International Inflation Linkages and Forecasting in the Presence of Structural Breaks

A thesis
submitted to The University of Manchester
for the degree of
doctor of philosophy (PhD)
in the faculty of Humanities

2013

Gantungalag Altansukh

School of Social Sciences/Economics

Contents

List of Tables

List of Figures

Abstract

The University of Manchester
Gantungalag Altansukh
Doctorate of Philosophy in Economics
International Inflation Linkages and Forecasting in the Presence of Structural Breaks
2013

This PhD thesis consists of three related chapters; each contributes to the study of inflation dynamics by examining different issues that have previously been raised in the relevant literature. In particular, the first chapter is concerned with the nature of different changes that have taken place in the conditional mean and variance of inflation. To shed light on this question, an iterative structural break testing methodology is developed which allows the possibility of distinct changes in the conditional mean and variance components by iterating tests between them, with outliers also identified in relation to regimes. This methodology is applied to models that link domestic and foreign inflation, and uncovers a positive and strengthening contemporaneous relationship between domestic and foreign inflation, adding to the literature that provides evidence of increasing globalization of inflation.

The second chapter sheds further light on the nature of the globalization of inflation by separating core, energy and food components of aggregate inflation, analyzing changes in the international links in these separate components. Comparison with the aggregate inflation reveals that the overall globalization is driven largely by the mean levels of core inflation being very similar across countries, especially from the early 1990s. Further, an increased synchronization of short-run movements in non-core (energy and food) components contribute to the overall globalization effect, but such short-run synchronization is less evident in core inflation.

The first and second chapters show that structural breaks either in the conditional mean or variance parameters of inflation are a common feature. Therefore, the third chapter focuses on the problem of forecasting in the presence of structural breaks. Specifically, chapter 3 proposes a forecast method which allows for break date uncertainty by employing a confidence interval estimate of the break date. A Monte Carlo simulation study and an empirical application to inflation time series demonstrate the usefulness of this approach. This chapter also proposes an algorithm that re-orders time series data based on the similarity of regimes. It is shown that such re-ordering can improve forecast accuracy when estimation exploits the additional information provided by the re-ordered series. These improvements are significant when there are multiple breaks which have the form of reverting coefficients.

Declaration

Candidate name: Gantungalag Altansukh
Faculty: Humanities
Thesis Title: International Inflation Linkages and Forecasting in the Presence of Structural Breaks

Declaration:

I declare that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Signed: Date:

Copyright

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialization of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy[1], in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations[2] and in The University's policy on Presentation of Theses.

[1](see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487)
[2](see http://www.manchester.ac.uk/library/aboutus/regulations)

Acknowledgments

Foremost, I would like to express my deepest gratitude to Prof. Denise R Osborn, my Principal Supervisor, and Dr. Ralf Becker, my Associate Supervisor, who always have been available for my questions and invested much more time and effort than I would dared to ask for. It would not have been possible to write this doctoral thesis without their continuous support, patience, enthusiasm, immense knowledge and experience. I would also like to thank Dr. George J Bratsiotis, my Associate Supervisor, for his valuable advice on my research and manuscripts.

I wish to acknowledge the generous financial support I received from the School of Social Sciences at the University of Manchester for my research. Without this support and the facilities provided to me, my last three years would not have been as enjoyable and successful. I also thank to all members of staffs, fellow students and friends for providing the stimulating intellectual environment, encouragement and all the fun we have had in the last three years.

Finally, I thank to my husband, Surenragchaa Nyamjav, for his continuous support throughout. He stood by me all of the ups and downs of my research. I also wish to thank my parents, Altansukh Ishee and Bolormaa Nachin, for their support and encouragement with their best wishes.

For any errors or inadequacies that may remain in this work, of course, the responsibility is entirely my own.

Introduction

A number of studies, including Monacelli and Sala (2009), Ciccarelli and Mojon (2010), Neely and Rapach (2011), Mumtaz and Surico (2012) have commented on the remarkable degree of similarity shown by inflation across many countries over the last few decades. Inflation was high and volatile during the 1970s until the mid 1980s, and subsequently declined, except around 2007 when inflation has increased for a few countries. The most notable decline occurred in the early to mid 1990s for most industrialized countries, bringing levels of inflation closer to each other, and it has remained low and stable in recent years. This increased similarity is referred as the 'Globalization of Inflation' and has attracted a good amount of attention. The main question posed in the literature is whether this development is due to the increased importance of international, as opposed to domestic factors. Indeed, the interpretation for this phenomenon is important as it may carry implications for the conduct of monetary policy. If the observed inflation dynamics is the result of a common global feature, then national central banks need to monitor carefully international price developments and analyze their implications for the domestic economy (Bernanke, 2007, Trichet, 2008, Mumtaz et al., 2011). Further, it raises the question of the extent to which any individual country can successfully pursue an inflation targeting policy.

This thesis contributes to this literature, from the perspectives of both understanding (possibly changing) international linkages and the importance of taking account of structural breaks for forecasting inflation. In particular, a structural break analysis is an integral part of all the contributions made in this thesis. This is due to the fact that structural instability is found to be a prevalent feature of inflation series, as documented in the context of univariate inflation models (see Benati and Kapetanios, 2002, Gadzinski and Orlandi, 2004, Corvoisier and Mojon, 2005, Cecchetti and Debelle, 2006, among others).

The first chapter of the thesis measures the degree of inflation co-movement across many industrialized countries. A number of relevant studies exist in the literature that focus on measuring international inflation dependence and analyzing its effects for shaping the domestic inflation rate. The main finding

of the chapter which is common to that in many independent studies, is the existence of a strong link between domestic inflation and the international environment (Pesaran et al., 2004, Ciccarelli and Mojon, 2010, Mumtaz and Surico, 2012, and many others).

Specifically, the contributions of chapter 1 are twofold. Firstly, as part of the methodological contribution, an iterative structural break testing methodology is developed which is later used to pin down the nature and date(s) of change in international inflation linkages. Detection of structural breaks has a long history and there are a number of available testing methods including Andrews (1993), Bai (1997), Bai and Perron (1998). The proposed methodology builds upon them but attempts to improve the identification of distinct changes in different components of a series by testing for breaks in one component conditional on another; see also Bataa et al. (2013b).

Secondly, we analyze the changing dynamics of domestic inflation in relation to the international environment to answer the question: Has the globalization of inflation deepened? This issue is often addressed using either subsample analysis or allowing for random coefficient variation in the model. Results from these studies indicate an increased role of the global factor in explaining variations of domestic inflation (Ciccarelli and Mojon, 2010, Mumtaz and Surico, 2012, Neely and Rapach, 2011, among others). Consistent with this result, our analysis based on the iterative structural break testing methodology uncovers a positive and strengthening contemporaneous relationship between domestic and country specific foreign inflation.

The second chapter of the thesis sheds further light on the nature of the globalization of inflation. Although important for policy purposes, understanding the sources of inflation globalization is not easy and little has been revealed to date. Perhaps, the most supported hypothesis is that the increased co-movement of inflation is a consequence of the adoption of similar monetary policies across countries. This includes a currency peg in many countries that are now part of the Euro Area (Altissimo et al., 2006, Borio and Filardo, 2007), and the conduct of tighter monetary policies such as inflation targeting by a number of countries in the 1990s (Altissimo et al., 2006, Mumtaz and Surico, 2012) and a strong response to expected inflation in the US (Clarida et al., 2000, Ahmed et al., 2004). However, recent studies show that monetary policy shocks are unable to fully explain inflation dynamics, suggesting that the co-movement of inflation should be explained beyond the conventional view that inflation is a pure monetary phenomenon (Wang and Wen, 2007, Canova and Ferroni, 2012). Other studies suggest alternative explanations, including a role for the foreign output gap (Borio and Filardo, 2007, Ihrig et al., 2010), cheaper

imported goods (Peacock and Baumann, 2008), openness to trade (Monacelli and Sala, 2009) for the co-movement of inflation. However, their corresponding empirical results do not largely support the observed inflation dynamics across countries. Perhaps, a number of sources and channels are jointly responsible for this and exploring them is beyond the scope of a single study alone, neither is this what chapter 2 strives to do.

Specifically, chapter 2 attempts to shed light on the existing studies through the use of inflation components while most inflation studies use 'headline' inflation data. The components studied are core, energy and food inflation. An analysis of core inflation may provide clearer evidence on the role of monetary policy, as it is frequently seen as the appropriate concept for monetary policy purposes (Mishkin, 2007). Further, energy inflation is anticipated to have a strong international dimension and also there have been large changes in food supply for developed economies over the last forty years, moving from predominantly domestically produced to being largely imported. Therefore, the possibility of increased international co-movement for these components is assessed.

Two types of analysis are made in this chapter, both of which are based on the examination of structural breaks through the use of the iterative algorithm, developed in chapter 1. Firstly, a dynamic model that links domestic and foreign series (aggregate or components) are analyzed for many countries. The most prominent characteristic from this analysis is a convergence in the mean rates for each of the aggregate and components for the countries in our data set. This provides evidence for the globalization of inflation, which appears to be clear also across components as well as for the aggregate. Secondly, we dissect the globalization of aggregate inflation, by examining whether breaks in international linkages at the aggregate level can be attributed to changes in the responses of domestic inflation to foreign components. Our analysis shows an increased role for the foreign core component for a number of European countries, pointing to the importance of monetary policy in the context of the formation of the Euro Area.

The subject of the third chapter is a little distinct from the previous two chapters. However, structural breaks remain a key aspect of this study. Particularly, chapter 3 focuses on the problem of forecasting in the presence of structural breaks. This is important because structural instability is a generic feature of many macroeconomic time series (Stock and Watson, 1996) and it is a major cause of forecast failure if unaccounted for (Hendry and Clements, 2003). The existing literature includes various forecast techniques which deal with different causes of forecasts failure such as structural break and model

uncertainties. The common objective of all methods is to deliver increased forecast accuracy, often measured as mean squared forecast error. Chapter 3 is no exception to this and proposes a forecast method to deal with historical discrete breaks that may have occurred during the estimation sample.

Ideally, if information such as the date and the size of a break is known, this can be used to select an optimal estimation sample by exploiting the bias-variance trade off (Pesaran and Timmermann, 2007) or optimal observational weighting scheme in the estimation sample (Pesaran et al., 2013). More simply, a post break window forecast which includes observations only after the break could be employed which often yields good forecast accuracy. However, the forecast accuracies of methods that exploit information on breaks heavily rely on how well the true break date is estimated. However, in practice, the estimates of break dates can be imprecise and this rules out achieving the full efficiency of these methods. Indeed, precise identification of break dates depends on a number of parameters that researchers have to assume but often do not have prior knowledge about, including the size, the number of breaks, and frequency of breaks.

On the other hand, it becomes attractive to adopt robust method such as averaging forecasts from different windows because it does not require exact break information, yet it yields consistently good performance in many cases (Pesaran and Timmermann, 2007, Pesaran and Pick, 2011, Clark and McCracken, 2009, Eklund et al., 2013, Tian and Anderson, 2011, and many others). Our proposed method comes from the same perspective that a combination technique is preferable rather than relying on potentially poor point estimates of breaks. However, we employ information on breaks for choosing the range of windows to be averaged. More specifically, we employ a confidence interval for the estimated break date where each date in the confidence interval is treated as one of a sequence of choices for the potential break date, and the corresponding post break window forecasts are averaged. In this way, forecasts from windows which include observations that are likely to belong to the current regime are averaged. Therefore, our approach can be seen as an improvement on existing methods that combine forecasts from all possible windows, many of which may yield large forecast errors, consequently leading to distortions in overall forecast accuracy. As a result of the extensive Monte Carlo simulations and empirical analysis of univariate inflation models, chapter 3 reports an overall good performance of the proposed method in the presence of large and small breaks that occur in the coefficients of the forecast model. Furthermore, the iterative structural break testing methodology is useful when dealing with distinct breaks occurring in the coefficients and variances of the

series as it yields an improvement in forecasting accuracy for the relevant cases.

Chapter 3 also considers a situation where time series may exhibit a regime-switching process, in which the detected structural breaks capture switches between two or more distinct (but recurring) regimes. When multiple breaks are present, most existing methods use the information on the most recent break only, which may not be ideal for regime-switching processes as it ignores the fact that data prior to any previous break can be informative with regard to the forecast. Therefore, this chapter proposes an algorithm that re-orders time series data based on the similarity of regimes to exploit the additional information. It is shown that such re-ordering can improve forecast accuracy of all methods considered, and the improvement is substantial when coefficient changes are reversed with multiple breaks.

To sum up, this thesis contributes to the literature that studies the role of structural breaks for inflation. In particular, it studies the nature of the globalization of inflation and proposes methods to deal with breaks when forecasting inflation or other variables.

Unless otherwise stated, all computations are performed using MATLAB version R2011a. Also, except where indicated otherwise, the associated programs have been written as part of the research for this PhD thesis.

Chapter 1

Structural Breaks in International Inflation Linkages

1.1 Introduction

Over the last decade or so, policymakers and researchers have documented and discussed the globalization of inflation, namely the apparently strong co-movement of inflation seen over the last two decades or more. Indeed, even in the context of the large economies of the US and Euro area, Bernanke (2007) and Trichet (2008), respectively, emphasize that their central banks now need to monitor carefully international price developments and analyze their implications for the domestic economy. The strong link between domestic inflation and the international environment is also recognized in the models of Pesaran et al. (2004), Ciccarelli and Mojon (2010), Mumtaz and Surico (2012) and many others. However, Bataa et al. (2013a) is, to our knowledge, the only paper that attempts to pin down the nature and dates of change in international inflation linkages between specific countries.

Studies of the globalization of inflation predominantly employ factor analysis to extract a common international inflation component. Changes in co-movement in relation to this factor are then studied for individual countries, using either subsample analysis or allowing for random coefficient variation (see Ciccarelli and Mojon, 2010, Monacelli and Sala, 2009, Mumtaz and Surico, 2012, Neely and Rapach, 2011). Nevertheless, implicit and untested assumptions about parameter constancy are required in order to extract the factor(s), and changing covariances could make these unreliable. This is established in a univariate context by Pitarakis (2004), who shows that serious size distortions arise in testing for mean (and/or persistence) change when volatility is assumed to be stable but is, in fact, subject to breaks.

This chapter studies the globalization of inflation by applying an iterative structural break testing methodology to model the link between domestic and country-specific foreign inflation. While the multiple break testing methodology of Bai and Perron (1998) provides the basic building block, our procedure allows the possibility of distinct changes in the model coefficients and in volatility. Not only does the presence of volatility breaks affect inference on coefficients, as shown by Pitarakis (2004), but inference on volatility breaks can be misleading if the computed residuals are contaminated by un-modeled mean breaks (Sensier and van Dijk, 2004). Further, ignoring the presence of outliers can lead to misspecification and biases in estimated parameters (see, among others, Giordani et al., 2007, Chen and Liu, 1993). Therefore, to avoid these problems, breaks in the conditional mean and variance parameters are identified by iterating between mean and variance tests, with outliers also identified in relation to conditional mean and volatility regimes. This methodology is closely related to, and builds upon, that of Bataa et al. (2013a,b).

In these models, country-specific foreign inflation is constructed as the bilateral trade weighted average of inflation in all other countries in our sample and is treated as weakly exogenous. As a preliminary step to the bivariate[1] inflation models linking domestic and foreign inflation, univariate inflation models are employed to examine the stability of domestic inflation and to assess the robustness of existing univariate findings. Although there is a substantial literature on breaks in univariate inflation models, including Cecchetti and Debelle (2006), O'Reilly and Whelan (2005), Levin and Piger (2003), Bataa et al. (2013b), the tests applied in almost all papers make the unrealistic assumption that the variance of inflation is constant over time. Our main focus, however, is analyzing changes in the linkage of domestic with international monthly CPI inflation for 19 OECD countries over the period January 1970 to September 2010. Further, to be clear, we do not attempt to identify potential reasons for or channels of international interdependencies in this chapter.

Our main findings can be summarized as follows. Firstly, univariate inflation models yield inference on breaks in the conditional mean that are broadly consistent with the existing literature. However, the number of conditional mean (that is, intercept and/or dynamic) breaks found in our analysis is fewer compared to other studies (see Bataa et al., 2013b, Cecchetti and Debelle, 2006). Secondly, we document clusters of variance breaks occurring around the mid 1970s, early 1980s and early 1990s, while only clusters of mean breaks

[1] All the models in this chapter are single equation models. In other words, we use the term "bivariate model" to refer to the model that shows the relationships between two variables, namely domestic and foreign inflation, where their lagged and contemporaneous terms are also allowed.

have been widely documented in the previous literature. These variance breaks typically reflect substantial declines in the volatility of inflation, casting doubt on the common claim in the literature that changes of inflation have been mainly in the mean. Thirdly, examining bivariate inflation models, we find a positive and strengthening contemporaneous relationship between domestic and country specific foreign inflation. Furthermore, the timing of break dates in conditional means and variances, identified using bivariate inflation models, also exhibit a clustering pattern around the mid 1970s, early to mid 1980s and early 1990s, suggesting commonality in changes to international inflation linkages. These bivariate inflation models also confirm a general pattern of declining persistence of domestic inflation.

The rest of the chapter is organized as follows. Section 1.2 reviews the literature on international inflation and structural break analysis. Section 1.3 describes our methodology, including our iterative procedure for structural break detection. Section 1.4 then presents the data and section 1.5 reports the results of both the univariate and bivariate inflation analyses. A sensitivity analysis is presented in section 1.6 and section 1.7 concludes.

1.2 Review of Literature

This chapter relates to two strands of the empirical literature, namely, modeling international inflation and testing for structural breaks. The reviews on both strands of literature are arranged in the following separate subsections.

1.2.1 International Inflation

Inflation dynamics across countries display a remarkable degree of similarity over the last few decades. Specifically, inflation becomes low and less volatile after the mid 1980s compared to a high and volatile period between the 1970s and the early 1980s for most developed countries investigated in this chapter. Understanding the reasons for the observed co-movement is not easy, despite it being a focus of many studies.

Recent studies seem to suggest that the co-movement of inflation should be explained beyond the conventional view that inflation is a pure monetary phenomenon. Wang and Wen (2007) document that neither money growth across countries is systematically correlated nor that country specific monetary shocks produce a co-movement across countries based on OECD countries' data. Although it may not necessarily point to the absence of a role for monetary policy in explaining the co-movements, yet it does not support the hypothesis

that the co-movement of inflation is due primarily to monetary policy coordination[2]. Similarly, the recent study by Canova and Ferroni (2012) shows that monetary policy shocks are unable to capture inflation dynamics fully in the US, although they partially explain volatility declines. They also note that the level of inflation becomes less responsive to monetary shocks over time.

A broader perspective is the increased globalization that may have affected domestic inflation through cheaper imported goods. Cheaper imported goods engender increased competition, lower production costs, and balance demands in different countries. In that respect, some studies include import prices or global output gaps into a structural model of inflation as a proxy for the global environment. For example, assuming perfect substitutability of domestic and imported goods, Peacock and Baumann (2008) consider cheaper imported goods which may exert downward pressure on domestic prices and hence lower inflation for being a main source of increased globalization. They test their hypothesis based on the structural New Keynesian Philips Curve (NKPC) by including an intermediate import price in the firm's marginal cost using data from the United Kingdom, United States and Japan. Overall results suggest that import prices help in explaining inflation to some extent, but the influence is relatively small and constant over time (see also Ihrig et al., 2010, for similar results). This seems to suggest that an ever deepening globalization has left little mark on international price linkages. It may be, according to Ball (2006), sharp changes in relative prices matter, and any small or smooth changes do not have any visible impacts on general price changes.

The literature also documents that the use of an output gap as a global proxy is less informative in explaining the increased co-movements, as opposing evidences arise from different studies. Borio and Filardo (2007) find an increased role of a global output gap[3] in explaining domestic inflation. However, their results are criticized as being sensitive to model specification (Mishkin, 2009, Ihrig et al., 2010). A number of studies employing similar data sets (OECD countries) find less significant evidences for the global output gap when different specifications are employed (see Ihrig et al., 2010, Ball, 2006, Calza, 2009, for industrial countries). For instance, Calza (2009) employs both

[2]Countries may adopt similar monetary policies due to a pegged currency. Specifically, trade and financial integration may cause countries to peg their exchange rates to a larger currency area which allows monetary shocks in larger economies to affect smaller economies, hence the co-movements of domestic inflations (Borio and Filardo, 2007). This hypothesis may be more relevant to Euro area countries. There has been a common monetary policy across Euro area countries since 1999, but exchange rates have been linked across many of these countries since 1979.

[3]The global output gap is measured by a weighted average of output gaps in other countries.

backward looking and forward looking Philips Curves augmenting them by a global output gap and finds the corresponding coefficient to be insignificant except in a backward looking Philips Curve where a level change is taken into account.

The approaches reviewed above account for only a single channel that links inflation across countries. Alternatively, Pesaran et al. (2004) propose a Global Vector Autoregressive (GVAR) method in which interactions of various macroeconomic variables across markets and across countries are examined simultaneously (see also Dees et al., 2007, for further details). Although it may be appropriate to consider more than one variable, the related studies are broader than the framework of our study and not particularly focused on inflation. Therefore, we do not discuss the GVAR methodology in detail. Further, we do not attempt to identify potential reasons for or channels of inflation co-movements in this chapter. Instead we attempt to measure the impact of international interdependencies on the domestic inflation and analyze its changes over time. In this respect, our study relates to a common factor approach.

There are several studies that address the international co-movement of inflation by extracting a common factor from various cross country data. This method is well suited for disentangling country specific and globally common shocks to inflation. The principal commonality measure of inflation they propose is to deduce the share of domestic inflation variance attributable to a common factor fluctuation. The corresponding results are largely consistent with each other, indicating that global inflation has a sizable role in explaining domestic inflation fluctuations with the implication of inflation being largely a global phenomenon (Ciccarelli and Mojon, 2010, Monacelli and Sala, 2009, Neely and Rapach, 2011, etc).

For instance, Ciccarelli and Mojon (2010) document that almost 70% of the variance of domestic inflation in 22 OECD countries is explained by one global factor using over 45 years data. A similar conclusion is drawn by Monacelli and Sala (2009), although with less explanatory power of the common factor, based on four OECD countries' sectoral level data. Further research, analyzing a variance decomposition of domestic inflation with respect to the regional factor in addition to the world and country specific components, is conducted by Neely and Rapach (2011). Employing the dynamic latent factor model on the extensive data set of 64 countries' inflation, they find on average 35%, 16% and 49% of the fluctuation in domestic inflation is explained by the world, regional and country specific factors respectively[4].

