%, ** of 5%, and * of 10%. LRGDP and LRRP - long-run income and price elasticities, respectively. BIC - Bayesian information criterion; SER - standard error of regression.

In addition, the Bai-Perron test suggests 1969Q4, 1979Q3 and 1989Q2 as the most likely points for breaks, and all three are confirmed by the Chow test. When we break the data in these years (Table 2.16), in the first sub-period (1960Q2-1969Q4), short-run price elasticity coefficient has the correct (negative) sign, but is insignificant.

Overall, the only general pattern is for long-run price elasticity to increase, starting from the first sub-period onwards. In addition, when we add a dummy variable, as in (2.9), to check for change in the G7 income-US exports relationship following the financial crisis 2007-2008, the short-run income coefficient is significant and comes to 2.53.

2.10.2 The US Imports

VAR

The imports demand equation, (2.1), for the US is estimated for a slightly longer period, 1950Q2 to 2018Q1, than for the UK (1955Q1-2018Q1). When we estimate the unrestricted VAR, (2.5), there appears to be a mutual effect of the UK imports and GDP on each other. Nonetheless, prices do not seem to be affected by the lagged values of imports and income, and prices are insignificant in the VAR equations for both. Hooper et al. (2000) found the lag length of nine to be optimal for the US imports equation. However, for our data the AIC suggests the lag length of two as optimal.

In the VAR(2) model everything is Granger-causal for imports, while imports predict the US GDP, but neither US imports nor GDP forecast prices, as was in case with the UK exports. Considering the GIRFs for VAR(2), all own effects are significant. A shock to imports has a significant positive effect on the US GDP and visa versa. However, prices seem to intact with GDP only. The correlation between GDP and imports is 0.293 while that of relative prices with income is much lower (0.060) and with imports is low and negative (-0.081).

Variables are confirmed to be all integrated of order one and according to the Johansen's cointegration test, there is one cointegrating vector among the variables when the relationship is modelled with an intercept and a quadratic trend.

VECM

Thus, we estimate a VECM, as in (2.6), with one lag and adding a quadratic trend to it. The trend coefficient is only significant in the US income equation, where the trend coefficient is negative and seems to be picking up a structural break in the US output, which was likely to occur around 1970s, time of a major recession in the country. When we compare this model with the VECM(1) with a trend in cointegrating equation, the latter seems to fit the data better and has a slightly higher adjusted R-squared of 0.23 (comparing to 0.21), although the cointegrating vector coefficients are very similar in these two models.

In the VECM(1) with a trend in cointegrating equation, the long-run income elasticity is estimated to be equal to 1.28 (Table 2.17), which is comparable to the estimation of 1.79 from Hooper et al. (2000).

Price elasticity in the long-run is -0.35, which is close to the value of -0.31 from Hooper et al. (2000). The estimated loading coefficient of -0.13 is again very close to -0.10 from Hooper et al. (2000). The trend coefficient in our model is 0.005 and significant. The income and prices speed of adjustment coefficients are insignificant, suggesting that the variables are

Table 2.17 Imports (The US)

	Estimation Summary		Hooper et al. (2000)	
	ARDL/ECM	**VECM**	**ARDL/ECM**	**VECM**
Estimation Period	1950Q2-2018Q1	1951Q2-2018Q1	1961Q4-1994Q4	
Income S-R El-ty[1]	1.50***		2.31**	-
	(0.22)		(0.30)	
Income L-R El-ty[2]	1.90***	1.28***	-	1.79**
	(0.03)	(0.25)		(0.15)
Price S-R El-ty	-0.11		-0.55	-
	(0.08)		(0.14)	
Price L-R El-ty	-0.40***	-0.35***	-	-0.31**
	(0.09)	(0.08)		(0.09)
Speed of Adj.[3]	-0.12***	-0.13***	-0.02	-0.10
	(0.03)	(0.03)	(0.01)	(0.03)
R-Squared	0.22	0.23	0.63	-
SER	0.03	0.03	2.48	-
Coin. Vectors[4]	-	1	-	1
Lag Length	-	1	-	9

Note: *** indicates a significance level of 1%, ** of 5%, and * of 10%; [1]S-R El-ty - short-run elasticity; [2]L-R El-ty - long-run elasticity; [3] speed of adjustment; [4]cointegrating vectors.

exogenous. Therefore, we proceed by estimating the relationship between the US imports, prices and income using the ARDL model.

ARDL/ECM

In the reduced form of the ARDL model, the ECM as in (2.8), income appears to determine imports in the short run and in the long run, while price short-run elasticity coefficient is insignificant (Table 2.17). The income elasticity is 1.50 in the short run, which is slightly higher than in the UK case (1.10), but lower than the estimation of 2.31 from Hooper et al. (2000). As for the long-run income elasticity, it is 1.90 in the US case, which is close to the UK estimate of 1.65 and to the estimation of 1.79 from Hooper et al. (2000).

The short-run price elasticity is insignificant, as in the UK imports case and in Hooper et al. (2000). The long-run price elasticity is -0.40, which is close to the estimation of -0.31 from Hooper et al. (2000). In the UK case, the price elasticity for imports in the long run came to -0.95, suggesting that prices have a more deteriorating effect on the UK imports than on the US imports. The speed of adjustment coefficient in the US case is -0.12, which is close to the UK estimation of -0.11 and slightly higher than -0.02 from Hooper et al. (2000).

Structural Stability

When the ECM is checked for the structural stability, the Bai-Perron and Chow tests identify breakpoints in 1965Q2 and 1978Q3. When data are split in three sub-periods, it slightly improves the estimates (Table 2.18).

Table 2.18 Structural Stability Analysis: Imports (The US). Dependent Variable is ΔlnM_t

	1950Q2-65Q2	**1965Q2-78Q3**	**1978Q4-2018Q1**	**Whole Period**
ΔlnY_t	1.05**	0.81	2.02***	1.50***
$\Delta lnPM_t$	-0.85***	-0.17	0.05	-0.11
lnM_{t-1}	-0.54***	-0.71***	-0.13***	-0.12***
lnY_{t-1}	0.52***	1.60***	0.25***	0.22***
$lnPM_{t-1}$	-0.52***	-0.23***	-0.04**	-0.05***
LRGDP	0.96***	2.26***	1.98***	1.87**
LRRP	-0.95***	-0.33***	-0.33**	-0.40*
BIC	-3.62	-3.3	-4.97	-3.9
SER	0.03	0.04	0.02	0.03

Note: *** indicates a significance level of 1%, ** of 5%, and * of 10%. LRGDP and LRRP - long-run income and price elasticities, respectively. BIC - Bayesian information criterion; SER - standard error of regression.

The short-run price elasticity coefficient seems to be fluctuating around zero, and the long-run one increases from -0.95 in 1950Q2-1965Q2 to -0.33 in 1965Q2-1978Q3 and remains at this level till the end of the whole estimation period. Using dummy variables for 1965Q2 and 1978Q3 does not affect the significance of other coefficients in the model.

In addition, estimation of (2.10) showed that there was no substantial change in the income-US imports relationship after the financial crisis 2007-2008. Hooper et al. (2000) included dummy variables for the first two quarters of the 1969 and 1972 as well as for 1974Q2, to account for dock strikes and price controls. The Chow test confirms these breakpoints for our data. Nonetheless, splitting data at these breakpoints or using dummy variables for them does not improve the estimates.

Overall, in contrast to the UK case, the sum of the US exports and imports price elasticities is less than unity (in absolute values), meaning the Marshall-Lerner condition does not seem to hold for the US data. Hence, overall results on whether Marshall-Lerner condition holds are mixed.

2.11 Conclusion

Analysis of the income-trade relationship for a sample of 15 countries using a general ARDL model demonstrates that there are three main issues when applying a general-to-specific modelling approach and doing structural stability analysis on a number of countries with short time series. Firstly, there are various patterns for different countries, making it difficult to arrive to a common model for the whole sample. Secondly, the general models that were estimated are badly determined either because they include too many variables or there are many structural breaks in the data. Thirdly, the restricted models that seem to be chosen for the individual countries are often difficult to interpret.

Estimation of a simple ECM for real exports with the world GDP being the only explanatory variable produces more sensible results. Using this model we estimate the average income elasticity of exports in the long run to be 1.26% and 4.08% in the short run. However, the results on significance of the long-run effect of income on exports are mixed. While Wald test confirms significance of the long-run elasticity coefficients for all but two countries, the F-test suggests that the lagged variables might be jointly insignificant for more than half of the countries in the sample.

After comparing the estimates obtained using the WTO and IMF datasets and estimating the income-trade relationship (using imports) over different sub-periods, several results were established.

Firstly, once again there seems to be a relationship between trade and income in the short run and in the long run, especially when we estimate it over sufficiently long period. However, even if the estimation period is broken into two sub-periods, the significance of the long-run and short-run income elasticity coefficients is not shattered for the majority of countries in our sample.

Secondly, the structural stability analysis suggests different breakpoints for different countries. However, when choosing year 2000 as a common breakpoint, the coefficients for long-run and short-run income elasticities remain significant in pre-2000 and post-2000 sub-periods and are not dramatically different from the ones for the whole period. Hence, year 2000 seems to be an appropriate choice for a common breakpoint for these data.

Thirdly, when comparing the results obtained using the WTO and IMF databases, the values for short-run and long-run elasticities do not exactly match up, but on many occasions are comparable. The magnitude of income-trade elasticity is what remains to be determined, as the values vary depending on the chosen dataset as well as from country to country.

One unusual finding, which in fact is supported by the results presented in Constantinescu et al. (2015), is that in most cases the short-run coefficients are bigger than long-run ones. The logic behind this empirical evidence is unclear, as it runs against intuition, because normally it takes time for a dependent variable to adjust to a change in the explanatory one. One possible explanation may be an endogeneity issue, as world output and trade are co-determined (Constantinescu et al., 2015). It is unlikely that world GDP would be shifted by the exports of any particular country, but exports could be driving domestic output, which in turn might affect the global income.

Moreover, when estimating changing income-trade relationship, it might be necessary to consider not only structural, but also cyclical factors that might be behind the trade slowdown in 2000s. However, it is very difficult to separate the two unless working with very long data.

When we consider special cases of the UK and US, for which longer quarterly data are available, several conclusions are drawn. Firstly, the estimation of short-run and long-run income and price elasticities of the UK and US trade flows is sensitive to the model specifications. We find that the Marshall-Lerner condition seems to hold for the UK, but not for the US.

Secondly, even though structural stability tests detect breaks in our data, they are numerous and the addition of time dummy variables or breaking data in corresponding sub-periods neither substantially alters, nor improves ECM estimations over the whole period.

Thirdly, some estimates we obtain using ECM are similar to the ones found in the literature, such as to those from Hooper et al. (2000). However, there are some discrepancies too, which might be due to focus on different estimation periods and differences in the data used for estimations.

In addition, in Granger causality tests for the UK and US trade flows, both, exports and imports, are predicted by income and prices. However, prices (except the UK imports prices) are not substantially affected by either income or trade flows. Moreover, the G7 GDP is predicted by the US exports and prices, but not by those of the UK. This might be a reflection on the fact that the US is the largest economy in the G7 and, hence, has substantial effect on other G7 members and, thus, on G7 GDP.

To conclude, the results on the Marshall-Lerner condition are mixed as it seems to hold for the UK, but not for the US. Overall, there seems to be a significant relationship between income and trade in the short run and in the long run, but our estimates of the income elasticity of exports and imports are sensitive to the model specifications, sample size and estimation period.

Chapter 3

Balance of Payments Sustainability

3.1 Introduction

The initial analysis of the balance of trade components, exports and imports, produced mixed results. Nonetheless, we believe that a balance of payments (BOP) deficit cannot explode. Even though the exact requirements for a solvency or sustainability condition for the balance of payments are disputed, it generally assumes a country cannot run a Ponzi scheme, which involves borrowing more to pay the interest on earlier loans.

In this chapter we aim to determine whether a balance of payments deficit is sustainable in the long run by using simple reduced form autoregressive specifications since the analysis of exports and imports in the previous (second) chapter showed that more structural models do not seem to work very well. Hence, in this chapter we look at the feedback directly using the balance of trade and current account as proxies for the balance of payments. We perform this analysis using data from Jordà-Schularick-Taylor Macrohistory Database for 1870-2016 for seventeen countries, namely Australia, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Japan, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, the UK and US.

In the empirical literature solvency of the balance of payments or government debt is often investigated by testing for unit roots (Schoder, 2014; Schoder et al., 2013) or cointegration among the variables (Trehan and Walsh, 1991). However, Bohn (2005, 2007) argued that while stationarity of a process is a sufficient condition to satisfy the intertemporal budget constraint (IBC), which says that the initial stock of debt equals the expected present value of future balance of payments surpluses if the transversality condition (TC), that is the discounted future debt converges to zero, holds. Bohn however argued that the transversality condition is very weak by its design and can be satisfied regardless of the order of integration of data.

Bohn (2016) argued that whether the balance of payments deficit of any particular country is sustainable is in fact an economic question as the solvency depends on the rational expectations of the lenders. If they believe that a country can repay what it owes, then they are willing to lend more and as a result the economy remains solvent. Hence, Bohn called the refinancing, which is a crucial condition for solvency, 'a game between investors' with a rational expectations equilibrium (Bohn et al., 2016, p. 19).

Thus, whether a country's balance of payments deficit is sustainable cannot be answered empirically, as it is a question of beliefs. Therefore, in this chapter we aim to analyse whether the balance of payments is a stabilising process and if so, what are the patterns of the adjustment process. In order to do that we estimate our models for a sample of seventeen countries for 1870-2016 and also over three shorter sub-periods, 1870-1914, 1915-1950 and 1951-2016. We also consider a possibility that the pressure to adjust may be different on deficit and surplus countries, but because balance of payments should sum to zero over the world, reduced deficits must be matched by reduced surpluses elsewhere.

Looking at the large global economies, such as the UK and US, it is plausible to maintain large balance of payments deficits over many years. However, empirical literature suggests that when it comes to the Organisation for Economic Co-operation and Development (OECD) countries, current account imbalances appear to be small relative to levels of saving and investment when estimated over sufficiently long period. Nonetheless, the increasing current account deficit in the US, and to a lesser extent in the UK, has become the source of increasing concern regarding sustainability of the balance of payments.

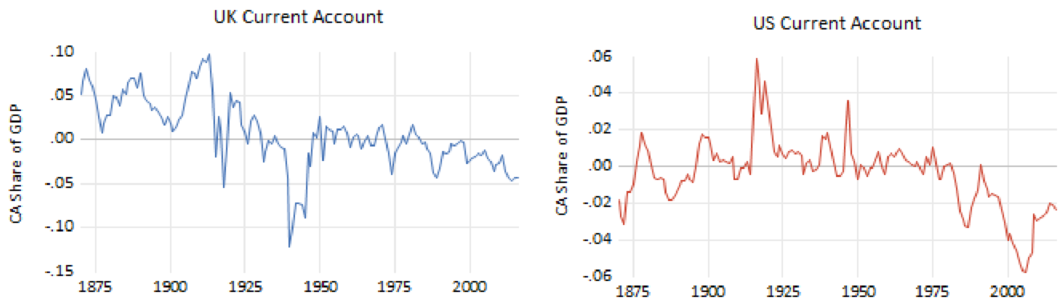
The outline is as follows, first section 3.2 reviews the literature. Section 3.3 describes the basic theory behind the estimations presented in this chapter. The data and descriptive statistics are presented in section 3.4. Section 3.5 describes the econometric approach adopted for the analysis of the solvency condition for the balance of payments. Section 3.6 provides the discussion of the estimates. In section 3.7 we discuss the possibility of the asymmetric adjustment when a country is running a deficit versus one that is running surplus. Section 3.8 contains some concluding comments.

3.2 Literature Review

There is a growing literature on the Great ratios, economic indicators that are supposed to be constant in the long run, such as government debt, employment rates and others. We expect the solvency to hold in the long run, but the question is how long is the long run? It might be a very long time indeed. Moreover, empirical papers on the topic found these ratios often behave like random walks at least over some periods of time, meaning solvency condition is not being satisfied. What are the reasons for that?

Balance of payments solvency is often analysed using data on the current account balance (change in a country's foreign net wealth, including unilateral transfers and investment incomes) and the balance of trade (net exports). Large current account deficit figures reported over last decade for countries such as the UK and US raise a question of solvency (Figure 3.1). Can an economy go on running a balance of payments deficit forever? If yes and the balance of payments is not stationary but instead is a random walk, that would imply that the current account deficit could increase forever and debt will eventually explode.

Fig. 3.1 UK and US Current Accounts (1870-2016)



Source: Jordà-Schularick-Taylor Macrohistory Database (2019).

The US however, might be considered to be somewhat of a special case. Gourinchas et al. (2019) analysed how international currencies are used for various purposes in the global economy and the diverse effects this has on individual economies through international monetary and financial systems. Focusing on the US case they argue that the US dollar hegemony is among factors that allow the country to run large external deficit as its liabilities are readily absorbed by the market. The securities issued by the hegemon, such as the US, are considered nominally safe and liquid, hence there is high demand for the US assets, such as government bonds, despite low interest rate payable on them. Also many countries have large portions of their exports and imports traded in US dollars not to mention the US dollar playing important role as a vehicle currency, being preferred by many investors worldwide for the financial transactions and investments as the US dollar is considered to be one of the safest currencies and is very liquid. All this allows the US to have the exorbitant privilege as it benefits from the valuation and trade channels of adjustment and being global banker and insurer at the same time. This position in the world economy is what ensures the US external solvency.

There are different explanations for why some countries, for example the US, run large trade deficits, while others, such as China, are subject to substantial trade surpluses. Two main sets of possible reasons include cross-country differences in the business cycle and differences in trade policies (Alessandria and Choi, 2019). For instance, Muellbauer and Murphy (1990) focused on the UK current account deficit and argued that it is a self-correcting mechanism with deficit arising from excessive private spending that is eventually corrected by the decumulation of foreign assets, which reduces private sector demand and eventually restores external balance. Meanwhile, focusing on the case of the US, Alessandria and Choi (2019) used two standard dynamic trade models and found that changing trade barriers have a strong influence on the trade balance, especially if it is measured as a GDP share. In addition, they considered the effect of trade policy. They analysed it by estimating the effect of an anti-dumping penalty that is modelled as a shock that leads to a ten percentage points difference between the costs of exports and imports. They found that trade policy in fact has its influence on trade balance mainly in the period of expansion of the trade flows. They explained, that the extent to which the US inward and outward trade barriers affect its balance of trade depends on the future policy which might or might not be designed to return these barriers to pre-Great Recession levels and, as a result, may cause a significant change in the US trade balance in the future.

We focus on the current account, but capital inflows might be important as well, and these are either portfolio or foreign investments. Caporale et al. (2017), using the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models for emerging Asian markets and a time-varying transition probability Markov-switching specification, found that there is a positive relationship between high exchange rate volatility and equity inflows from those markets towards the US (while bond inflows decrease exchange rate volatility). The uncertainty of the exchange rate is argued to reduce net equity and net bond flows with investors choosing to invest at home instead of abroad (Caporale et al., 2015a). Hence, it was suggested that credit controls and capital controls could be useful tools in achieving financial stability (Caporale et al., 2017, 2015a). In addition, Caporale et al. (2015a) found that the exchange rate uncertainty has different effect on the European Monetary Union members versus the non-members economies, and the effect is negative in the former case and positive in the latter. Hence, there might be substantial difference in the balance of payments and government deficits adjustment processes in euro and non-euro countries (this idea we explore in greater detail in the fourth chapter).

Bohn (2005, 2007) criticised a common approach of using stationarity and cointegration tests to analyse the solvency condition for the balance of payments as well as fiscal debt and deficit. He emphasised that these tests are aimed at checking whether the transversality condition holds, and the debt and deficit series satisfy the intertemporal budget constraint. However, Bohn argued that this condition for solvency is quite weak. The intertemporal budget constraint says that the initial debt equals to the expected present value of future surpluses (in case of balance of payments, that would mean current account surpluses) if the discounted future debt converges to zero, meaning if the transversality condition holds. The transversality condition is comprised of the discounting factor, $\frac{1}{(1+r)^n}$, and the expected value of the stock of debt, $E_t[b_{t+n}]$ (where b_{t+n} is a future debt-GDP ratio). Hence, the transversality condition can be presented as:

$$\lim_{n\to\infty} \frac{1}{(1+r)^n} E_t[b_{t+n}] = 0.$$

However, Bohn emphasised that the exponential decay of the discounting factor in the transversality condition dominates the polynomial growth of debt. Hence, the intertemporal budget constraint and transversality condition can be satisfied regardless of the order of integration of the data.

Durdu et al. (2013) adopted an error-correction reaction function outlined by Bohn in his 2007 paper to check whether the solvency condition holds for the net exports (NX) and net foreign asset (NFA) positions for the sample of 21 industrial countries (IC) and 29 emerging markets (EM) over the 1970-2006. In order to do so, they used Pesaran and Smith's (1995) mean group (MG) and Pesaran et al. (1999) pooled mean group (PMG) estimators. In addition, they also analysed the dynamic patterns of the process of adjustment in order to attain external solvency, including country specific characteristics, if any, that might affect patterns of adjustment. Using PMG estimators, they found that the dynamic NX and NFA relationship is the same for all countries in the long run, but there are some country-specific variations when it is estimated over shorter periods. They confirmed the negative relationship between the net exports and net foreign assets with one percentage point decrease in the later causing 0.07 percentage points increase in the former in the long run, confirming that both are stationary processes assuming a growth-adjusted real interest rate, (r-g), below 7%. The net exports response coefficient (to changes in the net foreign assets), however, vary with development. When the sample was split in two groups, industrial countries and emerging markets, it was found that the NX is less responsive to NFA movements in more developed countries, with the average long-run coefficient being -0.047 in the industrial countries and about twice higher in absolute value (-0.092) in the emerging economies. Moreover, it appears that countries with weaker fundamentals, such as poorer institutional quality and less-developed financial sector, have to adjust their NX positions more strongly to any variations in the NFA in order to maintain solvency. There is no substantial changes in the estimates when countries are divided into high and low leverage ones or when sample is split in more and less open economies based on their trade share of GDP. Hence, the authors concluded that solvency condition for NFA positions holds as long as any finite difference of the NFA is stationary. Their analysis also showed that the NFA series for all countries in the sample have low degree of integration.

Closely related to Durdu et al. (2013) is the paper by Schoder et al. (2013) on the current account imbalances of the European Monetary Union (EMU) countries. Their sample includes ten European countries that joined EMU from its foundation and four European stand-alone economies. They used the OECD Economic Outlook data for 1975Q2-2011Q2 and also opted for the MG (Pesaran and Smith, 1995) and PMG estimators (Pesaran et al., 1999). Adopting a theoretical framework of Bohn (1995; 1998), they argued that if the net exports response to a one-unit change in external debt is positive then the sufficient condition for solvency can be considered satisfied. This is what they call *TC sustainable* net external debt-GDP ratio. However, since this condition is rather weak, as it holds for any debt process integrated of finite order, they also analyse the persistence of the external imbalances and argued that if the net external debt-GDP process is mean-reverting, I(0), it is considered *operationally sustainable*. They also perform a non-parametric estimation that involves fitting the data to a simplified general additive model using a penalised spline estimator. They found that on average, considering all EMU and non-EMU countries over the whole period, the net external debt is consistent with the transversality condition and, hence, can be described as TC sustainable. After splitting their estimation period in two sub-periods, pre-euro, 1975Q2-1996Q4, and post-euro, 1997Q1-2011Q2, they found that the average long-run response coefficient for the EMU countries dropped and changed sign from 0.09 to -0.03, suggesting that the European economic integration might have impeded the adjustment of the net exports in EMU countries under consideration. The non-parametric estimation of the response coefficient suggests that the exchange rate mechanism is not important in the adjustment of trade imbalances in the Northern countries. However, the opposite is found to be true for the Southern countries, where the exchange rate adjustment is seen to be associated with external imbalances adjustments. Moving to the operational sustainability tests (stationarity tests), all countries failed it at 5% level over the whole period. When the data were split in the two sub-periods, for all EMU countries the unit root hypothesis is rejected on average over the first sub-period, but not after the introduction of euro. In the non-EMU countries external debt is on average operationally unsustainable over both sub-periods. Finally, the non-parametric analysis on the basis of the operational sustainability criteria suggests that the lower degree of exchange rate flexibility is associated with higher persistence of external debt imbalances. They argued that the introduction of a common currency is associated with slowing the current accounts adjustment in the EMU countries, and since its formation the current account imbalances appear to have become more persistent.

It appears that at least in the short run, trade deficit can increase sharply in response to rise in the international capital mobility. Considering the capital movements across national borders, the traditional economic theory suggests that rational investors choose to invest where the expected return is the highest. If that was true, the capital would flow from developed to developing countries where low levels of capital per worker translates into higher return on capital. Statistical data, however, show little movement of capital from rich to poor countries, and this gives rise to a puzzle first discussed by Lucas in his 1990 paper and since then known as Lucas Paradox. Furthermore, in condition of perfect capital mobility, meaning an absence of international financial markets regulations, domestic saving and investment rates should have low or zero correlation as investors are free to move their capital anywhere with highest return on capital. Nonetheless, in 1980 Feldstein and Horioka estimated the relationship between saving and investment rates being close to unity, suggesting that it can be interpreted as evidence of capital immobility with investors being more keen on investing at home. However, Coakley et al. (1996) argued that if the current account solvency condition holds, this correlation between shares of saving and investment in GDP comes from cointegration between the two variables in the long run and is not a sign of low mobility of capital.

Nell and Santos (2008) further analysed the long-run solvency constraint model proposed by Coakley et al. (1996) by estimating a long-run solvency constraint model for six OECD countries. The correlation between savings and investments tends to decrease during times of liberal capital control, the finding that essentially refutes the Feldstein-Horioka hypothesis. Some papers (De Vita and Abbott, 2002) find a cointegration between saving and investment rate even during periods of liberal capital movement, suggesting that saving-investment correlation is not very informative when it comes to measuring capital mobility. Nell and Santos argued that the causality between investment and saving rates is bidirectional and needs to be taken into account in order to confirm a solvency condition for the current account. They concluded that the Feldstein-Horioka approach is a useful but incomplete measure of sustainability of the current account deficit, at least when we consider the OECD countries.

Fostel et al. (2019) analysed the effect of differences in the ability to use an asset as a collateral on the volatility of capital flows between otherwise similarly-developed countries. They used a general equilibrium model with two countries, where a Home economy is selling assets that provide an investor with state-contingent promises and a Foreign country is offering assets which can only be used as a collateralised debt. The authors emphasised that this way of sharing collateral might be one of the incentives for capital flows between otherwise identical countries. Using a static model they analysed the effect of financial innovations on capital flows and found that as integration between two economies increases, so is the price difference between two assets, as a result of the gap in their collateral values. Then using a dynamic model they again confirmed that an increase in financial integration leads to asset price fluctuations as well as increase in the volatility of capital flows. Thus, they argued that a country can run a current account deficit if it has an ability to finance it by selling its assets that might be more attractive as it is in case of a Home economy in their settings. Since Home assets offer state-contingent promises, they are more attractive and, hence, more expensive than Foreign assets. Therefore, in their setting a Home country can run a current account deficit that is proportional to the gap in the collateral values. Thus, capital flows allow an economy to run a large deficit as long as it can be financed by the country earning sufficiently high intermediation rents.

Hence, one would assume that even though the balance of payments deficits or surpluses can occur over some periods of time, the current account is stationary in the long run. This chapter presents some support for this argument as well as for the findings from Coakley et al. (1996). Our analysis of the long data for seventeen countries demonstrates that the current account and balance of trade appear to be stablilising processes when we consider sufficiently long period. Hence, this chapter provides evidence for the balance of payments sustainability. It appears that the current account and balance of trade may diverge from the equilibrium for a period of time, but eventually they return to a stationary state. How long can they diverge from that state? This question is addressed by looking at the sub-periods estimations.

3.3 Basic Theory

For the theoretical framework we begin with Obstfeld and Rogoff's (2002) classic theory on intertemporal trade that is measured by the current account. This framework adopts Fisher's (1930) model of saving to a new setting of a small two-period economy that consumes one good. The model is then extended to an infinite-horizon one.

The individual that represents population maximises the following utility function:

$$U_t = u(C_t) + \beta u(C_{t+1}) + \beta^2 u(C_{t+2}) + ... + \beta^T u(C_{t+T}) = \sum_{s=t}^{t+T} \beta^{s-t} u(C_s), \qquad (3.1)$$

where U_t is utility, C_t is a consumption level, with the economy beginning on date t, while s is some future date.

Trade over time is represented by the current account, CA_t, which is equal to the change in the economy's net foreign assets (NFA). If B_t is the NFA at time t and B_{t+1} is the NFA at the end of the period t, then $CA_t = B_{t+1} - B_t$. This is also equal to exports, X_t, minus imports, M_t, plus income from abroad, $r_t B_t$ (where r_t is the interest earned (paid) on the foreign assets). Then the current account identity is given by:

$$CA_t = B_{t+1} - B_t = X_t - M_t + r_t B_t = BOT_t + r_t B_t, \tag{3.2}$$

where $BOT_t = X_t - M_t$, is the balance of trade.

Using the national income identity, we can also define the current account as

$$CA_t = B_{t+1} - B_t = Y_t + r_t B_t - C_t - G_t - I_t, \tag{3.3}$$

where G_t is the government expenditure and I_t is the new investment during period t with $I_t = K_{t+1} - K_t$, where K_t is the pre-existing capital, while K_{t+1} is the stock accumulated through the end of the period t (we ignore depreciation of capital).

From (3.3) we can see that the current account can also be defined as $CA_t = S_t - I_t$, where savings, S_t, are equal to income, $(Y_t + rB_t)$, minus expenditure, $(C_t + G_t)$. Even though in this chapter we focus on the current account definition as the balance of trade plus the rate of return on the foreign assets, (3.2), it is worth noting that the definition of the current account as the difference between a country's savings and investments may in fact help to explain a phenomenon known as the Feldstein-Horioka puzzle.

Feldstein and Horioka (1980) argued that if there was perfect capital mobility, investment could be financed from anywhere, so there should not be any relationship between national investment and national savings, but when they ran the cross-section regression to estimate the following equation:

$$\frac{S_t}{Y_t} = \alpha + \beta \frac{I_t}{Y_t} + \nu_t,$$

they found that β was not significantly different from one. This result they interpreted as the evidence for the lack of capital mobility, which became known as the Feldstein-Horioka puzzle. However, Coakley et al. (1996) argued that this result came from the current account solvency restriction on CA_t. If the current account is stationary, then national savings and investments, both integrated of the first order, are cointegrated with a unit coefficient irrespective of the degree of capital mobility. Hence, Feldstein and Horioka's (1980) result does not necessarily suggest a lack of capital mobility, but rather provides some support for the sustainability of the current account.

We can derive the traditional solvency condition for the current account starting from the current account identity, (3.3). If we rearrange the terms and run the identity forward, we arrive to the following constraint (Obstfeld and Rogoff, 2002, p. 60-61):

$$\sum_{s=t}^{t+T}\left(\frac{1}{1+r}\right)^{s-t}(C_s+I_s)+\left(\frac{1}{1+r}\right)^{T}B_{t+T+1}=(1+r)B_t+\sum_{s=t}^{t+T}\left(\frac{1}{1+r}\right)^{s-t}(Y_s-G_s). \tag{3.4}$$

If we assume that debt eventually always gets repaid and these resources are not left unused, then

$$B_{t+T+1}=0. \tag{3.5}$$

If we assume that (3.5) holds, then (3.4) simplifies to

$$\sum_{s=t}^{t+T}\left(\frac{1}{1+r}\right)^{s-t}(C_s+I_s)=(1+r)B_t+\sum_{s=t}^{t+T}\left(\frac{1}{1+r}\right)^{s-t}(Y_s-G_s). \tag{3.6}$$

Then, in the infinite-horizon case, the condition (3.5) becomes:

$$\lim_{T\to\infty}\left(\frac{1}{1+r}\right)^{T}B_{t+T+1}=0. \tag{3.7}$$

Assuming (3.7) holds, the infinite-horizon intertemporal budget constraint is:

$$\sum_{s=t}^{\infty}\left(\frac{1}{1+r}\right)^{s-t}(C_s+I_s)=(1+r)B_t+\sum_{s=t}^{\infty}\left(\frac{1}{1+r}\right)^{s-t}(Y_s-G_s). \tag{3.8}$$

With $T\to\infty$, the utility function (3.1) is generalised to

$$U_t=\lim_{T\to\infty}[u(C_t)+\beta u(C_{t+1})+\beta^2u(C_{t+2})+...]=\sum_{s=t}^{\infty}\beta^{s-t}u(C_s). \tag{3.9}$$

Using (3.3), we substitute for the consumption levels in (3.9) and with $T\to\infty$, the utility function becomes:

$$U_t=\sum_{s=t}^{\infty}\beta^{s-t}u[(1+r)B_s-B_{s+1}+A_sF(K_s)-(K_{s+1}-K_s)-G_s]. \tag{3.10}$$

Then, maximisation with respect to B_{s+1} and K_{s+1} yields two conditions that must hold for every period $s\geq t$ (including the infinite-horizon case):

$$u'(C_s)=(1+r)\beta u'(C_{s+1}), \tag{3.11}$$

which is the consumption Euler equation and:

$$A_{s+1}F'(K_{s+1}) = r, \tag{3.12}$$

which is the equality between the marginal product of capital and the world interest rate.

We can then determine the optimal level of consumption by combining the first order conditions (3.11) and (3.12) with the intertemporal budget constraint, (3.8). However, in this chapter we are focusing on analysing whether the balance of payments is stabilising rather than on what is the optimal level of consumption for an economy.

This is a more general version of the traditional economic theory of the current account solvency. For our analysis we use data for the current account and balance of trade. We can derive the intertemporal budget constrain and the transversality condition expressed in terms of the balance of trade and stock of debt (expressed in terms of net foreign assets) in a few simple steps, starting from the current account identity, (3.2). We can write future net foreign assets as

$$B_{t+1} = (1+r_t)B_t + X_t - M_t. \tag{3.13}$$

Then, following Bohn (2005), the future path of a country's stock of debt can be obtained by taking (3.13) n periods forwards and considering the arbitrary sequences of interest charges on the foreign liabilities and of the balance of trade surpluses,

$$B_{t+n} = \left(\prod_{k=0}^{n}(1+r_{t+k})\right)B_{t-1} + \sum_{j=0}^{n}\left(\prod_{k=j+1}^{n}(1+r_{t+k})\right)BOT_{t+j}. \tag{3.14}$$

Assuming that $b_t = \frac{B_t}{Y_t}$, $x_t = \frac{X_t}{Y_t}$, $m_t = \frac{M_t}{Y_t}$ and $bot_t = x_t - m_t$, then from (3.14) we can derive the intertemporal budget constraint and the transversality condition in three steps. First, we assume that rate of return, r_t, is constant and then take conditional expectations, $E_t[\cdot]$, of both sides:

$$E_t[b_{t+n}] = (1+r)^n b_t^* + \sum_{j=0}^{n}(1+r)^{n-j}E_t[bot_{t+j}], \tag{3.15}$$

where $b_t^* = (1+r_t)b_{t-1}$ is the stock of the foreign debt at the beginning of period t.

Then, after dividing both sides by $(1+r)^n$ and rearranging the terms we get:

$$b_t^* = -\sum_{j=0}^{n}\frac{1}{(1+r)^j}E_t[bot_{t+j}] + \frac{1}{(1+r)^n}E_t[b_{t+n}]. \tag{3.16}$$

Finally, assuming the discounted sum converges, the infinite-horizon version of (3.16) is:

$$b_t^* = -\sum_{j=0}^{\infty} \frac{1}{(1+r)^j} E_t[bot_{t+j}] + \lim_{n\to\infty} \frac{1}{(1+r)^n} E_t[b_{t+n}]. \tag{3.17}$$

Hence, according to the traditional theory, the balance of payments is considered solvent if the following intertemporal budget constrain is satisfied:

$$b_t^* = -\sum_{j=0}^{\infty} \frac{1}{(1+r)^j} E_t[bot_{t+j}], \qquad \text{[IBC]} \quad (3.18)$$

which is subject to the following transversality condition:

$$\lim_{n\to\infty} \frac{1}{(1+r)^n} E_t[b_{t+n}] = 0. \qquad \text{[TC]} \quad (3.19)$$

In (3.17) the external solvency implies an inverse relationship between balance of trade (net exports) and net foreign assets. For instance, if a country's net foreign assets are decreasing (meaning, its foreign liabilities are increasing relative to its foreign assets), the country has to increase its net exports in order to cover the interest payments on the increasing foreign liabilities. The systemic corrective actions of this sort will in theory ensure the balance of payments solvency.

Nonetheless, Bohn (2005, 2007) argued that the most credible evidence in favour of sustainability is the robust positive response of surpluses in response to the increase in the stock of debt. Bohn (2007) developed his theoretical framework considering fiscal policy sustainability. However, Durdu et al. (2013) presented a version of Bohn's (2007) model for the analysis of the balance of payments sustainability:

$$NX_t - \rho NFA_{t-1} = z_t \sim I(m),$$

for some $\rho < 0$, such that $|\rho| \in (0, 1+r]$, the interest rate is constant ($r_t = r$) and the process, z_t, can be integrated of any order m. NX_t is the net exports.

Thus, Bohn (2007) proposed to check for sustainability using an error-correction specification with surpluses and debt instead of using standard time-series tests, such as unit root and cointegration tests, that are commonly used to check whether the transversality condition holds and the intertemporal budget constraint is satisfied. As was mentioned before, Bohn (2005, 2007) argued that the transversality condition is a quite weak condition because while it ensures that the debt will eventually get repaid, it does not set any time limit as to when this will happen.

For the balance of payments sustainability analysis, the data for the stock of debt (net foreign assets) are not available for the long span covered in this chapter. Therefore, we analyse the balance of payments sustainability in terms of autoregressive feedback mechanism using the balance of trade and current account data. If these processes are stabilising on their own lagged values in a self-corrective way, we assume that the balance of payments remains sustainable. Before we move to a more detailed description of the empirical models used in this chapter, we first discuss the data.

3.4 Data

We check sustainability of the balance of payments using data for the balance of trade and current account from Jordà-Schularick-Taylor Macrohistory Database. Our sample includes seventeen countries, namely Australia, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Japan, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, the UK and US, and the estimation period is 1870-2016.

One complication with the JST data is that it come from various sources that use different definitions. For instance, data for Australia for the current account alone are from three sources: 1870-1945 from Jones and Obstfeld (1997), 1946-1979 from Mitchell (2013) and 1980-2016 is from International Monetary Fund. Jones and Obstfeld's (1997) measure of the current account excludes gold flows, net exports gold and change in the monetary gold stock. Meanwhile, Mitchell (2013) defines the current account as the overall current balance which equals to national accounts minus the balance of payments and is in level form. Finally, IMF current account data are in percentage of GDP times the nominal GDP from the JST dataset. In comparison, current account data for Belgium for 1870-1946, 1949, 1952 and 1953 come from Mitchell (1980) and the current account balance is set to be equal to the trade balance. Current account data for Canada for 1948-2016 come from the IMF, but for Canada, Germany, Portugal, Switzerland and the US it is defined slightly differently than for other countries (as a net value that excludes exceptional financing). Data for Finland are reported in old Finnish Markka, thus, was divided by 100 to redenominate into new one. The current account data for 1980-2016 for Finland and France come from the IMF World Economic Outlook Database and reported as percentage of GDP, then multiplied with the nominal GDP from the same source. The current account data for Germany for 1975-2016 have been redominated into Deutsch mark using exchange rate from the JST dataset. The data for Italy are in the US dollars and, as for Australia, Finland, Germany and Norway, the current account is defined as the overall current balance. For the Netherlands, the current account data for 1870-1913 are from Smits et al. (2000), where it is the sum of net merchandise exports, net service exports and net primary incomes, reported in level form. For Portugal the data for current account for 1870-1947 are from Mitchell (2013) and calculated as exports minus imports. In case of the UK, the current account data for 1870-2007 are from Hills et al. (2015), and the rest, as with most countries, are from the IMF.

The definitions for the imports and exports data also vary. For example, in case of Australia, for 1948-2017 data come from the IMF with trade flows reported as goods value of exports and imports which calculated as international transactions minus external trade. The trade data for Belgium are also based on values of exports and imports of goods only. Data for Germany from 1914 to 1924 are missing and even though the data for 1920-1923 are available from Mitchell (2013), the numbers are inconsistent with high inflation and, therefore, data from this source are only taken for 1924-1943. Italian data are from two sources and again, as for many other countries, such as Canada, Denmark, France, Germany, the Netherlands, Norway and Switzerland, for 1948-2016 it is from the IMF with exports and imports being defined as merchandise trade in national currency. Moreover, in case of Italy for 1870-2011 the trade data are from Baffigi (2011), and are chained-lined while for the remaining years these data are in the local currency and level form. In contrast to the majority of other countries in the sample, for Portugal trade data over most of the estimation period (1954-onwards) are exports and imports of not only goods, but also services. The data are in current prices and domestic currency. Another similar exception is the UK, for which the data for 1950-2016 are from the IMF and include exports and imports of services, not just goods. The data are in nominal term and seasonally adjusted.

In terms of definitions and units, the data for the nominal GDP are quite consistent from country to country. For later years, mid-20th century onwards, for vast majority of countries the data come from the same few sources, such as the IMF (various years), the Maddison (2013) and Mitchell (1980, 2013). However, there are some differences in the GDP data for the countries in our sample. For example, for Australia for the whole period (1870-2016) the data come from Hutchinson and Ploeckl (2012) with the nominal GDP being defined as the value of production at current market prices, measured in millions of US dollars and reported in level form. In comparison, the GDP data for Denmark come from three sources and are defined slightly differently. For example, data for 1950-2002 are from the IMF and the nominal GDP is in billions of local currency, not in the US dollars. For Switzerland, the nominal GDP for 1879-1889 is from the University of Zürich, and it is a chain linked measure in millions of Swiss frank. For the UK, the original data for 1948-2017 are GDP at current prices, seasonally adjusted and expressed in millions of British pounds. The figures in JST dataset are then converted into millions of local currency for Australia, Belgium, Finland, the Netherlands, Norway, Portugal, Spain, Sweden and Switzerland, into billions of local currency for Canada, Denmark, France, Germany, Italy, the UK and US and into trillions of Japanese yen for Japan.

Needless to say that numerous data sources for each variable for every country and, as a result, differences in definitions and units of measure as well as some missing values, might bring some discrepancies to the analysis.

The data for exports, imports and GDP are available for 1870-2016 for all seventeen countries. However, the current account data are more scarce. For instance, for Switzerland these data are only available from 1920. For Germany the current account data are missing from 1914 to 1924 and then from 1939 to 1947. Similar situation is for other countries, such as Spain, France, the Netherlands and Norway.

Alternatively, balance of trade can be used as a proxy for the balance of payments. Data for the current account, GDP, exports and imports come in nominal values and are expressed in local currency. Exports and imports shares are calculated by dividing exports and imports by GDP, respectively. Using these new variables, the balance of trade is defined as the difference between exports and imports shares. By the analogy with the trade flows shares, the current account share values are obtained by dividing the current account by the nominal national income.

Since raw data are expressed in local currency, trade flows, current account and income data are not directly comparable among the countries. Hence, we use means and standard deviations for exports (x), imports (m), balance of trade (bot) and current account (ca), all expressed as shares of GDP (Table 3.1).

Highest exports and imports shares are in Spain, over half of the national GDP (0.56 and 0.66, respectively) and in Belgium (0.40 and 0.43, correspondingly). Both countries import more than export and, as a result, on average, are running balance of trade deficit. In case of Belgium, balance of trade and current account deficits are equal (-0.02). In contrast, in Spain while the economy is running the largest in the sample balance of trade deficit (-0.09), it appears to have an average current account surplus of 0.04. While the balance of trade is the largest component of the current account, the current account encompasses other components as well, such as factor income and financial transfers. In case of Spain these additional components seem to keep the current account in surplus, despite imports being, on average, higher than exports. It is, however, somewhat unusual that Spain has the highest imports and exports shares in the sample. In general, the bigger the economy, the lower are the exports and imports shares. As can be seen from the Figure 3.2, in case of Spain the shares have been almost constantly increasing ever since the end of World War II, rising from approximately 0.02 in the late 1940s to about a third of the GDP in 2000s.

Table 3.1 Descriptive Statistics (1870-2016)

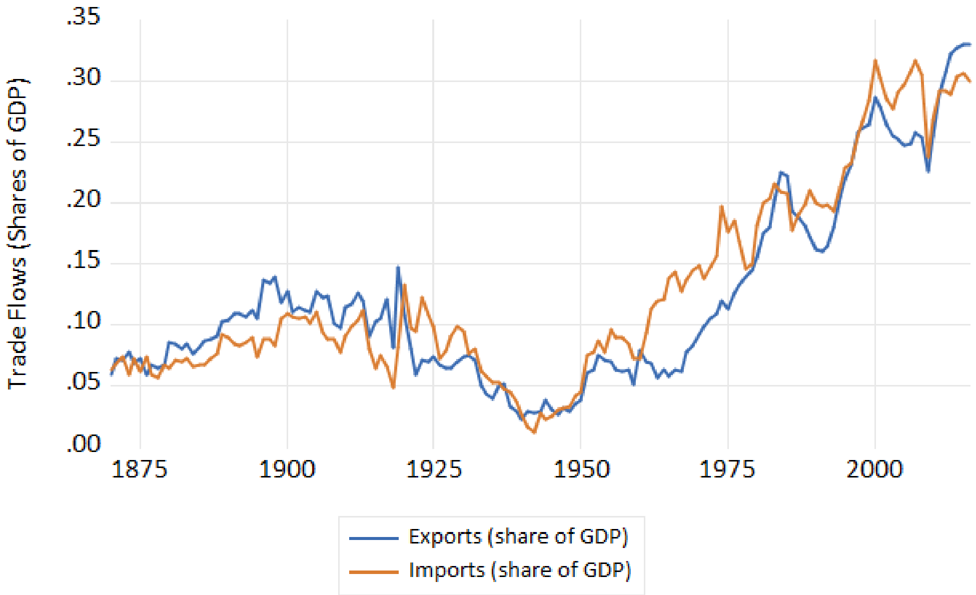
Country	x	m	bot	ca
Australia	0.15	0.14	0.00	-0.03
	0.03	0.03	0.03	0.04
Belgium	0.40	0.43	-0.02	-0.02
	0.22	0.19	0.06	0.05
Canada	0.21	0.20	0.01	-0.03
	0.06	0.05	0.04	0.04
Denmark	0.23	0.27	-0.04	0.03
	0.07	0.08	0.04	0.04
Finland	0.19	0.18	0.01	0.02
	0.08	0.07	0.04	0.03
France	0.23	0.27	-0.03	-0.01
	0.05	0.05	0.05	0.03
Germany	0.12	0.13	-0.01	-0.01
	0.08	0.08	0.03	0.02
Italy	0.21	0.23	-0.02	-0.02
	0.06	0.05	0.05	0.05
Japan	0.14	0.16	-0.03	0.01
	0.04	0.06	0.05	0.03
Netherlands	0.18	0.22	-0.04	0.01
	0.05	0.05	0.03	0.04
Norway	0.14	0.16	-0.02	0.00
	0.07	0.06	0.04	0.03
Portugal	0.10	0.11	0.00	0.01
	0.04	0.04	0.02	0.03
Spain	0.56	0.66	-0.09	0.04
	0.31	0.38	0.11	0.05
Sweden	0.20	0.25	-0.05	0.00
	0.06	0.05	0.09	0.06
Switzerland	0.13	0.19	-0.06	-0.04
	0.11	0.12	0.04	0.05
UK	0.20	0.21	-0.01	0.00
	0.07	0.05	0.03	0.03
US	0.07	0.07	0.00	0.00
	0.03	0.04	0.02	0.02

Notes: *x* and *m* are exports and imports shares respectively; *bot* and *ca* are the balance of trade and the current account shares of GDP, correspondingly. For each country: first row - mean, second row - standard deviation.

The lowest exports and imports shares are reported in the US (both are 0.07) and Portugal (0.10 and 0.11, respectively), with both countries having nearly zero balance of trade and current account. This is unexpected for the US, that has been consistently running balance of payments deficit since early 1990s.

Overall, 12 out of 17 countries (Belgium, Denmark, France, Germany, Italy, Japan, the Netherlands, Norway, Spain, Sweden, Switzerland and the UK) are running an average balance of trade deficit, which is the highest in Spain and Switzerland. In this sample net exports in Australia, Portugal and the US are closest to zero. Norway, Sweden, the UK and US have average current account shares of zero over this period, with Japan, the Netherlands

Fig. 3.2 Spain: Trade Shares (1870-2016)



Source: Jordà-Schularick-Taylor Macrohistory Database (2019).

and Portugal having a small average surplus of 0.01. Other countries in the sample, with exception of Denmark, Finland and Spain, on average, have current account deficits.

3.5 Empirical Models

Building on the theoretical framework presented in section (3.3), the outline of the empirical analysis performed in this chapter is presented below. Working with the panel data, we first estimate simple bivariate regression:

$$m_{it} = \alpha + \beta x_{it} + \varepsilon_{it}, \tag{3.20}$$

where m_{it} and x_{it} are imports and exports shares of GDP, correspondingly, and ε_{it} is the error term.

We estimate (3.20) for the whole period, 1870-2016, for pooled data with fixed cross-country effects. Then we present heterogeneous estimates for the whole period and for three sub-periods, 1870-1914, 1915-1950 and 1951-2016.

In fact it is not clear whether x_{it} or m_{it} should be used as a dependent variable when analysing the relationship between exports and imports shares of GDP using bivariate regression. We address this issue by also estimating the VECM for x_{it} and m_{it} which treats them symmetrically. We estimate the following VECM (1,1) model:

$$\Delta m_{it} = \alpha_m + \beta(m_{it-1} - \alpha x_{it-1}) + \varepsilon_{it}^m \tag{3.21}$$

$$\Delta x_{it} = \alpha_x + \gamma(m_{it-1} + \alpha x_{it-1}) + \varepsilon_{it}^x, \tag{3.22}$$

where m_{it} and x_{it} are import and export shares of GDP and x_{t-1} with m_{t-1} are their respective lagged values, ε_{it}^m and ε_{it}^x are the error terms.

Assuming that

$$(m_{it} - x_{it}) \sim stationary,$$

then in $(m_{it} + \alpha x_{it})$ we expect $\alpha = -1$. In addition, we expect $\beta < 0$ and $\gamma > 0$.

When we estimate a long-run imports-exports relationship using VECM, there are two parts to consider. First, if this analysis of x_{it} and m_{it} shows that their long-run relationship equals unity (in absolute value), then we can conclude that exports and imports balance to maintain a stable balance of payments. Second, a cointegration between the two variables can be confirmed if at least one of β and γ coefficients in (3.21) and (3.22) are not equal to zero.

In addition, we estimate a VECM with shares imposing the restriction $\alpha = -1$ to check what effect it will have on the feedback. Starting from the equations (3.21) and (3.22), if we subtract the later from the former:

$$\Delta m_{it} - \Delta x_{it} = (\alpha_m - \alpha_x) + (\beta - \gamma)(m_{it-1} - x_{it-1}) + (\varepsilon_{it}^m - \varepsilon_{it}^x),$$

we get a restricted VECM:

$$\Delta m_{it} - \Delta x_{it} = a + b(m_{it-1} - x_{it-1}) + \varepsilon_{it}, \tag{3.23}$$

where $a = (\alpha_m - \alpha_x)$, $b = (\beta - \gamma)$ and $\varepsilon_{it} = (\varepsilon_{it}^m - \varepsilon_{it}^x)$.

In addition, we try to estimate a VECM with imports and exports in logarithmic form instead of shares, to see whether it improves the estimation of the relationship under consideration. This version of the VECM can be presented as:

$$\Delta lnM_{it} = \alpha + \beta(lnM_{it-1} - \alpha lnX_{it-1}) + \varepsilon_{it}^M \tag{3.24}$$

$$\Delta lnX_{it} = \alpha + \gamma(lnM_{it-1} - \alpha lnX_{it-1}) + \varepsilon_{it}^X, \tag{3.25}$$

where lnM_{it} and lnX_{it} are logarithms of imports and exports, and ε_{it} is an error term.

If (3.24) and (3.25) produce sensible results, we also try to estimate the difference between logarithms of exports and imports as in equation below

$$\Delta(lnX_{it} - lnM_{it}) = a + b(lnX_{it-1} - M_{it-1}) + \varepsilon_{it}, \tag{3.26}$$

where $a = (\alpha_m - \alpha_x)$, $b = (\beta - \gamma)$ and $\varepsilon_{it} = (\varepsilon_{it}^M - \varepsilon_{it}^X)$.

We try different forms of similar equations for exports and imports relationship because one apparent issue with analysis of this relationship, is that it is not clear what is the right functional form and disaggregation. Possible options include having trade flows expressed in levels:

$$M_{it} = \alpha + \beta X_{it},$$

in which a one unit change in exports causes a β units change in imports.

Alternatively, we can use imports and exports as shares:

$$\frac{M_{it}}{Y_{it}} = \alpha + \beta \frac{X_{it}}{Y_{it}},$$

where a one percentage point increase in the share of exports increases the share of imports by β percentage points.

Most of the theory is expressed in terms of shares of GDP, but the relationship may also be proportional, therefore we also consider logarithmic form of both trade flows:

$$lnM_{it} = \alpha + \beta lnX_{it},$$

where a one percent increase in exports causes a β percent increase in imports.

After examining the relationship between trade flows shares, we move to estimation of an OLS model for the balance of trade and current account, both expressed as GDP shares. The purpose of this exercise is to determine whether these proxies for the balance of payments are stationary.

Hence, starting from the balance of trade, we estimate the equation:

$$\Delta bot_{it} = \alpha + \beta bot_{it-1} + \varepsilon_{it}, \tag{3.27}$$

where bot_{it} is a share of the balance of trade (bot_{it} = (nominal exports/nominal GDP) - (nominal imports/nominal GDP)). If $\beta < 0$ in (3.27), then the balance of trade share is stationary and the sustainability condition for the balance of payments is satisfied.

The other proxy we are using to check the balance of payments solvency condition is current account share of GDP. Similarly to the balance of trade share estimation we analyse the following model:

$$\Delta ca_{it} = \alpha + \beta ca_{it-1} + \varepsilon_{it}, \tag{3.28}$$

where ca_{it} is the current account share of GDP. As with the balance of trade share estimation, we can conclude that if $\beta < 0$, then the current account share is stationary, thus the balance of payments solvency condition is satisfied.

Finally we check an assumption that the pressure to adjust on the countries running a balance of payments deficit might be stronger than on those running a surplus. Hence, leading to an asymmetric adjustment. We check for this by regressing the change in ca_{it} on ca_{it-1} and $capos_{it-1}$, a new variable which is the product of ca_{it} and a dummy *cadum* that equals one when ca_{it} is positive. Hence, we augment (3.28) by including this new variable and estimate the resulting equation stated below:

$$\Delta ca_{it} = \alpha + \beta ca_{it-1} + \gamma capos_{it-1} + \varepsilon_{it}. \tag{3.29}$$

If in (3.29) β coefficient (speed of adjustment when the current account is negative) is different from $(\beta + \gamma)$ (speed of adjustment when the current account is positive), then we can conclude that there is evidence of the asymmetric adjustment. All results are presented and discussed in the following section.

3.6 Results

We first estimate each of the equations for the whole sample and report homogeneous fixed effect panel estimates for the whole period. Then we estimate each of the equations for individual countries for the whole period and for three sub-periods.

