EUROPEAN UNIVERSITY INSTITUTE
Department of Economics

Three Essays on EMU, Exchange Rates and Time Series Econometrics

Dmitri Boreiko

Thesis submitted for assessment with a view to obtaining the degree of Doctor of the European University Institute

Florence, March 2006

European University Institute

Department of Economics

Ph.D. dissertation

Three Essays on EMU, Exchange Rates and Time Series Econometrics

Dmitri Boreiko

Thesis committee:

Anindya Banerjee

Mike Artis

Agnes Benassy-Quere

Mark Salmon

Acknowledgments

I thank Professors Anindya Banerjee, Mike Artis and Helmut Luetkepohl for excellent supervision and guidance during recent years, for their useful comments, discussions and encouragement. I am also very grateful to the administrative staff of the Economics Department of the EUI for their support and service.

I acknowledge financial support from the Italian Ministry of Foreign Affairs, the European University Institute and the European Commission.

Contents

[1] This chapter is based on the author's article published in the *International Journal of Finance and Economics* 8, 2003 : 309–325.

Introduction

This introductory chapter to the dissertation firstly explains the motivation for focusing on this topic. Next, it outlines the main areas of research pursued and describes the main research findings and principal contributions made to economic knowledge. Lastly, it identifies the structure of thesis and summarises its major conclusions.

'May you live in interesting times' the old Chinese saying goes. This is certainly true for a young economist searching for fresh topics of future research. Undoubtedly, in the last 20 years of European history, we have witnessed many dramatic political and economic transformations. However, one of the most important events, from the European point of view, has been the creation of a single European currency. My work on this dissertation commenced during the first years of Stage III of EMU and just before the Euro's introduction. The sheer scale of this event, with eleven developed countries[2] abdicating their individual monetary policy discretion in favour of the new European Central Bank, was unprecedented in history and, as a result, attracted immense research interest, providing several research opportunities.

First, the creation of European Monetary Union encompassing 12 EU members raised the question of the irrevocable parity at which the members of the union locked

[2] With Greece joining the EMU two years later.

their currencies against the new Euro. As various historical examples show the choice of a target for nominal exchange rate is crucial, and misjudgment in this respect can have disastrous effects. Second issue has been the estimation not only of current or historic misalignments of currencies according to estimated equilibrium values, but also of future exchange rates movements. Furthermore, once EMU members exchanged their currencies for the Euro, one question naturally raised concerned the appropriate level of the Euro against other major world currencies. One strand of the literature followed the view that Purchasing Power Parity (PPP) is the only determinant of exchange rates movements, whereas an alternative strand held that this to be true in the very long run, but with regard to shorter time horizons, the exchange rates depend on certain economic processes identified as fundamentals. Driver and Westway (2004) have provided a succinct review of the main research contributions in this area, yet this topic still appears far from fully researched. Moreover, in this author's view, many of the recent studies using the new econometric cointegration methods have focussed on getting the right result, instead of conducting proper model formulation and testing.

In this context, attention has been turned to the specific question of Equilibrium Exchange Rates (EERs), analysed using a statistical, or as it is sometimes called the 'reduced-model' approach. Using a constructed dataset for nineteen OECD countries, including the eleven EMU members (excluding Luxembourg), for the time period 1974

- 2003, it is shown that real effective exchange rates were indeed cointegrated with economic fundamentals for each country in the study, and that it is possible to estimate cointegrated Vector Autoregressive (VAR) models for each country both by using relative productivity and net foreign asset position fundamentals (3-VAR models), and by adding real interest rate differentials and relative fiscal spending (5-VAR models). However, the results also showed that the level of heterogeneity across countries is quite large and, apart from relative productivity, fundamentals' coefficient estimates were of different signs. This appears to contradict the findings by Alberola *et al.* (1999), who however omitted to report statistics required to assess the statistical significance of the coefficient estimates, perhaps indicating that their constructed models were inappropriate. Hansen and Roeger's (2000) study, in which the results identified by Alberola *et al.* (1999) could not be replicated may serve as further confirmation of this finding.

This study's estimates, on the other hand, tracked real exchange rates quite well, showing a substantial level of adjustment of exchange rates to equilibrium trends (displaying on average half-lives of around one year). Moreover, given that the data used to construct the fundamentals and forecasts for the following periods are readily available, it is possible to construct forecasts of real effective and nominal bilateral exchange rates to be used in further economic analysis. Even in the face of the uncertainty surrounding statistical models of exchange rate determination, it is suggested,

the results are still useful and may prove superior to black box or PPP exchange rate estimates.

An equally important feature of initial years of EMU's existence was the imminent enlargement of the EU, with ten new members duly joining the EU in 2004 and taking on obligations to participate in the ERM2 system, and in EMU in the future. In light of this, the question of the new members' preparedness for joining EMU took on greater importance, with two main issues to be addressed. First, they have to meet the Maastricht Treaty criteria in order to qualify. The Maastricht criteria limit candidate countries' ability to conduct monetary and fiscal policy decisions at their discretion - a clear cost incurred by countries on their way to joining the Euro. Second, abdication of sovereign monetary policy has its own corollary costs and benefits, usually assessed in the framework of the Optimum Currency Area (OCA) theory, which advocates monetary union on the condition that adjustment of the bilateral exchange rate is either ineffective or unnecessary to stabilise output (Mundell, 1961).

As concerns the process of joining the EU, three groups of transition countries may be identified: Czech Republic, Estonia, Hungary, Poland and Slovenia, who started negotiations first and were referred to as 1998 Accession Group; a second cohort described as the 2000 Accession Group, consisting of Bulgaria, Latvia, Lithuania, Romania and the Slovak Republic, which were less advanced than the first group in the negotiation process; thirdly, countries such as Croatia, which, for various reasons,

were not a part of the 2001 negotiation process. Given segregation in this respect, a question naturally arising was whether a division along similar lines would apply to joining EMU. In response to this question, this study identified a group of countries which were "more EMU-ready", or at least better suited to enter into a currency board against the Euro.

Following the approach of Artis and Zang (1998) applied to the EMU members, the technique of cluster analysis was employed in order to check for the existence of homogeneous groups among the new EU members. This technique is used to examine the similarities and dissimilarities of economic structure, and to group countries according to varying sets of criteria. To counter the problems of incomplete and noisy data, the more powerful technique of fuzzy clustering was employed. This method splits country data into groups, by assigning to each object membership coefficients which indicate its degree of "belongingness" to each of the groups.

The analysis showed that, among the new EU members, there were indeed several clearly defined groups with regard to nominal and real convergence to EMU criteria. Estonia and Slovenia were the leaders in both nominal and real convergence, whereas other countries from the 1998 Accession Wave achieved substantial results only in real convergence. Poland was excluded from the leading group during the latest years due to deteriorated economic performance.

This chapter of the thesis was completed in 2002 when the issue of joining the EMU was rather distant. Nevertheless, the chapter's conclusions on Estonia and Slovenia leadership in nominal and real convergence were confirmed in 2006. Slovenia and Estonia are the most likely candidates to join the EMU in January 2007, whereas other countries (including Lithuania, that is discouraged to join at the same time due to the high inflaiton rate) will do that later.

In the course of the study on estimating equilibrium exchange rates for the OECD countries, a further question of structural stability of the estimated models presented itself as a promising avenue of inquiry. As one of the key assumptions in the econometric modelling, this assumption is used to estimate the parameters of the model, and allows the model's use for forecasting. Therefore any partial or overall parameter instability may have severe consequences on inference and model validity.

Given the importance of stability, it is surprising that so many empirical studies didn't pay sufficient attention to stability tests for their estimated models before proceeding to conclusions regarding the nature of the economic relations evidenced. A possible explanation for the scarcity of the stability tests in empirical work may be that, with the advent of new econometric models and estimation methods, time elapses before corresponding stability tests theory is developed. This was the case with single-equation regression models: by the time formal stability tests were developed, the focus of econometric research had shifted to multivariate regression models and to dy-

namic models such as VAR models. Development of new stability tests coincided with the emergence of non-stationary dynamic models analysis which in turn required new estimation methods, inference and of course new methods to check stability.

In the course of the conducting this research, the author was involved in programming the stability analysis section for econometric package JMulTi[3]. As a result, it was decided to analyze a small sample performance of Chow-type tests, looking for a single structural break under different specifications of multivariate models and types of breaks. The three most widely used types of Chow test were compared to each other, and to the generalised Chow-type tests with unknown break-point by Andrews and Ploberger (1994) in terms of size and power. For all the tests, the bootstrap versions were used to reduce the massive size distortions of the original tests. The study's results indicated the superior performance of a sample-split test against all other versions. The tests were also applied to EMU money demand dataset in order to test stability and were employed to control the stability of the real exchange rate cointegration systems in the first chapter.

The following three chapters will examine in greater detail the issues described in this introduction.

[3] JMulTi is an interactive software programme designed for univariate and multivariate time series analysis. At the time of writing JMulTi is distributed as freeware, and is available at www.jmulti.de.

References

[1] Alberola, E., Cervero, L., Ubide, A. and Lopez, H. (1999), "Global Equilibrium Exchange Rates: Euro, Dollar, "Ins", "Outs", and Other Major Currencies in a Panel Cointegration Framework", IMF Working Paper, 99/175.

[2] Andrews, D.W.K. and W. Ploberger (1994), "Optimal Tests When a Nuisance Parameter is Present only under the Alternative", Econometrica, 62,1383-1414.

[3] Artis, M. and Zhang, W. (1998), "Membership of EMU: A Fuzzy Clustering Analysis of Alternative Criteria", IUE Working Paper 98/52.

[4] Driver, R. L. and Westaway, P.F. (2004), "Concepts of Equilibrium Exchange Rates", Working Paper 248, Bank of England.

[5] Hansen, J. and Roeger, W. (2000), "Estimation of Real Equilibrium Exchange Rates", European Economy - Economic Papers 144, Commission of the EC, Directorate-General for Economic and Financial Affairs (DG ECFIN).

[6] Mundell, R. (1961), " A Theory of Optimum Currency Areas", The World Economy, 16, 663-685.

Chapter 1
EMU and Accession Countries: Fuzzy Cluster Analysis of Membership[4]

This chapter estimates the readiness of the Countries of Central and Eastern Europe for the EMU or unilateral euroisation using a fuzzy clustering algorithm. The variables to which the algorithm is applied are suggested alternately by the criteria in the Maastricht Treaty (nominal convergence) and by Optimum Currency Area theory (real convergence). The algorithm reveals that Estonia and Slovenia are the leaders in both nominal and real convergence, whereas the other countries from the 1998 Accession Wave have achieved substantial results only in real convergence. Moreover, Poland is excluded from the leading group in the most recent years due to its worsened economic performance.

KEY WORDS: CEECs, Optimum Currency Area , EMU, Fuzzy Cluster Analysis, Nominal and Real Convergence.

---

[4] This chapter is based on the author's article published in the *International Journal of Finance and Economics* 8, 2003 : 309–325.

1.1 Introduction

The successful accession to membership in the European Union (EU) by the transition economy applicants from the Central and Eastern Europe Countries (CEECs) have thrown up many challenges. One of the main ones is the eastward expansion of the euro area because the new members of the EU are not able to stay outside the European Monetary Union[5], as has been the case with the several countries of the Western and Northern Europe. Nevertheless, entry into the EU, which duly happened in 2004, didn't guarantee immediate acceptance into the monetary union because prior to this the candidates have to demonstrate for two years their ability to satisfy the convergence criteria of the Maastricht Treaty. Therefore, according to the most optimistic estimates, the countries of Central and Eastern Europe could only join the EMU in 2007. However, many economists and especially politicians have even been arguing that the new EU members and candidates should fix their currencies or enter into currency board arrangements based on the European currency or even introduce the euro unilaterally as a means of speeding up the accession and convergence processes (e.g. Nuti, 2001 and Coricelli, 2001). They put forward several reasons why the CEECs should join the EMU at an early date. First, if the CEECs join the EMU they will enjoy lower risk premiums and interest rates, as well as lower transaction costs. They will, moreover, have a say in shaping the ECB's

[5] It was one of the EU-accession criteria specified in Copenhagen in 1993 which explicitly stated that new EU members will have to "... take on the obligations of membership, including (...) the Economic and Monetary Union" meaning that no 'opt-out' provision exists for these countries.

monetary policy, whereas if they decide to stay out they will lose this privilege, although the independence from the ECB will become more imaginary than real once a small country has integrated into the economy of the euro zone. Second, it is often argued that they satisfy the Optimum Currency Area (OCA) criteria and therefore it is beneficial for them to join. Next, given the likely insistence of the EU members on adopting measures to limit the exchange rate variability of the new members there will be no alternatives for them but to fix the exchange rates within some band (an arrangement which could prove fragile and prone to crisis) or to enter the Estonian- or Bulgarian-style currency board (which is the second-best solution in respect to forming a monetary union). Moreover, the incumbent EU members might not be able to do much to keep the aspirants out (Eichengreen and Ghironi, 2001). In the light of these arguments, the question of the CEECs' readiness to join the EMU becomes even more important. Two main issues have to be addressed. First, there is the necessity of meeting the Maastricht Treaty criteria in order to qualify. These criteria limit the ability of the candidates to exercise monetary and fiscal policies at their discretion, which clearly represents a cost to be incurred by the countries on their way to the Euro. Second, abdication of sovereign monetary policy has its own costs and benefits, which have usually been assessed in the framework of the OCA theory (see Mundell, 1961), which advocated forming a monetary union if the adjustment of the bilateral exchange rate is either ineffective or unnecessary to stabilise output. As concerns the process of joining the EU three groups of transition countries may be

identified. The Czech Republic, Estonia, Hungary, Poland and Slovenia started negotiations first, and constitute what was called the 1998 Accession Group, which is argued to have made substantial progress towards satisfying the entry requirements[6] The other group, called the 2000 Accession Group, consisted of Bulgaria, Latvia, Lithuania, Romania and the Slovak Republic, which had not yet advanced as far as the first group in the negotiation process. The rest were countries such as Croatia, which for various reasons are not yet part of the negotiation process. Given this segregation, a natural question to ask is whether a similar division applies to the issue of joining the EMU. The subsequent analysis of this paper thus endeavours to identify a group of countries which are "more EMU-ready" or better suited to enter into a currency board against the Euro and whether these countries are from the 1998 Accession Group or have already implemented a currency board arrangement. In order to check for the existence of homogeneous groups the technique of cluster analysis is employed. This technique is used to examine the similarities and dissimilarities of economic structure in the data and to group the countries according to various sets of criteria. Given the problem of incomplete and noisy data, the more powerful technique of fuzzy clustering is employed. This method splits the data into groups by assigning membership coefficients indicating the degree of "belongingness" of each object to each of the groups, so that the highest coefficient would then indicate the group to which this country is most likely to belong. The accompanying statistics in-

[6] As the paper focuses on the transition countries, Malta and Cyprus are omitted from the analysis that follows.

dicates the existence of the clear-cut structure in the data. The first section briefly describes the algorithm of fuzzy clustering, clarifying its use for the problem at hand as well as the associated diagnostic statistics used in the paper. In the second section we look at the readiness of the applicants for the EMU from an institutional point of view according to their performance with respect to the Maastricht criteria. This section attempts to answer the question of the countries's ability to adopt the euro. The following section looks at the real convergence of the CEECs to the EU and to Germany in particular. This section therefore looks at the question of the desirability of their joining the EMU. The penultimate section compares the results of nominal vs. real convergence. The last section concludes.

1.2 Fuzzy clustering analysis

Cluster analysis is a well-known technique in the science of pattern recognition and is frequently applied in disciplines such as medicine, archeology etc., although its use in applied economic analysis is rather rare. In this paper fuzzy clustering analysis is used, which, unlike the hard clustering algorithms that assign each object to only one subgroup, is much better equipped to analyze the data where some ambiguity is present. The method is applied to uncover the similarities of economic structure in the data across countries and to identify homogeneous subgroups of countries with regard to sets of economic criteria.

The algorithm of fuzzy clustering is taken from Kaufman and Rousseuw (1990) and can briefly be described as follows. The dataset consists of n objects (countries) with p variables (various criteria used in our analysis) for each object and is denoted by $X_{np} = \{x_1, x_2, ..., x_n\}$, where each $x_i = \{x_{i1}, ..., x_{ip}\}$. Each variable is standardised with mean zero and standard deviation of one in order to treat them as having equal importance in determining the structure[7]. The dissimilarity coefficient between two objects is defined as a Euclidean distance[8]:

$$d(i,j) = \sqrt{\sum_{k=1}^{p}(x_{ki} - x_{kj})^2}$$

The algorithm minimizes the objective function C:

$$C = \sum_{v=1}^{k} \frac{\sum_{i,j=1}^{n} u_{iv}^2 u_{jv}^2 d(i,j)}{2\sum_{j=1}^{n} u_{jv}^2}$$

subject to:

[7] In some cases, the standardisation of the variables is important to keep a variable with high variance from dominating the cluster analysis. It is also needed in cases where the variables are of different magnitude and are not directly comparable (e.g. budget deficit and government debt level, the latter always being much higher).

[8] This is the special case of the Minkowski distance metric with argument equal to 2. There are several distance measures for continuous data that may be used such as other Minkowski distance metrics, the Canberra distance measure, the correlation coefficient similarity measure and some others.

$$u_{iv} \geq 0 \; for \; i = 1, ..., n; \; v = 1, ..., n$$

$$\sum_{v} u_{iv} = 0 \; for \; i = 1, ..., n$$

in which u_{iv} represents the unknown coefficient of membership of object i to cluster v, and k represents the exogenous number of clusters into which the data is partitioned. The algorithm produces the matrix of coefficients U_{nxk} with rows summing to one and showing the degree of belongingness of that object to each of the groups. If one of the coefficients is very high then it can be said that there is a high degree of certainty that this object belongs to that group, otherwise this object cannot be classified that easily.

In order to analyze how well the data is partitioned for a given number of clusters, several statistics are used. One is the normalized Dunn's partition coefficient:

$$F_k = \frac{\frac{k}{n} * \sum_{i=1}^{n} \sum_{v=1}^{k} u_{iv}^2 - 1}{k - 1}$$

which varies from 1 (indicating well-partitioned data) to 0 (indicating complete fuzziness of the data). It reaches one only if for each object there is one coefficient equal to one and the others to zero and zero when all the coefficients of belongingness are $\frac{1}{k}$.

Another useful statistic are the silhouette width for each object and average silhouette width for each cluster and for total dataset. Silhouette width for each object is defined as:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

where $a(i)$ is defined as average dissimilarity of object i to all objects in the same cluster (calculated as an average of all $d(i,j)$ for a given cluster) and $b(i)$ as the minimum across all other clusters of average dissimilarity of object i to all objects in each cluster. When $s(i)$ is close to one it indicates that the object is well classified. A value near zero indicates the ambiguity in deciding to which cluster the object might belong. Negative values indicate that the object is misclassified. The corresponding averages for each cluster and for the total dataset indicate how well each cluster's and the total dataset's partitioning has been done.

1.3 EMU and Maastricht Criteria

The Maastricht Treaty specifies a set of criteria to be fulfilled by countries aspiring to participate in the EMU. Their declared aim is convergence in both nominal and fiscal terms ensuring that monetary and fiscal policy converged in order not to disrupt functioning of the EMU in the future[9]. In formal terms, the criteria for nominal convergence say that a country must have an inflation rate within 1.5% of the average

[9] When supplemented by the Growth and Stability Pact.

inflation rate of the three members with the lowest inflation rates and a long-run bond yield within 2% of the average of the bond yields of the same three countries. Furthermore, the Treaty requires that the exchange rate must have been stable within the ± 15% ERM bounds for at least two years. As regards fiscal policy, the budget deficit should be no higher than 3% of the GDP and public debt less than 60% of the GDP.

The same set of qualifications will be applied to any future applicant. Although the earliest date for the new EU members to enter the EMU is estimated to be the year 2006 and criteria are to be complied with only for a year before admission, it is nevertheless useful to see whether the new EU members represent a uniform group with respect to stability orientation. First, it might indicate how easy it will be for the applicants to comply in the future with the provisions of the Stability and Growth pact, and second, it might show whether the countries obey the criteria when conducting their macroeconomic policies in order to show their commitment to the accession process[10].

Given that the criteria were criticised for focusing on the short one-year period of assessment before qualification[11] data for longer time periods is used here[12].

---

[10] As was pointed out by an external reader, the partition of the countries according to Maastricht criteria might not be useful at all as these criteria are very precise and missing one theoretically disqualifies a country. Nevertheless, we argue such a classification is still useful, as it shows the stage at which various countries are at the moment, in respect to obeying all the criteria. Indeed, a country with only a higher inflation level than required by Maastricht criteria has a better chance of meeting all of them in the future than a country with several limits breached.

[11] Two years for the exchange rate stability criterion.

[12] We split the data into three overlapping time periods of 1993-2001, 1997-2001, and 2001 and used the averages over the corresponding periods. Thus, it might be argued, a clearer picture of true

Table 1 shows the corresponding values for new EU members and Croatia as well as an average for 12 EMU countries. The casual inspection of the data reveals several things. First, most of the countries tried to keep their budget deficits low, which proved to be a hard task. During the last eight years five out of eleven countries in the sample had an average budget deficit lower than the three per cent requirement. In recent years the budget deficit has diminished in Bulgaria and Hungary but has increased in Croatia, the Czech Republic, Poland and Slovakia. Second, the debt levels are comfortably below the 60% criterion except in Bulgaria and Hungary (and the EU average itself). Third, volatility of exchange rates (as measured by the standard deviation of the log difference in bilateral exchange rates against the German mark, is low for countries which fixed their currencies against the DM. By the year 2001 it had reduced substantially for almost all countries with the notable exception of Poland. Next, inflation rates have dropped below ten per cent except in Romania. This has had an effect on the lending rates[13] although the difference between Polish lending rates and inflation is above ten per cent, indicating the commitment of the Central Bank of Poland to reduce inflationary expectations brought about by the recent inflation increase [14] .

"stability orientation" of the economy might be obtained and any progress in the development towards stability might be more evident.

[13] This assumption is made because of the data unavailability for the CEECs. The European Commission in its regular reports on countries' progress towards accession look at the lending rates of over one year when assessing the countries' performance, therefore we are using these rates as proxies for the yields on the long-term government bonds.

[14] We are grateful to Ryszard Kokoszczynsky for his remarks on this issue

We run the algorithm for three subsamples and in each case the optimal number of clusters was chosen by maximising the average silhouette width of the dataset (Table 2 reports only the best partitioning for each period). Dunn's coefficient is around 0.5, indicating the presence of some fuzziness in the data, and the average silhouette widths, showing the extent to which the groups formed are different from each other, are higher than 0.5 which is a sign that the structure is present in the data (Kaufman and Rousseuw, 1990).

For the sample of 1993-2001 the optimal number of groups is two - one comprising Bulgaria and Croatia whilst the other countries form the other group. This should come as no surprise because it has been quite a turbulent period for the transition countries and most of them have had to stabilise and restructure their economies which has had an effect on their economic and monetary performance. During that period Bulgaria and Croatia are characterised by extremely high levels of exchange rate volatility, inflation and interest rates compared to the other countries in the sample; therefore they were identified as a distinct group. For the rest of the countries no further conclusions can be made for this sample and, therefore, it is instructive to look at more recent periods.

During the period of 1998-2001 we observe several noticeable changes. The statistics indicate that the data is best partitioned into four groups. The first group is comprised of Estonia, Slovenia and Latvia and is characterised by low values of all criteria except for exchange rate volatility which varies from very low in Estonia

to high in Latvia. Apart from the latter criterion and high inflation rates, this group performs in line with the EU average. The second group consists of the EU average and Bulgaria which have been put together primarily due to the very low exchange rate volatility, low budget deficit and high level of public debt, which is above the 60 per cent limit. Disregarding the public debt criterion these two groups can be treated as one group, that is those Accession Countries which have performed in line with the EMU members according to stability orientation criteria. Interestingly, two of the three CEECs (Bulgaria and Estonia) that officially entered into currency board arrangements are in this group. On the other hand, Romania is a distinct outlier with very high values for all criteria except for the public debt and therefore it has been classified as a singleton (i.e. a group consisting only of one member). The rest of the countries were grouped together because they had high level of budget deficits, an average level of public debt, average to high exchange rate volatility but mixed results for inflation and interest rates.

Given that the Maastricht criteria are to be applied to assess the performance of would-be members one year before entry, it is, therefore, useful to look at the latest data and to see what the current economic and financial situation is. With that in mind we ran the algorithm for the data of year 2001 alone[15]. This time the best partition consists of five groups, although many regularities from the previous subsample are still present. Estonia and Slovenia again form the group with low values for all

[15] Subsample of 2000 - 2001 is used to calculate the exchange rate volatility.

criteria except for the interest rates. Considering the fact that we use lending rates instead of government bond yields as specified by the Maastricht Treaty, this group may be regarded as the best performing one. Latvia and Lithuania form the other group, which follows Estonia and Slovenia closely, although they have higher exchange rate volatility. As in the previous period, Bulgaria and the average EU member form the third group because of the high debt level, although the inflation rate in Bulgaria is too high by EMU standards. Allowing for some flexibility in interpreting the Maastricht criteria it may be argued that the countries from these three groups are the best performers and by now have managed to bring the government finances and domestic monetary situation under control. Again, the interesting fact is that this time all three countries which implemented the currency board arrangements are included[16]. Romania on its own forms another group again because of grossly breaching all the criteria and the rest of the countries constitute the last group, which is characterised by high budget deficit and average to high values for the other criteria.

Looking across all the subsamples the following conclusions can be reached (Table 3 summarises the findings). During the whole sample period of the eight years the countries have shown mixed performance, so that no detailed partitioning can be made except for separating the countries which have undergone some serious crisis during that period. Nevertheless, looking separately at the recent period there

[16] High level of exchange rate volatility of the Lithuanian Lit against the DM may be explained by the fact that it is fixed against the basket of the currencies, with the US dollar in sizeable proportion. As for Bulgaria, its public debt declines constantly each year, which may be regarded as a sign of convergence to the debt criterion limit.

appears to be a clear-cut segmentation among the CEECs. All three Baltic states, Bulgaria and Slovenia seem to make a group of countries, which is well ahead as concerns the stability orientation of the economy and expressed by the Maastricht criteria. An interesting fact is that all three CEECs who implemented the currency board arrangement against the Euro are in this group.

1.4 OCA Criteria and Economic Convergence

1.4.1 OCA Criteria Explained

It is often argued in the literature that although in the nineties the EU countries were converging in nominal terms as was manifested in general compliance with the Maastricht Criteria, real convergence was far from being achieved and one could even have pointed out to real divergence between some countries. As a consequence, the countries whose initial conditions are unfavourable and which are unable to use a national monetary policy to adjust to specific shocks will find themselves on low growth and high unemployment paths. As pointed out by Bayoumi and Eichengreen (1997b) and others, the Maastricht criteria do not ensure the real convergence which is required for the successful functioning of any monetary union. This idea of real convergence was first put forward by Mundell (1961) and later revived by Krugman (1990). Krugman developed the foundations of the OCA theory, which stated that two countries should form a monetary union in the case of prevalence of a high de-

gree of intra-trade among the members and the absence of any profound asymmetry in the pattern of shocks impacting their economies.

As OCA theory states, there are certain benefits and costs associated with adopting a single currency that depend on the degree of convergence of the economies. The benefits are associated with economising on exchange costs and with importing the credibility of the union's central bank, thus reducing the inflationary expectations and level of inflation. This point is illustrated by the example of Bulgaria, which entered the currency board arrangement in order to combat inflation and stabilise its economy. Another clear case is Estonia where inflation was substantially lower than in the other CEECs. As for the associated costs they are essentially the opposite of the benefits of having an independent monetary policy and exchange rate, which are useful as a means of coping with shocks that are asymmetric between the potential monetary union partners. The less effective the monetary policy is in counteracting the idiosyncratic shocks by adjusting the nominal exchange rate, the lower the costs. Other domestic conditions such as sufficient labour mobility or fiscal federalism also reduce the need for independent monetary policy.

The OCA criteria are a useful benchmark for evaluating the costs and benefits of any exchange rate arrangement. First, the qualitative analysis of the costs and benefits and comparative studies can be conducted. One of the examples for the European countries is by De Grauwe and Yunus (1999). On the sample of CEECs there are several papers by Boone and Maurel (1998 and 1999) and Habib (2000) as well

as by Fidrmuc and Schardax (2000). Second, the OCA theory was rendered operational through cross-country estimations of the effect on the variability of the bilateral exchange rates by the asymmetry of the business cycles and other explanatory variables. This was first done by Bayoumi and Eichengreen (1997a) for industrialised countries and later adopted to CEECs by Bénassy-Quéré and Lahrèche-Révil (1998).

Notwithstanding the popularity of the approach, recently there has been growing criticism of the classical OCA literature. Two basic points have been made. First, the OCA literature has allegedly failed to consider the dynamic and endogenous nature of the criteria because economists have often applied OCA criteria as if they were taking a snapshot of a motionless object. However, these characteristics could react to the very policy decision to fix the exchange rate, adopt another country's currency or join a currency union. In other words, the OCA literature does not take into account the Lucas Critique and considers the several criteria as exogenous parameters. Frankel and Rose (1998) claimed that the OCA criteria are in fact endogenous and found that greater integration resulted in more highly synchronised business cycles. According to this result, a country that does not satisfy the OCA criteria could join a currency union eliminating exchange risk and transaction costs. Reduced costs would foster trade integration, which, in turn, would increase the correlation of business cycles. Hence, the endogeneity of OCA criteria poses some limitations to a static application of the theory. Second, the OCA literature has not paid enough attention to the increased role of international financial markets and capital mobility.

These limitations contribute even more to the already complicated cost-benefit analysis of a common currency. However, for the purpose of the paper they have little relevance. Here we are more concerned with identification of the homogeneous groups among CEECs, so the analysis will indicate if there is a group of countries whose current nominal and real convergence with the EMU is at a higher stage. If the criteria are endogenous then these countries will have some competitive advantage over the other applicants and the likely structural changes and catch-up processes will be less dramatic.

1.4.2 Empirical results for economic convergence

Choice of variables

The choice of variables to analyze the economic convergence of CEECs was inspired by the OCA criteria following the work of Artis and Zhang (1998). For the sample of ten accession countries we collected the monthly and annual data (see Table 4) starting from 1993 from various sources which are described in the Appendix.

1. *Synchronisation of business cycles*

The popular choice to analyze the symmetry of output shocks is to study the cross-correlation of the cyclical components of output (e.g. Artis and Zhang, 1998). Due to the data unavailability of quarterly GDP growth rates[17] we decided to fol-

[17] Romanian National Statistical Office does not produce quarterly GDP estimates at all; Bulgaria has started to publish them only in 2000. For the other countries the qurterly data from 1993 would

low the approach of Artis and Zhang (1998), who identified the symmetry of output shocks with the cross-correlation of the cyclical components of monthly industrial production series[18]. Whereas the aggregate GDP estimates for the eurozone are available[19] this is not so for the industrial production data, and therefore for the purpose of the estimation the Germany monthly industrial production index was taken. The choice was justified on the grounds of the existence of what is called the "European business cycle" (see Artis and Zhang, 1995), and it is confirmed when we look at Figure 1, which shows the quarterly GDP and industrial production growth rates for the Eurozone, Germany and Estonia[20].

Given the close comovement of the series we decided to use German industrial production index as a proxy for EMU output. In the light of the heated debate as to what type of filtering is more appropriate we use two filtering techniques. First, the industrial production series were seasonally adjusted and detrended using the Hodrich-Prescott (H-P) filter with the value of the dampening parameter equal to 50,000[21]. Second, as an alternative, we used the twelfth differences of the logs of the series (i.e. the growth rate of each month relative to the same month of the previous

give only 30 observations.

[18] As pointed out by Ryszard Kokoszczynski the share of the industrial production in the GDP of the CEECs is rather low and therefore it can be erroneous to omit the agricultural production from our analysis. We fully agree with that comment and can only justify our choice by data availability.

[19] For example, from Beyer and Hendry (2001)

[20] We was unable to find the monthly industrial production index for Estonia and Bulgaria and therefore used monthly unemployment rate for Bulgaria and quarterly GDP growh rate for Estonia for which monthly unemployment rate is not available either.

[21] See Artis and Zhang (1998) for the discussion of the choice.

year). Both methods produced similar results although slightly higher values in the former case, which are used in the subsequent analysis. The cross-correlations *vis-'a-vis* Germany were calculated for the whole sample and subsamples. Figure 2 illustrates that the correlation between CEECs and German business cycles has grown considerably and has a tendency to converge to a very close range for all countries. However, the increased divergence after the beginning of the year 2001 merits special attention.