[4]See also Mumtaz et al. (2011) for the same statistical settings to extract world, regional

However, common factor models employed in the above studies assume that the dynamics between variables and underlying parameters are constant and as such that they are unable to identify time variations in the relationship between domestic and the international factors. This issue is often addressed using subsamples (see e.g Peacock and Baumann, 2008, Neely and Rapach, 2011). For instance, Neely and Rapach (2011) divides the full sample into subsamples where the first subsample covers the period until the early 1980s and the second starts around early 1980s onward when globalization is often assumed to accelerate. The comparisons of two subsamples indicate increased degree of international interdependencies. This is, albeit a crude, indication that inflation links have strengthened with increased globalization.

An exception that allows time variations in the factor based setting is Mumtaz and Surico (2012) who employ a dynamic factor model with time varying coefficients. Specifically, domestic inflation is decomposed into country specific and common factor components, which each follows an autoregressive process. Time variations come through the autoregressive coefficients and variances which evolve as random walks and geometric random walks respectively. However, the common factor extraction is based on a time invariant dynamic factor model to reduce the computational burden. Nevertheless, despite the more general methodology, their conclusion is largely consistent with other studies – the degree of co-movement is increased since the mid 1980s.

An alternative approach of examining changes in the dynamics of international inflations is proposed by Bataa et al. (2013a) who use structural break tests. They employ a Vector Autoregressive (VAR) specification for inflation across countries in which breaks in dynamics parameters and disturbance covariances are tested separately using their new iterative approach. Further, the breaks in the disturbance covariance matrix are disentangled in order to establish whether these breaks are associated with variances or correlations. The evidence for breaks in the correlation component indicates changes in the link of inflation across countries. Employing this approach on the Euro area VAR system including France, Germany and Italy, and a VAR for the G7 excluding Japan, they show that the international co-movement increases starting around the 1980s for the Euro area VAR system and the mid 1990s in the G7-VAR system.

This chapter closely relates to and builds upon Bataa et al. (2013a) when testing for breaks in the international links, but employs a simple bivariate model which includes a measure of international inflation as an explanatory

and country specific components from real and nominal variables, by jointly estimating co-movements in output growth and inflation.

variable. The system approach makes a strong assumption that all cross-country inflation linkages exhibit changes at the same date. Additionally, it considers only links across a small number of countries. Whereas, the analysis in this chapter is more general in examining each country in conjunction with an appropriate measure of foreign inflation and in allowing breaks to occur independently. Further, this chapter improves on their methodology, particularly in relation to the detection of outliers, which are important for inflation. For the next subsection, we assess structural break testing methodologies used in the literature and explain the importance of our methodological contribution.

1.2.2 Structural Break Analysis

Theoretical research highlights the sensitivity of structural break testing to the assumptions made. On the one hand, Pitarakis (2004) shows that tests for mean and persistence breaks are distorted substantially when volatility changes occur but are unaccounted for. On the other hand, inference on volatility breaks is misleading if the computed residuals are contaminated by neglected mean breaks as noted by Sensier and van Dijk (2004). However, empirical analyses for inflation are generally deficient in recognizing these theoretical findings.

1.2.2.1 Single break tests

The structural stability of inflation is well studied in relation to analysis of persistence change, which, in the context of an autoregressive model, is usually measured by the sum of the estimated autoregressive coefficients. A large number of studies note that the previously believed high inflation persistence is due to the failure of accounting for structural instability (Gadzinski and Orlandi, 2004, Clark, 2006, Levin and Piger, 2003, Altissimo et al., 2006, etc). On the contrary, other studies find persistence to be largely constant after allowing a structural change in the mean level of inflation (O'Reilly and Whelan, 2005, Pivetta and Reis, 2007).

A number of studies of inflation changes focus on its mean level, with changes in persistence of lesser interest and changes in the variance generally ignored. Perhaps, this is due to the fact that many studies examine the effects of monetary policy changes which are often associated with changes in the mean of inflation[5]. To date, the main methodology employed in these studies

[5]Specially, this is the case in those who are motivated to explore the link between monetary policy change and low levels of inflation observed at least for last two decades (Altissimo et al., 2006, Gadzinski and Orlandi, 2004).

is to conduct the Andrews (1993), test for a single break at an unknown date. This is applied to the intercept of a univariate autoregressive model assuming dynamics parameters are constant (see e.g Levin and Piger, 2003, Altissimo et al., 2006, on OECD and Euro area data). As a result, persistence is lower compared to models that impose a constant mean inflation level. However, in a finite sample, the asymptotic p values may over reject the null of stability when the true degree of persistence is high (Diebold and Chen, 1996, Hansen, 2000). Therefore, when assessing the significance of break dates, bootstrapped p values are used which follow the study by Diebold and Chen (1996), who showed these to be valid.

However, the superiority of the bootstrapping procedure revealed by Diebold and Chen (1996) is based on the case of simultaneous testing for mean and persistence breaks. But, the literature tends to test for mean break only or mean break first and persistence break second, conditionally or unconditionally on the obtained mean break. For instance, in the studies by Gadzinski and Orlandi (2004) and Clark (2006), models are re-estimated allowing for a previously obtained intercept change and then the stability of persistence parameters is tested using bootstrapped p values. After this sequential test, they conclude that the null of constant persistence is hard to reject based on Euro area and US data (Gadzinski and Orlandi, 2004, Clark, 2006). This may be due to low power for detecting small and even moderate size persistence breaks when testing for mean and persistence breaks separately even after employing bootstrapped p values, as pointed out by O'Reilly and Whelan (2005).

The majority of empirical studies assume a constant variance. However, in practice, we should expect changes in the volatility of inflation at least for some countries in relation to oil price shocks during the 1970s and 1980s, disinflation or inflation targeting policies in the early 1980s and 1990s respectively. If those events lead to omitted volatility change, this may contaminate the performance of the mean break and persistence break tests, since neither Andrews (1993) asymptotic distribution nor simple bootstrapped p values are robust to the presence of heteroskedasticity.

Few empirical studies employ a heteroskedasticity robust method when testing for a mean break. Levin and Piger (2003) use a "wild" bootstrap method when testing for a mean break using the OECD data. Furthermore, Hansen (2000) proposed two forms of "fixed regressor bootstrap", one being appropriate in the presence of heteroskedastic variance and another in the absence of heteroskedasticity, and compared them with the inference using the asymptotic distribution which assumes homoskedastic variance. Although the results from two forms of fixed regressor bootstrap yield a substantial size

improvement over the asymptotic distribution, these tests also tend to over reject in the presence of heteroskedasticity.

As a conclusion, many available results for inflation breaks are unreliable because they do not allow for structural breaks in the variance.

1.2.2.2 Multiple break tests

Inflation was high and volatile around 1970s until the mid 1980, and becomes low and less volatile after that. It declined further in the early 1990s and has remained low and stable over the last two decades. Perhaps, looking for multiple breaks in over forty years may be intuitive. Many studies find more than one structural break which they associate with the start of European Monetary System, disinflation policies in the US and UK during the early 1980s (Benati and Kapetanios, 2002, Corvoisier and Mojon, 2005) and inflation targeting during the 1990s (Benati and Kapetanios, 2002, Corvoisier and Mojon, 2005, Cecchetti and Debelle, 2006). Along with these breaks, significant declines of both mean and persistence are documented.

For example, Corvoisier and Mojon (2005) find 57 breaks in the mean of inflation across 22 OECD countries when testing for breaks in the unconditional mean in autoregressive models. That is a little less than three breaks for each country, on average. Similarly, Cecchetti and Debelle (2006) also document multiple mean breaks using data from 19 OECD countries. However, the numbers of mean breaks found in these studies may be the results of oversized tests due to the authors' failure to account for potential changes in variance and persistence.

As mentioned previously, testing for structural breaks in dynamic models requires consideration be given also to the nature of volatility. This is because omitted variance breaks could be interpreted as mean breaks leading to a conclusion of spurious mean breaks (Sensier and van Dijk, 2004, Pitarakis, 2004). On the other hand, testing variance breaks without accounting for existing mean breaks may cause an identification of too few variance breaks than there are. Because, omitted mean breaks make variance estimates larger and therefore the detection of variance breaks less likely. Whereas, in the empirical literature, the variance of residuals is constructed using the residuals obtained from the ordinary least squares estimates where mean breaks are not taken into account (Benati and Kapetanios, 2002, Clark, 2006). Therefore, variance breaks may be contaminated by omitted mean breaks which lead misleading inference.

Empirical support for the theoretical claim stated above seem to be reflected in the work by Benati and Kapetanios (2002) who note a reduced

power of separate coefficient break testing. They separately tested multiple breaks in the intercept, the AR coefficients and the innovation variances of univariate autoregressive model using the Andrews (1993), Andrews and Chen (1994) sequential methodology described in Bai (1997) assuming other parameters except the one being tested are constant. Based on real data, they find uniformly strong evidence of structural instability in the variance and uniformly non rejection of the null of stability in the intercept of almost all series considered,[6] and the rejection of stability in the AR coefficients in most cases. When they employ different test such as the Nyblom-Hansen test, the null of stability is rejected in the intercept of most autoregressive models while the number of variance and AR coefficient breaks remains large. This is indicative of the existence of breaks in the variances and persistence parameters, and one should be careful when assuming stability in all components other than those being tested.

Thus, testing for structural change in one component conditional on identified breaks in other components is vital. This issue is addressed in the recent paper by Bataa et al. (2013a,b). They propose an iterative methodology to examine breaks in each component sequentially, using the multiple structural break test by Qu and Perron (2007) together with an outlier detection procedure. Detail of this approach is provided in section 1.3.1 since it closely relates to our approach.

1.3 Methodology

1.3.1 Iterative methodology of structural break analysis

As a complement to the existing literature that often conducts break point tests under misspecification (omitting changes in either mean or variance of a time series), we employ an iterative approach which aims to avoid misspecification through the use of an iterative procedure. Our research adapts the iterative methodology by Bataa et al. (2013b,a) to analyze structural breaks in the mean, persistence (dynamics) and innovation variance (volatility) of univariate inflation series.

The iterative methodology proposed by Bataa et al. (2013b) tests for structural breaks in each of the components of inflation: seasonal, mean, dynamics and volatility one at a time conditional on previously found breaks in all other components. The testing procedure employed is that of Qu and Perron (2007),

[6] A similar approach is undertaken by Clark (2006) in a single break context. He finds strong evidence of intercept break and very small evidence of persistence and variance changes.

together with the outlier detection and removal procedure of Stock and Watson (2003). However, this procedure is quite complex and, as indicated by the Monte Carlo results in Bataa et al. (2013a) for the multivariate case, iteration is relatively unimportant in practice for the variance component. Further, their separation of mean and dynamics breaks can have relatively poor performance in practice, especially since the initial tests for mean breaks apply Heteroskedasticity and Autocorrelation robust (HAC) inference using the approach of Andrews (1991), which is known to be sometimes badly oversized (Bai and Perron, 2006). Finally, while their outlier detection procedure makes use of detected coefficient breaks, variance breaks are ignored for outlier detection. Therefore, we propose a simple, yet efficient version of the iterative approach of Bataa et al. (2013b) that also takes account of these concerns. It is more flexible in a number of respects, including re-specification of the model employed at each iteration, reflecting the effects of detecting and removing outliers.

Note that seasonality is not a particular focus of interest in this study. Since CPI data are typically available only in a seasonally unadjusted form, we use the widely applied X-12-ARIMA seasonal adjustment procedure[7] to deseasonalize the data prior to beginning our iterative procedure. The X-12-ARIMA procedure is particularly suitable in our context, as it allows for the presence of trend, deterministic seasonal patterns, holidays and trading day adjustment, additive outliers and level shifts (Osborn and Ghysels, 2001, p.106-127). Note, however, that while additive outliers are taken into account for the purposes of seasonal adjustment, they remain in the series after seasonal factors are removed (Census Bureau, 2011, p.123-127).

Here the discussion of methodology focuses on univariate inflation models although this chapter concerns changing dynamics in international links. This is because many studies are readily available in the context of univariate inflation and results from these studies can be compared to that of ours after applying the iterative testing procedure. In subsection 1.3.3, we will turn to the analysis of international inflation links.

A time-varying univariate AR model for monthly domestic inflation in a country, π_t^D, is given by

$$\pi_t^D = \mu_j + \sum_{i=1}^{n} \alpha_{ij} \pi_{t-i}^D + \upsilon_t, \tag{1.1}$$

[7]This is implemented using the EVIEWS 7 software (EVIEWS, 2009). We performed a small experiment by comparing official seasonally adjusted US data by Bureau Census with ones filtered by X-12-ARIMA. A graphical analysis indicated that the two series had very similar properties, thus we proceed with X-12-ARIMA for the inflation series of all countries.

where the subscript j indicates the coefficient regime and υ_t is a zero mean uncorrelated process whose variance $\sigma_k^2 = E[\upsilon_t^2]$ is allowed to change over variance regimes (indicated by the subscript k). Our interest, therefore, focuses on possible discrete breaks in the coefficients and the disturbance variance, while allowing for the presence of additive outliers in π_t^D, which could be due to (say) changes in indirect taxes. Denote m as the unknown number of coefficient breaks. Within each of $m+1$ coefficient regimes, $\delta_j = (\mu_j, \alpha_{1j}, \ldots, \alpha_{nj})'$ is time-invariant and all AR roots are assumed to lie strictly outside the unit circle. The jth regime extends over observations $t = T_{j-1}+1, \ldots, T_j$ using the convention that $T_0 = 0$ and $T_{m+1} = T$. All coefficients are allowed to change and the break dates $(T_1, \ldots, T_m)$ are treated as unknown. Similarly, σ_k^2 is constant within each volatility regime and is assumed to be conditionally homoskedastic. Our iterative approach to specifying the model in (1.1) is given by the following steps and a schematic illustration of the algorithm is provided in the Appendix A.2.

Step 1 - Outlier detection: The first iteration starts by identifying outliers in the deseasonalized full sample of data. Employing the outlier detection procedure by Stock and Watson (2003), outliers are defined as four times of the interquartile range from the median[8]. Detected outliers are replaced by the median of the six neighboring non outlier values.

Step 1* - Outlier detection for subsequent iterations: In subsequent iterations, outliers are examined separately within each coefficient regime and in data adjusted for volatility breaks (by standardizing the series using standard deviations of residuals in corresponding volatility regimes). Detected outliers are replaced by the median of the six neighboring non outlier standardized values. The data are then destandardized, to yield a series adjusted only for outliers.

Step 2 - Model selection: A univariate inflation model is selected using the Schwartz Information Criterion (thereafter SIC). Specifically, using the AR model and allowing a maximum lag of $n = 17$, all possible combinations of lags are considered, implying a total of 2^{17} models. Since "gaps" are permitted in coefficients, i is not necessarily consecutive in (1.1). To ensure comparability, all models for a given country are estimated over

[8]There is a trade-off for choosing between too small or too big number to multiply the interquartile range. If the number is chosen too large, then it is unable to pick up obvious outliers. If it is chosen too small, too many outliers are detected in a single series. In our judgment a value of four times the interquartile range seems appropriate for most inflation series as it allows obvious outliers to be identified and results in a reasonably small number of outliers.

a common set of data, and the choice among them is made based on minimum SIC. Persistence is measured by the sum of autoregressive coefficients, $\hat{\rho} = \sum_{i=1}^{n} \hat{\alpha}_i$, as it is the best scalar measure of the persistence, as indicated by Andrews and Chen (1994).

Step 3 - Preliminary coefficient break test: After having specified lags in (1.1), the Bai and Perron (1998) multiple structural breaks procedure is applied to the coefficient vector of the autoregressive model (including intercept and slope parameters of the regression). The possibility of heteroskedasticity in the variance is allowed by employing Heteroskedasticity Consistent (HC) inference[9]. Although HC inference can lead to oversized coefficient break tests when there is no heteroskedasticity, shown by the simulation analysis by Bai and Perron (2006), the estimates in each regime are consistent in a large sample. Further, coefficient breaks identified here are reconsidered in step 5 of the iteration.

Step 4 - Variance break test: Conditional on the coefficient break dates from step 3, variance breaks are examined through tests applied to the mean of the squared residuals (see section 1.3.2 for details). This is to mitigate the concern of misleading inference of variance breaks, caused by obtained residuals that may be contaminated by coefficient breaks (Sensier and van Dijk, 2004, Pitarakis, 2004).

Step 5 - Coefficient break test: To avoid the serious problems for coefficient break tests of omitted variance breaks (Pitarakis, 2004), we re-test breaks in the coefficients conditional on the variance breaks from step 4. That is, we apply the feasible GLS transformation[10] and, assuming homoskedasticity in the error term, the Bai and Perron (1998) procedure is performed again on the new transformed data in order to obtain volatility adjusted coefficients break dates for the model specified in step 2. If no volatility breaks are found from step 4, coefficient tests are applied to the original data with a homoskedastic variance assumption, and the iteration ends.

The iterative testing procedure outlined above differs from the methodology by Bataa et al. (2013b) in several respects. Firstly, Bataa et al. (2013b) test

[9]The procedure of Bai and Perron (1998) allows for the presence of disturbance heteroskedasticity and/or autocorrelation using the approach of Andrews (1991). Our implementation requires only HC inference, which follows Bai and Perron (1998) in using the Andrews (1991) method.

[10]This methodology is based on the findings by Pitarakis (2004) who revealed substantial improvement of this transformation in small samples by comparing bootstrap based test on both transformed and untransformed data.

for breaks in seasonal components as part of the iterative procedure whereas we apply seasonal adjustment procedure to the data once prior to beginning of our iterative procedure. Secondly, outlier detection procedure in step 1* takes account of the latest identified coefficient and variance breaks while variance breaks are ignored when detecting outliers in Bataa et al. (2013b). Thirdly, in step 2 we re-specify the model employed at each iteration, reflecting the effects of detecting and removing outliers. This is not a concern in Bataa et al. (2013b).

Fourthly and most importantly, the preliminary coefficient break test in step 3 (where mean and dynamics are jointly tested) employs HC inference to account for possible heteroskedasticity in the variance. For their initialization, HAC inference is employed when testing for mean breaks to account for un-modeled dynamics and variance, and later they employ HC inference when testing for breaks in dynamics on the demeaned data. However, as mentioned previously, this procedure can be substantially oversized, therefore and consequently we jointly test for mean and dynamic breaks. Finally, the iterative procedure by Bataa et al. (2013b) incorporates 'inner loop' that iterates between tests for breaks in the dynamics and the residual variance. However, as shown by their Monte Carlo simulation, variance breaks are detected well without iteration. Our variance break testing procedure in step 4 simplifies the iterations in respect to identification of variance breaks. In each iteration, possible breaks in the residual variance are tested once conditional on coefficient breaks detected from step 3.

A single iteration is composed of steps 1 to 5. The iterations proceed to convergence, with a maximum number of iterations set to 10. Convergence may be achieved in two different ways: firstly, the same set of break dates may be obtained from consecutive iterations; alternately, the iteration can cycle between two or three sets of break dates. In the later case, we choose the set which achieves the smallest SIC criterion among these local optima. When calculating SIC for this purpose, we use a fixed number of observations, T. The version of SIC is that proposed by Yao (1988) for structural break inference, which is applied to the GLS transformed data and calculated for m breaks as

$$SIC(m) = \ln \hat{\sigma}^2(m) + p^* \ln(T)/T, \tag{1.2}$$

where $\hat{\sigma}^2(m) = T^{-1} S_T(\hat{T}_1, \ldots, \hat{T}_m)$, in which $S_T(\hat{T}_1, \ldots, \hat{T}_m)$ is a sum of squared residuals over m breaks, and $p^* = (m+1)q + m$ in which q equals the number of coefficients (including the intercept) in (1.1). Thus, the penalty effectively treats each break date as a parameter to be estimated.

A single iteration accounts for the main issues that we address in this chapter - namely, the integrity of estimated mean, persistence and variance breaks. However, on the one hand, those break dates from steps 4 and 5 can have a considerable impact on the outlier detection procedure of step 1. For example, an outlier detected using the full sample may not be an outlier for a certain high volatile regime but appear as an outlier compared to a smooth part of the sample. Similarly, an outlier appearing in the relatively stable regime may be too small to be detected using the full sample compared to a volatile part of the sample. On other hand, a different set of outliers can be found from one iteration to another depending on the variance and coefficient breaks identified in the previous iteration, and newly identified outliers also can have an impact on the identification of coefficients and variance breaks in the following steps. Hence, the need for iteration.

1.3.2 Estimating the number of breaks

The heart of the iteration described in the above subsection is the multiple structural break testing procedure by Bai and Perron (1998)[11]. Say the model of (1.1) has a maximum of m coefficient breaks and hence $m + 1$ regimes, $j = 1, \ldots, m + 1$. The estimates of the parameters and the optimal break dates are computed using the dynamic programming algorithm of Bai and Perron (1998, 2003a), which searches for the minimum total residual sum of squares over all $m+1$ regimes. This yields m sets of possible break dates: that is, $1, 2, \ldots, m$ possible estimated break dates.

After m sets of possible estimated break dates are obtained, we employ two different tests: WDmax and sequential $Sup\ F(l + 1|l)$ to choose among those sets. First, we use WDmax[12] as an indication of the presence of at least one break. WDmax tests the null hypothesis of no breaks against the composite alternative of $1, \ldots, m$ breaks and failure to reject the null hypothesis then zero breaks are estimated to occur. As recommended by Bai and Perron (1998, 2003a), when the null hypothesis is rejected, their sequential $SupF(l+1|l)$ test is employed to estimate the appropriate number of breaks. That is, the null hypotheses of $l = 1, 2, 3, \ldots$ breaks (subject to a maximum of m breaks) are examined sequentially against the alternative of $l+1$ breaks, with the first non-rejection yielding l breaks. In particular, this test is applied first for 2 versus

[11]We adapt the MATLAB code for testing multiple structural breaks which is originally developed by Pierre Perron in the GAUSS program and translated later to MATLAB program by Yohei Yamamoto (2012).