3.6.1 Exports and Imports

Static Model (Exports and Imports Shares)

Starting with a basic model for exports and imports shares, we estimate (3.20):

$$m_{it} = \alpha + \beta x_{it} + \varepsilon_{it}$$

and obtain the following results (with standard errors in parentheses) for the pooled regression:

$$\begin{array}{lllllll} m_t & = & 0.029 & + & 0.972\, x_t & + & \varepsilon_t \qquad R^2 = 0.916 \\ & & (0.002) & & (0.010) & & \qquad SER^2 = 0.051 \end{array}$$

All coefficients are significant. It appears that the relationship between exports and imports shares is close to unity. When we estimate (3.20) for each country individually (an appendix, Table B.1) there is still a strong positive relationship between the trade flows shares (except for Norway β is negative), with the response coefficient being significant for all countries in the sample when the relationship is estimated over the whole period. However, the imports-exports relationship is not as close to unity as the pooled data estimate suggests.

There is still strong evidence of the significant effect of exports on imports when we estimate (3.20) over three sub-periods. In all but five countries the effect is significant in the first sub-period, and the significant coefficients range from 0.43 (the UK) to 1.29 (Belgium). In the second and third sub-periods all coefficients are significant (except Portugal for 1915-1950, and for 1951-2016 β is insignificant for Denmark and negative for Norway). The significant coefficients range from 0.31 (the US) to 1.13 (Italy) in the second sub-period and from 0.11 (Denmark) to 1.48 (the US) in the third sub-period.

Overall, Norway appears to be an exception in our sample, and the reason for mixed results for this country might be that the oil production has been large in Norway relative to its economy since the late 20th century. For oil producers the balance of payments constraint is different. Many of them just spend their export earnings and imports go up to match them. However, Norway has a Sovereign Wealth fund that breaks the link.

The static model analysis showed that it is somewhat easier to obtain robust and consistent with economic theory results when we average over the whole sample and long period, than when we focus on individual countries and shorter sub-periods.

VECM (Exports and Imports Shares)

We then estimate (3.21) and (3.22), a VECM(1,1) for imports and exports shares with an intercept (no trend) in the cointegrating equation:

$$\Delta m_{it} = \alpha_m + \beta(m_{it-1} - \alpha x_{it-1}) + \varepsilon^m_{it}$$

$$\Delta x_{it} = \alpha_x + \gamma(m_{it-1} + \alpha x_{it-1}) + \varepsilon^x_{it}.$$

The above equations, (3.21) and (3.22), outline the basic adjustment model we are interested in, but for estimation we also add some dynamics to it by including lagged changes in imports and exports shares. For the whole sample with the model estimated over the whole period the estimates are:

$$
\begin{array}{lcllllllllllllll}
\Delta m_t & = & 0.001 & - & 0.039\,(m_{t-1} & - & 1.111\,x_{t-1}) & - & 0.001 & - & 0.070\,\Delta\,m_{t-1} & + & 0.056\,\Delta x_{t-1} & + & \varepsilon_t^m & R^2 = 0.013 \\
 & & (0.001) & & (0.010) & & (0.037) & & & & (0.024) & & (0.029) & & & SER^2 = 0.026 \\
 & & & & & & & & & & & & & & & \\
\Delta x_t & = & 0.002 & + & 0.045\,(m_{t-1} & - & 1.111\,x_{t-1}) & - & 0.001 & + & 0.022\,\Delta\,m_{t-1} & + & 0.031\,\Delta x_{t-1} & + & \varepsilon_t^x & R^2 = 0.016 \\
 & & (0.001) & & (0.008) & & (0.037) & & & & (0.020) & & (0.024) & & & SER^2 = 0.022
\end{array}
$$

It appears that the speed of adjustment is significant in both equations. Change in exports shares lagged one period does not seem to have a significant effect on imports shares. However, previous-year imports shares seem to have effect on the change in m_t current value. The long-run relationship between trade shares is close to unity, confirming results obtained using the static model for pooled data. The speed of adjustment is -0.04 percentage points for imports shares, while exports shares loading coefficient is positive and slightly higher in the absolute value, 0.05 percentage points. Both feedback coefficients are relatively small, but neither appears to be exogenous, meaning both respond significantly to the disequilibrium.

The heterogeneous results (an appendix, Table B.2) are mixed. The α coefficients appear to be very dispersed across the sample for the whole period, suggesting it is not being estimated well. However, in most cases α has the correct sign, the only exception is Norway. The average α is -0.76 percentage points, but is further away from unity for sub-periods.

The β coefficient is significant in all but six countries and negative, as expected, for all economies (for which β is significant). The average imports speed of adjustment is -0.14 percentage points, and the sub-period average did not change as much as it was in case of α coefficient. The γ coefficient is significant and has the right (expected) sign in six out of 17 countries and the average speed of adjustment of exports to equilibrium is about about 0.02 percentage points.

Restricted VECM (Exports and Imports Shares)

Since the static model and VECM(1,1) estimates for pooled data suggest a close to unity relationship between exports and imports, but heterogeneous results are mixed, we now estimate the VECM(1,1) with an imposed restriction of $\alpha = -1$ (equation (3.23)):

$$\Delta m_{it} - \Delta x_{it} = a + b(m_{it-1} - x_{it-1}) + \varepsilon_{it}.$$

Below are the restricted VECM(1,1) results for pooled data for the whole period:

$$
\begin{array}{lcllllllllllllll}
\Delta m_t & = & 0.002 & - & 0.039\,(m_{t-1} & - & x_{t-1}) & - & 0.006 & - & 0.146\,\Delta\,m_{t-1} & + & 0.107\,\Delta x_{t-1} & + & \varepsilon_t^m & R^2 = 0.023 \\
 & & (0.001) & & (0.014) & & & & & & (0.039) & & (0.043) & & & SER^2 = 0.021 \\
\Delta x_t & = & 0.002 & + & 0.028\,(m_{t-1} & - & x_{t-1}) & - & 0.006 & - & 0.097\,\Delta\,m_{t-1} & + & 0.101\,\Delta x_{t-1} & + & \varepsilon_t^m & R^2 = 0.010 \\
 & & (0.001) & & (0.013) & & & & & & (0.036) & & (0.040) & & & SER^2 = 0.019
\end{array}
$$

The speed of adjustment coefficient for exports decreases from 0.05 to 0.03 when the restriction is applied, but for imports remains about the same (-0.04). However, both speed of adjustment coefficients are now close to the edge of significance.

The heterogeneous results (an appendix, Table B.3) for the whole period suggest the β coefficient is significant in nine countries in restricted VECM and in 11 in the VECM without the restriction, while the γ coefficient is significant in five in the VECM with restriction and in eight countries in the unrestricted VECM. There are even fewer significant coefficients when we estimate the restricted VECM over shorter sub-periods.

Overall, it appears that a VECM with this restriction on α does not fit the data well as more β and γ coefficients become insignificant when the relationship is estimated over the whole period and over three sub-periods comparing to the estimates for the VECM without restriction. It is especially noticeable in the second and third sub-periods estimations for restricted VECM, where β and γ become insignificant for most countries.

VECM (Logarithms of Exports and Imports)

In addition, we estimate a VECM(1,1) for the logarithms of imports and exports expressed in nominal values and local currency, (3.24) and (3.25):

$$\Delta lnM_{it} = \alpha + \beta(lnM_{it-1} - \alpha lnX_{it-1}) + \varepsilon_{it}^{M}$$

$$\Delta lnX_{it} = \alpha + \gamma(lnM_{it-1} - \alpha lnX_{it-1}) + \varepsilon_{it}^{X},$$

to check whether solvency condition will hold if the trade flows are in level form rather than expressed as shares of GDP.

The results for pooled data are:

$$\underset{}{\Delta lnM_t} = \underset{(0.014)}{0.044} - \underset{(0.067)}{0.274}\,(lnM_{t-1} - \underset{(0.014)}{1.029}\,lnX_{t-1}) + 0.237 + \underset{(0.081)}{0.023}\,\Delta lnM_{t-1} + \underset{(0.105)}{0.282}\,\Delta lnX_{t-1} + \varepsilon_t^m \qquad R^2 = 0.236,\ SER^2 = 0.149$$

$$\Delta lnX_t = \underset{(0.013)}{0.063} + \underset{(0.063)}{0.043}\,(lnM_{t-1} - \underset{(0.014)}{1.029}\,lnX_{t-1}) + 0.237 + \underset{(0.076)}{0.015}\,\Delta lnM_{t-1} - \underset{(0.098)}{0.025}\,\Delta lnX_{t-1} + \varepsilon_t^x \qquad R^2 = 0.007,\ SER^2 = 0.140$$

The long-run coefficient appears to be even closer to unity comparing to the VECM(1,1) with imports and exports shares. Meanwhile, logarithm of exports is exogenous now, as it does not seem to significantly respond to the disequilibrium. Speeds of adjustments, however, have not changed dramatically, the imports one is still close to -0.03, while the exports loading coefficient is about 0.04. SER is nonetheless higher for VECM(1,1) with logarithms, comparing to the one with shares.

The heterogeneous estimates (an appendix, Table B.4) suggest that α coefficient is significant and close to unity not only when the model is estimated over the whole period, but also if we estimate it over the three sub-periods. Over the whole period, only in case of the Netherlands there is no stabilisation feedback. In seven countries the model is stabilised thorough adjustment of imports and in 11 it is so through exports.

For the sub-periods, β coefficient is significant in eleven countries when the model is estimated over 1870-1914, in five economies over 1915-1950 and in six in the last sub-period. Meanwhile γ is significant only in two in the first sub-period, in nine countries in 1915-1950 and in eight in 1951-2016. In addition, we find that the average α is not significantly different from -1 when the model is estimated over three sub-periods.

In VECM(1,1) with logarithms (comparing to the VECM(1,1) with shares) more coefficients are significant, but t-statistics for these estimates seem to be unrealistic in some cases. For instance, for the α coefficient for Germany with t-statistic of -871.6 when the model is estimated over the whole period.

One possible reason for this is a strong upward trend in exports and imports data, which makes the estimates very precise as trade flows track very closely and this can potentially cause unrealistic t-statistic values in many cases. However, a significant deterministic trend in the VECM for imports and exports would mean that the debt can expand indefinitely. Nonetheless, when we try to include deterministic trend in the model, its coefficient is insignificant. Hence, we conclude that it is not fixed deterministic trend that is causing the issue. In fact, we find that the variables under consideration have large variances, and the standard errors reflect this as higher variance leads to smaller standard errors and bigger t-statistics.

We also try to estimate the model for change in logarithms of exports and imports (3.26):

$$\Delta(lnX_{it} - lnM_{it}) = a + b(lnX_{it-1} - lnM_{it-1}) + \varepsilon_{it}$$

and obtain the following results for the whole sample estimated over the whole period:

$$\begin{array}{lllllll} \Delta(lnX_t - lnM_t) & = & -0.022 & - & 0.218\,(lnX_{t-1} - lnM_{t-1}) & + \; \varepsilon_t & R^2 = 0.112 \\ & & (0.004) & & (0.013) & & SER^2 = 0.160 \end{array}$$

In the regression for the whole sample all coefficients are significant and β has correct (negative) sign, meaning that if trade balance (in this case proxied by $(lnX_{t-1} - lnM_{t-1})$) increased in the previous period, the current net exports are adjusted downwards, resulting in a stabilising process.

The heterogeneous estimates (an appendix, Table B.5) suggest that over the whole period the balance of trade is stationary in 13 out of 17 countries when (3.26) is estimated over the whole period. Over the first and second sub-periods the balance of payments is stabilising in seven countries and over the third one – in six. Since (3.26) is an example of an autoregressive model, we check the significance using the Augmented Dickey-Fuller (ADF) critical value of -2.9 (5% significance level; a model with a constant and no trend; Fuller, 2009), because the variables under consideration are unlikely to be trend stationary.

Overall, there does not seem to be a major difference between using shares or logarithms when analysing the balance of payments sustainability using exports and imports series. It also appears that there is more evidence of stabilisation of the balance of payments, and the estimates are more consistent with economic theory, when models are estimated for pooled data and over the whole period, rather than when we focus on individual countries or shorter sub-periods.

3.6.2 Balance of Trade

To analyse sustainability condition for the balance of payments we use two proxies, the balance of trade and current account. Starting from the balance of trade sustainability analysis, first we are interested whether the adjustment of components of the balance of trade, exports and imports, is symmetric or not. Hence, we estimate the following equations for the change in exports and imports shares of GDP:

$$\Delta m_t = a_1 + b_1 bot_{t-1} + u_{1t}$$

$$\Delta x_t = a_2 + b_2 bot_{t-1} + u_{2t}.$$

These equations are estimated for pooled data with fixed cross-country effects for the whole period, and we obtain the following results:

$$\begin{array}{lllllll} \Delta\, m_t & = & 0.003 & + & 0.051\, bot_{t-1} & + & u_{1t} & R^2 = 0.018 \\ & & (0.001) & & (0.011) & & & SER^2 = 0.027 \\ \Delta\, x_t & = & 0.001 & - & 0.054\, bot_{t-1} & + & u_{2t} & R^2 = 0.027 \\ & & (0.001) & & (0.009) & & & SER^2 = 0.023 \end{array}$$

We find that the adjustment is almost perfectly symmetric with both, exports and imports shares, adjusting by about 5% a year.

Since the adjustment of the balance of trade components is symmetric, rather than estimating separate equations for exports and imports shares, we can estimate the balance of trade as a whole. We estimate (3.27):

$$\Delta bot_{it} = \alpha + \beta bot_{it-1} + \varepsilon_{it},$$

where $bot_{it} = x_{it} - m_{it}$ (x_{it} and m_{it} are exports and imports shares of GDP), $\alpha = (a_2 - a_1)$, $\beta = (b_2 - b_1)$ and $\varepsilon_{it} = (u_{2it} - u_{1it})$.

For the whole sample with fixed cross-country effects for 1870-2016 the estimates are:

$$\begin{array}{lcccccccc} \Delta\, bot_t & = & -0.002 & - & 0.105\, bot_{t-1} & + & \varepsilon_t & & R^2 = 0.048 \\ & & (0.001) & & (0.001) & & & & SER^2 = 0.024 \end{array}$$

It appears that balance of trade is a stabilising process when the model is estimated for the whole sample over long period. One percentage point increase in the previous period balance of trade causes about 10-11 percentage points decrease in its current level per year. This adjustment is done through about equal changes in exports and imports shares.

Since the balance of trade and current account are unlikely to be trend stationary, we can potentially use non-standard critical values from the bounds testing approach proposed by Pesaran et al. (2001), which does not require us to identify whether the variables are trend stationary, I(0), or difference-stationary, I(1). For the autoregressive specifications, such as the balance of trade model, (3.27), and current account model, (3.28), the corresponding non-standard critical values for are -2.86 for I(0) and -3.22 for I(1). However, since the autoregressive models of this type, (3.27) and (3.28), only have a lagged dependent variable as the explanatory variable, we can simply check significance against the usual Augmented Dickey-Fuller critical value of -2.9 for a 5% significance level and a model with a constant and no trend (Fuller, 2009). Since this critical value, -2.9 is also in the range of critical values suggested by bounds testing approach (-2.86 and -3.22), for both these models, (3.27) and (3.28), we use critical value of -2.9 to determine the significance of the explanatory variable.

The Table 3.2 summarises the estimates for each country for the whole period and for three sub-periods. The sustainability condition for the balance of trade is confirmed for more than half of the sample, 10 out of 17 countries, when we estimate the model over the whole period. We are more likely to reject a unit root the longer the span of the data is and confirm that considering the whole period the balance of trade is stationary for Australia, Canada, Denmark, Finland, France, Japan, Portugal, Spain, Sweden and the UK.

When we estimate the model over shorter sub-periods, the estimates for the first sub-period confirm that the balance of trade is stabilising for six countries, Finland, Japan, Spain, Sweden, Switzerland and the US. Moving to the results for 1915-1950, there are four stationary series, namely Australia, Denmark, Sweden and Switzerland. Finally, in the third sub-period, 1951-2016, the balance of trade is stabilising for Australia and Italy. In addition, the average for all 17 countries for the whole period suggests a slightly faster adjustment to equilibrium (14.6% per year) than the estimate for pooled data (10.5% per year). However, the 14.6% estimate includes non-significant coefficients.

Table 3.2 OLS: Balance of Trade Share of GDP (Equation (3.27); 1870-2016 and Sub-Periods)

	1870-1914		**1915-1950**		**1951-2016**		**1870-2016**	
Country	β	**t-stat.**[1]	β	**t-stat.**	β	**t-stat.**	β	**t-stat.**
Australia	-0.238	-2.359	**-0.573**	**-3.652**	**-0.764**	**-6.385**	**-0.341**	**-5.400**
Belgium	-0.345	-2.568	-0.295	-2.409	-0.115	-2.048	-0.112	-2.693
Canada	-0.221	-2.268	-0.358	-2.716	-0.188	-2.508	**-0.174**	**-3.761**
Denmark	-0.431	-2.823	**-0.541**	**-3.604**	-0.064	-1.518	**-0.215**	**-4.099**
Finland	**-0.560**	**-4.053**	-0.338	-2.636	-0.141	-2.201	**-0.214**	**-4.156**
France	-0.278	-2.666	-0.147	-1.613	-0.186	-2.514	**-0.138**	**-3.276**
Germany	-0.231	-2.311	-0.459	-2.576	-0.146	-2.530	-0.078	-2.355
Italy	-0.268	-2.553	-0.114	-1.412	**-0.258**	**-2.979**	-0.101	-2.683
Japan	**-0.737**	**-4.681**	-0.245	-2.071	-0.166	-2.435	**-0.276**	**-4.683**
Netherlands	-0.311	-2.645	-0.259	-1.734	-0.087	-1.968	-0.025	-0.856
Norway	-0.241	-2.616	-0.422	-2.682	-0.049	-1.449	-0.063	-2.057
Portugal	-0.229	-2.216	-0.176	-1.829	-0.189	-2.282	**-0.132**	**-3.129**
Spain	**-0.457**	**-3.640**	-0.351	-2.708	-0.110	-1.702	**-0.136**	**-3.168**
Sweden	**-0.349**	**-3.001**	**-0.480**	**-3.311**	-0.081	-1.695	**-0.172**	**-3.696**
Switzerland	**-0.520**	**-3.291**	**-0.475**	**-3.242**	-0.054	-1.009	-0.079	-2.060
UK	-0.205	-2.306	-0.299	-2.247	-0.124	-1.872	**-0.157**	**-3.466**
US	**-0.312**	**-2.910**	-0.205	-1.913	-0.046	-1.243	-0.067	-2.137
Average:	**-0.349**		**-0.337**		**-0.163**		**-0.146**	

Notes: x and m are exports and imports shares respectively, D is the balance of trade (share of GDP). [1]t-statistic for β in equation (3.27). Significant coefficients are in bold (5% significance level; significance is determined using the Augmented Dickey-Fuller critical value of -2.9).

Our results suggest that the sustainability condition for the balance of trade is most likely to be confirmed when we average over the large sample and long period. Nonetheless, even when we consider heterogeneous estimates, there is still a lot of evidence of the balance of payments solvency. Although for eight out of 17 countries in our sample (Belgium, Canada, France, Germany, the Netherlands, Norway, Portugal and the UK), the balance of trade is non-stationary over individual sub-periods, but the sustainability condition is confirmed for the whole period.

3.6.3 Current Account

By analogy with the balance of trade estimations, we check the sustainability condition for the current account expressed as share of GDP.

Besides trade in goods and services (balance of trade), current account has other components, such as investment incomes and transfer payments. Hence, we create a variable z_t which captures these other elements of the current account. Thus, we define z_t as

$$z_t = ca_t - bot_t,$$

where ca_t is the current account share of GDP and bot_t is the balance of trade share of GDP, where $bot_t = x_t - m_t$ (x_t and m_t are exports and imports shares of GDP).

Hence, to see how the components of the current account adjust to changes in the current account share of GDP, if they do, we first estimate the following equations, one for exports:

$$\Delta x_t = a_1 + b_1 ca_{t-1} + u_{1t}$$

one for imports:

$$\Delta m_t = a_2 + b_2 ca_{t-1} + u_{2t}$$

and one for non balance of trade elements of the current account:

$$\Delta z_t = a_3 + b_3 ca_{t-1} + u_{3t}.$$

We find that all three processes confirm adjustment and the estimates are:

$$\begin{array}{lclclclcl}
\Delta\, x_t & = & 0.002 & - & 0.069\, ca_{t-1} & + & u_{1t} & & R^2 = 0.027 \\
 & & (0.001) & & (0.012) & & & & SER^2 = 0.022 \\
\Delta\, m_t & = & 0.002 & + & 0.010\, ca_{t-1} & + & u_{2t} & & R^2 = 0.032 \\
 & & (0.001) & & (0.014) & & & & SER^2 = 0.026 \\
\Delta\, z_t & = & -0.001 & - & 0.040\, ca_{t-1} & + & u_{3t} & & R^2 = 0.011 \\
 & & (0.001) & & (0.009) & & & & SER^2 = 0.016
\end{array}$$

There is a small adjustment to equilibrium by the non balance of trade components of the current account, 4% per year, larger amount by exports, about 7% per annum, and the largest is by imports, 10% per year.

There are some missing observations that result in different number of observations for these three regressions. There are 2,313 observations for exports and imports equations, but only 2,305 for Δz_t regression. Hence, when we estimate (3.28) for the current account (expressed as share of GDP) as whole

$$\Delta ca_{it} = \alpha + \beta ca_{it-1} + \varepsilon_{it},$$

where by design $\alpha = (a_1 - a_2 + a_3)$, $\beta = (b_1 - b_2 + b_3)$ and $\varepsilon_{it} = (u_{1t} - u_{2t} + u_{3t})$, assuming the number of observations for all variables involved (x_{it}, m_{it}, z_{it} and ca_{it}) is the same. It is not the case for our dataset, thus, due to missing observations these equalities may not hold exactly.

For the pooled regression estimated over the whole sample the results are presented below:

$$\begin{array}{lcccccccl} \Delta\, ca_t & = & -0.001 & - & 0.205\, ca_{t-1} & + & \varepsilon_t & & R^2 = 0.106 \\ & & (0.001) & & (0.013) & & & & SER^2 = 0.024 \end{array}$$

Similarly to the balance of trade results, we can confirm the sustainability of the current account when we estimate the model for pooled data over the whole period. The significant negative coefficient confirms that current account adjusts to its previous values in the self-corrective way, hence, this process is stabilising. The coefficient suggests that the current account adjusts to the equilibrium level by about 21% in a year. Most of the adjustment is done by imports and exports and a smaller bit by the non balance of payments components of the current account.

Heterogeneous estimates, that are summarised in the Table 3.3, suggest that the current account is stabilising in all but four countries (Canada, Germany, Switzerland and the US) considering the whole period. As was in case of the balance of trade estimations, there is less evidence of stabilisation over shorter sub-periods. Nonetheless, for six countries current account share of GDP is stabilising over two out of three sub-periods, such as in case of Australia, Denmark, Finland, France, Germany and Japan. For most countries, 14 out of 17, there is evidence of stabilisation of the current account at least over one sub-period.

There are three countries, Portugal, Switzerland and the US for which there is no evidence of stabilisation over any of the three sub-periods, but the current account is stabilising over the whole period for Portugal. Also, it is worth to note that in case of Switzerland we have an insufficient number of observations to estimate the model over the first sub-period, 1870-1914. In addition, the average annual adjustment for 17 countries for the whole period is the same as the estimate for the pooled data, about 21% per year.

Overall, the current account appears to be stabilising when we use large number of countries and sufficiently long estimation period. Heterogeneous estimates for the whole period also provide strong support for the sustainability of the current account. However, as in case with the balance of trade, it is more challenging to confirm stabilisation of the current account when we consider each country in a sample individually and use shorter estimation periods.

Table 3.3 OLS: Current Account Share of GDP (Equation (3.28); 1870-2016 and Sub-Periods)

	1870-1914		**1915-1950**		**1951-2016**		**1870-2016**	
Country	β	**t-stat.**[1]	β	**t-stat.**	β	**t-stat.**	β	**t-stat.**
Australia	-0.262	-2.518	**-0.712**	**-4.327**	**-0.650**	**-5.629**	**-0.541**	**-7.304**
Belgium	-0.345	-2.568	-0.315	-2.502	**-0.221**	**-2.981**	**-0.128**	**-2.888**
Canada	-0.120	-1.549	-0.217	-1.875	**-0.256**	**-3.055**	-0.060	-1.625
Denmark	**-0.635**	**-3.394**	**-0.594**	**-3.145**	-0.086	-1.464	**-0.178**	**-3.179**
Finland	**-0.619**	**-4.358**	-0.304	-2.467	**-0.307**	**-3.398**	**-0.249**	**-4.532**
France	-0.266	-2.566	**-0.564**	**-6.804**	**-0.439**	**-4.245**	**-0.463**	**-11.415**
Germany	**-0.687**	**-4.557**	**-0.979**	**-4.011**	-0.038	-0.824	-0.105	-2.366
Italy	-0.248	-2.475	-0.203	-1.953	**-0.414**	**-4.039**	**-0.200**	**-3.965**
Japan	**-0.736**	**-4.751**	-0.197	-1.761	**-0.288**	**-3.406**	**-0.335**	**-5.336**
Netherlands	-0.163	-1.710	-0.392	-2.335	**-0.254**	**-2.966**	**-0.195**	**-3.721**
Norway	-0.196	-2.322	**-0.491**	**-2.929**	-0.109	-1.914	**-0.166**	**-3.534**
Portugal	-0.229	-2.216	-0.169	-1.692	-0.155	-2.292	**-0.142**	**-3.283**
Spain	**-0.368**	**-3.163**	-0.668	-2.517	-0.162	-2.275	**-0.165**	**-3.185**
Sweden	-0.188	-1.826	**-0.481**	**-3.298**	-0.076	-1.495	**-0.177**	**-3.666**
Switzerland	X	X	-0.504	-2.626	-0.151	-2.222	-0.161	-2.703
UK	-0.150	-1.841	-0.310	-2.615	**-0.264**	**-3.152**	**-0.127**	**-3.102**
US	-0.153	-2.020	-0.310	-2.473	-0.052	-1.297	-0.091	-2.599
Average:	**-0.335**		**-0.436**		**-0.231**		**-0.205**	

Notes: ca_t is the current account (share of GDP); X - insufficient number of observations. [1]t-statistic for β in equation (3.28). Significant coefficients are in bold (5% significance level; significance is determined using the Augmented Dickey-Fuller critical value of -2.9).

3.7 Asymmetric Adjustment

Finally, we analyse the adjustment patterns by checking for any asymmetry in the adjustment towards equilibrium among the countries that are running deficits versus those running surpluses. We estimate (3.29) below

$$\Delta ca_{it} = \alpha + \beta ca_{it-1} + \gamma capos_{it-1} + \varepsilon_{it}.$$

For the pooled data and the whole period we obtain the following results:

$$\begin{array}{lllllllll} \Delta\, ca_t & = & -0.003 & - & 0.262\, ca_{t-1} & + & 0.131\, capos_{t-1} & + \quad \varepsilon_t & R^2 = 0.111 \\ & & (0.001) & & (0.019) & & (0.034) & & SER = 0.024 \end{array}$$

The adjustments are significantly different, the feedback is -0.26 when the current account balance is negative and -0.13 $(-0.262 + 0.131 = -0.131)$ when it is positive.

Hence, the annual adjustment to equilibrium is roughly twice higher (26% per year) when the country is running a deficit than when it is running a surplus (13% per annum), suggesting that there is likely to be more pressure on economies that are running deficits versus those running surpluses.

For the balance of trade, (3.27), and current account, (3.28), models we used non-standard critical value of -2.9 to check the significance of the coefficient of the lagged dependent variable. However, we cannot use this critical value to check significance of the β and γ coefficients in the asymmetric adjustment model. The bound testing approach non-standard values also will not work as it provides the critical value for joint significance of the explanatory variables, while we are interested in whether there is a feedback from either of the coefficients. Hence, we check significance against standard critical value of ± 1.96 (5% significance level).

Moving to the heterogeneous estimates, the results for individual countries over the whole period and three sub-periods are summarised in the Table 3.4.

Table 3.4 Asymmetric Adjustment (Equation (3.29); 1870-2016 and Sub-Periods)

	1870-1914				**1915-1950**				**1951-2016**				**1870-2016**			
	β		γ		β		γ		β		γ		β		γ	
Country	**Coeff.**[1]	**t-stat.**[2]	**Coeff.**	**t-stat.**	**Coeff.**	**t-stat.**	**Coeff.**	**t-stat.**	**Coeff.**	**t-stat.**	**Coeff.**	**t-stat.**	**Coeff.**	**t-stat.**	**Coeff.**	**t-stat.**
Australia	-0.303	-1.645	0.120	0.269	**-0.880**	**-3.836**	0.642	1.048	**-0.529**	**-3.718**	-0.858	-1.437	**-0.665**	**-6.362**	0.449	1.671
Belgium	X	X	X	X	-0.231	-1.515	-0.757	-0.988	**-0.445**	**-2.153**	0.319	1.161	**-0.149**	**-2.366**	0.094	0.476
Canada	X	X	X	X	**-0.574**	**-2.893**	**0.701**	**2.157**	**-0.380**	**-2.903**	0.590	1.228	-0.090	-1.948	0.149	1.079
Denmark	**-0.633**	**-2.783**	-0.016	-0.012	-0.477	-1.691	-0.360	-0.562	**-0.395**	**-2.761**	**0.502**	**2.354**	**-0.495**	**-4.679**	**0.611**	**3.479**
Finland	X	X	X	X	**-0.392**	**-2.467**	0.558	0.880	-0.128	-0.568	-0.281	-0.869	**-0.282**	**-3.696**	0.128	0.629
France	0.148	0.124	-0.432	-0.347	**-0.699**	**-6.478**	0.656	1.824	**-0.540**	**-2.429**	0.182	0.515	**-0.675**	**-13.444**	**0.583**	**6.037**
Germany	X	X	X	X	**-0.950**	**-2.384**	-0.072	-0.095	-0.368	-1.184	0.363	1.074	**-0.734**	**-3.708**	**0.715**	**3.253**
Italy	-0.375	-1.657	0.207	0.625	-0.177	-1.549	-0.775	-0.586	-0.407	-1.790	-0.013	-0.036	**-0.163**	**-2.553**	-0.172	-0.941
Japan	**-0.723**	**-3.385**	-0.038	-0.088	-0.274	-0.724	0.096	0.213	**-1.042**	**-3.057**	**0.925**	**2.280**	**-0.582**	**-4.446**	**0.400**	**2.142**
Netherlands	0.163	0.329	-0.367	-0.672	-0.591	-1.898	0.513	0.762	**-1.324**	**-1.998**	1.146	1.628	-0.207	-0.935	0.015	0.058
Norway	-0.210	-1.561	0.058	0.133	-0.272	-1.228	-0.851	-1.461	-0.242	-1.355	0.181	0.783	**-0.311**	**-3.045**	0.237	1.592
Portugal	X	X	X	X	-0.165	-1.379	-0.047	-0.061	-0.117	-1.490	-0.489	-0.965	**-0.118**	**-2.467**	-0.380	-1.165
Spain	-0.576	-1.889	0.326	0.739	-0.522	-0.722	-0.214	-0.218	-0.091	-1.018	-0.461	-1.313	-0.073	-1.059	**-0.382**	**-1.982**
Sweden	-0.272	-1.863	0.782	0.817	-0.594	-1.765	0.168	0.372	**-0.382**	**-2.346**	**0.394**	**1.975**	**-0.241**	**-2.340**	0.108	0.703
Switzerland	X	X	X	X	**-1.312**	**-2.422**	1.181	1.586	-0.340	-0.838	0.210	0.474	**-0.644**	**-1.978**	0.543	1.508
UK	X	X	X	X	-0.201	-1.180	-0.397	-0.901	-0.091	-0.773	**-0.730**	**-2.029**	**-0.231**	**-2.771**	0.180	1.432
US	-0.128	-0.969	-0.065	-0.236	-2.095	-1.543	1.874	1.320	-0.026	-0.528	-0.344	-1.014	-0.008	-0.159	**-0.217**	**-2.147**
Average	**-0.291**		**0.057**		**-0.612**		**0.172**		**-0.403**		**0.096**		**-0.333**		**0.180**	

Notes: where β is coefficient of ca_{it-1} and γ is the coefficient of $capos_{it-1}$; ca_{it} is the current account (share of GDP). [1] Coefficient, [2] t-statistic. X - missing coefficients (and t-statistic) due to the insufficient number of observations or near singular matrix. Significant coefficients are in bold (5% significance level; significance is determined using the standard critical values, ± 1.96).

We expect the coefficient β to be negative and the coefficient γ to be positive if there is more pressure on the deficit countries to adjust. Over the whole period $\beta < 0$ for all countries and $\gamma > 0$ for all economies in the sample with exception of four (Italy, Portugal, Spain and the US) for which γ coefficient is negative. Only in the first sub-period there are two positive β coefficients (for France and the Netherlands). The sign of the γ coefficient varies when the model is estimated over shorter sub-periods, but in most cases it is positive. The γ coefficient is the sum of the ca_{it} and $capos_{it}$ which should also be negative, so in absolute value the coefficient on $capos_{it-1}$ should be smaller than that on ca_{it}. This largely seems to be the case, even if they are not all significant. The lack of significance may result from the noise in the system as the equations do not fit well when estimated for individual countries over short time periods. Overall, there is strong evidence of asymmetric adjustment when we average over the whole sample and consider the whole period, but there is less evidence when we consider heterogeneous estimates and sub-periods.

3.8 Conclusion

When we estimate a relationship between exports and imports using a simple static model and a VECM for 17 countries for 1870-2016, one unit change in exports seems to cause approximately one unit change in imports, suggesting that exports and imports balance to maintain a stable balance of payments. We also find strong evidence of stabilisation for the balance of payments and current account when we average over the whole sample and estimation period. Nonetheless, it is more challenging to establish these results when we focus on shorter estimation periods.

First we consider the balance of trade components, exports and imports. Static model estimates for individual countries suggest that there is a strong positive relationship between exports and imports (expressed as shares of GDP) for all countries, except Norway. However, the average for heterogeneous estimates suggests that the relationship between exports and imports is not one-to-one, but rather imports share increase by 0.76 percentage point when export shares increase by one percentage point. When we estimate this model over shorter sub-periods it does not significantly alter this average as it remains in the 0.68-0.78 range.

Since export and import shares are likely to be co-determined, we also estimate the relationship between them using a vector error-correction model with one lag. The VECM(1,1) estimates for pooled data confirm a close to unity relationship between exports and imports shares. Nonetheless, once again, heterogeneous results are mixed. We find that over the whole period cointegration coefficient is significant in all but three countries, but it averages to -0.760, which is not very close to unity. When we try to impose a restriction on the cointegration coefficient to be equal -1, this restriction is rejected in case of about half of the sample over the whole period, and it did not improve significance of the imports and exports loading coefficients. We also try to use VECM(1,1) with logarithms of the trade flows expressed in nominal values and local currency. In this case, the long-run import-export relationship is much closer to unity not only for pooled data, but the average for the whole sample over the whole period is -0.997. Moreover, this is consistent when the model is estimated over three sub-periods. However, there are also many insignificant coefficients, such as the imports speed of adjustment coefficient is significant only in seven countries, comparing to eleven in the VECM(1,1) with trade shares. Hence, there does not seem to be a major advantage of using logarithms over shares or the other way around for the analysis of the imports-exports relationship.

We then analyse balance of payments sustainability using balance of trade and current account shares of GDP as proxies. Pooled estimates for the whole period suggest that both, balance of trade and current account are stabilising processes. In case of the balance of trade, its annual adjustment is about 10% and the adjustment done by exports and imports is almost symmetric, both adjust by about 5% per year. As for the current account, its annual adjustment is about 21% per year with most of the adjustment being done by imports (10% per year), a bit less by exports (7% per year) and the smallest bit (4% per year) by the non balance of trade elements of the current account.

As for the heterogeneous estimates, it appears that despite quite a few missing observations for the current account, there are more significant negative coefficients for this model than for the balance of trade one. Hence, there is more evidence of sustainability for the current account than for the balance of trade. When estimated over 1870-2016, the balance of trade is stabilising in 10 out of 17 countries, but in case of the current account there is evidence of stabilisation in 13 economies. In both cases, however, there does not seem to be a country for which the balance of trade or current account shares of GDP are stabilising over each of the sub-periods and whole period. Nonetheless, it is clear that in case of most countries in our sample, the balance of payments solvency can be confirmed as long as the estimation period is sufficiently long and the sample is relatively large.

In addition, we find that there is evidence of asymmetric adjustment when we analyse the whole sample over the whole period with there being more pressure on countries running deficit to adjust, versus those running surplus. There is, however, less evidence of the asymmetric adjustment when we consider individual countries or shorter estimation periods.

To summarise, when we estimate balance of payments sustainability condition using balance of trade or current account data for many countries and over long period there is strong evidence of stabilisation of the balance of payments. However, the data are very noisy, hence, there is less evidence of stabilisation over shorter estimation periods and for individual countries.

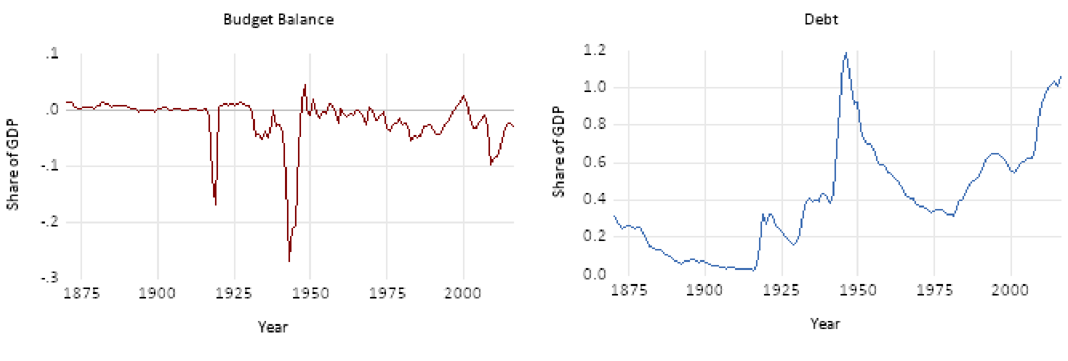
Chapter 4

Fiscal Sustainability

4.1 Introduction

A worrisome number of countries are running large deficits and have high debt-GDP ratios. For example, looking at the US government budget balance and debt-GDP ratio for 1870-2016 (Figure 4.1), especially large deficits occurred around World War I and World War II and, overall, the US has been running deficits since the 1970s with only temporary surpluses in the late 1990s, till the balance turned negative again in the early 2000s. Meanwhile, the debt-GDP ratio has been growing dramatically in the 1940s till it started falling in the 1950s, but has been on the rise again since the 1980s. It appears that quite often decreases in the fiscal deficit occur when the debt was accumulating. This might imply a corrective response from the government as it adjusted fiscal policy in such a way as to be able to meet increasing debt interest payments.

Fig. 4.1 US Government Budget Balance and Debt-GDP Ratio (1870-2016)



Source: Jordà-Schularick-Taylor Macrohistory Database (2019).

Nonetheless, economic theory suggests that a government cannot go on borrowing forever. This concept, known as a government solvency condition, is build on the theoretical frameworks that the government cannot run a Ponzi scheme against the representative agent, meaning paying the interest on old loans by borrowing more (Bohn, 1995). The government solvency is of central political and economic importance, as it is concerned with such issues as whether governments can repay their debts and whether they can continue borrowing and on what terms.

Traditionally deficits sustainability has been assessed by performing unit root and cointegration tests on fiscal and external deficit time series (Ahmed and Rogers, 1995; Quintos, 1995; Trehan and Walsh, 1988). These papers use tests for stationarity and cointegration to check whether the intertemporal budget constraint (IBC), which guarantees that the debt is backed by the expected present value of future primary surpluses, is satisfied. The issue however, is that the intertemporal budget constraint is a quite weak condition as it can be satisfied for series of any order of integration (Bohn, 2005, 2007). Hence, as Bohn argued, the standard time series tests can never reject sustainability. As an alternative approach, Bohn proposed to model surplus-GDP ratio as a positive linear function of the lagged debt-GDP ratio. He explained that if the fiscal policy is sufficiently "responsive" to accumulation of debt, meaning the government takes corrective actions by increasing surplus to meet increasing interest payments, then it can be concluded that the primary surplus has an error-correction representation (Bohn, 2007, p. 1846).

Nonetheless, Bohn et al. (2016) also argued that the whole question of the government solvency is in fact of economic nature and cannot be tested empirically. He argued that the necessary solvency condition is based on the rational expectations and is a question of whether the borrowing government can refinance its debt. Refinancing ability depends on whether the lenders believe the borrowing government is solvent and capable of repaying its debt out of its current revenues. If lenders doubt a nation's solvency, they are likely to refuse to buy this government's debt, making the government insolvent and by doing so, they confirm their expectations. Hence, government of too big to fail countries, such as Germany, France or the UK are able to run large deficits as everyone believes that they will always have the ability to refinance their debts. Nonetheless, this perception by lenders about the solvency of the government can change at any time, meaning there is a danger of a government becoming insolvent if it is running a large deficit-GDP.

Assuming that the true solvency condition depends on the expectations and cannot be tested statistically, this chapter focuses on the sustainability of the government debt and deficit by examining whether a government does in fact take corrective actions in response to the accumulating debt and how the surplus (deficit) adjusts to debt and its own lagged values. Hence, this chapter aims to shed some light on the patterns of adjustment and the feedback process from the lagged surplus and debt. The data for this analysis come from the Jordà-Schularick-Taylor Macrohistory Database. The sample includes 17 countries (Australia, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Japan, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, the UK and US), and we first consider the whole period, 1870-2016, and then focus on three shorter sub-periods, 1870-1914, 1915-1950 and 1951-2016. We find that surplus responds in the correct way to its own lagged values (adjustment is mainly done through revenue, not expenditure), but not to lagged debt-GDP ratio. Hence, there does not seem to be an equilibrium level of debt-GDP, meaning different countries may be able to maintain different debt-GDP ratios. In addition, we included GDP growth, long-term interest rate and inflation rate in the model in order to check whether this would alter the significance of the debt-GDP coefficient. Indeed, when the key macroeconomic variables are added to the model, there is a feedback coming from debt-GDP ratio.

In a previous chapter we analysed a solvency of the balance of payments (BOP), though there we did not have data on the net foreign assets (the stock of debt). Although the algebra in terms of theoretical framework is similar, the economics of two cases is somewhat different. In case of the balance of payments, the net foreign assets (NFA) have to add to zero across countries. So there are some countries who have positive NFA and others with negative NFA balance, hence, there are both borrowers and lenders in the international context. In the public sector context most governments have only debt.

There are also some differences between how the balance of payments and public deficit can be financed. A country can finance a balance of payments deficit through exchange rate adjustments or changing reserves and borrowing. A government can finance a public sector deficit through issuing new debt (there may be interest rate adjustment) and printing money. In the case of governments they are in practice just borrowers: governments can borrow from lenders within the country (internal debt) or from foreign investors (external or foreign debt). In addition, if sustainability fails either exchange rate collapses in the balance of payments case or interest rate shoots up in the government debt case. Thus, even though the nature of default is different, there are possibilities for default in both cases. However, algebra ignores default in a sense that it does not account for a possibility that a country or a government just decides not to repay its loans. Nonetheless, with government debts of more stable or large economies there is usually a high demand for it, for instance, either international foreigners want to hold US assets, financing US government deficit, or internally, pension funds want long duration assets.

To summarise, as Bohn (2007) puts it, the statistical analysis on the solvency has displaced the economic one. The solvency condtion based on the intertemporal budget constraint, commonly used in the empirical papers on the government solvency, is quite weak and the true necessary condition is based on the rational beliefs, making it an economic issue that is impossible to estimate statistically. Hence, in this chapter we analyse the long-run statistical properties of the public sector surplus and government debt and focus on the adjustment of the surplus, revenue and expenditure to changes in the lagged surplus and lagged debt. We check whether a feedback is coming with respect to the surplus, debt or both, and how does it change over different periods. In addition, as in the balance of payments sustainability chapter, we check for asymmetric adjustment in regards to governments running deficits versus those running surpluses. Finally, we discuss the twin deficits hypothesis that links balance of payments and fiscal sustainability together, and check whether our data provide empirical support for cross adjustments of the balance of trade and government surplus.

The outline is as follows, section 4.2 describes the basic theory behind the estimations presented in this chapter. Section 4.3 reviews the literature. The data and descriptive statistics are presented in section 4.4. Section 4.5 describes the econometric approach adopted for the analysis of the fiscal sustainability. The pooled data estimates are discussed in section 4.5.1. Section 4.6 provides the discussion of the country-specific estimates. Section 4.7 is dedicated to the extended model for surplus. Asymmetric adjustment is covered in section 4.8. In section 4.9 we discuss twin deficits hypothesis and analyse the adjustment of the government surplus and the balance of trade together. Finally, section 4.10 contains some concluding comments.

4.2 Theoretical Framework

Bohn (2005) considered sustainability conditions that most often occur in the relevant literature and critically reviewed them. After analysing a case of the US using relatively long data, 1792-2003, he concluded that the most credible evidence in favour of sustainability of fiscal policy is the robust positive response of primary surpluses to increase in the debt-GDP ratio.

Starting point for analysing solvency condition for government debt and deficit is the budget identity that links the deficit to revenues, spending and public debt:

$$DEF_t = E_t^0 - R_t + i_t B_{t-1}, \tag{4.1}$$

where DEF_t is the fiscal or total deficit (with-interest deficit) in year t, E_t^0 is non-interest spending (then E_t is the total expenditure), R_t is the government revenue, $i_t B_{t-1}$ is the interest charge on the stock of debt, B_t, and B_{t-1} is a previous period's debt.

The second budget identity is:

$$DEF_t = B_t - B_{t-1}. \tag{4.2}$$

The variables in the budget identities (4.1) and (4.2) are all in nominal values. In order to separate the stock of debt, B_t, from the flows of government expenditures and revenues, it is useful to define the primary deficit, DEF_t^0, as the non-interest spending minus revenue ($DEF_t^0 = E_t^0 - R_t$). Then by putting (4.1) and (4.2) together, the nominal budget equation can be defined as:

$$B_t = E_t^0 - R_t + (1+i_t)B_{t-1} = DEF_t^0 + (1+i_t)B_{t-1}. \tag{4.3}$$

Then the GDP-ratio version of the (4.3) is:

$$\frac{B_t}{Y_t} = \frac{DEF_t^0}{Y_t} + \left(\frac{1+i_t}{1+\gamma_t}\right)\frac{B_{t-1}}{Y_{t-1}}, \tag{4.4}$$

where $\gamma_t \approx \left(\frac{Y_t}{Y_{t-1}} - 1\right)$ is the GDP growth rate. The first right hand side (RHS) term is the primary deficit, and the second RHS term is the previous period's debt multiplied by a propagation factor.

If $S_t = R_t - E_t^0$ is the primary (non-interest) surplus (the difference between revenue and non-interest spending) and r_t is the "return" on debt (i.e. $r_t = \frac{(1+i_t)}{(1+g_t)} - 1 \approx (i_t - \gamma_t)$), then the (4.4) can be written as:

$$b_t = (1+r_t)b_{t-1} - s_t, \tag{4.5}$$

where, $b_t = B_t/Y_t$ and $s_t = -DEF_t^0/Y_t = -(E_t - R_t)/Y_t$.

Hence, from (4.5), fiscal decisions that affect tax and spending have impact on debt accumulation through their effect on primary deficit (surplus) and r_t.

There are two conditions for ad-hoc sustainability, the ad hoc intertemporal budget constraint (IBC) and transversality condition (TC) (these can be obtained from (4.5) in four steps; an appendix C.1).

The intertemporal budget constraint states that initial debt equals the expected present value of future primary surpluses if discounted future debt converges to zero:

$$b_t^* = \sum_{j=0}^{n} \frac{1}{(1+r)^j} E_t[s_{t+j}]. \qquad \text{[IBC]} \quad (4.6)$$

It is equivalent to the following no-Ponzi game condition, which is often referred to as the transversality condition:

$$lim_{n\to\infty} \frac{1}{(1+r)^n} E_t[b_{t+n}] = 0. \qquad \text{[TC]} \quad (4.7)$$

Most papers on sustainability are concerned with checking whether the intertemporal budget constraint is satisfied or not, and two main approaches to address this question is by running unit root tests (Ahmed and Rogers, 1995; Trehan and Walsh, 1988) and cointegration tests (Quintos, 1995).

However, Bohn (2005) argues that the ad hoc sustainability definition for the fiscal policy $B_t = E_t \left[\sum_{i=0}^{\infty} (1+r)^{-i} s_{t+1} \right]$, the present value of future primary surpluses equals the initial debt, is flawed.

The equation (4.7) is essentially saying that a borrower can repay the debt in any amount of time and that would satisfy the transversality condition, even if it is in a very long period indeed. This comes from the exponential growth in the discount factor in the transversality condition, (4.7). Taking as an example r = 5% (which is commonly used in the literature on solvency), then the above equation becomes

$$\lim_{n\to\infty} \frac{1}{(1+0.05)^n} E_t[b_{t+n}] = 0,$$

which is the same as:

$$\lim_{n\to\infty} (0.95)^n E_t[b_{t+n}] = 0,$$

but since $(0.95)^n$ goes to zero very quickly with n, so $(0.95)^n E_t[b_{t+n}] \to 0$ just because $\frac{1}{(1+r)^n}$ wipes this whole structure out. Hence, the actual deficit would have to grow exponentially fast to stop it from going to zero.

Thus, Bohn emphasised that exponential decay (division by $(1+r)^n$, where $n = j$, in (4.6)) over linear growth implies a zero limit for any $r > 0$ and, hence, dominates polynomial growth. Hence, the intertemporal budget constraint and transversality condition can be satisfied regardless of the order of integration of fiscal data. Bohn (2007) proved that the intertemporal budget constraint is satisfied if either the debt or the revenue and spending (with interest) are integrated of arbitrarily high order. Thus, unit root and cointegration tests cannot reject the consistency of data sets with the intertemporal budget constraint, as a broad class of stochastic processes can fail these tests for sustainability, but still satisfy the intertemporal budget constraint (Bohn, 2007, p. 1838). He argued that since higher order of integration cannot be completely ruled out, the possibility of sustainability cannot be ever rejected using unit root tests. Since neither the debt series difference-stationarity of any order nor a presence of cointegration between revenues and spending are necessary for the intertemporal budget constraint to be satisfied, the important conclusion then is that the above tests cannot be used to determine whether a particular fiscal policy is sustainable or not.

One of the influential papers on the topic, Trehan and Walsh (1991), confirms presence of cointegration between debt and primary surpluses. Building on this result, Bohn (2005) emphasised that it implies that some linear combination of the two, surplus and debt, is then stationary and a strictly positive linkage between the two is implied (an appendix C.2).

Hence, Bohn proposed to model the relationship between primary surplus and debt-GDP using an error-correction-type policy reaction function (fiscal reaction function). The proof that the approach based on the error-correction-type policy reaction function does not require stationarity driving processes is presented as a Proposition 3 in Bohn's paper (2007, p. 1844-45) and this mathematical proof is also replicated in an appendix to this chapter (C.3). The rate of return is assumed to be constant ($r_t = r$). Bohn presents the sufficient condition as an error-correction relationship. Suppose that

$$(s_t) - \alpha b_{t-1} = z_t \sim I(m), \tag{4.8}$$

for some $\alpha \in (0, 1+r]$ and where z_t is of any order of integration, it does not have to be a stationary process.

Hence, if surplus-GDP ratio is a positive linear function of debt-GDP ratio, then it can be interpreted as that a government takes adjusting actions (such as increases its revenue or decreases expenditure in response to increase in debt-GDP), so that the debt-GDP ratio remains bounded, the transversality condition (4.7) is satisfied and fiscal policy can be considered sustainable. However, the main idea behind it is if the debtor takes adjusting actions in response to accumulating debt (meaning $\alpha > 0$), and all roots are strictly less than $(1+r)$, then the transversality condition (4.7) holds and order of integration is irrelevant.

In this chapter we follow Bohn's (2007) analysis for sustainability and adopt equation (4.8). Whereas the JST macrohistory dataset does not have data on net foreign assets, it does have data on public sector debt-GDP ratio. The data on government revenues and expenditure are obtained from JST macrohistory dataset as well. The revenue, R_t, and expenditure, E_t, are in nominal in local currency, meanwhile the GDP, Y_t, is in nominal terms and local currency, and the stock of debt is used as a debt-GDP ratio, $b_t = B_t/Y_t$. Thus, forming $r_t = R_t/Y_t$ and $e_t = E_t/Y_t$, we can construct a public sector equivalent of (4.8):

$$(r_t - e_t) = \lambda b_{t-1} + z_t. \tag{4.9}$$

If debt is high, revenues must increase or expenditure must fall, causing the deficit to fall, and this returns the system back to equilibrium. We will augment this equation by adding the lagged surplus to allow the surplus to respond to its own lagged values as well as to the lagged debt-GDP:

$$(r_t - e_t) = \mu + \lambda b_{t-1} + \gamma s_{t-1} + z_t. \tag{4.10}$$

Then if we subtract $(r_{t-1} - e_{t-1})$ from both sides of (4.10), we get:

$$(r_t - e_t) - (r_{t-1} - e_{t-1}) = \mu + \lambda b_{t-1} + (\gamma - 1)s_{t-1} + z_t, \tag{4.11}$$

which simplifies to:

$$\Delta r_t - \Delta e_t = \mu + \lambda_1 s_{t-1} + \lambda_2 b_{t-1} + z_t, \tag{4.12}$$

where $\lambda_1 = (\gamma - 1)$ and $\lambda_2 = \lambda$.

Then we obtain the following equations for the change in the revenue

$$\Delta r_t = \mu_1 + \lambda_{11} s_{t-1} + \lambda_{12} b_{t-1} + z_{1t} \tag{4.13}$$

and expenditure:

$$\Delta e_t = \mu_2 + \lambda_{21} s_{t-1} + \lambda_{22} b_{t-1} + z_{2t}, \tag{4.14}$$

where $\mu = \mu_1 - \mu_2$; $\lambda_1 = \lambda_{11} - \lambda_{21}$; $\lambda_2 = \lambda_{12} - \lambda_{22}$. This allows feedback of both debt and deficit on the government revenue and expenditure. Thus, allows for asymmetric adjustment between the components of the surplus.

From a closer look at (4.12) we can see that there might not be an equilibrium level of debt-GDP ratio in the long run. If write the equation (4.12) as

$$\Delta s_t = \mu + \lambda_1 s_{t-1} + \lambda_2 b_{t-1} + \varepsilon_t$$

and if in the long run $s_t = s_{t-1} = s = 0$, so $\Delta s_t = 0$ and $b_t = b_{t-1} = b$, then

$$0 = \mu + 0 + \lambda_2 b,$$

and, thus, the equilibrium level of long-run debt is

$$b = \frac{-\mu}{\lambda_2},$$

but if $\lambda_2 = 0$, then there is no equilibrium debt-GDP ratio. As the descriptive statistics in the Table 4.1 show, average debt-GDP ratios vary a lot over time and countries, so it is possible that there is no equilibrium debt-GDP ratio. It might be that any ratio is sustainable as long as those lending to the government think it is. For instance, a number of countries including the UK and Japan have had no difficulty borrowing even when debt-GDP ratios were over 200%.