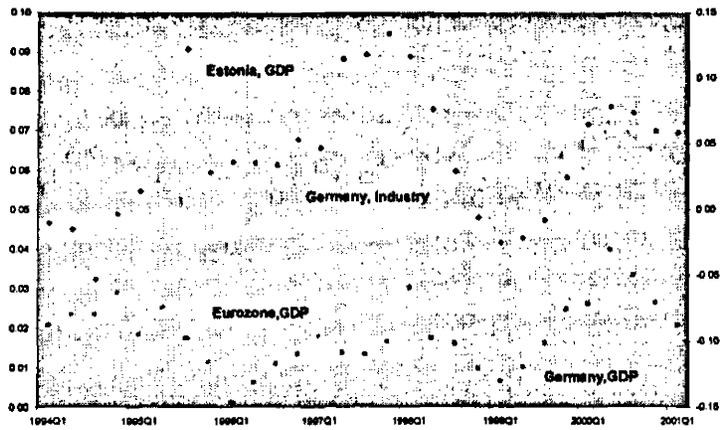


Figure 1.1: Figure 1. The quarter-on-quarter growth series of German industrial output, Estonian GDP (all right axis), Eurozone GDP and Germany GDP (left axis).

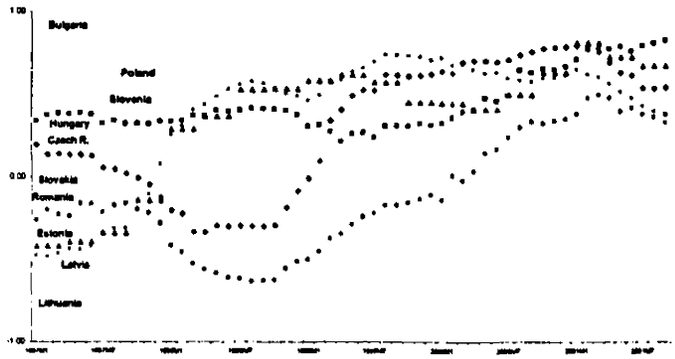


Figure 2. Time - varying correlations of industrial production (CEECs vis-'a-vis Germany over previous three years). Unemployment levels correlation for Bulgaria and GDP growth rates correlation for Estonia.

2. *Volatility of the real exchange rate (RER)*

According to the OCA criteria, the costs of monetary union are associated with the loss of a separate exchange rate. By influencing the nominal exchange rate the monetary policy presumably changes the real exchange rate which acts as a shock absorber. If there has been little cause for variation in the exchange rate then not much will be if single currency is adopted. In this study we represent the variation in the exchange rates as the standard deviation of the log-difference of real DM exchange rate, where deflation is accomplished using the relative wholesale price index.

As was pointed out by an external reader, theoretically high real exchange-rate volatility should be related to high occurrence of asymmetric shocks, but, in practice, this effect may be offset by dependence on a specific exchange rate regime. Therefore, it would have been more appropriate to choose a more structural measure of asymmetric shocks such as an index of sector specialisation. We agree on

this point but were forced to use the exchange rate volatility proxy due to problems with obtaining needed monthly data, but conducted robustness analysis by removing exhange-rate volatility from the analysis.

3. *Openness to trade*

This criterion is assumed to be represented by trade intensity between EMU members as a whole and each CEEC, i.e. for any country i as $(x_{iEMU}+m_{iEMU})/(x_i+m_i)$, which is the ratio of exports and imports to EMU members over total imports and exports of country i.

4. *Inflation criteria*

The recent addition to the classical OCA theory is that " a strong incentive for monetary union is created by an assurance that the union's inflation will be low" (Artis and Zhang, 1998). This criterion is measured by the annual inflation differential of each CEEC against average EMU inflation.

Estimation results

The data was split into four overlapping periods: 1993 -2001, 1995 -2001, 1997 - 2001 and 1999-2001. This was done in order to see whether the results for the total sample were influenced by the processes of economic restructuring and transition turbulence in the mid-nineties and to look at the recent development in the CEECs progress towards economic convergence to the EMU. Several points about the data, which is reported in Table 5, should be mentioned. One of the most important characteristics, that is of business cycle correlation, increased dramatically towards the end

of the estimation period and showed the tendency to converge to a close range for all countries by the beginning of the year 2001, as illustrated in Figure 2. Although this finding is confirmed by other studies (Boone and Maurel, 1998 and Fidrmuc and Schardax, 2000), the short time period of only one full business cycle and the presence of only a few supply and demand shocks makes it less robust and conclusive then we would like it to be. The high degree of trade with the EMU countries also merits attention, with the Czech Republic, Hungary and Poland already reaching the average of EMU intra-trade level (around 67%). The volatility of the real exchange rate has decreased for countries that used exchange rate arrangements close to the fixed rates, but stayed higher for Poland and Romania for all the sample and quite high for Latvia and Lithuania. The inflation differential was reduced to single digits since 1998 for all the countries except Romania.

Application of the clustering algorithm reveals a substantial level of fuzziness in the data (Dunn's coefficient is around 0.5) and slightly worse results than in the previous section (average silhouette width is around 0.5 for all subsamples).

Thus, the results of the estimation for the whole period of 1993-2001 show the presence of the three groups. The best performing group consists of the Czech Republic, Hungary, Estonia, Poland and Slovenia, which show low volatility of RER (with the exception of Poland and probably the Czech Republic), very high trade openness of above 60%, and relatively high degree of business cycles synchronisation (see Figure 3), although the group statistics for inflation rate is less uniform. The

second group is formed by Latvia, Lithuania and Romania, which have much worse values for all the parameters. The third group consists of Bulgaria and Slovakia.

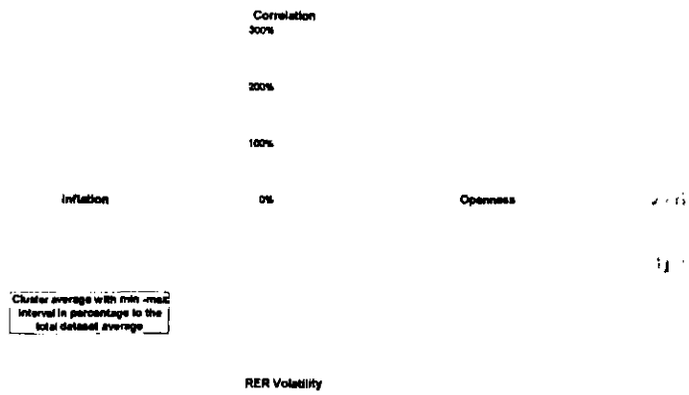


Figure 3. Statistics for the group of best-performers (the Czech Republic, Estonia, Hungary, Poland, Slovenia), 1993-2001.

The results for the 1995 - 2001 period are almost identical to the previous ones. The only exception is that the group of best performing countries is joined by Latvia, the performance of which has improved substantially according to all four criteria.

The analysis of the period 1997-2001 indicates the improved performance of several countries, particularly Slovakia, which is grouped with the best-performers. The algorithm has identified only two groups in this sample - one of the five countries from the first Accession wave along with Latvia and Slovakia and the second group consisting of the other three countries, which are lagging behind. Figure 4 shows the comparative statistics for the first group.

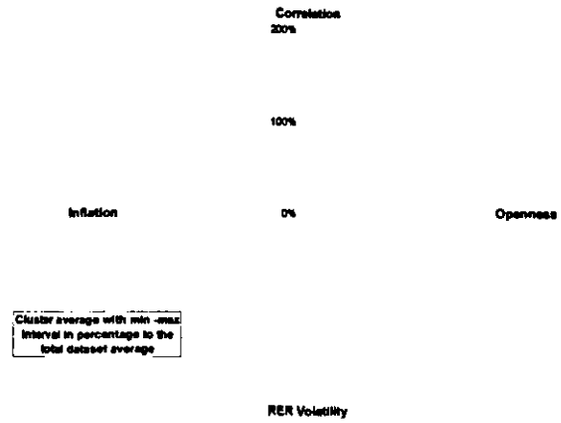


Figure 4. Statistics for the group of the best-performers (the Czech Republic, Estonia, Hungary, Latvia, Poland, Slovakia, Slovenia), 1997-2001.

In order to analyze any changes in the performance of CEECs after five of the countries negotiated an accession status (and were subsequently called the 1998 Accession Group) we look at the subsample of the 1999-2001 data. The results indicate that the best partition for this period was of four groups (see Figure 5), the best-performing group not including Poland this time due to its high RER volatility, which instead is grouped with Slovakia, exhibiting high volatility as well. The number two group consists of Latvia, Lithuania, and Romania, which form a cluster of countries with low business cycles correlation. Bulgaria is identified as a singleton.

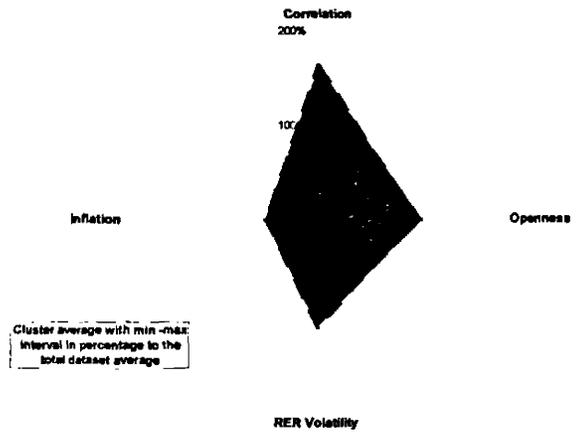


Figure 5. Statistics for the group of best-performers (Czech R., Estonia, Hungary, Slovenia) and Poland separately (the dashed line), 1999 - 2001.

The summary of the findings is as follows (see Table 6). It is possible to identify the group of countries which, by the OCA criteria, is more suited to join the EMU or to euroise. The criteria indicate that this group of countries is the 1998 Accession Group, joined sometimes by Latvia and Slovakia. It seems, therefore, that the progress in liberalising the economy and restructuring as prescribed, monitored and assessed by the European Commission is correlated with successful integration into the EU economic area if the OCA theory is used. Interestingly, the analysis confirmed that Poland in recent years is lagging behind the more successful applicants and its acceptance will become more a political issue unless it speeds up its reforms.

Given a small number of criteria used in the study there was no much scope for conducting robustness analysis of the findings. Nevertheless, we run the analysis for several overlapping periods to observe how classification changes over time and

reported the findings above. Moreover, we omitted real exchange-rate volatility from the analysis and looked at the resulting changes to the groups. The most noticeable difference is reclassification of Estonia from the group of best performers to other two Baltic states' group during the last time periods. Given that Estonia has currency board arrangement against Euro, it might indicate that apart from low exchange rate volatility, the other real convergence indicators are not so good.

1.5 Maastricht and OCA criteria compared

How do our identification of groups based on nominal convergence compare with the one that based on the OCA criteria? If we focus on the latest period, several interesting conclusions may be drawn. First, only Estonia and Slovenia are classified as the countries that achieved both nominal and real convergence with the EMU countries. In both parts of the analysis these two countries are grouped together in all the subsamples and they always form a part of the most advanced group. Therefore, Estonia and Slovenia may be regarded as the main candidates for joining the EMU among the applicants from the CEECs. Other countries from the 1998 Accession Wave (Czech Republic, Hungary, Poland) are leading only in real convergence while lagging behind in terms of inflation achievement, the stance of fiscal policy and exchange rate behaviour. Second, Poland is shown to lag behind the rest of the countries from the 1998 Accession Wave and it does not make a part of the leading group even by the OCA criteria during the last years. Third, all three countries (Bulgaria, Estonia,

Lithuania) that have implemented the currency board arrangements show considerable nominal convergence but only Estonia is leading in economic convergence as well. Fourth, Romania with its high inflation, interest rate, loose fiscal policy and volatile exchange rate is a clear outlier.

1.6 Conclusions

This paper has analysed the empirical evidence on the topic of readiness of CEECs to join the EMU or euroise their economies. The problem was split into two parts and cluster analysis to identify the groups of countries that are closer to being ready to do so was employed in each case.

First, the paper looked at the Maastricht Criteria as a set of requirements to be fulfilled by the applicants in order to qualify. Support was found for the existence of a clear-cut structure in the data. Several countries, among them all the CEECs that have implemented the currency board arrangement, joined by Slovenia and Latvia, consistently outperformed others in coming close to satisfying the Maastricht Criteria. Whether the stability orientation of an economy is improved by fixing its currency against the euro remains a question deserving further attention.

Second, the question of the economic convergence of CEECs to the EU was tackled by analysing their performance with respect to the OCA criteria. It turned out that only Slovenia and Estonia are the leaders both in nominal and real convergence. Additionally, the recent economic and restructuring performance of Poland

is identified as the main reason for associating it with the other group of countries, which are not converging at such a fast rate to the EU economic area.

Table 1: Maastricht Treaty Criteria and Transition Countries

	Deficit[1)]		Debt[1)]		Volatility of ER[2)]		Inflation[3)]		Interest rate[4)]	
	1993-2001	2001	1993-2001	2001	1993-2001	2001	1993-2001	2001	1993-2001	2001
Bulgaria	-3.6	-1.7	113.1	97.5	5.7	0.0	161.5	7.9	49.6	11.1
Czech R.	-2.2	-5.2	25.8	29	0.9	0.5	8.7	4.7	11.3	7.0
Estonia	-0.7	-0.8	8.0	6.1	0.2	0.1	24.6	5.9	15.6	7.7
Hungary	-5.8	-3.7	88.7	64.4	0.9	0.6	17.2	9.6	21.7	12.3
Latvia	-2.5	-2.2	9.3	10.2	1	1.0	23.2	3.0	30.0	11.5
Lithuania	-4.4	-1.4	22.5	25	1.7	1.4	62.6	1.2	29.4	10
Poland	-3.3	-4.3	49.1	42.8	1.2	1.4	18.7	6.0	25.9	19.3
Romania	-4.3	-4	22.6	32.2	2.4	1.3	89.2	34.4	57.7	45.8
Slovak R.	-4.3	-5	31.8	42.7	0.7	0.7	10.6	7.5	16.4	12.2
Slovenia	-1.2	-1.1	21.7	25.5	0.4	0.4	13.1	8.5	23.7	15.2
EU-12	-2.9	-1	71.2	67.4	0.4	0.0	2.4	2.0	9.0	7.9
Croatia	-2.6	-5.3	27.1	38	2.8	0.9	182.1	3.3	175.2	9.6

Source: see Appendix for data description[5)].

Notes:

1) Deficit and debt as % of GDP.

2) Volatility in exchange rate is measured by the standard deviation($x10^2$) of monthly differences.
of the log difference in bilaterial monthly average exchange rate against DM.

3) CPI index.

4) Lending rates of longest maturity are taken for accession countries and the average of the lending rate of France, Italy and Germany for EU-12

5) Data for 1998-2001 are not reported in the table but are available upon request from the author

Table 2. Partitioning by Maastricht Criteria

	1993 - 2001				1998 - 2001						2001						
	Coefficients[1]		Silhouette width[2]	Cluster	Coefficients				Silhouette width	Cluster	Coefficients					Silhouette width	Cluster
Bulgaria	.19	**.81**	0.18	2	.03	**.92**	.01	.04	0.85	2	.03	**.90**	.01	.03	.02	0.81	2
Croatia	.22	**.78**	0.12	2	.09	.04	.02	**.86**	0.66	4	.02	.05	.01	**.92**	.04	0.70	4
Czech Republic	**.98**	.02	0.88	1	.08	.03	.01	**.88**	0.56	4	.07	.05	.02	**.78**	.08	0.62	4
Estonia	**.86**	.14	0.79	1	**.89**	.05	.01	.06	0.77	1	**.85**	.04	.01	.04	.07	0.74	1
Hungary	**.64**	.36	0.58	1	.18	.20	.06	**.56**	0.57	4	.13	.22	.04	**.47**	.14	0.40	4
Latvia	**.95**	.05	0.86	1	**.53**	.07	.03	.37	0.29	1	.08	.02	.01	.05	**.84**	0.71	5
Lithuania	**.84**	.16	0.80	1	.08	.03	.02	**.86**	0.68	4	.05	.02	.01	.04	**.89**	0.82	5
Poland	**.97**	.03	0.87	1	.15	.08	.05	**.72**	0.59	4	.11	.08	.06	**.44**	.31	0.21	4
Romania	**.55**	.45	0.54	1	.00	.00	.00	**1.0**	0.00	3	.00	.00	**1.0**	.00	.00	0.00	3
Slovak Republic	**.97**	.03	0.89	1	.07	.03	.01	**.88**	0.67	4	.01	.01	.00	**.96**	.02	0.79	4
Slovenia	**.88**	.12	0.80	1	**.91**	.04	.01	.04	0.72	1	**.83**	.01	.04	.04	.07	0.62	1
EU-12 Average	**.89**	.11	0.83	1	.08	**.87**	.01	.04	0.71	2	.12	**.75**	.01	.06	.06	0.59	2
Number of clusters			2					4							5		
Silhouettes width[3]			0.78 0.15					0.68 0.78 0.00 0.62						0.68 0.70 0.00 0.54 0.77			
Average silhouette width			0.68					0.64							0.64		
Dunn`s coefficient			0.54					0.63							0.62		

Source: author's calculations

Notes:

1) The coefficients of belonginness of the country to each cluster with the highest in bold.

2) Individual silhouette width

3) Silhouettes widths for each cluster in ascending order

Table 3. Classification of the countries by nominal convergence[1)]

	ER volatility	Budget deficit	Public debt	Inflation	Interest rates
1993 - 2001					
{Bulgaria, Croatia}	High	Ave	Mixed	High	High
{Other countries}	Mixed	Mixed	Mixed	Mixed	Mixed
1998 - 2001					
{Estonia, Latvia,Slovenia}	Mixed	Low	Low	Low	Low
{Bulgaria, EU-12}	Low	Low	High	Mixed	Low - Ave
{Croatia, Czech R.,Hungary, Lithuania, Poland, Slovakia }	Ave - High	High	Ave	Mixed	Mixed
{Romania}	High	High	Ave	High	High
2001					
{Latvia, Lithuania}	Ave - High	Low	Low	Low	Low
{Estonia, Slovenia}	Low	Low	Low	Low	Mixed
{Bulgaria, EU-12}	Low	Low	High	Mixed	Mixed
{Croatia, Czech R., Hungary, Poland, Slovak R.}	Ave - High	High	Ave - High	Ave	Ave
{Romania}	High	High	Ave	High	High

Source: author's calculations

Notes:

1) Table shows groups' classification according to whether the countries in each group have low (Low), average (Ave) or high (High) values for each criteria. If the countries in a group have a large dispersion of values for a criterion, then Mixed is reported

Table 4. OCA criteria and economic convergence

	Correlation in business cycles[1)]			Exchange rate volatility[2)], ($x10^2$)			Trade openness[3)]			Inflation differential [4)], %		
	1993-2001	1997-2001	1999-2001	1993-2001	1997-2001	1999-2001	1993-2000	1997-2000	1999-2000	1993-2001	1997-2001	1999-2001
Bulgaria	0.12	0.34	-0.32	4.45	3.53	0.48	0.45	0.48	0.50	159.1	218.2	5.0
Czech R.	0.20	0.50	0.52	0.97	1.16	0.71	0.59	0.63	0.66	6.3	4.0	1.7
Estonia	0.15	0.38	0.64	0.60	0.25	0.20	0.59	0.58	0.59	22.2	4.5	2.5
Hungary	0.52	0.60	0.84	0.85	0.68	0.62	0.64	0.69	0.70	14.8	10.4	7.9
Latvia	0.14	0.41	0.29	1.38	0.83	0.90	0.47	0.53	0.52	20.7	2.3	0.8
Lithuania	-0.32	0.04	0.05	2.00	1.42	1.53	0.40	0.44	0.46	60.2	1.5	-0.9
Poland	0.39	0.59	0.68	1.10	1.18	1.18	0.66	0.66	0.67	16.3	8.1	5.9
Romania	-0.12	0.06	0.33	3.04	3.54	1.48	0.55	0.60	0.63	86.8	66.0	40.0
Slovak R.	0.30	0.53	0.58	0.92	0.79	0.79	0.43	0.52	0.54	8.1	6.7	8.1
Slovenia	0.49	0.45	0.56	0.49	0.44	0.45	0.66	0.67	0.67	10.6	6.8	6.7

Source : see Appendix for data description[5)]

Notes:

1) Measured as cross-correlation of differenced monthly industrial production indices with German one. For Bulgaria cross-correlation of differenced monthly unemployment rate with German unemployment is taken. For Estonia cross-correlation of quarterly GDP growth with German GDP growth is taken.

2) Standard deviation of log difference in bilateral real exchange rate against DM

3) Average for the period of the ratio of import and export to the EU over total imports and exports

4) Average for the period of the CPI indices less the EU-15 average inflation over the same period

5) Data for 1995 - 2001 is not reported in the table but available from the author upon request

Table 5. Partitioning by OCA Criteria

	1993 - 2001					1995 - 2001					1997 - 2001			1999 - 2001					
	Coefficients[1]			Silhouette width[2]	Cluster	Coefficients			Silhouette width	Cluster	Coefficients		Silhouette width	Cluster	Coefficients				Silhouette width	Cluster
Bulgaria	.01	.03	**.96**	0.13	3	.16	.38	**.46**	0.22	3	.21	**.79**	0.23	2	.00	.00	.00	**.99**	0	4
Czech R.	**.53**	.39	.08	0.29	1	**.50**	.32	.18	0.31	1	**.93**	.07	0.75	1	**.91**	.03	.06	.01	0.67	1
Estonia	**.65**	.27	.08	0.52	1	**.59**	.09	.32	0.68	1	**.81**	.19	0.72	1	**.44**	.14	.31	.12	0.23	1
Hungary	**.94**	.04	.02	0.81	1	**.91**	.04	.06	0.79	1	**.89**	.11	0.76	1	**.78**	.06	.12	.03	0.69	1
Latvia	.11	**.83**	.06	0.25	2	**.40**	.25	.35	0.39	1	**.71**	.29	0.57	1	.12	**.70**	.11	.07	0.35	2
Lithuania	.11	**.73**	.16	0.67	2	.11	**.68**	.22	0.42	2	.18	**.82**	0.27	2	.07	**.73**	.10	.11	0.49	2
Poland	**.70**	.16	.14	0.54	1	**.48**	.14	.39	0.44	1	**.85**	.15	0.75	1	.08	.05	**.85**	.02	0.14	3
Romania	.13	**.68**	.19	0.58	2	.08	**.82**	.10	0.47	2	.29	**.71**	0.29	2	.23	**.43**	.27	.07	0.33	2
Slovak R.	.32	.34	**.34**	-0.08	3	.06	.06	**.88**	-0.07	3	**.57**	.43	0.47	1	.15	.18	**.51**	.16	0.47	3
Slovenia	**.92**	.05	.02	0.81	1	**.94**	.03	.04	0.81	1	**.94**	.06	0.78	1	**.95**	.02	.03	.01	0.79	1
Number of clusters	3					3					2				4					
Silhouettes width[3]	0.59 0.50 0.02					0.57 0.44 0.08					0.69 0.26				0.60 0.39 0.31 0.00					
Average silhouette width	0.45					0.45					0.56				0.46					
Dunn`s coefficient	0.44					0.35					0.41				0.48					

Source: author's calculations

Notes:

1) The coefficients of belonginness of the country to each cluster with the highest in bold.

2) Individual silhouette width

3) Silhouettes widths for each cluster in ascending order

Table 6. Classification of the countries by OCA criteria[1)]

	Business cycles correlation	Real exchange rate volatility	Trade openness	Inflation differential
1993 - 2001				
{ Czech R., Estonia, Hungary, Poland, Slovenia}	High	Low - Ave	High	Low - Ave
{Latvia, Lithuania, Romania}	Low	High	Ave	High
{Bulgaria, Slovakia}	Ave	High	Low	Mixed
1995 - 2001				
{ Czech R., Estonia, Hungary, Latvia, Poland, Slovenia}	High	Low - Ave	High	Low - Ave
{Lithuania, Romania}	Low	High	Low - Ave	Mixed
{Bulgaria, Slovakia}	Low	Mixed	Low	Mixed
1997 - 2001				
{ Czech R., Estonia, Hungary, Latvia, Poland, Slovakia, Slovenia}	High	Low - Ave	High	Low - Ave
{Bulgaria, Lithuania, Romania}	Low	High	Low - Ave	Mixed
1999 - 2001				
{ Czech R., Estonia, Hungary, Slovenia}	High	Low	High	Low - Ave
{Poland, Slovakia}	High	High	Ave - High	Ave
{Latvia, Lithuania, Romania}	Low	High	Mixed	Mixed
{Bulgaria}	Low	Low	Low	Ave

Source: author's calculations

Notes:

1) Table shows groups' classification according to whether the countries in each group have low (Low), average (Ave) or high (High) values for each criteria. If the countries in a group have a large dispersion of values for a criterion, then Mixed is reported

References

[1] Artis, M. and Zhang, W. (1998), "Membership of EMU: A Fuzzy Clustering Analysis of Alternative Criteria", IUE Working Paper 98/52.

[2] Artis, M. and Zhang, W. (1995), " International Business Cycles and the ERM: Is there a European Business Cycle?", International Journal of Finance and Economics, Vol. 2., pp. 1-16.

[3] Bayoumi, T. and Eichengreen B. (1997a), "Ever Closer to Heaven? An Optimum-Currency-Area Index for European Countries", European Economic Review, vol. 41, pp. 761-770.

[4] Bayoumi, T. and Eichengreen B. (1997b), "Optimum Currency Areas and Exchange Rate Volatility; Theory and Evidence Compared", in (eds) Cohen, B., International Trade and Finance: New Frontiers for Research: Essays in Honour of Peter Kenen, Cambridge.

[5] Bénassy-Quéré, A. and A. C. Lahrèche-Révil (1998), "The Euro as a Monetary Anchor in the CEECs; in: Pegging the CEEC's Currencies to the Euro", CEPII Working Paper No. 98-04, August, Paris.

[6] Beyer, A., Doornik, J.A., and Hendry, D.F. (2001), "Constructing Historical Euro-Zone Data". Economic Journal, 111, 308-327.

[7] Boone, L. and Mathilde, M. (1998), "Economic Convergence of the CEECs with the EU". Discussion Paper 2118. CEPR, London.

[8] Boone, L. and Mathilde, M. (1999), "An Optimum Currency Area Perspective of the EU Enlargement to the CEECs". Discussion Paper 2119. CEPR, London.

[9] Brada, J., Kutan, A. (1999), "The Convergence of Monetary Policy Between Candidates Countries and the European Union", Working Paper of the Centre for European Integration Studies.

[10] Coricelli, F. (2001), "Exchange Rate Arrangements in the Transition to EMU: Some Argument in Favour of an Early Adoption of the Euro", paper presented at the conference " When is the National Currency a Luxury", London, 16 -17 March 2001.

[11] De Grauwe, P. and Aksoy, Y. (1999), "Are Central European Countries Part of the European Optimum Currency Area", In: Paul De Grauwe and Vladimir Lavrac (eds.) Inclusion of Central European Countries in the European Monetary Union. Dordrecht: Kluwer : 1336.

[12] Eichengreen, B., Ghironi, F. (2001), "EMU and Enlargement", paper presented at the conference on Economic and Monetary Union, Brussels, 21-22 March 2001.

[13] European Bank for Reconstruction and Development (2000), Transition Report.

[14] Fidrmuc, J., Schardax, F. (2000), "More Pre-Ins Ante Portas? Euro Area Enlargement, Optimum Currency Area, and Nominal Convergence," in: Österreichische Nationalbank, Focus on Transition, 2/2000, pp. 28-47.

[15] Frankel, J. and Rose, A. (1998), "The Endogeneity of the Optimum Currency Area Criteria", Economic Journal 108, (July): 1009-1025.

[16] Habib, M. (2000), "The Case for Euroisation in Central and Eastern European Countries", University College London, mimeo.

[17] Habib, M. (2001), "Is There a Case For Unilateral Euroisation? Evidence From Bulgaria, Hungary, The Czech Republic and Poland", Centre For New and Emerging Markets, Discussion Paper 19.

[18] International Monetary Fund, (1996). World Economic Outlook, Washington DC, October.

[19] Kaufmann, L. and Rousseuw. P. (1990), Finding Groups in Data - An Introduction to Cluster Analysis. Wiley series in Probability and Mathematical Statistics.

[20] Krugman, P. (1990), "Policy Problems of a Monetary Union", in P. De Grauwe and L. Papademos (eds) "The European Monetary System in the 1990's", Harlow: Longman.

[21] Mundell, R. (1961), " A Theory of Optimum Currency Areas", The World Economy, 16, 663-685.

[22] Nuti, M. (2001), "Costs and Benefits of Unilateral Euroisation in Central Eastern Europe", paper presented at RSCAS Conference "Currency Inde-

pendence in a World of Interdependent Economies", European University Institute, Florence, 15 - 16 February 2002.

2.A Data sources and description

2.0.1 *Maastricht criteria data*

Budget Deficit:

CEECs' data is taken from EBRD (2000) and Deutsche Bank Research;

EMU countries' data from Eurostat (2001).

Public Debt:

Croatia - from WDI database (World Bank);

Lithuania - NSO of Lithuania; the rest of CEECs' from Deutsche

Bank Research.

Exchange rates:

all from IFS (IMF).

Inflation:

CEECs' data - EBRD (2000), Deutsche Bank Research and

national statistics;

EMU average inflation is taken from OECD Economic Outlook (2001).

Interest rates:

Slovenia - Slovenian Central Bank; rest of CEECs from EBRD (2000);

EMU interest rates from Eurostat(2000).

2.0.2 *OCA criteria data*

Business cycles correlation:

Bulgarian unemployment, Polish, Slovak Republic, Slovenian industrial output from PlanEcon Monthly Report (various issues); the rest of CEECs' data and Germany industrial output and unemployment from IFS (IMF) and NSO of Estonia; Eurozone GDP from Beyer *et al.* (2001).

Real Exchange Rates:

all from IFS (IMF).

Trade openness:

all from European Commission Statistics.

Inflation differential:

CEECs' data - EBRD (2000); EMU average inflation is taken from OECD Economic Outlook (2001).

Unemployment:

all from EBRD (2000) and National Statistics offices.

Chapter 2
Estimating Equilibrium Exchange Rates of OECD Countries

This chapter presents a thorough statistical analysis of the real and nominal equilibrium exchange rates for sample of 19 OECD countries, including 12 EMU members. The analysis proceeds from well-established proposition that persistent deviations of exchange rates from PPP levels may be explained by the economic processes undergoing in the economy and influencing equilibrium exchange rates. On this foundation a set of fundamental-based models of real exchange rate determination was tested on a country-by-country basis. Using quarterly data from 1974 to 2003, the research focuses on so-called behavioural and permanent equilibrium exchange rate (BEER/PEER) approaches, devoting special attention to proper econometric model construction and testing. The estimation results are then used to predict both nominal and real exchange rate misalignments are estimated for each currency and for the "synthetic" Euro.

KEY WORDS: Real Exchange rates, Equilibrium Exchange Rates, Cointegration Analysis.

2.1 Introduction

Fluctuations of exchange rates in both developing and industrialised countries over the last 30 years, has stimulated considerable research activity dedicated to establishing whether exchange rates are driven by some quantifiable fundamental forces, or merely reflect market expectations, which are sometimes very volatile, and sometimes simply irrational. As the exchange rate is a crucial variable linking a nation's domestic economy to international markets, the choice of an exchange rate regime, and of its target, is a central component of economic policy in developing and transition countries, and a key factor affecting economic growth. Not surprisingly, estimation of the correct or equilibrium exchange rate has been important both for politicians, and for national monetary policy institutions. A number of factors have recently triggered yet greater interest in these issues in the European arena.

First, creation of the European Monetary Union, raised the question of the irrevocable parity at which members of the union locked their currencies against the new Euro. Moreover, the EU's recent expansion has drawn attention to the question of sustainable central parity for exchange rates of new members from Central and Eastern Europe, given their obligation to participate in the EMU in the very near future.

As historical example shows, the choice of a target nominal exchange rate is crucial. Misjudgment can have disastrous effects if its consequence is that the peg is not "...consistent with satisfactory and sustainable macroeconomic outcomes"

(Williamson, 1984). One of the often quoted examples of such a misjudgment is Britain's notorious decision to restore its prewar mint parity with the dollar in 1925 based on Purchasing Power Parity (PPP) calculations, which was not sustainable, either due to miscalculation of the appropriate PPP-based level, or to the failure of the PPP concept as a tool for assessing the appropriate exchange rate level (Officer, 1976). So, the question arises of the appropriate exchange rate level, and what methods should be used to estimate it with satisfactory accuracy. Beyond estimating current or historic misalignments of the currencies according to equilibrium levels, a critical issue is how to use the results of the estimation to forecast future exchange rate movements.

One of the earliest approaches, most respected by theoretical macroeconomists was the PPP method. Straightforward and intuitively appealing, under this method it is stated that the exchange rate level is the ratio of the domestic price level to the foreign price level, so that the real exchange rate of any country should be equal to unity. Prerequisite to the PPP concept is that the law of one price (or LOOP) holds - which is not in general the case. Moreover, even if LOOP does hold, the exchange rate may not still be equal its PPP estimate due to existence of different preferences, trade barriers and other market imperfections.