[12]The WDmax statistic is used in preference to UDmax because it embodies a set of weights that ensure the marginal p-values are equal for the null of no breaks against each specific number of breaks $1, 2, \ldots, m$ (Bai and Perron, 1998).

1 break (not 1 versus 0) due to the difficulty of rejecting the null hypothesis of zero versus a single break in the sequential test, especially in a case that the value of the coefficients returns to its original value after the second break when two breaks are present (Bai and Perron, 2003a, 2006). Sequential $Sup\ F(l+1|l)$ tests are conducted due to their good performance under both presence and absence of serial correlation and heterogeneity compared to the use of information criterion (Bai and Perron, 2006).

All tests are computed at a nominal 5 percent level of significance, with the maximum number of breaks considered being $m = 5$. Testing employs the asymptotic distributions obtained by Hall and Sakkas (2013), which are shown by these authors to more accurate than the critical values provided by Bai and Perron (2003b) and have the additional advantage of allowing computation of asymptotic p -values. The so-called trimming parameter, which defines the minimum distance between two consecutive breaks as a function of the total sample size T is set at 0.15.

More specifically, the testing procedure we describe in this section relates to steps 3, 4 and 5 of the iteration above. We first test $H_0 : \mu_j = \mu_{j+1}$ and $\alpha_{i,j} = \alpha_{i,j+1}$ for $j = 1, \ldots, m$ against the alternative of $H_A : \mu_j \neq \mu_{j+1}$ or $\alpha_{i,j} \neq \alpha_{i,j+1}$ for at least some $m \leq M$ (M is an upper bound), using

$$WD\ \max F_T(M,q) = \max_{1\leq m\leq M} a_m[\sup_{(\lambda_1,\ldots,\lambda_m)\in\Lambda_\epsilon} F_T(\lambda_1,\ldots,\lambda_m;q)] \tag{1.3}$$

where λ_j for $j = 1, \ldots, m$ are possible break dates as fractions of the sample size, and Λ_ϵ denotes the set of all possible sample partitions given ε which is the smallest fraction of the sample that must be included in each segment, satisfying $0 < \varepsilon < 1$. For $m > 1$, $a_m = c(q,\alpha,1)/c(q,\alpha,m)$ in which $c(q,\alpha,m)$ is the asymptotic critical value of the test $\sup_{(\lambda_1,\ldots,\lambda_m)\in\Lambda_\epsilon} F_T(\lambda_1,\ldots,\lambda_m;q)$ at a significance level α, where $\sup F_T$ is given as

$$\sup F_T(\lambda_1,\ldots,\lambda_m;q) = \sup[\frac{1}{T}(\frac{T-(m+1)q}{mq})\hat{\delta}'R'(R\hat{V}(\hat{\delta})R')^{-1}R\hat{\delta}] \tag{1.4}$$

where q is the number of regressors that are allowed to change and $\hat{\delta} = (\hat{\mu}_j, \hat{\alpha}_{1j}, \ldots, \hat{\alpha}_{nj})$. We allow the covariance matrix of $\hat{\delta}$ to evolve as $\hat{V}(\hat{\delta}_j) = \hat{\sigma}_j{}^2[(\Delta\widehat{T}_j)^{-1}\sum_{t=\widehat{T}_{j-1}+1}^{\widehat{T}_j} Z_t Z_t']^{-1}$ where $\widehat{\sigma}_j^2 = (\Delta\widehat{T}_j)^{-1}\sum_{t=\widehat{T}_{j-1}+1}^{\widehat{T}_j} \hat{\upsilon}_t^2$ for $j = 1, ..., m+1$, under the HC inference and $Z_t = (1, \pi_{t-i}^{D'})$ is the vector of regressors. The HC case here, however, only allows for variance breaks that coincide with coefficient breaks. R is a matrix of restrictions such that $(R\delta)' = \left(\delta_1' - \delta_2', ..., \delta_m' - \delta_{m+1}'\right)$.

Once the WDmax test rejects the null of no breaks, we employ $Sup\ F(l+1|l)$ to define the number of optimal breaks using

$$F_T(l{+}1|l) = \{SSR_T(\widehat{T}_1, \ldots, \widehat{T}_l) - \min_{1\le j\le l+1} \inf_{\tau\in\Lambda_{j,\varepsilon}} SSR_T(\widehat{T}_1, \ldots, \widehat{T}_{j-1}, \tau, \widehat{T}_j, \ldots, \widehat{T}_l)\}/\widehat{\sigma}_j^{\,2} \tag{1.5}$$

where $\Lambda_{j,\varepsilon} = \{\tau; \widehat{T}_{j-1} + (\widehat{T}_j - \widehat{T}_{j-1})\varepsilon \le \tau \le \widehat{T}_j - (\widehat{T}_j - \widehat{T}_{j-1})\varepsilon\}$.

Here one additional break is inserted, conditional on the break dates already uncovered and assessed whether additional break reduces the overall sum of squared residuals. For example, the null hypothesis of l breaks is rejected against the alternative of $l+1$ if its overall sum of squared residuals is sufficiently larger than the sum of squared residuals from the model with $l+1$, and it continues sequentially until the testing procedure fails to reject the null hypothesis.

At step 3 of the iteration, we obtain the estimated coefficient break dates under equations (1.3) to (1.5) and denote these as $\widehat{T}_1^C, \ldots, \widehat{T}_m^C$. After obtaining the estimates of $\hat{\delta} = (\hat{\mu}_j, \hat{\alpha}_{1j,\ldots,}\hat{\alpha}_{nj})$ and the corresponding coefficient break dates $\widehat{T}_1^C, \ldots, \widehat{T}_m^C$, we estimate the variance of residuals by first concatenating the squared residuals in each regime

$$\hat{\upsilon}_t^2 = (\pi_t^D - \widehat{\mu}_j - \sum_{i=1}^{n} \widehat{\alpha}_{i,j}\pi_{t-i}^D)^2 \tag{1.6}$$

where $j = 1, \ldots, m+1$ and $t = \hat{T}_{j-1}^C + 1, \ldots, \hat{T}_j^C, E(\upsilon_t) = 0$ are assumed.

Then at step 4 of the iteration, we run the tests described in equations (1.3) to (1.5) again on the variance of residuals through the regression

$$\hat{\upsilon}_t^2 = \gamma_j + u_t \tag{1.7}$$

where γ_j is a constant whose value is allowed to change over time.

At step 5, if any variance breaks[13], denoted as $\hat{T}_1^V, ..., \hat{T}_m^V$, are found in the equation (1.7), we calculate the standard errors in each regime as $\hat{\sigma}_j = \sqrt{(\Delta\widehat{T}_j^{\,V})^{-1} \sum_{t=\widehat{T}_{j-1}^V+1}^{\widehat{T}_j^V} \hat{\upsilon}_t^2}$. Then, the standard error in each regime is used to standardize the data that leads GLS transformation, $\bar{\pi}_t^D = \frac{\pi_t^D}{\hat{\sigma}_j}$ $\bar{\pi}_{t-i}^D = \frac{\pi_{t-i}^D}{\hat{\sigma}_j}$ $\bar{\mu}_j = \frac{\mu}{\hat{\sigma}_j}$ where $t = \hat{T}_{j-1}^V + 1, \ldots, \hat{T}_j^V$. Then coefficient break testing is applied to the model using GLS transformed data, but under the homoskedastic assumption so that the covariance matrix of $\hat{\delta}$ is obtained as $\hat{V}(\hat{\delta}_j) = \hat{\sigma}^2 \left[(\Delta\widehat{T}_j^{\,C})^{-1} \sum_{t=\widehat{T}_{j-1}^C+1}^{\widehat{T}_j^C} Z_t Z_t'\right]^{-1}$ with $\widehat{\sigma}^2 = (T)^{-1} \sum_{t=1}^{T} \hat{\upsilon}_t^2$.

[13]Note that although m is used to denote the number of both coefficient and variance breaks, in practice we allow different numbers of breaks to apply for these components.

1.3.3 Testing in the bivariate model

So far, we have focused on univariate inflation models to test for structural breaks using our iterative methodology. This subsection introduces the bivariate model of principal interest, which examines changes in the degree of interdependence of domestic and foreign inflation. For this purpose, a parsimonious representation of domestic inflation for country s in month t ($\pi^D_{t,s}$) is given by

$$\pi^D_{t,s} = \mu_j + \sum_{i=1}^{n} \alpha_{ij}\pi^D_{t-i,s} + \beta_{0j}\pi^F_{t,s} + \sum_{i=1}^{n} \beta_{ij}\pi^F_{t-i,s} + \varepsilon_t \tag{1.8}$$

where $\pi^F_{t,s}$ is foreign inflation in relation to country s at time t, and β_{0j} captures the contemporaneous co-movement between domestic and foreign inflation in coefficient regime j. Inflation in country s also depends on its own lags and the lags of foreign inflation, where the effects are captured through $(\alpha_{1j}, \ldots, \alpha_{nj})$ and $(\beta_{1j}, \ldots, \beta_{nj})$ coefficients respectively. Foreign inflation is treated as weakly exogenous for domestic inflation. Inflation persistence for country s in this model is measured by $\hat{\rho}^d_j = \sum_{i=1}^{n} \hat{\alpha}_{ij}$.

The motivation for the form of (1.8) is the Global-VAR (GVAR) analysis which examines international links using country-specific foreign variables. For instance, Pesaran et al. (2004) model each domestic macroeconomic variable considered in terms of its own lags, contemporaneous foreign variables and their lags. The US is a special case in their studies, with foreign inflation and output excluded from the US model as they assume that it violates weak exogeneity. Our bivariate model of inflation in (1.8) is similar, but in a single equation context in order to focus on international linkages of inflation. This allows us to test for time variations without losing too much power. Additionally, we include contemporaneous foreign inflation in the US model. This follows the arguments of Dees et al. (2007), that, in a foreign context and as the number of countries increases, this variable can be treated as weakly exogenous also for the US.

We anticipate breaks in the foreign coefficients $(\beta_{0j}, \beta_{1j}, \ldots, \beta_{nj})$, if there are changes in the way in which domestic inflation relates to foreign inflation. Additionally, the locations of breaks in the $(\mu_j, \alpha_{1j}, \ldots, \alpha_{nj})$ coefficients may differ from those found in the univariate models of equation (1.1), due to the inclusion of foreign variables. Although we do not employ tests to disentangle explicitly what elements of $\delta_j = (\mu_j, \alpha_{1j}, \ldots, \alpha_{nj}, \beta_{0j}, \beta_{1j}, \ldots, \beta_{nj})$ change at break dates, coefficient estimates in each regime are informative with regard to this.

Inference as to the presence and dates of the breaks in (1.8), including

breaks in the disturbance variance, is achieved by employing the iterative procedure outlined in subsection 1.3.1 and the multiple break testing methodology in subsection 1.3.2. Although the general procedure is the same as in the univariate analysis, some additional remarks should be made. In step 1 of the iteration (step 1* for subsequent iterations), the outlier detection and removal procedure runs only on domestic inflation because aberrant observations in the explanatory variables should not affect the size of the test[14]. Furthermore, we note that the presence of a break in the explanatory variable does not affect the size of the test[15].

In step 2, we choose bivariate models in a slightly different manner from the univariate models. Employing the same model selection method is computationally excessive, since the best model would be selected out of 2^{25} possible models, provided that the maximum lags allowed for domestic and foreign variables are 12 each plus a contemporaneous foreign variable. Therefore, we employ a general to specific methodology to remove irrelevant lags from the general model, but still decide the best model based on SIC. Precisely, we start by evaluating the model with 25 lags (12 lags for each of domestic and foreign inflation plus a contemporaneous foreign inflation), then the least significant lag using t-tests is eliminated and corresponding information criterion (SIC) is calculated. Continuing by sequentially dropping the least significant lag one at a time, until only the intercept remains, we choose the model which achieves the smallest SIC criterion across all 25 models.

However, the selected model is the optimum within a single path. There could be multiple paths that yield different optima depending on the starting point of elimination. Therefore, we check the sensitivity of the model selection to the starting point using the idea of the multipath search algorithm, proposed by Krolzig and Hendry (2001). To be specific, we proceed through 5 paths by initially eliminating the zth (where $z = 1, \ldots, 5$) least significant variable. Once the first variable is dropped, the least significant variable is dropped at all subsequent stages. At the end of the search, we have 5 sets of SIC values from which the final model is selected based on the smallest SIC criterion[16]

[14]We undertook a simulation study to examine the performance of the Chow test with explanatory variables having moderate and large size outliers. Based on the 10000 replications, on average, the size of the test is unaffected. The results are reported in the Appendix

[15]Allowing a single break in the process generating the explanatory variable occurring in the middle or towards the end of the sample, the test is well-sized at a 5% significance level, based on 5000 replications. The results are provided in the Appendix.

[16]We also compare our information criterion based models with a conventional testing down method, using a significance level of 1%. In the latter approach, all remaining coefficients are significant at 1% but this does not necessarily achieve the smallest information criterion. It yielded very similar lags to those selected by SIC, except for the inclusion of an additional lag in a few cases.

achieved among all values.

Moreover, in the sensitivity analysis (which we will discuss in detail in section 1.6), an additional explanatory variables is included in the bivariate model, with a contemporaneous and 12 lagged values added. The additional variables are oil price inflation and the change in trade weighted real effective exchange rates. The first is employed because a sudden increase in oil price can cause an exogeneous inflationary shock to domestic inflation, and omitting this variable may result upward bias in the estimated coefficients. The latter is included as it may be important in explaining domestic inflation, especially for open economies, through its influence on import and export prices. The approach, including the way SIC is used for model selection, is unchanged from that employed for the bivariate models.

1.3.4 Measuring foreign inflation

We construct foreign inflation for country s (where $s = 1, \ldots, N$) based on a weighted average of inflation series over the other $N - 1$ countries in the data set. Weights are computed based on bilateral trade statistics as,

$$w_{s,t}^{(i)} = \frac{(M_{s,t}^{(i)} + X_{s,t}^{(i)})}{\sum_{i=1, i \neq s}^{N} (M_{s,t}^{(i)} + X_{s,t}^{(i)})} \text{ and } \pi_{s,t}^{F} = \sum_{i=1, i \neq s}^{N} w_{s,t}^{(i)} \pi_{i,t} \tag{1.9}$$

where $\sum_{i=1, i \neq s}^{N} w_{s,t}^{(i)} = 1$ for $i = 1, \ldots, 19$ and $i \neq s$. The trade weight for country s with respect to country $i, w_{s,t}^{(i)}$, is given by the share of total trade between country s and i, in the total trade of country s with all its trading partners. Precisely, the total trade of country s with country i is measured by the sum of total imports from i ($M_s^{(i)}$) and exports to i ($X_s^{(i)}$). The weights are time varying and changes from month to month are relatively small, although this is not generally the case over the entire sample period. After computing trade weights, country specific foreign inflation is constructed as in (1.9) for each of the 19 countries in our sample.

1.4 Data

The data set we use in our analysis comprises of monthly aggregate series of Consumer Price Index (CPI) inflation for 19 OECD countries over the period between January 1970 and September 2010[17]. These include ten countries

[17] Although a few countries can be added if a shorter period is allowed, we prefer to focus on results using the larger sample.

that are members of the Euro Area (Austria, Belgium, Finland, France, Germany, Greece, Italy, Netherlands, Portugal, Spain), five other European countries (Denmark, Norway, Sweden, Switzerland, UK) and four other countries (Canada, Japan, Korea, US). All inflation series are calculated by differencing logged monthly indexes and multiplying by 100 where monthly CPI values are obtained from the OECD Main Economic Indicator database. Since we are using monthly series, seasonal oscillation is high and taken care of using the X12-ARIMA filter in EVIEWS 7 program with default options[18].

We also use monthly values of trade, which is defined by the sum of total exports and imports, by partner countries to construct trade weights using equation (1.9). According to the OECD Main Economic Indicator statistical website, all series are expressed in US dollars using (where appropriate) the exchange rates which adjust the rates before and after the start of the European Monetary Union (EMU). This adjustment facilitates a comparison within and across countries. The range of trade data is the same as CPI inflation although there are some missing data for Belgium, Korea and Portugal. Korea starts registering bilateral trade data from January 1988 and Belgium from January 1993. Portugal has missing trade data with respect to Italy between January 1971 and December 1973. Due to those missing observations, the trade weights corresponding to those periods are filled by the first available weight after the missing observations. This does not unduly distort the data since monthly weights are generally smooth over the 40 years of our sample.

Table 1.1 shows bilateral trade weights averaged over 40 years. In general, Germany is the biggest trade partner for most European countries, while the US is the main trade partner for non-European countries such as Japan, Korea and Canada. However, the UK does not have a dominant trade partner, although shares with respect to Germany, US and France are relatively large compared to others. Those weights are informative to construct a country specific foreign inflation, by taking account of contributions of trading partners' inflation. We should note, however, that weights based on bilateral trade statistics may be limited as they do not reflect trade effects of a third-country such as the big emerging economies of China and India. But, the limitation of data for those countries precludes their use.

For the sensitivity analysis in section 1.6, the world average crude oil price index, over the period between January 1970 and September 2010, is used to calculate oil price inflation which is added as an additional variable in

[18] Although there are official seasonally adjusted series available for the US and Germany, due to the consistency with other inflation series, we run the seasonal adjustment procedure on seasonally unadjusted data for all countries.

equation (1.8). This is available from the OECD Main Economic Indicator database. Another variable added in equation (1.8), although not at the same time with oil price inflation, is monthly averaged trade weighted real effective exchange rate indexes for individual country. This is obtained from the Bank of International Settlement database. Changes in these variables are computed by differencing logged monthly indexes and multiplying by 100, consistent with the construction of CPI inflation.

1.5 Results

This section presents the results. Section 1.5.1 provides a summary of results for the univariate inflation models. Section 1.5.2 presents the results for bivariate inflation models and discusses inferences with regard to the spillovers from foreign inflation to domestic inflation. All tests are conducted at the 5 percent significance level allowing a maximum of 5 breaks with value of trimming $\varepsilon = 0.15$, such that a minimum fraction of the sample in each regime equals to approximately 73 months. Asymptotic p-values are approximated using the method of Hall and Sakkas (2013).

1.5.1 Univariate inflation models

Table 1.2 represents the selected autoregressive lags of the univariate and bivariate inflation models; the latter are discussed in section 1.5.2. In the univariate models, we always find short lags to be present (say, 1, 2 and/or 3) when the maximum lag allowed is 17. This is not surprising as the recent past is more relevant. Also, longer lags (say 11, 12 and/or 13) are often found and this could indicate that some seasonal effects may still be present.

Both the WDmax and Sequential tests are conducted to estimate the number of breaks. Their test statistics and corresponding p-values when testing for coefficients and variance breaks are provided in Table A.3 and Table A.4 in the appendix. Generally, the WDmax and Sequential tests agree and point either to the existence or non existence of breaks. The latter is used to choose the number of breaks present. We note however, the concern of Bai and Perron (2003a) of low power of the $SupF(1|0)$ test in the presence of multiple breaks, which may be relevant to the case of Canada when testing for variance breaks; see Table A.4, in which the WDmax test rejects the null of no breaks, indicating the existence of some number of breaks. However, the sequential procedure reaches the conclusion of zero breaks, as $SupF(1|0)$ fails to reject. However, Canada is the only case where this occurs.

Table 1.3 reports the break dates uncovered in the univariate coefficients and residual variances. This table also indicates the number of iterations required for convergence of the testing procedure of subsection 1.3.1. All countries except Finland converge to a unique set of break dates, whereas for Finland the iterative procedure cycles between two local optima, in which the one with the smaller SIC is selected. We note that the iteration is necessary as convergence usually requires more than one iteration. However, our application requires no more than four iterations (except for Finland), highlighting the efficiency of our iterative methodology.

We also provide a figure for every country (figure 1.1.1-1.1.19, in alphabetical order), each comprising four graphs. The first two graphs in each figure correspond to the univariate specification and compare the difference between before and after iteration. Specifically, the first graph presents the break dates as well as some statistics relating to the corresponding regimes from applying the testing procedure of subsection 1.3.1 once, while the second graph reports the results after iterating the testing procedure multiple times until the convergence. For the majority of cases, the results in the first graphs can be seen as intermediate results to the second graphs as convergence usually requires more than one iteration, and thus the results in the second graphs are discussed in this subsection. However, in many cases identical results appear from employing the testing procedure once and iterating multiple times, indicating the effectiveness of the proposed testing procedure. Similarly, the third graph in each figure relates to the bivariate specification after iteration and these are discussed in the next subsection. The last graph in each figure plots country specific foreign inflation for each corresponding country. This series is also plotted in the third graph in order to compare dynamics between domestic and country specific foreign inflation.

To illustrate, refer to figure 1.1.18b for the univariate specification of UK inflation, for example. The vertical lines indicate the locations of the coefficient break dates with the estimated dates (June 1982 and December 1991) in the boxes next to these lines. Text arrows point to the locations of variance breaks and the corresponding changes in the variance of the consecutive regimes. The variance break occurring around April 1982 leads to a reduction of the variance from 0.13 to 0.03. Furthermore, the estimates of persistence and the unconditional mean in each regime, denoted by P and UcM respectively, are shown in the boxes. These estimates are indicative with regard to their changes over time. Outliers detected at the convergence of the iterations are indicated by black dots if any outliers are detected, with three outliers found in UK inflation over 40 years.

Overall, we find a total of 26 coefficient breaks across all 19 countries, with Austria and Switzerland having no breaks. This compares with the total of 23 mean and dynamics breaks obtained by Bataa et al. (2013b) for only 8 countries, despite the similar iterative approaches. For example, they uncover 4 mean breaks for France, whereas we find 2 significant breaks over a longer sample period[19]. This may point to their testing procedure being oversized if mean and dynamic breaks are considered separately.

Nevertheless, the timing of breaks presented in Table 1.3 is broadly consistent with the existing literature. We find clusters of coefficients breaks around the first half of the 1980s (although breaks for France, Spain and Norway are estimated to occur shortly after this) and early 1990s. The first cluster of breaks is widely considered to be a consequence of disinflation policies in a number of countries including the US and UK (Altissimo et al., 2006, Benati and Kapetanios, 2002, etc), and the currency peg in France, Italy and Netherlands which was designed to mimic the low inflationary experience in Germany (Altissimo et al., 2006). Consistent with this view, the unconditional mean declines to less than half of its pre-break value for most countries. The cluster in the early 1990s includes many European countries and may be related to the implementation of the Maastricht Treaty in 1992, in which inflation rates in the countries joining the Euro Area were required to converge. Additionally, break dates for the UK and Canada seem to relate closely with their introduction of inflation targeting policies in October 1992 and February 1991, respectively. In relation to these later breaks, further declines in the unconditional mean are observed with a smaller magnitude than the declines in the 1980s. The largest decline is in the mean of Japan after December 1991, pushing it to a negative value. These changes in unconditional mean can be seen in their respective country's figure.