4.3 Literature Review

We consider fiscal sustainability in order to check if government debt and deficit are sustainable in the long run. There is no consensus on how dangerous large government deficit and debt are. Boskin (2020) presents a useful overview of the traditional view on this question. On one side, expansionary fiscal policy and high debt might be a useful if not necessary remedy in time of recession to boost the economy, assuming of course that government is to finance investments that are most likely to be productive and worthy in the long run. Tax smoothing might be another macroeconomic benefit of large government deficits, for instance in times of severe recession or wars.

On the other side, large deficit and debt give incentive to central bank to finance government's liabilities by simply printing money, which may eventually lead to high inflation. Moreover, high government borrowings are likely eventually lead to increase in the nominal interest rates, making borrowing more expensive and essentially crowd out private investment. Furthermore, larger deficit can be associated with higher tax on the private sector and higher corporation tax which can also lead to fall in consumption and investment. Moreover, often government spending is financed by borrowing from private sector and while investors are buying government bonds, there is less funds available for the private sector investments. This can also contribute to the crowding out of the private investment. Mountford and Uhlig (2009) found that increase in government expenditure and tax do indeed lead to fall in investment, with the multiplier being much higher in the latter case (increase in tax). In most extreme cases, expectation of higher interest rates in the future and inflation might lead to capital flight and depreciation of the currency which might end with a crisis (Boskin, 2020).

Nonetheless, Mountford and Uhlig (2009) found that while deficit spending does seem to crowd out private investment, contrary to the traditional view, it does not cause increase in interest rates or real wages. They argue that deficit spending can in fact to a small extent stimulate the economy. Moreover, Blanchard (2019) argued that lately risk-free rate has been below growth rate; hence, large debt can come cost-free and be purely beneficial for the economy.

However, Boskin (2020) claimed that the long-term costs of deficit spending are often overlooked in the literature. He gave example of Uhlig's (2013) finding according to which each dollar of debt financed spending essentially translates in the cost of 3.40 US dollars in the current value. Boskin concluded that historical data as well as debt and fiscal variables projections suggested that increasing government deficit and debt are sources of major economic uncertainties and likely to lead to higher taxes and lower incomes as well as major intergenerational inequality in the future.

Whether large government deficits and debt are sustainable or not is a question open to debate. For instance, starting with a linearised version of the government debt flow identity and iterating backward, Cochrane (2019) performed an ex-post rational calculation to shed some light on the decomposition of the variation in the debt-GDP. For his empirical analysis he focused on the US case, using data for the US Federal debt, Treasury rates and US primary surplus, and considered various estimation periods for the decomposition of the value of debt variance, with the longest period being 1930-2018. Assuming that the market value of government debt was equal to the present discounted value of primary surpluses, he argued that the two factors that were equally important sources of the variation in the value of government debt were accuracy in prediction of the future primary surpluses and variation in discount rates. The latter came mostly from nominal returns. Meanwhile, future growth rate did not seem to substantially affect variations in the market value of debt to GDP ratio. He also emphasised, that when the full sample that went back to 1930 was used, the effect of change in discount rates on the variation in debt-GDP level was more dominant comparing to the effect of forecasts of the future surpluses.

In contrast, in this chapter we use longer data, 1870-2016, and focus not on the US specifically, but use a sample of 17 countries to account for possible heterogeneity in the long-run adjustment processes of government surplus and debt-GDP ratio. We find that while surplus tends to stabilise in majority of cases when we consider the whole period, 1870-2016, there does not seem to be an equilibrium level for the debt-GDP ratio and different governments can settle on various levels of debt-GDP in the long run.

Also focusing on the US case, Bohn (1998) argued that primary surplus is an increasing function of the debt-GDP ratio, and the government budget constraint ensures that the debt-

GDP ratio is stationary. He based his framework on the Barro's (1979) tax-smoothing model and using data for 1916-1995 (and six sub-periods), Bohn found that cyclical fluctuations and major economic and political events, especially war and the resulting increase in military spending can significantly alter this ratio, with war years, for instance, being clear negative outliers. Hence, those factors need to be controlled for when estimating the relationship between debt-GDP ratio and surplus. He also clarifies that for debt-GDP ratio to be stationary the interest rates need to be near or below the GDP growth rate (Bohn, 1998, p. 955). Bohn also criticised using standard unit root regressions to confirm mean-reversion in the debt-GDP ratio as those omit variables, such as the level of temporary government spending (GVAR) and a business indicator (YVAR)[1], and do not properly account for the systematic components in the error term, and are inconsistent if the autoregressive (AR) coefficient is close to one. Moreover, in his later papers (2005, 2007) Bohn argued that classic tests for sustainability do not properly address effect of uncertainty and risk aversion on the interest rate. If interest rates on government bonds are below the GDP growth rate, as was in case of the US, and which was true for many countries following the financial crisis 2007-2008, then primary deficits do not necessarily confirm insolvency, as government is likely to adopt sustainability policies. He also argued that, at least in case of the US, fiscal policy can be used to maintain stationarity of the debt-GDP ratio unless the interest and growth rates substantially deteriorate. Finally, he argued that using real levels of fiscal variables in cointegration analysis of this sort is inappropriate as the unit root might be from GDP series while the debt-GDP ratio is stationary.

Schoder (2014) again focused on the sustainability of the government debt. He analysed primary surplus response to the change in the debt-GDP ratio for nine European Monetary Union (EMU) and six non-member countries using quarterly data for 1981Q1-2010Q4 from the OECD Economic Outlook 89 database. He adopted a stochastic general equilibrium model and followed theoretical work by Bohn (1995, 1998). A positive response of the surplus to a one-unit change in the debt-GDP ratio is considered to be a sufficient condition to confirm the validity of the intertemporal budget constraint and, therefore, sustainability of the government debt. In his estimation he used the mean-group and pooled mean-group estimators developed by Pesaran and Smith (1995) and Pesaran et al. (1999). First he estimated the long-run relationship between surplus and government debt for individual countries using the error-correction form of the autoregressive distributed lag model. Since the two variables might be covariance stationary, Schoder chose ARDL bounds testing approach proposed by Pesaran et al. (2001). However, even though he found evidence of the

[1]GVAR and YVAR are measures of temporary government spending and of cyclical variations in output, respectively, from Barro (1986).

long-run relationship between surplus and debt for eight countries, the response coefficient that measured the effect of debt-GDP ratio on surplus was insignificant for most of the countries in the sample. Then he split the estimation period into years before and after introduction of the Euro Convergence Criteria, 1980Q1-1996Q4 and 1997Q1-2010Q4, and estimated pooled regression for the whole sample for the two sub-periods. He found that over both sub-periods the response coefficient was positive and significant at 1% level. Schoder also emphasised that when he split the sample in the EMU and non-EMU countries, the response coefficient decreased significantly for the later group (from 0.032 to 0.005), but not so much for the former (from 0.037 to 0.030), suggesting that tighter rules implied by the Euro Convergence Criteria had a positive effect on the sovereign debt sustainability in the EMU countries. In addition, he analysed the effect of the financial crisis 2007-2008 and found that non-member countries were affected more than EMU economies as investors showed more faith in government bonds of the monetary union member countries. He emphasised that in case of non-sustainable non-EMU economies, the debt accumulation appeared to have been sustainable before the shock, as the response coefficient was positive (0.031) when the estimation did not include the financial crisis, but turned to a negative -0.001 when it was included.

Another paper that builds on Bohn (1998) definition of the sufficient condition for solvency of the government debt is Greiner et al. (2007). They focused on countries with high debt-GDP ratio and those who exceeded the 3% deficit-GDP threshold and thus violated Maastricht treaty. As in Schoder (2014) paper, the debt is considered solvent if the response coefficient of surplus-GDP to changes in debt-GDP is positive. However, they model the relationship slightly differently from Schoder (2014) by adding a vector variable to their regression, which accounts for net interest payments on public debt relative to GDP as well as a business cycle variable, as these factors are related to the surplus and hence might have a significant effect on it. They analysed countries individually with their sample including five economies, namely Germany, France, Italy, Portugal and the US. The estimation periods vary from country to country, 1964-2003 for Italy, 1977-2003 for France and Portugal, 1960-2003 for Germany and the US. Their estimations confirmed sustainability of fiscal policy in France, Italy and the US. In case of Germany the solvency condition is satisfied as the response coefficient is significant and positive, however, it has been decreasing over the last four decades. Furthermore, they found evidence that in case of Portugal the response coefficient is negative, suggesting that the process does not stabilise, but it is insignificant. Overall, they argued that this inverse relationship between public investment and debt-GDP can be clearly observed empirically.

Among the extensive literature on the sustainability of US balance of payments deficit is the paper by Bohn (2005) that provided some empirical support for sustainability of the US fiscal policy. Looking at the real data for the US debt, Bohn identified that the debt-growth was entirely nominal in 1950-1980, but real in the post-1980. By noting periods of the positive deficits and the flat paths of debt, Bohn noted that the equating deficits with increases in debt does not apply to real values nor to GDP-ratios. He explained that the budget identity $DEF_t = B_t - B_{t-1}$, the equation (4.2), only holds in nominal terms, but changes in debt equal the real value of the deficit minus an inflation term. As for the GDP-ratios, changes in the debt-GDP equal to the deficit-GDP minus a nominal growth rate term, which he called a "growth dividend". He argued that the nominal growth term has historically covered the entire interest bill on the US debt. Hence, in contrast to many papers on the topic that focused on the role of government surpluses in keeping debt-GDP bounded, Bohn argued that it was growth rate rather than adjustment of the surpluses that kept the debt-GDP bounded. He also emphasised the role of major wars that are normally deficit-financed. In addition, Bohn argued that analysis of the sustainability condition is subject to specification issues, such as choosing the discounting rate and suitable explanatory variables.

In contrast, to analyse debt sustainability Wyplosz et al. (2007) used a four-step International Monetary Fund (IMF) approach known as a debt sustainability assessment (DSA). It is based on the accounting identity: $b_t - b_{t-1} = (r-g)b_{t-1} - \mathit{primary\ balance}_t$, where b_t is a the debt-GDP ratio, r and g are the real interest and GDP growth rates, respectively. He also discussed the IMF definition of debt sustainability. According to the IMF, debt is considered sustainable if the sum of current debt and the present discounted value of government spendings is below the present discounted value of revenues and if there is no need for any major corrective actions. Wyplosz, however, criticised this definition arguing that the first part of it implies that past trends will be true in the future which is highly unlikely, and the second one is a judgment on a debt being too large, which is subjective. Even if an economy is running a very large debt-GDP ratio, like the US or UK do, the debt can still be sustainable, which is true as long as there are opportunities to refinance. Thus, it is hardly possible to determine a debt ceiling, but Wyplosz argues it should be based on economic costs that are required to service the debt and political acceptability. He explains that debt distress is self-fulfilling process with interest rates on public debts increasing in response to increasing uncertainty about probability of default, creating a vicious circle. Like Bohn (2007, 2016) and Greiner et al. (2007), Wyplosz emphasised that solvency and, hence, sustainability are forward looking and since the future is unknown, one cannot draw long-term conclusion from the analysis of the debt sustainability. It is impossible to foresee the future fiscal policy and whether the government will value the solvency of debt-GDP ratio over economic growth. He argued that even though the DSA approach is simple and transparent, it is also subject to impossibility principle and requires strong assumptions about future evolution of some variables which is in reality unknown and cannot be accurately predicted. Hence, any conclusions from this approach is to be considered with care. As a concluding remark, he also suggests that instead of making assumptions about evolution of the key variables, the alternative is, given these assumptions, to determine what adjustments to government surplus (deficit) are required to achieve a certain debt path. Therefore, he suggested that the best option might be to combine both, debt-stabilising based approach and the IMF one, which is based on debt path projections.

After analysing IMF approach and discussing its faults, such as assumption of growth path being time invariant, Arnone et al. (2005) proposed an alternative approach for assessing debt sustainability. They suggested that existing approaches on the topic can be divided in *optimising* models, that equate marginal benefit and cost of borrowing, *non-optimising* ones that have a flaw of assuming a time invariant growth paths and fail to consider the effect of imports, *fiscal space* models that assume a decrease in government spending in order to finance debt service and *disincentive effects* ones that are based on the idea of debt overhang effect on economic performance. In turn they proposed a five-step framework, which is based on the debt thresholds, which should be country-specific, on debt servicing and on the analyses of the risk of default, of the public deficit financing and of the external financing gap. They argued that while various theoretical frameworks and empirical models on sustainability of public debt address these five issues to various extent, they fail to connect them together.

In addition, Chudik et al. (2017) argued that there is no equilibrium for the level of public debt-GDP ratio and various economies can maintain different levels of indebtedness, while enjoying economic growth. Thus, they emphasised heterogeneity among countries and that high debt-GDP ratio does not necessarily translate into financial distress if an economy has a good track record of repaying its debt obligations. What is of main importance, however, is how fast the debt is growing in respect to the GDP growth. They argued, that temporary jumps in debt-GDP ratio are not an issue, however, if the rises are persistent, then high debt-GDP ratio can have a deteriorating effect on the economy in the long run. Therefore, what matters is the ability of government to assure its potential lenders of its ability to cover debt payments, and that any noticeable rises in debt-GDP ratio are of temporary nature. They also found statistically significant threshold effects for the economies with increasing debt-GDP ratio. They argued that it is the trajectory of debt (how fast it is growing and how persistently) is what of the main importance. They also emphasised the key role of fiscal policy in maintaining the debt at sustainable level. Finally, they confirmed that in the long run there is indeed a negative effect of debt on GDP.

To summarise, building on the rich literature on the government debt solvency, this chapter presents some additional findings on the topic. Firstly, our analysis is based on the estimations obtained from a much longer period, 1870-2016, than was covered in other papers mentioned above. Having a long enough data also allows us to split the data in three sub-period and check how our results vary from one sub-period to another. However, while Bohn's (2007) empirical approach implies the use of debt-GDP ratio and *primary* surplus data, we use debt-GDP ratio and *total* (with interest) surplus as the JST dataset only has data for the government revenues and total (with interest) expenditures. We find that there is in general more evidence of stabilisation when we consider the whole period rather than shorter sub-periods. Secondly, instead of focusing on any particular economy, we use a sample of 17 countries to check whether surplus and debt-GDP ratios adjustment patterns vary from country to country. We find that while surplus is stabilising on its own lagged values in all but one country (Japan), there does not seem to be a significant feedback coming from the lagged debt-GDP ratio, suggesting that different governments can potentially maintain different debt-GDP ratios. Thirdly, in addition to estimating the surplus model, we also estimate separate equations for the government revenue and expenditure, and we use lagged surplus and debt-GDP ratio as independent variables. The data we use is discussed in the following section.

4.4 Data

To check whether the government debt and deficit are stabilising we use data for a sample of seventeen countries for 1870-2016 from the Jordà-Schularick-Taylor Macrohistory Database[2]. The sample includes Australia, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Japan, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, the UK and US.

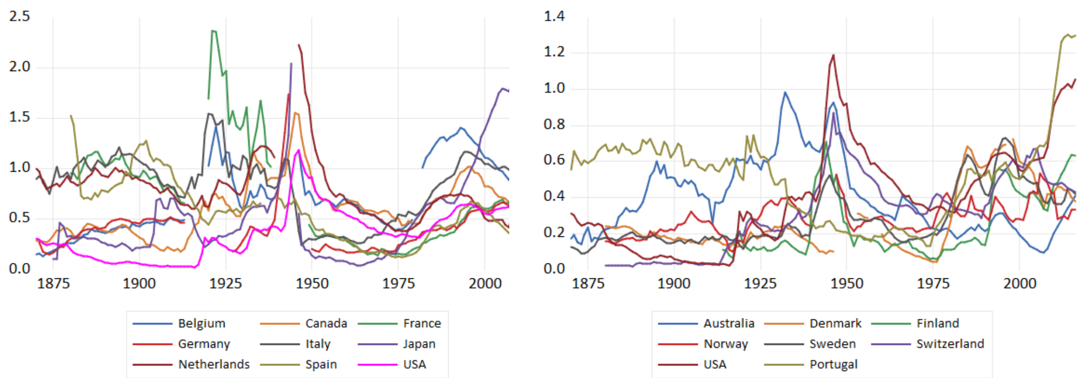
The data that we use for public surplus and debt-GDP are not the ideal measures of the government's liabilities. Boskin (2020) emphasised that there are numerous papers on the limitations of the official deficit and national debt measures. For instance, these measures do not fully account for inflation or intangible investments in R&D and education. We acknowledge these limitations, nonetheless, these measures are widely used in the literature on the sustainability of the government surplus and debt. Thus, for the lack of better alternatives and for simplicity, we also opt for these commonly used measures of deficit and debt. We use data from the JST database, as these data are sufficiently long and are

[2]Oscar Jordà, Moritz Schularick and Alan M. Taylor. 2017. "Macrofinancial History and the New Business Cycle Facts." in NBER Macroeconomics Annual 2016, volume 31, edited by Martin Eichenbaum and Jonathan A. Parker. Chicago: University of Chicago Press.

available for 17 countries. However, to some extent, the debt and deficit measuring issue is made worse from using long span data for various countries as definitions differ for different countries and over time. Nonetheless, our analysis confirms that there are substantial benefits of long span heterogeneous panel data to the analysis of balance of payments and fiscal sustainability.

We use the government surplus (which is calculated using the government revenue and expenditure) and debt-GDP ratio. The Figure 4.2 suggests that the highest increases in debt-GDP ratios over 1870-2016 were recorded in the UK, Japan and France, however, the jumps occurred at different times. In case of France, it was right after the end of World War I, in the UK – in the 1940s, while Japan has been noted to have exceptionally high debt to GDP level since the end of the estimation period covered in this chapter. Among the ones with relatively low debt-GDP ratios are Sweden and Switzerland, that mostly maintained quite low debt to GDP ratio levels throughout 1870-2016.

Fig. 4.2 Debt to GDP Ratios for the Sample (1870-2016)



Source: Jordà-Schularick-Taylor Macrohistory Database (2019).

Since debt is expressed as income share, we also try to include nominal GDP in the model to check if it is significant and should be kept as an independent variable to account for the fact that revenue and expenditure are in local currency, not shares of GDP. In addition, all variables are transformed into logarithms before starting the estimations. The raw data for revenue, expenditure and GDP are expressed in nominal values and local currency.

The data come from multiple sources, normally there is at least two sources for each variable. However, when it comes to definitions, the data sources generally agree on how government revenue, expenditure and debt-GDP are defined.

There are some minor definition specifics for some countries. For instance, for Australia government expenditure for early years of the 1949-1996 period is the combined expenditure of the Commonwealth Revenue Fund and the Loan Fund. In case of Portugal for 1870-1989 the revenue defined explicitly as the sum of tax revenue and other effective income, and from 1990 onwards it is a state revenue where state is a central government. The debt-GDP ratio is generally defined as gross public debt, expressed as a percentage of nominal GDP. For some countries, such as Norway or Switzerland, some data (for 1870-1979 for Norway and 1870-1945 for Switzerland) are for central government debt, and for the rest of the period the data are for a general government gross debt percentage of GDP. These data are in the level form.

There are quite a few missing values in the data for some countries. For instance, in Finland the data for debt-GDP are missing for 1870-1913. In addition, there are no data for revenue and expenditure from 1870 to 1881 for this country. Hence, there is an insufficient number of observations to estimate our models for Finland for the first sub-period, 1870-1914. Other data gaps are less wide, but some are still substantial. For instance, in Australia revenue and expenditure data are missing from 1870 to 1901. In Belgium these data are missing for 1913-1919 and the GDP values are not available for 1914-1919. For Germany the data for revenue, expenditure and GDP are not available for periods covering both World Wars. For Denmark there is a large data gap for revenue, 1936-1952 and three gaps for debt-GDP, 1870-1879, 1947-1952 and 1957-1959. In France, Norway, Spain and Switzerland the data for debt-GDP are missing for a decade from 1870 to 1879 and then for France only from 1914-1919 and 1939-1948, covering the World War I and II. However, for Norway there are gaps in all variables except expenditure for Second World War years. For Spain all data except GDP are missing for 1936-1939. In case of Japan the GDP and debt-GDP data are missing for 1870-1874. For the Netherlands the GDP data are not available for both world war periods.

The Table 4.1 summarises the descriptive statistics for revenue, expenditure and surplus, expressed as shares of GDP and the debt-GDP ratio.

Even though for analysis we use logarithms of revenue, expenditure, surplus, GDP and debt-GDP ratio, it is more useful to run descriptive analysis for the variables expressed as shares of GDP as this allows for comparison across countries.

The revenue-GDP ratio is the highest for the UK (0.24), but is matched by the highest in the sample expenditure-GDP ratio average of 0.25, hence, the average surplus comes close to zero. It is important to note that these are averages over very long period. The lowest average revenue and expenditure shares are for Switzerland (0.06) with one of the lowest debt-GDP ratio in the sample (0.32). Most of the countries, 12 out of 17, are running an average deficit, the exceptions are Germany, Japan, Switzerland and the UK, that have an average surplus of zero, while Norway is the only country with an average positive surplus of 0.01. The lowest minimum value of revenue share is for Switzerland (0.01) followed

Table 4.1 Descriptive Statistics (Shares of GDP; 1870-2016)

Country	Revenue				Expenditure				Surplus				Debt			
	Mean	SD[1]	Max[2]	Min[3]	Mean	SD	Max	Min	Mean	SD	Max	Min	Mean	SD	Max	Min
Australia	0.17	0.08	0.29	0.04	0.20	0.10	0.46	0.01	-0.03	0.05	0.03	-0.32	0.40	0.21	0.98	0.10
Belgium	0.18	0.09	0.32	0.04	0.22	0.11	0.43	0.05	-0.04	0.04	0.02	-0.25	0.71	0.35	1.41	0.13
Canada	0.12	0.06	0.25	0.04	0.14	0.08	0.45	0.04	-0.02	0.04	0.04	-0.22	0.63	0.28	1.55	0.18
Denmark	0.19	0.14	0.42	0.04	0.19	0.14	0.44	0.05	-0.01	0.02	0.05	-0.10	0.27	0.18	0.72	0.04
Finland	0.18	0.07	0.30	0.06	0.21	0.10	0.58	0.06	-0.03	0.06	0.04	-0.37	0.25	0.18	0.71	0.06
France	0.17	0.05	0.25	0.07	0.21	0.09	0.63	0.10	-0.04	0.08	0.05	-0.52	0.76	0.50	2.37	0.14
Germany	0.10	0.06	0.60	0.03	0.10	0.05	0.32	0.03	0.00	0.02	0.06	-0.12	0.42	0.23	1.74	0.15
Italy	0.18	0.06	0.33	0.08	0.23	0.10	0.44	0.10	-0.05	0.07	0.02	-0.36	0.86	0.32	1.54	0.25
Japan	0.13	0.03	0.28	0.07	0.13	0.04	0.27	0.06	0.00	0.03	0.06	-0.08	0.61	0.59	2.39	0.04
Netherlands	0.20	0.09	0.46	0.09	0.22	0.12	1.06	0.09	-0.02	0.07	0.09	-0.75	0.79	0.29	2.23	0.38
Norway	0.21	0.16	0.53	0.02	0.20	0.13	0.41	0.03	0.01	0.05	0.20	-0.05	0.28	0.09	0.53	0.10
Portugal	0.11	0.06	0.25	0.04	0.13	0.08	0.29	0.04	-0.02	0.03	0.01	-0.09	0.51	0.25	1.30	0.13
Spain	0.12	0.04	0.21	0.06	0.13	0.05	0.26	0.08	-0.01	0.02	0.02	-0.09	0.60	0.30	1.52	0.13
Sweden	0.16	0.10	0.42	0.06	0.18	0.11	0.44	0.05	-0.02	0.03	0.07	-0.14	0.28	0.16	0.73	0.09
Switzerland	0.06	0.04	0.13	0.01	0.06	0.04	0.19	0.01	0.00	0.02	0.02	-0.08	0.32	0.21	0.87	0.02
UK	0.24	0.13	0.44	0.06	0.25	0.15	0.64	0.06	0.00	0.10	0.13	-0.45	0.90	0.60	2.70	0.27
US	0.10	0.07	0.20	0.02	0.12	0.09	0.41	0.01	-0.02	0.04	0.04	-0.27	0.40	0.29	1.19	0.02

Notes: [1]standard deviation,[2]maximum, [3]minimum.

closely by Norway and the US (0.02), with the corresponding maximum values for these countries being 0.13, 0.53 and 0.20. Meanwhile, the lowest minimum value for expenditure share is for Australia, Switzerland and the US (0.01). However, the maximum values for expenditure shares for Australia and the US are quite high, around $0.4-0.5$, while it is very low for Switzerland, 0.19. Hence, it appears that Switzerland maintains more or less stable government revenue and spending, while most countries have periods of exceptionally high government surplus followed by periods of running large deficits.

The average debt-GDP ratio is the lowest for Finland (0.25), with a relatively low standard deviation of 0.18. Denmark (0.27), Norway and Sweden (0.28) follow Finland with only slightly higher average debt-GDP ratios. The highest debt-GDP ratio is for the UK (0.90), and it is nearly as high for Italy (0.86) and the Netherlands (0.79). The standard deviation for debt-GDP ratio is the highest for the UK (0.60) followed by Japan (0.59), suggesting that data points are very dispersed, and it is the lowest for Norway (0.09). It appears that Switzerland and the US have the lowest recorded debt-GDP ratios of 0.02 over 1870-2016, while the highest maximum for debt-GDP ratio was for the UK (2.70) and Japan (2.39) followed closely by France (2.37).

In addition, according to the Jarque-Bera test, the data are mostly non-normally distributed, with the only exception being expenditure-revenue ratio for Australia, Japan and the US; in all three cases we fail to reject the null hypothesis of normality at 5% level.

4.5 Empirical Models

For estimation we will work with the logarithms of the shares, rather than the shares themselves. So defining: $lnS_{it} = ln(R_{it}/Y_{it}) - ln(E_{it}/Y_{it}) = (lnr_{it} - lne_{it})$ and $lnb_{it} = ln(B_{it}/Y_{it})$, where R_{it} is the government revenue, E_{it} is the total expenditure, S_{it} is the total surplus, and B_{it} is the debt-GDP ratio; we can write (4.12) as:

$$\Delta lnr_{it} - \Delta lne_{it} = \mu + \lambda_1 lns_{it-1} + \lambda_2 lnb_{it-1} + z_{it},$$

or equivalently

$$[(lnR_{it} - lnY_{it}) - (lnR_{it-1} - lnY_{it-1})] - [(lnE_{it} - lnY_{it}) - (lnE_{it-1} - lnY_{it-1})] =$$

$$\mu + \lambda_1 ln[(lnR_{it-1} - lnY_{it-1}) - (lnE_{it-1} - lnY_{it-1})] + \lambda_2 [lnB_{it-1} - lnY_{it-1}] + z_{it},$$

which simplifies to

$$\Delta lnR_{it} - \Delta lnE_{it} = \mu + \lambda_1 (lnR_{it-1} - lnE_{it-1}) + \lambda_2 (lnB_{it-1} - lnY_{it-1}) + z_{it}. \tag{4.15}$$

To allow for cyclical effects we will also include logarithm of GDP. In addition, we allow for differential effects of revenue and expenditure, everything is in nominal terms. Thus, we estimate:

$$\begin{aligned} \Delta lnR_{it} - \Delta lnE_{it} &= \mu + \lambda_1 (lnR_{it-1} - lnE_{it-1}) + \lambda_2 \\ &(lnB_{it-1} - lnY_{it-1}) + \lambda_3 lnE_{it-1} + \lambda_4 lnY_{it-1} + z_{1it}. \end{aligned} \tag{4.16}$$

Then we split (4.16) into two equations, one for government revenue:

$$\Delta lnR_{it} = \mu_1 + \lambda_{11} lnS_{it-1} + \lambda_{12} lnb_{it-1} + \lambda_{13} lnE_{it-1} + \lambda_{14} lnY_{it-1} + z_{1it}, \tag{4.17}$$

and one for government expenditure:

$$\Delta lnE_{it} = \mu_2 + \lambda_{21} lnS_{it-1} + \lambda_{22} lnb_{it-1} + \lambda_{23} lnE_{it-1} + \lambda_{24} lnY_{it-1} + z_{2it}, \tag{4.18}$$

where $\mu = \mu_1 - \mu_2$, $\lambda_1 = \lambda_{11} - \lambda_{21}$, $\lambda_2 = \lambda_{12} - \lambda_{22}$, $\lambda_3 = \lambda_{13} - \lambda_{23}$, $\lambda_4 = \lambda_{14} - \lambda_{24}$.

We expect $\lambda_{13} = \lambda_{23} \approx 0$ if feedbacks from expenditure and revenue are similar, and $\lambda_4 \approx 0$ if cyclical effects are small.

In (4.16), we expect $\lambda_1 < 0$, as negative lagged surplus should encourage the government to increase the current one. Meanwhile, $\lambda_2 > 0$, as when debt-GDP increases, the government needs to increase current surplus to meet increasing debt payments. From equation (4.17), for ΔlnR_{it}, if surplus is high then the government can cut taxes, hence we expect $\lambda_{11} < 0$. In contrast, if debt is high, there should be an increase in taxes to finance it, thus we expect $\lambda_{12} > 0$. Then from the equation (4.18), for ΔlnE_{it}, if surplus is high, the government can increase expenditure, hence $\lambda_{21} > 0$, while if debt is high, the government needs to reduce expenditure, and we expect $\lambda_{22} < 0$.

We estimate the above equations for ΔlnS_{it} (equation (4.16)), ΔlnR_{it} (equation (4.17)) and ΔlnE_{it} (equation (4.18)) for pooled panel data with fixed country effects. The coefficients and standard errors (in parentheses) are:

$$\underset{}{\Delta \ln S_t} = \underset{(0.016)}{0.040} - \underset{(0.013)}{0.182 \ln S_{t-1}} + \underset{(0.005)}{0.008 \ln b_{t-1}} + \underset{(0.007)}{0.018 \ln E_{t-1}} - \underset{(0.007)}{0.020 \ln Y_{t-1}} + \varepsilon_t \qquad \begin{array}{l} R^2 = 0.110 \\ SER = 0.127 \end{array}$$

$$\Delta \ln R_t = \underset{(0.014)}{0.110} - \underset{(0.012)}{0.172 \ln S_{t-1}} - \underset{(0.004)}{0.036 \ln b_{t-1}} + \underset{(0.006)}{0.036 \ln E_{t-1}} - \underset{(0.006)}{0.039 \ln Y_{t-1}} + \varepsilon_t \qquad \begin{array}{l} R^2 = 0.154 \\ SER = 0.115 \end{array}$$

$$\Delta \ln E_t = \underset{(0.019)}{0.063} + \underset{(0.015)}{0.009 \ln S_{t-1}} - \underset{(0.005)}{0.044 \ln b_{t-1}} + \underset{(0.008)}{0.017 \ln E_{t-1}} - \underset{(0.008)}{0.016 \ln Y_{t-1}} + \varepsilon_t \qquad \begin{array}{l} R^2 = 0.045 \\ SER = 0.150 \end{array}$$

In the models covered in this chapters the variables are unlikely to be trend-stationary. In the previous chapter we considered simple autoregressive models of feedback for the balance of trade and current account and used Augmented Dickey-Fuller non-standard critical values to check the significance of the response coefficients. However, this approach will not work here as besides lagged dependent variable in the above model, (4.20) we also have a lagged independent variable. Furthermore, unit root tests, such as the ADF test, are quite sensitive to specifications, such as the choice of lags and estimation period. Hence, in this chapter we check the significance of the coefficients against the standard critical values for a 5% significance level, ± 1.96.

In the surplus model for pooled data, (4.16), the lagged surplus, λ_1, is negative and the lagged debt-GDP ratio coefficient, λ_2, is positive, as expected, but the λ_2 coefficient is insignificant. This suggests that surplus is stabilising on its own lagged values, but not on the lagged debt-GDP ratio.

In the revenue model, (4.17), all coefficients are significant. Coefficients λ_{11} is negative as expected since it means that when lagged surplus increases, the government can reduces the current revenue and spend more. Meanwhile, the negative sign of the λ_{12} is not what we would expect, as it suggests that when debt increases the government cuts its revenue instead of increasing it to meet higher interest payments. In the expenditure model, (4.18), the λ_{21} is positive and λ_{22} is negative, which is what we would expect as it suggests that the government increases expenditure in response to increase in surplus or decrease in debt-GDP ratio. However, the feedback from lagged surplus is insignificant.

As for the lagged expenditure and income coefficients, in both equations they are significantly different from zero, but very small. Also, taking into account a large number of observations, 2240, we proceed by assuming that the model can be estimated with surplus and debt-GDP ratio variables as they are, without expressing surplus as a share of GDP. In addition, we assume the effects of revenue and expenditure on the dependent variables are not different, and only thr difference between them is of the main importance.

Hence, the model above (4.16) can be simplified further. Starting with (4.17) and (4.18), the difference between ΔlnR_{it} and ΔlnE_{it} can be estimated as follows:

$$\Delta lnR_{it} - \Delta lnE_{it} = (\mu_1 - \mu_2) + (\lambda_{11} - \lambda_{21})lnS_{it-1} + (\lambda_{12} - \lambda_{22})lnb_{it-1} + (z_{1it} - z_{2it}), \quad (4.19)$$

or alternatively:

$$\Delta lnS_{it} = \alpha^s + \beta^s lnS_{it-1} + \gamma^s lnb_{it-1} + \varepsilon_{it}^s, \quad (4.20)$$

where $\alpha^s = \mu_1 - \mu_2$, $\beta^s = \lambda_{11} - \lambda_{21}$, $\gamma^s = \lambda_{12} - \lambda_{22}$ and $\varepsilon_{it} = z_{1it} - z_{2it}$.

That is what we proceed to estimate. However, we acknowledge that as in case with the balance of payments, besides lagged surplus and stock of debt there are other variables, z_{it}, that should be in (4.20) that explain change in surplus. These variables are likely to include macroeconomic variables, such as inflation rate, GDP growth and interest rate. Hence, (4.20) would become

$$\Delta S_{it} = \alpha + \beta S_{it-1} + \gamma b_{it-1} + \delta z_{it} + \varepsilon_{it}, \quad (4.21)$$

where z_{it} accounts for other potential explanatory variables. We can measure the effect of the omitted variables, z_{it}, as

$$z_{it} = a_1 + b_1 S_{it-1} + c_1 b_{it-1} + w_{it}. \quad (4.22)$$

However, putting (4.22) in (4.21) gets us

$$\Delta S_{it} = \alpha + \beta S_{it-1} + \gamma b_{it-1} + \delta(a_1 + b_1 S_{it-1} + c_1 b_{it-1} + w_{it}) + \varepsilon_{it}, \quad (4.23)$$

which is the same as

$$\Delta S_{it} = (\alpha + \delta a_1) + (\beta + \delta b_1)S_{it-1} + (\gamma + \delta c_1)b_{it-1} + (\delta w_{it} + \varepsilon_{it}). \quad (4.24)$$

This means that when we estimate (4.20),

$$\Delta lnS_{it} = \alpha^s + \beta^s lnS_{it-1} + \gamma^s lnb_{it-1} + \varepsilon^s_{it},$$

the coefficients α^s, β^s and γ^s pick up the average effect of the relationship between z_{it} and other variables over the period, although we acknowledge that these relationships can change over time.

When we estimate (4.20) we expect β^s to be negative, meaning that if there was a surplus in the previous year, the government could increase expenditure or reduce revenue by cutting taxes, and adjust its current surplus to remove the excess. The second coefficient, γ^s is expected to be positive, if there is an increase in the government debt, the government needs to reduce spending or increase taxes to pay for the increased debt.

In addition, we estimate the following, similar to (4.20), equations for change in revenue:

$$\Delta lnR_{it} = \alpha^r + \beta^r lnS_{it-1} + \gamma^r lnb_{it-1} + \varepsilon^r_{it}, \tag{4.25}$$

and for change in expenditure:

$$\Delta lnE_{it} = \alpha^e + \beta^e lnS_{it-1} + \gamma^e lnb_{it-1} + \varepsilon^e_{it}. \tag{4.26}$$

In (4.25) we expect $\beta^r < 0$, as in case of positive lagged surplus, the government can afford to reduce current revenue. Meanwhile, we expect $\gamma^r > 0$, as the government needs to raise current revenue to cover increasing debt-GDP ratio. Following the same logic, in (4.26) we expect $\beta^e > 0$ and $\gamma^e < 0$.

Moreover, we also try estimate an augmented version of (4.20) that includes three key macroeconomic variables (GDP growth, GDP_{it}, long-term interest rate, LR_{it}, and inflation, INF_{it}) as additional regressors. Preliminary analysis showed that there is barely any feedback coming from the lagged debt-GDP ratio when we focus on individual countries or shorter estimation periods. The three key macroeconomic variables are likely to be correlated with the debt-GDP ratio and, hence, might alter the significance of this coefficient. The data for GDP growth, long-term interest rate and inflation also come from the Jordà-Schularick-Taylor Macrohistory Database, as the rest of the data we use in the analysis in this chapter.

We try to estimate two versions of the extended model for the pooled panel data with fixed country effects. First version is with change in the surplus-GDP ratio ($s_{it} = (R_{it} - E_{it})/Y_{it}$), which is the dependent variable, as well as lagged surplus-GDP ratio, lagged debt-GDP ratio and three additional macroeconomic variables:

$$\Delta s_{it} = \alpha + \beta^s s_{it-1} + \gamma^s b_{it-1} + \mu^s GDP_{it} + \nu^s LR_{it} + \eta^s INF_{it} + \varepsilon_{1it}. \tag{4.27}$$

The second version is with logged change in surplus ($lnS_{it} = ln(R_{it}/E_{it})$), which is the dependent variable, as well as logarithm of lagged surplus and lagged debt-GDP ratio as well as three additional macroeconomic variables:

$$\Delta lnS_{it} = \alpha + \beta^s lnS_{it-1} + \gamma^s lnb_{it-1} + \mu^s GDP_{it} + \nu^s LR_{it} + \eta^s INF_{it} + \varepsilon_{2it}. \tag{4.28}$$

Income growth, long-term interest rate and inflation rate are already in proportionate form, hence, we do not need to take logarithmic transformations of these three variables for the second version of the model.

In addition, as in case with the balance of payments analysis, we check for asymmetric adjustment. It is possible that when a government runs a deficit it may be under more pressure to adjust than when it runs surplus. To test this we check whether the surplus responds differently to lagged surplus or lagged deficit. Therefore, we use a dummy variable, $defdum$, that equals one when surplus (S_{it}) is negative and zero otherwise. We then create a new variable $ldefpos_{it}$, which is a product of the dummy, $defdum$, and logarithm of surplus, ln S_{it}. We add $ldefpos_{it}$ to the equation (4.20) as another explanatory variable (we use lagged version of it, $ldefpos_{it-1}$) and estimate:

$$\Delta lnS_{it} = \alpha + \beta^s lnS_{it-1} + \gamma^s lnb_{it-1} + \mu^s ldefpos_{it-1} + \varepsilon_{it}. \tag{4.29}$$

Then if the coefficient of this new variable, $ldefpos_{it-1}$, is significant, we can conclude that the surplus adjusts to lagged deficit differently than to lagged surplus.

Finally, we consider whether balance of payments and fiscal sustainability are related. This part of the analysis is based on the twin deficit hypothesis which suggests that fiscal shocks that have negative effect on the government budget also deteriorate the current account (Corsetti and Müller, 2006).

Following Corsetti and Müller (2006), we start with the identity that links balance of payments, BOT_{it}, and government surplus, S_{it}, together. We derive this identity from the GDP equation:

$$Y = C + I + E + (X - M), \tag{4.30}$$

where Y is GDP, C is consumption, I is investment, E is government expenditure, X is exports and M is imports.

It follows, that the disposable (after tax) income can be defined as:

$$Y - R = C + I + E - R + X - M, \tag{4.31}$$

where R is government revenue. We rearrange (4.31) to get

$$(Y-R-C)-I=(E-R)+(X-M).$$

Then the difference between private saving, S, and investment equals the sum of government deficit and current account (balance of trade) surplus:

$$(S-I)=(E-R)+(X-M). \tag{4.32}$$

From (4.32) it follows that net private savings plus government surplus and current account (balance of trade) deficit sum to zero:

$$(S-I)+(R-E)+(M-X)=0,$$

where $(S-I)$ is net private savings, $(R-E)$ is government surplus and $(M-X)$ is current account deficit.

Hence, to analyse the link between balance of payments and fiscal sustainability we start from the following logarithmic equations for the government surplus:

$$lnS_{it}=ln(\frac{R_{it}}{Y_{it}})-ln(\frac{E_{it}}{Y_{it}})=lnR_{it}-lnE_{it},$$

and for the balance of trade:

$$lnBOT_{it}=lnX_{it}-lnM_{it}.$$

We estimate the equations that not only include a lagged dependent variable, but also measure the effect that one surplus has on the other one. Hence, for the public surplus we estimate:

$$\Delta lnS_{it}=a_{10}+a_{11}lnS_{it-1}+a_{12}lnBOT_{it-1}+\nu_{1it}, \tag{4.33}$$

and for the balance of trade:

$$\Delta lnBOT_{it}=a_{20}+a_{21}lnS_{it-1}+a_{22}lnBOT_{it-1}+\nu_{2it}, \tag{4.34}$$

where lnS_{it} and $lnBOT_{it}$ are logged government surplus and balance of trade, respectively, and ν_{1it} and ν_{2it} are the error terms.

We then discus the cross-surplus adjustments which are measured by the coefficients a_{12} and a_{21} and provide an overall summary for the empirical part of this chapter.

4.5.1 Pooled Estimations

Starting from the surplus equation, we obtain the following estimates for (4.20) using pooled data with fixed country effects:

$$\underset{}{\Delta \ln S_t} = \underset{(0.005)}{-0.007} - \underset{(0.012)}{0.191 \ln S_{t-1}} + \underset{(0.004)}{0.014 \ln b_{t-1}} + \varepsilon_t \qquad \begin{array}{l} R^2 = 0.107 \\ SER = 0.128 \end{array}$$

Both coefficients have the correct signs and both are significant. The lagged surplus coefficient is -0.19, suggesting that when previous period surplus increases by 1%, the government tends to decrease the current surplus by 0.19%. This can be done by reducing the revenue through cutting the taxes, hence bringing the surplus closer to zero. The coefficient of the lagged debt-GDP ratio is 0.01, thus as debt-GDP increases by one 1%, the government raises the surplus by 0.01% to meet increasing interest payments on debt, brining debt-GDP ratio closer to zero.

Moving to the components of surplus, revenue and expenditure, after estimating the equations (4.25) and (4.26) for pooled data we obtain the following results:

$$\Delta \ln R_t = \underset{(0.004)}{0.031} - \underset{(0.011)}{0.195 \ln S_{t-1}} - \underset{(0.004)}{0.024 \ln b_{t-1}} + \varepsilon_t \qquad \begin{array}{l} R^2 = 0.141 \\ SER = 0.116 \end{array}$$

$$\Delta \ln E_t = \underset{(0.006)}{0.037} - \underset{(0.015)}{0.004 \ln S_{t-1}} - \underset{(0.005)}{0.038 \ln b_{t-1}} + \varepsilon_t \qquad \begin{array}{l} R^2 = 0.040 \\ SER = 0.150 \end{array}$$

In the revenue regression (4.25) for the pooled data both coefficients are significant. The current revenue adjusts to the lagged surplus in the correct way, but moves in the wrong direction in response to change in the lagged debt-GDP ratio. The lagged surplus coefficient suggests that if previous surplus increased by 1%, the government would reduce the current revenue by 0.20%, brining the surplus closer to zero.

As for (4.26), the lagged surplus coefficient is insignificant and has the wrong sign. Nonetheless, the coefficient of the debt-GDP ratio is significant and negative, as expected, suggesting that as debt-GDP ratio increases by 1%, the government reduces spending by 0.04%. Hence, there is some evidence of stabilisation of the debt through expenditure.

Thus, overall, our estimations suggest that stabilisation of surplus on its lagged value is mostly achieved through adjustment in revenue. Mountford and Uhlig (2009) argue that deficit-financed tax cuts do indeed have better effect on growth than alternative fiscal policy tools, such as deficit-spending or balanced budget spending expansion, and are less costly. Mountford and Uhlig (2009) analysed the effect of the government revenue and expenditure shocks, controlling for business cycle and monetary policy shocks, using US quarterly data for 1955-2000. They did it by estimating ten-variable vector autoregressive model with six lags and analysing impulse response functions.

They found that the basic government revenue shock (with revenue rising for a year after the shock) had negative effect on GDP, consumption, investment and, surprisingly, reserves, but positive effect on interest rate. Meanwhile, the effect of a basic government spending shock on output and consumption was rather weak, which is also consistent with our finding that most of the adjustment of public surplus to its lagged values is done through revenue, not expenditure.

Mountford and Uhlig (2009) also argue that the long-run costs of the deficit-financed expansionary fiscal policy were likely to outweigh the benefits of its initial boosting effect on the economy. They calculated that a 2% government spending increase would eventually translate into an over 2% increase in taxes and more than 7% decrease in GDP. Hence, they concluded that while deficit-financed tax cuts appeared to be best fiscal policy out of the three scenarios they considered, the short-term stimulating effect to the economy was likely to come at a high long-run price. Therefore, it appears that increasing government deficit and debt might have potentially very dangerous consequences for the economy.

Above we estimated models for pooled data for the whole period. However, pooling assumes that the parameters are the same in each country, which is implausible. Thus, we now estimate a separate equation for each country and sub-period. Since, in this case we have fewer observations, we adopt a simpler model with fewer parameters: (4.20) for surplus, (4.25) for revenue and (4.26) for expenditure.

4.6 Country-Specific Estimates

We now estimate the equations for surplus (4.20), revenue (4.25) and expenditure (4.26) for each country individually over 1870-2016. In addition, we split sample in three sub-periods in order to analyse the adjustment processes of surplus and debt-GDP ratio.

The three sub-periods are 1870-1914, 1915-1950 and 1951-2016. The decision to split estimation period in these sub-periods is based on two considerations. Firstly, we split the data roughly around the historical breaks, World War I and World War II, making first sub-period a pre-World War I period, third sub-period is a post-World War II period. Meanwhile, second sub-period encompasses highly disturbed time with two world wars and depression in-between. Secondly, we aim to have three sub-periods of a roughly equal sizes in order to have sufficient data points to estimate three models (expenditure, revenue and surplus) for most countries in the sample. Finland is the only country for which we do not a sufficient number of observations to estimate any of the three equations over the first sub-period, 1870-1914. For all other countries the estimations for the whole period and three sub-periods are presented and discussed below.

4.6.1 Surplus

When we estimate (4.20) for each country individually for the whole period the picture looks somewhat different comparing to the pooled data estimates.

The heterogeneous estimates for (4.20),

$$\Delta lnS_{it} = \alpha^s + \beta^s lnS_{it-1} + \gamma^s lnb_{it-1} + \varepsilon_{it}^s,$$

for 1870-2016 and three sub-periods are summarised in the Table 4.2. As was mentioned before, in the surplus equation (4.20) we expect $\beta^s < 0$ and $\gamma^s > 0$.

Table 4.2 Government Debt Solvency: Surplus (Equation (4.20); 1870-2016 and Sub-Periods)

	1870-1914				**1915-1950**				**1951-2016**				**1870-2016**			
Country	β^s	**t-stat.**[1]	γ^s	**t-stat.**	β^s	**t-stat.**	γ^s	**t-stat.**	β^s	**t-stat.**	γ^s	**t-stat.**	β^s	**t-stat.**	γ^s	**t-stat.**
Australia	0.235	0.790	-0.254	-0.442	**-0.253**	**-2.703**	**0.406**	**2.500**	**-0.362**	**-3.978**	-0.009	-0.387	**-0.109**	**-3.144**	-0.002	0.934
Belgium	**-0.723**	**-4.914**	0.067	1.835	**-0.490**	**-3.084**	0.042	0.228	**-0.207**	**-3.244**	**0.032**	**2.261**	**-0.447**	**-7.827**	0.028	1.794
Canada	**-0.587**	**-3.960**	-0.117	-1.238	**-0.215**	**-2.067**	0.173	1.957	**-0.181**	**-2.664**	**0.100**	**2.513**	**-0.242**	**-4.437**	0.038	1.498
Denmark	**-0.669**	**-4.077**	0.366	1.840	**-1.173**	**-5.084**	-0.070	-0.384	**-0.247**	**-2.790**	0.004	0.468	**-0.576**	**-6.782**	0.021	1.015
Finland	X	X	X	X	-0.310	-1.844	0.082	1.109	**-0.132**	**-2.046**	0.006	0.464	**-0.268**	**-3.630**	0.039	1.603
France	-0.139	-0.412	**0.402**	**2.572**	-0.325	-1.446	0.124	0.832	**-0.314**	**-3.117**	**-0.037**	**-2.257**	**-0.292**	**-3.281**	-0.021	-1.246
Germany	**-0.603**	**-3.763**	0.112	1.735	0.016	0.039	-0.025	-0.238	**-0.526**	**-6.594**	-0.002	-0.140	**-0.213**	**-3.731**	0.025	1.280
Italy	-0.400	-1.218	**0.279**	**2.296**	-0.164	-1.674	-0.044	-0.444	-0.101	-1.804	0.016	1.134	**-0.125**	**-3.093**	0.005	0.207
Japan	**-0.266**	**-2.251**	0.003	0.124	-0.071	-0.938	0.003	0.137	-0.094	-1.329	-0.006	-0.499	-0.034	-1.195	0.001	0.152
Netherlands	**-0.722**	**-3.593**	**0.227**	**2.438**	**-0.419**	**-2.562**	0.139	1.395	**-0.321**	**-3.741**	0.029	1.038	**-0.368**	**-5.197**	0.047	1.742
Norway	**-0.293**	**-2.433**	-0.022	-0.426	**-0.553**	**-4.127**	0.048	0.770	**-0.177**	**-2.795**	0.078	1.891	**-0.204**	**-4.206**	**0.055**	**1.972**
Portugal	**-0.440**	**-4.368**	**-0.347**	**-2.112**	**-0.243**	**-2.244**	-0.088	-1.145	-0.122	-1.756	0.012	0.861	**-0.194**	**-3.962**	0.002	0.128
Spain	**-0.812**	**-4.764**	**0.152**	**2.595**	**-0.427**	**-2.548**	0.154	0.616	**-0.272**	**-2.847**	0.003	0.132	**-0.308**	**-4.841**	0.013	0.902
Sweden	**-0.454**	**-3.704**	**0.161**	**2.228**	**-0.479**	**-2.971**	0.043	0.536	**-0.148**	**-2.044**	0.039	1.864	**-0.325**	**-5.103**	0.028	1.448
Switzerland	0.275	0.772	0.052	0.416	**-0.246**	**-2.659**	0.082	1.329	**-0.398**	**-4.144**	0.010	0.258	**-0.207**	**-3.919**	0.013	1.268
UK	**-0.605**	**-2.883**	0.118	1.815	**-0.417**	**-4.692**	**0.392**	**4.960**	**-0.093**	**-2.081**	0.019	1.382	**-0.128**	**-3.200**	**0.047**	**2.807**
US	**-0.279**	**-2.812**	0.039	1.821	**-0.237**	**-1.997**	0.109	1.792	**-0.214**	**-2.640**	0.034	1.060	**-0.197**	**-3.726**	0.018	1.137

Notes: Significant coefficients are in bold (5% level); [1] t-statistic; X - insufficient number of observations.

Over the whole period only for Norway and the UK the surplus is stabilising on both, its lagged value and lagged debt-GDP, with both coefficients (β^s and γ^s) being significant and having correct (expected) signs. For all countries the lagged surplus coefficient, β^s, is negative, and for all but Japan it is significant, suggesting a stabilising process. However, with the exception of Norway and the UK, there is no evidence of the surplus stabilising on the lagged debt-GDP ratio, as γ^s coefficient is insignificant for all countries, but these two.

When we split the data in the three sub-periods. There is an insufficient number of observations to estimate the model for Finland over the first sub-period, hence, the sample size for the first sub-period is one country short (16 not 17). In the first sub-period both β^s and γ^s are significant and have correct signs only for three countries, namely the Netherlands, Spain and Sweden. For Portugal both lagged coefficients are significant, but the debt-GDP one has the wrong sign, suggesting that as debt increases, government decreases surplus, which would be evidence of destabilising process. The government surplus is stabilising on its lagged value for 12 out of 16 countries and on the lagged debt-GDP ratio for five economies.

Over 1915-1950, the surplus is stabilising on its lagged value for 12 out of 17 countries and on the lagged debt-GDP ratio only for two, Australia and the UK. Hence, once again, there does not seem to be feedback coming from the debt-GDP ratio for majority of countries.

In the third sub-period, for 14 countries the lagged surplus coefficient is significant and negative, as we would expect. For two countries, Belgium and Canada, both lagged surplus and debt-GDP coefficients are significant and have correct signs. However, in case of France, lagged debt-GDP ratio coefficient is significant, but has the wrong sign.

Overall, in case of most countries and over all sub-periods the government is taking corrective measures in response to changes in the lagged surplus, but not debt-GDP. Therefore, there is significantly less evidence of fiscal sustainability when we focus on individual economies instead of using pooled data for a large sample of countries.

In addition, the Table 4.3 presents the estimated equilibrium level of debt-GDP for each of 17 countries in our sample as well as historical average debt-GDP ratios for comparison. We estimate equilibrium level of debt-GDP for each country using fixed cross-country effect estimator of the lagged debt-GDP, which is 0.014, from the estimation of (4.20) in section 4.5.1.

In the model for surplus, (4.20):

$$\Delta lnS_{it} = \alpha^s + \beta^s lnS_{it-1} + \gamma^s lnb_{it-1} + \varepsilon^s_{it},$$

in equilibrium $\Delta S_{it} = S_{it} = 0$ and $b_{it} = b_{it-1}$. Hence, equilibrium level of debt-GDP for each country can be estimated as the exponential of

$$b^* = -\frac{\alpha^s_i}{\gamma^s}, \tag{4.35}$$

where α^s_i is the intercept of a country i, and γ^s is lagged debt-GDP estimate for pooled data from (4.20).

The implied equilibrium debt-GDP ratio (Table 4.3) is close to the average for some countries, such as Switzerland and Norway, but is very different for others, for instance, Italy and Portugal. It is not clear why this is. It may be because the country-specific coefficients are very different from the pooled estimates. For example, in the case of Italy the coefficient on lagged debt is 0.005 (from the Table 4.2), roughly a third of the pooled estimate of 0.014.

Table 4.3 Equilibrium Debt-GDP Levels (Equation (4.35))

Country	α_i^s	$(\gamma^s)^1$	b^*	Equilibrium Debt-GDP[2]	Hist. Aver. Debt-GDP[3]
Australia	-0.007	0.014	0.47	1.61	0.40
Belgium	-0.015	0.014	1.07	2.92	0.71
Canada	-0.014	0.014	1.00	2.71	0.63
Denmark	0.009	0.014	-0.66	0.52	0.76
Finland	-0.002	0.014	0.13	1.14	0.86
France	-0.013	0.014	0.98	2.65	0.61
Germany	0.014	0.014	-0.99	0.37	0.25
Italy	-0.033	0.014	2.44	11.42	0.28
Japan	0.017	0.014	-1.21	0.30	0.51
Netherlands	-0.002	0.014	0.17	1.18	0.60
Norway	0.021	0.014	-1.50	0.22	0.28
Portugal	-0.014	0.014	0.99	2.70	0.32
Spain	-0.001	0.014	0.06	1.06	0.42
Sweden	0.007	0.014	-0.54	0.59	0.90
Switzerland	0.019	0.014	-1.39	0.25	0.27
UK	0.012	0.014	-0.85	0.43	0.79
USA	0.005	0.014	-0.40	0.67	0.40

Notes: [1] lagged debt-GDP coefficient from estimation of (4.20) using pooled data; [2] Equilibrium Debt-GDP − is an exponential of (b^*); [3] Hist. Aver. Debt-GDP − historical average debt-GDP are averages for each country calculated using their debt-GDP series from the JST Dataset.