As a result of the evident inability of the PPP method to explain prolonged swings in exchange rates, subsequently alternative methods were advanced. Researchers developed models ranging from the purely statistical, to purely theoretical,

with some sharing characteristics of both[22]. In all cases the underlying premise was that the exchange rate, as the most important link between the domestic economy and the outside world, should be heavily dependent on the current state of the economy or, as the relevant literature phrased it, on "economic fundamentals".

This part tackles the issue of Equilibrium Exchange Rates (EERs) using a statistical approach, also sometimes called the 'reduced-model' approach. Using a constructed dataset for 19 OECD countries, a thorough empirical testing of previous research in this area is executed, paying special attention to model specification, model construction and the econometric methodology employed - issues often overlooked in the previous studies. The time span is substantial, starting in 1974 and ending in 2003, covering almost 30 years' data on exchange rates and a set of economic fundamentals for all countries studied.

The following section presents an overview of various approaches to exchange rate determination, and survey relevant research. Next, the data is described - its sources, and the choice of variables used as potential fundamentals. The following section presents econometric methodology and empirical findings. The last section presents the conclusions drawn from the foregoing analysis.

2.2 An overview of equilibrium exchange rate models

[22] See Driver and Westaway (2004) for a recent 'in-depth' review of the concepts of equilibrium exchange rates.

2.2.1 Exchange rates and Purchasing Power Parity

Attempts to predict the behaviour of nominal and real exchange rates have a long history. As listed above, one of the first hypotheses to address this question was the PPP concept. PPP was first defined and proposed as a tool for setting relative gold parities, which may be regarded as the exchange rate relationships of the time, by Swedish economist Cassel (1921). The concept is disarmingly simple, stating that prices for the same good converted to a common currency should be the same. The same applies to aggregate price levels for a sufficiently large range of individual goods, so that the nominal exchange rate should serve only as a means of equalizing the relative prices of two countries in one currency:

$$s_t = p_t - p_t^* \tag{2.1}$$

$$q_t = s_t - p_t + p_t^* \tag{2.2}$$

where s_t denotes the home currency price of a unit of foreign exchange, p_t is the price level, p_t^* is a foreign price level (all logarithms) and q_t is the definition of Real Exchange Rate (RER), which should be equal to zero, provided PPP holds. This is an absolute version of PPP and, as noted even by Cassel himself in 1921, it was unlikely, for various reasons, to hold true in practice. These were neatly summarized by Rogoff (1996), who addressing both practical and theoretical issues, including non-existence of the uniform international price index, possible sources of friction, such

as transportation costs or tariffs, existence of non-traded goods, product differentiability, etc. Moreover, the fact that the PPP-based exchange rate estimate is an index quoted against some base year, during which it is assumed RER was at equilibrium value, makes PPP an even less suitable tool. Therefore, the research has focused on a much less restrictive version of PPP, which assumes that the RER is not zero but is mean-reverting to it. The mean reversion is usually tested by unit root analysis of the series or, in more formal terms, whether the coefficient on the first lag of the analyzed variable is equal to 1:

$$q_t = \rho q_{t-1} + \varepsilon_t,\ 0 < \rho < 1 \tag{2.3}$$

The relative version of PPP claims that ρ should be significantly different from 1. This hypothesis has been tested with various versions of Dicker-Fuller and other tests; however, empirically, it proved difficult to confirm PPP validity when the period of flexible exchange rates since the mid-seventies was reviewed[23]. Yet during recent years, especially due to new econometric developments, further attempts have been made to overcome the low power of the earlier tests. Potential remedies addressing earlier tests defects included employing more powerful unit root tests, expanding the time span of the data, using panel data analysis and cointegration methods.

Basing the study on more than one country - especially with regard to exchange rates - to many researchers seemed to be a natural way of modelling RERs. By expanding the cross-section, the researcher might overcome the problem of short times-

[23] See, for example, Froot and Rogoff (1995).

pan of the sample and thereby increase the power of the tests. However, the panel data methods frequently used in studying the macroeconomic data, comprising large cross-section and few time-series observations, were of little use for cross-country macroeconomic data with a long time dimension and a restricted number of cross-sectional units. Only in 1993 were the first panel unit root tests developed. As Levin and Lin (1993) pointed out, this brought about a dramatic improvement in statistical power when using panel data. The tests they developed analyzed the joint null hypothesis that each individual series in the panel was non-stationary, against the alternative hypothesis that all the series were stationary. Therefore, rejecting the null seemed to imply that all the series were stationary, which, when applied to PPP, meant that it was true for all countries in the panel. The advantages of Levin and Lin's approach were that it allowed heterogeneity across units by modelling fixed effects and unit-specific time trends, as well as common time effects. They derived the asymptotic distribution for the test statistic, provided that the time dimension expanded more slowly than the cross-section. Later, Levin and Lin developed an extension of the model, with errors having a more general correlation and heteroscedastic structure, while retaining independence across cross-sectional units.

One objection to these tests is that rejection of the null hypothesis does not necessarily entail that all of the series are stationary (Taylor, 2000). Therefore, Im, Pesaran and Shin (1997) extended Levin and Lin's framework by allowing for hetero-

geneity under the alternative hypothesis, consequently obtaining better finite-sample properties.

Subsequent research has highlighted the importance of residual correlation, and identified other shortcomings in the panel unit root tests, which were not taken into account by Levin and Lin (1993) and Im *et al.* (1997). Maddala and Wu (1999) proposed a simple test intended to avoid the problems of unbalanced panel, choice of lag length in ADF regression, and presence of more complicated cross-correlation structures in the data. This test was based on averaging the p-values of the test statistics for a unit root in each cross-sectional unit.

Several further studies attempted to test the validity of PPP on the sample of emerging economies. By and large, this hypothesis was not supported by individual country studies (e.g. Devereux and Connoly, 1996 for Latin America, and Montiel, 1997 for East Asia) with the exception of a group of studies that used cointegration tests, such as Liu (1992) and Seabra (1995), which found evidence of the equilibrium relation between exchange rate and domestic and foreign prices for the majority of countries in the Latin America sample. The main flaw of the earlier research has been that no tests for the mean reversion properties of RER and, hence, for the speed of convergence to long-run PPP had been conducted. The cross-country studies utilizing panel methods were, to a degree, more successful, and found PPP deviations to be between 3 to 5 years, as in studies for industrial countries (Frankel and Rose, 1996, O'Connell, 1997), although some sources of bias, such as predominance of

monetary shocks in many high-inflation developing countries, cross-sectional dependence, and aggregation across different nominal exchange rate regimes render this an area deserving further future research.

However, even with the revival of PPP research, it was found that the speed of the mean reversion was extremely slow. Most findings indicated the half-life of deviations from PPP to be between 3 and 5 years (e.g. Frankel and Rose, 1995, Taylor, 2000) - too long a time period to accept PPP as a valid tool in assessing country exchange rate misalignment and drawing any policy conclusions in response. Moreover, Murray and Papell (2002) found that, in many cases, the upper bounds of the confidence intervals were infinite, implying that the estimated half-lives provided little information about the speed of mean reversion.

Notwithstanding these facts, and the various conceptual limitations cited above, PPP concept is computationally straightforward, imposes minimal data requirements and, therefore, is still used to indicate appropriate exchange rate, especially for countries in transition (IMF Economic Outlook, 1999) - despite the fact that it is arguably even less suitable for developing and transition economies than for industrial countries.

Bearing in mind that exchange rate and price levels are both endogenous variables determined by some other exogenous variables, one should realize that PPP is less a theory of exchange rate determination, than a relation between these variables to be satisfied in equilibrium (Chi-WaYuen, 1998). Given the size of distortions

in the real world, especially for developing and transition countries, the proposal is therefore to analyze the quantitative significance of the various economic factors that are likely to affect exchange rate levels.

2.2.2 Exchange rates and economic fundamentals

As a result of PPP's empirical failure and deficiencies noted in the previous section, a number of studies went beyond simple univariate analysis based on some version of PPP, by adding further explanatory variables to the exchange rate equation. While the empirical results of structural models for exchange rate determination (such as monetary models) were rather disappointing, other approaches that incorporated additional economic variables, called fundamentals, proved to be more successful. Alongside, the emphasis in EER's modelling shifted towards models analyzing the whole economy, instead of a single sector, so that the exchange rate can serve as an important link between the domestic economy and the foreign sector, determined both by stock and flow variables.

Foreign exchange rate fluctuations have long been a topic of debate, with numerous articles written on the subject. Two basic theoretical views have dominated the discussion since the 1970s. Of the theoretical models of exchange rate determination one, the monetary approach, emphasizes the impact of monetary-related variables. The second is the portfolio balance approach, which emphasizes the impact of balance of payment-related variables on exchange rates. After more than

twenty years of empirical studies testing these prepositions, results now indicate that neither of these models has consistently performed well. The simultaneity effects, globalization of markets, linkage effects, tremendously increased capital flows, market shocks and their continuous transformation have created greater exchange rate uncertainty and volatility. In order to identify a model which would better describe exchange rate behaviour, many empirical studies were undertaken, which aimed to confirm theoretical relations between exchange rate and explanatory variables.

Exchange rates and Balassa-Samuelson effect

In order to understand exchange rate behaviour, one strand of analysis has tried to identify a relatively small set of economic variables that can be used to explain movements in exchange rates based on one of the theoretical models developed. These models are denoted as 'EER models' as they sought to assess whether the current level exchange rate level was consistent with the underlying economic fundamentals which, according to their theoretical assumptions, should both directly and indirectly affect exchange rate, at least in the medium to long term.

Chronologically, the first model to consider long-run structural deviations from PPP was advanced by Balassa (1964) and Samuelson (1964). They argued that rich countries tend to have higher price levels because they are more productive in the traded goods sector. They further contended that a rise in productivity would have no effect on prices in the traded goods sector because the domestic price level for tradables was tied down by the world price level and competition. As wages must

rise in the traded goods sector as a result of increased productivity, nontraded goods producers would be forced to lift prices to match the absence of productivity advance, leading to an overall price level increase relative to the less productive country. This in turn would lead to RER appreciation. This impact is easily incorporated into the ERER model by assuming that the corresponding price levels for two countries are decomposed into the prices of both traded and nontraded goods:

$$p_t = ap_t^T + (1-a)p_t^N \tag{2.4}$$

$$p_t^* = \alpha^* p_t^{*T} + (1-\alpha^*)p_t^{*N} \tag{2.5}$$

where T and N denote the indices for traded and non-traded goods, a and α^* are the shares of the traded goods sector in the economy, which might be assumed as changing in time or across countries. Combined with the definition of the real exchange rate (2.2) for the economy in total, and for the traded goods sector alone, it is easy to derive an expression for the RER which captures the productivity differential effect as:

$$q_t = q_t^T + (1-a)(p_t^T - p_t^N) + (1-\alpha^*)(p_t^{*T} - p_t^{*N}) \tag{2.6}$$

If we assume that the law of one price holds continuously, at least for traded goods, then q_t^T will be zero, or, less restrictively, will be constant. Then, the trends in RER arise because of movements in the relative prices of goods between countries.

Although the productivity hypothesis appeared theoretically to be very sound, the earlier studies could not find any statistically significant relationship between productivity and exchange rate changes. For example, Rogers and Jenkins (1995) found the evidence of real exchange rate for traded goods or q_t^T in 2.6 being non-stationary for the OECD countries, a fact that seemed to contradict the Balassa-Samuelson hypothesis. In addition, they could not find supporting evidence even for highly traded disaggregated goods. Moreover, Rogoff (1996) showed that the relationship between income (assumed as a proxy for productivity)and relative prices is not uniform,as the hypothesis suggests; instead there is a clear segregation into two groups - developed industrial countries and the rest. Indeed, it is quite striking that Japan was the only industrialised country to experience development of the exchange rate in line with Balassa-Samuelson hypothesis. For other countries, results are less convincing. Froot and Rogoff (1991), for example, found little support for this hypothesis across EMS countries for the period 1979 - 1990.

The results of research on the Balassa-Samuelson effect concerning exchange rate deviations from PPP for developing and transition countries were rather mixed. Although these countries should exhibit large productivity growth as a consequence of catching-up or restructuring processes, cross-country differentials in productivity growth have been unable to account for persisting PPP deviations for the aggregate panel of developing countries (Cheung and Lai, 1998). As was shown by Ito, Isard and Symanski (1996), in relation to East Asian economies, this effect could be ap-

plicable only at a certain stage of development "when a resourceless, open economy is growing fast by changing industrial and export structures". These conditions do, however, apply to the transition economies of Eastern and Central Europe. Richards and Tersman (1996) examined this effect in the context of the general equilibrium model for the Baltic countries and found it to be the driving force of RER deviation from PPP values. A recent paper by Halpern and Wyplosz (2001) attempting to estimate directly the Balassa-Samuelson effect for nine European transition economies reached the conclusion that 1991-1999 data provided unequivocal evidence favouring the sector productivity effect on RERs.

Studies using cointegration methods for univariate series, or in the panel framework have only recently appeared. Faruqee (1995) and Strauss (1995) both employed the Johansen MLE method, and found cointegration between RER and their proxies of productivity for most countries in the study. When analysing single equation time series models, Chinn (1996), who used several cointegration estimators, found very weak statistical links between real exchange rate and total labour productivity (the latter frequently used as a proxy for the Balassa-Samuelson effect). Much more favourable results were obtained, however, when working with a panel of countries. Mark and Doo-Yull (1995), analysing the effect of various proxies over a longer time period, on the other hand found the relative productivity measure to be the only significant variable.

Exchange rates and interest rates

Another putative explanation of exchange rate behaviour was the interest parity condition. In the mid-1970s, two competing theories were developed - Dornbush's (1976) theory of exchange rate overshooting, and Frenkel's (1976) theory of the effect of expected inflation differentials. The two corresponding schools assumed that the Uncovered Interest Parity (UIP) held, and used this as one of models' main assumptions:

$$q_t = E_t[q_{t+1}] + (r_t - r_t^*) \tag{2.7}$$

where $E_t[q_{t+1}]$ is the expectation at time t of the RER at time $t+1$, and $(r_t - r_t^*)$ is the difference between real interest rates at home and abroad. Ignoring risk premia, and also assuming free cross-border capital movements, the arbitrage condition on the international foreign exchange markets insures that the difference in real interest rates between the two countries is reflected in the expected change of exchange rate.

Surprisingly, the initial tests almost uniformly rejected the existence of any relationship between real interest rates and RER. Meese and Rogoff (1988), Geweke and Feige (1979), Hakkio (1981), Hsieh (1984) all concurred in this pessimistic conclusion. Arguments in favour of UIP, for example, that the financial markets are much more liquid and efficient than the goods market, and that possibilities for legal arbitrage are minuscule, somehow didn't work in the real world.

However, as Baxter (1994) observed, these negative results were based on inappropriate usage of differencing of the data, in order to remove present non-

stationarity. This also removed most of the low frequency information from the data - the information which in this context was the most useful. Baxter showed that, whereas previous research relied on standard regression techniques and applied them to the following equation:

$$q_t = b_0 + b_1 r_t + b_2 r_t * + \varepsilon_t \tag{2.8}$$

the more appropriate specification would have decomposed RER into its transitory component q_t^T, which is stationary, and the permanent component q_t^P, which is not, so that:

$$q_t^T = a + b(r_t - r_t^*) + \varepsilon_t \tag{2.9}$$

On this specification by contrast, Baxter found that the majority of b's significant and correctly signed. MacDonald and Nagayasu (1999) later confirmed these findings. Moreover, on testing jointly with PPP, strong evidence of cointegration was found (Johansen and Juselius, 1992, MacDonald, Marsh and Nagayasu, 1996 and others). According to MacDonald and Swagel (2000), the interest rate effect reflects primarily the influence of business cycles on exchange rates and, therefore, is, by nature, of short to medium-run. This argument, in parallel with frequent findings of the stationarity of the interest rate differential then prompted researchers to exclude this variable from the list of fundamentals affecting RER in the long-run.

Exchange rates and indebtedness of the country

While frequently argued as being one of the major determinant of RER (Burda and Wyplosz 1997, p. 154), the absence of statistics on countries's financial position made it virtually impossible to analyze the effect of a country's indebtedness in practice. Initial studies attempted to confirm the relationship between sustained current account deficit and RER depreciation. Theory predicted that current account developments were likely to induce significant RER changes because they led to wealth-transfer across countries. Studies did find significant correlation between RER and accumulated current account or net foreign assets (e.g. Obstfeld and Rogoff 1995 for OECD countries). Moreover, recent work by Lane and Milesi-Ferretti (2004) provided time series of net foreign assets (including imputed capital gains and losses, and other adjustments) for a large number of industrial and developing countries and established the unambiguous effect on RERs of national indebtedness.

2.2.3 Fundamental-based models of exchange rate determination

Differences in sector productivity, real interest rates and a country's net foreign asset position have not been the only structural factors advanced in explanation of RERs' behaviour. The mixed empirical results described above served to justify the development of more general models, in which exchange rate was determined via interaction of various economic forces. Two principle approaches emerged. One strand of the literature attempted to build a full general equilibrium model, incorporating a concrete

normative concept of equilibrium exchange rates, for example, the "Fundamental Equilibrium Exchange Rates" models (FEER) developed by Williamson (1994) and Driver and Wren-Lewis (1998). In FEER, the main focus is on current account determinants, as a function of home and foreign output, and real effective exchange rate. Thus, the real exchange rate that is consistent with macroeconomic balance is the rate that brings the current account to equality with the normal, or sustainable, capital account, where the determinants of the current account are set at their equilibrium of full employment values.

Clark and MacDonald (1998) have observed that "... the FEER is calculated using the exogenously given estimate of sustainable net capital flows". Under their proposition, the FEER approach didn't embody the theory of exchange rate determination, rather, it implicitly assumed that a divergence of the exchange rate will set in motion unspecified forces in turn bringing it back to equilibrium. Thus, FEER's lack of dynamics is its main drawback, so that it primarily serves only as a means of exchange rate current value assessment, in a process of parameter estimation involving considerable judgement. Williamson (1994), for example, relied on a host of factors in order to arrive at current account targets in 1995 for fourteen countries. Wren-Lewis and Driver (1997) relied on an even wider set of factors and assumptions to calculate FEERs for 2001. Although Isard and Faruqee's (1998) extension of the model remedied some of the problems of the FEER approach, it nevertheless retained one of its unsatisfactory features, namely that the current account plays a leading role

in the model, without any feedback from the capital account - in particular including, neither saving nor investment as a function of the exchange rate.

A second strand in the literature attempted to explain actual trends in RER, this time by estimating reduced-form equations implied by the macroeconomic balance or other approaches. Several authors attempted to build general equilibrium models of the economy, focussing on medium to longer run RER behaviour, in contrast to the previously advocated short-run models of monetary dynamics with rational expectations. Here, ERER's main determinants were assumed to be the fundamentals of thrift, productivity, capital intensity, and net debt to foreigners - in other words, all the variables that might influence a country's long-term capital flows and clear the balance of payments (e.g. Faruqee, 1995). The various models in this mould shared one idea in common: ERER was the value, or path, consistent with internal and external macroeconomic balance, and the role of empirical studies was to confirm which of the economic fundamentals did indeed play a role in RER determination.

Given that most economic time series cannot reject the hypothesis of non-stationarity, any inference from the usual regression will give misleading results. The earlier research, therefore, transformed the data by differencing, in order to study the relations between EER and economic variables, or fundamentals. This, however, failed to produce any satisfactory results, which led researchers to claim that EER behaviour was unpredictable, and bore no relation to economic processes. Only with

the advent of cointegration theory, which allowed studying relations between non-stationary variables in the long run, were more optimistic results obtained.

First, Engle and Granger (1987) developed the single-equation method for estimating cointegration relations; this was based on a restrictive assumption of a single cointegration relationship. Later, however, Johansen (1988 and 1996) developed the maximum likelihood estimation (MLE) method of testing for cointegration, this time based on the reduced rank regression, allowing for identification of more than one cointegration vector in the data. This approach's most serious limitation was its need for data covering a sufficient timespan to allow meaningful inferences to be drawn, and thereby to increase the power of the tests. On this occasion, the remedy was to apply panel methods, given the absence of reliable data over a sufficient period for the majority of countries.

Techniques developed by Kao(1999) and Pedroni (1999) have extended static single equation regressions of the type indicated by Engel and Granger (1987), to pooled panel static regressions, producing test statistics for cointegration in panel data. However, the implicit assumption in such tests was the existence of unique cointegrating vectors, albeit, heterogeneous, across the panel units. Later studies by Groen and Kleibergen (1999), Larsson and Lyhagen (1999 and 2000), and Larsson, Lyhagen, and Loethgren (1998) developed techniques similar to Johansen's (1996) MLE method allowing for multiple cointegrating vectors in cross-sectional units.

In contrast, Banerjee, Marcellino and Osbat (2002 and 2004) have questioned the validity of panel methods for integrated series where cross-unit cointegrating relationships are present; a common condition in macroeconomic time series, which entails massive size and power distortion of the tests. In this instance, the remedy proposed was to use full system estimation when possible or, alternatively, unit by unit cointegration analysis. The following step was to apply the Gonzalo and Granger (1995) procedure, in order to test for the presence of cross unit cointegration and different ranks across units. Only with a negative answer could the panel method be used. This approach thus appears rather restrictive, and further research is needed to develop correct panel cointegration methods. Nevertheless, due to the short duration of data available for transition economies, these methods may be the only possible solution for exchange rate models cointegration analysis in the context of the study.

2.2.4 Empirical results of estimation-based exchange rate modelling

As anticipated, identification of clear-cut relations between fundamentals and RER, with uniform application across countries, was not an easy task. First, research has in almost all cases focused on only one fundamental at a time, instead of adopting the general economic model perspective - a recent progression. Secondly, data availability is an all too frequent problem with regard to data on capital flows and on transition and developing countries. Using proxies for the economic fundamentals at the same time exposed the analyses' conclusions more vulnerable to critique. Thirdly, apply-

ing conventional regression techniques for nonstationary data series analysis was also problematic.

As mentioned above, models incorporating both stock and flow approaches to ERER determination have risen in popularity in recent years. As proof that these models did explain RER behaviour, relevant studies attempted to find relationships between exchange rates and economic fundamentals. Faruqee (1995), for example, argued that RER determinants included factors affecting both the home country net trading position in world markets, as well as its underlying propensity to be a net lender or borrower of capital; in other words, that the interaction between permanent structural components in both current account and capital account determine the sustainable or equilibrium RER. From the trade side, the model predicted that determinants such as productivity growth differentials, or terms of trade, might play a role; from the finance side, fundamentals determining long-run net foreign asset position, such as demographic factors or government expenditure, would also be important. Using postwar data on Japan and the US, Faruqee found that net foreign assets and productivity differentials were, in long-run, correlated with RERs.

Using this approach to study RER for the major world economies, MacDonald (1997 and 2000) found cointegration between RER, terms of trade, relative sector productivity, risk premium and government spending. He termed his approach as Behavioural Equilibrium Exchange Rate (BEER), and advocated its advantages over Williamson's FEER normative model. The first empirical evidence in favour of the

BEER approach was provided by Clark and MacDonald (1997 and 1999). Using a set of fundamentals including the terms of trade, the relative price of traded to non-traded goods, and net foreign assets, and applying cointegration techniques to demonstrate a long-run relation between these fundamentals and exchange rates of the US, Germany and Japan.

Alberola *et al.* (1999) followed a similar approach in order to estimate bilateral equilibrium exchange rates for the major currencies in the panel cointegration framework and obtained results, that were strongly supportive of the presence of a long-run relation between RER and economic fundamentals. By contrast, Roeger and Hansen's (2000) study using the same countries' sample failed to reach statistically sound results for many countries, instead relying on pragmatic assumptions in order to produce estimated equilibrium exchange rates.

One recent application of this approach, by Clostermann and Schnatz (2000), has estimated the BEER for the Euro by constructing synthetic fundamentals for the period prior to the Euro's introduction. A subsequent refinement was undertaken by Maeso-Fernandez *et al.* (2001), developing four models, each explaining the behaviour of a "synthetic" Euro currency. An ECB study (Detken *et al.*, 2002) employing various approaches to Euro equilibrium rate determination, stressed the uncertainty accompanying all estimation methods.

A criticism of the BEER approach is that the fundamentals used in calculations of the equilibrium value of the exchange rate are not themselves set at deemed equi-

librium values. One method of measuring equilibrium exchange rates is to extract fundamentals' trends using Hodrick-Prescott, or another filter; alternatively decomposing the time series into its permanent and transitory components. The permanent component is expected to have a persistent effect, accordingly interpreted as a measure of equilibrium, and forms the Permanent Equilibrium Exchange Rate (PEER).

In recent years a number of studies focusing on estimating the relationships between economic fundamentals and RER for developing economies have appeared. Most of these have relied on cointegration techniques, however several theoretical models have been deployed (representative agent with intertemporal frameworks, portfolio balance or asset market with balance of payments synthesis model) from which a reduced form for ERER is derived. Almost all such studies included, as regressors, the terms of trade, some measures of net capital flows, government spending, openness and output growth. Soto and Valdes (1998), for example, performed the cointegrated VAR analysis for Chile, using terms of trade, productivity differentials and net foreign assets, and obtained a clear indication of the presence of a cointegrated relationship. Broner, Loayza and Lopez (1997) used the Johansen cointegration technique for seven Latin American countries, using approximately the same set of fundamentals, and reached a similar conclusion. Research conducted by the World Bank (Montiel and Hinkel, 1999) reviewed practical methodologies for assessing exchange rate misalignment, especially in low-income developing countries, where data, time, and professional capacity are limited. The general conclusion

in this case was that reduced-form single-equation econometric estimates of ERER, which make it possible to take into account the interaction of key macroeconomic variables in a full general equilibrium theoretical framework, was a promising avenue for further research on ERER estimation. Its usefulness in policy applications, however, was dependent on the availability of fairly long and reliable data series for the RER and its key determining variables.

As for transition economies of Eastern and Central Europe, over the last ten years a number of studies applying statistical and model-based methods of exchange rate estimation have been conducted. Halpern and Wyplosz (1997), for example, calculated ERER by means of estimating equilibrium dollar wages: a set of national characteristics, using a cross-section of 80 countries was constructed, so that the actual dollar wage in each country was used to measure the degree of RER misalignment. While their model was based on the Balassa-Samuelson effect, various proxies were used to test for structural adjustments and efficiency gains in the financial sector, labour market and trade sector, and support was found for the hypothesis that productivity and gains in economic efficiency in general contribute to RER deviation from PPP value. Drawbacks of their study, however, were that only five time-series observations for each state in the study were used, and the usual panel data regressions were applied to estimate the importance of economic factors on RER expressed in terms of dollar wages. A more recent survey of equilibrium exchange rates in Central and Eastern Europe by Egert (2003), summarizing various studies conducted

for eight new EMU members, and assessing the extent of misalignment of their exchange rates for ERM II entry, observes considerable divergence in the estimated misalignments, due to the different period under investigation and variation in sets of fundamentals. Nevertheless, the results produced still offer a much better indication of possible misalignments than the PPP concept suggested.

Motivated by the uncertainty of the empirical evidence cited above, this study now undertakes to estimate equilibrium exchange rates for a large set of developed countries, using a comprehensive dataset spanning 30 years of quarterly observations, and paying close attention to the needs for both proper econometric model construction and testing.

2.3 Motivation and choice of the variables

Here the exchange rates of 19 OECD countries, namely 11 EMU members (except for Luxembourg), 3 EU countries outside of EMU (Denmark, Sweden, United Kingdom), 2 European non-EU countries (Norway, Switzerland), and 3 large world economies (Canada, Japan and United States) are examined covering the period 1974Q1 to 2003Q4. Cointegration analysis on a country by country basis is performed, using quarterly data from the NIGEM database as a primary source with certain gaps filled by recourse to Datastream and IFS. In distinction from some studies by the ECB on the synthetic euro, where the data for individual countries are used to construct the aggregate euro area series (see Clostermann and Schnatz, 2000,

Maeso-Fernandez *et al.*, 2001, Schnatz *et al.*, 2003), this study assesses individual EMU currencies even after introduction of the Euro[24] - an approach perhaps in line with what, in ECB studies are termed "synthetic" EMU currencies. Justification for this approach, it is suggested, derives from the following grounds.

First, thorough analysis of the synthetic euro and Euro-area aggregated fundamentals had already been completed in the course of the studies mentioned above. As results yielded in these studies were rather weak, a logical progression was to repeat the exercise on an individual country basis, in order to highlight cross-country differences perhaps obscured by aggregate analysis. Secondly, the resulting misalignments of the currencies of individual EMU members may provide an indicator of internal Euro stability as a single currency for 12 different economies. Thirdly, using an individual countries framework, the issue of structural break after the Euro introduction and overall stability of the constructed models can be statistically assessed.

This analysis focuses on real effective exchange rates, defined here as the log of CPI-deflated indices for each country. For computation of the effective exchange rates, the trade weights (w_{ij}) stated by the IMF study (Zanello and Desruelle, 1997) are used. The log real effective exchange rate (REER) for country i is thus the trade-weighted average of the log bilateral real exchange rates (e_{ij}) against its trading partners:

[24] Which are easily obtainable from the Euro market rates - by dividing by the irrevecable parity conversion rates.

$$REER_i = \sum_j w_{ij} e_{ij} \tag{2.10}$$

Nominal values of EMU currencies after Euro introduction are then constructed using Euro market rate divided by irrevocable parity rates of individual member currencies against Euro, as set by the ECB in 1999.

Several factors were taken into account in the choice of economic fundamentals used in the present study. First, data from which the fundamentals are constructed were required to be readily available at a greater than annual frequency, to avoid the use of interpolated series. Secondly, the set had to be limited, in order to avoid inflated models, entailing subsequent problems of interpretation. Thirdly, the analysis needed to include fundamentals studied in previous empirical analyses of PEER/BEER approaches. Finally, sound theoretic justification of the effect exerted by these fundamentals on the exchange rates was required. On the basis of these conditions, the following economic fundamentals were selected.

Productivity differentials

The impact of differences in relative productivity on exchange rate is based on the Balassa-Samuelson theory, explained in detail above. This is a widely used fundamental in studies on equilibrium exchange rates (e.g. MacDonald (1997 and 2000), Chinn (1999) and others). Given that direct measures of productivity in tradable and non-tradable sectors are not readily available on a timely basis for many countries, indirect proxies have frequently been used. The relative price differential of consumer

to wholesale price indices of the home country against its trading partners is believed to capture the effect of productivity increases in the traded goods sector. This proxy for each country i was calculated as:

$$RPROD_t^i = \log\left(\frac{\left(\frac{CPI_t^i}{PPI_t^i}\right)}{\prod_{j=1}^{18}\left(\frac{CPI_t^j}{PPI_t^j}\right)^{W_t^{ij}}}\right) \tag{2.11}$$

Net foreign assets

In addition to relative sector productivity, many studies included in their analysis country's stock of foreign assets (e.g. Alberola *et al.*, 1999), justifying this choice by portfolio-balance considerations. Given the absence of estimates of stock of net foreign assets in all cases, many studies used the cumulated current account balance as a proxy; this, however, ignores the effect of debt reduction, valuation issues, and errors and omissions in countries' capital account data. In this study, two proxies for net foreign assets were used. The first was constructed by linear interpolation from Lane and Milesi-Ferretti's (2004) data. The second was constructed by accumulating current accounts for each country, using OECD (1996) data as estimates of initial stocks[25]. In order to account for country size, the estimates were normalized by each country GDP:

[25] As subsequent analysis showed, the proxy constructed by accumulating current accounts provides a better estimate than the proxy based on Lane and Milesi-Ferretti's (2004) data for all countries except Belgium.

$$NFA_t^i = \frac{NFA_0^i + \sum_{j=1}^{t} CA_j^i}{GDP_t^i} \tag{2.12}$$

Real interest rates differentials

Numerous studies have confirmed the relationship between interest rates and exchange rates (MacDonald and Nagayasu (1999), Johansen and Juselius (1992), and others). While many researchers argued that real interest rates should equalize in the long run, across countries, empirical evidence had suggested that not be the case, at least looking at the last 30 years' data. What was found, was that real interest rates showed mean-reverting behaviour in the medium- to long-run. However, like exchange rates, real interest rates also show prolonged deviations from equilibrium values, and since the BEER/PEER approach focuses on the short- to medium-run of exchange rates determination, this phenomenon served to justify inclusion of this variable in the analysis. Real interest rates were proxied by government bond yields, less CPI index inflation for the preceding year. Real interest rate differential was constructed as a difference between each country's real interest rate, and the trade-weighted sum of foreign real interest rates:

$$RIR_t^i = r_t^i - \prod_{j=1}^{18} (r_t^j)^{w_t^{ij}} \tag{2.13}$$

Fiscal positions

As a key component of national savings, fiscal balance can be shown to have an effect on exchange rate determination (Frenkel and Mussa, 1998). Several studies have investigated the effect of fiscal balance on exchange rates (e.g. Maeso-Fernandez *et al.*, 2001). Whereas data on budget deficits is readily available, it was decided to include this variable in the analysis. As with real interest rate differential, the difference between the budget deficit of country i, and the trade-weighted average of other countries, was included in the cointegration analysis[26].

$$RFBAL_t^i = fbal_t^i - \prod_{j=1}^{18} (fbal_t^j)^{w_t^{ij}} \tag{2.14}$$

2.4 Econometric methodology

The econometric methodology employed in this study uses cointegration analysis of the system of variables in order to identify long-run relationships between them. In this respect, it adopts the usual strategy, as developed by MacDonald (1997 and 2000), applied in many studies on BEER/PEER analysis of exchange rates. In light of Alberola *et al.* (1999) and Hansen and Roeger's (2000) contradictory results, the paper first estimates equilibrium exchange rates using only two fundamentals (productivity differentials and net foreign assets), in order to reconcile findings from

[26] The data on quarterly fiscal balances was taken directly from NIGEM database (variable **GBR, where '**' is the country code).

those two sources. Secondly, the models are reestimated, with two additional fundamentals (real interest rate differentials and relative fiscal balances) with the aim of improving on the base models.