A figure for each country (figure 1.1.1-1.1.19, in alphabetical order) also reports the estimates of persistence in each coefficient break regime. In common with the existing literature, the results show that estimated inflation persistence is generally smaller after the coefficient breaks, especially in the latter part of the sample. In most cases (12 out of 19 countries) estimated persistence is high, between 0.60-0.90 before the first break, but it falls substantially in later regimes and almost disappears in the last regime[20]. This contrasts with the finding by Cecchetti and Debelle (2006), O'Reilly and Whelan (2005) and others, who detect weak evidence for persistence change over time. However,

[19]Bataa et al. (2013b) use data between March 1973 and December 2007.

[20]This is consistent with findings by Bataa et al. (2013b) who noted zero inflation persistence for Canada.

in line with these studies, we find that relatively stable persistence applies in Norway, Germany and Netherlands. In the cases of Portugal and Spain, persistence declines after the first break, but increases back to previous high persistence levels after the early 1990s (see figures 1.1.14b and 1.1.15b).

Visual inspection indicates that every country experiences a highly volatile inflation period that lasts until either the late 1970s or mid-1980s depending on the country, with volatility decreasing afterwards. For Canada, Norway and the US we find volatility increases again around the early 2000s (see figures 1.1.3b, 1.1.13b and 1.1.19b). Consistent with this observed pattern, our results imply an equal number of variance and coefficient breaks, stressing the importance of variance break testing, which is largely absent from the existing literature. More importantly, although the clustering pattern of mean breaks is widely documented in the literature, we find also a clustering of variance breaks. For example, we find declines in variances around 1977 for eight countries, which may reflect the stabilization of inflation after the large oil price shocks of 1973-1974. An even larger number of breaks (14 in total) occur in the first half of the 1980s, reflecting "the great moderation". A few breaks also occur around 1992, which may be an effect of stabilization due to inflation targeting policies; for example, in Greece and Portugal.

Finally, we emphasize the importance of the outlier detection. Searching for outlier values in the coefficient break regimes using volatility standardized data yields more plausible outliers compared to those detected using the full sample. For example, an outlier in April 1991 in the UK does not appear as an outlier in the full sample when compared to the high inflation experienced during periods of the oil price shocks. However, our procedure distinguishes regimes with high and low levels of inflation in which this visually evident outlier is identified.

More importantly, the outlier detection procedure appears to have a considerable impact on inferences concerning coefficient and variance breaks. Each iteration hinges on the outlier detection procedure such that the only thing that changes from one iteration to another is a different set of outliers depending on the variance and coefficient breaks identified in the previous iteration. The difference between a single iteration (graph a in each figure) and multiple iterations (graph b in each figure) with an outlier detection procedure is sometimes striking. For example, see figures 1.1.5, 1.1.10 and 1.1.16 where outliers contaminate both variance and coefficient breaks, see figures 1.1.12 and 1.1.18 in which outliers complicate the detection of mean breaks and see figures 1.1.15 and 1.1.19 where variance breaks change after outlier iteration. Since the results taking account of outliers iteratively always visually appear

more reliable than those obtained using the full sample information with no account taken of breaks, we conclude that our conditional break point testing method with outlier iteration adds value to the existing literature.

1.5.2 Models with foreign inflation

As previously mentioned, Table 1.2 reports the models for the relationship between domestic and foreign inflation selected by our SIC-based approach. Furthermore, we note that employing multipath searches with different starting points does not change the models given by a single search. Generally, the bivariate models are more parsimonious than the univariate ones, with the number of domestic AR lags declining when the foreign variable is added; indeed, Austria and Germany now have no AR lags. Furthermore, contemporaneous foreign inflation plays a key role, with lags of this variable absent for most countries. Portugal is the only case where contemporaneous π_t^F is not selected. However, it is included in the estimated models for this case for comparability with other countries.

Corresponding structural break test statistics and approximate p-values are presented in Table A.5 and Table A.6 in the Appendix. The null hypothesis of no break is rejected for all series with the resulting coefficient and variance break dates reported in Table 1.4. We also turn to figures 1.1.1-1.1.19, where the third graph of each presents the results of the bivariate models. In each case, country-specific foreign inflation is represented by the red line. Based on these numerical and graphical illustrations, the results of our analysis can be summarized as follows.

Firstly and most importantly, we find a positive and increasing contemporaneous relationship between domestic and country-specific foreign inflation. It is particularly notable after 1990 for most countries (but could be after 1980 or 2000 for a few countries) and the corresponding estimated coefficient (β_{0j} in equation (1.8)), on average across countries, more than doubles compared to the pre-break regime. The third graph of each figure presents this coefficient. It is also visually evident in the graphs that the differences between domestic and foreign inflation gets smaller in the later period of the sample.

There are exceptions to this, however, in a small number of cases. For example, UK, Korea and Netherlands (figures 1.1.18c, 1.1.11c and 1.1.12c, respectively), show almost constant contemporaneous interactions over time, while their marked changes in the domestic indicators are evident. This suggests that the observed breaks are due to internal factors such as monetary policy changes. Another exception is Japan (figure 1.1.10c), where the con-

temporaneous effect is high during the oil price shocks in the first half of the 1970s, and declines afterwards. Regardless of these exceptions, a notable increase in the contemporaneous relationship may be informative with regard to co-movements of inflation. This is in line with Bataa et al. (2013a) who note increased contemporaneous international inflation linkages for the major G-7 economies they examine.

Secondly, inclusion of foreign inflation in the bivariate models can substantially change the identified break dates, pointing to the relevance of foreign inflation in explaining changes in domestic inflation. Specifically, there are three different patterns of break point changes compared to the univariate models. First, the number of coefficient breaks increases for a small number of countries, including Austria, Germany, Switzerland and the US. It is clear for Austria (figure 1.1.1c) and Switzerland (figure 1.1.17c) that the new breaks reflect changes in the relationship between domestic and foreign inflation, as the univariate models did not exhibit any breaks. For Germany, a coefficient break in 1981 is replaced by a variance break, and two more coefficients breaks are detected in 1976 and 1990 (see figure 1.1.7c). Following each break, an increase in the contemporaneous foreign inflation coefficient and a decrease in the unconditional mean are found. Additionally, the relatively high and constant persistence observed in the univariate specification seems to be knocked out by foreign inflation, as no lags are selected in the bivariate case. For the US, the break in 1990 is primarily domestic, leading to lower persistence and lower mean (see figure 1.1.19c). Although not detected in the univariate analysis, the 1977 break also appears as primarily domestic, whereas that in 2003 seems to be caused by an increased role of foreign inflation. In general, for Austria, Germany, Switzerland and US, the overall direction appears to be towards stronger linkages with foreign inflation.

Second, for some countries, the number of coefficient breaks decreases in the presence of foreign inflation. Interestingly, some coefficient breaks appear to be replaced by variance breaks, but not necessarily at the same date. For France and UK (figures 1.1.6c and 1.1.18c), for example, one coefficient break is replaced by a variance break while the remaining coefficient breaks hardly change their locations. Table 1.4 shows an increased number of variance breaks compared to the univariate models in Table 1.3, and corresponding large declines of variances, noted in the graphs. Indeed, newly obtained variance breaks seem to be located in more plausible places in a sense that they separate high and low volatility parts of the graph than the coefficient breaks of the univariate model. This may be indicative of the fact that the univariate models are essentially misspecified if foreign inflation is omitted. Possibly, an effect of the

omitted variable may have been interpreted as a coefficient break which then disappears once relevant variable is included. However, this also could be related to the difficulty, discussed by Pitarakis (2004), of distinguishing between coefficient and variance breaks.

The third pattern covers countries where previously identified univariate coefficient breaks are altered although the number of breaks is unchanged. This may also indicate misspecification of the univariate models, as they omit effects of foreign environments. The largest variation in terms of location shifts of the coefficient breaks occurs in Japan by almost two decades (see figure 1.1.10). Variance breaks in univariate inflation, on the other hand, remain more or less the at same locations. However, for a few instances such as in Korea, Japan, Sweden and Spain (figures 1.1.11c, 1.1.10c, 1.1.16c and 1.1.15c, respectively), a new variance break appears in addition to the breaks identified in the univariate models. Finally, previously identified coefficient breaks are unchanged in the bivariate models of Norway and Sweden (figures 1.1.13c and 1.1.16c).

Despite the implied misspecification of univariate models, the general results of declining persistence and mean of domestic inflation remain in the bivariate models.

1.6 Sensitivity analysis

To assess the sensitivity of the results presented above, we extend the bivariate inflation models by including an additional variable which potentially has an impact on domestic inflation. The additional variables, oil price inflation and the change in trade weighted real effective exchange rates (EER)[21], are added to the bivariate models of inflation one at a time. These variables are measured in terms of percentage changes, as for CPI inflation. The first is employed because a sudden increase in oil price can cause an exogeneous inflationary shock to domestic inflation, and omitting this variable may result in upward bias in the estimated coefficients. Further, the effective exchange rate may be important in explaining domestic inflation, especially for open economies, through its influence on import and export prices.

The selected models including these variables are presented in Table 1.5 and Table 1.6 with the estimated coefficient and variance breaks using the selected models reported in Table 1.7 and Table 1.8, respectively. For convenience, the

[21] We also employed the nominal effective exchange rates for the purpose of the robustness analysis. However, it is not picked up by the model selection procedure for any country, consequently yielding the same models as the bivariate inflation models.

estimated coefficient and variance breaks using bivariate models (previously presented in Table 1.4) are repeated in Table 1.7 and Table 1.8, respectively. The results suggest that including either of these variables does not make a qualitative change for most countries.

In Table 1.7, previously identified coefficient breaks in bivariate models remain in a qualitatively similar location for most countries, when including oil price inflation in Table 1.5. The few exceptions are Belgium, Denmark, and US where some coefficient breaks are dropped, and Germany where the number of breaks increases. This indicates the potential misspecification of bivariate inflation models corresponding to these countries. Perhaps, in the absence of oil price inflation in bivariate models, its omitted effects to domestic inflation may have interpreted as an extra break in the estimation. While the number of coefficient breaks using bivariate models decreases when such variable is included in the model. For example, the US is known as one of the biggest oil importers, and two of three breaks (in 1970s and 1990s) are dropped when oil price inflation is included in the model. Despite the importance of oil price inflation for these countries, a decline in the number of breaks may also be due to a loss of power when testing for all coefficients.

Table 1.7 also compares coefficient breaks in bivariate models to the models with EER variable. EER appears to have less impact on big open economies and most of the Euro area. However, relatively small economies, Finland, Greece, Norway and Sweden, are sensitive to the inclusion of the EER. These countries yield an additional coefficient break which occurs prior or running up to the introduction of the European Monetary Union (EMU). The weakened role of EER after the EMU for these countries is evident in the estimates of the corresponding coefficients (see Table 1.10). Indeed, we do not expect large exchange rate fluctuations to play a role in explaining domestic inflation after the introduction of the EMU, especially for those belonging to EMU and who trade mostly with Euro area countries. Spain and Austria, on the other hand, drop coefficient breaks around early 2000. This may point misspecified bivariate models where the effects of the omitted exchange rate was captured previously as a break.

Furthermore, variance breaks obtained from the models with oil price inflation do not show any substantive differences from the breaks detected in the bivariate models. But, there is some variation from the model with EER for a small number of cases (see Table 1.8 for further details).

Table 1.9 and Table 1.10 provide estimated coefficients for models with oil price inflation and EER, respectively. Graphical illustrations of the results shown in these tables and comparison with bivariate models are given in figures

1.2.1-1.2.19. Each figure consists of four graphs showing changes in the estimates of persistence, mean, contemporaneous foreign inflation, and the sum of the contemporaneous and (/or) lagged coefficients of third variables (EER and oil inflation). For Italy in figure 1.2.9, for instance, subplot (a) depicts changes in persistence that are estimated using the bivariate model (in black line), the model with oil price inflation (in red line) and the model with EER (in blue line). A similar interpretation applies to the contemporaneous foreign inflation coefficients and the subsample mean in subplots (b) and (c) respectively. Subplot (d) shows sum of estimated coefficients corresponding to the contemporaneous and lagged oil price inflation (in red line) and EER (in blue line). A missing line either in subplot (a) or (d) indicates the absence of the corresponding lags (and contemporaneous variable) in the model.

In general, despite the break point changes for a few cases above, conclusions drawn from the bivariate models largely carry over. Looking at the figures, estimated persistence and mean of inflation typically show substantial declining patterns regardless of the different models, represented by the lines in the graphs. Moreover, the increasing and positive contemporaneous relationship between domestic and foreign inflation remains robust. We should note, however, that there are some countries (Japan, Netherlands and the UK) where contemporaneous coefficients do not increase, but those are the same countries that show the constant contemporaneous effect in bivariate models of inflation. Finally, the figures also show that an impact of the lagged and (/or) contemporaneous third variables on domestic inflation is relatively small.

1.7 Concluding remarks

This chapter adds to the existing literature on international inflation by comprehensively examining the structural stability in the relationship between domestic and country specific foreign inflation. For this aim, we propose and employ an iterative structural break testing methodology which is designed to deliver reliable inferences on structural breaks. In the iteration, we account for breaks in the conditional mean (which comprises intercept, autoregressive coefficients and coefficients on foreign inflation) and variance parameters by iterating between tests for conditional mean and variance breaks, while also taking care of outliers.

We document evidence of structural breaks in the linkage of domestic and country specific foreign inflation. Furthermore, taking into account the identified breaks, we find positive and increasing contemporaneous relationships between domestic and foreign inflation for most countries. This finding is com-

patible with the co-movement of inflation in different countries, documented widely in the literature (see Ciccarelli and Mojon, 2010, Neely and Rapach, 2011, etc). This also verifies the finding by Bataa et al. (2013a) who note increased contemporaneous correlations of inflation in a much more restricted number of G-7 countries. Moreover, the timing of breaks in mean and variances across countries exhibit notable clusters around the mid 1970s, early 1980s and early 1990s. The presence of such clusters already suggests the dependence of domestic inflation on foreign economic environments.

It appears to be widely accepted that changes in inflation have been mainly in the mean, with clusters of mean breaks documented in the univariate context. When applied to univariate inflation models, our procedure indicates that almost all countries in the data set experience at least one variance break, leading to substantial volatility declines. Furthermore, these breaks also show a clustering pattern. Overall the results from both univariate and bivariate inflation models suggest, declining unconditional mean and persistence of domestic inflation, consistent with the existing findings. Results on changes in inflation co-movement are robust to the inclusion of either oil price inflation or real effective exchange rates.

Finally, we emphasize that the use of the iterative structural break testing procedure was important to establish these findings. As shown using some illustrated cases in subsection 1.5.1, not employing this iterative procedure would lead to potentially substantial changes in the detected structural breaks compared to using non-iterated testing procedure.

Tables

Table 1.1: Trade weights by partner countries (monthly average over 40 years)

	Aus	Belg	Can	Den	Fin	Fra	Ger	Gre	Ita	Jap	Kor	Net	Nor	Por	Spa	Swe	Swi	Uk	US	SUM
Aus	0.000	0.025	0.008	0.012	0.009	0.052	0.488	0.006	0.111	0.025	0.005	0.035	0.007	0.006	0.018	0.024	0.072	0.047	0.050	1.000
Belg	0.011	0.000	0.005	0.009	0.006	0.209	0.246	0.005	0.059	0.024	0.004	0.177	0.009	0.009	0.030	0.018	0.016	0.100	0.064	1.000
Can	0.002	0.007	0.000	0.002	0.002	0.012	0.022	0.001	0.011	0.053	0.010	0.008	0.006	0.001	0.003	0.004	0.005	0.035	0.816	1.000
Den	0.013	0.032	0.008	0.000	0.033	0.058	0.241	0.006	0.047	0.032	0.006	0.063	0.066	0.007	0.018	0.157	0.021	0.125	0.070	1.000
Fin	0.015	0.032	0.010	0.045	0.000	0.059	0.191	0.006	0.043	0.040	0.009	0.056	0.047	0.008	0.022	0.181	0.022	0.132	0.083	1.000
Fra	0.011	0.120	0.010	0.010	0.007	0.000	0.239	0.008	0.136	0.029	0.007	0.069	0.012	0.013	0.073	0.019	0.042	0.099	0.096	1.000
Ger	0.061	0.091	0.011	0.025	0.013	0.155	0.000	0.010	0.107	0.038	0.010	0.126	0.019	0.010	0.037	0.032	0.059	0.090	0.106	1.000
Gre	0.017	0.043	0.007	0.014	0.011	0.099	0.250	0.000	0.174	0.063	0.019	0.072	0.006	0.004	0.034	0.021	0.021	0.077	0.069	1.000
Ita	0.033	0.053	0.013	0.012	0.007	0.193	0.261	0.018	0.000	0.023	0.007	0.062	0.007	0.010	0.051	0.017	0.055	0.078	0.101	1.000
Jap	0.005	0.017	0.057	0.008	0.005	0.032	0.081	0.006	0.023	0.000	0.109	0.029	0.008	0.003	0.011	0.011	0.020	0.053	0.524	1.000
Kor	0.004	0.010	0.033	0.005	0.005	0.025	0.066	0.004	0.024	0.339	0.000	0.021	0.005	0.001	0.010	0.008	0.011	0.036	0.392	1.000
Net	0.012	0.160	0.007	0.016	0.010	0.107	0.322	0.007	0.057	0.022	0.006	0.000	0.015	0.007	0.025	0.025	0.018	0.107	0.080	1.000
Nor	0.008	0.027	0.027	0.066	0.031	0.062	0.159	0.005	0.030	0.035	0.008	0.070	0.000	0.008	0.015	0.151	0.012	0.212	0.075	1.000
Por	0.011	0.040	0.009	0.015	0.010	0.130	0.177	0.003	0.069	0.025	0.003	0.054	0.015	0.000	0.179	0.030	0.028	0.126	0.080	1.000
Spa	0.010	0.039	0.010	0.010	0.008	0.208	0.184	0.007	0.107	0.033	0.007	0.053	0.007	0.049	0.000	0.019	0.024	0.101	0.124	1.000
Swe	0.016	0.046	0.011	0.092	0.073	0.062	0.186	0.005	0.041	0.030	0.007	0.063	0.103	0.008	0.020	0.000	0.022	0.123	0.092	1.000
Swi	0.050	0.036	0.009	0.013	0.008	0.126	0.317	0.005	0.112	0.041	0.006	0.045	0.006	0.007	0.025	0.022	0.000	0.076	0.098	1.000
Uk	0.011	0.068	0.033	0.026	0.018	0.117	0.169	0.006	0.062	0.043	0.009	0.102	0.035	0.012	0.038	0.042	0.040	0.000	0.170	1.000
US	0.005	0.027	0.354	0.006	0.004	0.046	0.089	0.003	0.037	0.207	0.048	0.035	0.007	0.003	0.016	0.014	0.020	0.077	0.000	1.000

Note: The shares for each country with respect to partner countries are given in the rows of the table.

Table 1.2: Autoregressive lags of univariate and bivariate models

Country	Univariate Models (maxlag=17)	Bivariate models (maxlag=12)	
	Domestic lags	Domestic lags	Foreign lags
Austria	[1;2;3;6;7;11]	NA	[0,4]
Belgium	[1;3;4;9;11;12]	[1;4]	[0]
Canada	[1;2;3;4;7;9;12;16]	[5;7;9]	[0,9]
Denmark	[1;3;6;8;11]	[1;3]	[0]
Finland	[2;5;6;7;9]	[2;9]	[0,4]
France	[1;3;6;8;12;17]	[1;3;10]	[0]
Germany	[2;6;7;8;9;11]	NA	[0]
Greece	[1;2;5;8;17]	[1;2;5;8]	[0]
Italy	[1;3;6;12;17]	[1;3;6]	[0,2]
Japan	[1;3;5;7;9;11;12;16]	[3;5;9;11]	[0]
Korea	[1;3;9;12;15]	[1]	[0,6]
Netherland	[1;4;6;7;8;17]	[4;6;8]	[0]
Norway	[1;2;3;7;8]	[1;3;7;8]	[0]
Portugal	[1;6;9;16]	[1;6;9]	[0,2]
Spain	[1;2;8;10;12;13;15]	[1;8;10]	[0]
Sweden	[2;3;7;8;9]	[7;8;9]	[0]
Switzerland	[1;2;4;6;10;17]	[1;6;9]	[0]
UK	[1;2;3;13]	[1;2;3]	[0]
US	[1;7;9;11;12]	[1]	[0]

Note: Autoregressive lags are obtained at convergence of the iterations.
For the domestic models, all combinations of lags are considered as discussed in subsection 1.3.1, while the bivariate models compared are based on a testing down approach as described in subsection 1.3.3. For both approaches, the final model is selected based on SIC.

Table 1.3: Breaks in univariate models

Country	Breaks in the set of coefficients				Breaks in the variances				
	1970s	1980s	1990s	2000s	1970s	1980s	1990s	2000s	NI
Austria	.	.	.	.	1977-Dec	.	.	.	1
Belgium	.	.	1994-Aug	2003-Dec	.	1985-Sep	.	.	2
Canada	.	1982-Jul	1990-Dec	.	1978-Nov	.	.	2000-Mar	2
Denmark	.	1982-Nov	1990-Jan	.	.	1980-Oct	1990-Dec	.	3
Finland	.	.	1991-Mar	.	1976-Nov	1983-Jul	.	.	[]
France	.	1986-Jan	1991-Dec	.	.	.	.	.	2
Germany	.	1981-Nov	.	.	.	.	.	.	2
Greece	.	.	1992-Nov	.	1977-Jul	.	1992-Sep	.	3
Italy	.	1982-Sep	1995-Jul	.	.	1981-Feb 1987-Jan	.	.	3
Japan	.	1980-Aug	1992-Jun	.	1977-May	.	1992-Jan	.	4
Korea	.	.	1998-Mar	.	.	1981-Mar	.	.	2
Netherland	.	1982-Feb	.	.	1978-Aug	.	.	.	3
Norway	.	1988-Apr	.	.	.	1982-Feb	.	2000-Dec	2
Portugal	.	1984-Aug	1992-Jun	.	1977-May	1985-Apr	1992-May	.	3
Spain	.	1986-Feb	1995-Mar	.	1977-Aug	1986-Aug	.	.	4
Sweden	.	.	1991-Feb	.	.	1985-Aug	.	.	3
Switzerland	.	.	.	.	.	1983-Feb	.	.	1
Uk	.	1982-Jun	1991-Dec	.	.	1982-Apr	.	.	4
US	.	.	1991-Feb	.	.	1982-Jul	.	2004-Oct	2
Total	0	11	14	1	8	13	4	3	

Note: Column NI represents the number of iterations required to converge to a single set of break dates. [] indicates the set of break dates is selected by a minimum SIC criterion.