4.6.2 Revenue

We now estimate the revenue model, (4.25), for each country individually over the whole period and three sub-periods:

$$\Delta lnR_{it} = \alpha^r + \beta^r lnS_{it-1} + \gamma^r lnb_{it-1} + \varepsilon_{it}^r.$$

As was mentioned before, a stabilising process would imply $\beta^r < 0$ while $\gamma^r > 0$. The results are presented in the Table 4.4.

When we consider all countries individually over the whole period, the lagged surplus coefficient is significant and negative in all but four countries (Denmark, Germany, Japan and Norway), suggesting that revenue is stabilising on surplus lagged values. For none of the countries debt-GDP ratio is stabilising, as for most its coefficient is insignificant and for five countries, for which it is significant, the sign is wrong. The lagged debt-GDP ratio coefficient has the correct sign for Belgium, Germany, Sweden and Switzerland, but is insignificant.

Table 4.4 Government Debt Solvency: Revenue (Equation (4.25); 1870-2016 and Sub-Periods)

	1870-1914				**1915-1950**				**1951-2016**				**1870-2016**			
Country	β^r	**t-stat.**[1]	γ^r	**t-stat.**	β^r	**t-stat.**	γ^r	**t-stat.**	β^r	**t-stat.**	γ^r	**t-stat.**	β^r	**t-stat.**	γ^r	**t-stat.**
Australia	-0.073	-0.595	-0.011	-0.045	**-0.251**	**-8.018**	**0.113**	**2.089**	**-0.354**	**-3.052**	-0.032	-1.138	**-0.106**	**-5.532**	-0.016	-1.168
Belgium	-0.023	-0.510	-0.003	-0.291	-0.139	-1.679	**0.267**	**2.798**	**-0.162**	**-2.617**	**-0.069**	**-4.956**	**-0.174**	**-4.659**	0.009	0.374
Canada	-0.117	-1.251	**-0.139**	**-2.348**	**-0.233**	**-2.884**	0.000	-0.004	-0.126	-1.654	-0.082	-1.842	**-0.149**	**-3.561**	-0.006	-0.317
Denmark	-0.116	-1.670	0.048	0.572	-0.551	-1.787	-0.267	-1.100	**-0.259**	**-3.059**	**-0.054**	**-6.395**	-0.148	-1.919	-0.034	-1.764
Finland	X	X	X	X	**-0.432**	**-3.111**	0.036	0.592	-0.094	-1.455	**-0.057**	**-4.347**	**-0.422**	**-6.435**	-0.029	-1.344
France	-0.052	-0.487	-0.004	-0.075	-0.196	-1.294	**-0.202**	**-2.012**	**-0.272**	**-2.565**	**-0.106**	**-6.218**	**-0.156**	**-2.735**	**-0.057**	**-5.436**
Germany	**-0.697**	**-2.565**	**0.403**	**3.650**	**-1.380**	**-2.586**	0.060	0.439	-0.129	-1.777	**-0.048**	**-3.639**	-0.141	-1.570	0.025	0.815
Italy	-0.229	-1.075	-0.023	-0.298	**-0.403**	**-5.572**	-0.137	-1.881	**-0.406**	**-7.716**	**-0.068**	**-5.076**	**-0.367**	**-12.202**	**-0.075**	**-3.966**
Japan	0.088	0.321	0.034	0.547	-0.238	-0.720	0.028	0.339	**-0.206**	**-2.455**	**-0.067**	**-4.707**	0.069	1.039	-0.011	-0.676
Netherlands	-0.105	-0.568	-0.072	-0.838	**-0.421**	**-4.056**	0.080	1.270	0.061	0.675	-0.045	-1.560	**-0.271**	**-4.792**	-0.019	-0.877
Norway	-0.169	-0.817	0.007	0.080	0.084	0.483	**-0.291**	**-3.569**	**-0.179**	**-2.830**	0.019	0.461	0.041	0.638	**-0.076**	**-2.079**
Portugal	-0.106	-1.103	-0.123	-0.784	**-0.380**	**-3.839**	0.031	0.450	**-0.590**	**-9.357**	**-0.088**	**-6.708**	**-0.430**	**-9.294**	**-0.073**	**-4.538**
Spain	**-0.617**	**-3.655**	**0.134**	**2.316**	-0.022	-0.257	-0.129	-1.003	-0.114	-1.563	**-0.098**	**-6.079**	**-0.094**	**-2.073**	**-0.087**	**-8.383**
Sweden	-0.075	-0.797	-0.025	-0.443	0.066	0.590	0.048	0.866	**-0.336**	**-3.526**	-0.039	-1.418	**-0.117**	**-2.085**	0.018	1.058
Switzerland	0.085	0.586	0.005	0.091	**-0.327**	**-3.705**	-0.002	-0.031	-0.123	-1.324	**-0.073**	**-2.050**	**-0.273**	**-6.382**	0.000	-0.021
UK	0.058	0.459	**-0.091**	**-2.321**	-0.076	-1.158	**-0.153**	**-2.647**	**0.171**	**4.497**	**-0.025**	**-2.155**	**-0.092**	**-3.234**	-0.020	-1.684
US	**-0.187**	**-2.020**	0.003	0.171	**-0.357**	**-4.158**	**-0.100**	**-2.255**	-0.054	-0.750	-0.021	-0.738	**-0.291**	**-7.227**	-0.023	-1.888

Notes: Significant coefficients are in bold (5% level); [1] t-statistic; X - insufficient number of observations.

For the first sub-period most coefficients are insignificant. There is insufficient number of observations for Finland, hence, the sample is reduced to 16 countries. For two countries, Germany and Spain, both β^r and γ^r are significant and have correct signs. The lagged surplus coefficient for the US has the correct sign and is significant, but the lagged debt-GDP ratio one is insignificant. For the UK the lagged debt-GDP ratio coefficient is significant, but has the wrong sign. It is positive, as expected, for Denmark, Germany, Japan, Norway, Spain, Switzerland and the US, but is only significant for Germany and Spain.

In the second sub-period, revenue is stabilising on both lagged coefficients only for Australia. In addition, the lagged surplus coefficient is significant and has the correct sign for nine countries, namely Australia, Canada, Finland, Germany, Italy, the Netherlands, Portugal, Switzerland and the US. There is a significant feedback from the lagged debt-GDP for two economies, Australia and Belgium, while for France, Norway, the UK and US the coefficient is significant, but has the wrong sign.

Finally, in the third sub-period, for nine out of 17 countries the revenue is stabilising on the lagged surplus. The lagged debt-GDP coefficient is significant for 11 countries, but has the wrong sign.

Hence, it appears estimations of (4.25) for the whole sample and for individual countries are consistent in that revenue is stabilising on the lagged surplus, but not on the lagged debt-GDP ratio.

4.6.3 Expenditure

Finally, we estimate the expenditure model, (4.26), for each country over the whole period and three sub-periods:

$$\Delta lnE_{it} = \alpha^e + \beta^e lnS_{it-1} + \gamma^e lnb_{it-1} + \varepsilon_{it}^e,$$

where we expect $\beta^e > 0$ and $\gamma^e < 0$.

The estimates for the individual countries for the whole period (Table 4.5) show that expenditure is stabilising on the lagged surplus and debt-GDP for three countries, namely Denmark, Norway and Spain. In case of two economies, Italy and Portugal, both coefficients are significant, but the lagged surplus one has the wrong sign. Besides Denmark, Norway and Spain, expenditure is stabilising on the lagged surplus for Belgium and Sweden, but the lagged debt-GDP coefficients for these countries are insignificant. However, there is a stabilising feedback on the lagged debt-GDP ratio for nine out of 17 countries (Denmark, Finland, France, Italy, the Netherlands, Norway, Portugal, Spain and the UK).

Table 4.5 Government Debt Solvency: Expenditure (Equation (4.26); 1870-2016 and Sub-Periods)

	1870-1914				**1915-1950**				**1951-2016**				**1870-2016**			
Country	β^e	**t-stat.**[1]	γ^e	**t-stat.**	β^e	**t-stat.**	γ^e	**t-stat.**	β^e	**t-stat.**	γ^e	**t-stat.**	β^e	**t-stat.**	γ^e	**t-stat.**
Australia	-0.309	-1.137	0.244	0.464	0.002	0.020	-0.292	-1.627	0.008	0.077	-0.023	-0.900	0.003	-0.014	0.091	-0.538
Belgium	**0.700**	**4.735**	-0.070	-1.918	**0.339**	**2.076**	0.199	1.036	0.044	0.660	**-0.101**	**-6.717**	**0.264**	**4.243**	-0.019	-1.069
Canada	**0.471**	**3.592**	-0.022	-0.267	-0.019	-0.146	-0.174	-1.574	0.055	1.086	**-0.182**	**-6.149**	0.093	1.566	-0.044	-1.606
Denmark	**0.553**	**3.051**	-0.318	-1.447	**0.622**	**2.316**	-0.197	-0.933	-0.012	-0.143	**-0.058**	**-7.101**	**0.428**	**4.924**	**-0.055**	**-2.557**
Finland	X	X	X	X	-0.122	-0.624	-0.046	-0.536	0.039	0.700	**-0.063**	**-5.639**	-0.154	-1.858	**-0.068**	**-2.498**
France	0.087	0.329	**-0.406**	**-3.302**	0.129	0.604	**-0.326**	**-2.304**	0.043	0.370	**-0.069**	**-3.734**	0.136	1.523	**-0.037**	**-2.221**
Germany	0.153	1.025	**0.209**	**3.488**	-1.396	-1.627	0.085	0.385	**0.397**	**5.632**	**-0.046**	**-3.596**	0.077	1.075	-0.009	-0.392
Italy	0.172	0.570	**-0.302**	**-2.720**	**-0.238**	**-2.159**	-0.093	-0.837	**-0.305**	**-4.413**	**-0.084**	**-4.795**	**-0.242**	**-5.481**	**-0.080**	**-2.891**
Japan	0.354	1.291	0.031	0.490	-0.167	-0.512	0.026	0.311	**-0.112**	**-2.372**	**-0.061**	**-7.616**	0.104	1.632	-0.012	-0.779
Netherlands	**0.617**	**2.282**	**-0.298**	**-2.384**	-0.002	-0.013	-0.059	-0.563	**0.382**	**4.620**	**-0.074**	**-2.775**	0.096	1.241	**-0.065**	**-2.229**
Norway	0.123	0.603	0.030	0.332	**0.637**	**5.904**	**-0.340**	**-6.722**	-0.002	-0.056	**-0.059**	**-2.237**	**0.245**	**4.514**	**-0.131**	**-4.204**
Portugal	**0.334**	**2.802**	0.224	1.152	-0.136	-1.106	0.119	1.371	**-0.468**	**-6.956**	**-0.100**	**-7.167**	**-0.236**	**-3.958**	**-0.075**	**-3.625**
Spain	0.196	1.585	-0.018	-0.419	**0.404**	**2.376**	-0.283	-1.116	0.159	1.926	**-0.101**	**-5.523**	**0.214**	**3.637**	**-0.100**	**-7.390**
Sweden	**0.378**	**4.621**	**-0.186**	**-3.844**	**0.545**	**2.645**	0.005	0.048	**-0.187**	**-2.803**	**-0.077**	**-4.043**	**0.208**	**2.734**	-0.010	-0.431
Switzerland	-0.190	-0.753	-0.047	-0.536	-0.081	-1.016	-0.084	-1.578	**0.276**	**3.164**	**-0.083**	**-2.463**	-0.066	-1.490	-0.013	-1.531
UK	**0.663**	**2.424**	**-0.209**	**-2.467**	**0.342**	**2.755**	**-0.545**	**-4.950**	**0.265**	**7.512**	**-0.045**	**-4.091**	0.036	0.666	**-0.067**	**-2.948**
US	0.092	1.014	-0.036	-1.824	-0.120	-0.748	**-0.209**	**-2.533**	**0.160**	**2.771**	**-0.054**	**-2.408**	-0.094	-1.346	-0.041	-1.959

Notes: Significant coefficients are in bold (5% level); [1] t-statistic; X - insufficient number of observations.

In the first sub-period the lagged debt-GDP ratio is stabilising for five economies, namely France, Italy, the Netherlands, Sweden and the UK. In case of Germany this coefficient is significant, but has the wrong sign. Expenditure is stabilising on the lagged surplus for seven countries, namely Belgium, Canada, Denmark, the Netherlands, Portugal, Sweden and the UK.

In the second sub-period, the debt-GDP ratio lagged coefficient is significant for four countries, France, Norway, the UK and US, and in all of them it has the correct sign, suggesting a stabilising process. The lagged surplus coefficient is significant for Belgium, Denmark, Italy, Norway, Spain, Sweden and the UK, but for Italy it has the wrong sign.

In the final sub-period, expenditure is stabilising on both lagged coefficients for five countries, Germany, the Netherlands, Switzerland, the UK and US. The lagged surplus coefficient is also significant for Italy, Japan, Portugal and Sweden, but has the wrong sign. Nonetheless, the debt-GDP ratio is stabilising for all but one country (Australia). Overall, it seems that the adjustment to the lagged surplus is mainly done through revenue and, when there is in fact an adjustment to the lagged debt-GDP, it is mainly done through expenditure.

4.6.4 Summary

In addition, for each of the three equations, (4.20), (4.25) and (4.26), the Table 4.6 summarises the number of countries for which coefficients of the lagged surplus (β) and lagged debt-GDP ratio (γ) have the *wrong* signs (signs that suggest destabilisation). Hence, the smaller the number, the more there is evidence of stabilisation, with the ideal outcome being zero.

Table 4.6 Destabilised Cases: Out of 17 Countries (Equations (4.20), (4.25) and (4.26); 1870-2016 and Sub-Periods)

	1870-1914		**1915-1950**		**1951-2016**		**1870-2016**	
	β	γ	β	γ	β	γ	β	γ
Surplus - Equation (4.20)	2	4	1	4	0	4	0	2
Revenue - Equation (4.25)	3	9	2	8	2	16	2	13
Expenditure - Equation (4.26)	2	5	9	5	6	0	5	1

Note: the numbers represent the count of countries for which the coefficient has a *wrong* sign. β - coefficient for lnS_{it}; γ - coefficient for lnb_{it}. Expected signs: surplus (4.20): $\beta^s < 0$, $\gamma^s > 0$; revenue (4.25): $\beta^r < 0$, $\gamma^r > 0$; expenditure (4.26): $\beta^e > 0$, $\gamma^e < 0$.

It appears that surplus equation produces the most sensible results in terms of signs. The surplus lagged coefficient has the correct sign for all countries when (4.20) is estimated over the whole period and third sub-period. There is only one and two wrong signs in the second and first sub-periods, correspondingly. The count for the wrong signs for the γ coefficient is slightly higher. It appears that for four countries the lagged debt-GDP ratio has the wrong sign over the three sub-periods, but only for two economies over the whole period. The revenue equation, however, produces γ coefficients with the wrong sign for all but one country when it is estimated over the third sub-period, for 13 economies over the whole period and for nine and eight over the first and second sub-periods, respectively. In case of the expenditure equation, however, all γ coefficients have the correct signs in the third sub-period and only one wrong sign when the equation is estimated over the whole period.

Overall, there are more lagged debt-GDP ratio coefficients with the wrong sign than the lagged surplus one. It is worth to note that $\beta^s = \beta^r - \beta^e$, as for instance in case of Italy for the whole period we have $-0.125 = -0.367 - (-0.242)$. Therefore, if for instance β^r for a country i is destabilising, but β^r is stabilising and big enough, then the overall lagged surplus, β^s, coefficient will be stabilising. The same is true if β^s is destabilising, but β^r is stabilising and big enough. Hence, for instance in case of β estimates for the whole sample, there are two destabilising β^r and five β^s, but there are no β^s coefficients with the wrong sign, thus this condition must hold, meaning the stabilising coefficients can be large enough to offset the effect of destabilising coefficients (with the wrong sign).

By analogy, $\gamma^s = \gamma^r - \gamma^e$, and large enough stabilising coefficients can offset effect of destabilising ones resulting in overall stabilising feedback. In case of the γ coefficient for the whole period, it has the wrong sign for 13 countries in the revenue equation, but only for one in the expenditure equation. Hence, we can conclude that revenue is not responding to changes in the lagged debt-GDP ratio in the right way, but the expenditure is, and the expenditure effect means that overall the surplus adjusts to debt-GDP in the right direction in all but two countries.

However, the number of observations for revenue and expenditure equations do not match exactly for some countries. Hence, these two conditions for β ($\beta^s = \beta^r - \beta^e$) and γ ($\gamma^s = \gamma^r - \gamma^e$) may not hold exactly in some cases. Also, these two conditions do not account for the significance of the coefficients. Nonetheless, overall, heterogeneous estimates suggest that surplus adjusts in the right way most of the time. In fact, over the whole period it does so in all cases with respect to the lagged surplus and in all but two countries with respect to debt, providing some evidence of fiscal sustainability.

4.7 Extended Model: Growth Rate, Inflation and Long Interest Rate

Since pooled estimates for surplus model, (4.20), show that there is significant feedback coming from both, lagged surplus and debt-GDP, but heterogeneous results suggest that there is little stabilising feedback from lagged debt-GDP, we also try to estimate an augmented version of (4.20) with three key macroeconomic variables (GDP growth, GDP_{it}, long-term interest rate, LR_{it}, and inflation, INF_{it}) as additional regressors. We check whether inclusion of these additional three variables, which are likely to be correlated with the debt-GDP ratio, will help to establish significant feedback from lagged debt-GDP.

We estimate and compare two versions of the extended model for the pooled panel data with fixed country effects. First version is for the change in the surplus expressed as a ratio of GDP, (4.27):

$$\Delta s_{it} = \alpha + \beta^s s_{it-1} + \gamma^s b_{it-1} + \mu^s GDP_{it} + \nu^s LR_{it} + \eta^s INF_{it} + \varepsilon_{1it},$$

and the second one is a log version of (4.27), the equation (4.28):

$$\Delta lnS_{it} = \alpha + \beta^s lnS_{it-1} + \gamma^s lnb_{it-1} + \mu^s GDP_{it} + \nu^s LR_{it} + \eta^s INF_{it} + \varepsilon_{2it}.$$

The coefficients and standard errors (SE; in parentheses) for the first version, (4.27), of the model are:

$$\underset{}{\Delta s_t} = \underset{(0.002)}{-0.006} - \underset{(0.012)}{0.174\, s_{t-1}} + \underset{(0.002)}{0.006\, b_{t-1}} + \underset{(0.001)}{0.001\, GDP_{t-1}} + \underset{(0.001)}{0.000\, LR_{t-1}} - \underset{(0.001)}{0.001\, INF_t} + \varepsilon_t^1 \qquad \begin{matrix} R^2 = 0.109 \\ SER = 0.024 \end{matrix}$$

There is a negative relationship between previous and current surplus-GDP, as we would expect in a stabilising process. The effect of change in lagged debt-GDP ratio is significant and positive, while inflation rate seems to have significant negative effect on the change in surplus-GDP ratio. There is no feedback coming from GDP growth or long interest rate.

For the second version, (4.28) the coefficients and standard errors (in parentheses) are:

$$\Delta \ln S_t = \underset{(0.007)}{-0.013} - \underset{(0.012)}{0.199\, lnS_{t-1}} + \underset{(0.004)}{0.014\, lnb_{t-1}} + \underset{(0.001)}{0.004\, GDP_{t-1}} + \underset{(0.001)}{0.001\, LR_{t-1}} - \underset{(0.001)}{0.001\, INF_t} + \varepsilon_t^2 \qquad \begin{matrix} R^2 = 0.128 \\ SER = 0.125 \end{matrix}$$

Now all coefficients are significant, except long interest rate one. Signs are consistent with the estimates for the first version of the model. Income growth has significant positive effect on the change in the logarithm of surplus, and there is a positive feedback coming from the logarithm of lagged debt-GDP ratio. The R-squared is slightly higher than for the first version (0.128 versus 0.109), but so is the standard error of the regression (SER), 0.125 versus 0.024. Overall, while the linear model is closer to the theory, it appears that the log version fits the data better than the linear one.

Nonetheless, when we estimate the log model for each country individually over the whole period, the logarithm of the debt-GDP coefficient is significant only for two out of 17 economies, Belgium and the UK. In case of the linear version the feedback from the lagged debt-GDP is coming for four out of 17 countries (the Netherlands, Spain, Sweden and the UK).

The mean group estimator for the debt-GDP for the sample is 0.012 (SE is 0.003) for the linear version of the model and 0.024 (SE is 0.006) for the log one. Since in most cases γ^s is insignificant, we cannot reject the null hypothesis that it equals to zero. For the pooled panel data with fixed country effects the γ^s coefficient in the linear version of the model is 0.006 (SE is 0.002) and 0.008 (SE is 0.003) in the log version. In both cases debt-GDP coefficient is significant. When the data for 17 countries are pooled together, the standard error is small, there is a significant feedback coming from debt-GDP ratio in both, linear and log, versions of the model. Hence, overall, it is not clear whether the best estimate for the effect of the lagged debt-GDP on the dependent variable is zero or γ^s.

4.8 Asymmetric Adjustment

By analogy with the balance of payments analysis, we assume that there might be more pressure on governments running deficits to adjust than on those running surpluses. Hence, we test for asymmetric adjustment by estimating (4.29):

$$\Delta lnS_{it} = \alpha + \beta^s lnS_{it-1} + \gamma^s lnb_{it-1} + \mu^s ldefpos_{it-1} + \varepsilon_{it},$$

where if $ldefpos_{it-1}$ is significant, it means surplus adjusts to the lagged deficit differently than to the lagged surplus. For the pooled data for the whole period we obtain the following results:

$$\begin{array}{llllllllll} \Delta \ln S_t & = & -0.017 & - & 0.101 \ln S_{t-1} & + & 0.008 \ln b_{t-1} & - & 0.100\, ldefpos_{t-1} & + \ \varepsilon_t \quad R^2 = 0.100 \\ & & (0.005) & & (0.031) & & (0.003) & & (0.038) & \quad\quad SER = 0.128 \end{array}$$

The adjustments in case of surplus versus deficit are significantly different. If a government is running a surplus ($lnS_{it-1} > 0$) then the speed of adjustment is 10%. However, if a government is running a deficit ($lnS_{it-1} < 0$), then the speed of adjustment is about twice higher, 20% (from $-0.101 + (-0.100)) = -0.201$). From this we can conclude that there is indeed evidence of asymmetric adjustment with there being more pressure on governments running deficits to adjust than on those running surpluses.

Nonetheless, as in case of the asymmetric adjustment analysis for the balance of payments, there is little evidence of asymmetric adjustment if we shorten the estimation period or consider countries individually. In fact if we run this model for the whole period without fixing cross-country effects, it appears that the coefficient of $ldefpos_{it-1}$ is only significantly different from zero in Portugal. Hence, our conclusion that there is evidence of asymmetric adjustment is mostly based on the pooled data estimations for the whole period.

4.9 The Twin Deficit Hypothesis: Balance of Payments and Fiscal Sustainability

The twin deficit hypothesis builds on the idea that fiscal shocks that cause a deterioration in the government budget also worsen the current account (Corsetti and Müller, 2006). This explains why often when a country is running large balance of payments deficit, the fiscal consolidation is seen as one of the most important remedies, meaning worsening of the current account is often associated with an increasing pressure on the government to reduce fiscal deficit and debt.

Traditional explanation for this transmission is based on the Mundell–Fleming model which assumes that government deficit leads to the appreciation of the local currency and, hence, local goods and services become relatively more expensive than foreign ones. As a result, consumption shifts towards imports, meaning net exports are crowded out. However, increase in the government deficit is often associated with increase in interest rates, which can potentially lead to decrease in domestic investment, thus somewhat mitigating the effect of appreciation. This model assumes flexible exchange rates, exogeneity of the rate of return to investment and largely ignores intertemporal consumption smoothing. In contrast Corsetti and Müller (2006) opted for a general equilibrium model and a VAR framework. Focusing on Australia, Canada, the UK and US they found support for the twin deficit hypothesis. Their estimations of the structural VAR models provided support for the twin deficit hypothesis. They argue that the deteriorating effect of fiscal shocks on current account is muted due to the positive effect of budget cuts on the domestic investment. However, they also emphasise that the transmission of these shocks to government budget is smaller the less open the economy is.

In addition, Ahmed (1987) analysed the link between the government spending, balance of trade and terms of trade. He adopted a large open-economy infinite horizon intertemporal framework based on a two-good two-country model and focused on the UK. Military spending data were used as a measure of the government spending. Ahmed found that temporary government spending changes appeared to have a strong impact on the balance of trade deficit. It was especially evident when the relationship was estimated over 1732-1830, the period that includes the war years and, hence, the significant variation in the government spending data. Moreover, increase in the government spending had a deteriorating effect on the terms of trade, suggesting that some of that additional expenditure was spent on imports. As for the adjustment processes the effect of the government deficit on the trade flows did not change whether government increased its revenue by raising taxes or by increasing debt.

Ahmed also emphasised that the results were sensitive to specifications, such as estimation period and the method of decomposition of the government spending.

Thus, in this section we want to see how balance of trade and government surplus adjust to each other. Hence, we estimate (4.33) for change in government surplus, ΔS_{it}:

$$\Delta lnS_{it} = a_{10} + a_{11} lnS_{it-1} + a_{12} lnBOT_{it-1} + \nu_{1it},$$

and (4.34) for change in the balance of trade, ΔBOT_{it}:

$$\Delta lnBOT_{it} = a_{20} + a_{21} lnS_{it-1} + a_{22} lnBOT_{it-1} + \nu_{2it}.$$

The equations were estimated for the whole sample and then for 17 countries individually for 1870-2016. We expect both effects, own and cross-surplus, to be negative. If an economy is running a balance of trade deficit, it wants to increase the government surplus (by increasing taxes or by reducing expenditure). Meanwhile, if there is a balance of trade surplus, then a government can choose an expansionary fiscal policy which leads to fiscal surplus going down, hence, the inverse relationship between balance of trade and government surpluses. The results for the cross-surplus analysis are:

$$\begin{array}{lllllllll} \Delta \ln S_t & = & -0.021 & - & 0.193 \ln S_{t-1} & - & 0.003 \ln BOT_{t-1} & + \ \nu_{1t} & R^2 = 0.104 \\ & & (0.003) & & (0.012) & & (0.011) & & SER = 0.132 \end{array}$$

$$\begin{array}{lllllllll} \Delta \ln BOT_t & = & -0.015 & + & 0.066 \ln S_{t-1} & - & 0.221 \ln BOT_{t-1} & + \ \nu_{2t} & R^2 = 0.116 \\ & & (0.004) & & (0.014) & & (0.013) & & SER = 0.157 \end{array}$$

The estimates suggest very little cross-surplus adjustment. When we estimate (4.33) and (4.34) for the whole sample with fixed cross-country effects and over the whole period, in (4.33) the coefficient that measures cross-surplus effect is insignificant, and in (4.34) it is significant but positive, suggesting that as government surplus increases, the balance of trade surplus increases and visa versa.

Heterogeneous estimates (Table 4.7) have mixed signs and in most cases the coefficients that measure cross-surplus effects are insignificant. For instance, for (4.33), $a_{12,i}$ coefficient, which measures the effect of an increase in the balance of trade on the government surplus, is only significant in six countries and only in three has the expected negative sign. As for the equation (4.34), the effect of government surplus on the balance of trade, $a_{21,i}$, is significant in seven countries and has the expected negative sign in two of them. Since the cross-surplus effects are mostly insignificant when estimated over the whole period, we opted not to estimate equations (4.33) and (4.34) over shorter sub-periods.

Table 4.7 Cross-Surplus Effects (Equations (4.33) and (4.34); 1870-2016)

	Equation (4.33)				**Equation (4.34)**			
	lnS_{it-1}		$lnBOT_{it-1}$		lnS_{it-1}		$lnBOT_{it-1}$	
Country	a_{11}	*t-stat.*[1]	a_{12}	*t-stat.*	a_{21}	*t-stat.*	a_{22}	*t-stat.*
17 countries	**-0.19**	**-16.13**	0.00	-0.27	**0.07**	**4.61**	**-0.22**	**-17.38**
Australia	**-0.11**	**-3.56**	-0.05	-0.93	**0.12**	**2.77**	**-0.43**	**-5.72**
Belgium	**-0.43**	**-6.25**	0.00	-0.03	0.17	1.87	**-0.55**	**-6.60**
Canada	**-0.26**	**-4.48**	-0.04	-0.70	0.02	0.44	**-0.16**	**-3.40**
Denmark	**-0.53**	**-6.83**	0.12	1.94	0.09	1.34	**-0.20**	**-3.64**
Finland	**-0.33**	**-4.79**	0.01	0.15	0.05	0.68	**-0.23**	**-3.98**
France	-0.04	-0.78	**-0.17**	**-3.01**	**0.48**	**6.40**	**-0.68**	**-9.03**
Germany	**-0.23**	**-3.38**	-0.03	-0.62	-0.10	-1.63	**-0.21**	**-4.22**
Italy	-0.06	-1.58	**-0.16**	**-4.22**	**0.31**	**4.54**	**-0.43**	**-6.90**
Japan	**-0.07**	**-2.08**	-0.01	-0.22	**-0.16**	**-2.45**	**-0.39**	**-6.60**
Netherlands	**-0.22**	**-3.62**	**-0.09**	**-2.03**	-0.09	-1.88	**-0.10**	**-3.11**
Norway	**-0.31**	**-5.95**	**0.13**	**4.68**	0.01	0.18	-0.06	-1.74
Portugal	**-0.20**	**-4.11**	0.01	0.19	0.10	1.46	**-0.12**	**-3.01**
Spain	**-0.31**	**-5.58**	0.01	0.30	0.05	0.49	**-0.19**	**-3.77**
Sweden	**-0.35**	**-5.66**	**0.11**	**2.12**	0.08	0.98	**-0.38**	**-5.83**
Switzerland	**-0.19**	**-3.49**	**0.16**	**2.23**	**-0.08**	**-2.38**	**-0.14**	**-3.07**
UK	**-0.14**	**-2.58**	0.03	0.39	**0.13**	**2.56**	**-0.28**	**-4.51**
US	**-0.21**	**-4.07**	0.00	0.06	-0.01	-0.33	**-0.10**	**-2.73**
Average	-0.234		0.001		0.068		-0.271	

Note: the significant coefficients are in bold; [1] - t-statistic. *Expected signs: all coefficients are expected to be negative.*

4.10 Conclusion

In this chapter we analyse the public sector surplus and government debt of 17 countries over sufficiently long period in order to check whether government surplus and debt converge to stationary state, and if they do, what is the adjustment process and feedback mechanism behind it. We estimate separate equations for the change in public sector surplus and its components (revenue and expenditure), and include lagged surplus and lagged debt-GDP ratio as explanatory variables in each case. We find that when cross-country effects are fixed, and we estimate the models over the whole period, the surplus is stabilising on its own lagged values and on debt-GDP. Most of the adjustment of the surplus to lagged surplus is done through revenue, while the adjustment to the lagged debt-GDP is achieved through changes in expenditure. However, if we consider heterogeneous estimates there is feedback coming from the lagged surplus to the current surplus, but not from the lagged debt-GDP.

Nonetheless, by design of the models for surplus (4.20), revenue (4.25) and expenditure (4.26), if one of the lagged surplus adjustment coefficients (revenue one, β^r or expenditure one, β^e) is stabilising and is larger than the destabilising one, then the overall surplus is stabilising. The same logic is true for the debt-GDP adjustment coefficient. We then find that over the whole period the surplus is stabilising on its own lagged values for all 17 countries and is adjusting correctly to changes in debt-GDP for all but two economies. However, we find less evidence of stabilisation over shorter periods.

Considering shorter sub-periods, in 1870-1914 three β^r and two β^s coefficients have the wrong sign, suggesting that most of the time both revenue and expenditure are responding to changes in the lagged surplus correctly and, as a results, overall current surplus adjusts to the lagged surplus in the right direction for all but two countries. However, revenue does not seem to adjust correctly to changes in the lagged debt-GDP for more than half of the sample (nine countries), while expenditure fails to do so for five economies. Overall, surplus adjusts to changes in the lagged debt-GDP in the right way for all but four countries. In the second sub-period, surplus responds correctly to changes in its lagged value and in the lagged debt-GDP ratio for all but one and four economies, respectively. It appears, that during this period the adjustment to the lagged surplus is mostly done through revenue, while the adjustment to the lagged debt-GDP – through expenditure. Out of the three sub-periods, the third one somewhat stands out. In 1951-2016 there are two β^r and six β^e coefficients with the wrong sign, suggesting that in these cases revenue and expenditure are not responding to changes in surplus in the right way. Nevertheless, the stabilising coefficients are large enough to offset the effect of the destabilising ones, thus overall surplus adjusts to previous surplus in the right direction for all countries. However, the γ coefficients for this sub-period suggest that revenue does not seem to respond to changes in the lagged debt-GDP in the right way, with only one exception, but expenditure responds to the lagged debt-GDP correctly in all cases. The results for the surplus equation suggest that these stabilising γ^e coefficients offset the effect of destabilising γ^r ones in all but four countries.

Overall, the heterogeneous estimates suggest that there does not seem to be a strong feedback coming from the lagged debt-GDP ratio, with most of the γ (lagged debt-GDP adjustment coefficient) being insignificant in our estimations of the equations (4.20), (4.25) and (4.26). Hence, we try to estimate an augmented version of (4.20) with three key macroeconomic variables, namely GDP growth, long-term interest rate and inflation, as additional explanatory variables. These three key macroeconomic variables are likely to be correlated with debt-GDP ratio and, hence, we check whether their addition to the surplus model alters significance of the lagged debt-GDP coefficient. In the log version of the model all coefficients with the exception of the long interest rate one are significant when we estimate pooled data over the whole period. Nonetheless, when we consider the model for individual countries, the lagged debt-GDP coefficient remains insignificant for 15 out of 17 countries. Since we cannot reject the hypothesis that the lagged debt-GDP coefficient equals zero, it is not clear whether the best estimate of the effect of debt on surplus is zero or the estimated lagged debt-GDP coefficient, γ^s.

Thus, if we consider countries in the sample individually there does not seem to be an equilibrium level for the debt-GDP ratio for the economies under consideration. It is somewhat consistent with Bohn's argument (2007) that any debt-GDP ratio can be sustainable as long as lenders believe that the government is solvent and can meet its debt payments, as in this case lenders continue to lend to this government. Therefore, there does not need to be an equilibrium level of debt. This conclusion brings attention to the main issue with the empirical analysis of the solvency condition for the public sector surplus. Since sustainability of the government debt depends on the rational expectations, then the question of solvency is of an economic nature and, hence, the true condition for solvency cannot be empirically tested.

As in the balance of payments sustainability analysis, we also check whether there is more pressure to adjust on governments running deficits versus those running surpluses. Our pooled data estimations suggest that, on average, speed of adjustment for governments running budget deficits is about twice higher (20%) than for those running surpluses (10%). Hence, we conclude that, as in case with the balance of payments, there is indeed some evidence of asymmetric adjustment.

To conclude this chapter we consider cross-surplus adjustments assuming that shocks that affect government surplus might move the balance of payments in the same direction. However, we find little evidence of the cross-surplus adjustments regardless of whether we estimate pooled data for the whole period or if we consider estimates for 17 countries individually.

Overall, government surplus seems to stabilise on its own lagged values whether we consider pooled data estimations or heterogeneous results. However, there is significant feedback coming from the lagged debt-GDP only when we consider the whole sample and long period, but not in case of individual countries or shorter sub-periods. Nonetheless, since pooled estimates suggest that government surplus is stabilising, the potential area for future research on this topic may be the in-depth analysis of the mechanisms behind this stabilising adjustment.

Chapter 5

Effect of Exchange Rate: European Monetary Union

5.1 Introduction

Traditional economic theory suggests that open economy adjustments are mainly done through changes in interest rate and exchange rate. In this chapter[1] we compare stand-alone European countries which have control over setting their monetary targets to maintain balance of payments sustainability, to members of the European Monetary Union (EMU) which moved to common interest and exchange rate and cannot use these tools to adjust their balance of payments positions.

Therefore, this chapter attempts to examine the effect of the establishment of the euro on eight economies that joined the common currency in comparison to four European economies that did not join the euro. There are many ways to measure effect of the euro, and the most common way is to compare the outcome with some counterfactual that is an estimate of what would have happened to a euro country had it not joined the euro. Constructing macroeconomic counterfactuals raises difficult issues. There is also an issue as to whether to use a simple single equation approach of the sort discussed in Pesaran and Smith (2016) or a full system of the sort discussed in Pesaran and Smith (2018). We use both.

In this chapter we opt for a very specific measure of the effect. We examine the extent to which joining the euro changed the main national macroeconomic relationships, that is whether there was a structural break in particular equations. We address the following

[1]This chapter is an expanded version of the paper written jointly with my supervisor, Professor Ron P. Smith (Akhmadieva, V. and Smith, P. R. (2019). The Macroeconomic Impact of the Euro. *Scientific Annals of Economics and Business*, 66(SI2), pp. 229-249.). I am very grateful to Professor Ron for the opportunity to use a version of our joint paper for this chapter.

questions: whether there was a significant break in 1999 for the economies that joined the euro and whether it was bigger for those countries that joined than for countries that did not join?

We also consider: in which equations the largest break occurred, and whether the break was in 1999, or at another time. We cannot be sure when the effect of the euro happened. It could have been some years before 1999 when they started to prepare for entry or some years after 1999 as they adjusted to entry.

The answers to these questions will always be conditional on other influences that we control for. We always control for foreign variables, the 2008 crisis was global. Of course, to the extent that the foreign variables were also influenced by the formation of the euro we do not pick up that indirect effect. However, we are primarily interested in changes in the national economy.

We also consider controlling for policy instruments, interest rates and exchange rates, which, once the common currency was established, might be regarded as exogenous to many, if not all, individual euro countries.

Section 5.2 reviews the literature. Section 5.3 describes the data and provides some descriptive statistics and history. Section 5.4 uses single equations models and examines the timing of structural breaks. Section 5.5 describes the econometric approach adopted for the systems analysis. Section 5.6 provides systems estimates, and section 5.7 is devoted to the forecast. Section 5.8 contains some concluding comments.

5.2 Literature

There is a large literature on the euro which considers such issues as whether the euro is an optimal currency area, and the extent to which its members are subject to symmetric shocks. De Grauwe (2018) provides a text-book treatment of the economics of the monetary union. While a single currency means equality of nominal interest rates and exchange rates, it does not mean real equality. Both, real interest rates and real exchange rates, diverged substantially among the euro economies. Adding to that the differences in size, factor endowment productivity and political environment of the euro economies, and there is the danger that a "one size fits all" setting is unlikely to be optimal for all members. The interest rate chosen by the European Central Bank (ECB) may be too low for rapidly growing countries and too high for ones in recession.

Indeed, a common currency does not translate into equality among the European Monetary Union members. Creditor countries tend to choose the policies for the whole union, forcing budgetary policies and tight fiscal disciplines on debtor countries (De Grauwe, 2016). During the Euro-crisis at the end of the 2000s, this pushed the Southern countries, that were already suffering from liquidity crisis, into a deeper recession, as they were forced to reduce wages and price level relative to the creditor members of the union.

Aksoy et al. (2002) noted that the determinants of the optimal interest rate are country-specific, and this may raise tension within Euroland when it comes to choosing the optimal monetary policy for the system. Following Rudebusch and Svensson (1999) they developed a model which is an approximation for the ECB optimal linear feedback decision rule. They found that the spread between nationally desired interest rate and the one decided by monetary union governors is wider for smaller countries in the union. Moreover, this difference tends to be larger when countries that desire to stabilise their output find it impossible to choose their optimal interest rate and, hence, left in frustration, unable to achieve their objective. Nonetheless, the authors claimed that having the ECB to choose the optimal interest rate for the union based on economic conditions of all the countries in the system improves welfare by reducing losses from the volatility of inflation and variability of output. These tend to be higher when the choice of the interest rate is based on the nationalistic objectives.

Therefore, there are many issues about the implications of a common monetary policy for fiscal policy, and whether fiscal federalism is required is a question open for discussion (Farhi and Werning, 2017). For instance, single currency members are not able to issue debt in a national currency they would have control over. This makes them susceptible to self-fulfilling liquidity and solvency crises as investors are lacking a guarantee that cash will be available at the maturity date. Hence, they choose to invest somewhere else, pushing interest rates up and reducing liquidity available to the euro countries (De Grauwe and Ji, 2013). Moreover, since monetary union does not allow its members to devalue their currency, they cannot follow this path as a remedy to increase their relative competitiveness in case of need. Hence, governments are forced to push their price level down by reducing the wages leading them to a deeper recession. It appears the economy's ability to defend itself against asymmetric shocks is tightly connected to the flexibility of its wages and price level.

Later De Grauwe (2013) empirically tested and confirmed the validity of fragility hypothesis proposed by De Grauwe. They analysed the government bond markets of the EMU countries and used a control group of 14 stand-alone developed countries for comparison. Stand-alone countries demonstrated much higher ability to sustain their sovereign debts as increase in their debt to GDP ratios was not perceived by financial market participants as a sign of increased fragility. As a result these countries overcame 2010-2011 crisis without noticeable raise in the interest rate spread. In contrast, the euro countries experienced a break in the spreads-debt to GDP ratio in 2010-2011 as their financial vulnerability increased substantially when they accumulated public debt during times of financial distress. Thus, EMU countries are more prone to self-fulfilling liquidity crises.

Saka et al. (2015) provided further empirical support for the Paul De Grauwe's fragility hypothesis. Using a capital asset pricing model, they found evidence for the herding contagion that was effectively countered by the timely and reassuring ECB announcements that helped to reduce investors' fear of losses and effectively addressed self-fulfilling nature of the crisis during the late 2000s. Nevertheless, the analysis relies on the influence of Spain-specific news on market participants and might have been different if the estimations were based on the news of other EMU members.

Potjagailo (2017) noted that monetary shocks generated within the Eurozone tend to spillover on the stand-alone European countries. The size of the spillover effects on each individual country depends on country-specific characteristics, such as its openness to trade. However, in case of majority of countries under consideration, financial variables, including short interest rates and real effective exchange rates, are significantly affected by the shocks originated in the Eurozone. As for common currency unions, when member economies are pushed 'out of sync' by shocks that are permanent, not only nature (De Grauwe, 2016), but also direction and extent of these shocks matter when estimating their effect on the flexibility-symmetry trade-off (Campos and Macchiarelli, 2018).

Those are mainly macroeconomic issues. The benefits of a monetary union lie in the microeconomic level (De Grauwe, 2018). A single currency means free movement of capital inside the EMU and reduced uncertainty about expected exchange rate. Hence, an increased economic efficiency for all the members in the monetary union.

Campos and Macchiarelli (2018) argued that creation of the EMU improved the stability of the Euro Area. They distinguished between core, deep-rooted periphery and a mixed set of countries. The authors found that an economy is more likely to be a core country if it was a euro member and had strict product market regulations in place. The euro countries benefit from reduced transaction costs and reduced uncertainty about exchange rate as well as from price transparency, higher trade and competition. However, to fully enjoy the perks of being an EMU member, an economy needs to be able to achieve a minimum level of symmetry, flexibility and openness.

Pesaran et al. (2007) estimated what would have happened to the UK, Sweden and the euro area if the UK or Sweden had joined the euro using a Global Vector Autoregression, GVAR. Smith (2009) estimated the effects on the euro area issues. In both cases there was a relatively short sample of data after the establishment of the euro, and the samples ended prior to the financial crisis 2007-2008. In both cases the effects were not large. Now there are more data, and the large shock associated with the financial crisis provides extra identifying information.

Besides development of the monetary union, various global disturbances could have affected welfare of the euro members over the period covered in this chapter. The euro countries were not exempt from the harsh effects of the global crisis 2007-2008. Caruso et al. (2019) analysed the effects of the financial crisis 2007-2008 on the euro countries. Using a multivariate VAR model they performed conditional and unconditional forecasts in order to examine a special nature of a financial crisis that they argued is different from a regular recession. They found major deviations in output, private and public debt ratios as well as other macroeconomic and financial indicators of the euro countries, when the model was estimated over pre-crisis and post-crisis periods. Some of these deviations, such as persistent decline in investment, are atypical to the extent that they, as the authors argued, cannot be explained by the business cycle regularities. The crisis was also characterised by the record high fiscal deficit-GDP ratios followed by an adaptation of extremely tight fiscal policies, all of which make this crisis unprecedented and likely to cause a structural change in the macroeconomic indicators of some euro countries.

The effect of a single currency on trade flows has been studied in depth with conflicting results. Rose's paper (2000) was among the first ones that argued that single currency boosts trade between the monetary union members. Using an augmented gravity model (that includes a number of additional conditioning and monetary variables that a traditional gravity model does not) and panel data for 186 countries, Rose analysed bilateral trade between economies that share a common currency. He found a large positive effect of a currency union and a negative effect of exchange rate volatility on international trade, with the latter being relatively small. In contrast, Thom and Walsh (2002) found no such relationships, when they analysed Anglo-Irish trade. Ireland dependence on trade with the UK declined since 1950s, but mainly due to the decrease of British share in global trade and falling transportation costs, that allowed Ireland to trade with other countries.

There has been considerable dispute about the effect of monetary union on trade and Rose (2017) analyses the factors that cause the estimates to vary so much. It appears that estimates are sensitive to the sample size, with the effects of a single currency on trade being stronger if the analysis includes more observations by country and time period. The number of countries included seems to make especially substantial difference, which, as Rose suggested, might be explained by the bias in the estimation of the country-time fixed effects that arises if some smaller economies are omitted from the sample.

An interesting natural experiment occurred when Ireland broke its link with sterling in 1979. Thom and Walsh (2002) argued that Ireland joining European Monetary System in 1979 and consequent change in Ireland's exchange rate regime did not have any substantial effect on Anglo-Irish trade. Checking for structural break in 1970s and using rolling regressions they found that Ireland's break with British sterling in 1979 had no adverse effect on the economic relationships between the two countries and free bilateral trade between them was maintained after the end of this currency union. Moreover, the trade between these countries grew faster in the post-currency-union period, increasing from 77% to 114% in 1950-1978 and then to 167% over the next two decades (Thom and Walsh, 2002, p. 1115). In addition to this time-series analysis, they studied Ireland's trade with 19 countries and after examining its intra-union and extra-union trade relations, they confirmed that Ireland's change of exchange rate regime did not alter trading patterns of this economy. Hence, they concluded that if two economies are already trading freely, the dissolution of a monetary union will not have a significant effect on the trade between these two countries. Nonetheless, Thom and Walsh's analysis is based on a case study of two developed economies and considers the effect of ending a monetary union on bilateral trade. Therefore, one cannot necessarily generalise from their findings when considering a case of creating a monetary union, especially if it is a monetary union of large number of countries with different features.

Nevertheless, it is important to note that not all countries in our sample joined the EMU, however, all of them are members of the European Single Market. Therefore, any difference or lack of such between the EMU-members and stand-alone countries might be either due to the effect of common currency or benefit from a single market.

5.3 Data

We use data from the GVAR toolbox 1979Q3-2016Q4, Mohaddes and Raissi (2018), to estimate VARX* for a number of European countries that did or did not join the euro.

For countries $i = 1, 2, .., 12$ and $t = 1979Q2 - 2016Q4$, the variables are:

- y_{it} is a natural logarithm of real GDP volume index;
- π_{it} is the rate of inflation, calculated by taking the difference of the natural logarithm of the consumer price index;
- eq_{it} is a natural logarithm of the nominal equity price index deflated by CPI;
- ep_{it} is a natural logarithm of the exchange rate of country i at time t expressed in units of foreign currency per US dollar deflated by country i's CPI. We will refer to this as a real exchange rate, even though it is not adjusted for the US price level;

- r_{it} is a nominal short-term interest rate per quarter, in percent. It is computed as $0.25 \times \ln(1+R_{it}^{r})/100$ where R_{it}^{r} is the nominal short rate of interest per annum in percent;

- lr_{it} is a nominal long-term interest rate per quarter, in percent. It is computed as $0.25 \times \ln(1+R_{it}^{lr})/100$ where R_{it}^{lr} is the nominal long rate of interest per annum in percent. Long interest rate data are not available for Finland;

- $poil_{it}$ is a natural logarithm of the nominal price of oil in US dollars.

Our sample includes eight euro-member countries, which are Austria, Belgium, Finland, France, Germany, Italy, the Netherlands and Spain as well as four European countries that are not the EMU members, which are Norway, Sweden, Switzerland and the UK.

In addition to individual variables, say x_{it}, there are global equivalents x_{it}^{*} calculated as country-specific trade weighted averages of the corresponding variables of the other countries

$$x_{it}^{*} = \sum_{j=0}^{N} w_{ij} x_{jt}, \text{ with } w_{ii} = 0, \tag{5.1}$$

where w_{ij} is the share of country j in the trade (exports plus imports) of country i. Thus, for instance, if y_{it} is log real GDP, then y_{it}^{*} is the weighted average of the log GDP of trading partners.

For each country we break the data into three sub-periods, 1979Q4-1998Q4, 1999Q1-2008Q4 and 2009Q1-2016Q4. The Table 5.1 provides means and standard deviations for the growth rate, rate of inflation, change in the real exchange rate and long interest rate (excluding Finland).

In 1999 a three-year transition period from national currencies to euro and single monetary policy began. During that time majority of the countries suffered from a negative growth, comparing to the pre-euro period. The only exceptions were Finland, Spain, Sweden and Switzerland, who enjoyed a minor increase in the GDP growth rate. Moreover, growth was lower in the post-crisis period, but that was a global phenomenon, not just for euro countries. Nonetheless, we control for this by including global variables, x_{it}^{*}, in out model.

The average inflation rate dropped during the period of euro formation. For most countries the inflation rate more than halved, for instance, in Norway and Sweden, while in some cases, such as in France and Italy, it fell by more than two thirds. Hence, this trend is true for both, the EMU members and non-member countries. In the UK the inflation rate dropped from average of 5.37 in 1979-1998 to 2.28 in 1999-2008 and remained around 2% since then. This reflects the Bank of England's commitment to maintain the inflation rate around 2%, the inflation targeting adopted by the central bank in 1992 after exiting the European Exchange

Table 5.1 Descriptive Statistics (Percent Per Annum (x400))

	Δy^1			π^2			Δep^3			lr^4		
	1	2	3	1	2	3	1	2	3	1	2	3
Austria	2.28	2.00	1.08	3.14	1.99	1.70	-3.70	-2.28	0.93	7.19	4.10	2.15
	3.66	3.58	4.26	2.08	1.14	1.25	20.34	20.86	15.91	1.51	0.49	1.22
Belgium	1.94	1.91	1.14	3.49	2.23	1.55	-2.56	-2.51	1.07	8.72	4.38	2.47
	3.23	2.79	2.32	2.67	1.73	1.69	20.41	20.64	15.76	2.32	0.64	1.33
Finland	2.51	2.63	0.09	4.66	1.90	1.18	-3.13	-2.20	1.45	0.00	0.00	0.00
	5.65	4.42	6.01	3.55	1.60	1.71	20.63	20.62	15.80	0.00	0.00	0.00
France	1.84	1.82	0.77	4.52	1.77	0.92	-3.01	-2.05	1.70	9.20	4.32	2.20
	1.78	2.32	1.82	4.04	1.19	1.14	19.38	20.85	15.77	2.85	0.61	1.11
Germany	1.89	1.35	1.32	2.74	1.62	1.12	-3.10	-1.91	1.51	6.88	4.11	1.65
	3.78	2.67	3.79	2.18	1.17	1.03	20.21	20.74	15.79	1.43	0.46	1.07
Italy	1.89	0.97	-0.38	7.50	2.33	1.12	-3.79	-2.62	1.51	11.72	4.48	3.63
	2.65	2.31	3.12	5.50	0.83	1.37	20.11	20.74	15.72	3.66	0.49	1.46
Netherlands	2.31	1.90	0.45	2.68	2.18	1.44	-2.92	-2.48	1.19	7.42	4.35	1.97
	3.26	2.43	2.83	2.14	1.25	1.42	20.27	20.46	15.81	1.56	0.66	1.13
Norway*	3.21	2.05	1.43	5.30	2.15	1.99	-3.12	-3.06	0.60	9.55	4.76	1.94
	5.33	3.78	5.17	3.71	3.17	2.05	18.06	22.40	18.05	2.69	1.11	0.85
Spain	2.51	2.82	0.43	6.99	3.12	0.66	-2.99	-3.41	1.96	11.54	4.17	3.66
	1.88	2.20	2.48	4.07	1.67	3.28	20.00	21.11	16.11	3.15	0.61	1.54
Sweden*	1.91	2.30	2.63	5.51	1.61	0.65	-2.08	-1.89	1.17	10.10	4.40	1.91
	5.64	4.50	4.11	4.37	1.63	1.52	21.21	21.98	19.38	2.12	0.71	0.97
Switzerland*	1.64	1.82	1.36	2.78	1.09	-0.30	-3.61	-3.04	-1.26	4.62	2.88	0.87
	2.82	3.94	2.79	2.33	1.12	1.13	22.94	16.45	15.47	0.90	0.44	0.83
UK*	2.21	2.02	1.36	5.37	2.28	2.08	-3.80	-1.51	0.74	9.39	4.59	2.52
	2.86	2.86	2.01	3.74	1.31	1.68	21.36	17.42	16.13	2.10	0.24	0.85

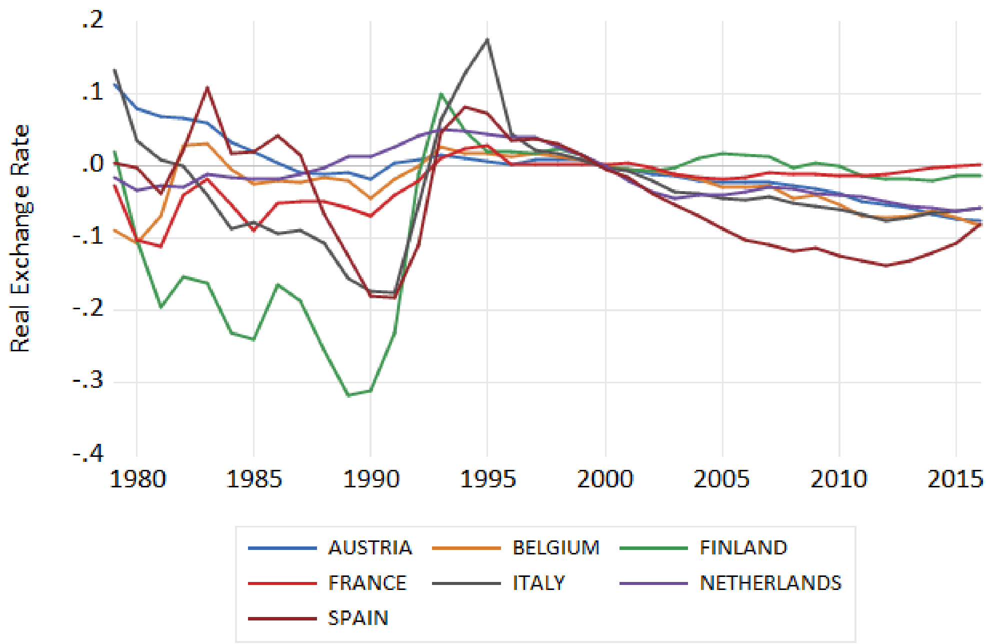
Notes: *non-member countries. Δy^1 is the GDP growth rate, π^2 is the rate of inflation, Δep^3 is the change in the exchange rate and lr^4 is the long interest rate. The information are given for periods: 1: 1979Q4-1998Q4; 2: 1999Q1-2008Q4; 3: 2009Q1-2016Q4.