Before estimating the VECM models, following the standard practice, the stochastic properties of the series are assessed using unit-root tests. This complete, the series of cointegrating tests are conducted for both 3-variable models and larger 5-variable models. Cointegration amongst variables in question established, the long-run relationships are estimated using several cointegration methods. Estimated models are then checked for adequacy, residuals behaviour, stability of the estimates, and plausibility of the estimated long-run relationships and exchange rate correction to equilibrium level. The estimated long-run relationships are used to construct the BEER estimates of the equilibrium exchange rate, whereas the estimated cointegrated parameters are used to perform permanent-transitory decompositions, as suggested by Gonzalo and Granger (1995), in order to obtain the PEER estimates. The resulting estimates show the degree of multilateral misalignments which can be used to obtain the bilateral misalignments of each currency against the others, in line with Alberola *et al.* (1999) algorithm.

2.4.1 Unit root tests

The order of the integration of the series is assessed using conventional ADF tests, with constant included. The lag length was selected according to information criteria,

in order to obtain uncorrelated residuals. For some of the series, where sufficient lag length failed to ensure uncorrelated residuals, seasonal dummies were included. In rare cases, where visual inspection of the data indicated the possible presence of the trending behaviour, the trend was also included in the ADF regression.

It is now fashionable to complement low power ADF tests with more powerful univariate alternatives, or panel unit roots tests. Given the favourable results obtained indicating the presence of unit roots in the series by univariate tests, the present study omits these tests.

2.4.2 Cointegration tests and VECM estimation

Selecting the VECM modelling framework requires prior choice of the cointegrating rank of the system. Widespread methods in this respect, are methods based on sequential likelihood-ratio tests. Two tests for cointegration have been employed in this study. First, the widely used trace test, developed by Johansen (1996), and second, the two-step test developed by Luetkepohl and Saikkonen (2000) which, in some cases, showed better power properties.

Having established the cointegrating rank of each system, the starting point of cointegration analysis is typically a VECM setup, specified as follows:

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \mu + \Phi D_t + \epsilon_t, \quad t = 1, ..., T \tag{2.15}$$

for fixed initial values of $X_{-k+1},\ldots, X_0$ and $\epsilon_1, \ldots, \epsilon_T$ being identically and independently distributed $N_p(0, \Lambda)$ errors, μ a constant, and D_t containing the deterministic terms of the model. The rank of the matrix Π determines the number of the cointegrating vectors. If it is of reduced rank, r, then there exist $(n \text{ x } r)$ matrices α and β such that:

$$\Pi = \alpha\beta' \tag{2.16}$$

where β's columns are linearly independent cointegrating vectors and α is the adjustment matrix of factor-loading vectors.

Once the cointegrating rank of the system is known, it becomes a straightforward matter to estimate the VECM using the reduced rank maximum likelihood (ML) estimation method (Johansen, 1996) - a method frequently used in cointegration analysis, as it is implemented in many econometric packages. However, as shown by Bruegmann and Luetkepohl (2004), at least on some datasets, it produces extreme estimation results. Consequently, here the models were additionally estimated using a simple-two-step (S2S) estimator, which has the same asymptotic distribution as the ML estimator (Reinsel, 1993), in order to assess the sensitivity of the results to the estimation method (see Luetkepohl (2004) for the a full description of S2S estimation method).

The estimated models were then checked for adequate representation of the data generation processes (DGP's) underlying the time series under analysis. For the model to be accepted, several criteria needed to be met. First, the residuals were

required to show no correlation both at low and high lags; to be close to normally distributed; and to show no conditional heteroscedasticity. The estimated system also had to be stable, showing no structural breaks. Additionally, the estimated beta vector coefficients had to be of plausible order. Moreover since we are looking at the equilibrium exchange rates this implies that the estimated system had to show significant and negative alpha vector coefficient for the exchange rate equation. As an additional check, the orthogonal complement of alpha vector ($\alpha_\perp$), identifying the variables that drive common stochastic trends in the model, had to show common stochastic trends driven by the variables other than exchange rate itself.

Only the models satisfying the above mentioned criteria were considered as suitable and used to construct BEER/PEER estimates. The BEER estimates were derived directly from the estimated long-run β vectors expressing real exchange rate in terms of the fundamentals. The PEER estimates were derived using Gonzalo-Granger (1995) decomposition that identifies the permanent component of VECM system as:

$$PEER_t = \beta_\perp(\alpha'_\perp\beta_\perp)^{-1}\alpha'_\perp y_t \tag{2.17}$$

where $\beta_\perp$ and $\alpha_\perp$ are orthogonal complements of the estimated α and β vectors.

2.5 Empirical results

2.5.1 Unit roots and cointegration analysis

Prior to cointegration analysis ADF tests for non-stationarity of the series were performed in order to identify the stochastic properties of the data. The ADF tests were run with a constant, and in some cases, the trend was included where its presence was suggested by visual inspection. Unit roots were implemented for the 1974Q1 - 2003Q4 period, spanning 30 years for the real exchange rates series and four fundamentals. Lag length was determined using information criteria, and checked for the absence of serial correlation of residuals. As most of the studies focused on shorter time periods, it was assumed that the ADF tests would have greater power in our case, and non-rejection of the null hypothesis of unit root - even by low-power ADF test - would be more convincing. Table 1 summarizes the results of the findings.

What emerges from the table, is that the only series rejecting the null of non-stationarity at the 1% level is relative fiscal balance (RFBAL) for Portugal. Indeed, visual inspection of this series does appear to indicate that the behaviour of the series is cyclical, with frequent swings around its mean. In light of this behaviour, and of subsequent model construction analysis (showing that there was no satisfactory model for Portugal REER, even assuming non-stationarity of the series in question) the series was excluded from the analysis.

With only a small number of exceptions, all series, except for real interest differentials (RIR), failed to reject the non-stationarity hypothesis at 5% significance level. Approximately half of the RIR series rejected the null at 5%, but not at 1%.

Overall, these results of the ADF tests provided sufficient evidence of non-stationarity of the REER and selected fundamentals, although in some cases only at 1%. Given that the question of non-stationarity issue can be further verified at the later stages, of cointegration tests and model specification, it was decided to proceed, relying on the ADF tests results as supporting the hypothesis of non-stationarity of relevant series.

The next step is to perform cointegration tests for each country. First, the tests were run for those models with only two fundamentals only (REER, RPROD and NFA), so as to replicate analyses of Alberola *et al.* (1999), and Hansen and Roeger (2000), each showing different results for the same dataset. Next, the cointegration tests were run for models with four fundamentals (introducing RIR and RFBAL variables into the analysis). Two cointegration tests were used: Johansen (1996) trace test, the second, Saikkonen and Luetkepohl (2000) two-stage test (hereafter 'SL-test') which in some cases, demonstrates better power properties (Luetkepohl and Saikkonen, 2000). Where no cointegration was found, the tests were re-run on a shorter time period (1980 Q1- 2003 Q4), as some of the series showed very volatile behaviour, or were interpolated during the 1970's. Indeed, on visual inspection, the analyzed series exhibit rather erratic behaviour, with frequent outliers, in certain cases of substantial magnitude. In the context of these observations, cointegration test results should be interpreted with caution, avoiding the purely subjective judgements made in some previous studies. Results of the cointegration tests are summarized in Table 2.

The results of cointegration tests for the 3-VAR systems show that these two methods do not in all cases indicate the same level of cointegration. Only in 9 cases out of 19 do the two methods simultaneously yield the cointegrating rank of 1; in 6 cases only the Johansen trace test finding cointegration of rank 1, whereas the SL-test indicates no cointegration at all; in 2 cases, the trace test displays the rank 2 and the SL-test 1; in 1 case the trace test finds no cointegration and the SL-test shows a rank of one, with rank 0 signalled by both tests in the final case. Notwithstanding these differences, overall, the results indicate - subject to a degree of uncertainty resulting from the small sample size, and presence of large outliers in the series - that the 3-VAR models have one cointegrating vector (except, as described earlier, for Portugal, where no cointegration can be established). To re-state, one or other of the tests indicates the presence of one cointegrated vector in the models in every case, except that of Portugal.

Looking at results for the 5-VAR models, again, the two tests do not always elicit identical results. In 11 out of 19 cases, two tests indicate the rank of 1; another 2 cases showing rank of 1 on the trace test, in parallel with the SL-test failing to find any cointegrating vector; 4 cases in which the opposite occurs; and 2 final cases for which both tests indicate rank 0 and 2. As mentioned above, due to deficiencies in data, the test results should be treated with caution and, overall, the conclusion should be that the tests yield one cointegrated vector, except in the cases of Spanish and Swedish data.

The selection of cointegrating rank is crucial for further analysis, and many researchers, faced with borderline results, have used economic theory justifications for their choice. Moreover, conclusions about the rank of each model will be subject to further analysis at the stage of model formulation and, as such, the results of the cointegration tests should not be regarded as a final decision. Given the results of the tests above, it was decided to accept the hypothesis of one cointegrating vector in all the models.

2.5.2 Estimated models

As previously mentioned, the analysis conducted was split in two parts. The first stage looked at exchange rate models with two fundamentals - relative price differential, and net foreign asset position. This was the model used by Alberola *et al.* (1999) for synthetic Euro currency and principal industrialised countries' currencies, on which basis they reported extremely clear and uniform results, although without indicating reporting any detailed findings apart from estimated beta vectors - making it difficult to assess the statistical adequacy of their models. A study by Hansen and Roeger (2000), using the same model and dataset, was unsuccessful at replicating the findings, having replaced statistical arguments with pragmatic assumptions that undermined their results. Given such conflicting conclusions, the present study attempts to reconcile these differences at the first stage. The second part of the analysis augments the basic 3-VAR model by adding two further fundamentals used in certain

studies on equilibrium exchange rates, and checking whether this improves exchange rate modeling in any way.

For 3-VAR and 5-VAR systems the following VECM specification was used. Each model was estimated with seasonal dummies, and a constant restricted to cointegrated space[27]. Number of lags was selected according to information criteria, and increased where necessary in order to obtain serially uncorrelated and normally distributed residuals. Given, firstly, the presence of multiple outliers, and secondly, the series' erratic behaviour , a number of impulse dummies were also introduced into the model. On the basis of the cointegration test results, the models were estimated with cointegrating rank of 1. In cases where the tests displayed borderline significance for rank of 2, the models were reestimated, and checked for stability of the second cointegrating vector and general model adequacy.

The 3-VAR models were estimated without any subset or long-run vector restrictions, using two estimation methods: namely, Johansen reduced rank regression (RR regression) and simple two stage (S2S) estimator. Whereas the two methods produced approximately the same beta vectors and short run coefficients, they gave very different estimates for alpha vector and short run coefficients for some countries in the datasets. The same problem was encountered during the estimation of 5-VAR models. As noted by Bruegmann and Luetkepohl (2004), certain problems were observed with the RR estimator on some samples. While this issue clearly warrants

[27] When the t-statistics for the deterministic term included in cointegraition space was low, the model was reestimated with an unrestricted constant.

further attention in future research, for purposes of the present study, S2S estimation results were used, provided the results passed the tests for no residual autocorrelation; otherwise, RR estimator was used. Results of the estimation are reported in Table 3.

It should be pointed out that, although the table reports only the best models, exhaustive analysis, relying on all the various permutations of number of fundamentals, different estimation methods and time periods was undertaken. For two countries (Netherlands and Portugal) it was not possible to find any model passing the tests for adequate DGP representation. For seventeen other countries, the constructed models had plausible beta vector coefficients, a significant and negative alpha vector coefficient for exchange rate (thus ensuring correction of exchange rate to equilibrium as defined by beta vector), and uncorrelated residuals. All the models were tested for structural stability, using Johansen's recursive eigenvalues test (Hansen and Johansen, 1999) and bootstrapped Chow tests (Luetkepohl, 2004). Johansen's recursive eigenvalues tests were used to check the stability of the cointegration relations and for most of the countries the test statistics were comfortably above rejection level, except for Norway data in 2000. Indeed, closer inspection of the data clearly showed abnormal behaviour in the Norway variables after year 2000; accordingly, the Norway system was estimated only to 2000 Q1[28]. Bootstrapped Chow tests were used to check the stability of the other parameters in the models, and in most cases

[28] The reasons for this phenomenon in the Norway data is unclear; while this may warrant further investigation, as this study's focus is on all-countries analysis, the matter was not pursued further.

the stability hypothesis was not rejected. As for the moment the precise distribution of the tests statistics for cointegration models is not known[29], we were not able to reject the hypothesis of stability with certainty.

Looking at the table, it can be observed that all of the models display negative coefficients for the relative productivity proxy - in line with the well-established relationship between relative productivity and currency appreciation. Nevertheless, the magnitude of the estimated elasticities is large, in contrast with Alberola *et al.* (1999) study, where the relative productivity coefficient was very close to one, for all countries examined. Other studies assessing the Balassa-Samuelson effect found a much greater dispersion of estimated coefficients, although all were positive (e.g. Canzoneri *et al.*, 1999, Kohler, 2000) - this result being confirmed by the present study.

Results for the effect of net foreign assets on the other hand, were with both negative and positive coefficients observed. Given portfolio-balance considerations, it follows that increased net foreign debt should lead to currency appreciation, in turn implying that a negative beta coefficient for NFA variable should be observed. Alberola *et al.* (1999) found a uniformly negative relation, whereas many studies on Eastern and Central European countries found the effect of net foreign assets to be ambiguous (see, e.g. Egert, 2000). One possible explanation for this may be that capital inflows into productive sectors materialize in the form of productivity

[29] The last chapter discusses this issue to some extent.

growth, or, for this time period, the real appreciation and net capital inflows occurred simultaneously. Notwithstanding differences in the coefficients' signs it is evident that, from a statistical point of view, if there is a specific relation between exchange rates and net foreign assets varying across countries, this should not be read as an indicator of model misspecification, but rather deriving from the heterogeneity of countries comprising the sample.

As a final check, the signs of the alpha coefficients for exchange rate equation are all negative and significant, suggesting that the exchange rate does adjust back to equilibrium level as determined by the beta vector. The speed of adjustment varies considerably from country to country, however, the estimated half-lives of the shocks are all around one year (alpha coefficient of around -0.1 with quarterly data), with some countries showing longer half-lives (around two years for US, Sweden, Finland and Canada). The orthogonal complements of alpha vector appeared to be in line with those of other studies, showing two stochastic trends originating mainly from two fundamentals time series (in some rare cases, alternatively, from relative productivity proxy and exchange rate itself)[30].

Given the estimation results, it is possible to derive the BEER and PEER estimates of equilibrium exchange rate for each country, as explained in the foregoing methodological section (Chart 1).

---

[30] Detailed results of the estimation, together with model specifications, are available from the author upon request.

Visual inspection reveals satisfactory results: estimates track the real exchange rate acceptably. In certain cases, some persistent swings away from equilibrium trends can be observed, and a natural question to pose is whether introducing new fundamentals would improve the models in any way. Therefore, in line with cointegration analysis, the models with additional fundamentals proxying long interest rate differentials and relative fiscal balances, were estimated using the same model setup and methodology. The only difference with 3-VAR models estimation was that, for 5-dimensional models, given a large number of parameters to estimate, the subset restrictions were used on alpha vector and short run coefficients, in order to reduce the number of estimated parameters. Therefore, the models were estimated in two stages, beta vector estimated during the first stage, using the S2S estimator and short run parameters and alpha vector during the second stage, using the feasible GLS estimation method. Results of the estimation are reported in Table 4.

Once again, only the best model results are reported for each country, except for Sweden and Spain for which no satisfactory model could be found. As in the 3-VAR systems, there is a clear and unambiguous relationship between relative productivity and exchange rate. Results for net foreign assets are mixed, with most coefficients having a positive sign. As regards interest rate differentials, the uncovered interest rate (UIP) parity is rejected for all countries except France, Germany and Netherlands. The rejection of UIP is well documented in many studies for the sample countries in the 1980's. The relative fiscal balance variable has a negative

sign in most of the countries, implying that an increase in fiscal balance (or reduction of budget deficit) leads to currency appreciation. Alpha coefficients of the exchange rate equations are all negative and significant, showing that half-lives of the shocks are around a year (with the noticeable exception of Japan, where it is rather large for the 5-VAR system). As with 3-VAR systems, it is possible to derive the BEER and PEER estimates; these are shown in Chart 2.

Comparing Chart 2 with Chart 1, it is noticeable that 5-VAR models are more successful in explaining RER swings, even in the 70s, whereas most 3-VAR models could be estimated only by excluding this time period. Additionally, on average, the adjustment of the exchange rate to equilibrium (as shown by the magnitude of the alpha vector coefficient) is more rapid. As expected, PEER and BEER estimates for most countries are quite similar, indicating moderate misalignments of fundamentals from their equilibrium values. In the context of all these observations, it was decided to proceed with further analysis using 5-VAR systems only[31].

Looking at exchange rate misalignments in the fourth quarter of 1998, and in the latest time period (fourth quarter of 2003) of the study, it can be observed that, prior to Euro introduction, the member states' currencies were scattered around their equilibrium values, with deviations ranging from -2.6% for Spain to +3.6% for Germany (Chart 3).

[31] The only exceptions being Spain and Sweden, where no 5-VAR model could be constucted, and Japan where the 3-VAR system displayed stronger performance.

The Japanese yen was the most undervalued currency (-13.4%), followed by the Canadian dollar and Norwegian krone (-6.5% and -6.2% respectively), the US dollar being the most overvalued currency (+3.8%). Applying the Euro-area trade weights drawn from Maeso-Fernandez *et al.* (2001), it is possible to estimate synthetic effective Euro misalignment as the weighted average of member currencies' misalignments - at that time only 1.3%. Looking at misalignments in the fourth quarter of 2003, the picture changes dramatically. Most Euro-area currencies moved into overvaluation, with the Euro weighted average overvalued by 5.1%, whereas the US dollar was overvalued to an even greater extent (10.3%). The Japanese yen, Swiss and Belgian francs were the only currencies undervalued at that time. Chart 3, showing misalignments of RER in 2003 Q4, suggests that the over- and under-appreciations are far from cancelling each other out; yet visual inspection of currencies' misalignments over the whole period prove that not to be the case. Only at the end of the sample do a substantial majority of the currencies go into significant overvaluation. One of the main explanations for this is that the current sample omits developing economies who frequently keep their currencies undervalued in order to promote exports[32].

Using the above-mentioned trade weights for the Euro-area it is possible to construct a historic series of synthetic effective Euro misalignments for the whole time period: this is shown in Chart 4, together with US dollar RER misalignments.

[32] An obvious example here is Chinese renminbi. In 2002, China's merchandise exports were $US 326 billion, or 5% of total world trade, and estimates of the renminbi's undervaluation ranged from 45 to 20% (see Funke and Rahn, 2005, for a in-depth discussion, and their use of BEER/PEER approach).

The synthetic Euro demonstrates rather modest level of misalignment, with a maximum of 5.1% reached precisely during the last quarter of 2003, whereas US dollar behaviour is characterized by prolonged deviations from equilibrium values, finally reaching the maximum overvaluation during the first quarter of 2003, and coming down almost to equilibrium value later on.

Chart 5 shows the standard deviation of Euro-area members' currencies misalignment from equilibrium values across the whole time period studied. As expected, it shows substantial variability in the 1970's and at the time of ERM's breakdown, but reaching a minimum just prior to introduction of the single the European currency. Nevertheless, later developments reveal some increased variability, which could perhaps indicate that Euro introduction had not yet brought about convergence of member countries' economic processes.

2.5.3 Estimating bilateral exchange rates misalignments

So far, the analysis has focused on multilateral real equilibrium exchange rates. Yet certain researches have gone one step further, using algebraic decompositions first formalized by Alberola *et al.* (1999) to extract equilibrium bilateral nominal exchange rates. As this study encompasses 19 countries, representing a large share of overall world trade, here the same methodology can be applied to extract nominal exchange rate misalignments for each country and for the synthetic Euro currency in addition, assuming that misalignments of exchange rates of the rest of the world to

be zero[33]. Though only estimates of bilateral exchange rates of each currency against the US dollar are reported, the analysis can be extended to any numeraire currency[34]. The nominal rate of the Euro against the US dollar is estimated using corresponding estimated exchange rates of EMU countries' currencies weighted by proportion of overall Euro-area trade. Estimated equilibrium and actual exchange rates of major EMU countries and the synthetic Euro are shown in Chart 6.

These bilateral exchange rates broadly follow the same path against the US dollar, with Euro estimates showing that EMU was created at a time when Euro was modestly undervalued against the US dollar, and, that, for the next three years, depreciated even further away from equilibrium level. Nevertheless, breaking this trend, in 2002-2003 the Euro rapidly returned to equilibrium value.

Chart 6 also shows estimated equilibrium and actual exchange rates of countries outside EMU, demonstrating that, whereas the US dollar was overvalued, the level of overvaluation was diminishing during the last two years of the sample under analysis.

---

[33] A subsequent sensitivity analysis showed that as not being too restrictive an assumption, as its effect on bilateral exchange rates was minimal. This is partially explained by the fact that the trade weight matrix from Zanello and Desruelle, 1997 was, first, based on the trade manufacturing data for 21 developed countries only; and second, the data was for 1989-1991 period only. As a result, the trade weights for each country were rebased to add up to 100 per cent, thus omitting the share of trade with the countries not included in the table - countries, that in the last 15 years have increased their share of world trade considerably. Due to the absence of exact trade weights, this table was used, thus introducing a bias into the calculations of bilateral misalignments.

[34] As noted by Benassy-Quere *et al.* (2004), the choice of numeraire currency is not trivial, because when moving to bilateral rates, the numeraire's misalignments in effective terms will not be accounted for, consequently leading to biased results. One solution might be to use as numeraire the country with the smallest share of trade with the rest of the sample. Alternatively, one can take as numeraire the currency which has the minimum misalignment in effective terms.

2.6 Conclusion

In the face of substantial fluctuations in exchange rates in many developing and industrialized countries there has been sustained interest over recent decades, in establishing whether there is an equilibrium level at which the exchange rates would rest in the long run. While the PPP approach has always seemed a likely candidate in the determination of exchange rates, research findings demonstrated the speed of mean-reversion to be very slow.

Thus, the focus of interest shifted to the analysis of economic processes, affecting trade flows and balance of payments, as alternative explanatory variables potentially responsible for persistent structural deviations in RERs. Based on various theoretical models and empirical regularities observed in the data, researchers have failed to identify evidence supporting a single model of exchange rate behaviour valid for all countries and time periods. As a result, the attention of recent studies has been devoted to identification of economic fundamentals affecting exchange rates in the long run, and to the development of cointegration methods applicable to data of short time span.

Research on industrialised countries identified several economic fundamentals as the principal variables moving in line with exchange rates. First, differences in sector productivity between traded and nontraded goods, known as the Balassa-Samuelson effect, demonstrated to provide a partial explanation of the trend movements in long-run RERs. This effect was clearer for countries passing through the

initial stages of restructuring in industrial and export sectors - a condition which certainly applies to the transition economies of Eastern and Central Europe. Secondly, some evidence favoured the conclusion that interest rate differentials primarily reflect the effect of business cycles on exchange rates; though there is now ample evidence of this relationship, it is by nature of medium term, and applies to the analysis of short-to-medium run fluctuations of exchange rates. Additionally, evidence of convergence of interest rates makes this effect less important. Recent attempts to combine stock and flow variables, in order to analyze exchange rate behaviour, have identified a country's net foreign assets position as the most important variable affecting the capital account, by influencing the amount and direction of capital flows between the country and the world - which should exert pressure on exchange rate.

Using the BEER/PEER reduced model approach to exchange rates determination for 19 OECD countries, and for a time sample of 1974 - 2003, it was shown that the real effective exchange rates were indeed cointegrated with economic fundamentals for each country in the study; moreover, it was possible to estimate cointegrated VAR models for each country by using relative productivity and net foreign asset position fundamentals (3-VAR models), and adding real interest rate differentials and relative fiscal spending (5-VAR models). However, the results demonstrated that the level of heterogeneity across countries was quite large, and apart from relative productivity, other fundamentals' coefficients estimates were of different signs. This result contradicts findings by Alberola *et al.* (1999) who omitted to report statis-

tics to assess statistical significance of the estimates, perhaps indicating that their constructed models were inappropriate. Hansen and Roeger's study (2000) in which they were unable to replicate Alberola *et al.* (1999) results, perhaps serves further to confirm these findings.

Notwithstanding those issues, the estimates constructed from the models described tracked the real exchange rate well, and displayed a substantial level of adjustments of exchange rates to equilibrium trends (with half lives of around one year on average). Moreover, given that the data used to construct the fundamentals and forecasts for the following periods are readily available, it is possible to construct the forecasts of real effective and nominal bilateral exchange rates to be used in further economic analysis. Facing the model uncertainty surrounding statistical models of exchange rate determination, the results are still useful and might prove to be superior to black box or PPP exchange rate estimates, one further issue that warrants further attention in future studies.

Table 1. Testing for unit roots with ADF test

	REER		RPROD		NFA		RIR		FBAL	
	t-stat	lags	t-stat	lags	t-stat	lags	t-stat	lags	t-stat	lags
AUSTRIA	-2.18	1	-0.85	2	-2.19	2	-2.37	4	-2.98	4
BELGIUM	-1.62	1	-2.31	1	**-3.42**	5	-2.41	1	-0.23	5
CANADA	-2.18	3	-2.25	3	-1.61	3	**-3.27**	1	-1.15	1
DENMARK	-2.53	1	-2.23	0	-0.09	1	-1.28	4	**-3.09**	2
FINLAND	-1.99	1	-2.69	1	-0.79	3	-2.56	1	-2.86	1
FRANCE	-2.64	0	-2.39	1	-1.68	4	-2.41	2	-1.60	4
GERMANY	-2.67	1	-2.61	2	-1.97	3	**-3.00**	1	-1.87	3
GREECE	-2.10	0	-1.08	0	-2.22	4	-2.53	4	-1.57	3
IRELAND	-1.94	1	-1.14	1	-0.76	2	**-2.86**	1	-1.72	2
ITALY	-2.24	3	-1.60	2	-1.83	3	**-2.97**	4	-0.17	3
JAPAN	-2.11	1	-1.99	0	-2.15	3	**-3.14**	3	-1.22	2
NETHERLANDS	-1.86	1	-1.15	2	-1.21	3	-1.75	4	-1.92	4
NORWAY	**-2.99**	1	-1.49	1	-1.79	1	**-3.03**	4	-2.27	1
PORTUGAL	-1.06	3	-2.50	4	-2.09	2	**-2.93**	3	**-3.57**	3
SPAIN	-2.17	0	-0.67	0	-2.12	2	-2.84	1	-2.64	2
SWEDEN	-1.88	1	-1.56	1	-1.22	3	**-3.18**	4	-2.48	3
SWITZERLAND	**-3.38**	2	-3.27	5	-0.30	3	-2.44	1	-2.83	2
UK	-2.21	0	-2.11	2	-0.58	1	**-3.02**	4	-2.60	1
USA	-1.85	1	**3.12**	4	-3.16	2	**-3.34**	3	-2.06	3

NOTES:
5% critical value: -2.86 (-3.41 with trend); 1% critical value: -3.43 (3.96 with trend)
Series significant at 5% are in bold; only Portugal FBAL series is significant at 1%
Number of lags was determined using information criteria and ensuring non-autocorrelated residuals
For some RPROD and NFA series the ADF test included seasonal dummies
Models with trend are estimated for Switzerland RPROD, USA, Norway, Japan NFA series

Table 2. Systems cointegration tests: p-values

			3-VAR			5-VAR			
	Start	H0	Trace	SL-test	Lags	Start	Trace	SL-test	Lags
AUSTRIA	1974-Q1	r = 0	0.04	0.79	2	1974-Q1	0.00	0.00	2
		r = 1	0.23	0.75			0.01	0.08	
BELGIUM	1980-Q1	r = 0	0.03	0.62	2	1980-Q1	0.00	0.03	2
		r = 1	0.55	0.17			0.04	0.85	
CANADA[1]	1974-Q1	r = 0	0.00	0.00	4	1974-Q1	0.03	0.01	3
		r = 1	0.19	0.27			0.23	0.21	
DENMARK	1974-Q1	r = 0	0.00	0.01	1	1974-Q1	0.04	0.04	2
		r = 1	0.45	0.97			0.23	0.13	
FINLAND	1974-Q1	r = 0	0.05	0.10	2	1974-Q1	0.06	0.04	3
		r = 1	0.33	0.58			0.42	0.13	
FRANCE	1980-Q1	r = 0	0.08	0.28	2	1980-Q1	0.08	0.05	2
		r = 1	0.23	0.74			0.38	0.27	
GERMANY	1974-Q1	r = 0	0.03	0.29	2	1974-Q1	0.00	0.01	2
		r = 1	0.36	0.86			0.38	0.61	
GREECE	1985-Q1	r = 0	0.00	0.14	1	1974-Q1	0.00	0.07	2
		r = 1	0.14	0.83			0.07	0.04	
IRELAND	1980-Q1	r = 0	0.00	0.00	1	1974-Q1	0.02	0.02	2
		r = 1	0.00	0.63			0.10	0.29	
ITALY	1980-Q1	r = 0	0.04	0.03	2	1974-Q1	0.01	0.01	4
		r = 1	0.37	0.84			0.52	0.73	
JAPAN	1980-Q1	r = 0	0.00	0.00	1	1974-Q1	0.00	0.00	2
		r = 1	0.21	0.10			0.05	0.35	
NETHERLANDS	1974-Q1	r = 0	0.06	0.05	1	1974-Q1	0.00	0.00	2
		r = 1	0.35	0.79			0.02	0.09	
NORWAY	1974-Q1	r = 0	0.00	0.00	1	1974-Q1	0.16	0.02	2
		r = 1	0.85	0.57			0.51	0.14	
PORTUGAL	1974-Q1	r = 0	**0.72**	**0.87**	2	1974-Q1	0.06	0.31	2
		r = 1	**0.97**	**0.96**			0.85	0.93	
SPAIN	1974-Q1	r = 0	0.05	0.22	2	1974-Q1	**0.45**	**0.35**	2
		r = 1	0.65	0.40			**0.60**	**0.63**	
SWEDEN	1974-Q1	r = 0	0.00	0.00	2	1974-Q1	**0.00**	**0.00**	2
		r = 1	0.09	0.01			**0.00**	**0.00**	
SWITZERLAND	1974-Q1	r = 0	0.01	0.04	4	1974-Q1	0.01	0.08	3
		r = 1	0.08	0.17			0.10	0.90	
UK	1980-Q1	r = 0	0.12	0.01	2	1980-Q1	0.04	0.01	2
		r = 1	0.51	0.20			0.19	0.08	
USA	1974-Q1	r = 0	0.00	0.04	3	1974-Q1	0.00	0.02	2
		r = 1	0.01	0.15			0.00	0.09	

NOTES:

Table reports p-values of the cointegration tests with null hypotheses that rank of the systems is at most zero or one.
3-VAR system is {REER, RPROD, NFA}, 5-VAR is {REER, RPROD, NFA, RIR, RFBAL}.
Cointegration tests: JOH-test - Johansen trace test, SL-test - Saikkonen and Luetkepohl test.
No NFA is Canada 3-VAR, Ireland and Japan 5-VARs, No RFBAL in Portugal 5-VAR.
No RIR in Canada and Norway 5-VAR systems.
Time series run from 1974-Q1 or 1980-Q1 to 2003-Q4

Table 3. Estimation results of 3-VAR systems

	Beta vector			Alpha vector			Residual diagnostics				Lags	Sample start	Estimation method
	REER	RPROD	NFA	REER	RPROD	NFA	Norm	Skew	Port (16)	BG (4)			
AUSTRIA	1	-0.538	-0.100	-0.118	0.124	-0.111	0.36	0.11	0.71	0.03	1	1974Q1	S2S
		-7.2	-3.6	-2.9	2.8	-1.4							
BELGIUM	1	-0.314	-0.160	-0.145	-0.037	-0.014	0.00	0.65	0.96	0.10	1	1980Q1	JOH
		-2.1	-1.3	-4.2	-1.0	-1.4							
CANADA	1	-1.821	-	-0.055	0.073	-	0.22	0.74	0.93	0.82	3	1974Q1	JOH
		-19.7	-	-1.2	3.9	-							
DENMARK	1	-0.797	-0.270	-0.091	0.043	0.078	0.02	0.56	0.72	0.14	1	1974Q1	S2S
		-2.7	-4.1	-1.7	1.5	1.6							
FINLAND	1	-0.954	0.294	-0.040	0.016	-0.094	0.00	0.01	0.80	0.03	1	1974Q1	S2S
		-3.0	3.2	-2.3	1.8	-6.5							
FRANCE	1	-0.291	0.466	-0.165	-0.004	-0.021	0.06	0.18	0.23	0.09	3	1980Q1	JOH
		-1.4	3.4	-3.8	-0.1	-2.1							
GERMANY	1	-1.272	-0.224	-0.078	0.041	-0.003	0.03	0.27	0.52	0.04	2	1974Q1	S2S
		-7.7	-1.1	-2.1	3.2	-0.3							
GREECE	1	-1.179	0.052	-0.086	-0.029	-0.035	0.64	0.90	0.31	0.72	0	1985Q1	S2S
		-3.4	0.3	-3.6	-0.9	-0.6							
IRELAND	1	-0.767	0.143	-0.212	-0.107	-0.185	0.50	0.50	0.93	0.18	0	1980Q1	S2S
		-6.1	5.4	-3.6	-3.2	-2.7							
ITALY	1	-2.501	-0.165	-0.055	0.046	-0.027	0.20	0.15	0.24	0.03	1	1980Q1	S2S
		-7.2	-0.6	-1.3	2.5	-2.6							
JAPAN	1	-7.367	1.104	-0.109	0.020	-0.004	0.00	0.02	0.99	0.70	3	1980Q1	S2S
		-11.6	5.1	-2.1	3.3	-0.3							
NORWAY	1	-0.365	0.111	-0.104	0.004	-0.105	0.09	0.20	0.27	0.01	1	1974Q1	S2S
		-1.7	1.7	-2.0	0.2	-2.8							
SPAIN	1	-0.413	-0.920	-0.130	-0.023	-0.016	0.00	0.01	0.95	0.21	1	1974Q1	S2S
		-2.3	-3.2	-3.4	-1.6	-1.0							
SWEDEN	1	-1.442	0.343	-0.053	0.032	-0.047	0.02	0.08	0.40	0.32	3	1974Q1	JOH
		-7.6	2.9	-1.5	2.4	-5.4							
SWITZERLAND	1	-1.320	0.084	-0.264	-0.013	0.035	0.05	0.07	1.00	0.40	3	1974Q1	S2S
		-9.4	3.0	-4.5	-0.9	0.4							
UK	1	-1.588	0.503	-0.112	0.043	-0.005	0.00	0.07	0.80	0.01	1	1980Q1	S2S
		-3.6	2.6	-2.3	3.5	-0.6							
USA	1	-2.441	-0.265	-0.044	0.005	-0.016	0.20	0.32	1.00	0.06	3	1974Q1	S2S
		-3.0	-4.9	-1.2	0.7	-5.9							

NOTES:

Table reports estimated beta and alfa vectors with t-values underneath, models residuals diagnostics and estimation method selected

In residual diagnostics NORM stands for joint test for normality (Luetkepohl, 1993), SKEW - for normality of the skewness only

Port (16) is Portmanteu test for residual autocorrelation with 16 lags, BG(4) is Breusch-Godfrey test for autocorrelation with 4 lags.