Table 1.4: Breaks in bivariate models

Country	Breaks in the set of coefficients				Breaks in the variances				
	1970s	1980s	1990s	2000s	1970s	1980s	1990s	2000s	NI
Austria	1976-Sep	.	.	2000-Apr	1976-Sep				3
Belgium	.	1982-Apr	1995-Nov	.	.	1983-Nov	.	.	2
Canada	.	.	1990-Dec	.	1978-Nov	.	.	.	2
Denmark	.	1982-Nov	1994-Jun	.	.	1980-Oct	1991-Feb	.	3
Finland	.	.	1990-Mar	.	1976-Nov	1983-Jul	.	.	2
France	.	1985-Aug	.	.	.	1983-Jan	.	.	3
Germany	1976-Mar	.	1990-Sep	.	.	1982-Jul	.	.	[]
Greece	.	.	.	2000-Dec	1976-Oct	.	1993-Apr	.	2
Italy	.	1986-Jan	1996-May	.	.	1981-Feb 1987-Feb	.	.	2
Japan	1977-Jan	.	.	.	1977-Jan	1985-Nov	1993-Nov	.	2
Korea	.	1985-Sep	.	.	.	1982-Mar	.	2003-Apr	5
Netherland	.	1989-Apr	.	.	1978-Aug	.	.	.	3
Norway	.	1988-Apr	.	.	.	1980-Mar	.	.	3
Portugal	.	1985-Mar	1992-Jul	.	1978-May	1985-Mar	1992-May	.	2
Spain	.	1986-Jul	.	2004-May	1977-Nov	1986-Feb	1992-Nov	.	4
Sweden	.	.	1991-Feb	.	1977-Jul	.	1993-Jan	.	3
Switzerland	.	1984-Oct	.	.	.	1982-Jun	.	.	2
Uk	.	1980-May	.	.	.	.	1991-Aug	.	[]
US	1977-Nov	.	1990-Oct	2003-Feb	.	1983-May	.	2004-Sep	3
Total	4	11	9	4	9	14	7	2	

Note: Column NI represents the number of iterations required to converge to a single set of break dates. [] indicates the set of break dates is selected by a minimum SIC criterion.

Table 1.5: Sensitivity analysis: Selected models with oil price inflation

	Domestic lags	Foreign lags	Oil lags
Austria	N/A	[0,4]	[3]
Belgium	[1,4,9]	[0]	[0,5]
Canada	[5,7,9]	[0,9]	N/A
Denmark	[1,6,11]	[0]	[0,5]
Finland	[2,9,11]	[0,4]	[1]
France	[1,3,8,10]	[0]	[0]
Germany	[6]	[0]	[1]
Greece	[1,2,5,8]	[0,5]	N/A
Italy	[1,3,6]	[0,2]	[8]
Japan	[3,5,9,11]	[0]	[1]
Korea	[1]	[0,6]	[12]
Netherland	[4,6,8]	[0]	N/A
Norway	[1,6,8,10]	[0]	[12]
Portugal	[1,6,9]	[0*,2]	[1]
Spain	[1,8,10]	[0]	N/A
Sweden	[7,8,9]	[0]	[9]
Switzerland	[1,6,9]	[0]	[0,4]
Uk	[1,2,3]	[0]	N/A
US	[1,7]	[0]	[0,1]

Note: * indicates that the model selected by SIC does not include contemporaneous foreign variable. However, we test for breaks on the model including contemporaneous foreign inflation.

Table 1.6: Sensitivity analysis: Selected models with EER

	Domestic lags	Foreign lags	EER lags
Austria	[11]	[0]	[0]
Belgium	[1,4]	[0]	[0]
Canada	[5,7]	[0,9]	[0]
Denmark	[1,6]	[0]	[0]
Finland	[9]	[0,4]	[0]
France	[1,3,10]	[0]	N/A
Germany	N/A	[0]	[0,5]
Greece	[1,2,6,8]	[0]	[0]
Italy	[1,3,6]	[0,2]	[4]
Japan	[5,9,11]	[0,1]	[0]
Korea	[1]	[0,6]	N/A
Netherland	[4,6,8]	[0]	[0]
Norway	[1,2,6,8]	[0,11]	[0]
Portugal	[1,6,9]	[0*,2]	[0]
Spain	[1,2,8,9]	[0]	[5]
Sweden	[3,8,9]	[0]	[0]
Switzerland	[1,6,9]	[0]	N/A
Uk	[1,2,3]	[0]	N/A
US	[1]	[0]	N/A

Note: * indicates that the model selected by SIC does not include contemporaneous foreign variable. However, we test for breaks on the model including contemporaneous foreign inflation.

Table 1.7: Sensitivity analysis: Coefficient Breaks

Country	Bivariate models				Models with oil price inflation				Models with EER			
	1970s	1980s	1990s	2000s	1970s	1980s	1990s	2000s	1970s	1980s	1990s	2000s
Austria	1976-Sep	.	.	2000-Apr	1976-Sep	.	1999-Aug	.	.	.	.	.
Belgium	.	1982-Apr	1995-Nov	.	.	.	1995-Oct	.	.	1981-Nov	1995-Nov	.
Canada	.	.	1990-Dec	.	.	.	1990-Dec	.	.	1987-Feb	.	.
Denmark	.	1982-Nov	1994-Jun	.	.	1985-Apr	.	.	.	1989-Sep	1995-Sep	.
Finland	.	.	1990-Mar	.	.	.	1990-Feb	.	.	1983-Apr	1991-Jul	.
France	.	1985-Aug	.	.	.	1987-Jan 1980-Nov	.	.	.	1985-Aug	.	.
Germany	1976-Mar	.	1990-Sep	.	.	1987-Dec	.	2001-Feb	1976-Jul	.	1990-Sep	.
Greece	.	.	.	2000-Dec	.	.	.	2000-Dec	.	.	1994-Aug	2000-Dec
Italy	.	1986-Jan	1996-May	.	.	1986-Jan	1996-May	.	.	1986-Jan	1996-May	.
Japan	1977-Jan	.	.	.	1977-Jun	.	.	.	1977-Jan	.	.	.
Korea	.	1985-Sep	.	.	.	1981-Oct	.	.	.	1985-Sep	.	.
Netherland	.	1989-Apr	.	.	.	1989-Apr	.	.	.	1987-Jun	.	.
Norway	.	1988-Apr	.	.	.	1988-Apr	.	.	.	1981- May	.	.
Portugal	.	1985-Mar	1992-Jul	.	.	1985-Mar	1992-Jul	.	.	1984-Aug	1990-Oct	.
Spain	.	1986-Jul	.	2004-May	.	1986-Jul	.	2004-May	.	1986-Jul	.	.
Sweden	.	.	1991-Feb	.	.	.	1991-Feb	.	.	1983-Jan	1991-Feb	.
Switzerland	.	1984-Oct	.	.	.	1984-Oct	.	.	.	1984-Oct	.	.
Uk	.	1980-May	.	.	.	1980-May	.	.	.	1980-May	.	.
US	1977-Nov	.	1990-Oct	2003-Feb	.	.	.	2003-Feb	1977-Nov	.	1990-Oct	2003-Feb
Total	4	11	9	4	2	12	7	4	3	15	9	2

Table 1.8: Sensitivity analysis: Variance Breaks

Country	Bivariate models				Models with oil price inflation				Models with EER			
	1970s	1980s	1990s	2000s	1970s	1980s	1990s	2000s	1970s	1980s	1990s	2000s
Austria	1976-Sep				1976-Sep	.	.	.	1977-Jan	.	.	.
Belgium	.	1983-Nov	.	.	.	1984-May	.	.	.	1983-Jan	.	.
Canada	1978-Nov	.	.	.	1978-Nov	.	.	.	1978-Nov	.	.	.
Denmark	.	1980-Oct	1991-Feb	.	1977-Oct	1984-Apr	.	.	.	1983-Apr	1991-Mar	.
Finland	1976-Nov	1983-Jul	.	.	1977-Jan	1983-Jul	.	.	.	1983-Jul	.	.
France	.	1983-Jan	.	.	.	1983-May	.	.	.	1983-Jan	.	.
Germany	.	1982-Jul	.	.	1976-Sep	.	.	.	.	.	.	.
Greece	1976-Oct	.	1993-Apr	.	1976-Oct	.	1993-Jun	.	1976-Oct	1986-Mar	1994-Aug	
Italy		1981-Feb				1981-Feb				1981-Feb		
	.	1987-Feb	.	.	.	1987-Feb	.	.	.	1987-Feb	.	
Japan	1977-Jan	1985-Nov	1993-Nov	.	1977-Jan	1986-Mar	1993-Dec	.	1977-Jan	1985-Dec	1995-Nov	.
Korea	.	1982-Mar	.	2003-Apr	.	1981-Apr	.	2003-Apr	.	1982-Mar	.	2003-Apr
Netherland	1978-Aug	.	.	.	1978-Aug	.	.	.	1979-Feb	.	.	.
Norway	.	1980-Mar	.	.	.	1982-Aug	.	.	.	1982-Aug	.	2002-Dec
Portugal	1978-May	1985-Mar	1992-May	.	1978-May	1985-Feb	1992-May	.	1977-Mar	1985-Mar	1993-Sep	.
Spain	1977-Nov	1986-Feb	1992-Nov	.	1977-Nov	1986-Feb	1992-Nov	.	1977-Sep	1986-Feb	1992-Nov	.
Sweden	1977-Jul	.	1993-Jan	.	1977-Jul	.	1993-Jan	.	.	.	1993-Jan	.
Switzerland	.	1982-Jun	.	.	.	1983-Feb	.	.	.	1982-Jun	.	.
Uk	.	.	1991-Aug	.	.	.	1991-Aug	.	.	1982-Jul	.	.
US	.	1983-May	.	2004-Sep	.	1983-May	.	2004-Oct	.	1983-May	.	2004-Sep
Total	9	14	7	2	11	13	6	2	7	15	6	3

Table 1.9: Sensitivity analysis: Estimated coefficients in regimes

	Persistence	Mean	Foreign inflation (lags)	Foreign inflation (contemporaneous)	Oil price change (lags)	Oil price change (contemporaneous)
Austria	N/A	0.57; 0.27; 0.15	0.04; 0.29; 0.04	0.72; 0.35; 0.75	0.004; -0.0009; 0.004	N/A
Belgium	0.52; -0.02	0.43; 0.16;	N/A	0.51; 1.36	0.002; 0.003	0.003; 0.002
Canada	0.38; 0.008	0.55; 0.17	0.29; 0.08	0.23; 0.68	N/A	N/A
Denmark	0.25; 0.21	0.76; 0.21	N/A	0.55; 0.32	0.008; -0.0008	0.006; 0.007
Finland	0.36; 0.11	0.70; 0.17	0.07; 0.14	0.55; 0.70	0.004; 0.0007	N/A
France	0.68; -0.02	0.74; 0.16	N/A	0.34; 0.77	N/A	0.004; 0.002
Germany	0.28; -0.17; 0.07; 0.02	0.40; 0.23; 0.19; 0.13	N/A	0.08; 1; 0.65; 1.01	0.004; 0.003; 0.006; -0.006	N/A
Greece	0.75; -0.32	1.1; 0.27	0.04; 0.25	0.45; 0.93	N/A	N/A
Italy	0.59; 0.36; 0.22	1.06; 0.44; 0.18	0.63; 0.16; 0.16	0.05; 0.22; 0.41	0.007; 0.001; 0.0009	N/A
Japan	0.38; 0.23;	0.84; 0.12	N/A	0.43; 0.45	0.01; 0.0004	N/A
Korea	0.41; 0.19	1.24; 0.40	0.59; 0.17	0.42; 0.52	N/A	N/A
Netherland	0.56; 0.35	0.42; 0.18	N/A	0.46; 0.41	N/A	N/A
Norway	0.55; 0.30	0.68; 0.21	N/A	0.14; 0.51	0.006; -0.0008	N/A
Portugal	0.34; 0.01; 0.42	1.51; 0.99; 0.27	0.27; 0.52; 0.04	(-1.2; 0.22; 0.66)	0.02; 0.0007; -0.0008	N/A
Spain	0.41; 0.45; 0.19	1.09; 0.37; 0.21	N/A	0.46; 0.52; 1.2	N/A	N/A
Sweden	0.31; 0.14	0.67; 0.16	N/A	0.42; 0.82	0.008; -0.0009	N/A
Switzerland	0.60; 0.24	0.40; 0.14	N/A	0.20; 0.87	0.009; 0.001	0.0008; 0.002
UK	0.68; 0.33	1.03; 0.31	N/A	0.60, 0.58	N/A	N/A
US	0.42; 0.21	0.40; 0.19	N/A	0.37; 1.18	0.005; 0.003	0.005; 0.006

Table 1.10: Sensitivity analysis: Estimated coefficients in regimes

	Persistence	Mean	Foreign inflation (lags)	Foreign inflation (contemporaneous)	Change in EER (lags)	Change in EER (contemporaneous)
Austria	0.10	0.29	N/A	0.6	N/A	0.11
Belgium	0.51; 0.28; -0.01	0.58; 0.30; 0.16	N/A	0.53; 1.10; 1.39	N/A	0.07; 0.06; 0.04
Canada	0.25; 0.12	0.60; 0.20	0.31; 0.09	0.27; 0.69	N/A	0.08; 0.02
Denmark	0.21; -0.02; -0.01	0.66; 0.19; 0.17	N/A	0.79; 0.23; 0.77	N/A	0.09; 0.01; 0.05
Finland	0.28; -0.04; -0.04	0.85; 0.56; 0.14	0.24; -0.03; 0.18	0.49; 0.65; 0.76	N/A	0.16; 0.19; -0.0002
France	0.69; -0.04	0.77; 0.17	N/A	0.30; 0.81	N/A	N/A
Germany	N/A	0.47; 0.25; 0.16	N/A	0.18; 0.62; 0.92	0.07; -0.01; 0.01	0.04; 0.02; 0.03
Greece	0.47; 0.63; -0.13	1.26; 0.50; 0.27	N/A	0.32; 0.80; 0.88	N/A	0.09; -0.03; 0.007
Italy	0.61; 0.37; 0.18	1.06; 0.44; 0.18	0.62; 0.13; 0.16	0.08; 0.23; 0.42	0.03; 0.005; 0.01	N/A
Japan	0.24; 0.16	0.84; 0.13	0.36; 0.23	0.74; 0.35	N/A	0.12; 0.01;
Korea	0.43; 0.17;	1.05; 0.36	0.69; 0.21	0.60; 0.63	N/A	N/A
Netherland	0.57; 0.38	0.47; 0.17	N/A	0.51; 0.39	N/A	0.03; 0.04
Norway	0.22; 0.47; 0.23	0.66; 0.72; 0.21	0.45; -0.06; 0.05	0.20; 0.17; 0.39	N/A	N/A
Portugal	0.45; -0.13; 0.39	1.48; 1.11; 0.32	0.37; 0.48; 0.10	(-0.72; 1.06;0.65)	N/A	N/A
Spain	0.56; 0.30	1.08; 0.32	N/A	0.35; 0.88	0.04; 0.0007	N/A
Sweden	0.40;0.12; 0.22	0.75; 0.56; 0.16	N/A	0.40; 0.83; 0.85	N/A	0.03; 0.33; -0.0005
Switzerland	0.62; 0.20	0.40; 0.14	N/A	0.30; 0.94	N/A	N/A
UK	0.69; 0.39	1.03; 0.31	N/A	0.59; 0.59	N/A	N/A
US	0.02; 0.44; 0.10; 0.13	0.52; 0.49; 0.22; 0.19	N/A	0.66; 0.53; 0.39; 1.44	N/A	N/A

Annotation page for Figures 1.1

graph (a) – graph (a) in each figure presents the break dates as well as some statistics relating to the corresponding regimes from applying the testing procedure of subsection 1.3.1 once to the univariate model of (1.1).

graph (b) – graph (b) in each figure presents the break dates in Table 1.3 as well as some statistics relating to the corresponding regimes from iterating the testing procedure of subsection 1.3.1 multiple times to the univariate model of (1.1) until its convergence.

graph (c) – graph (c) in each figure reports the break dates in Table 1.4 as well as some statistics relating to the corresponding regimes from iterating the testing procedure of subsection 1.3.1 multiple times to the bivariate model of (1.8) until its convergence.

graph (d) – graph (d) in each iteration plots country specific foreign inflation (in red line) for each corresponding country. This series is also plotted in graph (c) in order to compare dynamics between domestic (in blue line) and country specific foreign inflation.

Vertical Lines – The vertical lines indicate the locations of the coefficient break dates with the estimated dates in the boxes next to these lines.

Text arrows – Text arrows point to the locations of variance breaks and the corresponding changes in the variance of the consecutive regimes.

Black dots – Outliers detected at the convergence of the iterations are indicated by black dots if any outliers are detected.

P_D – This shows the estimates of persistence of domestic inflation in each coefficient break regime, measured by the sum of autoregressive coefficients of its own lags.

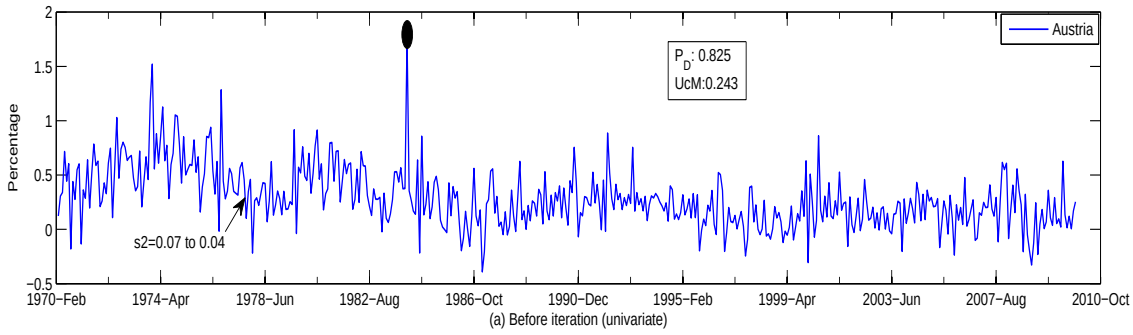
P_F – This shows the estimates of total lagged effect of foreign inflation in each coefficient break regime, measured by the sum of autoregressive coefficients of lagged foreign inflation series.

UcM – This reports unconditional mean of domestic inflation in each coefficient break regime.

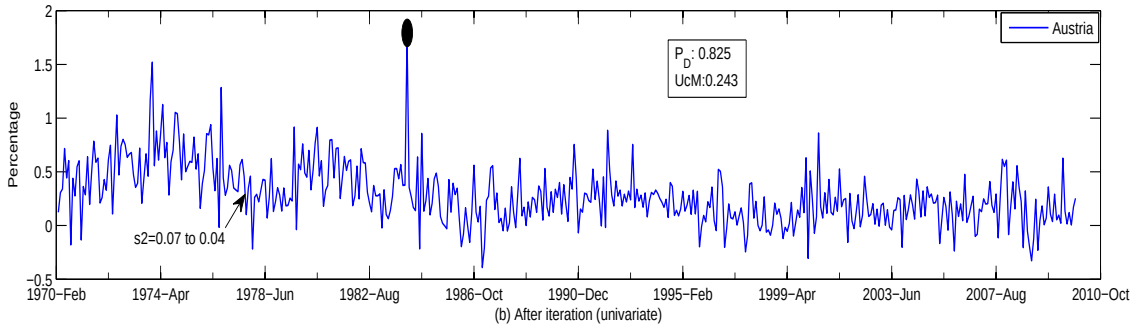
Cont – This shows estimated contemporaneous relationship between domestic and country specific foreign inflation in each coefficient break regime (β_{0j} in equation (1.8)).