Rate Mechanism. In case of the euro countries, since the EMU foundation, money supply and inflation were closely coordinated by the European Central Bank. However, it does not explain the decrease in inflation rates in the stand-alone countries. The inflation rate was maintained at a relatively stable rate after the financial crisis 2007-2008.

The average exchange rate growth over the first two sub-periods was negative, suggesting an appreciation against US dollar. However, following the financial crisis 2007-2008, the dollar strengthened against national currencies of the countries in the sample. The long interest rate declined significantly during the years of euro formation, partially as a result of the strict anti-inflationary policy adopted by the ECB, and fell further in the third period, possibly due to low expected inflation and returns on investments. Average growth rates fell over the three sub-periods both for the eight euro countries: 2.15 to 1.92, to 0.62, as well as for the four non-euro countries: 2.24 to 2.05, to 1.69. While the non-euro average is higher, the differences are quite small for the first two sub-periods, but in the post-crisis period the four non-euro countries grew faster than any of the eight euro countries. Meanwhile, inflation fell over the three periods in both groups.

In addition, we plot the real exchange rates (Figure 5.1) and real long interest rates (Figure 5.2) for the euro countries. Among euro-members, Germany's currency (in real terms) was the most stable throughout the period, hence, we use this economy's indicators as the base ones. Therefore, to compare real exchange rates we first use natural logarithm of the exchange rate of country *i* at time *t* expressed in US dollars deflated by country *i*'s CPI and then deflate it by this of Germany. To get a clearer pattern we take average of four quarters to calculate average annual rate. Looking at Figure 5.1, we can see high volatility among the countries prior to setting up the euro that temporarily decreases around 1999. However, shortly after the EMU formation the divergence continues, with most countries having weaker exchange rate than Germany. Finland has a stronger real currency from around 2003 to 2010 after which its exchange rate falls below zero.

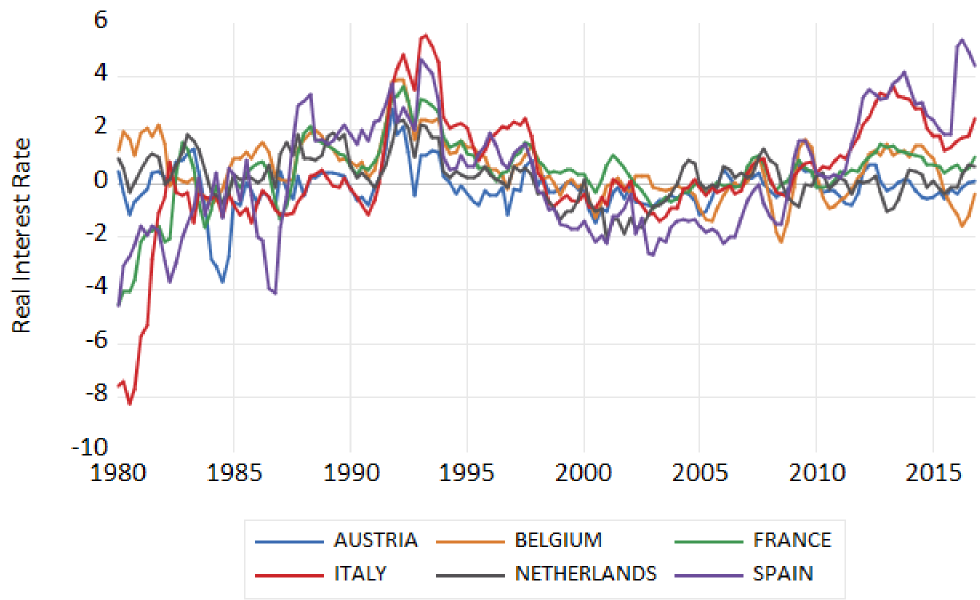
Fig. 5.1 Real Exchange Rate (Euro Members)



Source: Jordà-Schularick-Taylor Macrohistory Database.
Note: Germany is a base country.

Moving to the real interest rate, we use long-term rate and, therefore, leave Finland out as data for long interest rate are not available for this country. Starting from the nominal long-term interest rate per quarter, expressed in percent, we subtract the inflation rate to get the real interest rate. As with the real exchange rate, we use average annual rates and Germany's real interest rate as a base. Figure 5.2 shows that the real interest rates are very jumpy, although the variance is noticeably larger prior creation of the EMU. Following 1999 there is a reduction in dispersion of the real interest rates, but after the financial crisis 2007-2008, some of them, such as those of Italy and Spain diverge.

Fig. 5.2 Real Long Interest Rate (Euro Members)



Source: Jordà-Schularick-Taylor Macrohistory Database.
Note: Germany is the base country.

5.4 Single Equation

Pesaran and Smith (2016) suggest using simple single equation reduced-form or final-form "policy response equations" to provide counterfactuals for policy evaluation. In their empirical example, they use an equation in which UK growth is made of a function of foreign growth and interest rates, the policy variable. We do not have an explicit policy variable, instead we have a possible structural break caused by the euro. Nonetheless, the basic approach is the same. We need to construct a counterfactual, based on the pre-euro relationship, to compare with the actual post-euro realisation. There is the difficulty that we cannot be completely sure whether the difference is due to joining or not joining the euro or something else, including misspecification of the forecasting equation. We control for other factors, by conditioning on exogenous variables and comparing euro and non-euro countries which may be subject to similar exogenous shocks.

5.4.1 Econometric Techniques

Pesaran and Smith (2016) do not provide any real guidance about how one should choose the simple model, beyond indicating that it should pass diagnostic tests; therefore, it has to be a matter of experimentation. The issue is also more complicated here than in their case. They were looking at a single country, whereas we are considering 12 countries.

There are two main issues. Should we use a levels relationship, treating the variables as trend stationary, or difference stationary, as Pesaran and Smith (2016) did, estimating a growth rate relationship. Secondly, there is a questions of what other conditioning variables should we use. Hence, we proceed by estimating a levels relationship and then first difference model to compare the results. As for the latter issue, in this analysis we will use the GVAR measure of foreign income.

We evaluate the equations using diagnostic tests. This includes the F-version of a Lagrange Multiplier (LM) test for up to fourth order serial correlation and a Jarque-Bera test for normality. We also conduct the Ramsey Regression Equation Specification Error (RESET) tests for non-linearity and hereroskedasticity.

We then perform structural stability analysis. We use the Chow's first structural stability or breakpoint test (SS test) for parameter equality between the two periods and check if the variances are the same in both periods with $k = 5$ parameters estimated. It uses the unrestricted residual sum of squares of the regressions over the two periods $(\widehat{u}_1'\widehat{u}_1 + \widehat{u}_2'\widehat{u}_2)$ with degrees of freedom $T - 2k$ and restricted residual sum of squares from the regression over the whole period $\widehat{u}'\widehat{u}$ with degrees of freedom $T - k$. The null hypothesis is that the parameters are equal in the two periods implies k restrictions and the test statistic is

$$\frac{[\widehat{u}'\widehat{u} - (\widehat{u}_1'\widehat{u}_1 + \widehat{u}_2'\widehat{u}_2)]/k}{(\widehat{u}_1'\widehat{u}_1 + \widehat{u}_2'\widehat{u}_2)/(T - 2k)} \sim F(k, T - 2k). \tag{5.2}$$

Chow also suggested a second predictive failure test (PF test) for the hypothesis that the first period predicts the second one, and the test statistic for PF test is:

$$\frac{[\widehat{u}'\widehat{u} - \widehat{u}_1'\widehat{u}_1)]/T_2}{\widehat{u}_1'\widehat{u}_1/(T_1 - k)} \sim F(T_2, T_1 - k). \tag{5.3}$$

This tests the hypothesis that in

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ X_2 & I \end{bmatrix} \begin{bmatrix} \beta_1 \\ \delta \end{bmatrix} + \begin{bmatrix} u_1 \\ 0 \end{bmatrix} \tag{5.4}$$

δ the $T_1 \times 1$ vector of forecast errors is not significantly different from zero. This has a dummy variable for each observation in the second period.

These two Chow tests are for very different hypotheses: the SS test checks that the parameters are equal while the PF test hypothesis is that the actual is within the confidence interval of the forecast. Since the confidence intervals of the forecast are very wide, the PF test may not have a lot of power. This is also reflected in the degrees of freedom: for the SS test it is k, the number of parameters, for the PF test it is T_2, the number of observations in the forecasted period.

The two Chow tests assume a known break point, the time of the formation of the euro. We can also use the Quandt-Andrews test, which has an estimated break point, to see if it indicates a significant break and if so, whether this is at the time of the formation of the euro; and because this searches over all possible breakpoints calculating the F-statistic for each date, the critical value for the F-test is much larger than for the known break-point Chow tests.

5.4.2 Level of GDP Equation

In the comparisons of growth rates in the descriptive statistics of section 5.3 we did not control for other factors, particularly the crisis. We now do that. In the first part of this section we estimate a very simple equation that makes the log of real GDP an ARDL function of foreign GDP and trend over the pre-euro period 1979Q2-1998Q4, say $t = 1, 2, ..., T_1$:

$$y_{it} = \alpha_0 + \alpha_1 y_{i,t-1} + \beta_0 y^*_{it} + \beta_1 y^*_{i,t-1} + \gamma t + u_t, \tag{5.5}$$

where y_{it} and y^*_{it} are logarithms of real domestic and foreign GDP, respectively, y_{it-1} and y^*_{it-1} are lagged logarithms of real domestic and foreign GDP, correspondingly, t is a trend, and u_t is the error term.

This equation is then used to forecast GDP over the following 72 quarters, where $t = T_1 + 1, T1 + 2, ..., T$, with $T2 = T - T1$.

In levels equations for Austria, Belgium and Italy current and lagged foreign output are significant and the total effect is positive. The trend, however, is insignificant in Austria and negative in case of Belgium and Italy, which is not what we would expect. Nonetheless, the equations pass all diagnostic tests, except functional form. For Finland current and lagged foreign outputs are significant, but the lagged one and the total effect is negative. The trend is positive but insignificant. The equation passes the normality and heteroskedasticity tests, but fails the serial correlation and the functional form ones.

In case of France the current and lagged foreign outputs are significant, but the lagged one is negative. However, the overall effect is positive. The trend is insignificant, and there is an issue of serial correlation present. The equation passes all other diagnostic tests. For Germany all variables are significant, however lagged foreign income is negative again. The total effect of lagged and current foreign income is positive, though. The trend is significant and negative. Equation fails the serial correlation test.

In case of the Netherlands, the lagged foreign income is negative and only significant at 10% level, however, the overall effect of global GDP is positive. The trend is insignificant, and the equation fails all diagnostic tests, except the serial correlation one. Hence, it does not seem to be well-specified. For Norway the lagged foreign income and the overall effect are negative. Trend is positive, but significant only at 10% level. The equation passes all diagnostic tests, but the p-value for the serial correlation test (0.072) is above 0.05, hence, there might be an issue of autocorrelation.

The equation for Spain has negative lagged foreign GDP coefficient and negative overall effect. The trend is insignificant, but the equation passes all diagnostic tests except the serial correlation one. In case of Sweden, lagged foreign income is once again negative, but the total effect is positive. The trend is insignificant, and the equation fails all diagnostic tests except the functional form one. The p-value for the heteroskedasticity test is 0.055, which is just slightly above the chosen significance level.

For Switzerland, the overall effect of current and lagged foreign income is positive, and both variables are significant, but the lagged coefficient is negative. The trend is negative and the equation passes all diagnostic tests. Finally, for the UK the current foreign output is not significant, and lagged foreign output is significant, but with a negative coefficient. The total effect of foreign output is negative, which is not what we would expect. The trend is positive, and equation for the UK passes all diagnostic tests, except the heteroskedasticity test.

Overall, the lagged foreign income coefficient is negative and significant for all countries, except the Netherlands. However, as one would expect, the overall effect of global GDP is positive for all countries, except for Finland, Norway, Spain and the UK. In contrast, a rather unusual result is a negative trend coefficient for Belgium, Germany, Italy and Switzerland. Moving to the diagnostic tests, all of the equations pass the normality test. Nonetheless, the equations for six countries (Finland, France, Germany, Norway, Spain and Sweden) have an issue of autocorrelation, five countries (Austria, Belgium, Finland, Italy and the Netherlands) fail the functional form test, while only two equations, for the Netherlands and UK, fail the heteroskedasticity test. The equations might be misspecified, and the growth rate relationship might fit the data better.

We then perform a structural stability analysis. The Table 5.2 gives the Chow Predictive Failure, PF, and Structural Stability, SS, tests p-values, mean and root mean square prediction errors (RMSPE) for a break in 1998Q4. The plots of actual and predicted values are given in an appendix (D.1). The last column consists of potential breakpoints identified by the Quandt-Andrews test.

From the Table 5.2 we can see that the hypothesis of parameter equality between the two periods is rejected at the 10% level for all countries. It is not rejected at the 5% level for Austria, Finland and the Netherlands. In contrast, the predictive failure tests are not significant for any country except Switzerland. These will have low power, since the confidence interval for predictions up to 72 quarters ahead estimated from 72 observations are very large. Thus, the forecast failures are not significant. This is despite the fact that there are massive forecast errors for many countries, but particularly for the UK, where the forecast explodes so the actual is much worse than one would have expected.

Table 5.2 Structural Stability Analysis (Income; Level)

	PF[1] F(72, 72)	SS[2] F(5, 139)	Mean PE[3]	RMSPE[4]	Q-A Test[5]
Austria	0.49	0.05	0.01	0.02	1988Q3***
Belgium	1.00	0.00	0.05	0.06	1987Q3***
Finland	0.99	0.08	0.10	0.11	1990Q2***
France	0.94	0.00	0.04	0.04	1998Q2***
Germany	1.00	0.04	0.00	0.08	1988Q3***
Italy	0.99	0.00	-0.07	0.09	2003Q1***
Netherlands	1.00	0.09	-0.06	0.08	1986Q1***
Norway*	0.91	0.03	-0.11	0.15	1987Q3***
Spain	0.14	0.02	-0.09	0.15	2008Q2***
Sweden*	1.00	0.02	0.15	0.17	1990Q1***
Switzerland*	0.01	0.00	0.13	0.16	2008Q2***
UK*	0.68	0.00	-0.59	0.83	2008Q2***

Notes: *non-member countries; [1] - Chow Predictive failure test; [2] - Chow Structural stability test; [3] - Mean Prediction Errors; [4] - Root Mean Squared Prediction Errors; [5] - Quandt-Andrews Test. *** indicates a significance level of 1 percent, ** of 5 percent, and * of 10 percent.

For instance, for Austria the predicted is close to the actual with a small mean error and RMSPE. Neither the PF nor SS tests reject the hypothesis of structural stability at the 5% level. The countries which performed less well than expected from the pre-euro relationship are Italy, the Netherlands and Norway. Their mean errors and RMSPE's are on a higher end, comparing to the rest of the sample. Moreover, the Chow SS test statistics suggests possible structural instability in models for Italy and Norway at 5% level and for the Netherlands at 10% level.

The countries which performed better than expected from the pre-euro relationship are Belgium, Finland, France, Sweden, and Switzerland. The Chow SS test suggests presence of structural instability in time series of all these countries, except Finland. In case of Switzerland this conclusion is supported by the Chow PF statistics as well. Sweden and Switzerland have the highest mean prediction errors and RMSPE's in the sample, excluding the UK. The UK results are unreliable because the coefficient of the lagged dependent variable is greater than one on the pre-euro period.

The countries for which the evidence is mixed are Germany and Spain. For Germany the predicted growth rate was higher than the actual one and was expected to decline around 2011. However, contrary to this prediction, ever since a sharp drop during the global crisis 2007-2008, the actual growth rate has been slowly increasing. For Spain the actual was close to the predicted till the global crisis, after which the actual growth rate dropped below the predicted one and continued to increase at a much slower pace than was predicted.

Overall, for majority of countries in the sample their actuals and forecasts are not necessarily close to each other, but exhibit similar general patterns. Only in case of Spain, Germany, Switzerland and the UK the actual and predicted income growth rates diverge substantially from each other.

The Chow's first (structural stability) test assumes a known break-point, but there is also Quandt-Andrews test for an unknown break point, which searches over the possible dates for a single break. The trimming percentage is set to 15%. All the breaks are significant at 1% level. We also use the Bai-Perron sequential test that searches for multiple breaks. For every country the Bai-Perron test confirms the breakpoint identified by the Q-A test, however, it also suggests possible breaks in other years with results being mixed and quite sensitive to the exact sample used. Therefore, we proceed with the Q-A test only.

In Finland the break occurred in early 1990s (second quarter of 1990), time of Finnish banking crisis and collapse of the Soviet Union, with which Finland had strong trading ties. In case of Sweden the Q-A test identified a break in 1990Q1, which might be connected to the Swedish banking crisis that erupted in 1992. For Spain, Switzerland and the UK, the Q-A test statistics suggests a breakpoint in 2008Q2, year of the global financial crisis. As for the rest of the countries, the identified breakpoints are not characterised by any major economic or financial events.

5.4.3 Growth in GDP Equation

Now we estimate a growth rate relationship, presented below:

$$\Delta y_{it} = \alpha_0' + \alpha_1' \Delta y_{i,t-1} + \beta_0' \Delta y_{it}^* + \beta_1' \Delta y_{i,t-1}^* + \varepsilon_t, \tag{5.6}$$

where Δy_{it} and Δy_{it}^* are real domestic and foreign GDP growth rates, respectively, Δy_{it-1} and Δy_{it-1}^* are the lagged logarithms of real domestic and foreign GDP growth rates, correspondingly, and ε_t is the error term.

This equation is then used to forecast GDP growth rate as with the levels relationship over the following 72 quarters, where $t = T_1 + 1, T_1 + 2, ..., T$, with $T_2 = T - T_1$.

The results for the growth equation are quite different from the levels relationship ones. For Austria and Belgium the overall foreign income growth effect is positive, but the lagged foreign GDP is only significant at 5% level for Belgium and is insignificant for Austria. In addition, a coefficient of the lagged national income is negative and in case of Belgium, insignificant. The equations, however, pass all diagnostic tests, except the equation for Belgium fails the serial correlation test at 10% level.

In case of Finland, France, Germany, Italy and the Netherlands only the foreign income growth coefficient is significant and positive. The equation for France passes all diagnostic tests, while equations for Finland and Germany fail the serial correlation test. As for Italy, the equation fails the serial correlation and normality tests. The equation for the Netherlands fails the normality and functional form tests, the latter only at 10% level.

In the Norway equation the foreign income growth has significant positive effect on home GDP growth. However, the lagged domestic income growth has negative and significant effect on the dependent variable. This equation passes all diagnostic tests.

In case of Spain the lagged domestic income growth and foreign income growth have positive effect on the domestic GDP growth, but the foreign coefficient is only significant at 10% level. However, the equation fails the serial correlation test and normality test, the latter only at 10% level. For Sweden the foreign income growth and lagged domestic GDP growth coefficients are significant, although the former is positive, while the latter is negative. The equation passes all diagnostic tests. In case of Switzerland and the UK, the foreign income growth and lagged domestic income growth have positive and significant effect on the dependent variable. The equation for Switzerland passes all diagnostic tests, while the UK one fails the test for serial correlation at 5% level as well as the functional form test and normality test at 10% level.

The lagged foreign income growth is insignificant for all countries with an exception of Belgium where its coefficient is significant at 10% level. Nonetheless, comparing to the levels model, the lagged foreign income growth coefficient is positive in all equations, except for Finland, France and Norway, with the overall foreign income growth effect being positive for all countries. All equations pass the functional form test (equations for the Netherlands and UK fail this test, but only at 10% level) and the heteroskedasticity test, but the equations for Belgium (at 10% level), Finland, Germany, Italy, Spain and the UK fail the serial correlation test. In addition, equations for four countries fail the normality test (for Italy and the Netherlands at 5%, while for Spain and the UK at 10% level).

Moving to the structural stability analysis, similarly to the levels equations, the Table 5.3 summarises the p-values for the PF and SS tests, mean errors and RMSPE for a break in 1998Q4 as well as the Q-A test breakpoints. Plots of actual and predicted values are given in an appendix (D.2).

Table 5.3 Structural Break Analysis (Income Growth)

	PF[1] F(72,73)	SS[2] F(4,141)	Mean PE[3] Annual %	RMSPE[4] Annual %	Q-A Test
Austria	0.49	0.75	-0.28	3.22	1988Q2
Belgium	1.00	0.59	0.35	1.76	1986Q3**
Finland	1.00	0.05	-0.62	3.89	1988Q1***
France	0.94	0.34	-0.20	1.35	2001Q4
Germany	1.00	0.88	0.09	1.89	1988Q3*
Italy	1.00	0.01	-1.17	1.88	2003Q1***
Netherlands	1.00	0.14	-0.64	1.85	1986Q1**
Norway*	0.94	0.31	-1.27	4.36	1986Q2*
Spain	0.00	0.12	-0.59	2.14	2000Q2
Sweden*	0.99	0.03	1.22	3.30	1994Q2**
Switzerland*	0.06	0.08	0.62	2.95	2008Q4
UK*	1.00	0.69	-0.18	1.84	1986Q1

Notes: *non-member countries; [1] - Chow Predictive Failure test; [2] - Chow Structural Stability test; [3] - Mean Prediction Errors; [4] - Root Mean Squared Prediction Errors; [5] - Quandt-Andrews test. *** indicates a significance level of 1 percent, ** of 5 percent, and * of 10 percent.

For Germany the predicted is the closest to the actual with the smallest mean prediction error in a sample, 0.09%. However, RMSPE is the smallest for France, 1.35%, and its mean PE is also relatively small, -0.20%, suggesting that the country performed slightly worse than expected from the pre-euro relationship. In both cases, Germany and France, neither the PF nor SS test rejects the hypothesis of structural stability. However, the Quandt-Andrew test identified potential breakpoint for Germany in 1988Q3, but it is only significant at 10% level.

Similarly to France, Austria and the UK also performed a bit worse than expected, with the mean PE of -0.18% and -0.28%, respectively. The tests do not indicate either parameter instability or predictive failure. Spain, Finland, the Netherlands, Italy and Norway performed even worse with their negative mean errors and RMSPE's being on a higher end comparing to the rest of the sample. Furthermore, in case of Spain both, the Chow PF and SS tests, suggest possible structural instability. For Italy and Finland there is indication of parameter change, but not of predictive failure (in case of Finland, at 10% level only). The Q-A test suggests a breakpoint in 1988Q1 for Finland and in 2003Q1 for Italy and in 1986Q1 for the Netherlands as well as in 1986Q2 for Norway (only significant at 10% level).

Moving to the countries that performed better than expected from the pre-euro relationship, these are Belgium, Switzerland and, most of all, Sweden with the highest mean PE of 1.22% in the sample. While the Chow PF and SS tests find no signs of structural instability in Belgium time series, the Q-A test suggest a possible breakpoint in 1986Q3. Both tests, the Chow PF and SS, reject the null hypothesis of structural stability for Switzerland, but only at 10% level, and the Q-A test does not identify any structural breaks in this series. Finally, the Chow SS suggests structural instability in case of Sweden and the Q-A finds a break in 1994Q2 for this country.

Comparing the growth rate and levels relationship estimates, the Q-A test results are similar in case of Belgium and Norway, but very different for the Netherlands and Sweden. Interestingly the Q-A test does not identify a structural break around 1998-1999, the third stage of the implementation of the EMU, when the relationship estimated either in levels or in first differences. However, for levels models the Q-A finds significant breakpoints for all countries, while in growth rate models only for seven economies, even including breakpoints that are only significant at 10%.

Furthermore, if in levels relationships the SS test suggests structural instability for all countries (three of them, Austria, Finland and the Netherlands, are only significant at 10% level), in the growth rate estimations the SS test rejects the null only for four countries (for Finland and Switzerland at 10% level). Overall, it appears the growth rate form fits the data better.

5.4.4 Level of Inflation Equation

We now estimate a level of inflation equation to check whether there are any patterns that clearly distinguish the EMU countries from the economies that did not join the monetary union. Therefore, we proceed by estimating the following equation:

$$\pi_{it} = \alpha_0' + \alpha_1' \pi_{i,t-1} + \beta_0' \pi_{it}^* + \beta_1' \pi_{i,t-1}^* + \varepsilon_t, \tag{5.7}$$

where π_{it} and π_{it}^{*} are the home and foreign inflation rates, respectively; π_{it-1} and π_{it-1}^{*} are the lagged home and foreign inflation rates, correspondingly, and ε_t is the error term. This equation is then used to forecast inflation rate as with the levels relationship over the following 72 quarters, where $t = T_1 + 1, T_1 + 2, ..., T$, with $T_2 = T - T_1$.

In the equation for Austria level and lagged foreign coefficients are positive and significant and the equation passes all diagnostic tests except one for normality. In case of Belgium, all variables are significant, but lagged foreign inflation rate coefficient is negative. The equation does not pass the serial correlation and heteroskedasticity tests. For Finland and France, however, foreign inflation (at 10% level) and lagged domestic inflation have significant positive effect on the dependent variable. As with Belgium, equations for these countries fail the serial correlation and heteroskedasticity tests. Moreover, equation for France fails the normality test as well.

In case of Germany, the lagged domestic inflation and current foreign inflation (at 10% level) have positive and significant effects. The equation passes all tests, except the normality test. In the equation for Italy only the coefficient of the lagged domestic inflation is significant, and it is positive. As for the diagnostic analysis, the equation passes the functional form test, but fails all the others.

For the Netherlands, the current foreign inflation and lagged domestic coefficients are positive and significant (the former is so only at 10% level). The equation fails the serial correlation and functional form tests as well as the normality test at 10% level. In equation for Norway, only the lagged domestic inflation coefficient is significant. The equation fails the serial correlation and normality tests.

In the case of Spain, the foreign inflation (at 10% level) and lagged domestic inflation have significant positive effect on the dependent variable. However, as it was with Norway, this equation also fails the serial correlation and normality tests. For Sweden, Switzerland and the UK the previous period domestic and current foreign inflation rates have significant positive effect on the dependent variable. The Swedish equation fails the serial correlation and normality tests, while the one for Switzerland does not pass the serial correlation and heteroskedasticity tests. The UK equation fail the normality and heteroskedasticity tests.

Overall, the lagged domestic inflation rate coefficient is significant and positive for all, but one country (Austria). The current foreign inflation rate is significant and positive for six countries (Austria, Belgium, Finland, Sweden, Switzerland and the UK) and for four more (France, Germany, the Netherlands and Spain) is so at 10% level. However, the lagged foreign inflation rate coefficient is only significant for Austria and Belgium. It is negative for Belgium, while positive for Austria. As for the diagnostic tests, only the Netherlands fail the functional form test. Nonetheless, nine out of 12 countries fail the serial correlation test (Belgium, Finland, France, Italy, the Netherlands, Norway, Spain, Sweden and Switzerland) and normality test (Austria, France, Germany, Italy, Norway, Spain, Sweden and the UK as well as the Netherlands at 10% level). In addition, half of the sample fail the heteroskedasticity test (Belgium, Finland, France, Italy, Switzerland and the UK).

We then proceed by running the structural break analysis (Table 5.4). Only equations for Norway and Spain (at 10% level) fail the Chow Predictive failure test. However, as was mentioned, the confidence interval for this test is quite wide, the power of this test is rather low. As for the SS test, the equations for most countries, with exception of Italy and three non-members (Sweden, Switzerland and the UK), are subject to structural breaks. It is then confirmed by the Q-A test, which also identifies breaks for all but one country (Switzerland). For Belgium, France, Italy and Sweden the break is suggested in 1985 (third or forth quarter), the beginning of the Great Moderation period, and around same time for the Netherlands (1987Q2) and Norway (1988Q3). Then in the equations for Germany and the UK the breaks occurred around mid-1990 and about a decade later in the equations for Austria (1999Q4) and Spain (2001Q4). Finally, in case of Finland, the Q-A suggests a break in 1997Q2. Therefore, even though there is a break almost in every equation, the breaks do not seem to be connected to the formation of the EMU and there is no clear difference between the euro and non-member countries.

Table 5.4 Structural Stability Analysis (Inflation; Level)

	PF[1] F(72,73)	SS[2] F(4,141)	Mean PE[3] Annual %	RMSPE[4] Annual %	Q-A Test
Austria	1.00	0.00	1.25	1.51	1999Q4***
Belgium	0.97	0.00	1.51	1.99	1985Q3***
Finland	0.81	0.00	0.97	1.72	1997Q2***
France	1.00	0.00	2.08	2.29	1985Q3***
Germany	1.00	0.00	1.11	1.42	1990Q2***
Italy	1.00	0.46	0.60	1.06	1985Q3***
Netherlands	0.85	0.00	1.12	1.65	1987Q2***
Norway*	0.00	0.00	0.03	2.77	1988Q3***
Spain	0.05	0.00	-0.22	2.59	2001Q4***
Sweden*	1.00	0.32	1.70	1.98	1985Q4***
Switzerland*	1.00	0.25	0.49	1.17	1986Q1
UK*	1.00	0.25	1.58	1.96	1990Q3***

Notes: *non-member countries; [1] - Chow Predictive failure test; [2] - Chow Structural stability test; [3] - Mean Prediction Errors; [4] - Root Mean Squared Prediction Errors; [5] - Quandt-Andrews test. *** indicates a significance level of 1 percent, ** of 5 percent, and * of 10 percent.

Looking at the mean PE and RMSPE values, most countries, except Spain, had higher inflation that forecast would have suggested. The highest difference between the predicted and actual inflation rate was in France, with mean PE of 2.09 and one of the highest RMSPE's in the sample (2.29). However, the largest RMSPE's are for Norway (2.77) and Spain (2.59). As can be seen from the graphs (an appendix, D.7) the forecast for Norway is very smooth, while the actual has large spikes around early 2000s, then during and after the financial crisis 2007-2008. In case of Spain the forecast fails to predict the large drop in the inflation rate around the financial crisis 2007-2008 and in 2015-2016, even though the average prediction errors are close to zero for both of them, 0.03 for Norway and -0.22 for Spain. Overall, the actual inflation rates for most countries are more jumpy than forecasts predicted.

5.4.5 Change in Inflation Equation

Furthermore, following the same procedure as with income level and growth analyses, we also estimate the inflation growth equation. The equation we estimate is

$$\Delta\pi_{it} = \alpha_0' + \alpha_1'\Delta\pi_{i,t-1} + \beta_0'\Delta\pi_{it}^* + \beta_1'\Delta\pi_{i,t-1}^* + \varepsilon_t, \tag{5.8}$$

where $\Delta\pi_{it}$ and $\Delta\pi_{it}^*$ are the changes in home and foreign inflation rates, respectively; $\Delta\pi_{it-1}$ and $\Delta\pi_{it-1}^*$ are the lagged changes in home and foreign inflation rates, correspondingly, and ε_t is the error term.

This equation is then used to forecast change in inflation rate as with the levels relationship over the following 72 quarters, where $t = T_1+1, T_1+2, ..., T$, with $T_2 = T - T_1$.

For Austria both lagged changes in inflation are significant, the domestic one is negative, while the foreign one is positive. The current change in the foreign inflation is positive, but significant only at 10% level. The equation fails the serial correlation test and heteroskedasticity test (at 10% level). In the equations for Belgium, Finland and Germany, however, the lagged change in foreign inflation is insignificant. The current change in foreign inflation is positive and significant, while the lagged change in domestic inflation has significant negative effect on the dependent variable. As in case of Austria, the equations for Belgium and Finland fail the serial correlation (for Finland 10% level) and the heteroskedasticity tests. The equation for Germany fails the serial correlation and normality tests.

In equations for France, Norway and Spain only the lagged change in the domestic inflation coefficient is significant, and it is negative. The equations for Norway and France fail the normality test, and the latter one fails the serial correlation test as well. The equation for Spain fail the serial correlation and heteroskedasticity (at 10% level) tests. For Italy all coefficients are insignificant, and the equation fails the serial correlation and normality tests.

For the Netherlands, current and lagged changes in foreign inflation are positive, the lagged one is significant only at 10% level. The lagged change in the domestic inflation is also significant, but negative. The equation passes all diagnostic tests.

For Sweden and Switzerland the change in the foreign inflation coefficient is significant and positive, while the lagged differenced domestic inflation coefficient is negative. The equation for Sweden fails the serial correlation, heteroskedasticity and functional form (at 10% level) tests, while the one for Switzerland fails the functional form and heteroskedasticity tests. For the UK only lagged coefficients are significant, the foreign one is positive, while the domestic one is negative. The equation fails the serial correlation, functional form and heteroskedasticity tests.

Overall, the current change in foreign inflation coefficient is significant and positive for seven countries, for Austria (at 10% level), Belgium, Finland, Germany, the Netherlands, Sweden and Switzerland. The lagged one is positive and significant for Austria, the Netherlands and UK. The lagged change in domestic inflation is negative and significant for all countries but Italy, where it has no significant effect on the dependent variable. In terms of how well-specified the equations are, nine countries fail the serial correlation test, namely Austria, Belgium, Finland (at 10% level), France, Germany, Italy, Spain, Sweden and the UK. In addition, eight economies do not pass the heteroskedasticity test (Austria and Spain at 10% level as well as Belgium, Finland, Sweden, Switzerland and the UK). However, only four countries fail the normality test (France, Germany, Italy and Norway), and three economies do not pass the functional form one (Sweden at 10% level, Switzerland and the UK).

Moving to the structural break analysis, the results for the level (Table 5.4) and change (Table 5.5) in inflation rate are similar for some countries. For instance, once again only the equations for Norway and Spain fail the PF test while all equations, except for Italy, the Netherlands and Sweden, fail the SS test, and in case of Austria and the UK the null hypothesis of structural stability is rejected at 10% level.

However, the Q-A test only finds structural breaks for seven countries (in Belgium only at 10% level). However, for Spain and the UK the breaks are in the last quarter of 2001 and in 1990Q3, respectively, exactly the same as was identified in the levels inflation equations for these countries. The mean PE are positive, but close to zero for many countries, such as for Switzerland (0.02), Belgium (0.03), Austria (0.04), Finland, Germany and Norway (0.07 for all three).

Overall, all countries appear to have higher growth of inflation rate than forecasts predicted (graphs are presented in an appendix, D.8) with the Netherlands being the only exception with the mean PE of -0.01. The forecast for Norway is least precise (RMSPE is

Table 5.5 Structural Stability Analysis (Inflation Growth)

	PF[1] F(72,73)	SS[2] F(4,141)	Mean PE[3] Annual %	RMSPE[4] Annual %	Q-A Test
Austria	1.00	0.09	0.04	1.03	2003Q2
Belgium	0.80	0.01	0.03	1.39	1997Q2*
Finland	0.80	0.00	0.07	1.49	1994Q4***
France	1.00	0.00	0.16	1.24	1996Q1***
Germany	1.00	0.03	0.07	1.16	1997Q2
Italy	1.00	0.49	0.14	0.84	1995Q3
Netherlands	0.67	0.26	-0.01	1.45	1987Q1
Norway*	0.00	0.00	0.07	4.17	2001Q3**
Spain	0.02	0.00	0.22	3.05	2001Q4***
Sweden*	1.00	0.53	0.08	1.43	1985Q4
Switzerland*	1.00	0.01	0.02	1.26	1999Q2**
UK*	1.00	0.08	0.20	1.29	1990Q3***

Notes: *non-member countries; [1] - Chow Predictive failure test; [2] - Chow Structural stability test; [3] - Mean Prediction Errors; [4] - Root Mean Squared Prediction Errors; [5] - Quandt-Andrews test. *** indicates a significance level of 1 percent, ** of 5 percent, and * of 10 percent.

4.17), with Spain having second largest RMSPE of 3.05. Italy and Austria, however, have the lowest RMSPE's in the sample, 0.84 and 1.03, correspondingly.

5.4.6 Short Interest Rate Equation

To consider how interest rate determination changed with the formation of the euro and the crisis, a Taylor Rule type model is estimated for the pre-euro period, 1979Q4-1998Q4, the early pre-crisis euro period, 1999Q1-2008Q4, and the post-crisis period 2009Q1-2016Q4. The estimated equation for each country made the short interest rate for each country, r_t, a function of its lagged value, the lagged rate of inflation, π_{t-1}, and the lagged log output gap, $(y_{t-1} - y^*_{t-1})$. This takes the form:

$$r_t = (1-\lambda)r_{t-1} + \lambda(\theta_* + \theta_\pi \pi_{t-1} + \theta_y(y_{t-1} - y^*_{t-1})) + u_t, \tag{5.9}$$

where u_t is the error term.

Potential output is approximated by a linear trend, t, so the estimated equation takes the form

$$r_t = \alpha_0 + \alpha_1 r_{t-1} + \beta_\pi \pi_{t-1} + \beta_y y_{t-1} + \gamma t + u_t. \tag{5.10}$$

In the pre-euro period $\beta_y > 0$ for every country and is significant for all but three, while $\beta_\pi > 0$ for 10 countries, significantly so for three and insignificantly negative for the Netherlands and Spain.

The average value of θ_π is 0.17 with a standard deviation of 0.26 with a range from -0.53 in the Netherlands, to 0.47 in Belgium. The average value of θ_y is 0.15 with a standard deviation of 0.07 with a range from 0.06 for Sweden to 0.29 for the Netherlands. In the second, early euro, period five countries show negative β_π coefficients (Belgium, France, Italy, the Netherlands and UK) none significantly negative, only Finland and Austria are significantly positive. The β_y coefficient is significantly positive for all countries except Switzerland. The results for the third post-crisis period are subject to the fact that interest rates moved relatively little over that period. There are two negative β_π, for Switzerland and the UK, neither significantly so, and three significantly positive, for Sweden, Italy and Austria. There are now six negative β_y, for Belgium, the Netherlands, Norway (significant), Spain, Sweden and Switzerland, and only for the UK it is significantly positive.

For most countries the match between actual and predicted is fairly close until the crisis when the predicted interest rate falls sharply and the actual interest rate constrained by the zero lower bound cannot follow (graphs are presented in an appendix, D.5). With the exception of Sweden and Switzerland, predicted interest rates go sharply negative at the end of the sample, being below the actual. For Sweden and Switzerland, predicted is above actual for the whole period. For Germany the predicted is below the actual for the whole period. For Italy and the Netherlands the predicted is below the actual from about 2002.

The hypothesis of structural stability, no change in the parameters before and after the establishment of the euro, is not rejected for Finland, Norway, Spain and Switzerland. Though the test is conditional on equality of variances, which is unlikely to be the case here. The standard error of regression for the pre-euro period is large relative to that for the euro and post-crisis periods.

Overall, there is little in these results that would suggest a big difference between the euro members and non-members. Moreover, the effect of the zero lower bound disturbs the forecasting tests (thus, we do not present detailed tables for Taylor Rule estimations).

5.4.7 Long Interest Rate Equation

We also estimate a Taylor Rule type model using long interest rates with the estimated equation taking the following form:

$$lr_t = \alpha_0 + \alpha_1 lr_{t-1} + \beta_\pi \pi_{t-1} + \beta_y y_{t-1} + \gamma t + u_t, \tag{5.11}$$

where lr_t is a nominal long-term interest rate per quarter.

Our sample is one country short, because the long-term interest rate data are not available for Finland.

In the pre-euro period β_y is positive for every country and is significant for all but four. The average value of θ_y is 0.11 with a standard deviation of 0.13 and with a range from 0.02 for the UK to 0.49 for Italy.

As for inflation coefficients, β_π is positive for all but one country, significantly so for three and insignificantly negative for Norway. The average value of θ_π is 0.20 with a standard deviation of 0.30. The coefficients of θ_π range from -0.11 for Norway to almost unity (0.97) for Italy. In the second period β_π is negative for all countries except Germany, Italy and Sweden, but none are significant. In contrast, β_y is positive for all countries, but significantly so only for Austria, Germany, the Netherlands and Norway. As for the post-crisis period, β_π is positive for all countries except Norway, Switzerland and the UK, but is insignificant for the whole sample. Moreover, β_y, is now insignificant for all countries and is negative for Germany, the Netherlands, Sweden, Switzerland and the UK.

As before, the results do not vary substantially between the EMU members and stand-alone European countries in our sample. The differences between actuals and forecasts are smaller for all countries when a Taylor Rule type equation is estimated using long-term interest rates instead of short-term ones (graphs are presented in an appendix, D.6).

5.5 Systems approach

5.5.1 Theoretical Model

Although we are not going to use a structural model, we will set out how our estimated model relates to a fully specified structural model. Consider the following rational expectations (RE) model for a small open economy in the $k \times 1$ vector $\mathbf{x}_t$ of endogenous variables, determined in terms of their expected future values and past values, a $k_* \times 1$ vector of corresponding foreign variables, $\mathbf{x}_t^*$, which are treated as exogenous and a $k_d \times 1$ vector of deterministic elements like trend and intercept:

$$\mathbf{A}_0(\varphi)\mathbf{x}_t = \mathbf{A}_1(\varphi)E_t(\mathbf{x}_{t+1}) + \mathbf{A}_2(\varphi)\mathbf{x}_{t-1} + \mathbf{A}_3(\varphi)\mathbf{x}_t^* + \mathbf{A}_4(\varphi)\mathbf{d}_t + \mathbf{u}_t. \tag{5.12}$$

The expected future values $E_t(\mathbf{x}_{t+1}) = E(\mathbf{x}_{t+1} \mid \mathfrak{I}_t)$. The information set is

$$\mathfrak{I}_t = \left(\mathbf{x}_t, \mathbf{x}_{t-1}, ...; \mathbf{x}_t^*, \mathbf{x}_{t-1}^*, ...\right),$$

and $\mathbf{A}_i(\varphi)$ are matrices of coefficients. For $i = 0,1,2$, they are dimension $k \times k$, for $i = 3$ dimension $k \times k_*$, for $i = 4$ dimension $k \times k_d$.

$\mathbf{A}_0(\varphi)$ is non-singular, φ is a vector of deep parameters, and $\mathbf{u}_t$ is a $k \times 1$ vector of structural shocks. The exogenous variables are assumed to follow the VAR(1) model:

$$\mathbf{x}_t^* = \mathbf{a}(\rho) + \mathbf{R}(\rho)\mathbf{x}_{t-1}^* + \eta_\mathbf{t}, \tag{5.13}$$

where $\mathbf{a}(\rho)$ is a $k_x \times 1$ vector of intercepts, and $\mathbf{R}(\rho)$ is the $k_x \times k_x$ matrix of coefficients that depend on a vector of unknown coefficients, ρ. This marginal model is required because forecasts of $\mathbf{x}_t^*$ are required to construct the expectations $E_t(\mathbf{x}_{t+1})$. The errors, $\mathbf{u}_t$ and $\eta_\mathbf{t}$ are assumed to be serially and cross-sectionally uncorrelated, with zero means and constant variances, Σ_u and Σ_η, respectively.

If the quadratic matrix equation,

$$\mathbf{A}_1(\varphi)\Phi^2(\varphi) - \mathbf{A}_0(\varphi)\Phi(\varphi) + \mathbf{A}_2(\varphi) = \mathbf{0}, \tag{5.14}$$

has a solution, $\Phi(\varphi)$, with all its eigenvalues inside the unit circle, then, the RE model, (5.12) and (5.13), has the unique solution[2]

$$\mathbf{x}_t = \Phi(\varphi)\mathbf{x}_{t-1} + \Psi(\varphi,\rho)\mathbf{x}_t^* + \mu_a(\varphi,\rho)\mathbf{d}_t + \Gamma(\varphi)\mathbf{u}_t. \tag{5.15}$$

The variance matrix of the reduced form shocks, $\varepsilon_t = \Gamma(\varphi)\mathbf{u}_t$, is

$$\Sigma_\varepsilon(\varphi) = E(\varepsilon_t\varepsilon_t') = \Gamma(\varphi)\Sigma_u\Gamma'(\varphi). \tag{5.16}$$

Equation (5.15) is labelled a VARX* in the GVAR literature. It corresponds to the reduced form of a standard simultaneous equations model, when $\mathbf{A}_1(\varphi) = 0$ and there are no forward looking terms. It corresponds to a vector autoregression when there are no exogenous variables, so $\mathbf{A}_3(\varphi) = \Psi(\varphi,\rho) = 0$. What is relevant for the case of the euro is that the parameters of (5.15) may change either because the parameters of the process generating the endogenous variables, φ, change say from φ_1 to φ_2; or of the process generating the exogenous variables, ρ, changes from ρ_1 to ρ_2. Changes in the process driving the exogenous variables could be important because they may change how people form their expectations $E_t(\mathbf{x}_{t+1})$.

[2]See, for instance, Pesaran (2015, Chapter 20).

5.5.2 VARX*

The VARX* (5.15) was for a single country, now consider a set of countries $i = 0, 1, 2, ..., N$, with country 0, say the US, as the numeraire country: we use the exchange rate against the dollar. Suppressing the dependence on the deep parameters, a second-order country-specific VARX*(2,2) model with deterministic trends can be written as

$$\mathbf{x}_{it} = \mathbf{B}_{id}\mathbf{d}_t + \mathbf{B}_{i1}\mathbf{x}_{i,t-1} + \mathbf{B}_{i2}\mathbf{x}_{i,t-2} + \mathbf{B}^*_{i0}\mathbf{x}^*_{it} + \mathbf{B}^*_{i1}\mathbf{x}^*_{i,t-1} + \mathbf{B}^*_{i2}\mathbf{x}^*_{i,t-2} + \mathbf{u}_{it}, \tag{5.17}$$

where $\mathbf{x}_{it}$ is a $k_i \times 1$ (usually six) vector of domestic variables, $\mathbf{x}^*_{it}$ is a $k^*_i \times 1$ vector of foreign variables specific to a country i and $\mathbf{d}_t$ is a $s \times 1$ vector of deterministic elements as well as observed common variables, oil prices in our case: $(1, t, p^o_t)'$. The $\mathbf{x}^*_{it}$ are calculated as country-specific trade weighted averages of the corresponding variables of the other countries,

$$\mathbf{x}^*_{it} = \sum_{j=0}^{N} w_{ij}\mathbf{x}_{jt}, \text{with} w_{ij} = 0, \tag{5.18}$$

where w_{ij} is the share of country j in the trade (exports plus imports) of country i.

In the case of small open economies it is reasonable to assume that the $\mathbf{x}^*_{it}$ are "long-run forcing " or $I(1)$ weakly exogenous, and then estimate the VARX* models separately for each country, allowing for cointegration both within $\mathbf{x}_{it}$ and across $\mathbf{x}_{it}$ and $\mathbf{x}^*_{it}$.[3] Tests for the weak exogeneity of $\mathbf{x}^*_{it}$ generally do not reject the hypothesis. The $\mathbf{x}^*_{it}$ would typically refer to the same variables as $\mathbf{x}_{it}$, thus, there is a symmetric structure to the model.

The cointegrating VARX* can be written as a VECM:

$$\Delta\mathbf{x}_{it} = \mathbf{B}_{id}\mathbf{d}_t + \Pi_i\mathbf{z}_{i,t-1} + \mathbf{B}^*_{i0}\Delta\mathbf{x}^*_{it} + \Gamma_i\Delta\mathbf{z}_{i,t-1} + \mathbf{u}_{it}, \tag{5.19}$$

where $\mathbf{z}_{it} = (\mathbf{x}'_{it}, \mathbf{x}^{*\prime}_{it})'$. Restricting the deterministic terms and assuming that $rank(\Pi_i) = r_i < k_i + k^*_i$, we have $\Pi_i = \alpha_i\beta'_i$, where β_i is the $(k_i + k^*_i) \times r_i$ matrix of the cointegrating coefficients and

$$\Delta\mathbf{x}_{it} = \alpha_i\beta'_i\left(\mathbf{z}_{i,t-1} - \Upsilon_i\mathbf{d}_{t-1}\right) + \mathbf{B}^*_{i0}\Delta\mathbf{x}^*_{it} + \Gamma_i\Delta\mathbf{z}_{i,t-1} + \Pi_i\Upsilon_i\Delta\mathbf{d}_t + \mathbf{u}_{it}. \tag{5.20}$$

The r_i error correction terms of the model can now be written as

[3] This is unlikely to apply to a large economy like the US which may influence world interest rates. However, it seems reasonable for the European countries we consider.

$$\xi_{it} = \beta_i' \mathbf{z}_{it} - \beta_i' \Upsilon_i \mathbf{d}_t = \beta_{ix}' \mathbf{x}_{it} + \beta_{ix*}' \mathbf{x}_{it}^* + \gamma_i' \mathbf{d}_t, \tag{5.21}$$

where ξ_{it} are mean zero $r_i \times 1$ vector of disequilibrium deviations from the long-run relationships. Forecasts and counterfactuals are invariant to the just-identifying restrictions used to identify β_i'. To establish whether there are changes in the reduced form parameters, we do not need to identify either the structural shocks or the cointegrating vectors.

Notice that if $r = 0$ in (5.20), this gives a first difference model and if $r = k$ this gives an unrestricted levels VARX* (5.17).

5.5.3 Model Selection

We wish to examine whether there was a break at time T_0, the end of 1998Q4. To do this, we estimate (5.20) using the whole sample: 1979Q4-2016Q4, which we call period 0; then for period 1 (1979Q4-1998Q4) and period 2 (1999Q1-2016Q4) and examine the extent to which allowing for a structural break improves the fit. In estimating (5.20), for each country we have to (a) choose lag-length for endogenous and exogenous variables, p_{ei}, p_{mi}; which are set to a maximum of (2,2), (b) choose the number of cointegrating vectors r_i and (c) judge the significance of any structural breaks. Although there are tests for each of these, some of which are non-standard, it seems better to make the choices for the various elements within a consistent framework. This can be done using information criteria (IC).

If model i, has k_i estimated parameters and maximised log likelihood MLL_i, the Akaike information criterion is $AIC_i = MLL_i - k_i$. The Schwarz Bayesian information criterion is $BIC_i = MLL_i - 0.4 * k_i \ln T$; where T is the sample size.[4] Two models are estimated, one using p_{ei}^A, p_{mi}^A and r_i^A, chosen on the basis of AIC and one using p_{ei}^B, p_{mi}^B and r_i^B chosen on the basis of BIC. We use period 1, pre-euro, data to make the choice.

In the case of nested models, which is what we will be concerned with, the information criteria can be interpreted as likelihood ratio tests. Suppose that the unrestricted model with MLL_U has one more parameter than the restricted model with MLL_R. Then a standard likelihood ratio test with probability of type I error, $\alpha = 5\%$, would choose the unrestricted model if $LR = 2(MLL_u - MLL_R) > 3.84$. The AIC would choose the unrestricted model if $2(MLL_u - MLL_R) > 2$, corresponding to roughly $\alpha = 15\%$. The BIC would choose the unrestricted model if $2(MLL_u - MLL_R) > \ln T$, which for $T = 100$ is 4.6, which corresponds to rougly $\alpha = 3\%$. The AIC and LR keep α constant and use any extra information to increase the power. They will reject any deviation from the null, however small, for a sufficiently large sample size. The BIC reduces α with the sample size, so the probabilities of both type

[4] Some programs report -2 times these numbers.

I and type II errors fall with sample size. The BIC is consistent in that it will choose the true model as the sample size gets larger, if the true model is in the set being considered. If the true model is not in the set being considered, the AIC (by including more parameters) may provide a better approximation to it. By using both we can judge how robust our results are to possible over-fitting or under-fitting.

We will use the difference between the sum of the IC for periods 1 and 2 and the IC for the whole period as an indication of the extent of the structural break. For the AIC, the difference $D_A = AIC_1 + AIC_2 - AIC_0$, and $2D_A = LR - K$, where K is the number of parameters. Hence, the AIC choice corresponds to an LR test with a critical value of K. Similarly, $D_B = BIC_1 + BIC_2 - BIC_0$ which corresponds to an LR test with a critical value of $K(\ln T1 + \ln T2 - \ln T)$. The dimensions are $T1 = 77$, $T2 = 72$, $T = 149$, and with $p_{ei}, p_{mi} = (2,2)$ and $r_i = 6$, there are 25 parameters in each equation.

5.5.4 Conditioning

In order to identify where the structural changes originate, we will condition on some elements of $\mathbf{x}_{it}$, say $\mathbf{x}_{2,it}$, in particular interest rates and exchange rates, and treat them as exogenous, in explaining $\mathbf{x}_{1,it}$. The interpretation of this process follows Pesaran and Smith (1998). For clarity of exposition we abstract from the country identifier, i, the deterministic terms $\mathbf{d}_t$ and the other exogenous foreign variables, $\mathbf{x}_{it}^*$. Then (5.19) can be written as:

$$\Delta \mathbf{x}_t = \Pi x_{t-1} + \Gamma \Delta \mathbf{x}_{t-1} + \mathbf{u}_t. \tag{5.22}$$

We now partition $\mathbf{x}_t = (\mathbf{x}_{1t}', \mathbf{x}_{2t}')'$ to give

$$\Delta x_{1t} = \Pi_{11}\mathbf{x}_{1,t-1} + \Pi_{12}\mathbf{x}_{2,t-1} + \Gamma_{11}\Delta\mathbf{x}_{1,t-1} + \Gamma_{12}\Delta\mathbf{x}_{2,t-1} + u_{1t}$$
$$\Delta x_{2t} = \Pi_{21}\mathbf{x}_{1,t-1} + \Pi_{22}\mathbf{x}_{2,t-1} + \Gamma_{21}\Delta\mathbf{x}_{1,t-1} + \Gamma_{22}\Delta\mathbf{x}_{2,t-1} + u_{2t},$$

where the covariance matrix of the reduced form disturbances is given by:

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

This partition does not impose any restrictions in itself, but provides a framework for examining how exogenous variables relate to the structure of the *VAR*.

To condition $\mathbf{x}_{1t}$ on current values of $\mathbf{x}_{2t}$, define $\mathbf{E}(\mathbf{u}_{1t} \mid u_{2t}) = \Sigma_{12}\Sigma_{22}^{-1}\mathbf{u}_{2t} = \Theta u_{2t}$, with $\mathbf{u}_{1t} = \Theta u_{2t} + \eta_t$. The system for $\mathbf{x}_{1t}$, can then be written as

$$\begin{aligned}\Delta x_{1t} \quad = \quad & (\Pi_{11} - \Theta\Pi_{21})x_{1,t-1} + (\Pi_{12} - \Theta\Pi_{22})x_{2,t-1} + \Theta\Delta x_{2,t} + \\ & + (\Gamma_{11} - \Theta\Gamma_{21})\Delta x_{1,t-i} + (\Gamma_{12} - \Theta\Gamma_{22})\Delta x_{2,t-i} + \eta_t.\end{aligned} \tag{5.23}$$

$$\Delta x_{1t} = \mathbf{B}_1\mathbf{x}_{1,t-1} + \mathbf{B}_2\mathbf{x}_{2,t-1} + \mathbf{C}_{20}\Delta x_{2t} + \mathbf{C}_{11}\Delta x_{1,t-1} + \mathbf{C}_{21}\Delta x_{2,t-1} + \eta_t.$$

By construction $E(\eta_t \mid \Delta x_t) = \mathbf{0}$, and the parameters of (5.23) can be estimated efficiently by OLS.