Estimation was done both using Johansen method (JOH) and Simple-Two-Stage estimation method (S2S) and the better model was reported.

No adequate model was found for Netherlands and Portugal

Table 4. Estimation results of 5-VAR systems

	Beta vector					Alpha vector					Residual diagnostics				Lags	Sample start
	REER	RPROD	NFA	RIR	FBAL	REER	RPROD	NFA	RIR	FBAL	Norm	Skew	Port (16)	BG (4)		
AUSTRIA	1	-0.448	-0.092	-0.004	0.009	-0.150	-	-	-	-3.531	0.97	0.97	0.27	0.01	1	1974Q1
		-5.5	-3.5	-1.1	3.3	-4.9	-	-	-	-2.8						
BELGIUM	1	-0.411	0.039	-0.020	-0.003	-0.279	-	-0.043	-	-	0.00	0.04	0.74	0.22	1	1980Q1
		-6.0	0.7	-7.5	-2.9	-6.2	-	-2.8	-	-						
CANADA	1	-1.491	-0.077	-	0.001	-0.081	0.048	-0.045	-	-7.089	0.06	0.32	0.99	0.78	3	1974Q1
		-8.4	-0.6	-	0.3	-2.1	3.1	-2.8	-	-3.7						
DENMARK	1	-0.695	-0.244	-0.006	-0.007	-0.189	0.002	-	7.833	-	0.01	0.85	0.06	0.27	2	1974Q1
		-3.0	-4.2	-2.6	-2.7	-4.3	2.3	-	2.7	-						
FINLAND	1	-1.092	0.362	-0.014	-0.019	-0.055	0.016	-0.144	-	-	0.61	0.62	0.63	0.10	2	1974Q1
		-2.9	5.0	-2.4	-5.6	-3.2	2.0	-10.2	-	-						
FRANCE	1	-0.689	0.670	0.016	0.006	-0.126	-	-0.016	-1.967	-	0.10	0.88	0.59	0.05	1	1980Q1
		-3.0	4.5	1.7	1.4	-4.8	-	-2.0	-2.0	-						
GERMANY	1	-1.725	-0.380	0.016	0.010	-0.065	0.066	-	-	-	0.00	0.02	0.25	0.03	2	1974Q1
		-7.5	-2.3	2.3	2.1	-2.4	6.2	-	-	-						
GREECE	1	-0.104	0.186	-0.004	-0.010	-0.065	-	-0.096	4.097	1.887	0.00	0.52	0.87	0.04	1	1974Q1
		-0.6	2.0	-1.4	-3.7	-2.3	-	-3.0	2.1	3.1						
IRELAND	1	-0.373	-	-0.006	0.001	-0.128	-	-	3.721	-	0.14	0.52	0.21	0.17	2	1974Q1
		-3.1	-	-1.6	0.6	-5.2	-	-	2.8	-						
ITALY	1	-1.385	-1.044	-0.015	0.008	-0.174	0.042	-	11.070	-	0.00	0.69	0.03	0.05	2	1974Q1
		-6.5	-5.2	-4.2	4.5	-4.7	3.3	-	6.0	-						
JAPAN	1	-3.610	-	-0.034	-0.011	-0.037	0.032	-	0.905	-	0.03	0.07	0.06	0.21	1	1974Q1
		-9.5	-	-2.9	-2.1	-1.7	7.7	-	2.3	-						
NETHERLANDS	1	-0.198	0.314	0.013	-0.020	-0.090	-	-0.083	-	2.206	0.14	0.23	0.02	0.04	1	1974Q1
		-1.4	3.4	2.0	-5.0	-2.9	-	-2.0	-	2.5						
NORWAY	1	-0.325	0.129	-	-0.003	-0.145	-	-0.097	-	-	0.00	0.12	0.59	0.07	1	1974Q1
		-2.0	2.8	-	-1.5	-3.2	-	-2.8	-	-						
PORTUGAL	1	-0.991	0.359	-0.032	-	-0.045	-0.017	-	5.087	-	0.00	0.03	0.01	0.01	2	1974Q1
		-2.0	2.7	-7.2	-	-4.2	-1.9	-	6.6							
SWITZERLAND	1	-1.865	0.179	-0.015	-0.018	-0.308	-0.001	0.002	9.629	-	0.18	0.97	0.93	0.16	3	1974Q1
		-16.0	8.3	-6.6	-5.2	-7.4	-3.4	2.0	4.5	-						
UK	1	-0.974	0.528	0.002	-0.026	-0.138	0.023	0.018	6.060	2.257	0.00	0.04	0.05	0.19	1	1980Q1
		-3.0	3.8	0.3	-3.9	-3.1	1.8	2.1	4.3	1.5						
USA	1	-0.736	-0.174	-0.027	-	-0.055	-0.010	-0.013	2.550	-	0.00	0.25	0.99	0.67	3	1974Q1
		-0.8	-2.5	-2.3	-	-3.1	-1.6	-5.2	4.3	-						

NOTES:

Table reports estimated beta and alfa vectors with t-values underneath, models residuals diagnostics and estimation method selected In residual diagnostics NORM stands for joint test for normality (Luetkepohl, 1993), SKEW - for normality of the skewness only. Port (16) is Portmanteu test for residual autocorrelation with 16 lags, BG(4) is Breusch-Godfrey test for autocorrelation with 4 lags Estimation was done using two stage feasible GLS method. No adequate model was found for Spain and Sweden. Due to the structural break the model for Norway was estimated on 1974Q1 - 2000Q1 sample only.

Chart 1. Estimation results for 3 VAR models.

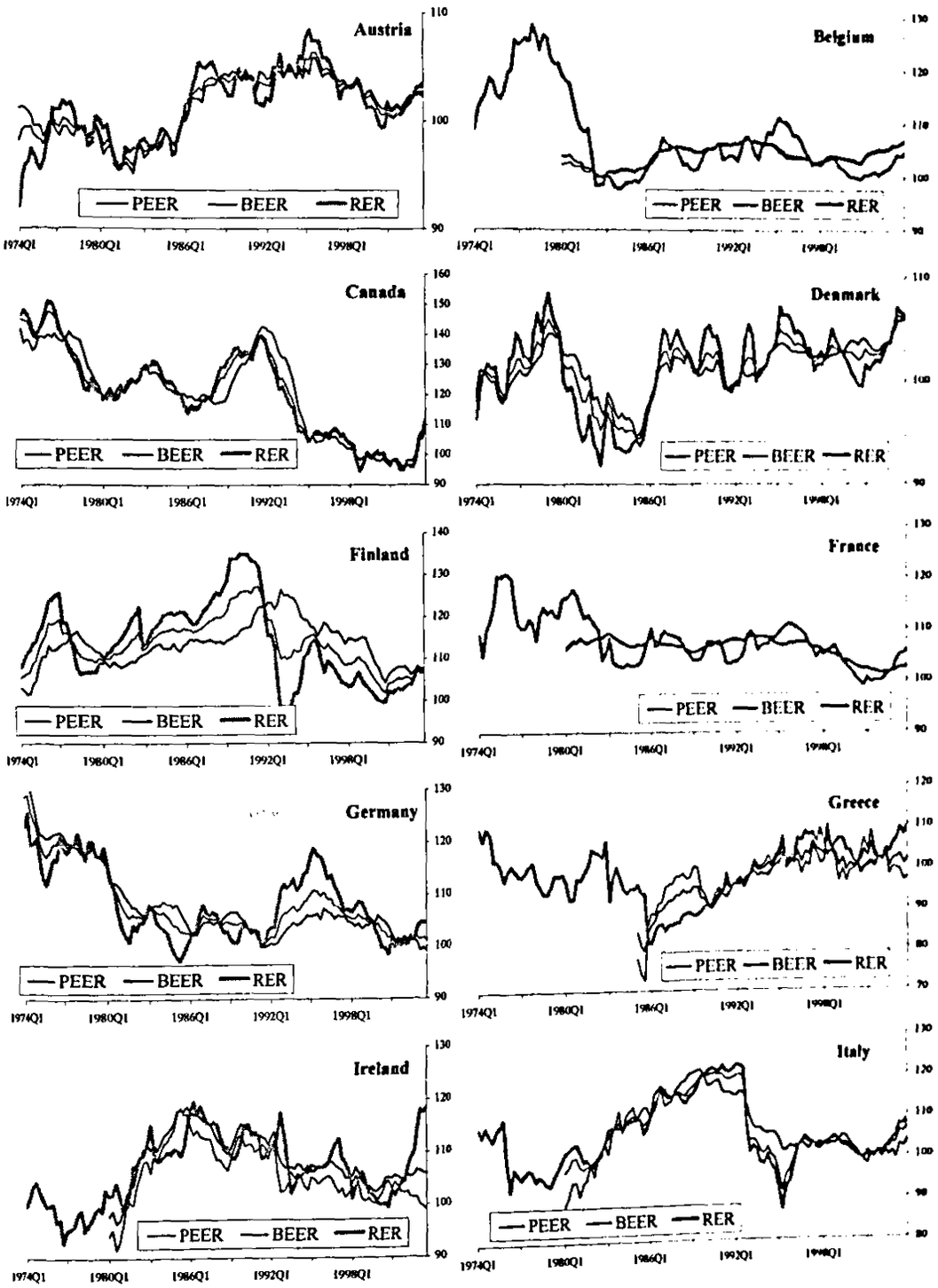


Chart 1 (continued). Estimation results for 3 VAR models.

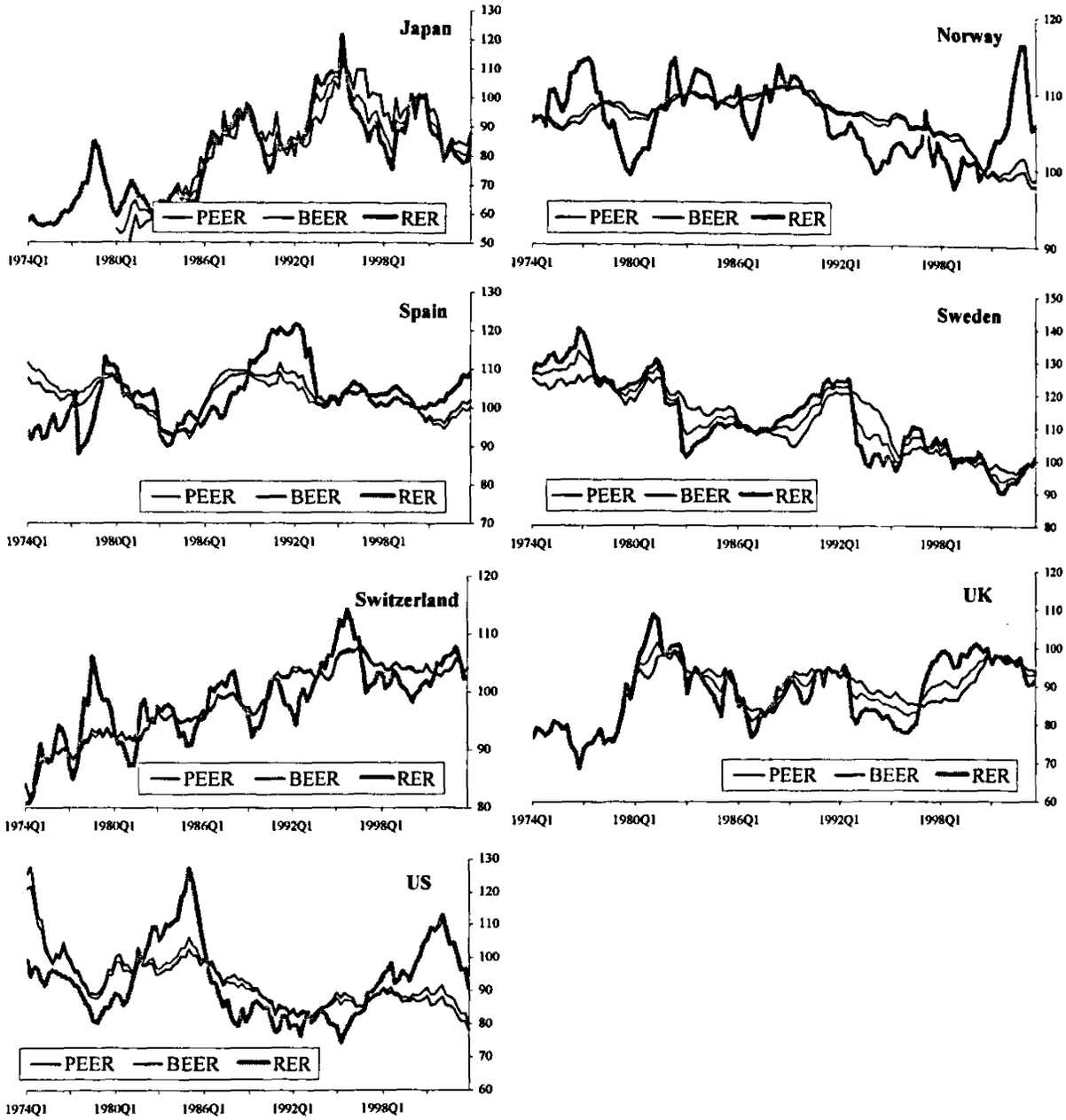


Chart 2. Estimation results for 5 VAR models.

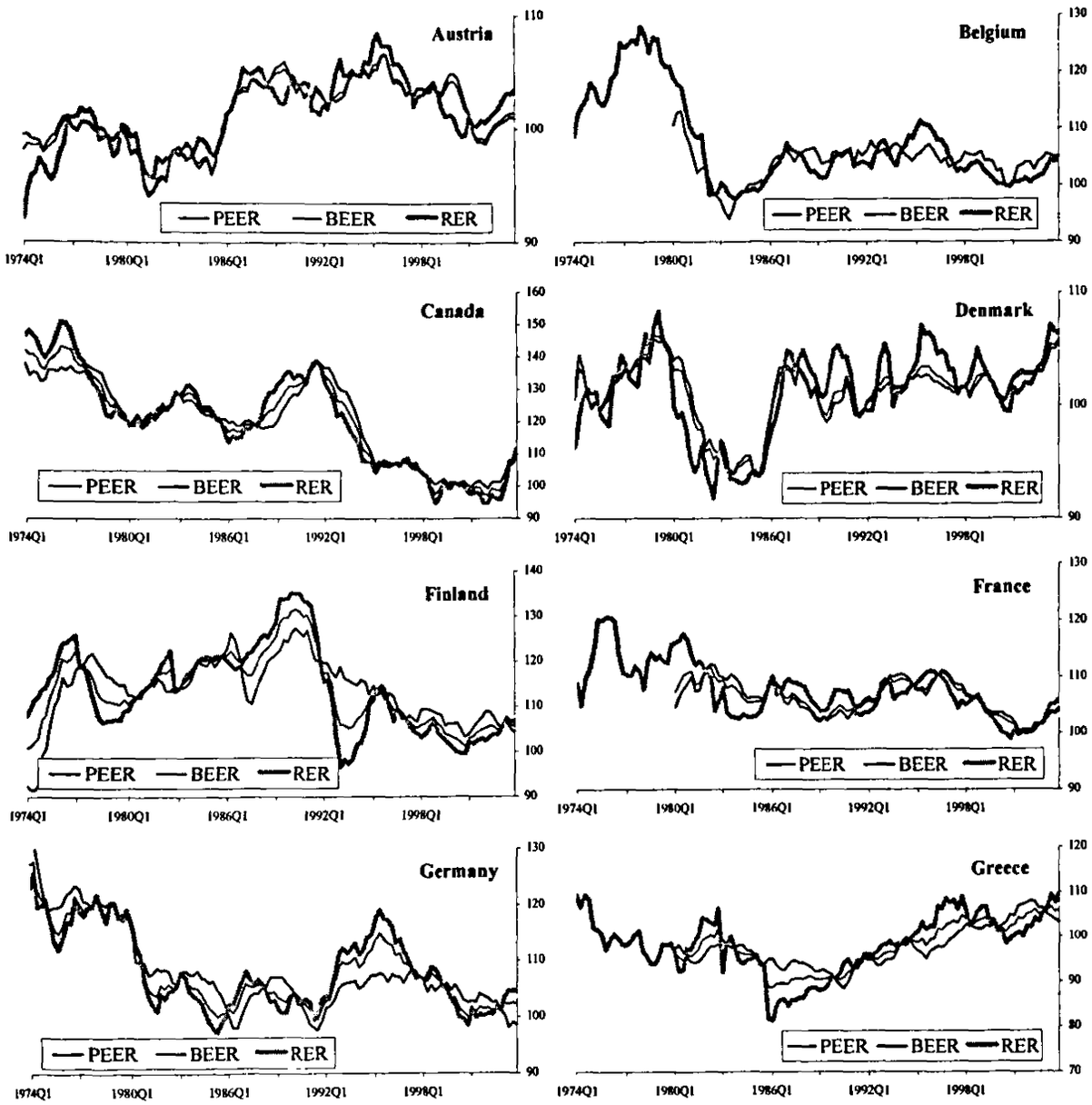


Chart 2 (continued). Estimation results for 5 VAR models.

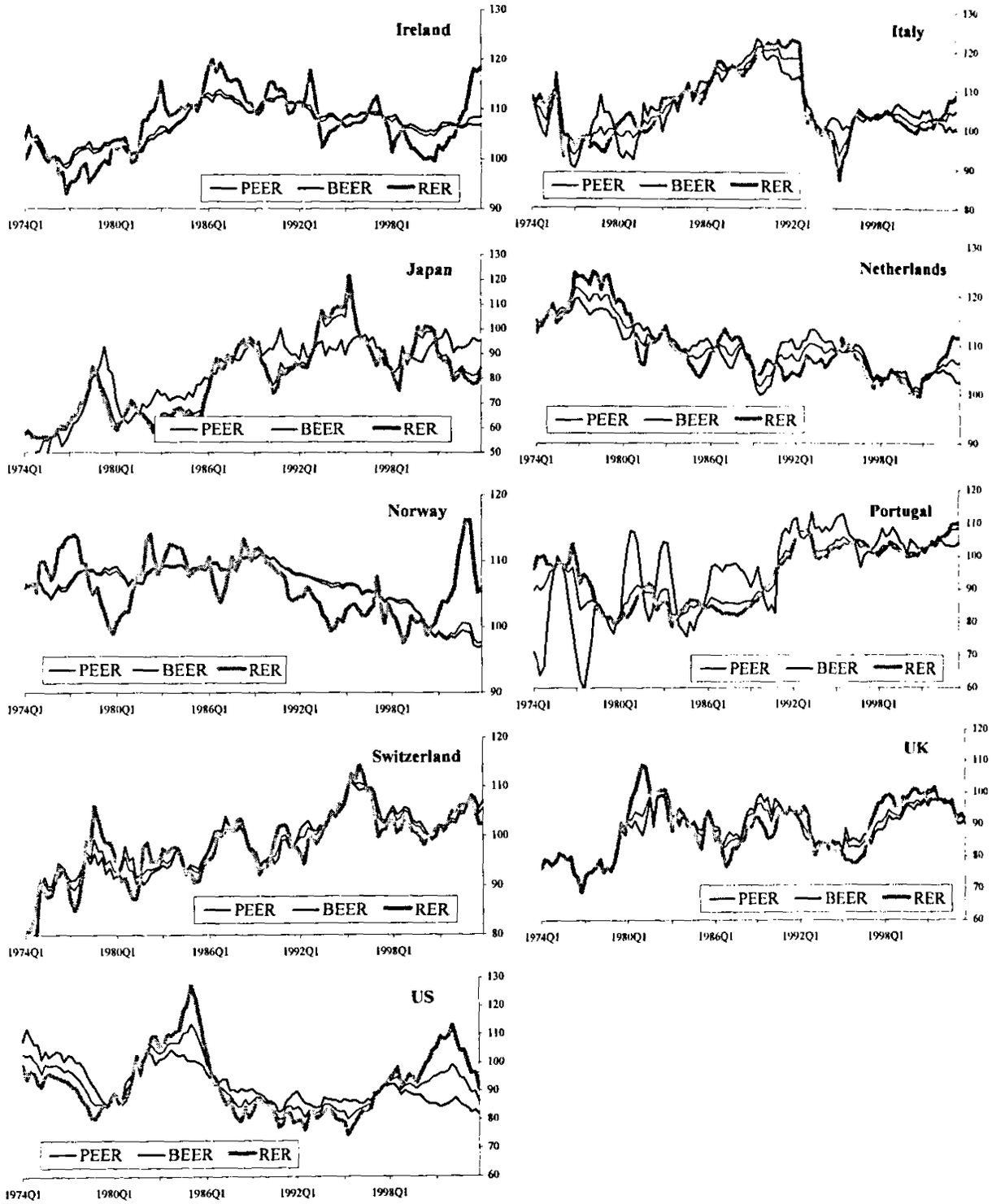
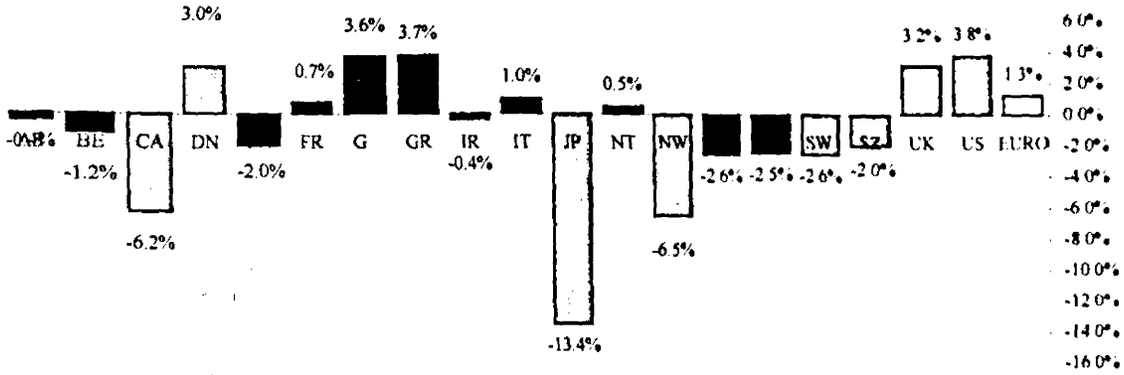


Chart 3. RER misalignments from BEER models estimation

RER misalignments as at 1998.4
(BEER models)



RER misalignments as at 2003.4
(BEER models)

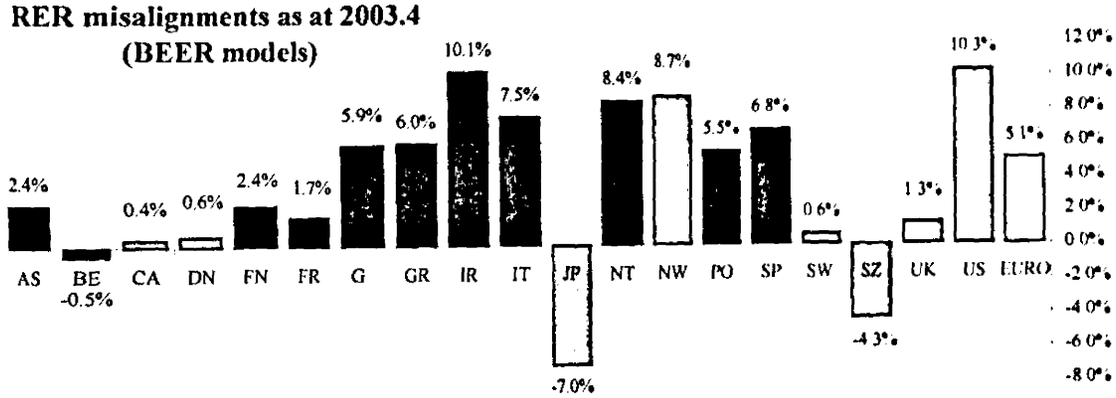


Chart 4. Synthetic Euro and US dollar RER misalignments (BEER models)

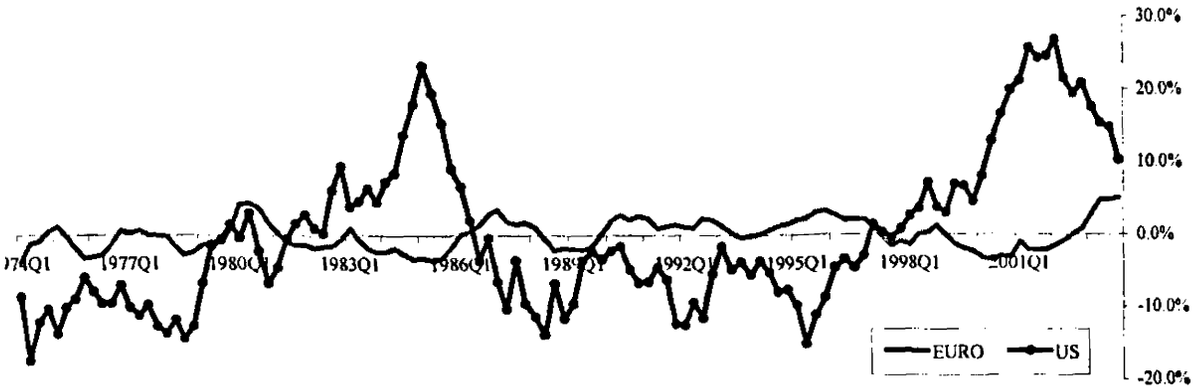


Chart 5. Euro-area members' currencies variability.
(standard deviation of EMU members' currencies misalignments from BEER estimates)