Figures 1.1: Inflation Dynamics

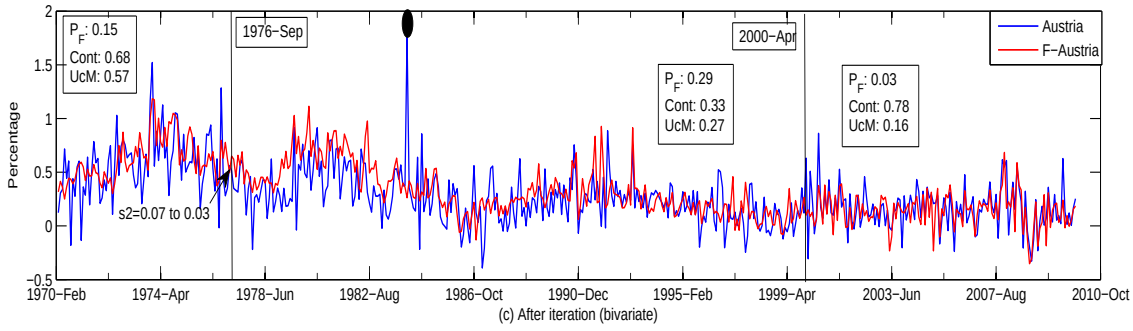
Figure 1.1.1: Inflation Dynamics: Austria



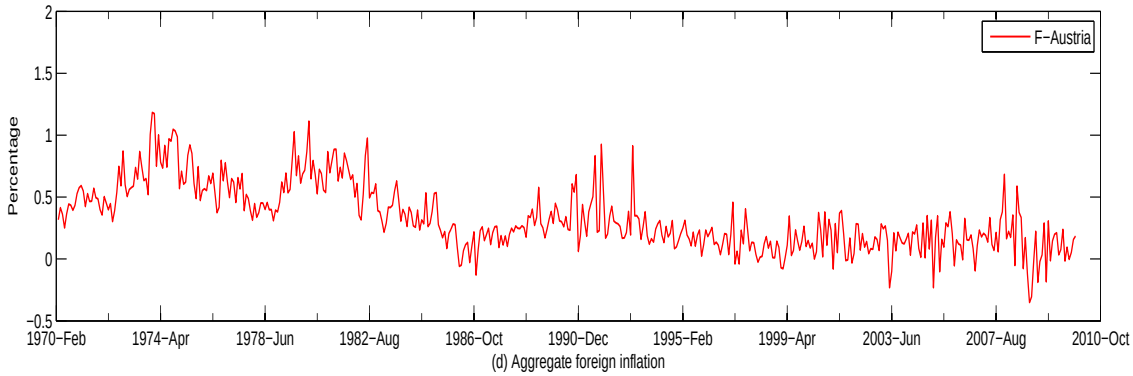
(a) Before iteration (univariate)



(b) After iteration (univariate)



(c) After iteration (bivariate)