Also, denoting the (conditional) variance of η_t by $\Sigma_{\eta\eta}$, it is easily seen that

$$\Sigma_{\eta\eta} - \Sigma_{11} = -\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \leq \mathbf{0}.$$

The variance of η_t will generally be smaller than that of $\mathbf{u}_{1t}$, so the parameters in the conditional model, (5.23), are likely to be estimated more precisely than the parameters of the unconditional model (5.22). Whether this is an advantage depends on what the economic parameters of interest are. If the parameters of interest are $\Pi = (\Pi_{11}, \Pi_{12})$, it is clear from equation (5.23) that Δx_{2t} will be weakly exogenous for Π_{11} only if either Σ_{12}= 0, so that $\Theta = 0$, or if Π_2= (Π_{21}, Π_{22})= 0.[5]

In either of these cases the coefficient matrix on $(\mathbf{x}_{1,t-1}, \mathbf{x}_{2,t-1})$ in the conditional model (5.23) will provide an estimate of Π, otherwise it will not. In other cases, the economic parameters of interest may be the long-run effects of $\mathbf{x}_{2t}$ on $\mathbf{x}_{1t}$ so one might be interested in $(\Pi_{12} - \Theta\Pi_{22})$ directly, in which case the model conditional on $\mathbf{x}_t$ is appropriate whether or not Π_2= $\mathbf{0}$.

For some purposes we are interested in the complete system, but for other purposes we are interested in the responses to particular policy variables and how these responses changed with the introduction of the euro. In this case, the parameters of interest are the parameters of the conditional model, (5.23). Of particular interest is the case where the parameters of the marginal model, the process generating the policy variables, interest rates and exchange rates, changed, shifting Π_{21}, Γ_{22} and Θ, but the parameters of the conditional model, $\mathbf{B}_i$, $\mathbf{C}_{ij}$ did not change.[6]

[5] When the restrictions Π_2= 0 hold, $\mathbf{x}_{2t}$ is referred to as "long-run forcing" for $\mathbf{x}_{1t}$. This is different from Granger non-causality (GNC). $\mathbf{x}_{2t}$ is said to be GNC for $\mathbf{x}_{1t}$ if $\Pi_{12} = \mathbf{0}$ and $\Gamma_{12} = \mathbf{0}$; $\mathbf{x}_{2t}$ does not predict $\mathbf{x}_{1t}$. If $\Pi_2 = \mathbf{0}$, $\mathbf{x}_{2t}$ cannot themselves be cointegrated.

[6] As is clear from (5.15) the Lucas critique says that any change in the marginal model determining policy will change the conditional model.

5.6 Systems Results

The lag orders for the endogenous (p_e) and exogenous (p_m) variables as well as the number of cointegrating vectors (r) could be chosen from the pre-euro sample; the post-euro sample or the whole sample. We determined them on the basis of the pre-euro sample, since any subsequent change will appear as a structural break.

Thus for each country we estimate a cointegrating VARX* for period 1: 1979Q4-1988Q4, and use the AIC to determine p^A_{ei}, p^A_{mi} and r^A_i, and the BIC to determine p^B_{ei}, p^B_{mi} and r^B_i. Using these values we estimate the AIC model and the BIC model for period 2: 1999Q1-2016Q4 and for period 0, with no break: 1979Q4-2016Q4. This gives results summarised in the Table 5.6.

Table 5.6 Akaike and Bayesian Information Criteria for a Euro Structural Break, m=6

	Model	r	Order e,m	IC by Period 0	1	2	Difference	Rank
Austria	A	5	2,1	3362.7	1799	1842.6	278.9	1
	B	1	1,1	3238.5	1683	1727.3	171.8	1
Belgium	A	6	2,2	3377.8	1728.6	1866.4	217.2	6
	B	1	1,1	3254.3	1625.6	1746.4	117.7	7
Finland	A	4	2,1	2240.2	1131.4	1256	147.2	9
	B	0	1,1	2153.4	1054.6	1145.8	47	11
France	A	4	2,1	3524.1	1746.5	2024.3	246.7	3
	B	0	1,1	3397.9	1655.3	1899.1	156.5	2
Germany	A	6	2,2	3508.3	1775.7	1956.5	223.9	4
	B	3	1,1	3365.5	1658	1831.9	124.4	5
Italy	A	6	2,2	3182.2	1580.4	1800.7	198.9	7
	B	1	1,1	3034.6	1460.4	1678.5	104.4	8
Netherlands	A	6	2,1	3547.7	1792.9	1975.4	220.6	5
	B	1	1,1	3435.2	1698.8	1856.6	120.2	6
Norway*	A	5	2,1	3041.2	1569.7	1615.6	144.1	10
	B	2	1,1	2871.6	1451.8	1507.2	87.4	9
Spain	A	6	2,1	3141.1	1620.1	1770.2	249.2	2
	B	1	1,1	2973.7	1474.7	1647.7	148.7	3
Sweden*	A	6	2,1	3093.8	1507.1	1768.9	182.2	8
	B	0	1,1	2948.8	1404.6	1670.3	126.1	4
Switzerland*	A	4	2,1	3423.9	1765.5	1740.2	81.8	12
	B	0	1,1	3294.3	1649.3	1660.2	15.2	12
UK*	A	5	1,1	3409.9	1711.7	1839.1	140.9	11
	B	2	1,1	3283.1	1596.6	1749.1	62.6	10

Notes: *non-member countries. Model A is chosen by AIC, Model B by BIC; r is the number of cointegrating vectors; e, m give the lag orders on endogensous and exogenous variables, respectively. The information criteria are given for periods: 0: 1979Q4-2016Q4; 1: 1979Q4-1998Q4; 2: 1999Q1-2016Q4. Difference gives the value for IC0−(IC1+IC2), rank gives rank of the difference. Endogenous variables are y_{it}, π_{it}, eq_{it}, ep_{it}, r_{it} and lr_{it}; except Finland where lr_{it} is not available. Exogenous variables are y^*_{it}, π^*_{it}, eq^*_{it}, r^*_{it}, lr^*_{it} and $poil_{it}$.

As one would expect the AIC model tends to have larger values of p_{ei}, p_{mi} and r_i. In some cases, such as Sweden, $r_i^A = 6$, indicating that all the variables are $I(0)$, while $r_i^B = 0$, indicating that all the variables are $I(1)$ and not cointegrated. The differences of AIC and BIC are all positive, indicating that the model with a break at the time of the euro formation is preferred. The differences are always smaller for the BIC than for the AIC because the BIC imposes a heavier penalty for the extra parameters in the break model. If we rank AIC by differences, the break seems smaller for the non-euro countries: Sweden had the 8th smallest difference, Norway 10th, the UK 11th and Switzerland 12th. Among the euro countries, Finland (ranked 9th) had the smallest difference. The ranking by BIC is similar, the difference in ranks is small except for Sweden which goes from 8th by AIC to 4th by BIC. The Table 5.6 indicates that there is evidence for a break and it seems larger in the euro countries than the non-euro countries.

We next repeat the exercise treating the short interest rate, r_{it}, as exogenous (since it is controlled externally by the ECB for the euro countries) in the second period (Table 5.7).

Table 5.7 Akaike and Bayesian Information Criteria for a Euro Structural Break, m=6

	Model Criterion	k=6	k=5	k=4
Austria	A	278.90	95.90	75.10
	B	171.80	-2.80	-10.10
Belgium	A	217.20	106.70	82.20
	B	117.70	-1.70	-14.30
Finland	A	147.20	56.17	39.50
	B	47.00	-27.67	-13.50
France	A	246.70	132.10	110.70
	B	156.50	51.60	51.10
Germany	A	223.90	152.90	120.80
	B	124.40	59.90	33.30
Italy	A	198.90	152.70	123.60
	B	104.40	69.90	72.10
Netherlands	A	220.60	113.00	66.20
	B	120.20	30.20	21.30
Norway*	A	144.10	92.50	73.10
	B	87.40	14.60	13.00
Spain	A	249.20	98.90	71.50
	B	148.70	-16.80	-7.60
Sweden*	A	182.20	97.60	78.60
	B	126.10	45.00	15.40
Switzerland*	A	81.80	63.70	36.60
	B	15.20	-13.50	-16.40
UK*	A	140.90	72.50	50.40
	B	62.60	5.60	-6.10

Notes: *non-member countries. Difference in AIC, A, and BIC, B, between whole period and 2 sub-periods, for 3 models; k=5, interest rates exogenous, k=4, interest rates and exchange rates exogenous. Negative value suggests no structural break.

We thus see how large the structural break is controlling for interest rates. Since the size of the system changes from $k=6$ to $k=5$, (except for Finland where it changes from $k=5$ to $k=4$) we again need to choose p^A_{ei}, p^A_{mi} and r^A_i, and p^B_{ei}, p^B_{mi} and r^B_i on the first period data. Assuming the real exchange rate, ep_{it}, is controlled by the ECB, to an extent after the formation of the euro, we continue the analysis treating the exchange rate as exogenous. Hence, further reducing the number of endogenous variables in the system, from $k=5$ to $k=4$, except for Finland (in which case $k=4$ is reduced to $k=3$). The Table (5.7) gives the differences in the AIC and BIC between the whole period and the sum of the two sub-periods for the three cases. For $k=6$, the first column, the differences are the same as in the previous table. When one controls for the short interest rate and exchange rate, there is clearly less evidence for a structural break, suggesting that the main breaks came in interest rate and exchange rate equations. For half of the sample, namely for Austria, Belgium, Finland, Spain, Switzerland and the UK, the difference of BIC is negative, meaning the model estimated over the whole period is preferred. The big reduction appears to come from the interest rate equation.

5.7 Systems Forecasts

In addition, we perform multivariate dynamic forecasts by estimating the cointegrating VAR model. It includes trend, and the forecast is conditional on the supplied values of the exogenous variables namely the foreign GDP, rate of inflation, equity price index, short-term and long-term interest rate per quarter as well as the nominal price of oil. We estimate cointegrating VAR up to last quarter of 1998, using BIC values for specification (to choose optimal number of lags and cointegrating vectors) from the Table 5.6.

We perform these system forecasts for level of GDP and then for the income growth. We then use these estimates to compare with forecasts produced using OLS for single equation and then for system one. There is difference in interpretation of these four estimates, summarised in the Table 5.8, for euro countries and for non-members. For instance, for a euro member, such as Italy, these would be interpreted as four different measures of the effect of the euro on GDP, plus prediction errors. For a non-member country, such as the UK, they are just prediction errors. For Italy, the Netherlands, Norway and Spain, the errors are all negative, in each case the actual was below the forecast. Meanwhile, for Belgium, Sweden and Switzerland the errors are positive, suggesting that countries performed better than forecast suggested they would.

The levels are cumulative effects, hence, for instance, Italian GDP is between 7% and 9% below what would have been predicted. The Italian estimates of growth are very similar

Table 5.8 Mean Error of Forecast Percent (Level), Percentage Points at Annual Rates (Growth)

	Level		Growth	
	Single	System	Single	System
Austria	0.84	-0.05	-0.07	-0.05
Belgium	5.20	1.32	0.09	0.04
Finland	10.33	-1.96	-0.16	-0.15
France	3.88	0.22	-0.05	-0.06
Germany	-0.07	-2.61	0.02	0.05
Italy	-6.79	-8.94	-0.29	-0.31
Netherlands	-5.95	-4.79	-0.16	-0.14
Norway*	-11.02	-9.32	-0.32	-0.28
Spain	-8.80	-1.67	-0.15	-0.17
Sweden*	14.65	10.94	0.31	0.37
Switzerland*	13.10	3.05	0.15	0.08
UK*	-58.78	0.74	-0.04	-0.13

Notes: *non-member countries; all values are in percentages.

between the system and the single equations, both about 1.2% per annum lower: slightly below pre-crisis and substantially below post-crisis. As in case of Italy, for the Netherlands the system and single equation estimates are also very similar for both, level and growth.

For France, Germany and the UK, there are two negative and two positive estimates. For example, the UK level is either 59% below (this is an explosive forecast), or 0.7% above (this is an average of the pre-crisis period, where the actual was above the forecast, and a post-crisis period, where it was below).

For some countries the single and system level forecasts are completely different. For instance, for Finland the single level estimate is 10.33%, while the system one is smaller in magnitude and of the opposite sign (-1.96%). Meanwhile, growth forecasts for Finland are quite similar, -0.64% per annum using single equation and -0.6% per annum using the VAR. Both of them are only proportion of a 1%, as levels based forecast estimates are accumulated, while growth ones add up over time, but are quite small in any particular year. In case of Austria, level single and system estimates are not as different as in case of Finland, however, they are also of the opposite signs. Nonetheless, the growth estimates are once again closer to each other, around 0.2% per annum below the forecast.

Overall, more than half of the sample performed worse than forecasted. From the graphs (an appendix, D.3 and D.4) we see that the growth rate is slightly below predicted in the pre-crisis period and substantially below that in the post-crisis period. However, looking at the Table 5.8, there does not seem to be any clear pattern in these estimates and no obvious difference between the euro and non-member countries.

Unlike the levels case, the single equation and systems estimates for growth are quite similar and all have the same sign. Two of the four non-euro countries and six of the eight euro countries did worse than forecast. Meanwhile, Belgium, Germany, Sweden and Switzerland did better.

5.7.1 Sensitivity to Exogeneity

In addition, we check whether treating short interest rate and exchange rate as exogenous will significantly change our results. In order to do so, we estimate a model with now four endogenous variables (y_{it}, π_{it}, eq_{it} and lr_{it}) instead of six (where r_{it} and ep_{it} were considered endogenous) and use model specifications suggested by BIC from the Table 5.6. We then compare the mean errors of forecast errors for level of income and growth for six-variable model with those for four-variable model. The results are summarised in the Table 5.9.

Table 5.9 Effect of Interest and Exchange Rates: Mean Error of Forecast 1999Q1-2016Q4, GDP (System)

	6-Variable Model[1]		**4-Variable Model**[2]		**3-Variable Model**[3]	
	Level (%)	**Growth (%)**	**Level (%)**	**Growth (%)**	**Level (%)**	**Growth (%)**
Austria	-0.05	-0.05	0.18	-0.04	-0.14	-0.05
Belgium	1.32	0.04	1.91	0.05	1.90	0.05
Finland	-1.96	-0.15	-1.25	-0.13		0.00
France	0.22	-0.06	0.37	-0.06	0.12	-0.07
Germany	-2.61	0.05	-4.32	-0.01	-8.77	-0.10
Italy	-8.94	-0.31	-8.27	-0.29	-8.36	-0.29
Netherlands	-4.79	-0.14	-4.99	-0.15	-4.73	-0.14
Norway*	-9.32	-0.28	-13.56	-0.33	-11.15	-0.33
Spain	-1.67	-0.17	-0.48	-0.13	-0.41	-0.13
Sweden*	10.94	0.37	11.03	0.37	12.02	0.39
Switzerland*	3.05	0.08	2.81	0.08	3.78	0.10
UK*	0.74	-0.13	-5.96	-0.53	-5.14	-0.48

Notes: *non-member countries. Level in percent; Growth in percentage points at annual rates. [1]6-Variable Model: y_{it}, π_{it}, eq_{it}, ep_{it}, r_{it} and lr_{it} are endogenous variables; [2]4-Variable Model: y_{it}, π_{it}, eq_{it} and lr_{it} are endogenous variables; [3]3-Variable Model: y_{it}, π_{it} and eq_{it} are endogenous variables.

For most countries mean values do not seem to change much whether short interest rate and exchange rate are treated as endogenous or exogenous variables. This is especially true for the growth equations, such in case of France (-0.06), Sweden (0.37) and Switzerland (0.08) the mean errors of forecast do not change at all after switching from six-variable to four-variable model. In addition, if we consider long interest rate to be exogenous as well, the estimations for now three-variable model are still not very different from the original ones (Table 5.9). The values for the three-variable model for Finland are missing since the long-term interest rate data are not available for this country.

We arrive to the same conclusion when analyse the root mean sum of squares for the three models. Results are presented in the Table 5.10. In some cases the estimates are somewhat unusual, for instance, for the level of income for Austria the coefficient jumps from 1.59 in a six-variable model to 0.18 in a four-variable model and then jumps up again to 1.44 when we treat long interest rate as exogenous as well.

Table 5.10 Effect of Interest and Exchange Rates: RMSPE[1] of Forecast 1999Q1-2016Q4, GDP (System)

	6-Variable Model[2]		**4-Variable Model**[3]		**3-Variable Model**[4]	
	Level (%)	**Growth (%)**	**Level (%)**	**Growth (%)**	**Level (%)**	**Growth (%)**
Austria	1.59	0.82	0.18	0.81	1.44	0.81
Belgium	1.62	0.56	2.16	0.51	2.19	0.51
Finland	4.60	1.00	4.19	1.00		
France	2.84	0.36	2.80	0.35	2.92	0.35
Germany	4.12	0.53	5.18	0.51	9.82	0.59
Italy	12.09	0.52	11.51	0.53	11.60	0.53
Netherlands	6.39	0.64	6.58	0.65	6.27	0.56
Norway*	11.32	1.13	14.96	1.25	15.59	1.24
Spain	6.99	0.62	6.55	0.64	6.50	0.64
Sweden*	13.04	0.84	13.15	0.84	14.27	0.86
Switzerland*	4.15	0.73	3.92	0.72	5.05	0.78
UK*	5.32	0.62	14.18	1.07	12.99	1.02

Notes: *non-member countries. Level in percent; Growth in percentage points at annual rates. [1]RMSPE - Root Mean Squared Prediction Errors. [2]6-Variable Model: y_{it}, π_{it}, eq_{it}, ep_{it}, r_{it} and lr_{it} are endogenous variables; [3]4-Variable Model: y_{it}, π_{it}, eq_{it} and lr_{it} are endogenous variables; [4]3-Variable Model: y_{it}, π_{it} and eq_{it} are endogenous variables.

In addition, in case of the UK the mean error of forecast for the level income changes from positive (0.74) to negative (-5.96 and -5.14 in four-variable and three-variable models, respectively). Meanwhile, the RMSPE increases from 5.32 to 14.18 in four-variable model and drops to 12.99 in three-variable one. Moreover, in case of Norway the change is not as substantial, the mean error for level decreases from -9.32 to -13.56 while the RMSPE increased from 11.32 to 14.96, which is again usual. However, for both these countries BIC suggested a presence of two cointegrating vectors. If instead we assume just one, the estimates appear to be more consistent. For instance, in case of the UK the mean prediction error for the level income changes from 2.86 (instead of 0.74) in six-variable model to 2.04 (instead of -5.96) in four-variable and three-variable ones, while the RMSPE remains around 4% throughout.

Overall, the variance does not change much from model to model and this, once again, is especially pronounced for income growth output. For instance, for some countries, such as for Belgium the RMSPE decreases from 0.56 in a six-variable model to 0.51 in a four-variable and six-variable ones. In other countries, for instance, for Italy it increases slightly from 0.52 to 0.53. Overall, there is no consistency among the countries, hence, there is no evidence that using actual values for short and long interest rates as well as for the exchange rate substantially changes the estimates.

To summarise, we find that there is more evidence for a structural break in the interest rate and exchange rate equations when $k = 6$ (Table 5.7) than when they were treated as exogenous (as in case of $k = 5$ and $k = 4$ the BIC changed to negative for most countries or decreased substantially, meaning splitting the estimation period in two sub-periods did not improve the fit of the model). Thus, we check the sensitivity of our estimates to exogenity by conditioning on these additional variables (short and long interest rates as well as exchange rate). Nonetheless, predictions do not change much when we change k and, hence, we conclude that there is no big difference between using interest and exchange rate values as they are being predicted from the period before the euro versus their actual values.

5.7.2 Structural Break in 2008Q1 (Forecast Based on 1979Q2-2007Q4)

Since some of the above results seem unusual, we consider possibility that the structural break occurred not in 1999Q1, but during the financial crisis 2007-2008. Hence, we repeat the same estimations for a structural break in 2008, and the forecast in this case is based on the estimations for 1979Q2-2007Q4.

Table 5.11 Lag Orders and Cointegration Rank Chosen by BIC for a Structural Break in 2008Q1 (m=6, 1979Q2-2007Q4)

	r	Order e,m
Austria	1	1,1
Belgium	1	1,1
Finland	0	1,1
France	1	1,1
Germany	2	1,1
Italy	0	1,1
Netherlands	1	1,1
Norway*	2	1,1
Spain	1	1,1
Sweden*	0	1,1
Switzerland*	1	1,1
UK*	3	1,1

Notes: *non-member countries; Model A is chosen by AIC, Model B by BIC; r is the number of cointegrating vectors; e and m give the lag orders on endogensous and exogenous variables, respectively; endogenous variables are y_{it}, π_{it}, eq_{it}, ep_{it}, r_{it} and lr_{it}; except Finland where lr_{it} is not available; exogenous variables are y^*_{it}, π^*_{it}, eq^*_{it}, r^*_{it}, lr^*_{it} and $poil_{it}$.

First, we use BIC to identify the number of cointegrating vectors (r) and to choose the optimal lag length for the endogenous (e) and exogenous (m) variables (Table 5.11).

For all countries lag length of one seems to be optimal for both, exogenous and endogenous variables. However, there are some differences in r across the sample. According to BIC there is no cointegration among the variables in case of Finland, Italy and Sweden. However, BIC confirms presence of one cointegrating vector for Austria, Belgium, France, the Netherlands, Spain and Switzerland. Three unusual cases, as before, are Germany and Norway with $r = 2$ and the UK with three cointegrating vectors.

Then we estimate mean error of forecast (Table 5.12) and the root mean squared prediction errors of forecast 2008Q1-2016Q4 (Table 5.13).

Table 5.12 Effect of Change in Exogenity on Real GDP: Mean Error of Forecast 2008Q1-2016Q4 (System)

	6-Variable Model[1]		**4-Variable Model[2]**		**3-Variable Model[3]**	
	Level (%)	**Growth (%)**	**Level (%)**	**Growth (%)**	**Level (%)**	**Growth (%)**
Austria	-2.79	-0.16	-2.87	-0.16	-2.88	-0.15
Belgium	0.51	0.03	1.26	0.06	1.33	0.06
Finland	-7.48	-0.37	-6.73	-0.33		
France	-4.64	-0.19	-4.62	-0.20	-4.70	-0.20
Germany	6.42	0.37	4.62	0.26	3.26	0.19
Italy	-9.26	-0.40	-6.41	-0.28	-9.21	-0.40
Netherlands	-5.23	-0.20	-5.25	-0.21	-5.64	-0.22
Norway*	-3.38	-0.16	-2.83	-0.17	-3.01	-0.18
Spain	-14.11	-0.55	-13.63	-0.53	-13.32	-0.51
Sweden*	5.79	0.40	5.76	0.39	6.39	0.42
Switzerland*	4.95	0.18	4.36	0.14	4.47	0.15
UK*	-23.09	-1.31	-21.94	-1.03	-28.10	-1.85

Notes: *non-member countries. Level in Percent; Growth in Percentage Points at Annual Rates
[1]6-Variable Model: y_{it}, π_{it}, eq_{it}, ep_{it}, r_{it} and lr_{it} are endogenous variables; [2]4-Variable Model: y_{it}, π_{it}, eq_{it} and lr_{it} are endogenous variables; [3]3-Variable Model: y_{it}, π_{it} and eq_{it} are endogenous variables.

However, changing the year of structural break (from 1999Q1 to 2008Q1) to account for the effect of the financial crisis 2007-2008 does not seem to improve forecasts. It still appears that the UK, Italy and Spain did noticeably worse than was predicted, with these three countries having especially large negative mean errors of forecast. Among those who did better are Germany, Sweden and Switzerland, which is also consistent with forecasts for 1999Q-2016Q4 with an exception of Germany (that had a negative mean error of -2.61 before, which changed to 6.42 when the structural break was moved to 2008Q1). The mean error of forecast for the UK level of GDP changed significantly, from 0.74 in a six-variable model for forecast 1999Q1-2016Q4 (Table 5.9) to the corresponding value of -23.09 for

Table 5.13 Effect of Change in Exogenity on Real GDP: RMSPE[1] of Forecast 2008Q1-2016Q4 (System)

	6-Variable Model[2]		4-Variable Model[3]		3-Variable Model[4]	
	Level (%)	**Growth (%)**	**Level (%)**	**Growth (%)**	**Level (%)**	**Growth (%)**
Austria	3.30	0.93	3.35	0.93	3.32	0.93
Belgium	0.92	0.53	1.50	0.50	1.59	0.50
Finland	8.53	1.03	7.72	1.08		
France	4.93	0.33	4.91	0.33	5.00	0.34
Germany	7.81	0.60	5.64	0.54	4.19	0.49
Italy	10.29	0.62	7.12	0.52	10.23	0.63
Netherlands	6.22	0.66	6.30	0.65	6.52	0.60
Norway*	3.74	1.20	3.53	1.23	3.70	1.26
Spain	15.79	0.82	15.28	0.83	14.93	0.81
Sweden*	7.38	0.90	7.32	0.90	7.91	0.91
Switzerland*	5.18	0.81	4.52	0.77	4.64	0.81
UK*	26.93	1.37	24.44	1.18	33.83	2.01

Notes: *non-member countries. Level in percent; growth in percentage points at annual rates. [1]RMSPE - Root Mean Squared Prediction Errors. [2]6-Variable Model: y_{it}, π_{it}, eq_{it}, ep_{it}, r_{it} and lr_{it} are endogenous variables; [3]4-Variable Model: y_{it}, π_{it}, eq_{it} and lr_{it} are endogenous variables; 433-Variable Model: y_{it}, π_{it} and eq_{it} are endogenous variables.

forecast 2008Q1-2016Q4. The root mean squared errors are also quite different (Table 5.10 and Table 5.13). Hence, we graph the forecast for the UK to check these estimations.

As can be seen from the Table 5.11, the BIC suggests presence of three cointegrating vectors, which is the only r=3 case in the sample, with all other countries having up to two cointegrating vectors. We plot forecast for the UK level of income with r=3 (Figure 5.3) and compare it with such for r=1 (Figure 5.4).

Fig. 5.3 UK GDP (level; r=3)

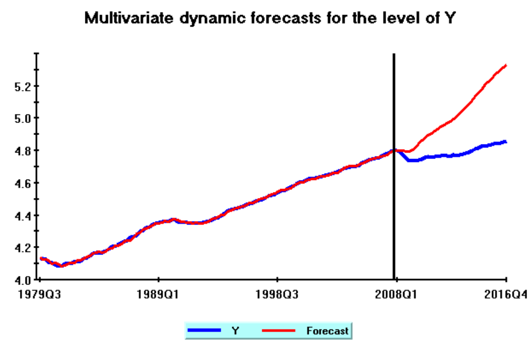
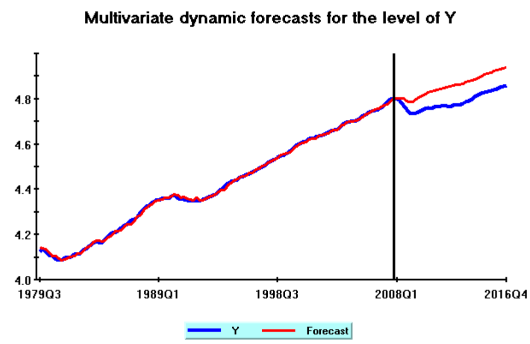


Fig. 5.4 UK GDP (level; r=1)



Note: r = number of cointegrating vectors.

It appears, that the reason for such unusual forecast for the UK might indeed be a result of misspecification rather than anything else. As we can see from the r=3 case, the forecast wonders off after the structural break in 2008Q1. In r=1 case, the actual series is still below the forecast, meaning that in both cases the UK did worse than was predicted, however, the growth level in forecast series is much more stable. The same can be found if we test for no cointegration at all (case of r=0). This suggests, that the unusual forecast for the UK is likely to be due to misspecification. Overall, we conclude that changing structural break from 1999Q1 to 2008Q1 does not significantly alter our estimations and forecasts.

5.8 Conclusion

In this chapter we compare our counterfactuals with the actual post-euro realisation and find that in most cases actuals diverge from forecasts quite substantially. However, this might not be necessarily due to joining or not joining euro, but due to some misspecification of the equations estimated. We attempt to condition on various exogenous variables, but there does not seem to be a clear difference between euro and non-euro countries' estimates.

We start with single equation estimates for GDP, inflation and interest rates and then perform multivariate dynamic forecasts for the level of income and for GDP growth. We then compare these system forecast estimates with the single equation forecast ones.

Starting with level of GDP equation, lagged foreign income appears to have a significant negative effect on domestic income for all countries with the exception of the Netherlands, while the overall effect of world income is positive for most countries, except four. The unexpected result is a negative trend coefficient for four cases, which might be due to some misspecification. Even though all equations passed normality test, about half of the countries failed the serial correlation and functional form tests and two failed the heteroskedasticity test. In general, the graphs for actuals and forecasts do not appear to be very close to each other, but exhibit similar general patterns.

As for the growth in GDP equations, the lagged foreign income growth coefficient is insignificant, but positive for all countries except three, and the overall foreign income growth effect is positive as well for all countries. All equations passed the functional form and heteroskedasticity tests, however, some failed the serial correlation and normality tests. Overall, comparing to the level form, the growth rate model seems to fit the data better. Although in both cases structural stability analysis identifies breaks in different years for different countries, hence, there does not seem to be a common break around the foundation of euro.

As for the change in inflation equation, the lagged change in domestic inflation is negative for the whole sample, except one country. However, most countries did not pass the serial correlation and heteroskedasticity tests. Overall, in the level of inflation equations actual inflation rates were more jumpy than the forecasted ones for most countries. Meanwhile, in the change of inflation equation, the actual growth rates of inflation are mostly higher than the forecasted ones.

When we estimate short and long interest rate equations we find that in case of majority of economies the actual and predicted short interest rates are quite close until the crisis 2007-2008. For majority of countries after the crisis the predicted rate went sharply negative while the actual one was constrained by the zero lower bound and, hence, could not follow. We find that the difference between actual and predicted rates is smaller when we use long-term rather than short-term interest rates to estimate Taylor Rule. Nonetheless, even in the long-term case, the estimates for euro members are not much different from the results for non-member countries.

When we compare system and single equations forecast estimates for the level of GDP and the income growth we find that in the first case for more than half of the sample the actuals are below forests, suggesting performance worse than forecast suggested. The difference is especially dramatic following the financial crisis, with the actuals going much more below forecasts. However,it appears that the growth single equation and system results for GDP are more similar than the level ones.

The main structural breaks seem to be in the interest and exchange rate equations, where there was a clear institutional change in their determination with the introduction of the euro. The formation of euro is not identified as the most likely date for a structural break in the GDP equations, and the GDP growth rate equation shows no structural break for many countries. In addition, changing year of a structural break to 2008Q1, to account for the financial crisis 2007-2008, does not significantly improve our estimates.

Overall, the results are sensitive to a range of specification choices and the confidence intervals around counterfactuals, that were based on pre-euro relationship, are large. There do not appear to be obvious differences in the patterns of structural breaks when comparing euro and non-member countries. The effects of the euro are apparent everywhere except in the main macroeconomic equations. This may be significant. The formation of the euro was a major break which required a change in the patterns of economic relationships to provide alternative methods of economic adjustment to changes in interest and exchange rates. The fact that the economic relationships did not seem to have changed may have been a source of tensions for the euro.

To summarise, we conclude that even though according to traditional economic theory the interest and exchange rates are seen as two key tools for the balance of payments adjustments, we do not find significant difference between those countries that have control over setting their monetary targets (stand-alone countries) versus those who do not (EMU-members).

Hence, we conclude that there must be other adjustment mechanisms at play. What they are is an area for future research.

Chapter 6

Conclusion

This thesis uses various panels of cross-country data to examine the types of open economy adjustments undertaken to maintain a sustainable or solvent position with respect to the balance of payments and public finance. In the beginning of this thesis we ask relatively general questions: how do exports and imports respond to balance of payments deficits? Does the balance of payments adjust in a stabilising way to maintain long-term solvency? Do exchange rate adjustments stabilise the balance of payments?

In the first chapter we provide the introduction for this thesis. Then in the second chapter we begin our empirical analysis by first focusing on the exports and imports demand equations. We find that there is a lot of evidence of the significant effect of income on trade, but the results on the Marshall-Lerner condition are mixed. Endogeneity and other issues make empirical analysis of the demand equations for exports and imports complicated. The results vary substantially depending on the sample size, estimation period and the dataset used.

We begin the second chapter by first analysing a general ARDL model (for the real exports) that includes exchange rates, consumer price indices, the national and global GDP. We find that this general model produces results that are very sensitive to the model specifications and does not lead to a common model for different countries in the sample. In contrast, a more restrictive error-correction model, with income being the only explanatory variable, produces results that are more in line with traditional economic theory. We find that the average income elasticity of trade for 15 countries is 1.26% in the long run and 4.08% in the short run. Nonetheless, this model produces mixed findings on the effect of income on trade in the long run. According to the Wald test long-run elasticity coefficients are significant for most countries in the sample, except two (Australia and India). However, the F-test results suggest a possible joint redundancy of the lagged variables in the model, leaving a doubt about whether there is a long-run relationship between income and trade.

In addition, after analysing the same model using two different datasets, WTO and IMF, we find that results vary depending on a dataset, however, there appears to be a positive relationship between the IMF and WTO estimates of short-run and long-run elasticities (considering the whole period, 1980-2013) and this relationship is stronger in the former case. Thus, we conclude that while both datasets confirm significant effect of income on trade both, in the short and in the long run, the actual estimates are sensitive to the dataset used.

We also check sensitivity of these estimates using structural stability analysis. We assume that if we account for possible structural breaks in the data it might help to produce more robust estimates. However, structural breaks appear to be in different years in different countries. Nonetheless, when we try to estimate this relationship using WTO data for our sample over 1980-2013 and break this period into pre-2000 and post-2000, both short-run and long-run coefficients are significant for the vast majority of countries. Hence, year 2000 appears to be if not the optimal, then a reasonably good choice for a common breakpoint for our data. Nevertheless, the structural stability analysis still shows that estimates of income-trade relationship are also sensitive to the choice of the estimation period.

Finally, quite unusual result is that the short-run elasticity coefficients are larger than the respective long-run ones. It is however, consistent with findings in the relevant literature, including the Constantinescu et al. (2015) paper. Turning to the literature for an explanation, one of the possible reasons for such counter-intuitive results is that income-trade relationship analysis is subject to an issue of endogeneity. It is unlikely that world GDP would be shifted by the exports of any particular country, but exports could be driving domestic output, which in turn might affect global income. The presence of lagged endogenous variables on the right-hand side of a regression model is likely to bias OLS estimates (Felbermayr, 2005).

Overall, after analysing income-trade relationship using 15 countries and two datasets, WTO and IMF, we can note that results are sensitive to the specifications of the model, estimation period and the datasets used. We find that the analysis of income-trade relationship is subject to a few issues and limitations. Besides the endogeneity issue, there is also likely to be an issue of omitted variables, as income and trade flows prices might be not the only variables that affect demand for exports and imports. However, what the other variables are is a complicated question. Furthermore, ideally the analysis of changing income-trade relationship requires long data, but for many countries sufficiently long data are not available, hence the data availability is another issue of this analysis.

Furthermore, it appears that besides structural factors, cyclical ones, such as temporary negative shocks like a recession, might also be behind the changing income-trade relationship. Nonetheless, it is quite challenging to separate the two in relatively short data series. Hence, we consider income-trade relationship for the UK and US for which longer quarterly data are available. After estimating VECM models we establish that prices and income can be considered weakly exogenous. Hence, we try to estimate ARDL and error-correction models that assume exogeneity of explanatory variables. We find that our ECM estimates for imports are comparable to the ones found in a paper by Hooper et al. (2000) that used similar estimation techniques to ours. There are also some discrepancies between our and their results, which may be due to different sample sizes and estimation periods.

We also find that the Marshall-Lerner condition holds only for the UK, but not for the US. Hence, it is not clear what mechanisms maintain balance of payments sustainability. To investigate this, rather than looking at the whole model we consider reduced form autoregressive equations for the balance of payments and government surplus. In the third chapter we focus on the balance of payments solvency and look directly at what is the feedback using autoregression without specifying the intermediate variables, income and the exchange rates.

To get more precise estimates of the adjustment process we use long span data for many countries. The data for a panel of 17 countries come from JST dataset[1]) and covers a period from 1870 to 2016.

First we focus on the relationship between exports and imports shares of GDP. Homogeneous fixed effect panel estimate is significant and comes to 0.97, suggesting that the change in exports and imports shares of GDP is about one-to-one. When we break the estimation period in three sub-periods, the new estimates confirm the fixed effect result with most coefficients, except few in the first sub-period, being significant and close to one.

Since imports and exports are likely to be co-determined, we then try to estimate a VECM in shares with one lag. Pooled sample estimates once again suggest that imports-exports relationship is close to unity, and both feedback coefficients even though small, respond to the disequilibrium, suggesting endogeneity. Meanwhile, heterogeneous estimates suggest that for the whole period cointegration coefficient is significant in all but three countries. It averages to -0.76 with adjustment being done through imports for most economies. Speed of adjustment coefficient of imports is significant for 11 countries and averages to 14% per year, while exports speed of adjustment coefficient is only significant for 8 economies and averages to 2%. Nonetheless, there is less evidence of cointegration over shorter sub-periods, and the coefficients do not seem to be close to unity. When we check whether cointegration coefficient is in fact equal to one, this restriction on VECM is rejected for about half of the sample. Nonetheless, estimating the model for the whole period with logs of imports and

[1]Oscar Jordà, Moritz Schularick, and Alan M. Taylor. 2017. “Macrofinancial History and the New Business Cycle Facts.” in NBER Macroeconomics Annual 2016, volume 31, edited by Martin Eichenbaum and Jonathan A. Parker. Chicago: University of Chicago Press.

exports in nominal values and local currency produces an average coefficient of -0.997, much closer to unity, and estimating that model over shorter sub-period does not substantially change this result. Hence, trade flows are clearly cointegrated when the relationship is estimated over sufficiently long period. However, it is still not clear whether exports and imports should be expressed as shares of GDP, or it is better to model this relationship using logarithms of the exports and imports nominal values.

In order to estimate balance of payments solvency we use two proxies, current account and balance of trade, both are expressed as shares of GDP. We analyse sustainability of the process by simply regressing change in the balance of trade (current account) share of GDP on its own lagged values, and if the response coefficient is negative, it confirms that the process is stabilising. Homogeneous fixed effect panel estimates suggest that both, the balance of trade and current account shares of GDP are examples of stabilising processes. The current account adjustment to equilibrium is about 21% per year with approximately 10% of the adjustment been done by imports, 7% by exports and roughly 4% by the other components of the current account. The balance of payments annual adjustment is about 10% with exports and imports adjustment being symmetric, about 5% per year.

When we consider heterogeneous estimates for the whole period, balance of payments is stabilising for 10 out of 17 economies. Even though current account series has more missing observations, the estimates suggest that the process is stabilising for 13 out of 17 countries. There is less evidence of stabilisation when we estimate the autoregressive models over shorter periods.

Fixed cross-country effects estimates also suggest that for the balance of payments (measured by the current account share of GDP) the annual adjustment to equilibrium is 26% when in deficit and 13% when in surplus, suggesting that there is about twice as much pressure to adjust on countries running deficits versus those running surpluses. Meanwhile, heterogeneous estimates provide some support for the presence of the asymmetric adjustment, but these evidence are mixed and are not be very significant.

Nonetheless, even though we find strong evidence of stabilisation for both, current account and balance of trade, it is still not clear how the adjustment is done. A traditional view is that in case of the balance of payments solvency, the adjustment to net foreign assets of a country should play an essential role (Bohn, 2007). However, these data are not available for the long span covered by the JST dataset. Nonetheless, we have sufficient data to analyse sustainability for public finances, and we consider this useful since the balance of payments and government debt crises often go hand in hand, as we saw for instance from the European sovereign debt crisis that peaked between 2010 and 2012. Moreover, theoretical framework of the balance of payments solvency and the government surplus sustainability is similar.

Hence, we then move to analysing sustainability of the public sector surplus and government debt, using the same dataset (JST) and similar theoretical framework. We emphasise that even though the algebra in terms of theoretical framework is similar for analysis of the balance of payments solvency and government surplus sustainability, the economics of the two and the way the balance of payments and government spending are financed are different.

Building on a theory proposed by Bohn (2007) we estimate a model for surplus and separate models for its components, revenue and expenditure. This allows us to shed some light on the adjustment process as well as the feedback mechanism of these processes converging to stationary state.

We first estimate models for change in surplus, revenue and expenditure using lagged values of debt-GDP, surplus, expenditure and income as explanatory variables. Fixed effects estimates suggest surplus and revenue do not seem to adjust to debt-GDP in the right way. Moreover, lagged expenditure coefficient is quite small allowing us to assume that effects of revenue and expenditure are not significantly different, meaning it is the difference between the two that is important. Since lagged income coefficient is very small too, we conclude that we can estimate models without expressing surplus as a share of GDP.

We then estimate models for change in surplus, revenue and expenditure with only two explanatory variables, lagged surplus and lagged debt-GDP. Moreover, as in the balance of payments chapter, to get some insights on the adjustment patterns we split our data in three approximately equal sub-periods, 1870-1914, 1915-1950 and 1951-2016.

Homogeneous fixed-effect panel estimates suggest that surplus is stabilising on both, its previous value and debt-GDP. It appears that surplus adjustment to its own lagged values is done through revenue and to the lagged debt-GDP − through expenditure. Heterogeneous estimates are consistent with these results in terms of how adjustment is done, although according to country-specific results, surplus is stabilising on its own lagged values, but not on debt-GDP ratio. Hence, we conclude that there might not be an equilibrium level of debt, which is consistent with Bohn's idea that different ratios of debt-GDP can be sustainable depending on the credibility of the borrowing government (Bohn, 2005, 2016).

Moreover, fixed-country effect estimates and heterogeneous results suggest that in most cases stabilising lagged surplus coefficient in the revenue equation offsets effect of the distabilising one in the expenditure equation. Meanwhile, more lagged debt coefficients have the correct sign in the expenditure equation than in the revenue model, once again offsetting the effect of the latter ones.

Furthermore, including key macroeconomic variables, namely GDP growth, long-term interest rate and inflation in the surplus model does not substantially alter these results. The pooled data estimates suggest that all coefficients except long interest rate are significant and that there is significant feedback coming from debt-GDP. Nonetheless, heterogeneous estimates still suggest that debt-GDP ratio coefficient is different from zero only in two out of 17 countries. Hence, it is not clear whether the best estimate of the feedback from debt-GDP is its coefficient or zero.

Moreover, as in case with the current account adjustment, we find evidence of the asymmetric adjustment when we estimate data with fixed cross-country effects over the whole period. We find that if the government is running deficit the annual adjustment to equilibrium is twice higher (20%) than when it is running surplus (10%), suggesting that there is more pressure to adjust on countries running deficits than those running surpluses. There is somewhat less evidence of the asymmetric adjustment if we consider countries individually, which, again, is consistent with the results for the current account analysis.

In addition, before we treated balance of payments and government surplus adjustments separately, however in the last section of chapter four we also consider whether they are linked through the identity which we derive starting from the GDP equation. The decision to check for cross-surplus adjustment comes from what known as a twin-deficit hypothesis, that suggests that shocks that deteriorate government budget, shift the current account balance in the same direction. Hence, we check if there any links between the balance of trade and public sector surplus. We find that even pooled data estimates for the whole period suggest very little cross-surplus adjustment. We also report heterogeneous estimates, but they do not provide any more support for the twin-deficit hypothesis.

Overall, we find strong evidence that large deficits do tend to adjust themselves, but there does not seem to be an equilibrium level of debt. Hence, even very high debt-GDP ratios may potentially be sustainable as long as a government has sufficient opportunities to refinance its debt. This conclusion is linked to the main issue with the empirical analysis of the solvency condition for the balance of payments and for the government debt, which is the fact that solvency is a question of economic nature and is based on lenders' expectations (Bohn, 2005, 2016). Therefore, debt-GDP can grow rapidly as long as the lenders expect to be paid. However, if lenders believe that the government is not a credible borrower they will not be willing to acquire this government's debt without a risk premium to compensate for possible default. Therefore, the debt acquisition can have different effect on interest rates for borrowers with different credibility. Thus, there does not need to be an equilibrium level of debt-GDP and various governments can remain solvent with different levels of debt-GDP depending on lenders beliefs in their credibility. Since the true necessary condition for solvency depends on the lenders' beliefs about a government's ability to repay its debts, it cannot be tested empirically (Bohn, 2005, 2016).

Another issues with these estimations are the gaps in data for some countries. For instance, in case of Finland there is an insufficient number of observations to estimate surplus, revenue or expenditure models. It is also not clear whether surplus and debt should have been used as shares or in nominal values. A number of transformations on these data is possible and there does not seem to be a clear preference for one over another when we consider different options.

While we find evidence of stabilisation for both, balance of payments and government surplus, it is still not clear what are the mechanism behind these adjustments. Most economic theories assume that open economy adjustments are mainly done though changes in interest rate and exchange rate. However, for some countries these adjustment mechanisms are not available. In case of the European Monetary Union, in moving to common interest rates and exchange rates, some of the adjustment mechanisms were removed, thus, the real economy had to adjust. In the fifth chapter we consider what happens in case when the nominal interest rate and exchange rate adjustments are removed on an example of the European Monetary Union members. We consider a range of equations for EMU members and non-member countries in order to identify whether there any significant differences in the estimates for the two groups.

Despite having sufficient data for this analysis, one of the issues was to consider what conditioning variables we should have included. Also, it was not clear whether we should model the relationships using levels, meaning treating the variables as trend stationary, or assume they are difference stationary and use growth rates. Thus, we do both. The third challenge is that working with counterfactuals, we cannot be certain that whatever differences between EMU and non-member countries we found are due to joining EMU or some possible misspecifications of the forecasting equation.

First we estimate GDP equation using levels and growth rates. We find that in the growth rate relationship the actuals and predicted growth rates for many countries in the sample exhibit similar general patterns and modelling the relationship among variables using the growth rates fits the data better. Nonetheless, there does not seem to be a major difference between estimates of the EMU members and those of stand-alone countries.

Then we estimated level of inflation and find the actual inflation rates in general appear to be more jumpy than the forecasted ones. We also estimate change in inflation equations in a hope that more equations pass diagnostic tests. But majority of the countries still fail serial correlation and heteroskedasticity tests. In addition, structural stability analysis suggests breaks in various years for different countries. Hence, change in inflation equation does not seem to fit data much better than levels relationship.

Then we estimated Taylor Rule equations using short and long interest rates and find that for most countries the actual levels are quite close to the predicted ones until the crisis. After the crisis the actuals remain constrained by the zero lower bound, while predicted ones fall sharply. Estimations for the Taylor Rule using long interest rate also do not provide any clear distinction between EMU and non-member countries with zero lower bound once again creating a major difference between actuals and predicted values. Nonetheless, the difference between the two is smaller when we use long rather than short interest rates.

Moving to the VARX systems we find that there is less evidence of the structural instability when we control for the short interest rate and exchange rate. We then estimate the cointegrating VAR model in order to perform multivariate dynamic forecast. These forecasts are done for level of GDP and for GDP growth. In the levels case for more than half of the sample the actual income growth is below the predicted one, especially after the crisis. Hence, these economies performed worse than forecasted. In the growth case the estimates for the single equation and system are more similar, but still there does not seem to be any clear difference between euro and non-euro countries.

Finally, we check sensitivity of our estimates to exogeneity, but find that treating short interest rate and exchange rate as exogenous does not produce major differences. We also consider whether a more suitable breakpoint for the data is not in 1999Q1 but in 2008Q1, during the crisis. Nonetheless, using 2008Q1 as a breakpoint does not help to establish any more consistency among our estimates. Overall, after estimating VARX with various specifications we find that results are very sensitive to the specification choices, and the confidence intervals around counterfactuals are too large, not allowing for any specific conclusions about possible differences between EMU and non-member countries.

Our conclusion that there does not seem to be any particular distinction between economic performance of EMU members and non-members suggests that becoming part of this monetary union at least did not make these countries worse off. It might be that adjustment is more complicated than simple models suggest, and it is not clear what the intermediate mechanisms are. We assumed that interest rates and exchange rates might be regarded as exogenous for EMU countries, however, it might not be the case for all individual euro countries. There might be other influences that we should have controlled for, but what they are is an open question.

To summarise, models that were used in this thesis are mainly descriptive, not structural. Hence, they provided us with only general ideas about the long-run interactions among the variables we considered in this thesis. The main limitations of the analysis performed in this thesis are sensitivity of the results to model specifications, exogeneity issue and limited data availability.

Overall, we find that there seems to be a significant effect of income on trade in the short and in the long run, but it is challenging to produce robust measures of these effects. In fact, it appears that the data required for the analysis of the sustainability of the balance of payments and public finances are very noisy. In most cases numerous shocks in the data make it rather challenging to confirm sustainability of the balance of trade, current account and government surplus when the models are estimated for individual countries or over short periods. Our analysis shows that there is, however, strong evidence of sustainability of the balance of trade, current account and public surplus when the models are estimated over sufficiently long period for a large number of countries. Using pooled data with fixed cross-country effects we find that in case of the balance of trade the adjustment of exports and imports is about symmetric (5% per year each), while in case of current account the largest portion of the adjustment is done by imports (10% per year) and slightly smaller by exports (7% per year). In case of public finances, deficits adjust mainly through revenue. To the best of our knowledge the use of long span heterogeneous panel data to the analysis of sustainability is an original contribution of this study.

In addition, we find that there does not seem to be an equilibrium level of debt, hence, different debt-GDP ratios may be sustainable depending on the borrower's credibility. Finally, we do not find major differences between countries that have full control over their interest rate and exchange rate, that are traditionally considered to be two main adjustment mechanisms, and economies that do not have such autonomy. Thus, while we find evidence of stabilisation of the balance of payments and public sector surplus (deficit), the mechanisms behind these adjustments remain to be an open question. Monetary policy decisions, changes in regulation, structure of taxes and uncertainty are likely to be important factors in the adjustment processes of the balance of payments and government deficits towards equilibrium level. These are possible areas for future research.

Appendix A

Chapter II: Trade Elasticities

A.1 History of Trade: Overview

Trade data used in the second chapter are very noisy and covers a long period. Hence, in addition to the material covered in the chapter we provide some historical context. There were significant changes in the trade growth rate in the 19th-beginning of the 21st centuries. The reasons for that include, but are not limited to, the technological progress, political and economical changes. Technological innovations have been dramatically changing the ways nations trade, increasing not only volume of goods traded, but giving strong boost for invisible trade. Starting from the creation of telephone in 1870s to introduction of the World Wide Web in 1990, the world has been getting smaller with transport costs falling dramatically. Innovations, such as telegraph, developed in the 19th century, enormously simplified the movement of capital across the world.

From economic side, period 1870-1914 was characterised by the Gold Standard system, later replaced by Bretton Woods in 1944-1973, that was in turn abandoned for the Flexible Exchange Rates system in 1973-1998. The Gold Standard reduced volatility of exchange rates further increasing trade among countries that were on gold (Meissner et al., 2003). End of the 19th century can also be seen as the beginning of the re-industrialisation of the developing countries with more developed countries exporting manufactured industrial goods and importing primary products from Asia, Africa, Latin America, and Australasia. This trend in global trade, that is known as the Great Specialisation, was among the major reasons of the world becoming significantly more asymmetrical in economic sense and of increasingly high income inequality in the world (Maddison, 2003).

From political side, the first half of the 20th century was marked by the two world wars and the Great Depression. Trade growth was increasing rapidly up till the World War I, which caused closure of Europe's major exchanges and effectively deteriorated trade around the world. Then, in 1929, the Great Depression started and lasted till late 1930s, leading to more severe stagnation. The real global trade declined by 14% during the interwar period, mostly due to falling output and increasing tariff rates (Madsen, 2001). The Great Depression was quickly followed by the World War II which was another great exogenous shock to global trade and world economy. However, the exact exports and imports figures for some countries in our sample for the second chapter (namely, Germany, Japan and the Netherlands) for 1944-1945 are missing. The Great War led to the reduced imports of manufactured goods and grain from the Eastern Europe while boosted exports of agriculture goods from Australia, Canada and the US. Moreover, the US, once a debtor, now was the main supplier of capital in the world (Litman, 1926).

Nevertheless, the pre-World War II level of global trade was recovered by 1949. The Second World War ended with establishment of the Bretton Woods system that largely promoted free trade in order to boost world economy recovery after the war. In 1947 23 countries signed the General Agreement on Tariffs and Trade (GATT) that removed many trade barriers and quotas. The GATT was eventually replaced by the Marrakesh Agreement that was signed by 124 countries on 15 April in 1994 and led to creation of the World Trade Organisation (WTO), which further boosted global trade growth by dismantling protectionist tariffs and managing negotiations among nations.

This boost to the trade growth were significantly magnified by the number of technological innovations, such as the containerisation, specialised ships equipped to transport chemicals and other goods, big vessels and open-registry shipping that reduced regulatory costs. Even though freight rates remained high throughout the century, after the World War II the world experienced unprecedented 185% GDP per capita growth from 1950 to 2000 (Maddison, 2003) and trade growth reached an average rate of 5.93% comparing to the 3.49% which was the average prior World War I (Findlay and O'rourke, 2009).

Another important effect of globalisation was emphasised by Helleiner (1973) who discussed the increasing trend in the vertical specialisation. By the end of the 20th century the production of a good could involve different stages to be based in different countries, making it somewhat difficult to measure the true extent of the effect of vertical specialisation on the trade growth. Among those who tried, Yeats (1999) estimated that in 1995 the intermediate goods trade accounted for 30% of global trade in machinery and transportation equipment. In addition, another major characteristic of trade in the 20th century, long-waited by some, was the unravelling of the Great Specialisation with declining dependence of the Southern (developing) markets on the Northern ones (Europe and Northern America). According to UNISTAD developing countries share of global output in manufacturing increased from 12% in 1970 to 20% in 1995 (UNCTAD., 1997, p. 82).

The foundation of the European Union in 1993 changed the landscape of trade in Europe by creating a single internal market among its members allowing free flow of labour, capital, visible and invisible trade. The union led to creation of a single currency, euro, that not only significantly simplified trade within the EU, but made the euro area an attractive region for non-member countries to do business, with euro becoming one of the most traded currencies in the world. We focus on the effect of euro in a separate chapter.

Moving away from the Europe, in 1994 the United States with its neighbours, Canada and Mexico, entered their separate trade agreement, the North American Free Trade Agreement (NAFTA). The NAFTA liberalised the trade in agriculture, apparel, textiles and automotive goods among its three members. Regional trade increased sharply during 1994-2014, the US merchandise trade with its NAFTA partners increased from around \$300 billion in 1993 to over \$1.1 trillion in 2014 (Villareal and Fergusson, 2017). The NAFTA has become one of approximately three hundred free trade agreements (FTA's), some regional (RTA's), others bilateral (BTA's), that were negotiated globally to this day.

However, global trade growth experienced slowdown following the financial crisis 2007-2008 that led to increase in protectionism, lower consumer demand worldwide, unstable exchange rates, fewer expanding markets and lower credit amounts available to households and firms. Trade growth slowdown is also explained by the imbalances and asymmetric distribution of income that was brought about by globalisation. Following Brexit-related uncertainty, escalating trade tensions in the world and various country-specific cyclical and structural factors, in October 2019 WTO economists revised their prediction of 3.0% trade growth increase in 2020 down to 2.7% (WTO., 2019).