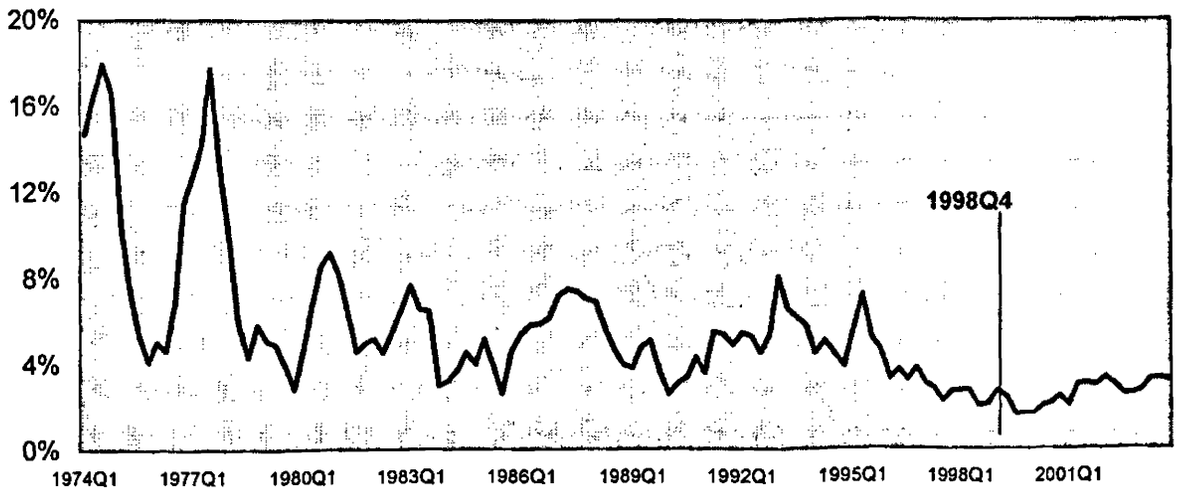


Chart 6. Estimated bilateral exchange rates against US dollar

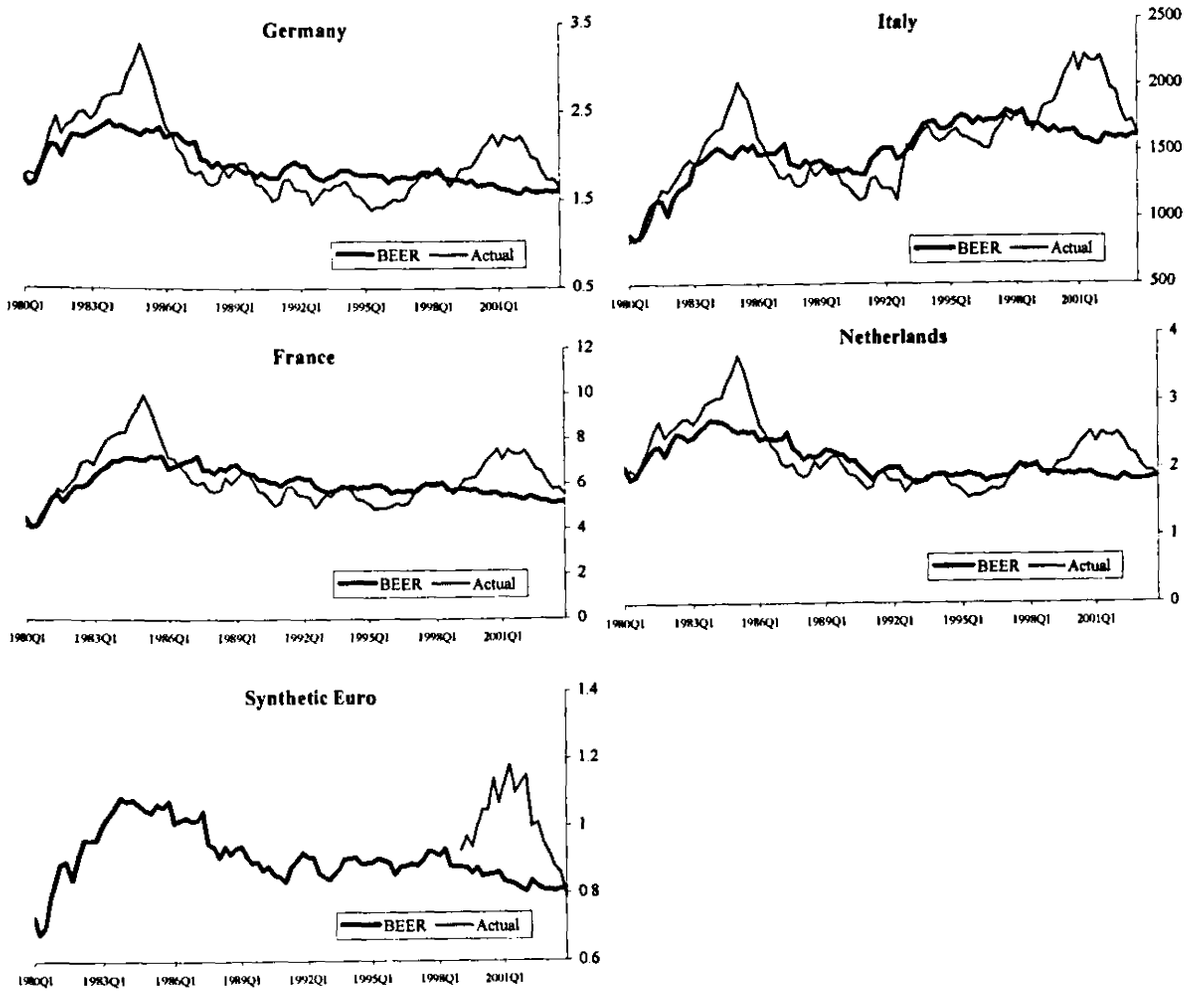
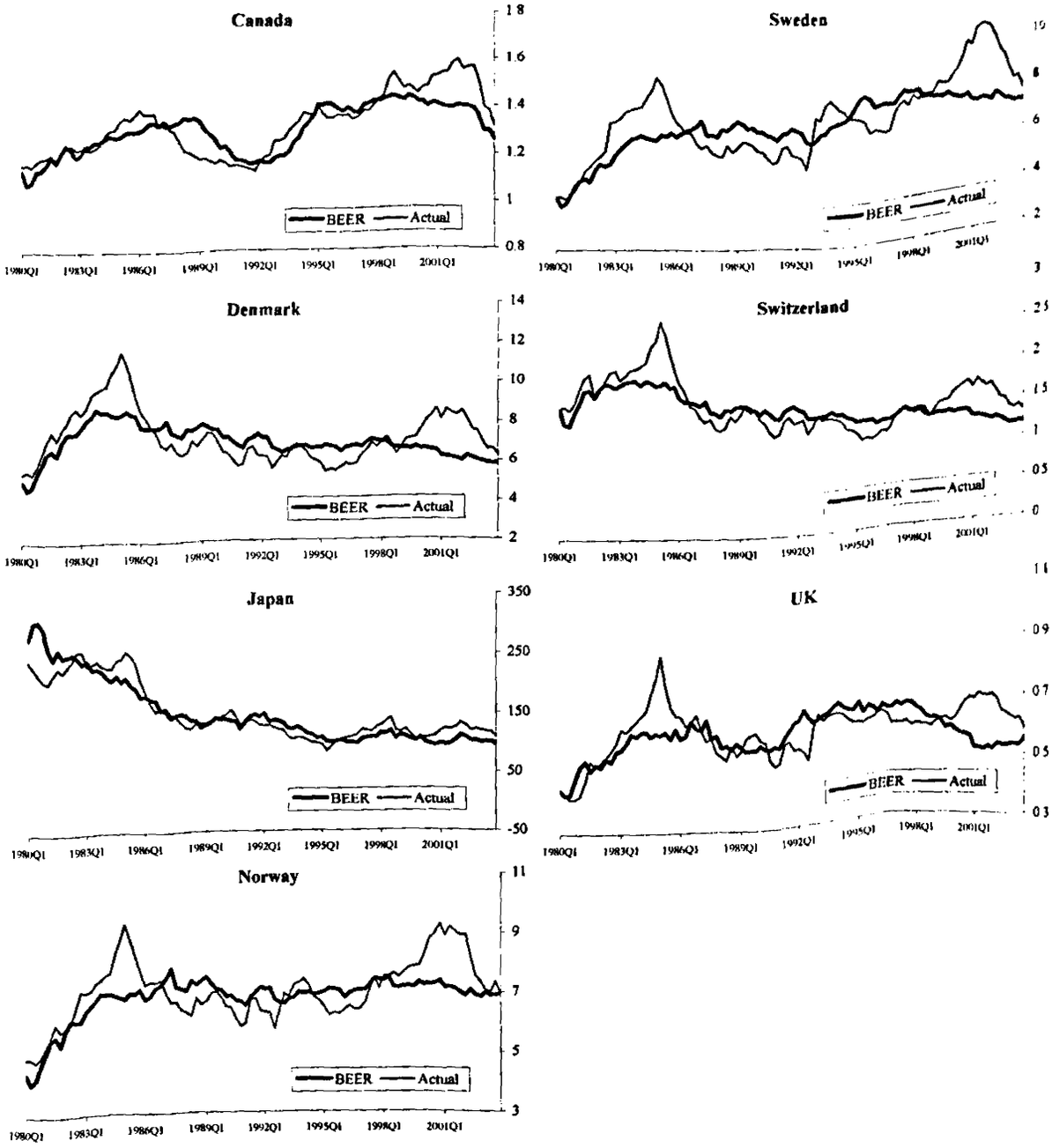


Chart 6 (continued). Estimated bilateral exchange rates against US dollar.



References

[1] Alberola, E., Cervero, L., Ubide, A. and Lopez, H. (1999), "Global Equilibrium Exchange Rates: Euro, Dollar, "Ins", "Outs", and Other Major Currencies in a Panel Cointegration Framework", IMF Working Paper, 99/175.

[2] Balassa, B. (1964), "The Purchasing Power Doctrine: A Reappraisal", *Journal of Political Economy*, 72, 584-96.

[3] Banerjee, A., Marcellino, M., Osbat, C., (2002), "Testing for PPP: Should We Use Panel Methods?," Royal Economic Society Annual Conference 2002 13, Royal Economic Society.

[4] Banerjee, A., Marcellino, M., Osbat, C., (2004), "Some Cautions on the Use of Panel Methods for Integrated Series of Macroeconomic Data," *Econometrics Journal*, Royal Economic Society, vol. 7(2), pages 322-340.

[5] Baxter, M. (1994), "Real Exchange Rates and Real Interest Differentials, Have we Missed the Business-Cycle Relationship? ", *Journal of Monetary Economics* 33, 5-37.

[6] Begnassy-Quere, A., Duran-Vigneron, P., Lahreche-Revil, P., Mignon, V. (2004), in C.F. Bergsten and J. Williamson eds., Dollar Adjustment: How Far? Against What?. Institute for International Economics special report 17, Washington D.C.

[7] Broner, F., N. Loayza, and J. Lopez, (1997), "Misalignment and Fundamentals: Equilibrium Exchange Rates in Some Latin American Countries", unpublished draft, The World Bank, November.

[8] Bruegmann, R., Luetkepohl, H. (2004), "Practical Problems with Reduced Rank ML Estimators for Cointegration Parameters and a Simple Alternative", mimeo.

[9] Burda, M.C. and Wyplosz, C. (1997), Macroeconomics - European Text. Oxford: Oxford University Press.

[10] Canzoneri, M. B., Cumby, R.E. and Diba, B. (1999), "Relative Labour Productivity and the Real Exchange Rate in the Long Run: Evidence for a Panel

of OECD Countries", Journal of International Economics, Vol. 47, pages 245-266.

[11] Cassel, G. (1921), The World's Money Problems. New York: E. P. Dutton and Co.

[12] Cheung, Y., Lai, K.S. (1998), "On Cross-country Differences in the Persistence of Real Exchange Rates", unpublished.

[13] Chinn, M., (1996), "Whither the Yen? Implications of an Intertemporal Model of the Yen/Dollar Rate," Mimeo.

[14] Chinn, M., (1999), "Productivity, Government Spending and the Real Exchange Rate: Evidence for OECD Countries", in R. MacDonald and J.L. Stein (eds) Equilibrium Exchange rates, Kluwer Academic Publisher, UK, 163-190.

[15] Chi-Wa, Y.(1998), "A Comment in "Changes in Exchange Rates in Rapidly Developing Countries", NBER - East Asia Seminar on Economics, Volume 7, Chicago: The University of Chicago Press.

[16] Clark, P., MacDonald, R. (1998), "Exchange Rates and Economic Fundamentals: A Methodological Comparison of BEERs and FEERs," IMF Working Paper, May.

[17] Clostermann, J. and Schnatz, B. (2000), "The Determinants of the Euro-Dollar Exchange Rate," Discussion paper 2/00: ERG of DB.

[18] Devereux, J. and M. Connolly (1996), "Commercial Policy, the Terms of Trade and the Real Exchange Rate Revisited", *Journal of Development Economics*, Vol 50: 81-99, 1996.

[19] Dornbush, R. (1976), "Expectations and Exchange Rate Dynamics," *Journal of Political Economy*, 84. p. 1161-176.

[20] Driver, R. L. and Wren-Lewis, S., (1998), "Real Exchange Rates for the Year 2000", Institute for International Economics, Washington DC: Policy Analyses in International Economics No 54.

[21] Driver, R. L. and Westaway, P.F. (2004), "Concepts of Equilibrium Exchange Rates", Working Paper 248, Bank of England.

[22] Edwards, S. and Miguel Savastano (1999), "Exchange Rate in Emerging Economies: What Do We Know, What Do We Need to Know", NBER Working Paper 7228.

[23] Egert, B. (2003), " Assessing Equilibrium Exchange Rates in CEE Acceding Countries: Can We Have DEER with BEER without FEER? A Critical Survey of the Literature", Focus on Transition, Oesterreichische Nationalbank.

[24] Engle, R. F. and Granger, C. W. J. (1987) "Cointegration and Error Correction: Representation, Estimation and Testing", *Econometrica*, 55, 251-76.

[25] Faruqee, Hamid, (1995), "Long Run Determinants of the Real Exchange Rate: A Stock-Flow Perspective," IMF Staff Papers 42(1) (March): 80-107.

[26] Frankel, J. and Rose, A. (1995), "Empirical Research on Nominal Exchange Rates," In Handbook of International Economics (eds. G. Grossman and K. Rogoff), pp. 1689-1729. Amsterdam: Elsevier Science Publishers B.V.

[27] Frankel, J., Rose, A. (1996), "A Panel Project on Purchasing Power Parity: Mean Reversion within and between Countries", *Journal of International Economics*, Vol. 40, 209-224.

[28] Frenkel, J. A. (1976), "A Monetary Approach to the Exchange Rate: Doctrinal Aspects and Empirical Evidence," *Scandinavian Journal of Economics*, vol. 78, no. 2, May, pp. 200-224.

[29] Frenkel, J. A. and M. Mussa (1988), "Exchange Rates and the Balance of Payments", in R. Jones and P. Kenen, eds, Handbook of International Economics, 2, Elsevier Science.

[30] Froot, K. A. and K. Rogoff (1991), "The EMS, the EMU, and the Transition to a Common Currency," in: O. J. Blanchard and S. Fischer, eds., 1991 NBER Macroeconomic Annual (The MIT Press, Cambridge, MA) 269-317.

[31] Froot K. A. and K. Rogoff (1995), "Perspectives on PPP and Long-Run Real Exchange Rates", In Handbook of International Economics (eds. G. Grossman and K. Rogoff), Amsterdam: Elsevier Science Publishers B.V.

[32] Funke, M. and Rahn, J. (2005), "Just How Undervalued is the Chinese Renminbi?" , The World Economy, Vol. 28, No. 4, pp. 465-489.

[33] Geweke J and Feige F. (1979), "Some Joint Tests of Efficiency or Markets for Foreign Exchange", *The Review of Economics and Statistics*, Vol 61, no3, 334-42.

[34] Gonzalo, J. and C. Granger, (1995), "Estimation of Common Long-memory Components in Cointegrated Systems", *Journal of Business and Economic Statistics* , 13, 1-9.

[35] Groen, J. J., and F. R. Kleibergen, (1999), "Likelihood-Based Cointegration Analysis in Panels of Vector Error Correction Models", Tinbergen Institute Discussion Paper TI 99- 055/4.

[36] Hakkio, C. S. (1981), "Expectations and the Forward Exchange Rate", *International Economic Review*, 22, 663-78.

[37] Halpern, L. and C. Wyplosz (1997), "Equilibrium Exchange Rates in Transition Economies", IMF Staff Papers Vol.44, No4, IMF.

[38] Halpern, L. , Wyplosz, C. (2001), "Economic Transformation and Real Exchange Rates in the 2000s: The Balassa-Samuelson Connection", unpublished.

[39] Hansen, H., Johansen, S. (1999), "Some Tests for Parameter Constancy in Cointegrated VAR-models", Econometrics Journal 2, 306-333.

[40] Hansen, J. and Roeger, W. (2000), "Estimation of Real Equilibrium Exchange Rates", European Economy - Economic Papers 144, Commission of the EC, Directorate-General for Economic and Financial Affairs (DG ECFIN).

[41] Hendry, D. F. and K. Juselius (2001), "Explaining Cointegration Analysis: Part II", unpublished.

[42] Hsieh, D. A. (1984), "Tests of Rational Expectations and no Risk Premium in Forward Exchange Markets", *Journal of International Economics*, 17, 174-84.

[43] Im, K., Pesaran, M. and Y. Shin (1997), "Testing for Unit Roots in Heterogeneous Panels", Department of Applied Economics, University of Cambridge, December.

[44] Isard, P and H. Faruqee (1998), "Exchange Rate Assessment: Extensions of the Macroeconomic Balance Approach," IMF Occasional Paper 167, Washington DC.

[45] Ito, T. P. Isard, S. Symanski and T. Bayoumi (1996), "Exchange Rate Movements and Their Impact on Trade and Investment in the APEC Region", Occasional Paper 145, IMF.

[46] Johansen, S. (1988), "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control* 12, 231-254.

[47] Johansen, S. (1996). Likelihood-based Inference in Cointegrated Vector Autoregressive Models. Oxford: Oxford University Press.

[48] Johansen, S. and K. Juselius (1992), "Testing Structural Hypotheses in a Multivariate Cointegration Analysis of the PPP and the UIP for UK", *Journal of Econometrics* 53, 211-244.

[49] Johansen, S (2001), " Common Trends in the Cointegrated VAR model", unpublished.

[50] Kao, C. (1999), "Spurious Regression and Residual-Based Tests for Cointegration in Panel Data", *Journal of Econometrics*, 90, 1-44.

[51] Kao, C. and Chiang, M.-H., (1999), "On the Estimation and Inference of a Cointegrated Regression in Panel Data", Working Paper, Centre for Policy Research, Syracuse University.

[52] Kazaks, M. (2000), "Real Exchange Rate Appreciation and Loss of Competitiveness. Case of Latvia", mimeo.

[53] Kohler, M. (2000), " The Balassa-Samuelson effect and Monetary Targets", in Mahadeva, L. and Sterne, G. (eds) Monetary policy framework in a global context, Routledge.

[54] Kopits, G. (1999), "Implications of EMU for Exchange Rate Policy in Central and Eastern Europe", IMF Working Paper WP/99/9.

[55] Lane, P., Milesi-Ferretti, G. M., (2004), "The Transfer Problem Revisited: Net Foreign Assets and Real Exchange Rates," The Review of Economics and Statistics, MIT Press, vol. 86(4), pages 841-857.

[56] Larsson, R., Lyhagen, J., (1999), "Likelihood-Based Inference in Multivariate Panel Cointegration Models", Working Paper Series in Economics and Finance, No. 331.

[57] Larsson, R., Lyhagen, J., (2000), "Testing for Common Cointegrating Rank in Dynamic Panels", Working Paper Series in Economics and Finance, No. 378.

[58] Larsson, R., Lyhagen, J. and Lothgren, M., (1998), "Likelihood-Based Cointegration Tests in Heterogeneous Panels", Working Paper No. 250, Stockholm School of Economics.

[59] Levin, Andrew, and Chien-Fu Lin (1993), "Unit Roots in Panel Data: Asymptotic and Finite Sample Properties," Discussion Paper 93-56 San Diego: University of California at San Diego.

[60] Liu, P. (1992), "Purchasing Power Parity in Latin America: A Cointegration Analysis", Weltwirtschaftliches Archiv, Vol. 128, No. 4: 662-680.

[61] Luetkepohl, H., Kraetzig, M. (2004). Applied Time Series Econometrics, Cambridge University Press.

[62] Luetkepohl, H., Saikkonen,P.,(2000), "Testing for the Cointegrating Rank of a VAR Process with an Intercept", Econometric Theory , 16 373-406.

[63] MacDonald, R. (1995), "Long-Run Exchange Rate Modeling: A Survey of the Recent Evidence," IMF Staff Papers, Vol 42, No 3, 437-489.

[64] MacDonald, R. (1997), "What Determines Real Exchange Rates: The Long and Short of It," IMF Working Paper 97/21.

[65] MacDonald, R. (2000), "Filtering the BEER: A Permanent and Transitory Decomposition," IMF Working Paper 00/144.

[66] MacDonald, R. Nagayasu, J. (1999), "The Long-Run Relationship Between Real Exchange Rates and Real Interest Rate Differentials: A Panel Study,"IMF Staff Paper 47/1.

[67] MacDonald, R., Marsh, I. and Nagayasu, J. (1996), "Revisiting the Long-Run Relationship Between Real Exchange Rates and Real Interest Rate Differentials," ICMM Discussion Paper Series, No. 39 (Glasgow: Economics Department, University of Strathclyde).

[68] MacDonald, R. and P. Swagel (2000), "Real Exchange Rates and the Business Cycle" in IMF Economic Outlook, June.

[69] Maddala, G. S. and Wu, S., (1999), "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test", *Oxford Bulletin of Economics and Statistics*, Special Issue, November, 61,631-652.

[70] Maeso-Fernandez,F. & Osbat,C. & Schnatz,B. (2001), "Determinants of the Euro Real Effective Exchange Rate : a BEER/PEER Approach," Working paper series 85, European Central Bank.

[71] Mark, N.C. and Doo-Yull, C. (1995), "Real Exchange Rate Prediction over Long-Horizons," Mimeo. (Ohio State University).

[72] Meese, R. and Rogoff, K. (1983), "Empirical Exchange Rate Models of the Seventies: do They Fit out of Sample?", *Journal of International Economics*, February.

[73] Meese, R. A. and Rogoff, K. (1988), "Was it Real? The Exchange Rate-interest Differential Relation over the Modern Floating-rate Period", *Journal of Finance* 43, 933-948.

[74] Montiel, P. (1997), "Exchange Rate Policy and Macroeconomic Management in ASEAN Countries", in J. Hicklin *et al.* (eds) Macroeconomic Issues Facing ASEAN Countries, Washington, D.C.: IMF.

[75] Montiel, P., Hinkle L.E. (1999). Exchange Rate Misalignment: Concepts and Measurement for DevelopingCountries. Oxford: for the World Bank by Oxford University Press.

[76] Murray, C. J., Papell, D. H., (2002), "The Purchasing Power Parity Persistence Paradigm," Journal of International Economics, Elsevier, vol. 56(1), pages 1-19.

[77] Obstfeld, M. and Rogoff , K. (1995), "The Intertemporal Approach to the Current Account," in G. Grossman and K. Rogo (eds), Handbook of International Economics, Vol. 3, Elsevier Science, Amsterdam, pp. 1731-1799.

[78] O'Connell, Paul G.J., (1997), "The Overvaluation of Purchasing Power Parity," *Journal of International Economics*, Vol. 44 (February), pp. 1-19.

[79] Officer, L. H. (1976), "The Purchasing Power Parity Theory of Exchange Rates", IMF Staff Papers, 23(1).

[80] Patel, J.(1990), "Purchasing Power Parity as a Long Run Relation", *Journal of Applied Econometrics,* 5, 367-379.

[81] Pedroni, P. (1999), "Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors," *Oxford Bulletin of Economics and Statistics*, 61 Special Issue, 653-70.

[82] Perron, P. (1989), "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis", *Econometrica*, 99, 1361-1401.

[83] Reinsel, G. C. (1993). Elements of Multivariate Time Series Analysis, Springer-Verlag, New York.

[84] Richards, A. and G.Tersman, (1996), "Growth, Nontradables and Price Convergence in the Baltics", *Journal of Comparative Economics* 23, 121-145.

[85] Rogers, J., and Jenkins, M. (1995), "Haircuts or Hysteresis? Sources of Movements in Real Exchange Rates", *Journal of International Economics* 38, 339-360.

[86] Rogoff, K. (1996), " The Purchasing Power Parity Puzzle", *Journal of Economic Literature.*

[87] Samuelson, P. A. (1964), "The General Theory," *Econometrica* 14 (July).

[88] Seabra, F.(1995), "Short-run Exchange Rate Uncertainty in Latin America",*Applied Economics,* Vol. 27: 441-450.

[89] Schnatz,B. & Vijselaar,F. & Osbat,C., (2003), "Productivity and the ("Synthetic") Euro-dollar Exchange Rate," Working paper series 225, European Central Bank.

[90] Soto, C., Valdes, R., (1999), "Exchange Volatility and Risk Premium," Working Papers Central Bank of Chile 46, Central Bank of Chile.

[91] Strauss, Jack, (1995), "Real Exchange Rates, PPP and the Relative Price of Nontraded Goods," *Southern Economic Journal* (April): 991-1005.

[92] Taylor, A. M. (2000), " A Century of Purchasing Power Parity," NBER Working Paper Series, Working Paper 8012.

[93] Thomas, A. (1997), " Is the Exchange Rate a Shock Absorber? The Case of Sweden", IMF Working Paper 97/176.

[94] Williamson, J. (1984). The Exchange Rate System, Washington, DC: Institute for International Economics.

[95] Wren-Lewis, S., Driver, R. L., (1998). Real Exchange Rates for the Year 2000. Washington: Institute for International Economics.

[96] Wu, Y. (1996), "Are Real Exchange Rates Non-Stationary? Evidence from a Panel Data Unit Root Test", *Journal of Money Credit and Banking*, 28, 54-63.

[97] Zanello, A., Desruelle, D. (1997), "A Primer on the IMF's Information Notice System", IMF Working Paper 97/71.

Chapter 3
Parameters instability in multivariate dynamic models: performance of the Chow-type tests

This chapter analyses small sample performance of the Chow-type tests for single structural break, under different types of models and breaks. The three types of Chow test most widely used in empirical research are compared against each other, and the generalised Chow-type tests with unknown breakpoint developed by Andrews and Ploberger (1994) in terms of size and power. For all the tests, the bootstrap is used to reduce massive size distortions in the original tests. The results obtained indicate superior performance by a sample-split test against a break-point and forecast tests. The tests are applied to the EMU money demand dataset in order to test stability .

KEY WORDS: Parameters stability, structural break, Chow-type tests, bootstrap.

3.1 Introduction

A key assumption in econometric modelling is parameter constancy, or stability. This assumption is used to estimate a model's parameters, and,with regard to the use of a model for forecasting, any form of partial or overall parameter instability may have severe consequences, negatively impacting both inference and model validity. Given the importance of stability, it seems surprising that so many empirical studies pay inadequate attention to the stability tests of their estimated models before proceeding to draw conclusions concerning the nature of the economic relations found. To an extent, this may be explained by authors' apprehension that results of stability analysis would undermine their models, and would cast doubts on the validity of their conclusions. Without such tests, however, the empirical findings may not acceptable to readers.

A second possible explanation for the relative scarcity of stability tests in empirical work could be that, with the advent of new econometric models and estimation methods, development of corresponding stability tests theory lags somewhat behind. This was indeed the case with single-equation regression models: by the time several formal stability tests were developed, the focus of econometric research had already shifted to multivariate regression models, and dynamic models, such as Stationary Vector Autoregressive (VAR) models. Development of new stability tests coincided with emergence of non-stationary dynamic models analysis which required new estimation methods, inference and, of course, new procedures to check their stability.

In spite of increased interest in model stability in the academic literature over the last decade[34] this area of research still remains a challenging one. During the recent years, the most important contributions in this area have included the emergence of tests for structural change of unknown timing; estimation of timing of a break; tests for multiple breaks; and tests for stability of stationary and cointegrated VAR models. A range of formal tests have been developed to address these issues. However, the accompanying requirement for a high level of knowledge of programming languages, authors, instead, in conducting stability analysis,had to rely on existing econometric packages. These offered only basic stability tests, such as visual inspection of the recursively estimated data or basic Chow tests, which suffer from massive size and power distortions (see, e.g. Diebold and Chen, 1996, or Candelon and Luetkepohl, 2001).

Against this background, this chapter attempts to conduct a comparative study of the properties of several tests for structural stability, namely, Chow-type tests using bootstrap methods to correct for massive size distortions observed in the multidimensional dynamic models. The performance of the three most widely used versions of Chow tests for a break with known timing and three generalised Chow-type tests with unknown break-point by Andrews and Ploberger (1994), is analysed using a Monte

[34] See, e.g., special issues of the Journal of Business and Economic Statistics (1992), and two issues of the Journal of Econometrics (1996 and 2004) devoted to the developments in the econometrics of structural change.

Carlo study. As an empirical example, the tests are then applied to the EMU money demand system, in order to check stability over the last years.[35]

The chapter is structured as follows. The next section reviews the existing literature and methods for testing model stability. In the third section, the tests used in this study are introduced. The fourth section contains a simulation study. The fifth section applies the tests to the model of M3 aggregate stability in the EMU, and final section describes the study's conclusions.

3.2 Review of structural stability literature and methods

3.2.1 Testing stability of stationary models

Since econometric science's earliest years, there has been debate on the role of econometrics in economics, centring on the issues of parameter constancy, and structural stability in modelling of empirical economic relationships (for a historical review see, e.g., Epstein, 1987). As highlighted by Andrews (1993), the statistical literature on change point problems had always been extensive, whereas the econometric literature small by relation, but rapidly growing - an observation that is still valid today. In this context, this section attempts to summarize the most important developments and contributions to the analysis of time series models stability.

[35] I am gratefull to Kai Carstensen for sharing his dataset on EMU money demand.

One of the earliest attempts demonstrating concern over the practical application of parameter constancy was Tinbergen's (1940) study. He tested equations on several sub-periods, used forecasting tests to evaluate model performance, and tested the robustness of regression coefficients when adding other variables. Nevertheless, the tests he applied still represented an informal construction, whereas in econometrics an hypothesis can be accepted or rejected solely on the basis of formal testing.

Informally, one straightforward way of assessing model stability could be to estimate the model recursively, adding one observation at a time, and constructing the series of recursive residuals, that is, the standardised 1-step forecast errors. These might perhaps be informative with respect to possible structural changes during the sample period. The series of cumulative sum of recursive residuals or squared residuals, CUSUM or CUSUM-of-squares, can also reveal structural changes, when the series extend to a point at long distance from zero. This technique was first proposed for this purpose by Brown, Durbin and Evans (1975), subsequently refined by Kraemer and Sonnberger (1986), and later modified by Ploberger and Kraemer (1992 and 1996) to take account of trending regressors and dynamic models. One of the principal advantages of these tests is that they can be used where the timing of a structural change is uncertain. On the other hand, these tests are of much lower power than any other tests of structural stability and, as a result, were not used to any significant extent within the mainstream analysis of structural stability, instead, relevant research took new and different directions.

Even before the use of CUSUM-type statistics had been considered, many authors had suggested recursive estimation of the parameters, i.e. starting from some initial subsample, extending the estimation sample by adding one observation at a time, to obtain an impression of model stability over time. To illustrate, informally, stability might be analysed by inspecting recursively estimated parameters, and basing a decision about model stability on their behaviour over time. These informal approaches were later implemented in formal tests. One, the so-called fluctuation test, operated on the basic idea of rejecting the null hypothesis of stability when the recursive estimates fluctuated too much. Ploberger *et al.* (1989) were the first to construct the test statistics, and to derive the limiting distribution for univariate linear regression models, including dynamic models; these were, however, only valid for regressors with no trends. At the same time, Nyblom (1989) put forward a Langrange Multiplier (LM) type test, addressing the behaviour of a recursively estimated LM-type statistics over the estimation period. He showed that the test based on the cumulative sum of the recursively estimated score function had a numerically tractable distribution. While intuitively rather plausible, these results required a sophisticated econometric theory; once this was developed, at the end of the 1980's, these tests were applied to various types of models and estimation methods.

Aside from recursive estimation, to indicate the possible timing of the break, a different strand of the literature looked at stability on the assumption that the approximate time of the break is known. The classical test for structural change was

developed by Chow (1960). His suggestion was to estimate the model on different sub-periods, with the equality of the two sets of parameters analysed using conventional F- or χ^2- distribution statistics. This test was popular for many years, and was extended to cover most simple econometric models of interest.

Despite its simplicity and intuitive appeal, the Chow-type tests were subject to two major limitations. First, in samples of common size, the F- and χ^2- approximations to the actual distributions may be very poor, and the actual rejection probabilities much larger than the desired Type I error (see Diebold and Chen, 1996 and Candelon and Luetkepohl, 2001). Indeed, this problem becomes even more severe in the case of the multivariate and dynamic models that have become the main tools of econometric macroanalysis. A second deficiency, is that the break date must be known *a priori*. Accordingly, the researcher has only two options - to pick an arbitrary date, or to pick the break based on some inspection of the data. In the former case, the test might miss the true break and, in the latter case, it might be misleading, as selection of the break is conditional on data, and conventional critical values are invalid (due to the fact that the nuisance parameter is identified only under the alternative hypothesis).

The necessary solution, as advanced by Quandt (1960), was to treat the break date as unknown: the alternative hypothesis was to be specified as a single structural break of unknown timing. The largest Chow test statistics would be drawn from

amongst all the candidate break dates. The Quandt or Sup test statistics is of the form:

$$Sup\ F_n = \sup_{k_1 \leq k \leq k_2} F_n(k)$$

where $F_n(k)$ is the Wald, LM, or LR statistics of the hypothesis of no structural change at date k which is known to lie in the range $[k_1, k_2]$. If the break date is known *a priori*, then the χ^2- distribution can be used to assess statistical significance, otherwise the χ^2 critical values are inappropriate. For many years, the question of what critical values should be used in the latter case remained unanswered. An analytical solution was then identified by Andrews (1993) and Andrews and Ploberger (1994). In addition to the $SupF_n$ test they suggested using Exp and Ave versions of the test with stronger optimality properties, with the following form:

$$ExpF_n = \ln(\frac{1}{k_2 - k_1 + 1}\sum_{k=k_1}^{k_2} \exp(\frac{1}{2}F_n(k))$$

$$AveF_n = \frac{1}{k_2 - k_1 + 1}\sum_{k=k_1}^{k_2} F_n(k)$$

It was shown that under a wide set of regulatory conditions, these statistics display asymptotic nonstandard null distributions; Hansen (1997) later provided a method to calculate p-values.

These tests represented a significant advance in the econometrics of structural change because they successfully relaxed the i.i.d. assumptions, by allowing for

dependent and heterogeneously distributed data. Although the model did not permit stochastic or deterministic trends, Vogelsang (1997) developed the Andrews (1993) version of the tests for detecting a break at an unknown date in the trend function of a dynamic univariate time series, allowing trend and unit root regressors and providing simulated critical values of the tests.

In the past, most research relating to structural change was devoted to the case of a single break, whereas the problem of multiple structural changes had received considerably less attention. More recently, the latter received greater attention with relevant literature including Andrews, Lee and Ploberger (1996), Bai and Perron (1998 and 2003) amongst others. Bai and Perron (1998), for example, proposed a sequential method for structural breaks that starts by testing for single break. If the test result is to reject the null of no break, the sample is split in two, and the test reapplied to each subsample.

3.2.2 Extensions of the basic stability tests

The tests above were developed principally for univariate models which are not widely used in modern empirical work. Currently, by contrast, published work addresses economic relations in the context of multivariate dynamic models, such as VAR models. While it is still possible in this setting to apply the earlier tests to each single equation, it is more desirable to test the stability of system as a whole. In distinction to the extension of CUSUM-type tests to multivariate models, Chow-type

tests are easily generalised to the multidimensional context, with the qualification that, as demonstrated by Candelon and Luetkepohl (2001), in the small sample simulations to which these tests are often applied, the distribution of test statistics under the null of stability may be substantially different from the assumed asymptotic F- or χ^2- distributions. The bootstrap versions of the tests, when the distribution of the estimator or test statistics is estimated by resampling one's data, or a model estimated from the data, emerged as much more reliable in the case of small samples (see Horowitz, 2001 for an overview of bootstrapping techniques). Further bootstrap techniques are constantly being reported (see Haerdle *et al*, 2003 and Buehlmann, 2002 for review of block, sieve, local, nonparametric autoregressive and periodogram bootstrap methods); amongst these developments has been the study by Diebold and Chen (1996) that applied bootstrap methods to the Andrews-type tests for structural break of unknown timing, thus greatly improving the asymptotic approximation to the finite-sample distribution.