(d) Aggregate foreign inflation

Figure 1.1.2: Inflation Dynamics: Belgium

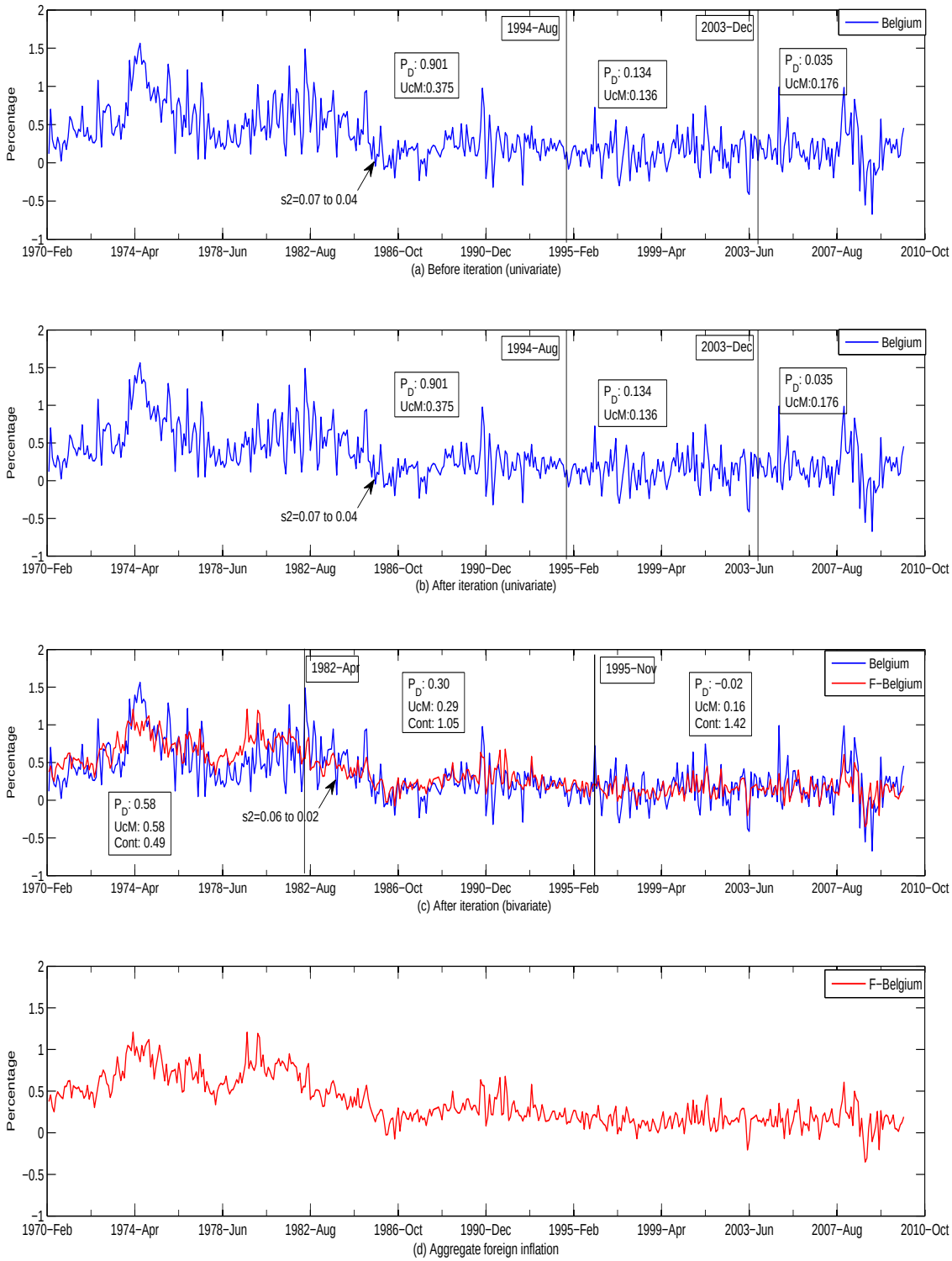


Figure 1.1.3: Inflation Dynamics: Canada

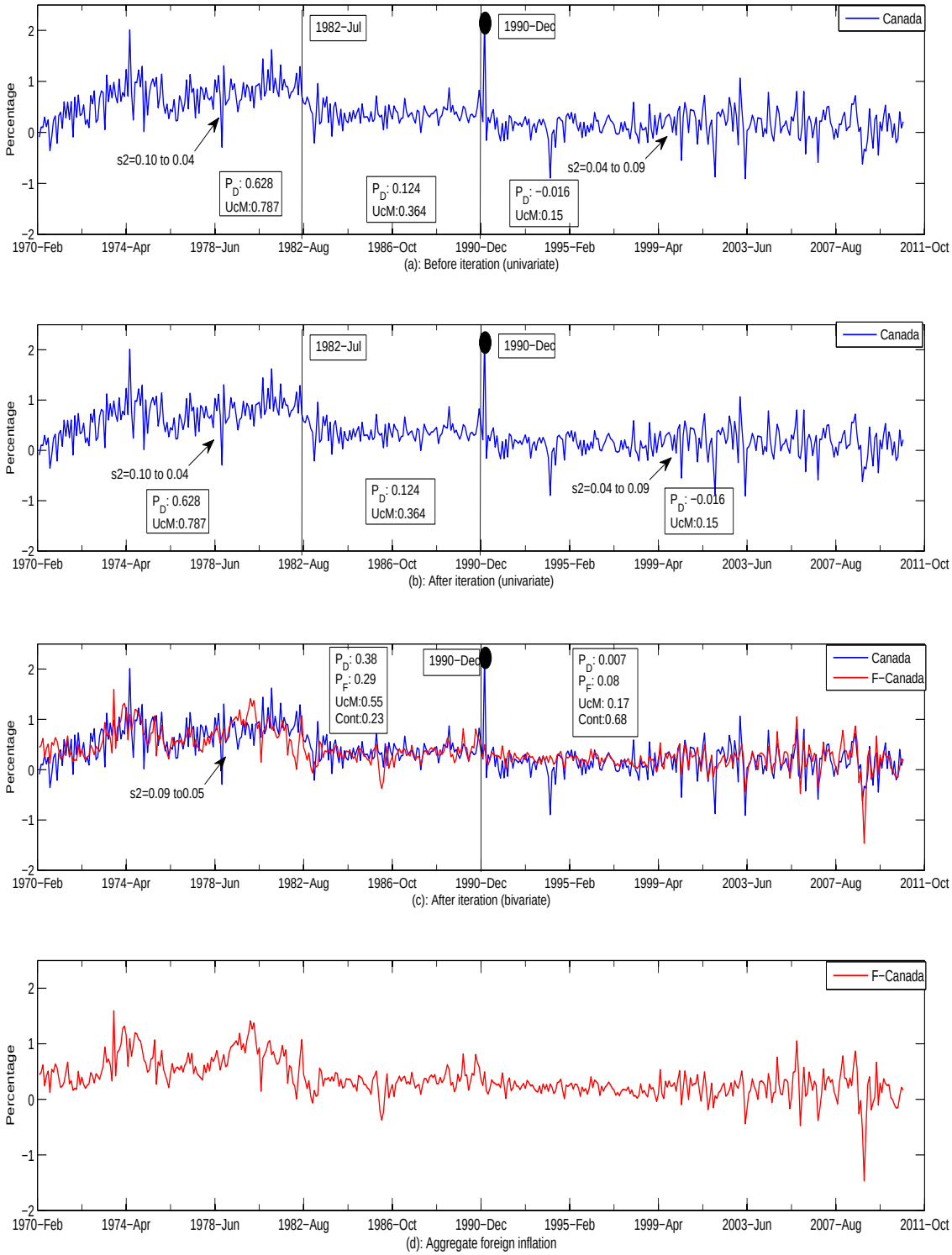


Figure 1.1.4: Inflation Dynamics: Denmark

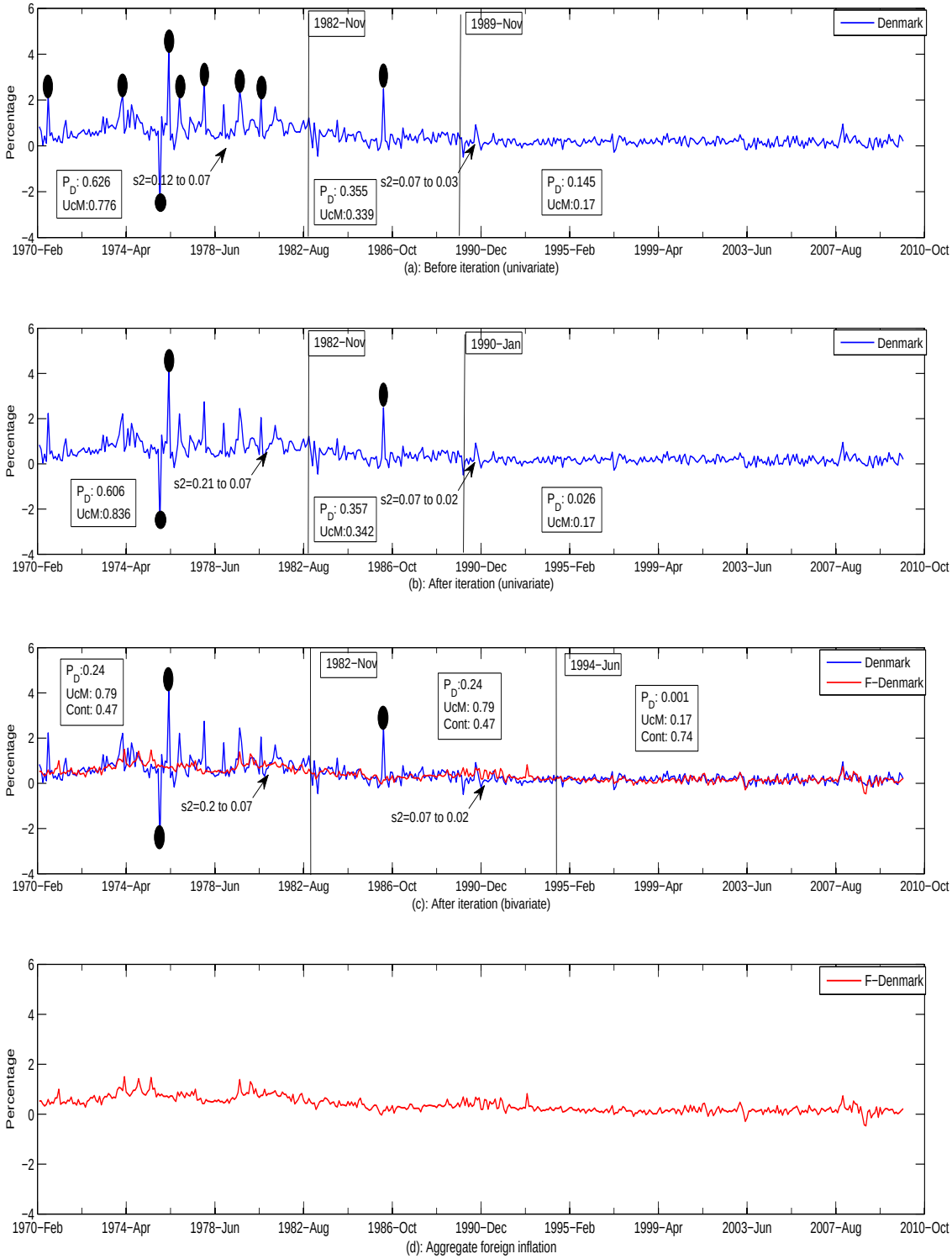


Figure 1.1.5: Inflation Dynamics: Finland

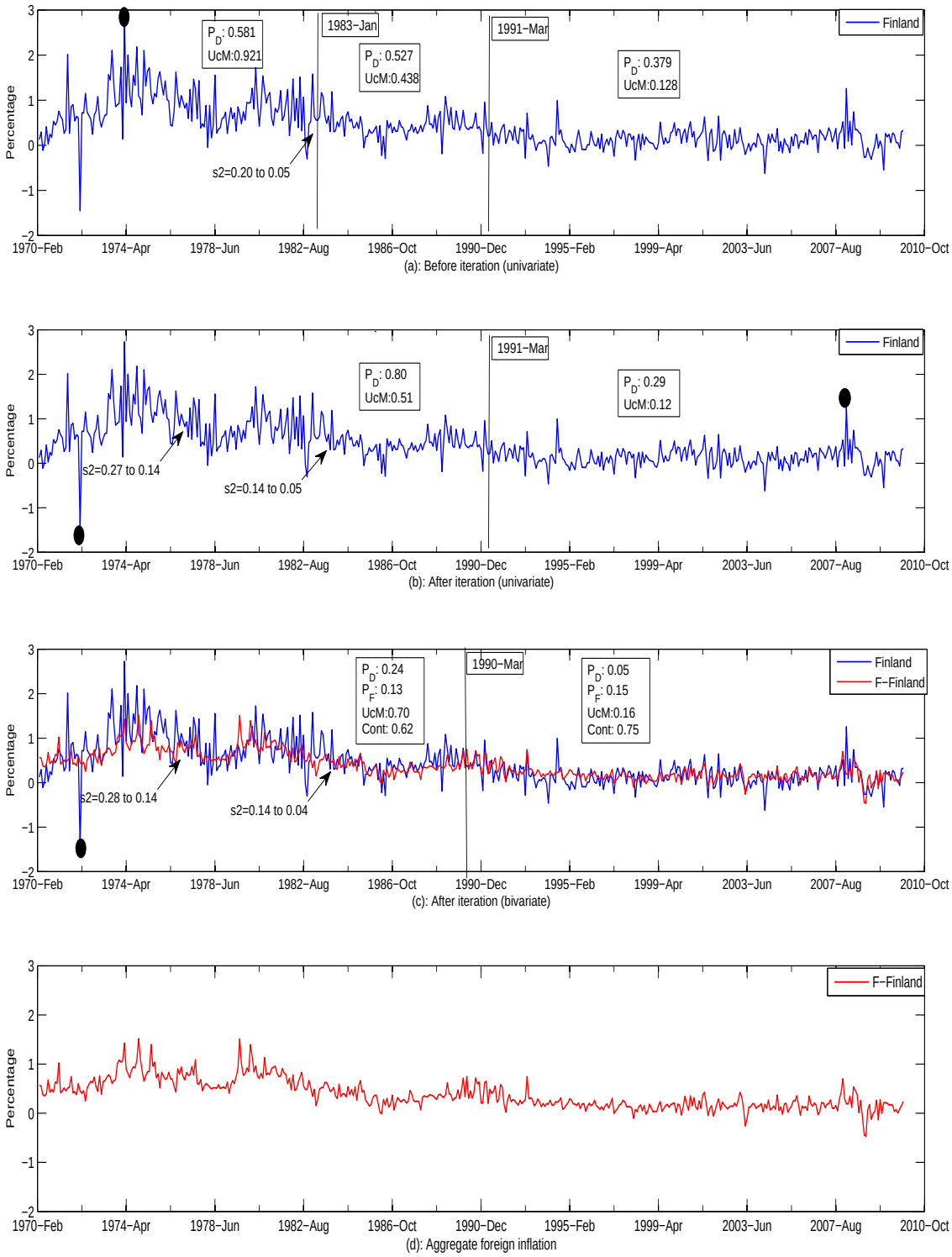


Figure 1.1.6: Inflation Dynamics: France

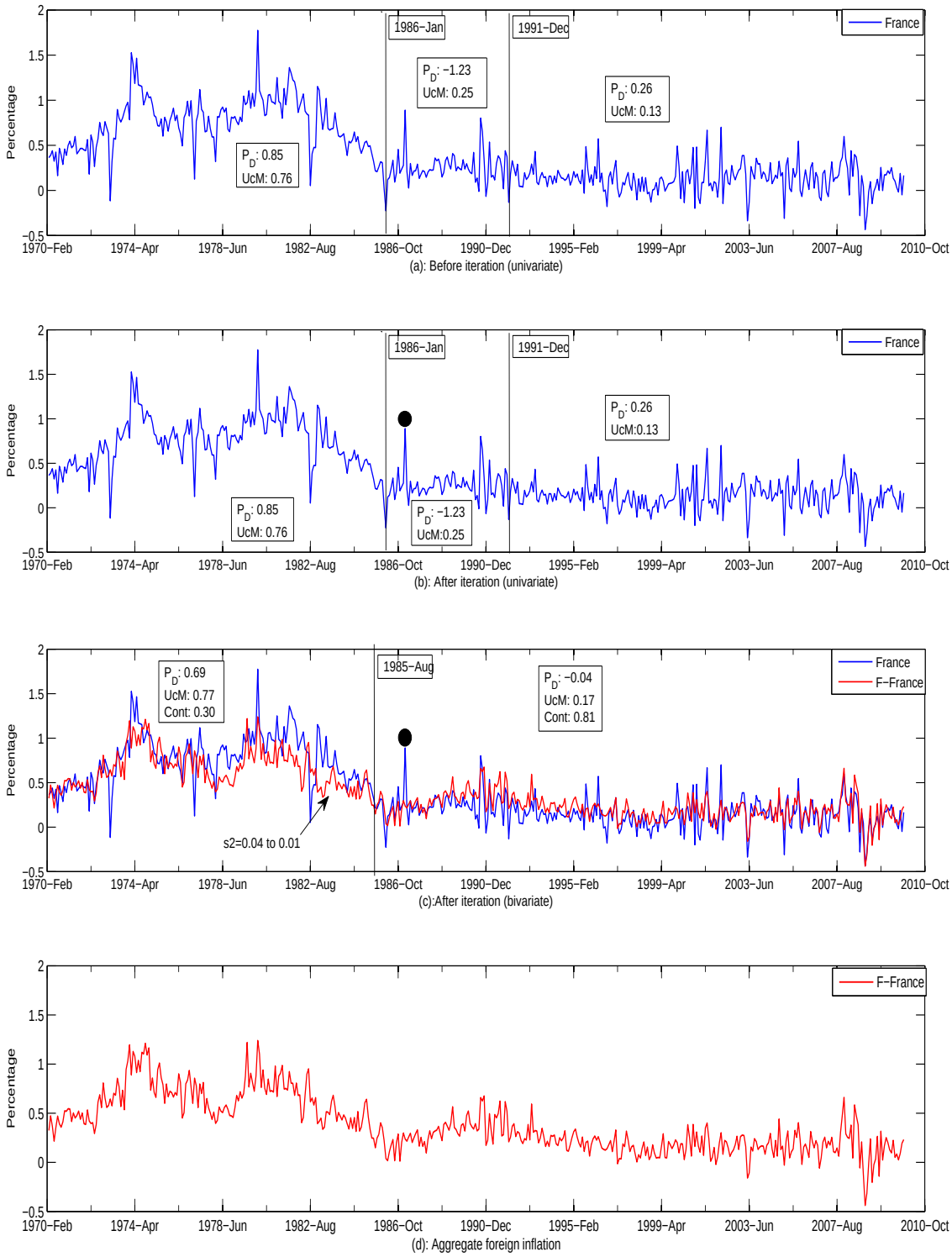


Figure 1.1.7: Inflation Dynamics: Germany

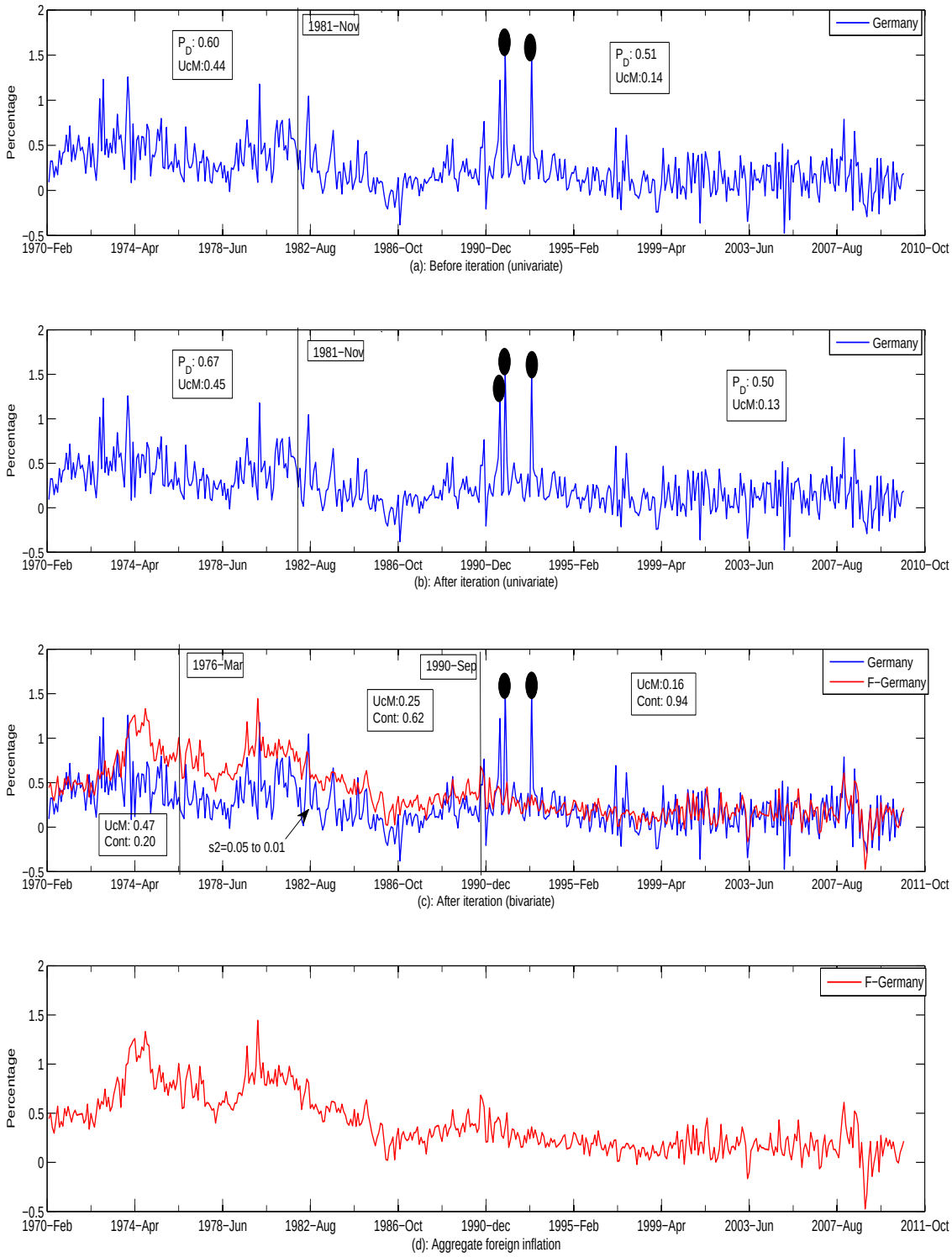


Figure 1.1.8: Inflation Dynamics: Greece

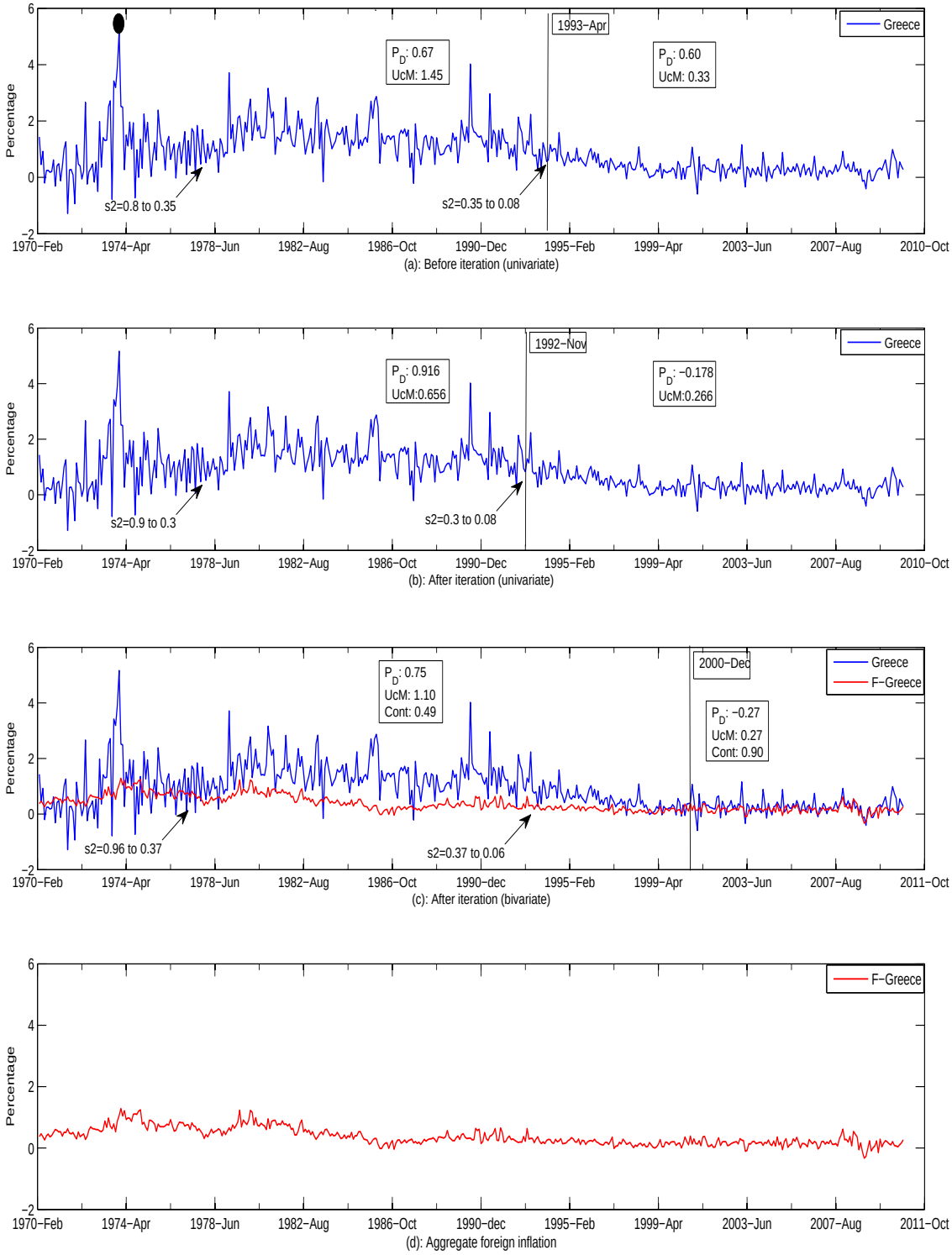


Figure 1.1.9: Inflation Dynamics: Italy

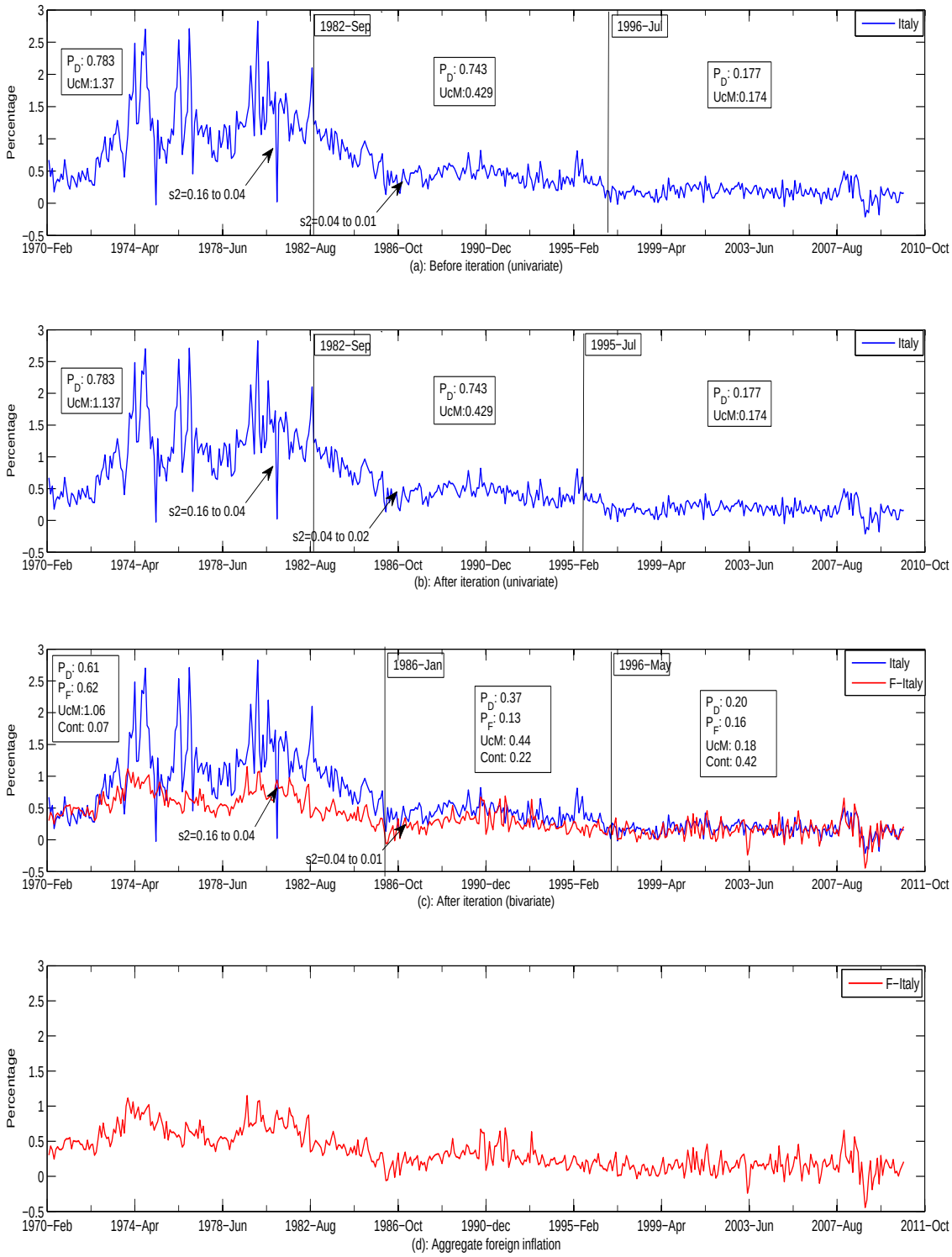
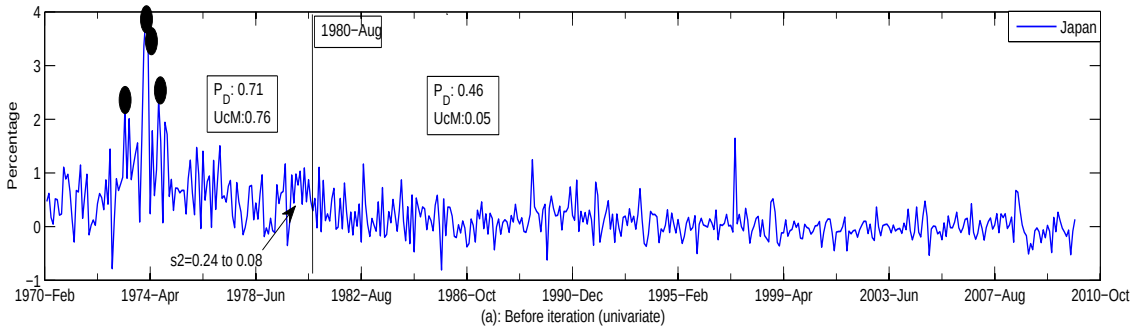
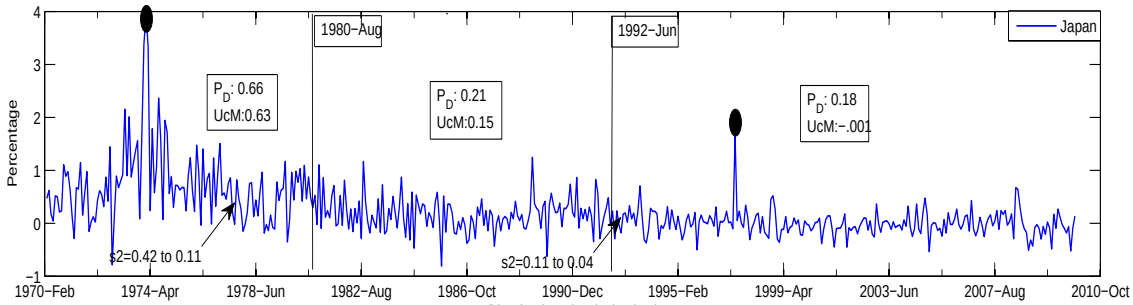


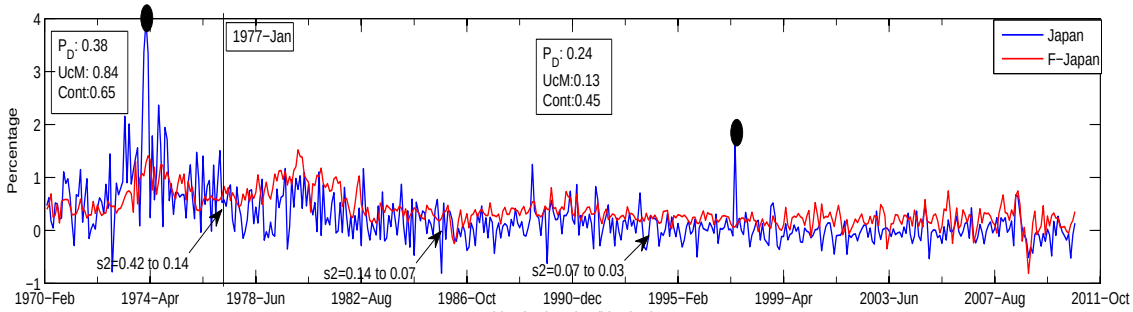
Figure 1.1.10: Inflation Dynamics: Japan



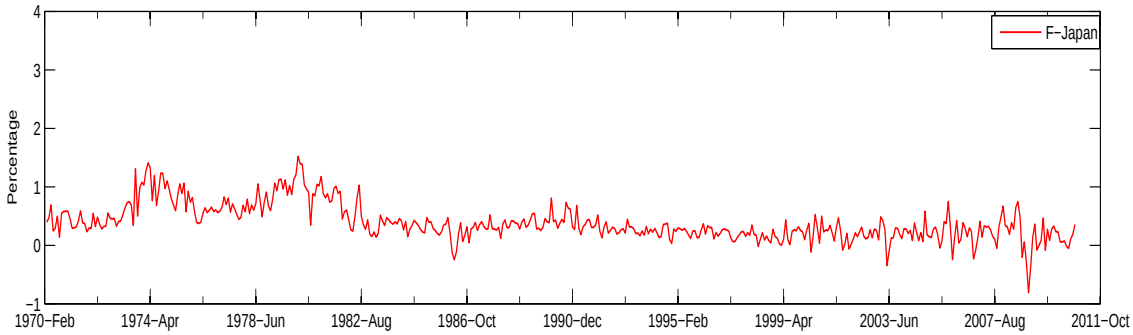
(a): Before iteration (univariate)



(b): After iteration (univariate)



(c): After iteration (bivariate)