A.2 General-to-Specific Modelling: Estimates

General model (2.13):

$$ARDL(1,1): lnX_{it} = \alpha_0 + \beta_0 lnN_{it} + \beta_1 lnN_{it-1} + \alpha_1 lnX_{it-1} + \delta t + u_{it}.$$

Table A.1 Restricted Model (Australia, Austria, Canada, Chile, Denmark, Finland and France)

	Australia	Austria	Canada	Chile	Denmark	Finland	France
$Constant$	-45.87***	-45.79***	-9.65***	-233.03***	-44.32**	-17.18***	-24.30***
	(10.36)	(2.54)	(1.53)	(38.86)	(17.59)	(6.25)	(6.59)
lnX_{it-1}	0.26**		0.399***			0.62***	0.399***
	(0.17)		(0.09)			(0.08)	(0.06)
lnY_{it}							1.29***
							(0.27)
ΔlnY_{it}		1.12**	1.30**	1.07**		2.52***	
		(0.53)	(0.54)	(0.398)		(0.37)	
lnY_{it-1}	-1.17*	1.77***			0.393*		
	(0.60)	(0.32)			(0.24)		
lnY_t^*				8.24***	1.72**	0.81***	
				(1.27)	(0.64)	(0.23)	
ΔlnY_t^*	2.48***	2.12***	3.18***				1.93***
	(0.88)	(0.65)	(0.77)				(0.58)
lnY_{t-1}^*	3.04***	0.68***	0.65***				
	(0.87)	(0.24)	(0.11)				
lnE_{it}	0.24**	0.397***		-0.68***			0.390***
	(0.09)	(0.07)		(0.09)			(0.05)
ΔlnE_{it}					0.390***	0.35***	
					(0.08)	(0.11)	
lnE_{it-1}	0.399***	0.15*		-0.20**	0.66***	0.24**	
	(0.13)	(0.07)		(0.10)	(0.06)	(0.10)	
$lnCPI_{it}$	0.392*					0.61***	0.63***
	(0.23)					(0.19)	(0.13)
$\Delta lnCPI_{it}$				1.37***			
				(0.16)			
$lnCPI_{it-1}$				0.77***			
				(0.12)			
$lnCPI_{USt}$	0.94***						
	(0.34)						
$\Delta lnCPI_{USt}$		1.73***	1.71***		0.95*		
		(0.393)	(0.36)		(0.52)		
$lnCPI_{USt-1}$		-0.397***			0.32***	-0.86**	-0.92***
		(0.07)			(0.08)	(0.35)	(0.23)
δt				-0.18***	-0.04**		
				(0.04)	(0.02)		

Source: WTO.

Note: Standard errors in parentheses; *** indicates a significance level of 1 percent, ** of 5 percent, and * of 10 percent.

Table A.2 Restricted Model (Germany, Greece, India, Indonesia, Israel, Italy, Japan and the UK)

	Germany	Greece	India	Indonesia	Israel	Italy	Japan	UK
$Constant$	-80.27***	-28.88***	-31.34***	-6.53	-14.89***	-20.86**	6.64***	-9.86**
	(15.05)	(9.02)	(8.97)	(8.93)	(4.61)	(7.81)	(2.28)	(4.76)
lnX_{it-1}		0.57***	0.82***	0.22*	0.65***	0.59***	0.394***	0.30***
		(0.10)	(0.06)	(0.12)	(0.10)	(0.11)	(0.11)	(0.10)
lnY_{it}								
ΔlnY_{it}				2.30***				
				(0.63)				
lnY_{it-1}	-1.21***			1.04**				
	(0.23)			(0.391)				
lnY_t^*	2.74***		3.38***					
	(0.51)		(0.94)					
ΔlnY_t^*		4.38***			4.40***	2.87***	3.96***	2.64***
		(1.17)			(0.60)	(0.74)	(0.67)	(0.62)
lnY_{t-1}^*	1.64***	1.22***	-2.20**		0.68***	1.03***	0.16**	0.81***
	(0.55)	(0.33)	(1.07)		(0.21)	(0.32)	(0.07)	(0.19)
lnE_{it}	0.392***	0.26*						0.30***
	(0.06)	(0.13)						(0.08)
ΔlnE_{it}						0.37***	0.38***	
						(0.10)	(0.07)	
lnE_{it-1}	0.20***			-0.399***		0.28***	0.29***	
	(0.06)			(0.11)		(0.09)	(0.07)	
$lnCPI_{it}$								0.61***
								(0.18)
$\Delta lnCPI_{it}$	2.09***	1.83***		1.81***	0.06**	1.44***	2.35***	
	(0.67)	(0.393)		(0.31)	(0.03)	(0.39)	(0.38)	
$lnCPI_{it-1}$	1.44***	0.23*		1.21***		0.81***	0.398***	
	(0.28)	(0.12)		(0.29)		(0.23)	(0.11)	
$lnCPI_{USt}$						-1.46**		
						(0.54)		
ΔCPI_{USt}				3.03***	1.53***			1.76**
				(1.00)	(0.399)			(0.65)
$lnCPI_{USt-1}$	-1.08***	-0.60*	-0.50***	-1.02***				-0.72**
	(0.14)	(0.32)	(0.16)	(0.35)				(0.33)
δt	-0.07***			-0.05*				
	(0.01)			(0.03)				

Source: WTO.
Note: Standard errors in parentheses; *** indicates a significance level of 1 percent, ** of 5 percent, and * of 10 percent.

A.3 Specific-to-General Modelling (Speed of Adjustment)

Table A.3 Speed of Adjustment

	Speed of Adjustment		**Half-life**
	Coefficient	SE	t(half-time) = ln(2)/($-\alpha_1(lnX_{t-1})$)
Australia	-0.03	0.06	21.30
Austria	-0.22*	1.03	3.19
Canada	-0.32***	0.11	2.14
Chile	-0.29***	0.10	2.42
Denmark	-0.21**	0.10	3.32
Finland	-0.15	0.10	4.55
France	-0.17*	0.08	4.17
Germany	-0.18*	0.10	3.91
Greece	-0.20**	0.10	3.45
India	-0.03	0.04	26.37
Indonesia	-0.19**	0.07	3.71
Israel	-0.17*	0.09	4.14
Italy	-0.20**	0.10	3.45
Japan	-0.16**	0.08	4.22
UK	-0.21**	0.11	3.23
Average	-0.21		6.24

Source: WTO.
Note: *** indicates a significance level of 1 percent, ** of 5 percent, and * of 10 percent. SE stands for standard error.

A.4 IMF and WTO datasets: Comparison

A.4.1 IMF: Dataset Used in Constantinescu, et al. (2015)

Table A.4 Raw Data (Data Source: Dr. Ileana Cristina Constantinescu)

Year	Gross Domestic Product, Constant Prices (2000=100)	Volume of Total Imports of Goods and Services (2000=100)
1970	35	17
1971	37	18
1972	39	20
1973	41	22
1974	42	24
1975	43	23
1976	45	25
1977	47	27
1978	49	28
1979	51	30
1980	52	31
1981	53	31
1982	54	31
1983	55	31
1984	58	34
1985	60	35
1986	62	37
1987	65	40
1988	68	43
1989	70	47
1990	73	50
1991	74	52
1992	76	55
1993	78	56
1994	80	62
1995	83	68
1996	86	73
1997	90	80
1998	92	84
1999	96	89
2000	100	100
2001	102	100
2002	105	104
2003	109	110
2004	115	123
2005	120	132
2006	126	144
2007	133	156
2008	137	160
2009	136	143
2010	143	161
2011	149	171
2012	154	176
2013	158	181

A.4.2 Constantinescu, et al. (2015): Replicated Estimates

Model used in Constantinescu et al. (2015):

$$ECM : \Delta lnM_{it} = \alpha + \beta \Delta lnY_{it} + \delta lnY_{it-1} + \gamma lnM_{it-1} + \varepsilon_{it}.$$

Table A.5 Constantinescu, Mattoo and Ruta (2015)

	Whole Period	**Separate Periods**		
	1970-2013	1970-1985	1986-2000	2001-2013
α	-0.393**	-0.35	-3.17***	-0.52**
	(0.17)	(0.53)	(0.64)	(0.19)
Short-Run Elasticity (β)	2.82***	2.13***	2.77***	3.43***
	(-0.36)	(0.6)	(0.35)	(0.21)
Speed of Adjustment ($-\gamma$)	0.12**	0.18	0.58***	0.31**
	(0.05)	(0.31)	(0.13)	(0.13)
δ	0.20**	0.23	1.26***	0.390**
	(0.09)	(0.39)	(0.26)	(0.17)
Long-Run Elasticity3 ($-\delta/\gamma$)	1.70***	1.31***	2.18***	1.31***
R-Squared	0.74	0.957	0.957	0.957
Sample Size	43	43	43	43

Source: Dataset used in Constantinescu, et al. (2015). Received Directly from Dr. Ileana Cristina Constantinescu.
Note: *** indicates a significance level of 1 percent, ** of 5 percent, and * of 10 percent.

A.4.3 WTO Data (Imports): 1970-2013 and 3 Sub-Periods

As a preliminary analysis, following Constantinescu et al. (2015), (2.16) and (2.17) are estimated using WTO data. The results are summarised in the Table A.6.

Table A.6 WTO Data (Imports): 1970-2013

	Short-Run Income Elasticity			Long-Run Income Elasticity			Short-Run	Long-Run
	1970-1985	1986-2000	2001-2013	1970-1985	1986-2000	2001-2013	Whole Period (1970-2013)	
World	2.38	3.20**	5.58***	48.81	1.34***	2.57***	4.30***	1.61***
	(1.81)	(1.25)	(1.09)	(1049.27)	(0.23)	(0.28)	(0.88)	(0.41)
15 Countries	3.41	3.70***	4.82***	194.25	1.30***	3.68***	4.32***	1.94**
	(2.18)	(1.17)	(1.29)	(17882.52)	(0.23)	(0.43)	(0.98)	(0.77)
Australia	3.12	2.02	9.23**	1.09	0.85*	3.19***	2.93**	1.35***
	(2.34)	(1.31)	(3.15)	(1.27)	(0.40)	(0.21)	(1.26)	(0.22)
Austria	3.45**	2.78	4.91***	-10.83	0.78*	3.25***	3.69***	1.91***
	(1.35)	(1.69)	(1.23)	(72.67)	(0.42)	(0.72)	(0.81)	(0.26)
Canada	2.78**	1.66***	4.73***	0.31	1.45***	2.57***	2.93***	1.29***
	(0.91)	(0.48)	(0.87)	(0.88)	(0.24)	(0.33)	(0.47)	(0.11)
Chile	2.57***	3.51***	6.88**	-85.53	1.33***	2.75***	2.96***	1.15***
	(0.70)	(0.63)	(2.14)	(1498.97)	(0.18)	(0.49)	(0.48)	(0.37)
Denmark	1.93	2.57	3.97***	-5.76	-0.08	5.41***	2.09**	1.33***
	(1.64)	(1.43)	(0.87)	(14.77)	(1.14)	(0.60)	(0.85)	(0.45)
Finland	4.05**	2.52***	4.44***	-88.56	-0.35	3.95***	3.03***	1.38***
	(1.77)	(0.69)	(0.47)	(5211.58)	(1.28)	(0.24)	(0.59)	(0.36)
France	4.55**	2.29*	5.07***	-1712.79	0.87*	4.10***	4.15***	1.77***
	(1.94)	(1.16)	(1.51)	(1777626.00)	(0.42)	(0.59)	(0.95)	(0.42)
Germany	3.64**	3.49**	3.33**	30.05	1.24***	1.56	3.40***	1.87***
	(1.35)	(1.23)	(1.09)	(199.27)	(0.36)	(6.06)	(0.70)	(0.34)
Greece	0.58	-0.28	0.88	0.15	2.06***	2.28*	0.93**	1.84***
	(0.93)	(1.18)	(1.06)	(2.48)	(0.60)	(1.03)	(0.44)	(0.20)
India	0.87	2.70**	2.68	0.22	1.05***	2.36***	1.17	1.48***
	(1.31)	(0.87)	(2.45)	(2.22)	(0.17)	(0.30)	(0.86)	(0.39)
Indonesia	0.77	2.32***	20.40***	142.47	1.39***	1.60***	2.56***	1.07***
	(2.46)	(0.39)	(4.78)	(10559.02)	(0.13)	(0.37)	(0.68)	(0.17)
Israel	-1.24	2.48***	3.43**	-0.61	0.80***	1.04***	1.39**	0.81***
	(1.33)	(0.76)	(1.26)	(1.18)	(0.19)	(0.21)	(0.63)	(0.08)
Italy	4.68***	5.07**	5.40***	-7.98	0.90	59.33	4.64***	5.17
	(0.87)	(1.75)	(1.23)	(43.71)	(0.70)	(681.37)	(0.75)	(8.69)
Japan	-1.74	5.24***	4.75***	0.53	1.99***	7.64***	1.53	1.19***
	(2.22)	(1.07)	(1.21)	(1.12)	(0.20)	(1.55)	(0.96)	(0.42)
UK	0.25	1.85*	3.90***	0.80	1.00**	2.83***	1.48**	1.33***
	(1.47)	(0.92)	(0.94)	(1.72)	(0.45)	(0.34)	(0.70)	(0.15)
Average	3.67	3.01	6.16	-	1.23	3.31	2.78	1.41

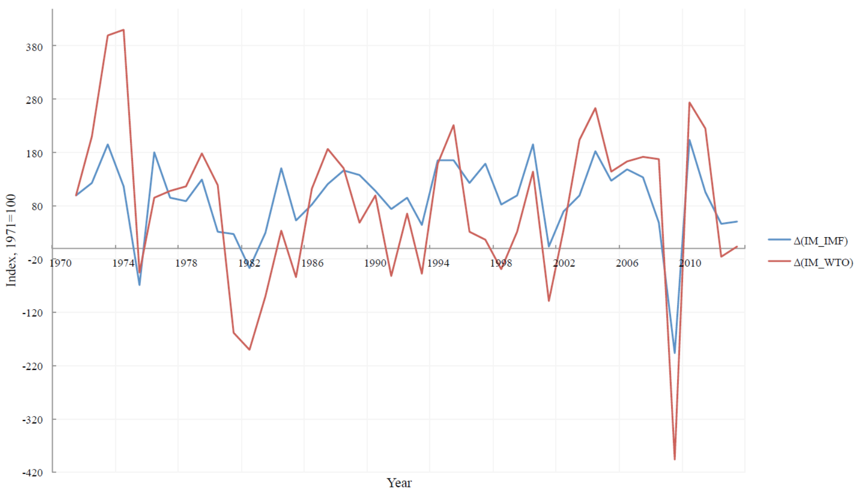
Source: WTO.
Note: Standard errors in parentheses; *** indicates a significance level of 1%, ** of 5%, and * of 10%.
15 Countries - aggregated.

When considering the whole world, the short-run elasticity is 2.82 using IMF data and 4.30 using WTO one, which are quite different. Nonetheless, the estimates for the long-run elasticity are close, 1.70 and 1.61, using IMF and WTO data, respectively.

Overall, both datasets confirm a relationship between income and trade in the short run and in the long run when the model estimated over the whole period. However, the actual values vary depending on the data used in the estimation. Breaking the data into three sub-periods, suggested in Constantinescu et al. (2015), worsens the significance of the estimated parameters and produces mixed heterogeneous results.

There are many reasons for the differences in the results obtained using IMF versus WTO datasets, including definitional differences and various data transformations. For instance, when we look at WTO and IMF data for world imports, the correlation coefficient, ρ, for the two import growth rates is relatively low (0.779), meaning two datasets are measuring different things when it comes to trade data (while GDP growth of IMF and WTO are highly correlated, $\rho = 0.977$). For comparison, IMF and WTO imports growth rates are re-based and plotted on the same graph (Figure A.1). The substantial difference between them confirms that they represent two different measures of world imports.

Fig. A.1 IMF and WTO Comparison (World Imports)



Source: WTO and IMF.

A.4.4 WTO and IMF Data: Structural Breakpoint in 2000

Table A.7 WTO and IMF Data: Structural Breakpoint in 2000

	WTO Data			IMF Data		
	Q-A test		Chow Test (2000)	Q-A Test		Chow Test (2000)
	LR F-stat.	Wald F-stat.	F-statistic	LR F-stat	Wald F-stat	F-statistic
World	-	-	-	1994***	1994***	Breakpoint***
15 Countries	-	-	-	1994*	1994*	-
Australia	2003***	2003***	Breakpoint***	-	-	-
Austria	-	-	-	1989***	1986***	-
Canada	1999***	1999***	Breakpoint***	1991***	1986***	Breakpoint***
Chile	1991**	1991**	Breakpoint**	2008***	2008***	Breakpoint**
Denmark	-	-	Breakpoint**	1996***	2009***	-
Finland	1997*	1997*	Breakpoint**	1992***	1986***	-
France	-	-	-	1988**	2009***	-
Germany	-	-	-	1995***	1986***	-
Greece	-	-	-	2001***	2009***	Breakpoint***
India	2008***	2008***	-	-	2009***	-
Indonesia	1986***	1986***	Breakpoint***	-	1987***	-
Israel	2008***	2008***	-	2006**	2008***	Breakpoint***
Italy	2002*	2002*	Breakpoint**	1997*	2009***	-
Japan	-	-	-	1994***	1987***	-
UK	-	-	-	-	1986***	Breakpoint*

Source: WTO and IMF World Economic Outlook April 2014. Note: Q-A test stands for Quandt-Andrews unknown breakpoint test. LR F-stat - Maximum Likelihood Ratio F-statistic. Wald F-stat stands for Maximum Wald F-statistics. *** indicates a significance level of 1 percent, ** of 5 percent, and * of 10 percent. 15 countries - aggregate.

A.4.5 IMF Data for 15 countries (1980-2013)

The IMF data for import flows and GDP are available for all 15 countries in the sample and for the whole world for 1980-2013. In addition, it was considered important to estimate the income-trade relationship for the sample as a whole. However, given that the IMF data for import flows and for GDP are in the percentage change form, it was not possible to add up the values for 15 countries to calculate total imports and income for the sample. First, the weights were calculated using imports data (in current dollars) for these countries from WTO, and only then the aggregate imports and GDP were calculated using these weights.

Appendix B

Chapter III: Balance of Payments Sustainability

B.1 Static Model (Exports and Imports Shares; Equation 3.20)

Table B.1 OLS: Imports and Exports (Equation (3.20); 1870-2016 and Sub-Periods)

	1870-1914		**1915-1950**		**1951-2016**		**1870-2016**		**Average for**
Country	**Coefficient**	**t-stat.**[1]	**Coefficient**	**t-stat.**	**Coefficient**	**t-stat.**	**Coefficient**	**t-stat.**	**3 Sub-Periods (Whole Sample)**
Australia	0.051	0.622	**0.540**	**4.274**	**0.907**	**10.017**	**0.433**	**7.074**	**0.499 (0.723*)**
Belgium	**1.290**	**22.297**	**1.069**	**12.924**	**0.821**	**33.263**	**0.834**	**46.795**	**1.060**
Canada	0.161	0.832	**0.414**	**5.232**	**0.802**	**31.206**	**0.562**	**12.411**	**0.459 (0.608*)**
Denmark	**0.754**	**6.307**	**0.936**	**6.726**	0.106	1.669	**0.523**	**7.954**	**0.599**
Finland	**0.795**	**5.287**	**0.504**	**3.270**	**0.547**	**9.457**	**0.456**	**7.542**	**0.615**
France	**0.486**	**2.241**	**0.764**	**2.559**	**1.124**	**29.669**	**0.763**	**8.891**	**0.791**
Germany	0.337	1.606	**0.960**	**8.755**	**0.775**	**49.800**	**0.695**	**22.425**	**0.691 (0.867*)**
Italy	**1.167**	**6.430**	**1.134**	**4.239**	**0.879**	**25.290**	**0.747**	**19.207**	**1.060**
Japan	**1.103**	**13.369**	**0.870**	**10.635**	**0.873**	**8.728**	**0.923**	**19.862**	**0.948**
Netherlands	**1.123**	**37.281**	**1.106**	**9.750**	**0.381**	**6.641**	**1.191**	**44.595**	**0.870**
Norway	**1.260**	**6.843**	**0.789**	**2.409**	**-0.720**	**-11.847**	**-0.237**	**-3.691**	**0.443**
Portugal	0.227	1.271	0.630	1.702	**0.820**	**15.542**	**1.086**	**34.947**	**0.559 (0.725*)**
Spain	**0.564**	**8.954**	**0.706**	**5.416**	**0.834**	**24.832**	**0.987**	**31.281**	**0.701**
Sweden	**0.693**	**5.069**	**0.635**	**5.880**	**0.613**	**16.864**	**0.651**	**20.293**	**0.647**
Switzerland	**0.833**	**10.329**	**0.881**	**13.555**	**0.506**	**11.181**	**0.919**	**15.935**	**0.740**
UK	**0.427**	**5.583**	**0.927**	**11.502**	**1.007**	**7.427**	**0.975**	**21.079**	**0.787**
US	0.207	1.087	**0.308**	**4.362**	**1.475**	**26.925**	**1.391**	**21.831**	**0.663 (0.891*)**
Average	**0.675 (0.875*)**		**0.775**		**0.691**		**0.759**		**0.714 (0.769*)**

Notes: [1] t-statistic; significant coefficients are in bold (5% level; * - average including significant coefficients only.

B.2 VECM (Exports and Imports Shares; Equations 3.21 and 3.22)

Table B.2 VECM: ms, xs (1,1) (Equations (3.21) and (3.22); 1870-2016 and Sub-Periods)

	1870-1914			**1915-1950**			**1951-2016**			**1870-2016**		
Country	α	β	γ	α	β	γ	α	β	γ	α	β	γ
Australia	0.033	**-0.331**	**-0.561**	**-1.521**	**-0.286**	0.129	**-1.156**	**-0.330**	**0.394**	**-0.482**	**-0.296**	-0.079
	[0.296]	**[-2.147]**	**[-3.328]**	**[-3.478]**	**[-2.641]**	[1.024]	**[-11.41]**	**[-2.443]**	**[3.488]**	**[-3.201]**	**[-4.522]**	[-1.102]
Belgium	**-1.453**	-0.222	0.237	**-1.083**	0.178	**0.520**	**-0.847**	-0.145	0.047	**-0.836**	-0.087	0.088
	[-13.473]	[-1.745]	[1.760]	**[-8.628]**	[0.595]	**[2.132]**	**[-10.940]**	[-1.185]	[0.395]	**[-13.500]**	[-1.050]	[1.156]
Canada	-0.684	**-0.263**	0.037	**-0.810**	-0.212	**0.415**	**-0.834**	-0.205	0.079	**-0.895**	**-0.113**	**0.129**
	[-1.454]	**[-2.775]**	[0.448]	**[-5.231]**	[-1.709]	**[2.437]**	**[-11.977]**	[-1.377]	[0.464]	**[-6.777]**	**[-2.623]**	**[2.553]**
Denmark	**-0.731**	**-0.483**	-0.119	**-1.403**	-0.031	**0.222**	0.073	**-0.440**	**-0.347**	-0.272	**-0.248**	-0.033
	[-3.794]	**[-3.346]**	[-1.288]	**[-6.702]**	[-0.159]	**[2.486]**	[0.705]	**[-4.022]**	**[-4.334]**	[-1.357]	**[-3.741]**	[-0.858]
Finland	**-0.872**	**-0.505**	-0.042	0.305	**-0.274**	**-0.259**	-0.309	**-0.260**	**-0.289**	-0.088	**-0.271**	**-0.216**
	[-3.347]	**[-3.372]**	[-0.278]	[0.775]	**[-2.027]**	**[-2.800]**	[-1.932]	**[-2.722]**	**[-2.806]**	[-0.547]	**[-4.191]**	**[-4.059]**
France	**6.991**	-0.005	**-0.057**	**-1.725**	-0.068	**0.085**	**-1.159**	-0.238	-0.026	**-1.032**	**-0.120**	**0.057**
	[3.002]	[-0.199]	**[-3.041]**	**[-2.592]**	[-0.900]	**[2.207]**	**[-10.240]**	[-1.666]	[-0.238]	**[-4.739]**	**[-2.916]**	**[2.288]**
Germany	-0.016	**-0.267**	0.073	**-1.235**	0.534	**0.620**	**-0.784**	**-0.390**	0.160	**-0.768**	**-0.097**	0.034
	[-0.038]	**[-2.545]**	[1.125]	**[-14.727]**	[1.819]	**[3.993]**	**[-33.271]**	**[-2.176]**	[0.810]	**[-6.738]**	**[-2.041]**	[0.788]
Italy	**-2.758**	-0.116	0.133	**-4.930**	0.079	**0.116**	**-0.903**	**-0.363**	-0.005	**-0.753**	**-0.161**	-0.013
	[-5.985]	[-1.283]	[1.882]	**[-5.122]**	[1.740]	**[4.056]**	**[-13.087]**	**[-2.762]**	[-0.044]	**[-5.646]**	**[-2.788]**	[-0.292]
Japan	**-1.245**	**-0.498**	0.149	**-1.299**	0.073	**0.295**	**-0.641**	-0.168	0.024	**-1.123**	-0.067	**0.162**
	[-10.282]	**[-3.067]**	[0.956]	**[-6.010]**	[0.538]	**[2.655]**	**[-2.095]**	[-1.664]	[0.334]	**[-8.104]**	[-0.959]	**[2.751]**
Netherlands	**-1.145**	-0.106	**0.363**	**-1.440**	0.092	**0.389**	-0.215	**-0.390**	**-0.268**	**-1.245**	0.029	**0.094**
	[-21.720]	[-0.635]	**[2.458]**	**[-10.064]**	[0.342]	**[2.180]**	[-1.554]	**[-3.254]**	**[-2.217]**	**[-10.544]**	[0.638]	**[2.321]**
Norway	**-1.732**	**-0.268**	0.061	**-5.501**	-0.048	**0.065**	**0.934**	**-0.219**	**-0.320**	**0.611**	**-0.130**	-0.073
	[-4.537]	**[-2.875]**	[0.770]	**[-2.869]**	[-0.795]	**[2.118]**	**[9.331]**	**[-2.981]**	**[-3.006]**	**[2.591]**	**[-2.908]**	[-1.959]
Portugal	1.661	-0.071	-0.067	-0.394	**-0.219**	-0.033	**-0.791**	-0.207	0.067	**-1.055**	**-0.115**	0.044
	[1.532]	[-1.124]	[-1.907]	[-0.416]	**[-2.206]**	[-0.825]	**[-5.973]**	[-1.935]	[0.816]	**[-10.220]**	**[-2.322]**	[1.446]
Spain	**-0.628**	**-0.474**	-0.037	**-0.914**	**-0.239**	-0.055	**-0.823**	**-0.231**	0.001	**-0.895**	-0.056	0.062
	[-5.097]	**[-2.781]**	[-0.171]	**[-2.619]**	**[-2.014]**	[-0.430]	**[-10.317]**	**[-2.597]**	[0.019]	**[-6.303]**	[-1.331]	[1.591]
Sweden	-0.281	**-0.329**	-0.090	**-0.607**	**-1.036**	-0.022	**-0.462**	**-0.387**	**-0.451**	**-0.633**	**-0.410**	-0.067
	[-0.867]	**[-2.737]**	[-0.749]	**[-7.412]**	**[-4.195]**	[-0.098]	**[-5.788]**	**[-2.792]**	**[-3.590]**	**[-9.834]**	**[-4.654]**	[-0.895]
Switzerland	**-0.933**	**-0.763**	-0.051	**-0.880**	**-0.572**	-0.052	**-0.449**	**-0.455**	-0.155	-0.359	**-0.121**	**-0.085**
	[10.135]	**[-2.466]**	[-0.324]	**[-8.626]**	**[-2.112]**	[-0.188]	**[-5.019]**	**[-3.524]**	[-1.830]	[-1.203]	**[-2.839]**	**[-2.405]**
UK	**-0.787**	-0.081	**0.271**	**-1.129**	-0.239	0.195	3.167	**-0.060**	**-0.038**	**-1.374**	-0.056	**0.114**
	[-4.448]	[-0.681]	**[2.146]**	**[-9.576]**	[-1.453]	[0.890]	[1.824]	**[-2.364]**	**[-2.303]**	**[-9.375]**	[-1.153]	**[2.236]**
US	-1.555	-0.063	**0.141**	**-0.569**	-0.132	0.260	**-1.646**	0.067	**0.153**	**-1.730**	0.027	**0.131**
	[-1.902]	[-1.036]	**[2.114]**	**[-2.641]**	[-1.188]	[1.300]	**[-12.207]**	[0.887]	**[2.738]**	**[-8.763]**	[0.949]	**[4.111]**
Average:	**-0.361**	**-0.285**	**0.026**	**-1.479**	**-0.141**	**0.170**	**-0.403**	**-0.260**	**-0.057**	**-0.760**	**-0.135**	**0.020**
Average (Sign. Only*):	**-0.481**	**-0.418**	**0.032**	**-1.670**	**-0.438**	**0.247**	**-0.735**	**-0.320**	**-0.146**	**-0.872**	**-0.189**	**0.048**

Notes: equation: VECM (1,1) (or VAR(2)) for ms_t and xs_t, with intercept (no trend) in cointegrating equation and VAR ; XS and MS are exports and imports shares, respectively; α is a xs_{t-1} coefficient in the cointegration vector; β is the speed of adjustment for imports; γ is the speed of adjustment for exports; significant coefficients are in bold (5% level); coefficient with t-statistic (below the coefficient)

Average (Sign. Only*) - average including significant coefficients only.

B.2.1 Restricted VECM (Exports and Imports Shares; Equation 3.23)

Table B.3 VECM (1,1) for *ms*, *xs* with a Restriction $\alpha = -1$ (Equation (3.23); 1870-2016 and Sub-Periods)

	1870-1914			1915-1950			1951-2016			1870-2016		
Country	$H_0 : \alpha = -1$	β	γ	$H_0 : \alpha = -1$	β	γ	$H_0 : \alpha = -1$	β	γ	$H_0 : \alpha = -1$	β	γ
Australia	X[1]***	**-0.156**	0.029		**-0.382**	0.164		**-0.466**	**0.324**		**-0.200**	0.064
		[-2.258]	[0.332]		**[-2.517]**	[0.93797]		**[-3.289]**	**[2.53652]**		**[-3.809]**	[1.141]
Belgium		0.080	**0.23**		0.080	0.437		-0.047	0.07	X**	-0.045	0.057
	X***	[0.727]	**[2.052]**		[0.277]	[1.82203]		[-0.522]	[0.82287]		[-0.713]	[0.985]
Canada		**-0.222**	0.061		-0.134	**0.371**		-0.041	0.133		**-0.093**	**0.128**
		[-2.572]	[0.832]		[-1.306]	**[2.73722]**		[-0.402]	[1.17065]		**[-2.346]**	**[2.789]**
Denmark		**-0.441**	-0.168	X*	-0.229	0.164	X***	-0.054	-0.006	X**	-0.102	0.052
		[-2.773]	[-1.739]		[-0.982]	[1.42288]		[-0.934]	[-0.14123]		[-1.679]	[1.509]
Finland		**-0.483**	0.01	X*	-0.300	-0.020	X**	-0.104	-0.003	X***	**-0.176**	-0.006
		[-3.202]	[0.064]		[-1.879]	[-0.16425]		[-1.322]	[-0.03253]		**[-3.026]**	[-0.128]
France	X**	**-0.235**	-0.030		-0.105	0.077		-0.244	-0.073		**-0.122**	**0.056**
		[-2.169]	[-0.303]		[-1.250]	[1.72662]		[-1.828]	[-0.72361]		**[-2.953]**	**[2.246]**
Germany	X*	-0.134	0.104	X**	0.094	**0.585**	X***	-0.053	0.099		-0.063	0.026
		[-1.338]	[1.811]		[0.259]	**[2.95781]**		[-0.631]	[1.10412]		[-1.718]	[0.771]
Italy	X***	-0.248	-0.041	X***	-0.101	0.051		**-0.291**	0.024		**-0.115**	0.004
		[-1.904]	[-0.377]		[-0.877]	[0.59783]		**[-2.371]**	[0.23535]		**[-2.232]**	[0.093]
Japan	X*	**-0.376**	0.110		-0.048	0.286		-0.102	0.066		-0.114	**0.137**
		[-2.509]	[0.789]		[-0.273]	[1.89156]		[-0.993]	[0.93870]		[-1.551]	**[2.203]**
Netherlands	X**	-0.122	0.188	X***	-0.137	0.182	X***	-0.05	0.031		0.064	**0.079**
		[-0.875]	[1.455]		[-0.452]	[0.80945]		[-0.674]	[0.43298]		[1.685]	**[2.263]**
Norway	X*	**-0.241**	0.003	X*	-0.319	0.011	X***	-0.007	0.044	X***	-0.027	0.026
		[-2.703]	[0.036]		[-1.836]	[0.10582]		[-0.363]	[1.51885]		[-1.072]	[1.237]
Portugal	X*	-0.097	0.036		**-0.219**	-0.019		-0.132	0.100		**-0.108**	0.044
		[-0.848]	[0.548]		**[-2.199]**	[-0.46235]		[-1.237]	[1.24970]		**[-2.256]**	[1.481]
Spain		-0.192	0.137		-0.226	-0.034	X**	-0.154	0.019		-0.062	0.057
		[-1.616]	[0.973]		[-1.958]	[-0.27370]		[-1.865]	[0.31530]		[-1.409]	[1.416]
Sweden		-0.207	0.145	X***	**-0.486**	0.204	X***	-0.091	-0.027	X***	**-0.153**	0.039
		[-1.451]	[1.098]		**[-2.094]**	[1.13765]		[-1.027]	[-0.32490]		**[-2.320]**	[0.735]
Switzerland		**-0.634**	-0.004		-0.416	0.090	X***	-0.087	-0.018	X*	**-0.121**	-0.046
		[-2.148]	[-0.028]		[-1.546]	[0.33319]		[-1.041]	[-0.35503]		**[-2.093]**	[-0.953]
UK		-0.022	**0.246**		-0.334	0.078	X**	-0.110	-0.01	X**	**-0.059**	-0.063
		[-0.233]	**[2.470]**		[-1.919]	[0.32769]		[-1.019]	[-0.13867]		**[-2.278]**	[0.673]
US		-0.101	0.155		-0.030	0.209	X***	-0.011	0.026	X***	0.021	**0.079**
		[-1.383]	[1.886]		[-0.430]	[1.75991]		[-0.214]	[0.62513]		[0.785]	**[2.523]**
Average:		**-0.225**	**0.071**		**-0.194**	**0.167**		**-0.120**	**0.047**		**-0.087**	**0.043**
Average (Sign. Only*):		**-0.349**	**0.238**		**-0.362**	**0.478**		**-0.379**	**0.324**		**-0.127**	**0.096**

Notes: equation VECM (1,1) (or VAR(2)) for MS and XS, with intercept (no trend) in cointegrating equation and VAR. The applied restriction: $\alpha = -1$. [1] - reject $H_0 : \alpha = -1$ at 1% (X***), 5% (X**) or 10% (X*); for alpha report: Chi-square statistics and probability. XS and MS are exports and imports shares, respectively; α is a XS_{t-1} coefficient in the cointegration vector; β is the speed of adjustment for imports; γ is the speed of adjustment for exports; significant coefficients are in bold (5% level); coefficient with t-statistic (below the coefficient). Average (Sign. Only*) - average including significant coefficients only.

B.2.2 VECM (Logarithms of Exports and Imports; Equations (3.24) and (3.25)

Table B.4 VECM (1,1) (Equation (3.24); 1870-2016 and Sub-Periods)

	1870-1914			1915-1950			1951-2016			1870-2016		
Country	α	β	γ	α	β	γ	α	β	γ	α	β	γ
Australia	**-0.662**	**-0.264**	-0.062	**-1.376**	**-0.446**	0.010	**-1.006**	**-0.515**	0.203	**-1.029**	**-0.274**	0.043
	[-3.500]	**[-2.477]**	[-0.584]	**[-6.750]**	**[-3.011]**	[0.069]	**[-95.652]**	**[-3.5370]**	[1.461]	**[-71.945]**	**[-4.108]**	[0.694]
Belgium	**-1.075**	-0.154	**0.408**	**-0.950**	-0.490	**0.717**	**-0.995**	0.133	**0.323**	**-0.968**	-0.218	**0.620**
	[-21.576]	[-0.926]	**[2.349]**	**[-18.708]**	[-0.942]	**[2.069]**	**[-64.340]**	[0.878]	**[2.023]**	**[-267.296]**	[-1.135]	**[4.356]**
Canada	**-1.171**	**-0.332**	-0.070	**-1.019**	-0.312	0.205	**-0.975**	**-0.426**	-0.192	**-0.991**	**-0.188**	0.028
	[-9.577]	**[-3.024]**	[-0.767]	**[-10.892]**	[-1.647]	[1.049]	**[-82.188]**	**[-2.535]**	[-1.119]	**[-71.795]**	**[-3.017]**	[0.454]
Denmark	**-0.839**	**-0.357**	0.012	**-0.901**	-0.179	0.260	**-0.958**	0.091	**0.210**	**-0.951**	-0.139	**0.192**
	[-14.381]	**[-2.832]**	[0.083]	**[-4.455]**	[-0.547]	[1.243]	**[-53.656]**	[0.836]	**[2.673]**	**[-111.521]**	[-1.114]	**[2.343]**
Finland	**-0.940**	**-0.516**	-0.053	**-0.917**	0.114	0.409	**-1.279**	0.048	**0.074**	**-0.969**	0.053	**0.346**
	[-19.366]	**[-2.966]**	[-0.233]	**[-8.270]**	[0.480]	[1.617]	**[-10.253]**	[1.784]	**[2.553]**	**[-106.543]**	[0.514]	**[3.143]**
France	**-1.361**	-0.110	0.090	**-0.778**	-0.329	**0.342**	**-1.013**	-0.194	0.155	**-0.980**	**-0.163**	**0.252**
	[-4.582]	[-1.221]	[0.938]	**[-6.714]**	[-1.731]	**[1.999]**	**[-91.488]**	[-1.026]	[0.965]	**[-90.679]**	**[-2.110]**	**[3.684]**
Germany	**-0.851**	-0.097	0.116	**-1.142**	0.551	**1.202**	**-0.987**	**-0.347**	0.266	**-0.990**	**-0.202**	**0.213**
	[-5.296]	[-0.964]	[1.686]	**[-21.635]**	[1.115]	**[4.288]**	**[-129.992]**	**[-2.240]**	[1.928]	**[-871.645]**	**[-2.196]**	**[2.809]**
Italy	**-1.248**	**-0.435**	-0.082	**-0.944**	0.077	**0.662**	**-0.960**	**-0.496**	-0.226	**-0.969**	0.090	**0.555**
	[-10.622]	**[-2.671]**	[-0.551]	**[-10.509]**	[0.435]	**[3.609]**	**[-84.992]**	**[-3.520]**	[-1.889]	**[-147.489]**	[1.114]	**[6.752]**
Japan	**-1.030**	**-0.351**	0.107	**-1.192**	0.180	**0.497**	**-0.985**	-0.013	**0.182**	**-0.989**	-0.124	**0.171**
	[-18.720]	**[-2.853]**	[1.088]	**[-12.394]**	[1.033]	**[2.776]**	**[-29.108]**	[-0.113]	**[2.240]**	**[-126.372]**	[-1.529]	**[2.305]**
Netherlands	**-0.930**	-0.221	0.219	**-1.351**	0.335	**0.587**	**-0.918**	-0.453	0.269	**-0.945**	-0.056	0.163
	[-26.794]	[-1.201]	[1.190]	**[-9.611]**	[1.044]	**[2.000]**	**[-213.576]**	[-1.580]	[1.012]	**[-84.091]**	[-0.376]	[1.158]
Norway	**-0.980**	**-0.319**	-0.051	**-1.396**	0.121	**0.444**	**-0.802**	-0.030	**0.409**	**-0.887**	-0.025	**0.148**
	[-10.471]	**[-2.291]**	[-0.328]	**[-5.544]**	[0.533]	**[2.166]**	**[-51.763]**	[-0.319]	**[2.951]**	**[-39.129]**	[-0.422]	**[2.314]**
Portugal	**-2.173**	**-0.345**	0.208	**-0.525**	**-0.190**	**-0.136**	**-0.983**	0.133	**0.440**	**-0.952**	-0.023	**0.159**
	[-11.630]	**[-2.922]**	[1.819]	**[-3.709]**	**[-2.865]**	**[-1.999]**	**[-103.128]**	[1.070]	**[3.642]**	**[-74.208]**	[-0.452]	**[3.193]**
Spain	**-0.882**	**-0.387**	0.057	**-1.648**	**-0.444**	-0.026	**-0.736**	**-0.096**	**-0.078**	**-1.001**	-0.065	**0.121**
	[-8.524]	**[-2.292]**	[0.325]	**[-5.689]**	**[-3.828]**	[-0.233]	**[-12.336]**	**[-3.636]**	**[-3.031]**	**[-54.272]**	[-1.184]	**[2.529]**
Sweden	**-0.829**	**-0.385**	-0.108	**-1.081**	**-0.829**	-0.019	**-0.924**	**-0.372**	**-0.326**	**-0.962**	**-0.680**	-0.067
	[-9.381]	**[-2.754]**	[-0.749]	**[-7.620]**	**[-3.189]**	[-0.082]	**[-50.844]**	**[-2.283]**	**[-2.072]**	**[-173.891]**	**[-5.772]**	[-0.616]
Switzerland	**-1.045**	-0.149	0.180	**-0.647**	**-0.895**	**-0.638**	**-0.950**	-0.074	0.123	**-0.930**	**-0.454**	-0.141
	[-14.478]	[-0.466]	[0.838]	**[-6.114]**	**[-3.677]**	**[-2.635]**	**[-53.184]**	[-0.503]	[1.372]	**[-85.878]**	**[-3.815]**	[-1.320]
UK	**-1.009**	0.012	**0.270**	**-0.545**	-0.307	0.343	**-1.038**	-0.221	-0.021	**-0.993**	-0.073	**0.152**
	[-9.521]	[0.124]	**[2.202]**	**[-4.845]**	[-1.664]	[1.208]	**[-50.897]**	[-1.479]	[-0.176]	**[-54.640]**	[-1.496]	**[2.433]**
US	**-0.874**	**-0.286**	0.026	**-4.213**	-0.012	0.032	**-1.083**	-0.060	0.160	**-1.101**	**-0.124**	0.059
	[-6.401]	**[-2.670]**	[0.313]	**[-3.044]**	[-0.609]	[1.644]	**[-53.413]**	[-0.554]	[1.502]	**[-46.765]**	**[-2.723]**	[1.307]
Average:	**-1.053**	**-0.276**	**0.075**	**-1.213**	**-0.180**	**0.288**	**-0.976**	**-0.170**	**0.116**	**-0.977**	**-0.157**	**0.177**
St. Error:	**0.081**	**0.034**	**0.035**	**0.202**	**0.094**	**0.100**	**0.027**	**0.055**	**0.053**	**0.011**	**0.045**	**0.047**
t-stat. (for α = -1):	**-0.653**			**-1.056**			**0.873**			**2.104**		
Average (sing. only*):	**-1.053**	**-0.361**	**0.339**	**-1.213**	**-0.561**	**0.409**	**-0.976**	**-0.375**	**0.154**	**-0.977**	**-0.298**	**0.266**

Notes: equation VECM (1,1) (or VAR(2)) for LM and LX, with intercept (no trend) in cointegrating equation and VAR; x and m are logarithms of exports and imports, respectively; α is a x_{t-1} coefficient in the cointegration vector; β is the speed of adjustment for imports; γ is the speed of adjustment for exports; significant coefficients are in bold (5% level); coefficient with t-statistic (below the coefficient). Average (Sign. Only *) - average including significant coefficients only.

B.2.3 OLS (Logarithms of Exports and Imports; Equation (3.26)

Table B.5 OLS: Logarithms of Exports and Imports (Equation (3.26); 1870-2016 and Sub-Periods)

	1870-1914		**1915-1950**		**1951-2016**		**1870-2016**	
Country	**Coefficient**	**t-stat.**[1]	**Coefficient**	**t-stat.**	**Coefficient**	**t-stat.**	**Coefficient**	**t-stat.**
Australia	-0.221	-2.262	**-0.486**	**-3.264**	**-0.735**	**-6.150**	**-0.348**	**-5.463**
Belgium	**-0.643**	**-4.378**	**-0.783**	**-4.688**	**-0.193**	**-2.908**	**-0.517**	**-7.438**
Canada	-0.223	-2.278	-0.374	-2.789	**-0.252**	**-2.982**	**-0.169**	**-3.718**
Denmark	-0.362	-2.475	**-0.607**	**-3.888**	-0.063	-1.534	**-0.255**	**-4.532**
Finland	**-0.614**	**-4.316**	-0.283	-2.364	-0.197	-2.659	**-0.220**	**-4.237**
France	-0.311	-2.802	**-0.446**	**-3.096**	**-0.331**	**-3.563**	**-0.364**	**-5.669**
Germany	-0.212	-2.165	-0.353	-2.517	**-0.638**	**-7.997**	**-0.160**	**-3.946**
Italy	**-0.353**	**-3.025**	**-0.468**	**-3.213**	**-0.272**	**-3.216**	**-0.330**	**-5.314**
Japan	**-0.683**	**-5.372**	-0.332	-2.856	-0.139	-2.236	**-0.341**	**-6.030**
Netherlands	-0.331	-2.816	-0.168	-1.422	-0.107	-2.533	**-0.104**	**-3.080**
Norway	**-0.321**	**-2.957**	**-0.626**	**-3.972**	-0.047	-1.443	-0.085	-2.463
Portugal	-0.207	-2.056	-0.150	-1.625	-0.182	-2.279	-0.108	-2.779
Spain	**-0.642**	**-4.618**	-0.356	-2.719	-0.104	-1.769	**-0.197**	**-3.960**
Sweden	**-0.353**	**-3.066**	**-0.712**	**-4.365**	-0.093	-1.835	**-0.373**	**-5.765**
Switzerland	-0.305	-2.761	-0.306	-2.458	-0.060	-1.145	-0.116	-2.623
UK	-0.166	-2.055	-0.262	-2.193	-0.115	-1.770	**-0.174**	**-3.693**
US	-0.292	-2.815	-0.199	-1.886	-0.070	-1.562	-0.101	-2.735
Average	**-0.367**		**-0.407**		**-0.212**		**-0.233**	

Notes: [1] t-statistic. The significant coefficients are in bold. Significant coefficients are in bold (5% significance level; significance is determined using the Augmented Dickey-Fuller critical value of -2.9).

Appendix C

Chapter IV: Fiscal Sustainability

C.1 Derivation of the Intertemporal Budget Constraint and Transversality Condition.

Starting from the government budget constraint, (4.5):

$$b_t = (1+r_t)b_{t-1} - s_t,$$

the intertemporal budget constraint (IBC) and transversality condition (TC) can be derived in 4 steps.

Step 1: Take the equation (4.5) n-periods forward:

$$b_{t+n} = \left(\prod_{k=0}^{n}(1+r_{t+k})\right) b_{t-1} - \sum_{j=0}^{n}\left(\prod_{k=j+1}^{n}(1+r_{t+k})\right) s_{t+j}. \tag{C.1}$$

Step 2: In (C.1) assume r to be fixed and take the conditional expectations ($E_t[\cdot]$) of both sides:

$$E_t[b_{t+n}] = (1+r)^n b_t^* - \sum_{j=0}^{n}(1+r)^{n-j}E_t[s_{t+j}], \tag{C.2}$$

where $b_t^* = (1+r_t)b_{t-1}$ is debt at the start of period t.

Step 3: Discount the equation (C.2) by $(1+r)^n$ and rearrange:

$$b_t^* = \sum_{j=0}^{n}\frac{1}{(1+r)^j}E_t[s_{t+j}] + \frac{1}{(1+r)^n}E_t[b_{t+n}]. \tag{C.3}$$

Step 4: Assume the discounted sum converges and the limit $n \to \infty$:

$$b_t^* = \sum_{j=0}^{n} \frac{1}{(1+r)^j} E_t[s_{t+j}] + lim_{n\to\infty} \frac{1}{(1+r)^n} E_t[b_{t+n}]. \tag{C.4}$$

From (C.4), the intertemporal budget constrain, (4.6), is:

$$b_t^* = \sum_{j=0}^{n} \frac{1}{(1+r)^j} E_t[s_{t+j}], \qquad \text{[IBC]} \tag{C.5}$$

and the transversality condition, (4.7), is:

$$lim_{n\to\infty} \frac{1}{(1+r)^n} E_t[b_{t+n}] = 0. \qquad \text{[TC]} \tag{C.6}$$

C.2 Strictly positive relationship between surplus and debt.

Building on Trehan-Walsh's (1991) result that surplus and debt are cointegrated, Bohn (2005) argued that this implies a strictly positive relationship between the two (Bohn, 2005, pp. 8-9).

1. Cointegration between surplus and debt means that some linear combination of surplus and debt is stationary. Hence, in the equation below:

   $$s_t - \alpha b_{t-1} = u_t, \tag{C.7}$$

   u_t is stationary (and α is a constant).

2. Then next period's debt can be written as:

   $$b_{t+1} = (1+r)b_t - s_{t+1} = (1+r-\alpha)b_t + u_{t+1}. \tag{C.8}$$

   The Trehan-Walsh's (1991) cointegration result is based on the assumption that the debt is difference-stationary, that is $(b_t - b_{t-1})$ is stationary. This implies the asymptotic rate of debt growth (meaning, $(\lambda - 1)$), and the IBC, (4.6), holds as long as $\lambda < 1 + r$.

   In (C.8) $\lambda = 1 + r - \alpha$. Since $\lambda < 1 + r$ and $\alpha = 1 + r - \lambda > 0$, then $\alpha > 0$.

C.3 Proposition 3 from Bohn (2007, pp. 1844-45).

This mathematical proof (that the error-correction-type specification of the form presented in (C.7) can be integrated of an arbitrarily high order and still imply fiscal sustainability) is based on the Proposition 3 from Bohn (2007, pp. 1844-45).

Suppose

$$DEF_t + \alpha B_{t-1} = z_t \sim I(m) \tag{C.9}$$

for some $\alpha \in (0, 1+r]$. If we also assume a constant interest rate, $r_t = r$, then debt satisfies transversality condition, (4.7).

Starting from the government budget constraint, (4.3):

$$B_t = E_t^0 - R_t + (1+i_t)B_{t-1} = DEF_t^0 + (1+i_t)B_{t-1},$$

then taking (4.3) one period forward (and assume $i_t = r_t = r$), we get

$$B_{t+1} = (1+r-\alpha)B_t + z_{t+1} = \lambda B_t + z_{t+1}, \tag{C.10}$$

where $\lambda = 1 + r - \alpha \in [0, 1+r)$.

For $\lambda < 1$ this implies $B_t \sim I(m)$, and for $\lambda = 1$ this implies $B_t \sim I(m+1)$. For $\lambda > 1$, consider

$$\rho^n E_t[B_{t+n}] = (\rho\lambda)^n B_t + (\rho\lambda)^n E_t[\sum_{i=1}^{n} \lambda^{-i} z_{t+i}].$$

Because $z_t \sim I(m)$, the expression $\sum_{i=1}^{n} \lambda^{-i} z_{t+i}$ can be expanded into a linear combination of t-dated differences $\Delta^k z_t$ and stationary mth differences $\Delta^m z_{t+i}$[1]. Because $\lambda^{-i} < 1$, the weights in the linear combination are bounded from above by polynomials. Because $\rho\lambda < 1$, discounting by $(\rho\lambda)^n$ implies $\rho^n E_t[B_{t+n}] \xrightarrow{\rho} 0$.

Bohn also emphasised the validity of the argument made by Trehan and Walsh (1991) that the most important mathematical condition for transversality condition to hold is that all roots in the debt process are strictly less than $(1+r)$, but this condition may hold even if we allow for unit roots.

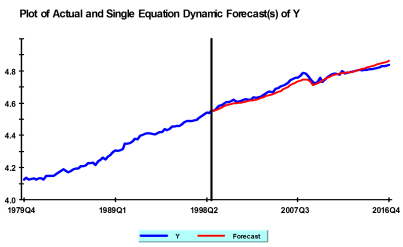
[1] Analogous to the expansion in the proof of Proposition 1 (Bohn, 2007, pp. 1840-42).