The above-mentioned approaches are based on the assumption that the alternative to parameter constancy is one or several structural breaks. The tests assumed no break under the null hypothesis, so the model under study was considered to be stable, although with stochastic trends present (Nelson and Plosser, 1982). However, Rappoport and Reichlin (1989) and Perron (1989) argued that only few events have any permanent effect on economic series. Consequently, they suggested representing such shocks as breaks in the underlying deterministic trends. In this context, much

attention was devoted to estimation and testing of models assuming *a priori* one or more breaks. Given the stochastic nature of the majority of macroeconomic series the issue of testing for unit roots in the presence of structural breaks in the data has been addressed by many recent papers[36].

Apart from assumptions about the location of the break, it can frequently happen that model misspecification shows not as a sharp break, but as continuous change in parameters[37]. The works of Lin and Terasvirta (1994, 1999) and He, Teraesvirta and Gonzalez (2002), in developing tests for the constancy of parameters with an alternative of continuous change are worth mentioning here, along with Busetti and Harvey's (2001, 2003) papers. Since the current study is restricted to tests that assume model stability under the null, the interested reader is referred to the above-mentioned research for more thorough treatment of these issues.

3.2.3 Testing stability of the models with cointegrated regressors

It is now well-established that most macroeconomic series exhibit non-stationarity, so that statistical inference within a conventional VAR model would be misleading. As a result, demand has grown for tests of stability of cointegrated models, whether the problem under study was long-run stability manifested in constant long-run relations, or, alternatively, short-run stability of adjustment coefficients, of short-run dynamics.

---

[36] See, for example, Perron (1989), Perron and Vogelsang (1992), Banerjee *et al.* (1992), Zivot and Andrews (1992), Amsler and Lee (1995), Leybourne *et al* (1998), Saikkonen and Luetkepohl (2001, 2002).

[37] See Perron and Vogelsang (1992) and Vogelsang and Perron (1998) for thorough reviews of this issue.

Initially, research focused solely on linear or nonlinear multiple regression models with a general error structure. For example, Banerjee *et al.* (1992), Gregory and Hansen (1996) and Hansen (1992) proposed tests for a shift in the cointegrated processes. Hansen (1992) constructed LM-type tests using a fully modified estimator of Phillips and Hansen (1990), based on the tests of Quandt (1960) and Nyblom (1989). Later, Seo (1998) derived the LM test for a structural change in cointegrating relations and adjustment coefficients with an unknown change point. This latter test was applicable to maximum likelihood estimation, and was found to have the same nonstandard asymptotic distribution as in Hansen (1992).

Further developments paid special attention to the error-correction framework, of which the main feature is the reduced rank of coefficients matrix for cointegrated regressors. A pioneer in this area was Quintos (1997), who applied a fluctuation test of Ploberger *et al.* (1989) to nonstationary reduced rank models, developing tests not only for the stability of long-run relations, but also a test for cointegrating rank stability of potentially wide application. In addition, based on the Nyblom LM-test and Ploberger fluctuation test, Hansen and Johansen (1999) offered some graphical procedures, based on recursively estimated eigenvalues, and their discovered asymptotic distribution, as well as constructed formal tests of the constancy of the long-run parameters in the cointegrated VAR models. These computationally straightforward and easily-coded tests have found subsequent use in many studies on structural stability,

although their small-sample properties have not yet been scrutinised; a live issue for many recently constructed tests.

Hansen (2003) generalised Johansen's cointegrated VAR model estimation method, to allow for structural changes; taking the time of change points and the number of cointegration relations as given. The resulting estimation technique (' generalised reduced rank regression') allows for linear restrictions on all parameters apart from the variance parameter, and did not require a constant covariance matrix. This provided a major breakthrough in the time-series econometrics of structural change. Indeed, using this estimation method, in addition to testing for structural changes, one is also able to construct and estimate models that take account of those changes, and to draw policy conclusions with regard to the whole sample under study even in presence of structural breaks.

While all the tests for cointegration breakdown described so far assume that post-breakdown period to be relatively long, testing for stability at the very end of the sample represents an attractive challenge for time-series econometrics. With this in mind, Andrews and Kim (2003) developed several tests, based on the sum of squared post-break residuals evaluated at a pre-break estimator, or, alternatively, on the sum of squared reverse partial sums of post-break residuals, with similarity to the Nyblom test. It was shown that their constructed tests were not consistent, but asymptotically unbiased.

One of the latest contributions to the field, by Andrade *et al.* (2005), has been the development of a statistical procedure capable of consistently identifying the number of cointegration relationships when a break occurs at a known date and affects the cointegrated space. The authors argued that all previous tests imposed stable cointegration condition under the null hypothesis and were therefore useless in answering the question of whether a rejection of cointegration stemmed from an undetected instability. This work extended earlier research by Inoue (1999) and Luetkepohl *et al.* (2003), who had considered the break affecting only deterministic components.

In sum, the foregoing discussion reveals that the topic of the structural stability continues to evolve rapidly, and is still far from being a mature and established econometric topic.

3.2.4 Empirical tests of the stability of the EMU money demand

So far, this review has examined only theoretical aspects of the tests. While proper modelling would require proof of the stability of the estimated model, particular empirical puzzles concerning the stability of certain relations between economic variables have endured. For example, many studies focused on the stability of the predictive power of the yield curve on output growth (Estrella and Hardouvelis, 1991, Candelon and Cubadda, 2004 to name only a few). Many authors looked at stability of the US term structure of interest rates (Hansen and Johansen, 1999 and Hansen,

2003), stability of the Phillips curve (Alogoskoufis and Smith, 1991 and many others), behaviour of the aggregate consumption function (Hansen, 1992) and many other stylized facts of economic theory.

However, one of the most popular areas of stability analysis has been, and remains, the stability of various money aggregates or stability of money demand. A host of papers address money demand in individual countries, and the creation of the ECB has drawn considerable interest towards stability of the demand for euro area M3. These studies use synthetic data for the pre-EMU period, and tend to use informal tests for stability, such as visual inspection of recursive coefficients, comparison of estimates across subsamples, visual inspection of 1-step Chow forecast tests[38] or predictive failure tests. (see Hayo, 1999, Coenen and Vega, 2001, Kontolemis, 2002). Overall, the studies concur that euro-area M3 demand was rather stable, much more so stable than individual members' money demands.

Recently, with development of formal stability tests for cointegrated models, researchers have been able to come to more solid conclusions. In a comprehensive study by Bruggeman *et al.* (2003), applying the fluctuation and Nyblom-type stability tests from Hansen and Johansen, the authors obtained mixed results with regard to the stability of euro area money demand. In the contention of Carstensen (2004), the limited dataset prevented the tests from indicating non-stability potentially resulting from the ECB's revision of monetary policy strategy in May 2003 (ECB, 2003).

[38] Usually performed in software package PcFiml. However, as will be shown later this test, even its bootstrapped version has inferior power properties to other tests analysed.

Carstensen (2004) therefore applied the Andrews and Kim (2003) new tests for detecting breaks at the end of the sample, both to the conventional money demand system and, secondly, to augmented model with stock market variables, to account for recent portfolio shifts from equities into safe and liquid assets (due to equity market downsizing). His finding was that instability of money demand disappeared once the augmented system was analysed.

As an empirical application of the bootstrapped Chow-tests, this study used the dataset constructed by Carstensen (2004), to test stability of the money demand model for the euro area M3 for 1980-2003.

3.3 Tests description

The principal purpose of the study is to address relative performance of three Chow-type tests frequently used in empirical research. First, the so-called Chow *forecast test,* which is included in the econometric package PcFiml, which denotes it as "$N\downarrow$ " or "$N\uparrow$-step Chow test"[39] Though widely employed in empirical research, unfortunately, as revealed below, this test is inferior, in terms of power performance, to the other two versions of the Chow test, here identified as the *sample-split* (SS-) and *break-point* (BP-) tests. Below, a description of the tests, closely following Luetkepohl and Kraetzig's (2004) notation is provided.

[39] Formula 14.28 or 14.29 in Doornik and Hendry (1997).

Given n time series $y_t = (y_{1t}, ..., y_{nt})'$, the basic VAR($p$) model considered in the derivation of the test has the following form:

$$y_t = \Phi D_t + A_1 y_{t-1} + ... + A_p y_{t-p} + u_t \qquad (t = 1, ..., T) \tag{3.1}$$

where D is a matrix $(m \times T)$ of deterministic terms and dummies, Φ is $(n \times m)$ matrix, the A_i are $(n \times n)$ coefficient matrices and $u_t = (u_{1t}, ...u_{nt})'$ is an unobservable zero mean white noise process with time invariant positive definite covariance matrix Σ, i.e. $u_t \backsim iid(0, \Sigma)$[40]. The derivation of the Chow tests for a break in period T_B proceeds as follows. The model under consideration is estimated from the full sample of T observations and from the first T_1 and the last T_2 observations, where $T_1 < T_B$ and $T_2 \leq T - T_B$. Denoting the resulting residuals by $\hat{u}_t$, $\hat{u}_t^{(1)}$ and $\hat{u}_t^{(2)}$, respectively, and using the following notation:

$$\widetilde{\Sigma}_u = T^{-1} \sum_{t=1}^{T} \hat{u}_t \hat{u}_t' \tag{3.2}$$

$$\widetilde{\Sigma}_{1,2} = (T_1 + T_2)^{-1} (\sum_{t=1}^{T_1} \hat{u}_t \hat{u}_t' + \sum_{t=T-T_2+1}^{T} \hat{u}_t \hat{u}_t') \tag{3.3}$$

$$\widetilde{\Sigma}_{(1,2)} = T_1^{-1} \sum_{t=1}^{T_1} \hat{u}_t \hat{u}_t' + T_2^{-1} \sum_{t=T-T_2+1}^{T} \hat{u}_t \hat{u}_t' \tag{3.4}$$

$$\widetilde{\Sigma}_{(1)} = T_1^{-1} \sum_{t=1}^{T_1} \hat{u}_t^{(1)} \hat{u}_t^{(1)'} \tag{3.5}$$

$$\widetilde{\Sigma}_{(2)} = T_2^{-1} \sum_{t=T-T_2+1}^{T} \hat{u}_t^{(2)} \hat{u}_t^{(2)'} \tag{3.6}$$

The sample-split (SS) test statistic has the form

[40] As presence of the ARCH effects in the residuals would invalidate the tests results, it is assumed that these are absent in the model being tested.

$$\lambda_{SS} = (T_1 + T_2)[\log\det\widetilde{\Sigma}_{1,2} - \log\det\{(T_1 + T_2)^{-1}(T_1\widetilde{\Sigma}_{(1)} + T_2\widetilde{\Sigma}_{(2)})\}] \quad (3.7)$$

This statistics has $\chi^2(k)$-distribution where k is the difference between the sum of the number of coefficients estimated in the first and last subperiods, and the number of coefficients in the full sample model. The null hypothesis of constant parameters is rejected if λ_{SS} is large. The SS statistic is derived under the assumption that the residual covariance matrix is constant over the whole sample.

The break-point (BP) test statistic has the form

$$\lambda_{BP} = (T_1 + T_2)\log\det\widetilde{\Sigma}_{1,2} - T_1\log\det\widetilde{\Sigma}_{(1)} - T_2\log\det\widetilde{\Sigma}_{(2)} \quad (3.8)$$

This statistic has $\chi^2(k)$-distribution where k is, once again, the difference between the sum of the number of coefficients estimated in the first and last subperiods, and the number of coefficients in the full sample model. The null hypothesis of constant parameters is rejected if λ_{SS} is large. This version of the test allows for changing residual covariance matrix across two subsamples.

The forecast (F) test statistic is distributed asymptotically as F-distribution and, given the recurrent problem of oversizing in small samples, Rao (1973) suggested the use of the following F-approximation for likelihood-ratio based tests, such as the forecast test:

$$\lambda_F = \frac{1-(1-R_r^2)^{1/s}}{(1-R_r^2)^{1/s}} \cdot \frac{Ns-q}{nk} \approx F(nk, Ns-q), \quad (3.9)$$

where

$$s = \left(\frac{n^2k^2 - 4}{n^2 + k^2 - 5}\right)^{1/2}, \quad q = \frac{nk}{2} - 1, \quad N = T - k_1 - k - (n - k + 1)/2. \quad (3.10)$$

Here k_1 is the number of regressors in the time invariant model and

$$R_r^2 = 1 - \left(\frac{T_1}{T}\right)^n |\widetilde{\Sigma}_{(1)}|(|\widetilde{\Sigma}_u|)^{-1}. \quad (3.11)$$

The forecast test also rejects the null hypothesis of constant parameters for large values of the test statistic.

Because the actual small sample distributions of the test statistics under H_0 may be quite different from the asymptotic χ^2- or F-distributions, Candelon and Luetkepohl (2001) suggested using bootstrap techniques to calculate the tests' empirical p-values. They are computed as follows. From the estimation residuals $\hat{u}_t$, centered residuals $\hat{u}_1 - \bar{\hat{u}}, \ldots, \hat{u}_T - \bar{\hat{u}}$ are computed. Bootstrap residuals $u_1^*, \ldots, u_T^*$ are generated by random drawing, with replacement from centered residuals. Based on these quantities, bootstrap time series are calculated recursively, starting from the given presample values $y_{-p+1}, \ldots, y_0$. Then the model is reestimated, with and without allowing for a break, and bootstrap versions of the statistics of interest, say λ_{SS}^* and λ_{BP}^*, are computed. The p-values of the tests are estimated as the proportions of values of the bootstrap statistics that exceeds the corresponding test statistic, based on the original sample. As demonstrated by Horowitz (2001), in many cases bootstrap tests performed quite satisfactorily in small samples and, as will later be shown, certainly addresses Chow tests problem of oversizing.

As discussed in the previous section, Andrews (1993) and Andrews and Ploberger (1994) suggested using the generalised Chow tests with unknown change point to investigate system instability. Here, we use the LM versions of these tests, and calculate the critical values of the test statistics following the procedure of Hansen (1997). To correct for possible size distortions, the bootstrapped versions of these tests are also used, and their performance is compared to the performance of the asymptotic tests, and to the three versions of known-breakpoint Chow tests (SS-, BP- and F-tests, in our notation).

3.4 Simulation study

3.4.1 Design of the experiment

It was decided to study the tests' relative performance, in terms of size and power, using Monte Carlo methods. On other occasions, experimental can be rather easily formalised (e.g. when testing autocorrelation of the residuals (Edgerton and Shukur, 1999); testing the Granger causality tests performance (Zapata and Rambaldi, 1996); or for cointegrated rank of a VAR process with a structural shift (Luetkepohl, Saikkonen and Trenkler, 2003). In this case, however, any experimental design could seem ambiguous due to the infinite number of break specifications. From this starting point, the experiment was designed in the following way.

The sample size was chosen to be $T = 100$[41]. Given the variety of models used in empirical work, it was decided to focus on univariate, bivariate and five-dimensional models as those being most representative. For example, the DGPs for univariate models are simulated on the basis of US investment and Polish productivity series available in the software package JMulTi. Bivariate models were generated artificially because their simplicity facilitates the study of test performance - especially true in the case of power studies - and because they appear frequently in applied work related to purchasing power parity, market efficiency etc. The stationary model with five variables was simulated using the estimated parameters of the EMU monetary system, taken from Brand and Cassola (2000). All model specifications can be found in the Appendix to this chapter. The number of Monte Carlo simulations in each experiment was 1000, and the nominal significance level of the tests equal to five per cent[42]. Every bootstrap sample used in generating the artificial data consisted of 1000 observations.

3.4.2 Size simulations

Tables 1 and 2 report the empirical sizes of the tests for the DGP's described in the Appendix. In each Monte Carlo simulation, the DGP process was generated accord-

---

[41] Macroeconomics and monetary models quite commonly use annual data available only from the beginning of the 20th century, or even after the WWII in majority of cases; quarterly data for which the time span in the mid 1970's. In all cases this leaves the researcher with approximately 100 time realisations of each variable in the model.

[42] The tests were run with 1, 5 and 10 % significance levels: all the results are available in full from the author upon request.

ing to the model parameters, and the tests for structural break were applied with timing of the break at 20% , mid-way, and at 80% of the sample size (labelled in the tables as 0.2T, 0.5T and 0.8T). The tests are featured by columns, and different the DGP's by rows.

Table 1 shows the sizes of the three tests under study - sample-split (SS), break-point (BP) and forecast (F) tests. It can be seen that the sizes of the asymptotic tests are badly biased when the number of parameters under the null hypothesis is large, compared to the sample size, confirming previous findings by Edgerton and Shukur (1999) and Candelon and Luetkepohl (2001). This is true for all the models studied, and with higher dimensions, or more lags, the size distortions grow dramatically. Thus, it is clear that asymptotic versions of the tests cannot be recommended for typical macroeconomic applications, as they will tend to overreject the null hypothesis of stability.

Careful inspection of the results for asymptotic tests yields several findings worthy of mentioning. First, increased residual variance does not appear to influence the size of the tests (indicated by a comparison of DGP1 with DGP1-A, the latter with ten-times higher variance). Inclusion of the trend in the model, however, leads to bigger oversizing of the tests (as seen by comparing DGP1 with DGP1-B)[43]. Secondly, the closer the process to the unit root, the bigger are the size distortions (see DGP3 with autoregressive coefficients of 0.4, DGP3-B with coefficients of 0.8 and

[43] Similar tests were run for all of the DGPs 1-5; results are not reported in the table due to space constraints.

DGP3-C with coefficients of 0.98), a result that is emphatically true for the BP-test. The same statement holds valid for the models with correlated residuals (see model DGP3-A). Thirdly, an increase in model dimension leads to higher size distortions, although the F-test is much more robust than the other two tests. Lastly, the F-test performs better at the middle, and at the end of the sample, whereas the SS- and BP-tests show roughly the same results for various break dates.

Given the size distortion of the asymptotic tests, performance of the bootstrapped version of the tests is interesting. Relevant figures are listed in the next columns of the table. It is clearly visible that, for all the models, the bootstrap versions of the tests have size close to true level of 5% (taking into account that the standard error of an estimator of a true rejection probability, P, based on 1000 replications of the experiment, is $\sqrt{P(1-P)/1000}$ or 0.007 in our case). The only case where the bootstrap tests are still oversized is DGP3, which is quite close to the unit root process.

Table 2 shows the performance of the Andrews-type tests SupF, AveF and ExpF, which test each DGP for the presence of a structural break in the interval of [0.15T, 0.85T] of the sample. It can be observed that these tests perform more strongly than the asymptotic single-break tests, and in line with their bootstrap versions. Unlike the basic tests, they are slightly undersized in all the tested models, except for the near unit root processes of DGP3A and DGP4, in which case they exhibit massive size distortions, even by comparison with asymptotic SS, BP and

F-tests. Interestingly, the tests are undersized most when the model includes a time trend (DGP3 and DGP3A). The same bootstrap techniques were applied to these tests, with the bootstrap versions showing slightly stronger performance than their asymptotic counterparts.

Overall, the size distortions are quite large for both tests and, given the computational power of modern computers, it is clear that the bootstrap method should be applied whenever possible, in order to correct for size distortions. Next, we turn to the analysis of the power properties of the tests. In light of the size distortions, the focus will be restricted to the bootstrap versions of the tests, although results were also obtained for the asymptotic tests.

3.4.3 Power simulations

In order to analyze the power of the tests, DGP's were generated under an alternative hypothesis of a break in the series, observing the empirical rejection rate for the nominal size of 5%. Adjustment for size was not performed, given the unavailability of such adjustment in practice. One of the most difficult issues in power analysis is the generation of breaks, because the number of types of breaks is infinite. Not having encountered a coherent classification of the break types, an attempt was made to create breaks along the following dimensions. First, the breaks were generated in the middle, and at the end of the sample, in order to see how the tests performed with respect to the break location. Secondly, the breaks were generated in the mean,

and in the trend, as these cases are often both encountered in empirical research, and observed in the data. Thirdly, a constant was added to all the coefficients of the model, following the approach of He, Teraesvirta and Gonzalez (2002). Next, one of the coefficients in the univariate model, and a column of coefficients corresponding to a specific variable were dropped or added to the generation of some DGP's. This type of break could occur when there is a breakdown in the economic relations between variables, or, conversely, the creation of a new relation. As a special case, it was also decided to inspect the break when the model process shifts closer to the unit root processes, by increasing the diagonal elements in the corresponding coefficient matrix.

Tables 3 and 4 show the empirical power of the tests, grouped by the location of the breaks (Table 3 reports results for the breaks generated in the middle of the sample, and Table 4 - at 80% of the sample size). The SS- ,BP- and F-tests are organized by columns, and for each DGP by rows, giving the test statistics for the breaks tested at 20, 50 or 80% of the sample size. Accordingly, the sensitivity of the tests to misspecification of the location of the break can be tested. The Andrews-type tests are featured by rows, since these test for the break in the prespecified range, rather than at a single point[44].

[44] [0.15 0.85] cut-off points were used for the location of the break, as these are most often used in empirical tests, although a tighter or broader interval might have been selected.

The principle objective of the analysis was to assess comparative performance of the tests in terms of size and power, rather than absolute performance. Several relevant which emerged are as follows.

1. First, and most strikingly, F-version of the test performed very poorly in terms of power. For all break locations, and for all DGP's, this was outperformed by the SS- and BP-test. This is an alarming finding, as the F-test is frequently used in empirical research and, given its low power, it must come as little that relevant research papers usually failed to find any sign of instability of analysed models. F-test results, it would appear, should always be checked against the other two tests.

2. Andrews-type tests perform in line with the SS- or BP-test, although with slightly lower power in all cases. This result is not surprising and suggests that these tests could be used together. First, Andrews-type tests could indicate presence of instability for the whole sample, with the SS- or BP-tests investigating potential break locations.

3. SS, BP and F–tests are very sensitive to misspecification of break location, especially when the test assumes the break to occur at the beginning of the sample. There, breaks were generated in the middle, and at the end of the sample, and an attempt was made to run the tests under two assumptions. Firstly, that the researcher has correctly identified the time of the break (testing

for a break at time t, when the true break is generated at time t); secondly, that the researcher runs a test that misspecifies the timing of the break (testing for a break at 20% of the sample when it is generated at 80%). The results obtained show that the tests' performance depends crucially on the researcher's ability to identify the break correctly.

4. SS-tests shows slightly stronger performance in terms of power than the BP-test in all cases except where a break is modelled in the covariance matrix of the residuals. Since the BP-break is designed to test the constancy not only of regression coefficients, but also of the white noise covariance matrix, so their conflicting results in empirical applications could serve as a sign of additional instability of the covariance matrix or, alternatively, of the presence of ARCH effects.

5. In certain cases when the break is generated at the end of the sample, testing for the break in the middle shows higher power. At present, it is not clear what might be responsible for such a result, and further analysis of this phenomenon is needed.

3.4.4 Chow tests in cointegrated VAR systems

As the findings described in previous sections have shown, the performance of all versions of the tests deteriorates quite dramatically as the processes in the model

become more similar to unit root processes (DGP 3 used for size and power simulations). Indeed, it is tempting to take the residuals from the cointegrated model estimated, for example, by reduced rank regression and calculate the tests statistics ignoring the cointegration. However, this would lead to erroneous conclusions, analogous to the case of statistical inference from cointegrated models estimated by O.L.S. While a possible algorithm might estimate the long-run relations on the total sample, and fix them on the sub-sample estimation, so far, there is no formal theoretical justification for such a procedure.

An additional complication attendant on the presence of cointegration is that it invalidates conventional arguments for the asymptotic validity of the bootstrap approach, such as the efficiency conditions presented by Beran and Ducharme (1991, Prop. 1.3). As demonstrated by Inoue and Kilian (2002), the bootstrap achieves the correct first-order asymptotic distribution for the non-unit root parameters, but not for the estimated unit root parameter (nor for deterministic regressors). Several other bootstrapped methods have been developed to take account of the time series nature of the data (Haerdle *et al.*, 2003 and Buehlmann, 2002 for a review of block, sieve, local, nonparametric autoregressive and periodogram bootstrap methods), and it is hoped that ongoing research in this area will produce more favourable results for cointegrated series bootstrapping.

Notwithstanding the limitations cited above, it remains useful to conduct the tests on cointegrated models, in order to explore relative performance of the asymp-

totic and bootstrapped versions of the tests. With this in mind, cointegrated models were added to the present studies: the two-dimensional cointegrated VAR model, either with one or two lags in the short-run dynamics, and a model with 5 variables simulated on the basis of the euro area money demand dataset from Carstensen (2004), with one lag. It must be stressed, however, that the analysis of the tests' performance for cointegrated models is far from complete, and these first steps serve only as rough indications of their application to non-stationary dynamic systems.

Table 1 displays the empirical rejection probabilities both for asymptotic and bootstrapped Chow tests[45]. In contrast with the stationary case, the asymptotic versions perform quite well for two-dimensional models, but, on the other hand, are badly oversized for five variables systems. Nevertheless, the bootstrap again helps to achieve the empirical sizes of the tests close to nominal 5% levels. Mediocre performance of the forecast version of the test (F-test) deserves special attention. In fact, even its asymptotic version is substantially undersized - whether this is due to inappropriate treatment of the degrees of freedom for the non-stationary case, or another as yet unidentified factor, remains an open question. But overall, it seems that using the bootstrapped version of Chow tests even for cointegrated models is, on empirical grounds at least, well justified.

Table 3 displays empirical rejection probabilities when the models are generated under the null of various break types. For cointegrated models, the focus was on

[45] The tests were run only for the known-time break-points, omitting the Andrews generalised tests.

the model with five variables, in order to explore the power of the tests for breaks in the deterministic trend, alpha vector and short-run coefficient matrix. As no proper tests with varying intensity of breaks were completed, the main issue addressed became that of tests' relative performance. As in stationary cases, SS-tests has better power properties for all the models, except for a model with a break in the deterministic component, where it shows worse performance than the BP- and F-tests.

3.4.5 Stability of the euro area M3 demand.

As an empirical application of the bootstrapped Chow-tests, the dataset constructed by Carstensen (2004) was used to test the stability of the money demand model for euro area M3 for 1980-2003. As mentioned above, in the course of the section on review of the literature concerning money demand stability, this topic has risen in practical prominence due to its unquestionable importance in the determination of ECB policy. Carstensen (2004) has suggested augmenting the money demand equation with variables accounting for stock market downsizing and a shift to safer investments such as M3 components. He uses the baseline long-run money demand function of the form

$$mp_t = \beta_1 y_t + \beta_2(r_t^s - r_t^0) + \beta_3(r_t^e - r_t^0) + \beta_4 v_t + u_t \qquad (3.12)$$

where mp_t denotes real M3, y_t is real GDP, r_t^0 is the own rate of M3, r_t^s is the short-term interest rate, r_t^e denotes equity returns and r_t^0 denotes stock market volatility.

A re-run was executed of the above money demand system estimation, using the same model specification with one lag, one cointegration relation and a constant restricted to the cointegrated space. Two questions can be asked in this framework. First, whether the creation of EMU in 1999Q1 triggered any signs of structural change in the system; and, secondly, whether the start of excessive M3 growth around the end of 2001, had any effect on model stability.

Analysis of the residuals of the estimated model showed that there was no residual autocorrelation or ARCH effects, and that the normality tests were passed, using both multivariate and univariate statistics Prior to running the tests, Hansen and Johansen (1999) recursive eigenvalues tests were applied to assess the stability of the long-run relations in the model. Figure 3.1 shows that, apart from initial unstable behaviour, the tests cannot reject the hypothesis of stable money demand for the euro area. If long-run relations are fixed at their full-sample estimates[46], bootstrapped Chow tests can be run to test the stability of the short-run dynamics of the model and cointegration adjustment coefficients. Two possible candidates are the potential breaks in 1999 and 2001. Table 5 shows p-values of the tests.

Date of the break	SS-test	BP-test	F-test
1999 Q1	0.011	0.018	0.440
2001 Q1	-	-	0.069

[46] The results of the recursive eigenvalues tests non rejecting the hypothesis of stable long-run relations, serve as a justification for this step.

Table 5. P-values of the bootstrapped tests for structural break.

These results indicate some signs of parameter instability in 1999 and in 2001, although the F-test p-value in the former case is above the rejection level. However, given its low power, as illustrated in the previous section, this test result should be treated with caution. For the break in 2001 Q1 we can only run F-tests, as for the 5-dimensional model, the second subsample should be sufficiently large, which is not the case here. In any event, even the F-test shows some instability around 2001 Q1.

As an informal check, it is possible to plot the graph of Chow tests p-values for the whole sample (Figure 3.2). Indeed, the BP-test p-value, for most of the time, falls below 10 %, indicating an unstable covariance matrix for the whole system; the SS-test p-value declines rapidly only in 1999 Q1, being sufficiently high before that; and F-test's instability starts only around 2001 Q1. Thus it indeed seems to be the case that euro area money demand system shows some structural instability in the recent years, while further research is needed to resolve this issue with a higher degree of certainty.

3.5 Conclusions

In this part, the analysis has been conducted of the performance of single break-point Chow-type tests. As suggested by previous studies, asymptotic distribution theory

does not work well in small samples, and the use of an empirical null distribution, obtained by a recursive bootstrap, appears necessary in order to avoid massive size distortions. Simulation studies were conducted to ascertain the size and power properties of the sample-split, break-point and forecast test versions of the Chow tests in the systems, with different lag length, number of variables, location of breaks, presence of near unit root processes, and degree of correlations between variables. The results obtained indicate that the SS-test outperforms BP- and F-tests in all cases, except for varying covariance matrix; and that the bootstrap version of the test has both correct size and reasonable power against the alternative of a structural break generated according to the extensive, though incomplete, types of breaks classification.

References

[1] Alogoskoufis, G.S. and Smith, R. (1991), "The Phillips Curve, the Persistence of Inflation, and the Lucas Critique: Evidence from Exchange-Rate Regimes", American Economic Review, American Economic Association, vol. 81(5), 1254-75.

[2] Andrade, P., Bruneau, C. and Gregoir, S. (2005), "Testing for the Cointegration Rank when some Cointegrating Directions are Changing", Journal of Econometrics, Volume 124, Issue 2, 269-310.

[3] Andrews, D.W.K. (1993), "Tests for Parameter Instability and Structural Change with Unknown Change Point", Econometrica, 61, 821-856.

[4] Andrews, D.W.K. and W. Ploberger (1994), "Optimal Tests when a Nuisance Parameter is Present only under the Alternative", Econometrica, 62,1383-1414.

[5] Andrews, D.W.K., Lee, I. and Ploberger, W. (1996), "Optimal Changepoint Tests for Normal Linear Regression," Journal of Econometrics, 70, 9-38.

[6] Andrews, D. W. K., Kim, J (2003), "End-of-Sample Cointegration Breakdown Tests", Cowles Foundation Discussion Papers 1404, Cowles Foundation, Yale University.

[7] Bai, J. (1997), "Estimating Multiple Breaks One at a Time", Econometric Theory, 13, 551-563.

[8] Bai, J., and P. Perron (1998), "Estimating and Testing Linear Models with Multiple Structural Changes," Econometrica, 66, 47-78.

[9] Bai, J., and P. Perron (2003), "Computation and Analysis of Multiple Structural Change Models", Journal of Applied Econometrics, John Wiley & Sons, Ltd., vol. 18(1), pages 1-22.

[10] Banerjee, A., Lumsdaine, R. L & Stock, J. H. (1992), "Recursive and Sequential Tests of the Unit-Root and Trend-Break Hypotheses: Theory and International Evidence", Journal of Business & Economic Statistics, American Statistical Association, vol. 10(3), pages 271-87.

[11] Beran, R., Ducharme, G.R., (1991), "Asymptotic Theory for Bootstrap Methods in Statistics", Les Publications CRM, University of Montreal, Canada.

[12] Brand, C. and N. Cassola, (2000), "A Money Demand System for Euro Area M3". ECB Working Paper Series No. 39.

[13] Brown, R. L., Durbin, J. and Evans, J.M. (1975), "Techniques for Testing for the Constancy of Regression Relationships over Time", Journal of the Royal Statistical Society B, 37, 149-192.

[14] Bruggeman,A., Donati,P. and Warne,A. (2003), "Is the Demand for Euro Area M3 Stable?", Working paper series 255, European Central Bank.

[15] Buhlmann, P. (2002), "Bootstraps for Time Series", Statistical Science, 17, no. 1, 52–72.

[16] Candelon, B. and Lutkepohl, H. (2001), "On the Reliability of Chow-type Tests for Parameter Constancy in Multivariate Dynamic Models", Economics Letters, Elsevier, vol. 73(2), 155-160.

[17] Candelon,B. and Cubadda, G. (2004), "Testing for Parameter Stability in Dynamic Models across Frequencies", mimeo.

[18] Carstensen, K., (2004)", Is European Money Demand still Stable?", Kiel Working Paper No. 1179, revised version.

[19] Chow, G.C. (1960), "Tests of Equality Between Sets of Coefficients in two Linear Regressions", Econometrica, 28, 591-603.

[20] Coenen, G. and Vega, J (2001), "The Demand for M3 in the Euro Area", Journal of Applied Econometrics, 6, 727-748.

[21] Diebold, F. X. and Chen, C. (1996), "Testing Structural Stability with Endogenous Breakpoint: a Size Comparison of Analytic and Bootstrap Procedures", Journal of Econometrics, Elsevier, vol. 70(1), 221-241.

[22] Doornik, J.A., Hendry, D.F. (1997). Modelling Dynamic Systems using Pc-Fiml 9.0 for Windows. International Thomson Business Press, London.

[23] ECB (2003). The ECB's monetary policy strategy, ECB Press Release, 8 May.

[24] Edgerton, D., Shukur, G. (1999), "Testing Autocorrelation in a System Perspective", Econometric Reviews, 18, 343-386.

[25] Epstein,R., 1987. A History of Econometrics. Amsterdam: North-Holland.

[26] Estrella, A. , Hardouvelis, G. (1991), "The Term Structure as Predictor of Real Economic Activity" , Journal of Finance,X LVI, 555-576.

[27] Gregory, A.W., Hansen, B. E. (1996), "Residual-based Tests for Cointegration in Models with Regime Shifts", Journal of Econometrics, Elsevier, vol. 70(1), pages 99-126.

[28] Hansen, B.E. (1992), "Test for Instability with I (1) Processes", Journal of Business and Economic Statistics 10, 321-335.

[29] Hansen, B.E. (1997), "Approximate Asymptotic p-values for Structural Change Tests", Journal of Business and Economic Statistics , 15, 60-67.

[30] Hansen, H., Johansen,S.(1999), "Some Tests for Parameter Constancy in Cointegrated VAR-Models", Econometrics Journal 2, 306-333.

[31] Hansen, P.R. (2003), "Structural Changes in the Cointegrated Vector Autoregressive Model", Journal of Econometrics 114, Pages 261-295.

[32] Hayo B., (1999), "Estimating a European Demand for Money", Scottish Journal of Political Economy 46(3), 221-244.

[33] He, C., Teräsvirta, T., and González, A. (2002), "Testing Parameter Constancy in Stationary Vector Autoregressive Models against Continuous Change", Working Paper Series in Economics and Finance 507, Stockholm.

[34] Hendry, D.F. (1995). Dynamic Econometrics, Oxford: Oxford University Press.

[35] Horowitz, J.L. (2001), "The Bootstrap", in: Heckman, J.J., Leamer, E.E. (Eds.). Handbook of Econometrics, Vol. 5.North-Holland, Amsterdam.

[36] Huerdle, W., Horowitz, J. and Kreiss, J.P. (2003), "A Review of the Bootstrap for Time Series", International Statistical Review, 71, 435-459.

[37] Inoue, A. (1999), "Tests of Cointegrating Rank with a Trend-break", Journal of Econometrics, Elsevier, vol. 90(2), 215-237.

[38] Inoue,A., Kilian, L. (2002), "Bootstrapping Autoregressive Processes with Possible Unit Roots", Econometrica, Econometric Society, vol. 70(1), 377-391.

[39] Kontolemis G. (2002), "Money Demand in the Euro Area: Where Do We Stand (Today)?", IMF Working Papers 02/185, International Monetary Fund.

[40] Krämer, W. und Sonnberger, H. (1986), "The Linear Regression Model under Test", Physica, Heidelberg.

[41] Lutkepohl, H., Saikkonen, P. and Trenkler, C., (2003), "Comparison of Test for the Cointegrating Rank of a VAR Process with a Structural Shift",Journal of Econometrics 112, 201-229.

[42] Nyblom, J. (1989), "Testing for the Constancy of Parameters Over Time," Journal of the American Statistical Association, 84, 223-230.

[43] Phillips, P., Hansen, B. E. (1990), "Statistical Inference in Instrumental Variables Regression with I(1) Processes", Review of Economic Studies, Blackwell Publishing, vol. 57(1), pages 99-125.

[44] Ploberger, W. and Kramer, W (1992). "The CUSUM Test with OLS Residuals", Econometrica, 60(2), 271-285.

[45] Ploberger, W. und Kramer. W. (1996), "A Trend-resistant Test for Structural Change Based on OLS Residuals", Journal of Econometrics, 70, 175-185.

[46] Ploberger, W., Kramer, W. und Kontrus, K. (1989), "A New Test for Structural Stability in the Linear Regression Model", Journal of Econometrics, 40, 307-318.

[47] Quandt, R. (1960), "Tests of the Hypothesis that a Linear Regression Obeys two Separate Regimes", Journal of the American Statistical Association, 55, 324-330.

[48] Quintos, C. E. (1997), "Stability Tests in Error Correction Models," Journal of Econometrics, 82, 289-315.

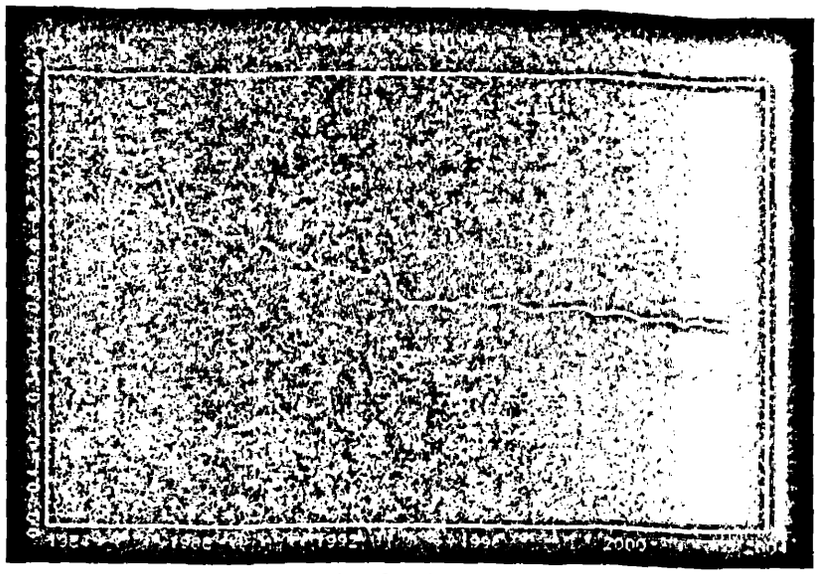
[49] Rao C.R., (1973). Linear Statistical Inference and its Applications. John Wiley and Sons, Inc.

[50] Seo, B. (1998), "Tests for Structural Change in Cointegrated Systems", Econometric Theory 14, 222-259.

[51] Tibergen, J. (1940). Statistical Testing of Business-Cycle Theories, Geneva: League of Nations, VOl I: A method and Its Application to Investment Activity.

[52] Vogelsang, T.J. (1997), "Wald-Type Tests for Detecting Shifts in the Trend Function of a Dynamic Time Series", Econometric Theory 13, 818-849.

[53] Zapata, H., Rambaldi, A. (1996), "Monte Carlo Evidence On Cointegration And Causation", Staff Papers 96-08, Louisiana State University, Agricultural Economics and Agribussines 1383-1414.



Recursive eigenvalue for the euro area M3 demand.

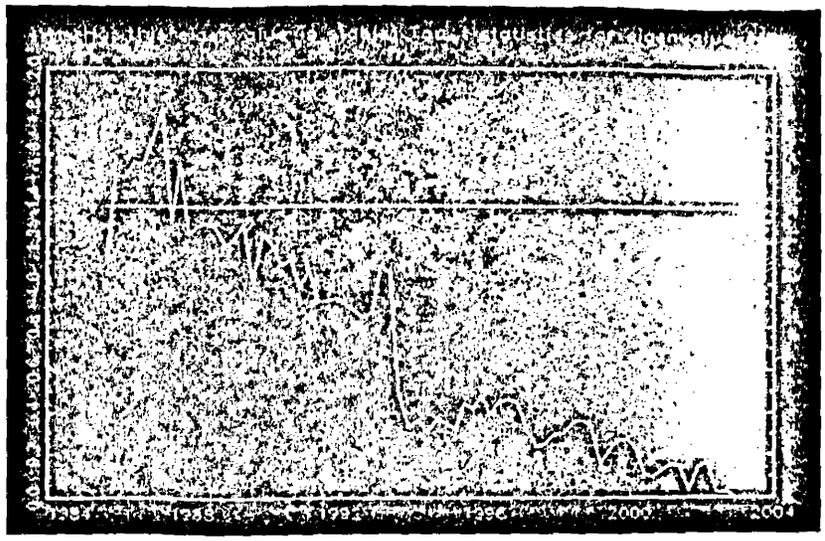


Figure 3.1. Recursive eigenvalue test for stability of the cointegration relation.

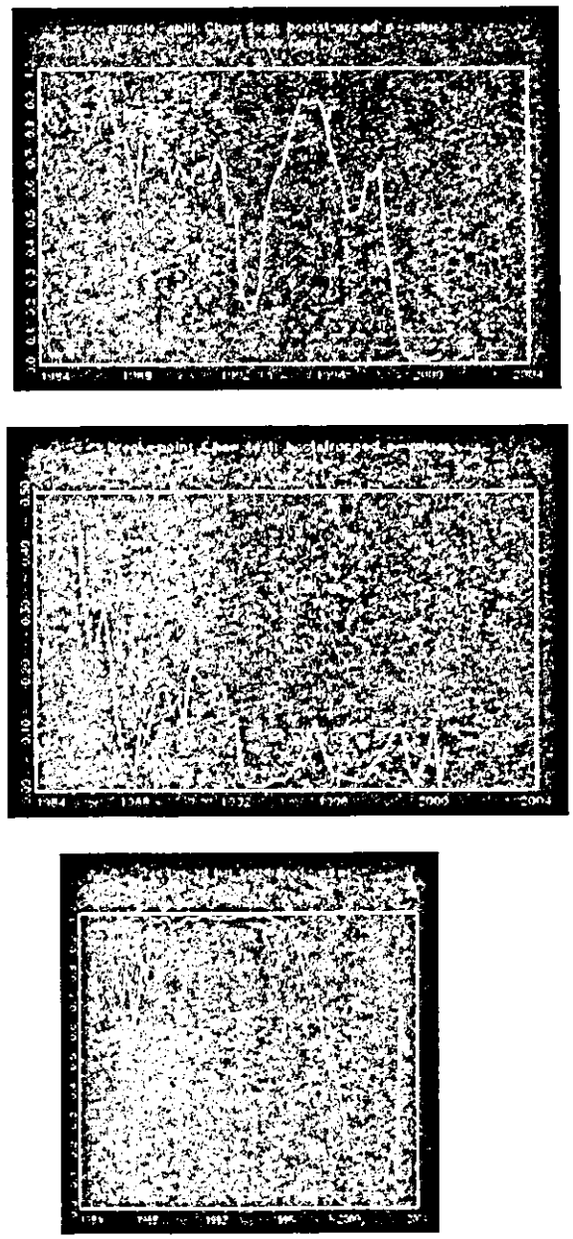


Figure 3.2. Graphs of bootstrapped Chow tests p-values.

Table 1. Relative rejection frequencies of the known-date breaks tests with nominal significance level of 5%

		Asymptotic			Bootstrapped		
		SS	BP	F	SS	BP	F
DGP 1	0.2T	0.112	0.224	0.223	0.053	0.045	0.043
(n=1,	0.5T	0.109	0.174	0.138	0.046	0.053	0.046
p=4)	0.8T	0.078	0.218	0.185	0.041	0.046	0.042
	0.2T	0.111	0.223	0.223	0.052	0.045	0.043
DGP 1A	0.5T	0.109	0.174	0.138	0.046	0.053	0.046
	0.8T	0.078	0.219	0.185	0.041	0.046	0.042
DGP 2	0.2T	0.073	0.154	0.190	0.045	0.041	0.042
(n=1,	0.5T	0.090	0.141	0.130	0.059	0.051	0.046
p=4)	0.8T	0.057	0.168	0.178	0.037	0.044	0.047
	0.2T	0.073	0.154	0.190	0.046	0.044	0.043
DGP 2A	0.5T	0.090	0.142	0.131	0.059	0.051	0.046
	0.8T	0.057	0.168	0.179	0.037	0.046	0.048
DGP 3	0.2T	0.111	0.387	0.099	0.035	0.036	0.044
(n=2,	0.5T	0.109	0.263	0.075	0.054	0.049	0.043
p=1)	0.8T	0.141	0.451	0.087	0.074	0.059	0.054
	0.2T	0.253	0.593	0.164	0.053	0.049	0.051
DGP 3A	0.5T	0.247	0.455	0.092	0.071	0.065	0.050
	0.8T	0.304	0.651	0.107	0.081	0.075	0.063
	0.2T	0.252	0.592	0.164	0.052	0.049	0.051
DGP 3B	0.5T	0.248	0.456	0.092	0.071	0.065	0.050
	0.8T	0.304	0.652	0.107	0.081	0.075	0.063
	0.2T	0.527	0.836	0.264	0.089	0.079	0.061
DGP 3C	0.5T	0.602	0.779	0.147	0.122	0.135	0.066
	0.8T	0.550	0.847	0.136	0.136	0.114	0.080
DGP 4	0.2T	0.225	0.972	0.166	0.056	0.039	0.053
(n=5,	0.5T	0.220	0.819	0.109	0.058	0.048	0.044
p=1)	0.8T	0.241	0.969	0.125	0.055	0.039	0.063

Notes:

1. Table shows the size of the sample-split (SS), break-point (BP) and forecast (F) tests. The sample size is 100 observations. Number of Monte Carlo simulations is 1000 with 1000 bootstrap replications. For the descriptions of the generated process DGP1-5 see the appendix.

Table 1 (continued). Relative rejection frequencies of the known-date breaks tests with nominal significance level of 5%.

	Break at	Asymptotic SS	Asymptotic BP	Asymptotic F	Bootstrapped SS	Bootstrapped BP	Bootstrapped F
DGP 5 I(1)-process n=2, p=1	**0.5T**	0.054	0.060	0.030	0.048	0.048	0.058
DGP 5A	**0.5T**	0.062	0.092	0.052	0.060	0.078	0.076
DGP 5B	**0.5T**	0.052	0.063	0.024	0.045	0.051	0.049
DGP 6 I(1)-process n=5, p=1	**0.5T**	0.187	0.229	0.012	0.060	0.049	0.062
DGP 6A	**0.5T**	0.469	0.677	0.006	0.063	0.051	0.031

Notes:

Table shows the size of the sample-split (SS), break-point (BP) and forecast (F) tests. The sample size is 100 observations. Number of Monte Carlo simulations is 1000 with 1000 bootstrap replications. For the descriptions of the generated processes DGP1-6 see the appendix.

Table 2. Relative rejection frequencies of the generalized Chow tests with nominal significance level of 5%

	Asymptotic			Bootstrapped		
	Sup	Ave	Exp	Sup	Ave	Exp
DGP 1	0.057	0.039	0.042	0.056	0.046	0.045
DGP 1A	0.056	0.039	0.042	0.056	0.046	0.045
DGP 2	0.030	0.026	0.026	0.046	0.047	0.036
DGP 2A	0.030	0.026	0.026	0.042	0.048	0.038
DGP 3	0.086	0.079	0.077	0.057	0.062	0.055
DGP 3A	0.248	0.248	0.237	0.084	0.073	0.064
DGP 3B	0.248	0.248	0.237	0.084	0.073	0.064
DGP 3C	0.694	0.718	0.731	0.146	0.150	0.178
DGP 4	0.043	0.043	0.041	0.057	0.058	0.043

Notes:

1. Table shows the size of the sample-split (SS), break-point (BP) and forecast (F) tests. The sample size is 100 observations. Number of Monte Carlo simulations is 1000 with 1000 bootstrap replications. For the descriptions of the generated process DGP1-5 see the appendix.

Table 3. Power of the bootstrapped tests when the break is generated in the middle of the sample.

	Break at:	SS	BP	F	Sup	Ave	Exp
DGP 2B	0.2T	0.500	0.446	0.088			
(n=1,	0.5T	0.990	0.990	0.478	0.960	0.984	0.990
p=4)	0.8T	0.236	0.202	0.080			
	0.2T	0.204	0.208	0.088			
DGP 2C	0.5T	0.998	0.998	0.586	0.984	0.996	0.980
	0.8T	0.062	0.102	0.038			
	0.2T	0.068	0.052	0.036			
DGP 2D	0.5T	0.062	0.064	0.032	0.062	0.066	0.048
	0.8T	0.046	0.052	0.062			
	0.2T	0.176	0.106	0.010			
DGP 2E	0.5T	0.044	0.248	0.002	0.108	0.098	0.044
	0.8T	0.008	0.064	0.000			
DGP 3B	0.2T	0.114	0.076	0.046			
(n=2,	0.5T	0.538	0.478	0.098	0.372	0.424	0.462
p=1)	0.8T	0.238	0.224	0.076			
	0.2T	0.066	0.058	0.044			
DGP 3C	0.5T	0.088	0.076	0.050	0.074	0.078	0.092
	0.8T	0.064	0.058	0.050			
	0.2T	0.078	0.070	0.050			
DGP 3D	0.5T	0.080	0.140	0.056	0.090	0.088	0.076
	0.8T	0.054	0.080	0.042			
DGP 4	0.2T	0.194	0.166	0.118			
(n=5,	0.5T	0.946	0.866	0.096	0.758	0.788	0.778
p=1)	0.8T	0.206	0.158	0.028			

Notes:

Table shows the empirical rejection probabilities of the tests when the data is generated under the alternative hypothesis of various structural breaks.The sample size is 100 observations. Number of Monte Carlo simulations is 1000 with 1000 bootstrap replications. For the descriptions of the generated process see the appendix to this chapter.

Sup, Ave and Exp-tests are generalised Chow-type tests from Andrews (1993) and show the probility of the break on trimmed subsample (0.15:0.85).

Table 3 (cont.). Power of the bootstrapped tests when the break is generated in the middle of the sample for I(1) processes.

	Break at:	SS	BP	F
DGP 6C (break in alfpha)	**0.5T**	0.080	0.075	0.115
DGP 6D (break in alfpha)	**0.5T**	0.700	0.620	0.120
DGP 6E (break in SR coeff.)	**0.5T**	0.097	0.066	0.057
DGP 6F (break in SR coeff.)	**0.5T**	0.130	0.090	0.060
DGP 6G (break in const)	**0.5T**	0.075	0.860	0.000
DGP 6H (break in const)	**0.5T**	0.085	0.985	0.000

Notes:

Table shows the empirical rejection probabilities of the tests when the data is generated under the alternative hypothesis of various structural breaks. The sample size is 100 observations. Number of Monte Carlo simulations is 1000 with 1000 bootstrap replications. For the descriptions of the generated processes see the appendix to this chapter.

Table 4. Power of the bootstrapped tests when the break is generated at the end of the sample.

	Break at:	SS	BP	F	Sup	Ave	Exp
DGP 2B	0.2T	0.106	0.136	0.120			
(n=1,	0.5T	0.700	0.694	0.386	0.922	0.952	0.948
p=4)	0.8T	0.974	0.942	0.802			
	0.2T	0.094	0.128	0.134			
DGP 2C	0.5T	0.804	0.848	0.618	0.990	1.000	0.994
	0.8T	1.000	0.998	0.988			
	0.2T	0.078	0.060	0.044			
DGP 2D	0.5T	0.080	0.082	0.050	0.088	0.090	0.100
	0.8T	0.082	0.070	0.056			
	0.2T	0.100	0.064	0.002			
DGP 2E	0.5T	0.062	0.090	0.124	0.062	0.060	0.038
	0.8T	0.044	0.010	0.000			
DGP 3B	0.2T	0.108	0.080	0.044			
(n=2,	0.5T	0.548	0.486	0.106	0.352	0.402	0.428
p=1)	0.8T	0.178	0.150	0.048			
	0.2T	0.062	0.056	0.058			
DGP 3C	0.5T	0.124	0.112	0.044	0.088	0.106	0.096
	0.8T	0.066	0.076	0.044			
	0.2T	0.094	0.078	0.042			
DGP 3D	0.5T	0.052	0.108	0.036	0.068	0.068	0.054
	0.8T	0.046	0.088	0.026			
DGP 4	0.2T	0.172	0.146	0.108			
(n=5,	0.5T	0.936	0.888	0.304	0.792	0.810	0.776
p=1)	0.8T	0.202	0.160	0.094			

Notes:

Table shows the empirical rejection probabilities of the tests when the data is generated under the alternative hypothesis of various structural breaks.The sample size is 100 observations. Number of Monte Carlo simulations is 1000 with 1000 bootstrap replications. For the descriptions of the generated process see the appendix to this chapter.

Sup, Ave and Exp-tests are generalised Chow-type tests from Andrews (1993) and show the probility of the break on trimmed subsample (0.15:0.85).

4.A Models used in the study

The data-generating model to study the test performance in VAR models is of the form

$$y_t = \Phi D_t + A_1 y_{t-1} + ... + A_p y_{t-p} + u_t \qquad (t = 1, ..., T)$$

where D is a matrix $(m \times T)$ of deterministic terms and dummies, Φ is $(n \times m)$ matrix, the A_i are $(n \times n)$ coefficient matrices and $u_t = (u_{1t}, ...u_{nt})'$ is an unobservable zero mean white noise process, with time invariant positive definite covariance matrix Σ, i.e. $u_t \backsim iid(0, \Sigma)$.

To investigate the size and power of the tests, several different DGP's were used. The first DGP is a univariate system with four lags, and a constant simulated on the basis of quarterly Polish productivity series (1970 Q1 - 1998 Q4), where

$$\begin{aligned} \Phi &= 0.11, \ a_1 = 0, \ a_2 = -0.147, \qquad \text{(DGP1)} \\ a_3 &= -0.218, \ a_4 = 0.423, \ \Sigma = 0.002 \end{aligned}$$

The second DGP is, again a univariate system with four lags, a constant and a trend simulated on the basis of the quarterly US investment series (1947 Q2-1972

Q4) where

$$\Phi = \{0.819,\ 0.5\},\ a_1 = 0.514,\ a_2 = -0.098, \qquad \text{(DGP2)}$$

$$a_3 = -0.064,\ a_4 = -0.218,\ \Sigma = 6.637$$

The third DGP is an artificially generated bivariate system with one lag and a constant term, where

$$\Phi' = \{0.2\ 0.2\},\ A_1 = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix},\ \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{(DGP3)}$$

The fourth DGP is a five-dimensional VAR model, with one lag based on the short-run dynamics of the model of the EMU aggregate monetary system ($m - p$, π, i_l, i_s, y) from Brand and Cassola (2000) where

$$\Phi' = \{-0.212, 0.023, 0, -0.047, 0.002\} \qquad \text{(DGP4)}$$

$$A_1 = \begin{pmatrix} 0.481 & 0 & -0.99 & -0.385 & -0.183 \\ 0 & -0.217 & 0 & 0.508 & 0 \\ 0.034 & 0 & 0.586 & 0 & 0 \\ 0.039 & 0 & 0.284 & 0.266 & -0.039 \\ 0.333 & 0.544 & 0 & 0 & 0 \end{pmatrix},$$

$$\Sigma = \begin{pmatrix} 1 & -0.487 & 0.319 & 0.156 & 0.091 \\ & 1 & 0.004 & -0.011 & 0.183 \\ & & 1 & 0.596 & 0.187 \\ & & & 1 & 0.171 \\ & & & & 1 \end{pmatrix}$$

When testing the size of the tests apart from DGP 1-5, some of their modifications were used in order to acquire better understanding of tests performance un-

der different conditions. First, for DGP1, variance was increased tenfold. Next, the tests were run for the model as in DGP2, but with a higher trend coefficient. For two-dimensional models, we added cross-correlation of the residuals (DGP3A), finally increasing the diagonal elements of the first-lag coefficient matrix, in order to see the performance of the tests when the DGPs behave more like non-stationary processes (DGP3B and DGP3C).

$$as\ DGP1,\ with\ \Sigma = 0.02 \qquad \text{(DGP1A)}$$

$$as\ DGP2,\ with\ \Phi = \{0.819,\ 0.9\} \qquad \text{(DGP2A)}$$

$$as\ DGP3\ with\ \ \Sigma = \begin{pmatrix} 1 & -0.3 \\ -0.3 & 1 \end{pmatrix} \qquad \text{(DGP3A)}$$

$$as\ DGP3\ with\ \ A_1 = \begin{pmatrix} 0.8 & 0 \\ 0 & 0.8 \end{pmatrix} \qquad \text{(DGP3B)}$$

$$as\ DGP3\ with\ \ A_1 = \begin{pmatrix} 0.98 & 0 \\ 0 & 0.98 \end{pmatrix} \qquad \text{(DGP3C)}$$

For testing the power, the following models were used. First, a break was simulated in the trend for the US investment series DGP (DGP2):

$$as\ DGP2\ with\ \Phi_{<B} = \{0.819\ 0.5\},\ \Phi_{B>} = \{0.819\ 0.9\} \qquad \text{(DGP2B)}$$

$$as\ DGP2\ with\ \Phi_{<B} = \{0.819\ 0.5\},\ \Phi_{B>} = \{0.819\ 0.1\} \qquad \text{(DGP2C)}$$

Next, with the same model, a break in the mean was modelled as follows:

$$as\ DGP2\ with\ \Phi_{<B} = \{0.819\ \ 0.5\},\ \Phi_{B>} = \{1.6\ \ 0.5\} \qquad \text{(DGP2D)}$$

Further, simulation of a break in the variance for the DGP2 was undertaken:

$$as\ DGP2\ with\ \Sigma_{<B} = 6.637,\ \Sigma_{B>} = 3 \qquad \text{(DGP2E)}$$

And for the DGP3 process, the power of the tests was analysed by adding a constant to the diagonal elements of the coefficient matrix:

$$as\ DGP3\ with\ A_{1<B} = \begin{pmatrix} 0.4 & 0 \\ 0 & 0.4 \end{pmatrix},\ A_{1B>} = \begin{pmatrix} 0.8 & 0 \\ 0 & 0.8 \end{pmatrix} \qquad \text{(DGP3B)}$$

The penultimate process, again obtained from DGP3, was to add off-diagonal elements to the coefficient matrix after the break:

$$as\ DGP3\ with\ A_{1<B} = \begin{pmatrix} 0.4 & 0 \\ 0 & 0.4 \end{pmatrix},\ A_{1B>} = \begin{pmatrix} 0.4 & 0.1 \\ -0.2 & 0.4 \end{pmatrix} \qquad \text{(DGP3C)}$$

For the DGP3, lastly, correlation among the residuals after the break was added:

$$as\ DGP3\ with\ \Sigma_{1<B} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ \Sigma_{1B>} = \begin{pmatrix} 1 & 0.3 \\ 0.3 & 1 \end{pmatrix} \qquad \text{(DGP3D)}$$

Finally, for the five variable system, the break was modelled by removing the third column from the coefficient matrix, except for the main diagonal element:

$$As\ DGP4\ with\ A_{1B>} = \begin{pmatrix} 0.481 & 0 & 0 & -0.385 & -0.183 \\ 0 & -0.217 & 0 & 0.508 & 0 \\ 0.034 & 0 & 0.7 & 0 & 0 \\ 0.039 & 0 & 0 & 0.266 & -0.039 \\ 0.333 & 0.544 & 0 & 0 & 0 \end{pmatrix} \tag{DGP4A}$$

The data-generating model to study test performance in CVAR models is of the form:

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \mu + \Phi D_t + \epsilon_t, \quad t = 1, ..., T \tag{4.13}$$

for fixed initial values of $X_{-k+1}, ..., X_0$ and $\epsilon_1, ..., \epsilon_T$ being identically and independently distributed $N_p(0, \Lambda)$ errors, μ being a constant, and D_t containing the deterministic terms of the model. The rank of the matrix Π determines the number of cointegrating vectors and is of the form:

$$\Pi = \alpha\beta' \tag{4.14}$$

where β's columns are linearly independent cointegrating vectors, and α is the adjustment matrix of factor-loading vectors.

For bivariate systems the following specifications were applied:

$$\begin{aligned} \alpha &= \{-0.3\ 0.1\},\ \beta = \{1 - 1\} \\ \Gamma_1 &= \{ \begin{matrix} 0.5 & 0 \\ 0 & -0.5 \end{matrix} \},\ \epsilon_t \tilde{} N_2(0, 1) \end{aligned} \tag{DGP5}$$

$$\alpha = \{-0.3\ 0.1\},\ \beta = \{1 - 0.2\} \qquad \text{(DGP5A)}$$
$$\Gamma_1 = \{ \begin{matrix} 0.5 & 0 \\ 0 & -0.5 \end{matrix} \},\ \epsilon_t \ \tilde{}\ N_2(0,1)$$

$$\alpha = \{-0.3\ 0.1\},\ \beta = \{1 - 0.5\} \qquad \text{(DGP5B)}$$
$$\Gamma_1 = \{ \begin{matrix} 0.5 & 0 \\ 0 & -0.5 \end{matrix} \},\ \Gamma_2 = \{ \begin{matrix} -0.2 & 0.1 \\ -0.2 & 0 \end{matrix} \},\ \epsilon_t \ \tilde{}\ N_2(0,1)$$

For five-variables systems, the DGP based on the euro area money demand system drawn from Carstensen (2004) was used, with one lag, cointegrating rank of one, and a constant restricted to the cointegrated space:

$$\alpha = \{-0.151 - 0.052 \quad 0.453 \ -0.082 \quad -0.577\},\ \Phi = 8.312 \qquad \text{(DGP6)}$$
$$\beta = \{1 \quad -1.249\ 0.124 \quad 1.866 \quad -0.040\}$$
$$\Gamma_1 = \begin{pmatrix} 0.46 & -0.087 & 0.001 & -0.173 & -0.008 \\ 0.221 & 0.243 & 0.021 & 0.172 & -0.006 \\ 0.502 & 1.289 & 0.229 & -1.433 & 0.018 \\ -0.165 & -0.193 & 0.029 & 0.426 & -0.006 \\ -1.124 & -0.080 & 0.435 & 0.831 & 0.507 \end{pmatrix},$$
$$\Lambda = \begin{pmatrix} 1.1e-05 & 8.3e-07 & 2.3e-05 & -2.1e-06 & 6.6e-05 \\ & 2.5e-05 & 1.8e-05 & 1.3e-06 & -1.6e-05 \\ & & 1.1e-03 & 2.8e-06 & -3.7e-04 \\ & & & 1.2e-05 & -3.4e-05 \\ & & & & 6.2e-03 \end{pmatrix}$$

For process DGP 6A the same specification as DGP6 was employed, but instead of fixing the beta at its full-sample estimate, an estimate was made of long-run vector on each subsample.

For power studies the following specifications were used:

$$as\ DGP6\ with\ \alpha_{B>} = \alpha_{B<} * 1.1 \qquad \text{(DGP6C)}$$

$$as\ DGP6\ with\ \alpha_{B>} = \{0 \quad -0.052 \quad 0.453 \quad -0.082 \quad -0.577\} \qquad \text{(DGP6D)}$$

$$As\ DGP6\ with\ \Gamma_{1B>} = \begin{pmatrix} 0.46 & -0.087 & \mathbf{0} & -0.173 & -0.008 \\ 0.221 & 0.243 & \mathbf{0} & 0.172 & -0.006 \\ 0.502 & 1.289 & \mathbf{0.229} & -1.433 & 0.018 \\ -0.165 & -0.193 & \mathbf{0} & 0.426 & -0.006 \\ -1.124 & -0.080 & \mathbf{0} & 0.831 & 0.507 \end{pmatrix} \qquad \text{(DGP6E)}$$

$$As\ DGP6\ with\ \Gamma_{1B>} = \begin{pmatrix} 0.46 & -0.087 & 0.001 & \mathbf{0} & -0.008 \\ 0.221 & 0.243 & 0.021 & \mathbf{0} & -0.006 \\ 0.502 & 1.289 & 0.229 & \mathbf{-1.433} & 0.018 \\ -0.165 & -0.193 & 0.029 & \mathbf{0} & -0.006 \\ -1.124 & -0.080 & 0.435 & \mathbf{0} & 0.507 \end{pmatrix} \qquad \text{(DGP6F)}$$

$$as\ DGP6\ with\ \Phi_{B>} = \Phi_{B<} * 0.8 \qquad \text{(DGP6G)}$$

$$as\ DGP6\ with\ \Phi_{B>} = \Phi_{B<} * 0.5 \qquad \text{(DGP6H)}$$