Appendix D

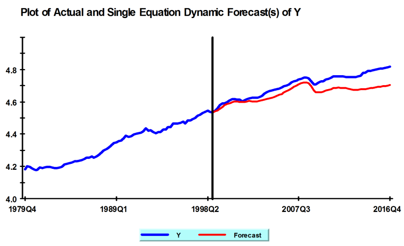
Chapter V: Effect of Exchange rate: European Monetary Union

D.1 Forecasts (Income; Levels Relationship)

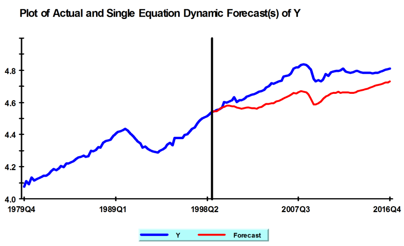
Austria



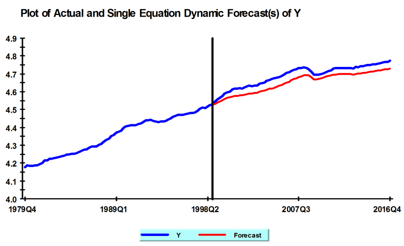
Belgium



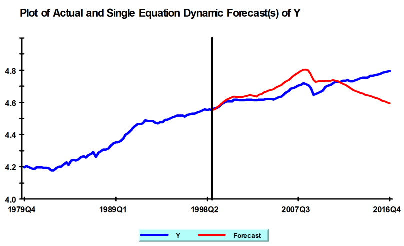
Finland



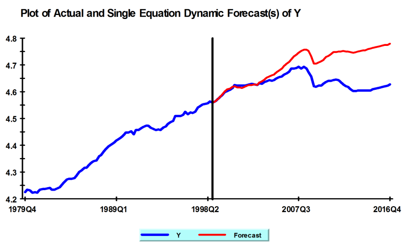
France



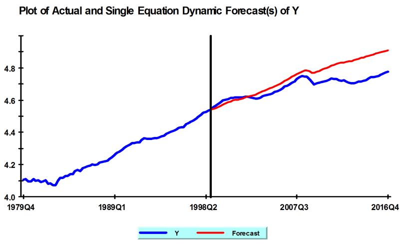
Germany



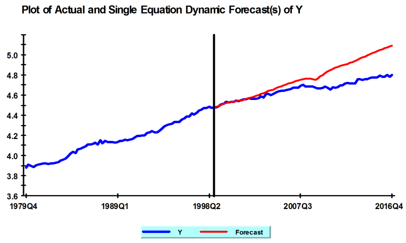
Italy



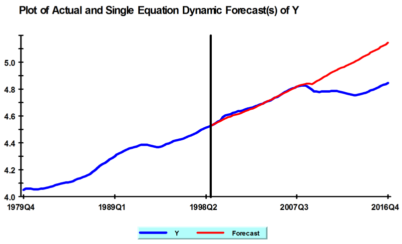
Netherlands



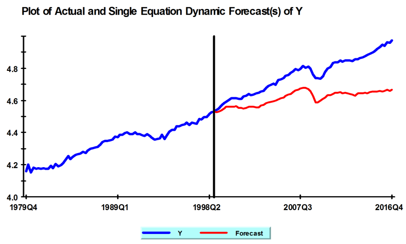
Norway



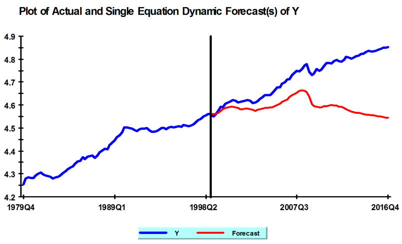
Spain



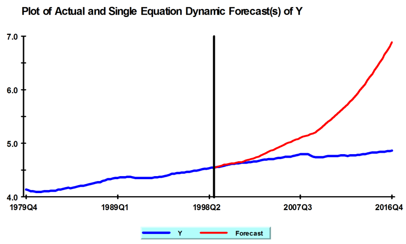
Sweden



Switzerland

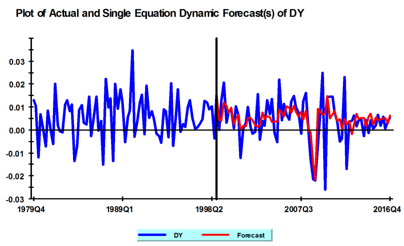


UK

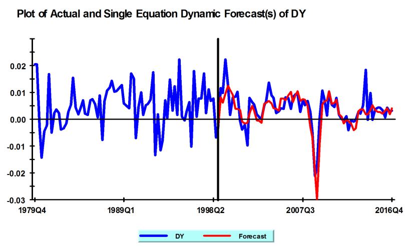


D.2 Forecasts (Income; Growth Rate Relationship)

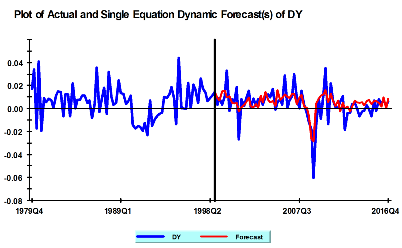
Austria



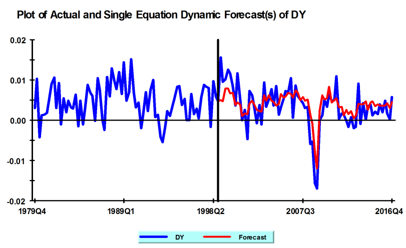
Belgium



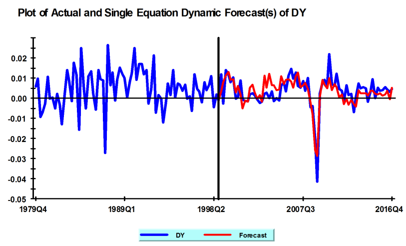
Finland



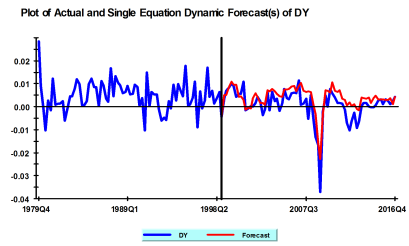
France



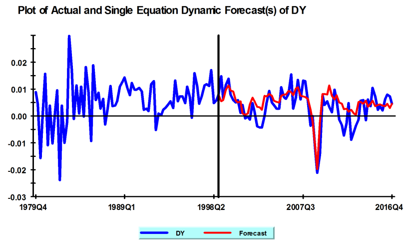
Germany



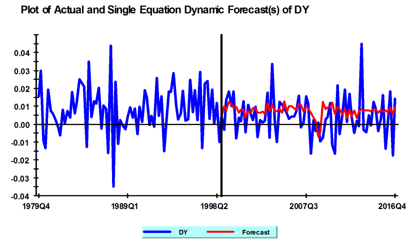
Italy



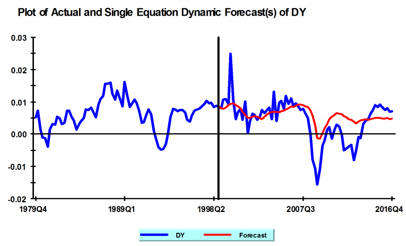
Netherlands



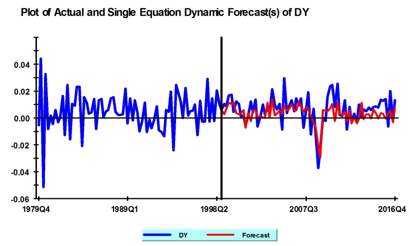
Norway



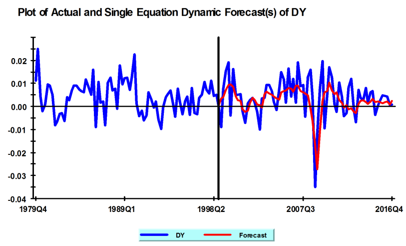
Spain



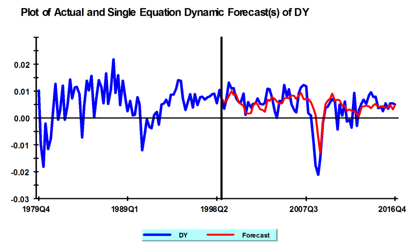
Sweden



Switzerland

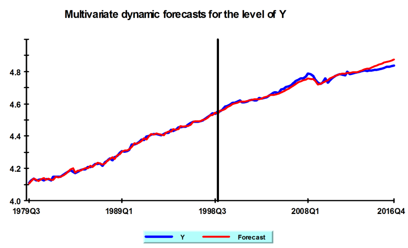


UK

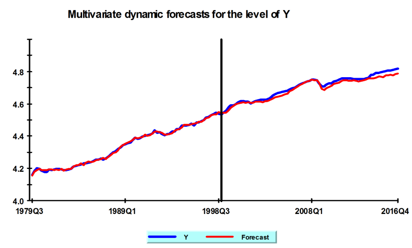


D.3 Forecasts (System Level)

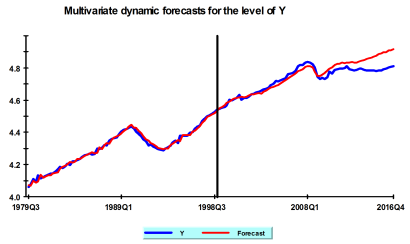
Austria



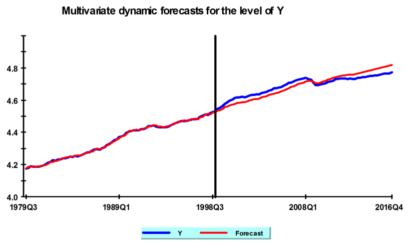
Belgium



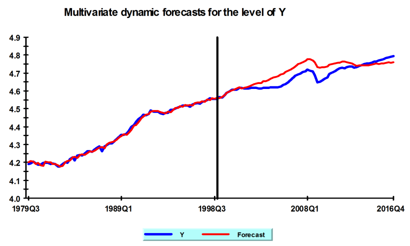
Finland



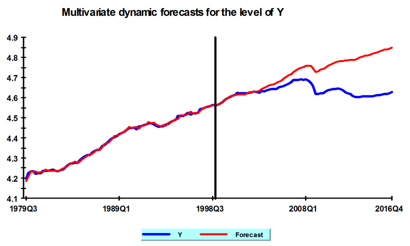
France



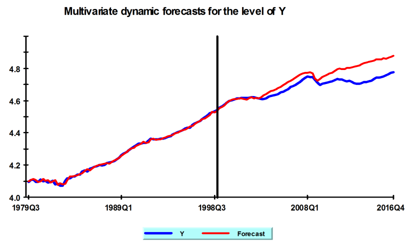
Germany



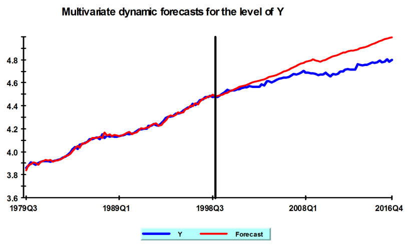
Italy



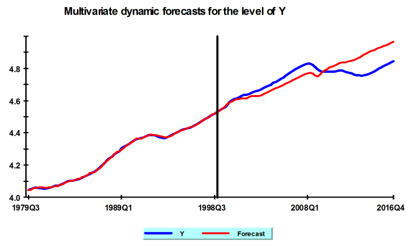
Netherlands



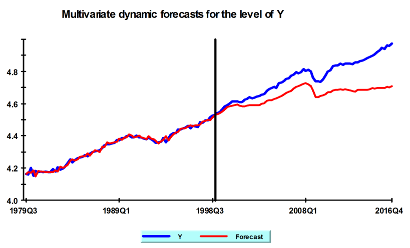
Norway



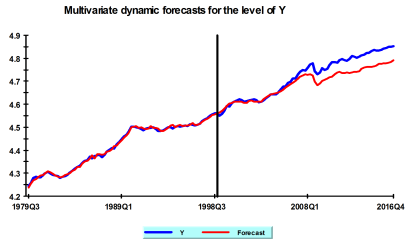
Spain



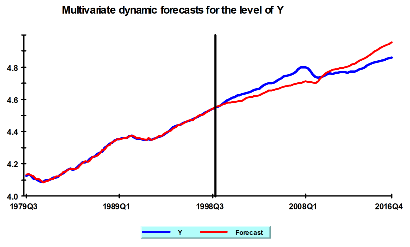
Sweden



Switzerland

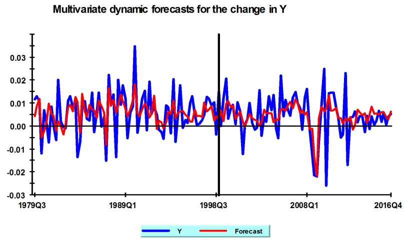


UK

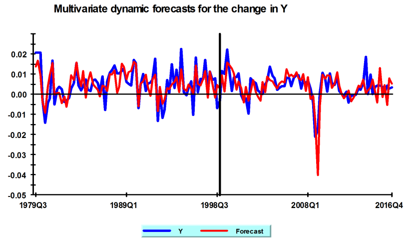


D.4 Forecasts (System Growth)

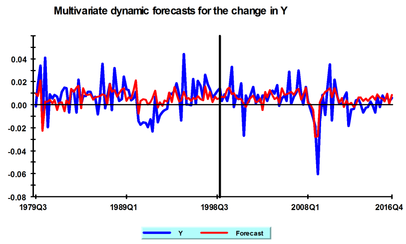
Austria



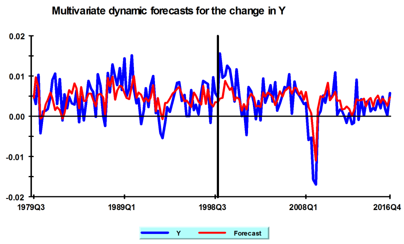
Belgium



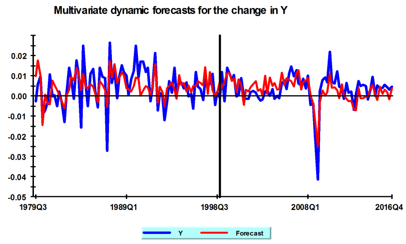
Finland



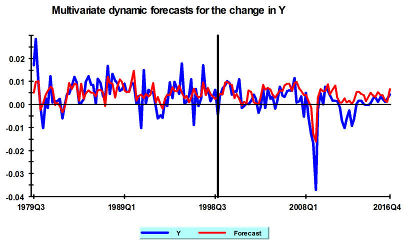
France



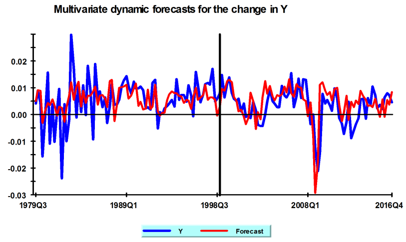
Germany



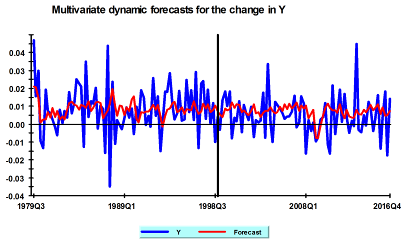
Italy



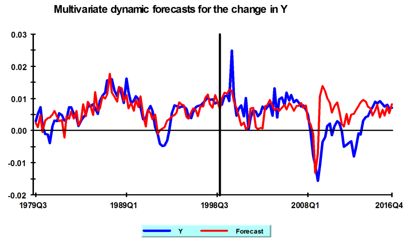
Netherlands



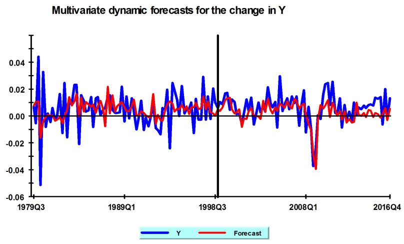
Norway



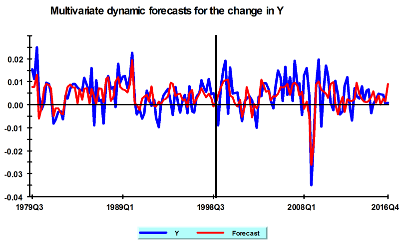
Spain



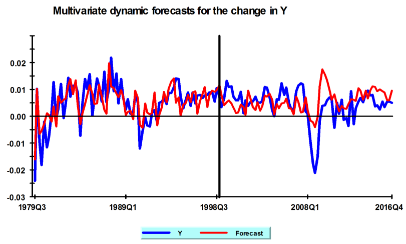
Sweden



Switzerland

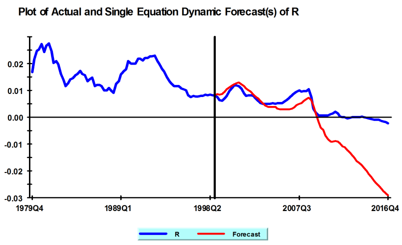


UK

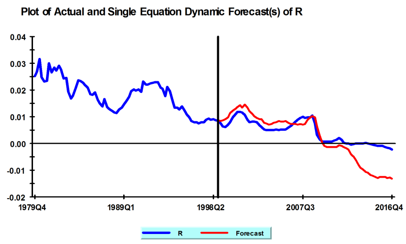


D.5 Forecasts (Short-Run Interest Rates - Taylor Rule)

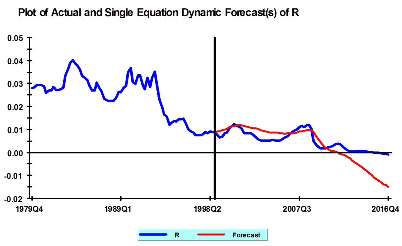
Austria



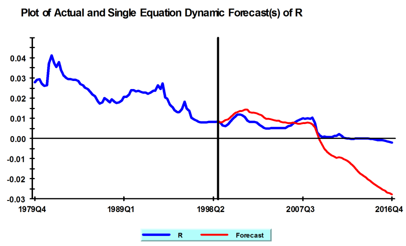
Belgium



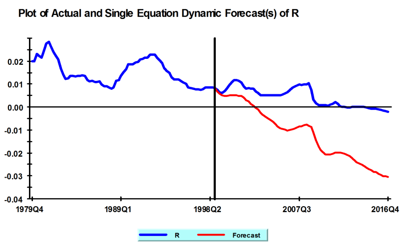
Finland



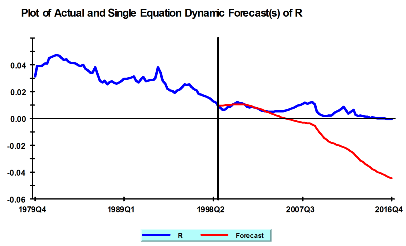
France



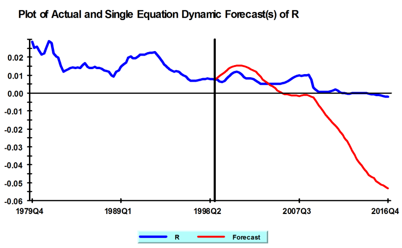
Germany



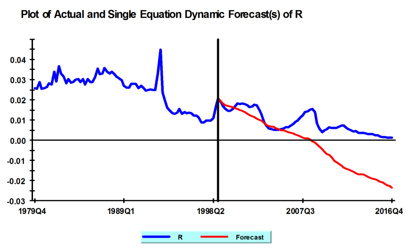
Italy



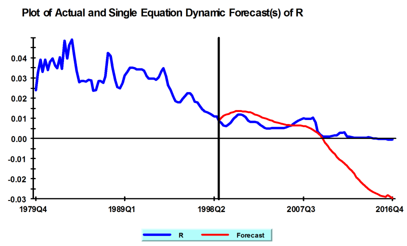
Netherlands



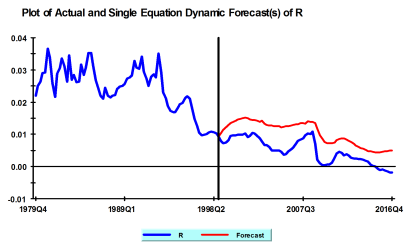
Norway



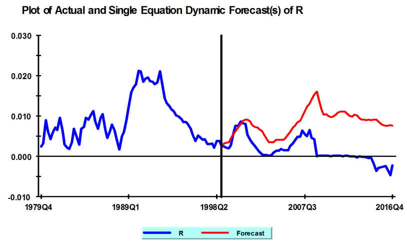
Spain



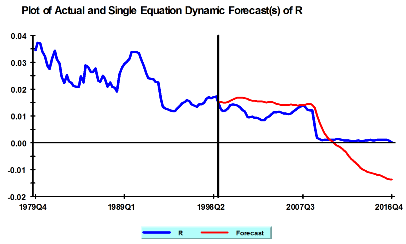
Sweden



Switzerland

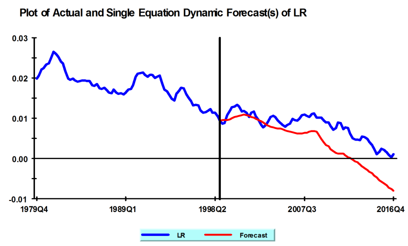


UK

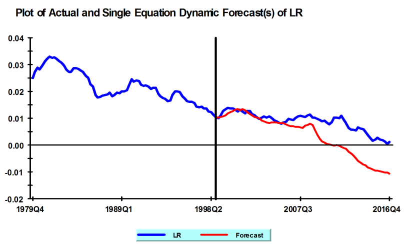


D.6 Forecasts (Long-Run Interest Rates - Taylor Rule)

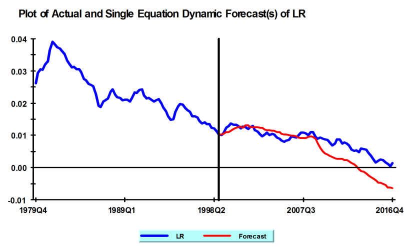
Austria



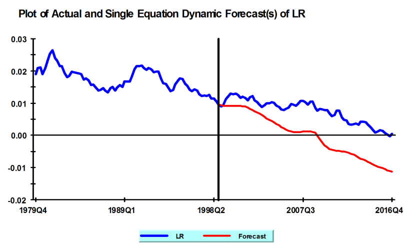
Belgium



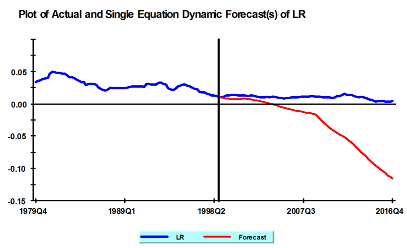
France



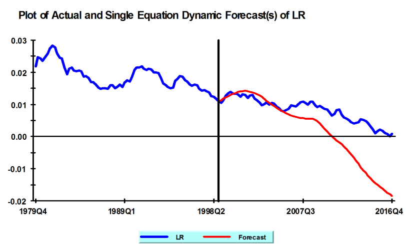
Germany



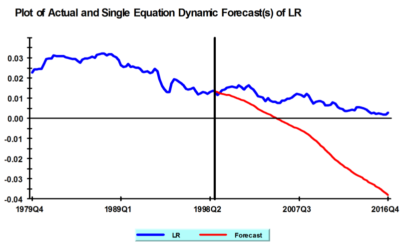
Italy



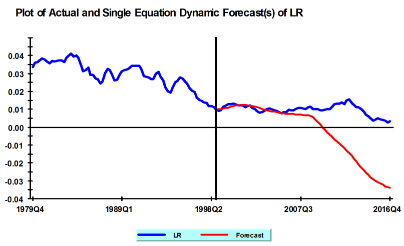
Netherlands



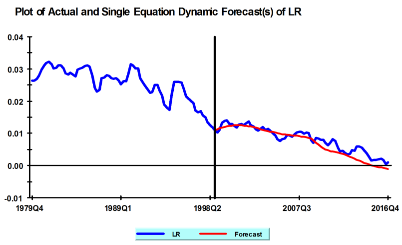
Norway



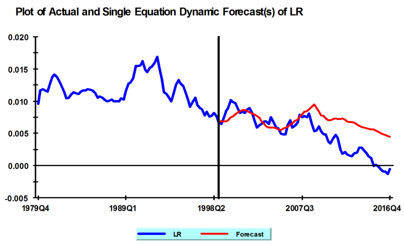
Spain



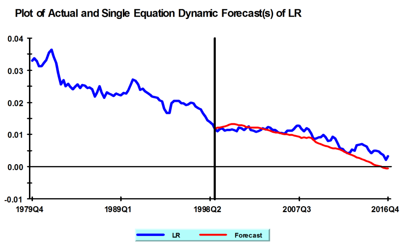
Sweden



Switzerland

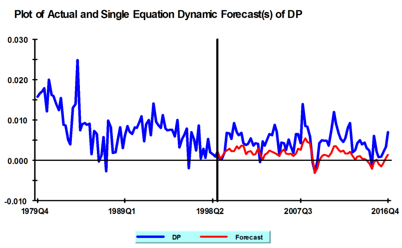


UK

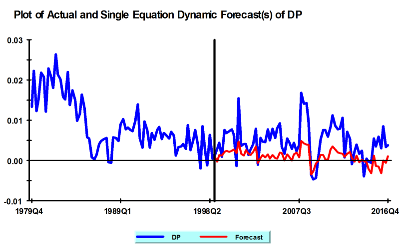


D.7 Forecasts (Inflation; Levels Relationship)

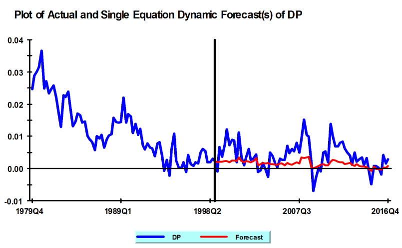
Austria



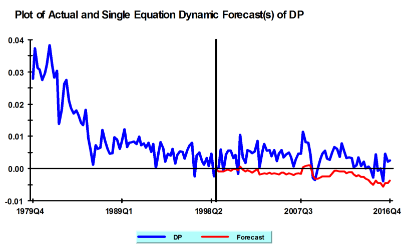
Belgium



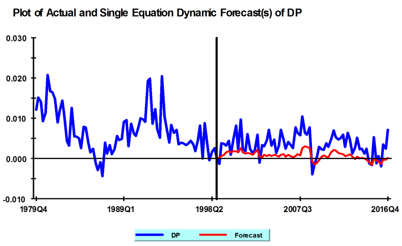
Finland



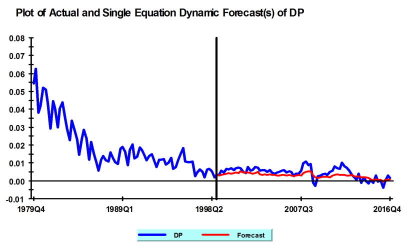
France



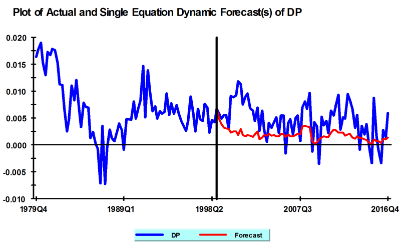
Germany



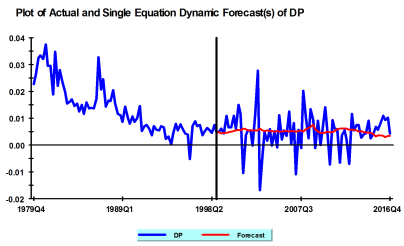
Italy



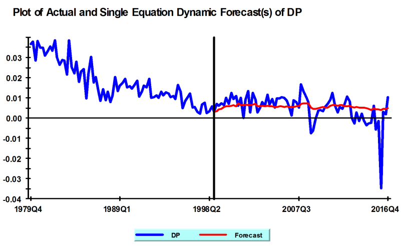
Netherlands



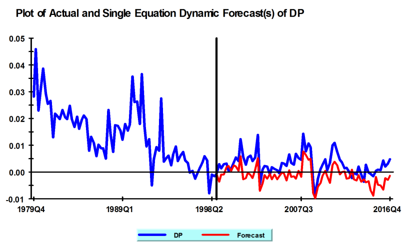
Norway



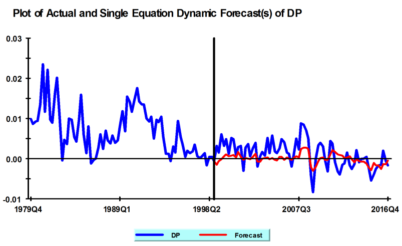
Spain



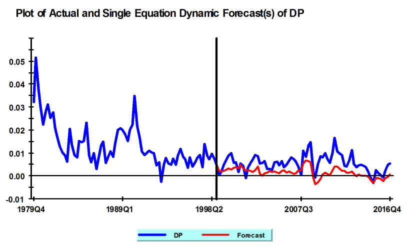
Sweden



Switzerland

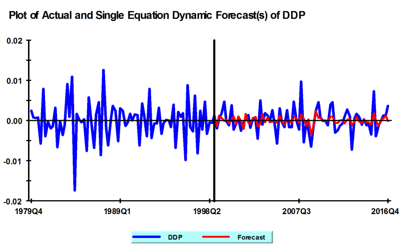


UK

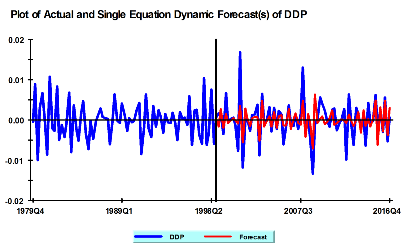


D.8 Forecasts (Change in Inflation Rate)

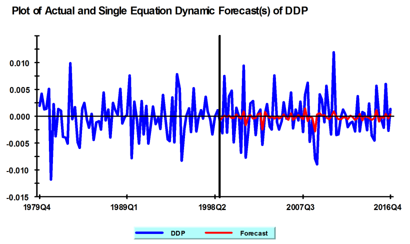
Austria



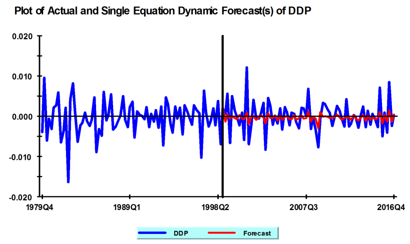
Belgium



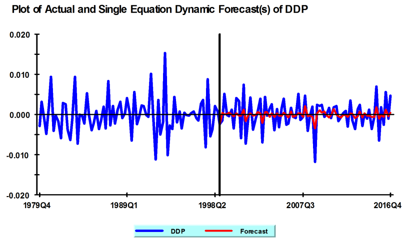
Finland



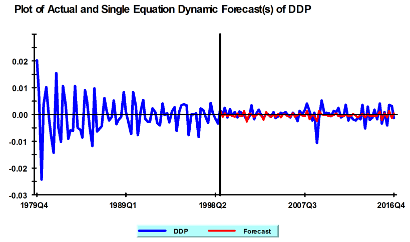
France



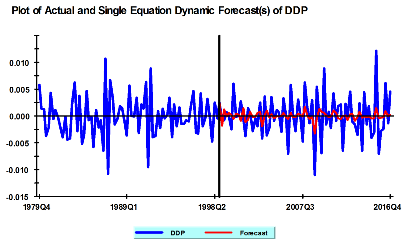
Germany



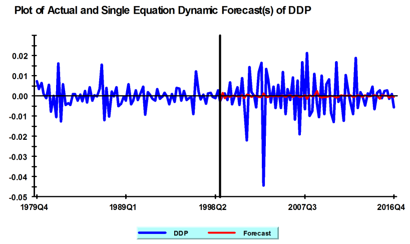
Italy



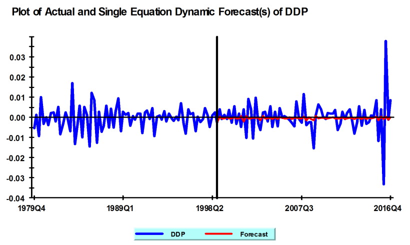
Netherlands



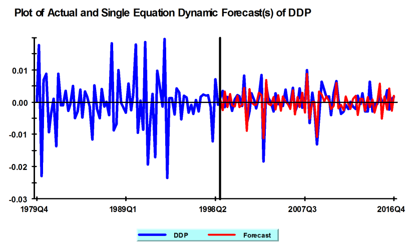
Norway



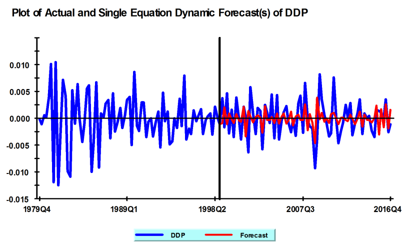
Spain



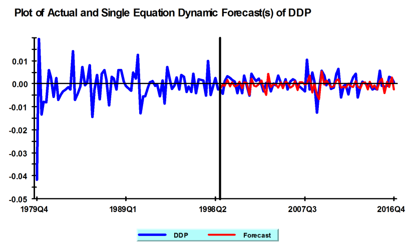
Sweden



Switzerland



UK



References

Afzal, M. and Ahmad, I. (2004). Estimating Long-Run Trade Elasticities in Pakistan: A Cointegration Approach [with Comments]. *Pakistan Development Review*, 43(4):757–770.

Ahmed, S. (1987). Government Spending, the Balance of Trade and the Terms of Trade in British History. *Journal of Monetary Economics*, 20(2):195–220.

Ahmed, S. and Rogers, J. H. (1995). Government Budget Deficits and Trade Deficits are Present Value Constraints Satisfied in Long-Term Data? *Journal of Monetary Economics*, 36(2):351–374.

Akhmadieva, V. and Smith, P. R. (2019). The Macroeconomic Impact of the Euro. *Scientific Annals of Economics and Business*, 66(SI2):229–249.

Aksoy, Y., De Grauwe, P., and Dewachter, H. (2002). Do Asymmetries Matter for European Monetary Policy? *European Economic Review*, 46(3):443–469.

Alessandria, G. A. and Choi, H. (2019). The Dynamics of the US Trade Balance and Real Exchange Rate: The J-Curve and Trade Costs? *National Bureau of Economic Research, Working Paper No. 25563*.

Altinay, G. (2007). Short-Run and Long-Run Elasticities of Import Demand for Crude Oil in Turkey. *Energy Policy*, 35(11):5829–5835.

Anderson, J. E. and Van Wincoop, E. (2003). Gravity with Gravitas: a Solution to the Border Puzzle. *American Economic Review*, 93(1):170–192.

Anderton, R. and Tewolde, T. (2011). The Global Financial Crisis: Understanding the Global Trade Downturn and Recovery. *World Economy*, 34(5):741–763.

Arize, A. (1986). The Supply and Demand for Imports and Exports in a Simultaneous Model. *Pakistan Economic and Social Review*, 24(2):57–76.

Armington, P. S. (1969). A Theory of Demand for Products Distinguished by Place of Production. *Staff Papers*, 16(1):159–178.

Arnone, M., Bandiera, L., Presbitero, A. F., et al. (2005). External Debt Sustainability: Theory and Empirical Evidence. *Catholic University of Piacenza Economics Working Paper*, 33:1–47.

Baffigi, A. (2011). Italian National Accounts, 1861–2011. *Bank of Italy Economic History Working Paper*, (18).

Balassa, B. (1964). The Purchasing-Power Parity Doctrine: a Reappraisal. *Journal of Political Economy*, 72(6):584–596.

Baldwin, R. and Taglioni, D. (2006). Gravity for Dummies and Dummies for Gravity Equations. *National Bureau of Economic Research, Working Paper No. 12516*.

Baldwin, R. E. (2008). *The Development and Testing of Heckscher-Ohlin Trade Models: A Review*. MIT Press.

Bank of England. (2018). *Bank of England. A Millennium of Macroeconomic Data*. Accessed at:. http://www.bankofengland.co.uk/statistics/research-datasets.

Barro, R. J. (1979). On the Determination of the Public Debt. *Journal of Political Economy*, 87(5, Part 1):940–971.

Barro, R. J. (1986). US Deficits since World War I. *The Scandinavian Journal of Economics*, pages 195–222.

BEA. (2018). *US Trade Data. Bureau of Economic Analysis*. Accessed at:. http://www.bea.gov.

Bems, R., Johnson, R. C., and Yi, K.-M. (2013). The Great Trade Collapse. *Annual Review of Economics*, 5(1):375–400.

Binh, D. T. T., Duong, N. V., Cuong, H. M., et al. (2011). Applying Gravity Model to Analyze Trade Activities of Vietnam. In *Forum for Research in Empirical International Trade Working Paper*.

Blanchard, O. (2019). Public Debt and Low Interest Rates. *American Economic Review*, 109(4):1197–1229.

Bohn, H. (1995). The Sustainability of Budget Deficits in a Stochastic Economy. *Journal of Money, Credit and Banking*, 27(1):257–271.

Bohn, H. (1998). The Behavior of US Public Debt and Deficits. *Quarterly Journal of Economics*, 113(3):949–963.

Bohn, H. (2005). The Sustainability of Fiscal Policy in the United States. *CESifo Working Paper Series 1446*.

Bohn, H. (2007). Are Stationarity and Cointegration Restrictions Really Necessary for the Intertemporal Budget Constraint? *Journal of Monetary Economics*, 54(7):1837–1847.

Bohn, H., Issing, O., and Westermann, F. (2016). The European Monetary System: Deficits and Cures. *Ifo Schnelldienst*, 69(09):18–24.

Boskin, M. J. (2020). Are Large Deficits and Debt Dangerous? *American Economic Association Papers and Proceedings*, 110:145–148.

Bowen, H. P., Leamer, E. E., and Sveikauskas, L. A. (1986). Multicountry, Multifactor Tests of the Factor Abundance Theory. *National Bureau of Economic Research, Working Paper No. 1918*.

Boyd, D., Caporale, G. M., and Smith, R. (2001). Real Exchange Rate Effects on the Balance of Trade: Cointegration and the Marshall–Lerner Condition. *International Journal of Finance and Economics*, 6(3):187–200.

Bridgman, B. (2008). Energy Prices and the Expansion of World Trade. *Review of Economic Dynamics*, 11(4):904–916.

Burke, P. J. and Liao, H. (2015). Is the Price Elasticity of Demand for Coal in China Increasing? *China Economic Review*, 36:309–322.

Bussière, M., Chudik, A., and Sestieri, G. (2009). Modelling Global Trade Flows: Results from a GVAR Model. *European Central Bank, Working Paper Series No. 1087*.

Caliendo, L., Feenstra, R. C., Romalis, J., and Taylor, A. M. (2015). Tariff Reductions, Entry, and Welfare: Theory and Evidence for the Last Two Decades. *National Bureau of Economic Research, Working Paper No. 21768*.

Campos, J., Ericsson, N. R., and Hendry, D. F. (2005). General-to-Specific Modeling: an Overview and Selected Bibliography. *FRB International Finance Discussion Paper*, (838).

Campos, N. F. and Macchiarelli, C. (2018). Symmetry and Convergence in Monetary Unions. *LEQS Paper No. 131/2018*, (131).

Caporale, G. M., Ali, F. M., Spagnolo, F., and Spagnolo, N. (2017). International Portfolio Flows and Exchange Rate Volatility in Emerging Asian Markets. *Journal of International Money and Finance*, 76:1–15.

Caporale, G. M., Ali, F. M., and Spagnolo, N. (2015a). Exchange Rate Uncertainty and International Portfolio Flows: A Multivariate GARCH-in-Mean Approach. *Journal of International Money and Finance*, 54:70–92.

Caporale, G. M., Gil-Alana, L. A., and Mudida, R. (2015b). Testing the Marshall–Lerner Condition in Kenya. *South African Journal of Economics*, 83(2):253–268.

Caruso, A., Reichlin, L., and Ricco, G. (2019). Financial and Fiscal Interaction in the Euro Area Crisis: This Time was Different. *European Economic Review*, 119:333–355.

Chaney, T. (2011). The Gravity Equation in International Trade: An Explanation. University of Chicago. *National Bureau of Economic Research, Working Paper No. 19285*.

Chinn, M. and Johnston, L. (1996). Real Exchange Rate Levels, Productivity and Demand Shocks: Evidence from a Panel of 14 Countries. *National Bureau of Economic Research, Working Paper No. 5709*.

Chor, D. and Manova, K. (2012). Off the Cliff and Cack? Credit Conditions and International Trade during the Global Financial Crisis. *Journal of International Economics*, 87(1):117–133.

Chudik, A., Mohaddes, K., Pesaran, M. H., and Raissi, M. (2017). Is there a Debt–Threshold Effect on Output Growth? *Review of Economics and Statistics*, 99(1):135–150.

Coakley, J., Kulasi, F., and Smith, R. (1996). Current Account Solvency and the Feldstein–Horioka Puzzle. *Economic Journal*, 106(436):620–627.

Cochrane, J. H. (2019). The Value of Government Debt. *Stanford Institute for Economic Policy Research, Working Paper No. 19-024*.

Constantinescu, C., Mattoo, A., and Ruta, M. (2015). The Global Trade Slowdown: Cyclical or Structural? *International Monetary Fund, Working Paper No. 15/6*.

Corsetti, G. and Müller, G. J. (2006). Twin Deficits: Squaring Theory, Evidence and Common Sense. *Economic Policy*, 21(48):598–638.

De Grauwe, P. (2013). The Political Economy of the Euro. *Annual Review of Political Science*, 16:153–170.

De Grauwe, P. (2016). The Legacy of the Eurozone Crisis and How to Overcome It. *Journal of Empirical Finance*, 39:147–155.

De Grauwe, P. (2018). *Economics of Monetary Union*. Oxford University Press.

De Grauwe, P. and Ji, Y. (2013). Self-Fulfilling Crises in the Eurozone: An Empirical Test. *Journal of International Money and Finance*, 34:15–36.

De Vita, G. and Abbott, A. (2002). Are Saving and Investment Cointegrated? an ARDL Bounds Testing Approach. *Economics Letters*, 77(2):293–299.

Disdier, A.-C. and Head, K. (2008). The Puzzling Persistence of the Distance Effect on Bilateral Trade. *Review of Economics and Statistics*, 90(1):37–48.

Dollar, D. and Kraay, A. (2003). Institutions, Trade, and Growth. *Journal of Monetary Economics*, 50(1):133–162.

Dornbusch, R., Fischer, S., and Samuelson, P. A. (1977). Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods. *American Economic Review*, 67(5):823–839.

Dumas, B. (1992). Dynamic Equilibrium and the Real Exchange Rate in a Spatially Separated World. *Review of Financial Studies*, 5(2):153–180.

Durdu, C. B., Mendoza, E. G., and Terrones, M. E. (2013). On the Solvency of Nations: Cross-Country Evidence on the Dynamics of External Adjustment. *Journal of International Money and Finance*, 32:762–780.

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

Eaton, J., Kortum, S., Neiman, B., and Romalis, J. (2011). Trade and the Global Recession. *National Bureau of Economic Research, Working Paper No. 16666*.

Edwards, S. (1986). Are Devaluation Contractionary? *Review of Economics and Statistics*, 68(3):501–508.

Eichengreen, B. (2010). Out–of–the–Box Thoughts about the International Financial Architecture. *Journal of International Commerce, Economics and Policy*, 1(1):1–20.

Engle, R. F. and Granger, C. W. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica: Journal of the Econometric Society*, 55(2):251–276.

Farhi, E. and Werning, I. (2017). Fiscal Unions. *American Economic Review*, 107(12):3788–3834.

FED. (2017). *Federal Reserve Economic Data*. Accessed at:. http://fred.stlouisfed.org.

Felbermayr, G. J. (2005). Dynamic Panel Data Evidence on the Trade-Income Relation. *Review of World Economics*, 141(4):583–611.

Feldstein, M. and Horioka, C. (1980). Domestic Saving and International Capital Flows. *Economic Journal*, 90(358):314–329.

Findlay, R. and O'rourke, K. H. (2009). *Power and Plenty: Trade, War, and the World Economy in the Second Millennium*. Princeton University Press.

Fisher, I. (1930). *Theory of Interest: as Determined by Impatience to Spend Income and Opportunity to Invest It*. Augustusm Kelly Publishers, Clifton.

Fostel, A., Geanakoplos, J., and Phelan, G. (2019). Global Collateral and Capital Flows. *National Bureau of Economic Research, Working Paper No. 25583*.

Frankel, J. A. and Romer, D. H. (1999). Does Trade Cause Growth? *American Economic Review*, 89(3):379–399.

Freund, C. (2009). *The Trade Response to Global Downturns: Historical Evidence*. The World Bank.

Fuller, W. A. (2009). *Introduction to Statistical Time Series*, volume 428. John Wiley and Sons.

Ghodsi, M., Grübler, J., and Stehrer, R. (2016). Import Demand Elasticities Revisited. *wiiw Working Paper No. 132*.

Goldberg, P. K. and Pavcnik, N. (2016). The Effects of Trade Policy. In *Handbook of Commercial Policy*, volume 1, pages 161–206. Elsevier.

Gourinchas, P.-O., Rey, H., and Sauzet, M. (2019). The International Monetary and Financial System. *Annual Review of Economics*, 11:859–893.

Graham, C., Helliwell, J., Kano, T., Murray, J., and Schembri, L. (2004). The Euro and Trade: is there a Positive Effect. In *Proceeding of Conference of the 38th Annual Meetings of the Canadian Economics Association*.

Granger, C. J. (1986). Developments in the Study of Cointegrated Economic Variables. *Oxford Bulletin of Economics and Statistics*, 48(3):213–228.

Greiner, A., Koeller, U., and Semmler, W. (2007). Debt Sustainability in the European Monetary Union: Theory and Empirical Evidence for Selected Countries. *Oxford Economic Papers*, 59(2):194–218.

Griliches, Z. (1967). Distributed Lags: A Survey. *Econometrica: Journal of the Econometric Society*, 35:16–49.

Hanson, G. H., Lind, N., and Muendler, M.-A. (2015). The Dynamics of Comparative Advantage. *National Bureau of Economic Research, Working Paper No. 21753*.

Hasan, M. A. and Khan, A. H. (1994). Impact of Devaluation on Pakistan's External Trade: An Econometric Approach. *The Pakistan Development Review*, 33(4):1205–1215.

Haynes, S. E. and Stone, J. A. (1983). Secular and Cyclical Responses of U.S. Trade to Income: an Evaluation of Traditional Models. *Review of Economics and Statistics*, 65(1):87–95.

Helleiner, G. K. (1973). Manufactured Exports from Less-Developed Countries and Multinational Firms. *Economic Journal*, 83(329):21–47.

Helpman, E. (1987). Imperfect Competition and International Trade: Evidence from Fourteen Industrial Countries. *Journal of the Japanese and International Economies*, 1(1):62–81.

Hills, S., Thomas, R., and Dimsdale, N. (2015). Three Centuries of Data–Version 2.2. *Bank of England*.

Hooper, P., Johnson, K., and Marquez, J. R. (2000). Trade Elasticities for the G-7 Countries. *Princeton Studies in International Economics*, (87).

Hoover, K. D. and Perez, S. J. (1999). Data Mining Reconsidered: Encompassing and the General-to-Specific Approach to Specification Search. *Econometrics Journal*, 2(2):167–191.

Houthakker, H. S. and Magee, S. P. (1969). Income and Price Elasticities in World Trade. *Review of Economics and Statistics*, pages 111–125.

Hutchinson, D. and Ploeckl, F. (2012). *What Was the Australian GDP or CPI Then?* Accessed at:. http://www.measuringworth.com/datasets/australiadata/.

Imbs, J., Mejean, I., et al. (2010). Trade Elasticities: a Final Report for the European Commission. *European Commission, European Economy, Economic Papers 432*.

IMF. (2014). *World Economic Outlook Database. International Monetary Fund*. Accessed at:. http://www.imf.org/external/pubs/ft/weo/2014/02/weodata/index.aspx.

Irwin, D. A. (2002). Long-Run Trends in World Trade and Income. *World Trade Review*, 1(1):89–100.

Johansen, S. (1988). Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control*, 12(2-3):231–254.

Johansen, S. and Juselius, K. (1990). Maximum Likelihood Estimation and Inference on Cointegration—with Applications to the Demand for Money. *Oxford Bulletin of Economics and Statistics*, 52(2):169–210.

Jones, M. T. and Obstfeld, M. (1997). Saving, Investment, and Gold: A Reassessment of Historical Current Account Data. *National Bureau of Economic Research, Working Paper No. 6103*.

JST. (2019). *The Jordà-Schularick-Taylor Macrohistory Database*. Accessed at:. http://www.macrohistory.net/data/.

Kee, H. L., Nicita, A., and Olarreaga, M. (2008). Import Demand Elasticities and Trade Distortions. *Review of Economics and Statistics*, 90(4):666–682.

Khan, A. H. and Khanum, S. (1997). Exports and Employment: A Case Study of Pakistan. *Economia Internazionale*, 50(2):261–282.

Khan, M. S. (1974). Import and Export Demand in Developing Countries. *Staff Papers*, 21(3):678–693.

Kohli, U. (1991). *Technology, Duality, and Foreign Trade: the GNP Function Approach to Modeling Imports and Exports*. Harvester Wheatsheaf.

Kortum, S., Neiman, B., Eaton, J., et al. (2016). Obstfeld and Rogoff's International Macro Puzzles: A Quantitative Assessment. *National Bureau of Economic Research, Working Paper No. 21774*.

Krugman, P. (1979). Increasing Returns, Monopolistic Competition, and International Trade. *Journal of International Economics*, 9(4):469–479.

Krugman, P. R. (2014). *International Economics: Theory and Policy*. Prentice Hall.

Lee, J. and Chinn, M. D. (1998). The Current Account and the Real Exchange Rate: a Structural VAR Analysis of Major Currencies. *National Bureau of Economic Research, Working Paper No. 6495*.

Leontief, W. (1953). Domestic Production and Foreign Trade; the American Capital Position Re-Examined. *Proceedings of the American Philosophical Society*, 97(4):332–349.

Litman, S. (1926). The Effects of the World War on Trade. *Annals of the American Academy of Political and Social Science*, 127(1):23–29.

Lucas, R. E. (1990). Why Doesn't Capital Flow from Rich to Poor Countries? *American Economic Review*, 80(2):92–96.

Maddison, A. (2003). *The World Economy: Historical Statistics*. Development Centre Studies, OECD Publishing.

Maddison, A. (2013). *The Maddison-Project Database*. Groningen Growth and Development Centre. Accessed at:. http://www.rug.nl/ggdc/historicaldevelopment/maddison/releases/maddison-project-database-2018.

Madsen, J. B. (2001). Trade Barriers and the Collapse of World Trade during the Great Depression. *Southern Economic Journal*, 67(4):848–868.

Meissner, C. M. et al. (2003). Exchange-Rate Regimes and International Trade: Evidence from the Classical Gold Standard Era. *American Economic Review*, 93(1):344–353.

Micco, A., Stein, E., and Ordoñez, G. (2003). The Currency Union Effect on Trade: Early Evidence from EMU. *Economic Policy*, 18(37):315–356.

Mitchell, B. R. (1980). *European Historical Statistics 1750–1975*. Palgrave Macmillan UK.

Mitchell, B. R. (2013). *International Historical Statistics, 1750–2010*. Palgrave Macmillan UK.

Mohaddes, K. and Raissi, M. (2018). Compilation, Revision and Updating of the Global VAR (GVAR) Database, 1979Q2-2016Q4.

Mountford, A. and Uhlig, H. (2009). What are the Effects of Fiscal Policy Shocks? *Journal of Applied Econometrics*, 24(6):960–992.

Muellbauer, J. and Murphy, A. (1990). Is the UK Balance of Payments Sustainable? *Economic Policy*, 5(11):347–396.

Naqvi, S. N. H. (1982). *The P.I.D.E. Macro-Econometric Model of Pakistan's Economy*, volume 1. Pakistan Institute of Development Economics.

Nell, K. S. and Santos, L. D. (2008). The Feldstein–Horioka Hypothesis Versus the Long-Run Solvency Constraint Model: A Critical Assessment. *Economics Letters*, 98(1):66–70.

Obstfeld, M. and Rogoff, K. (2002). *Foundations of International Macroeconomics*. MIT Press, Massachusetts.

ONS. (2017). *UK Trade Data. Office for National Statistics*. Accessed at:. http://www.ons.gov.uk/.

Patterson, K. D. (2000). *An Introduction to Applied Econometrics: a Time Series Approach*. Macmillan.

Pesaran, M. H. (2015). *Time Series and Panel Data Econometrics*. Oxford University Press.

Pesaran, M. H., Shin, Y., and Smith, R. J. (2001). Bounds Testing Approaches to the Analysis of Level Relationships. *Journal of Applied Econometrics*, 16(3):289–326.

Pesaran, M. H., Shin, Y., and Smith, R. P. (1999). Pooled Mean Group Estimation of Dynamic Heterogeneous Panels. *Journal of the American Statistical Association*, 94(446):621–634.

Pesaran, M. H. and Smith, R. (1995). Estimating Long-Run Relationships from Dynamic Heterogeneous Panels. *Journal of Econometrics*, 68(1):79–113.

Pesaran, M. H. and Smith, R. P. (1998). Structural Analysis of Cointegrating VARs. *Journal of Economic Surveys*, 12(5):471–505.

Pesaran, M. H. and Smith, R. P. (2016). Counterfactual Analysis in Macroeconometrics: An Empirical Investigation into the Effects of Quantitative Easing. *Research in Economics*, 70(2):262–280.

Pesaran, M. H. and Smith, R. P. (2018). Tests of Policy Interventions in DSGE Models. *Oxford Bulletin of Economics and Statistics*, 80(3):457–484.

Pesaran, M. H., Vanessa Smith, L., and Smith, R. P. (2007). What if the UK or Sweden had Joined the Euro in 1999? an Empirical Evaluation Using a Global VAR. *International Journal of Finance and Economics*, 12(1):55–87.

Potjagailo, G. (2017). Spillover Effects from Euro Area Monetary Policy across Europe: A Factor-Augmented VAR Approach. *Journal of International Money and Finance*, 72:127–147.

Pritchett, L. (1996). Measuring Outward Orientation in LDCs: Can It be Done? *Journal of Development Economics*, 49(2):307–335.

Quintos, C. E. (1995). Sustainability of the Deficit Process with Structural Shifts. *Journal of Business and Economic Statistics*, 13(4):409–417.

Reyes-Heroles, R. et al. (2017). The Role of Trade Costs in the Surge of Trade Imbalances. 2017 Meeting Papers 212, Society for Economic Dynamics.

Rogoff, K. (1992). Traded Goods Consumption Smoothing and the Random Walk Behavior of the Real Exchange Rate. *National Bureau of Economic Research, Working Paper No. 4119*.

Rose, A. K. (1991). The Role of Exchange Rates in a Popular Model of International Trade: Does the 'Marshall–Lerner' Condition Hold? *Journal of International Economics*, 30(3-4):301–316.

Rose, A. K. (2000). One Money, One Market: the Effect of Common Currencies on Trade. *Economic Policy*, 15(30):08–45.

Rose, A. K. (2017). Why Do Estimates of the EMU Effect on Trade Vary so Much? *Open Economies Review*, 28(1):1–18.

Rudebusch, G. and Svensson, L. E. (1999). Policy Rules for Inflation Targeting. In *Monetary Policy Rules*, pages 203–262. University of Chicago Press.

Saka, O., Fuertes, A.-M., and Kalotychou, E. (2015). ECB Policy and Eurozone Fragility: Was De Grauwe Right? *Journal of International Money and Finance*, 54:168–185.

Samuelson, P. A. (1964). Theoretical Notes on Trade Problems. *Review of Economics and Statistics*, pages 145–154.

Schnatz, B., Bussiere, M., et al. (2006). Evaluating China's Integration in World Trade with a Gravity Model Based Benchmark. *European Central Bank, Working Paper Series No. 693*.

Schoder, C. (2014). The Fundamentals of Sovereign Debt Sustainability: Evidence from 15 OECD Countries. *Empirica*, 41(2):247–271.

Schoder, C., Proaño, C. R., and Semmler, W. (2013). Are the Current Account Imbalances between EMU Countries Sustainable? Evidence from Parametric and Non-Parametric Tests. *Journal of Applied Econometrics*, 28(7):1179–1204.

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

Smith, R. (2009). EMU and the Lucas Critique. *Economic Modelling*, 26(4):744–750.

Smits, J. P. H., Horlings, E., Zanden, J. L., et al. (2000). *Dutch GNP and Its Components, 1800–1913*. Groningen Growth and Development Centre Groningen.

Soderbery, A. (2015). Estimating Import Supply and Demand Elasticities: Analysis and Implications. *Journal of International Economics*, 96(1):1–17.

Sousa, J. (2014). Estimation of Price Elasticities of Demand for Alcohol in the United Kingdom. *London: Her Majesty's Revenue and Customs, Working Paper 16*.

Thom, R. and Walsh, B. (2002). The Effect of a Currency Union on Trade: Lessons from the Irish Experience. *European Economic Review*, 46(6):1111–1123.

Tinbergen, J. (1966). Shaping the World Economy; Suggestions for an International Economic Policy. *Economic Journal*, 76(301):92–95.

Toda, H. Y. and Yamamoto, T. (1995). Statistical Inference in Vector Autoregressions with Possibly Integrated Processes. *Journal of Econometrics*, 66(1-2):225–250.

Trading Economics. (2018). *Trading Economics. United Kingdom Balance of Trade*. Accessed at:. http://tradingeconomics.com/united-kingdom/balance-of-trade.

Trehan, B. and Walsh, C. E. (1988). Common Trends, the Government's Budget Constraint, and Revenue Smoothing. *Journal of Economic Dynamics and Control*, 12(2-3):425–444.

Trehan, B. and Walsh, C. E. (1991). Testing Intertemporal Budget Constraints: Theory and Applications to US Federal Budget and Current Account Deficits. *Journal of Money, Credit and Banking*, 23(2):206–223.

Uhlig, H. (2013). Some Fiscal Calculus. *American Economic Review*, 100(2).

UNCTAD. (1997). *Trade and Development Report 1997*. United Nations Conference on Trade and Development.

Upadhyaya, K. P. and Upadhyay, M. P. (1999). Output Effects of Devaluation: Evidence from Asia. *Journal of Development Studies*, 35(6):89–103.

Villareal, M. and Fergusson, I. F. (2017). The North American Free Trade Agreement (NAFTA). *CRS Report R42965*.

Warner, D. and Kreinin, M. E. (1983). Determinants of International Trade Flows. *Review of Economics and Statistics*, 65(1):96–104.

Wood, A. et al. (2009). *A Practical Heckscher-Ohlin Model*. University of Oxford, Department of International Development.

World Bank. (2017). *World Bank Data*. Accessed at:. http://data.worldbank.org.

WTO. (2017). *Time Series on International Trade. World Trade Organisation*. Accessed at:. http://timeseries.wto.org.

WTO. (2019). WTO Lowers Trade Forecast as Tensions Unsettle Global Economy. Press Report 1 October 2019. Technical report, World Trade Organisation.

Wyplosz, C. et al. (2007). *Debt Sustainability Assessment: The IMF Approach and Alternatives*. Graduate Institute of International Studies Geneva.

Yeats, A. J. (1999). *Just How Big is Global Production Sharing?* The World Bank.

Yi, K.-M. (2003). Can Vertical Specialization Explain the Growth of World Trade? *Journal of Political Economy*, 111(1):52–102.