(d): Aggregate foreign inflation

Figure 1.1.11: Inflation Dynamics: Korea

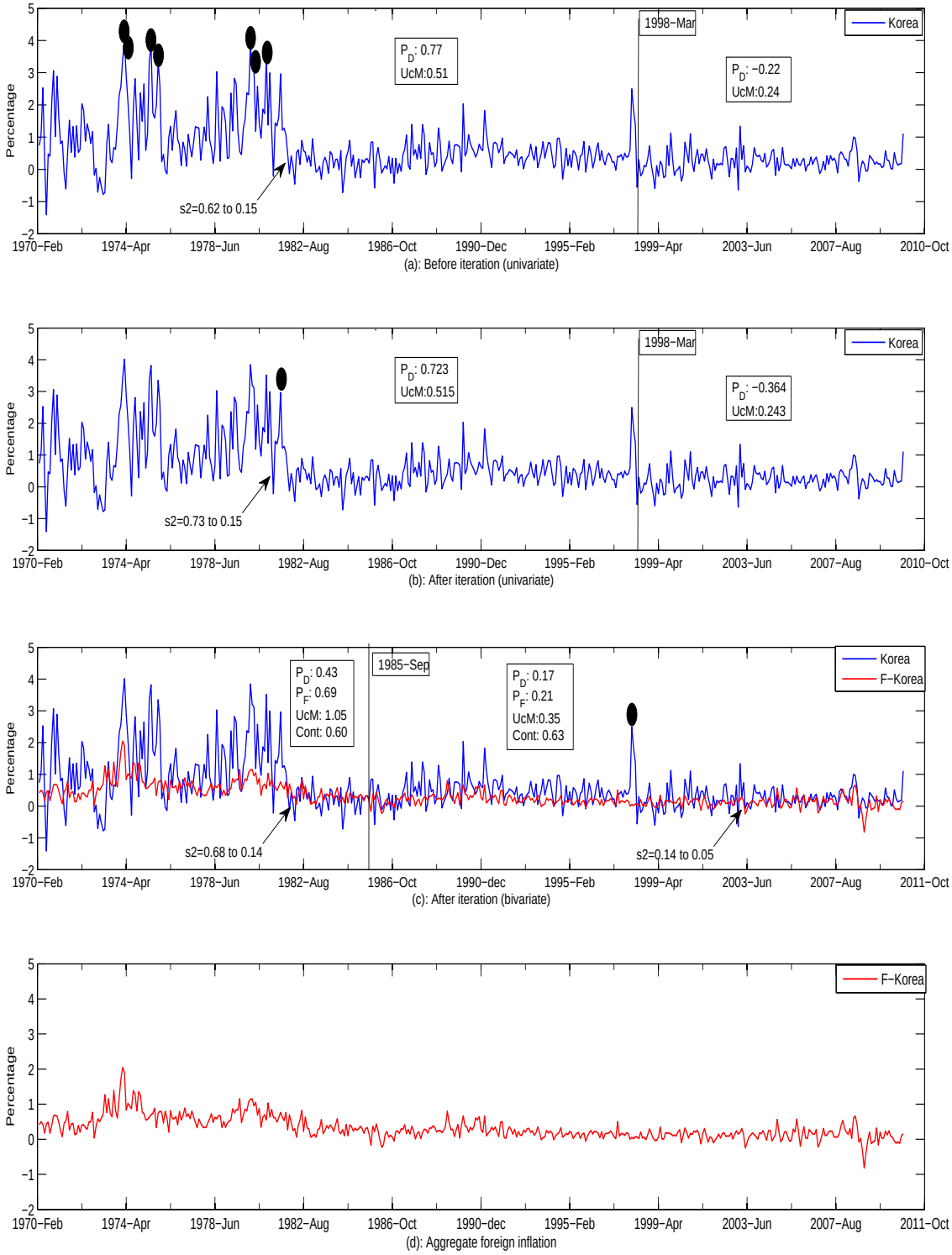


Figure 1.1.12: Inflation Dynamics: Netherlands

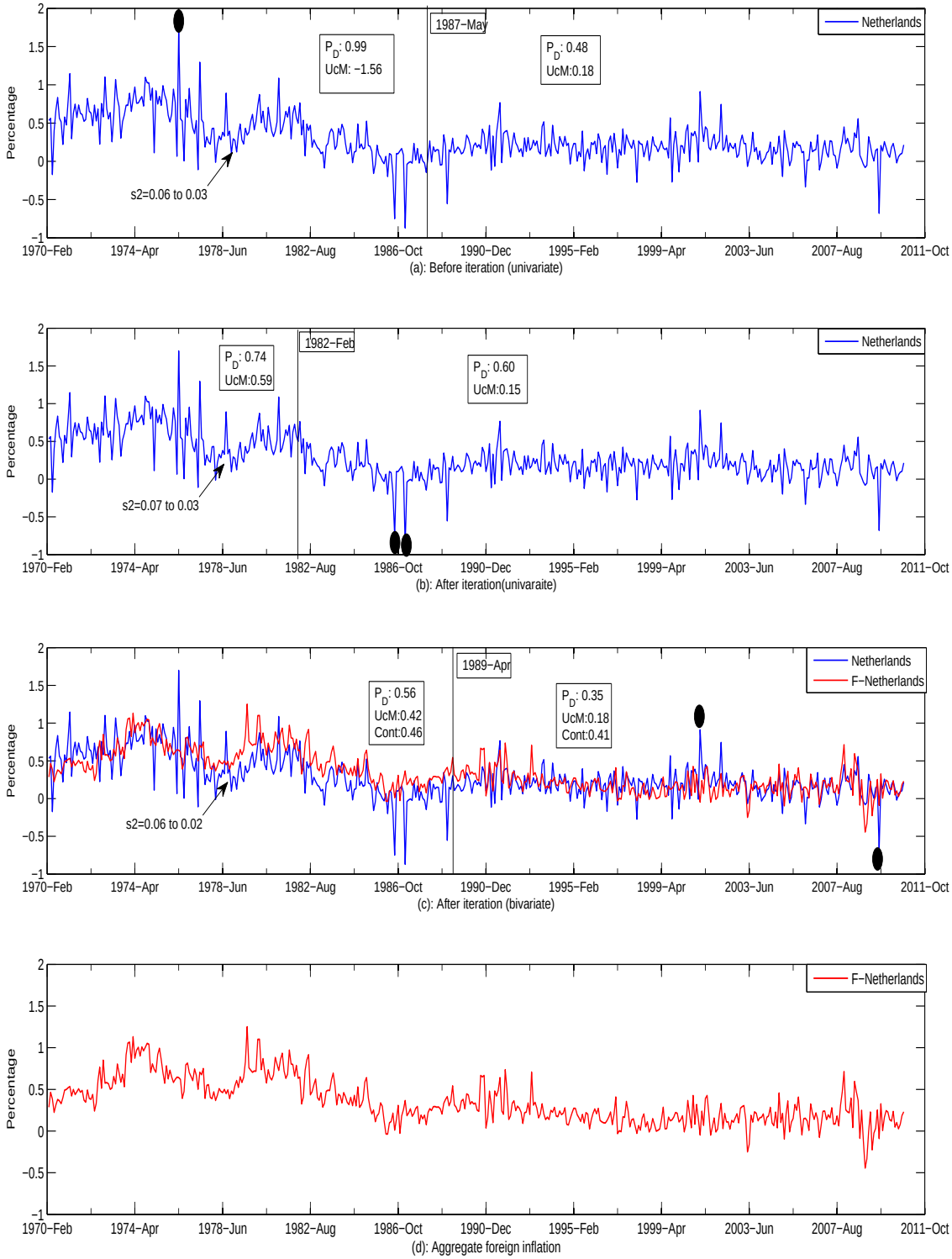


Figure 1.1.13: Inflation Dynamics: Norway

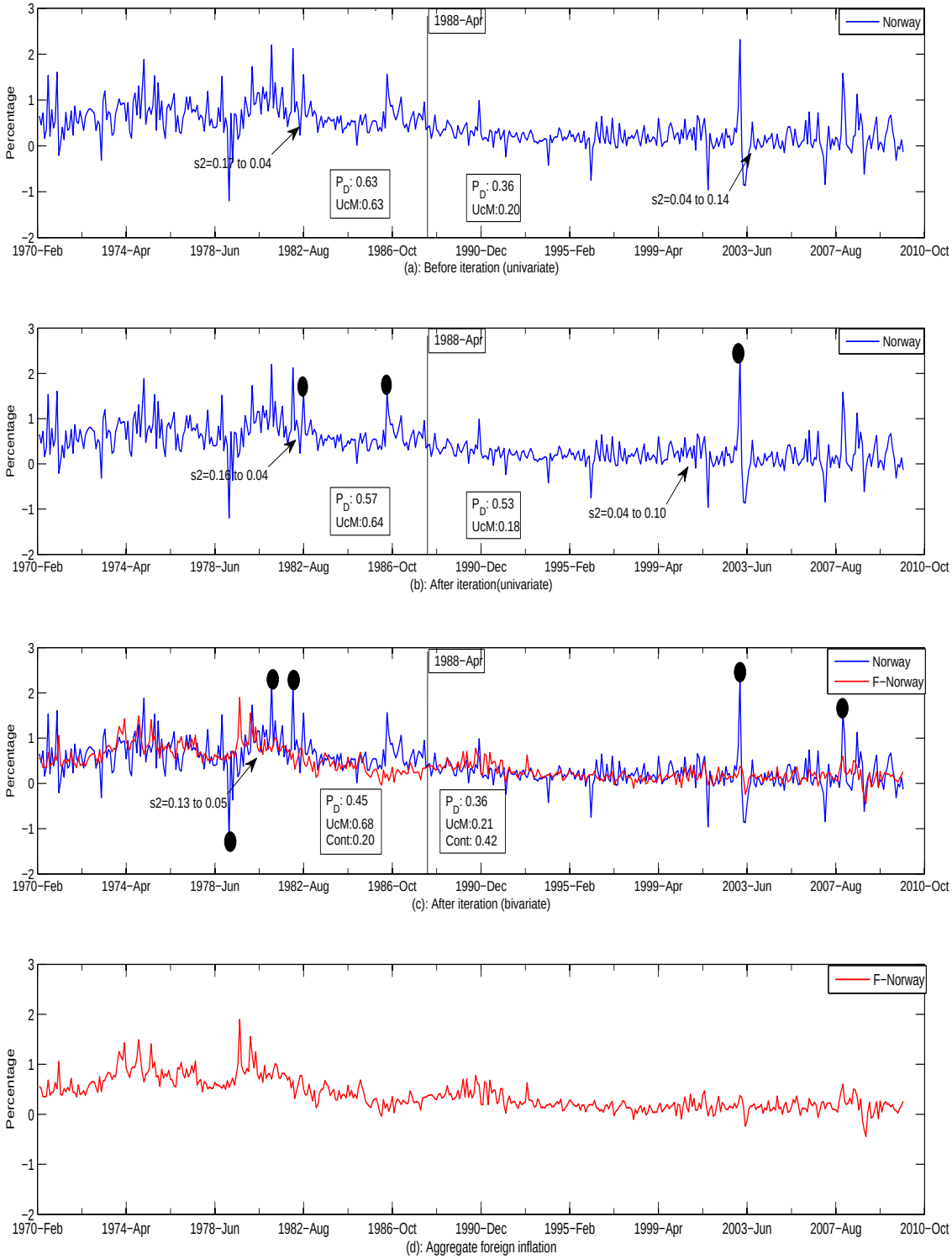


Figure 1.1.14: Inflation Dynamics: Portugal

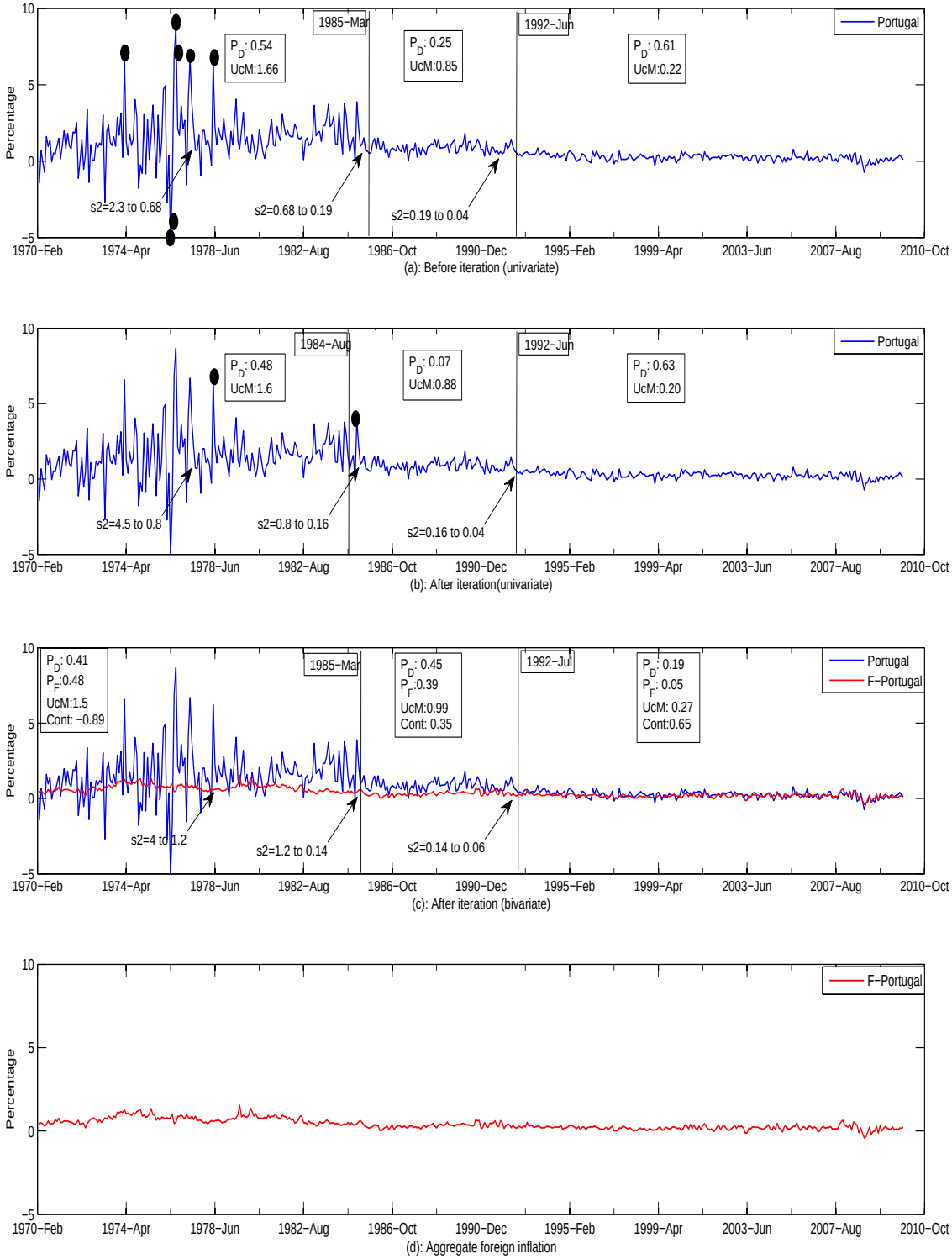


Figure 1.1.15: Inflation Dynamics: Spain

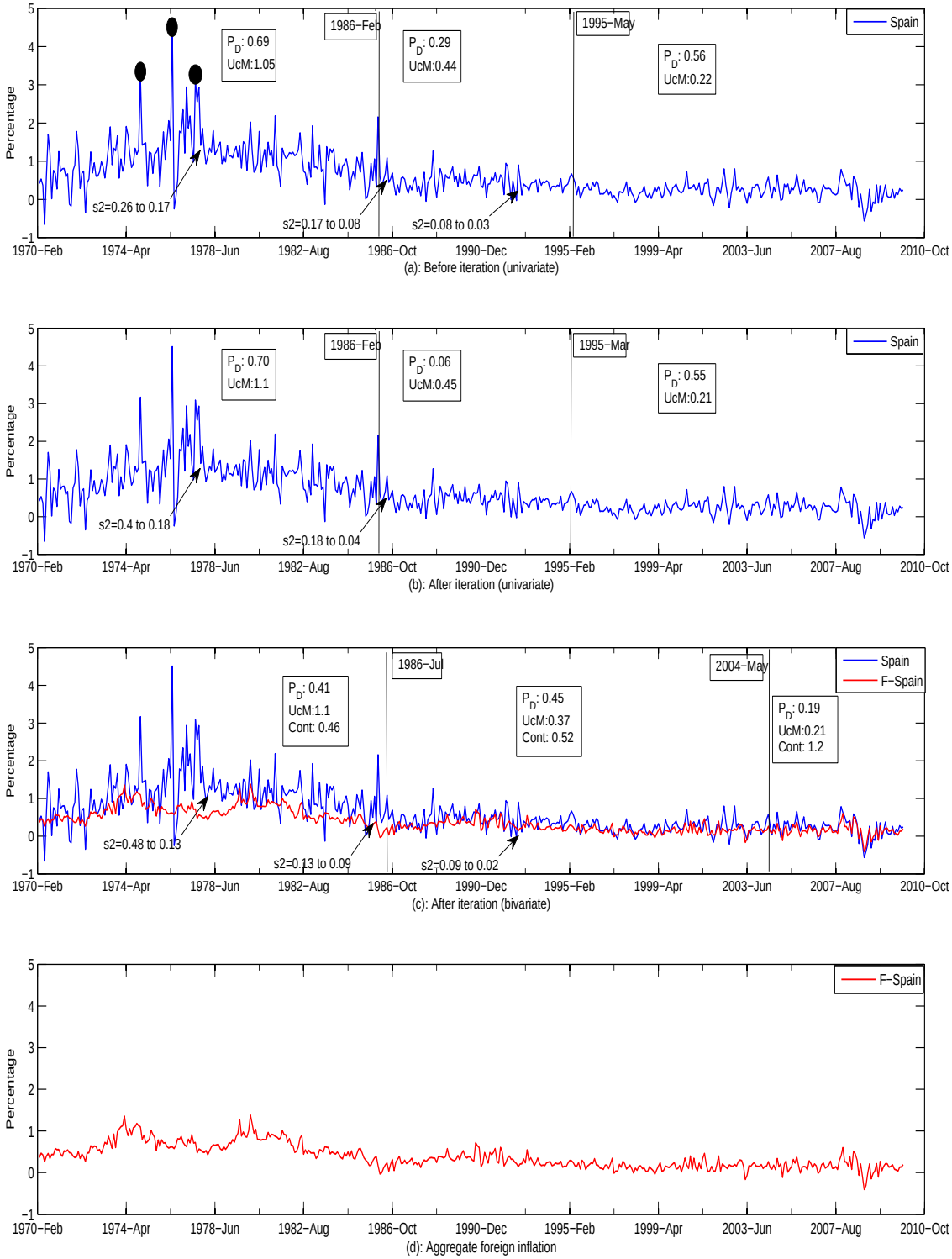


Figure 1.1.16: Inflation Dynamics: Sweden

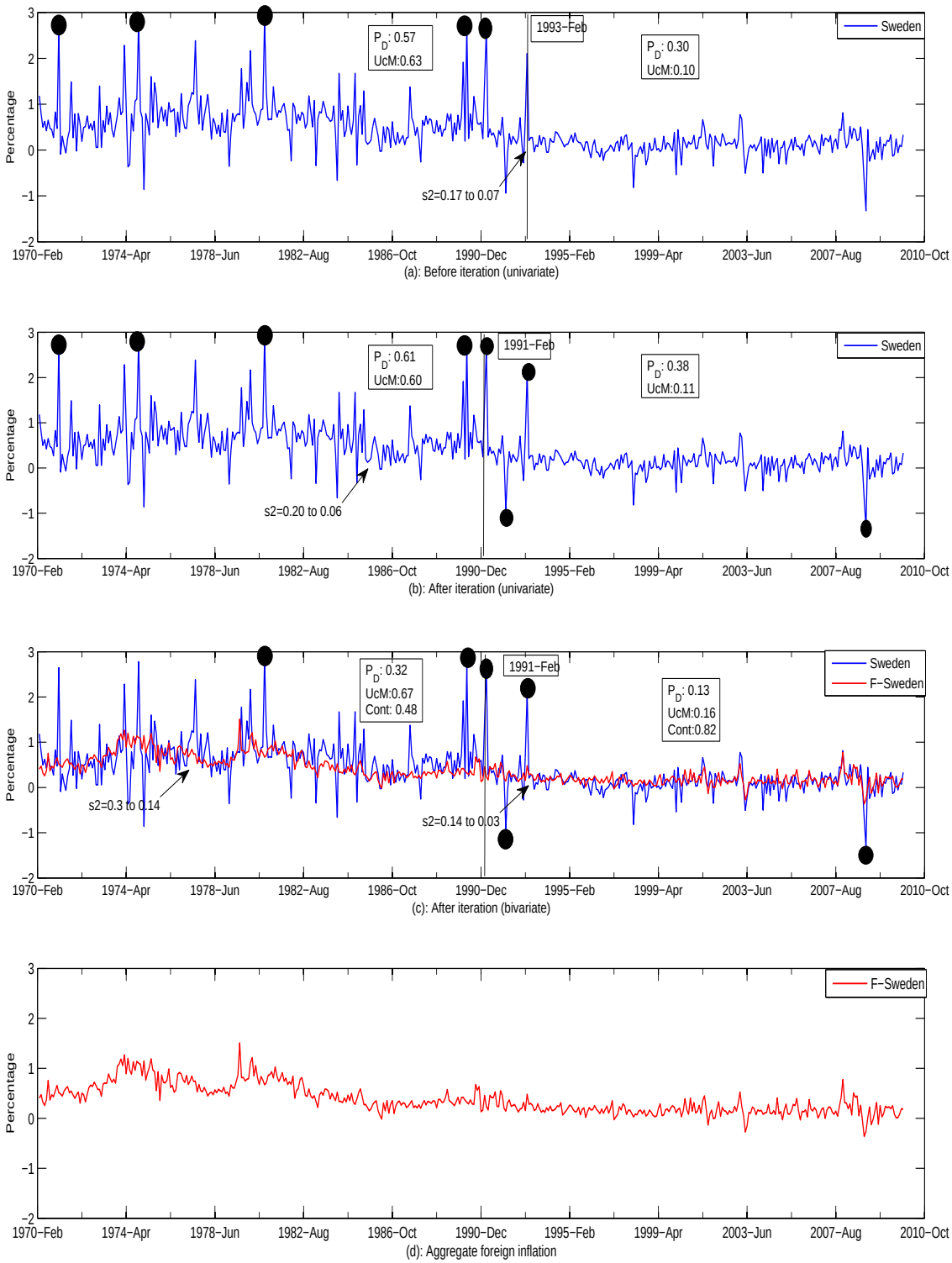


Figure 1.1.17: Inflation Dynamics: Switzerland

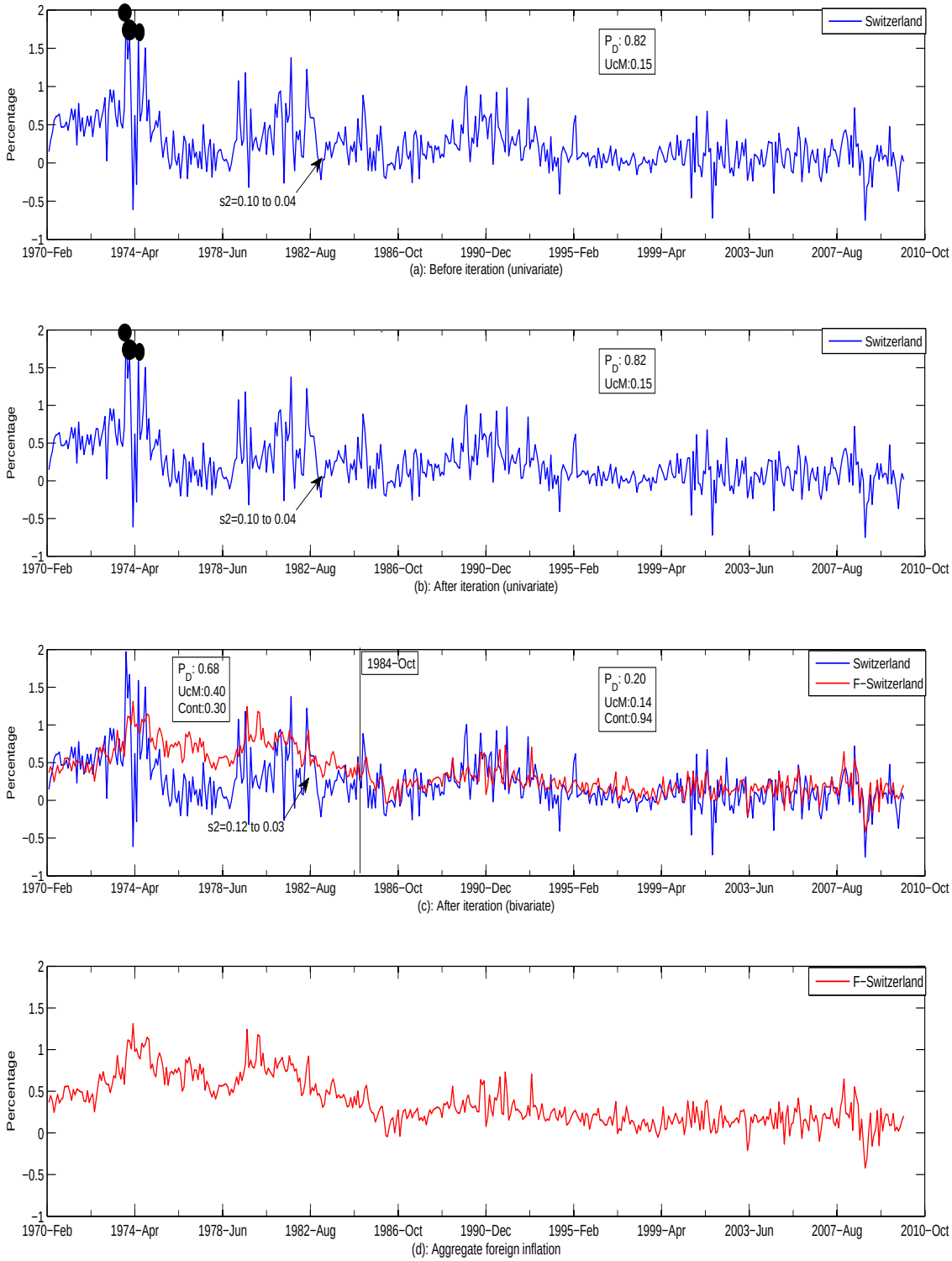


Figure 1.1.18: Inflation Dynamics: United Kingdom

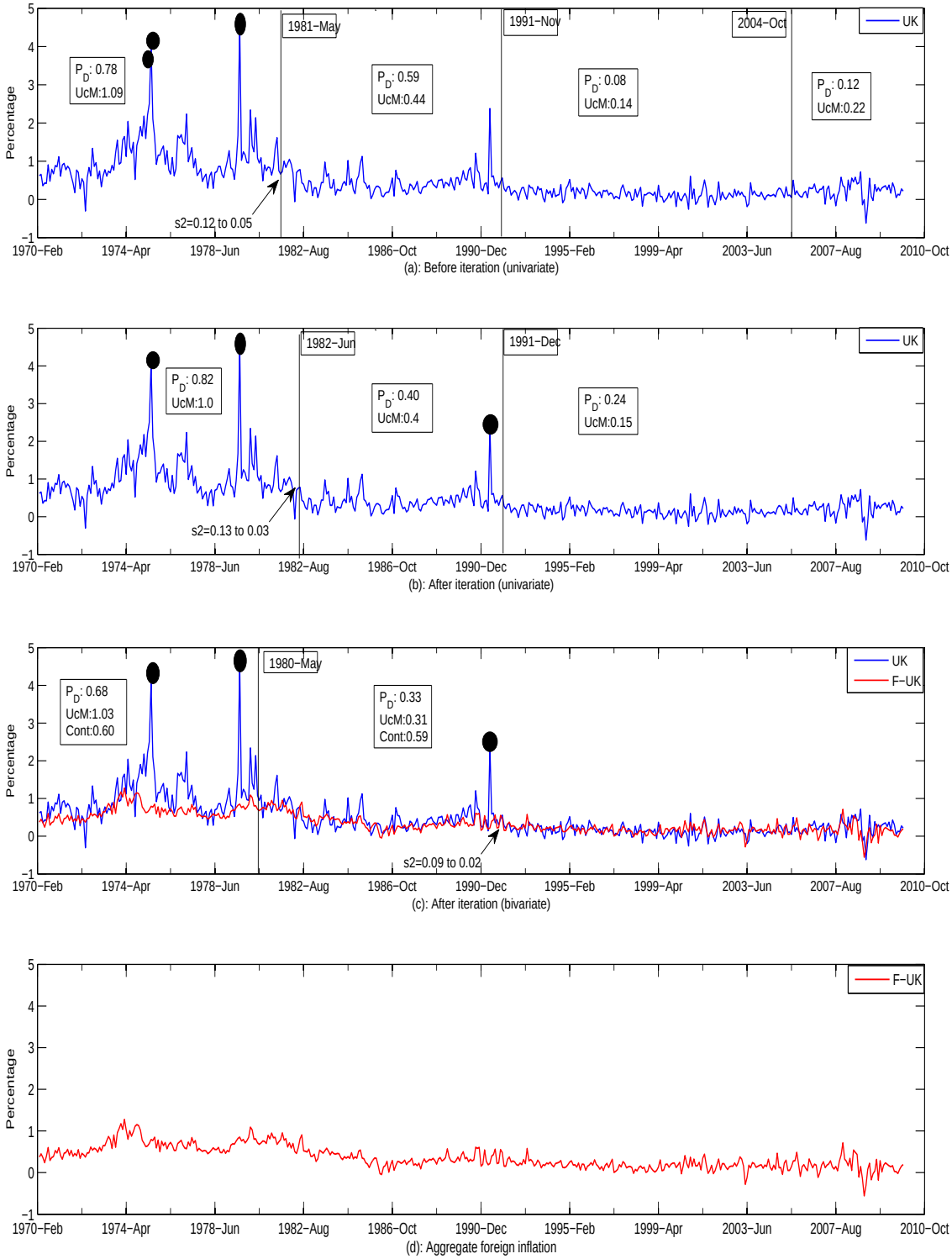
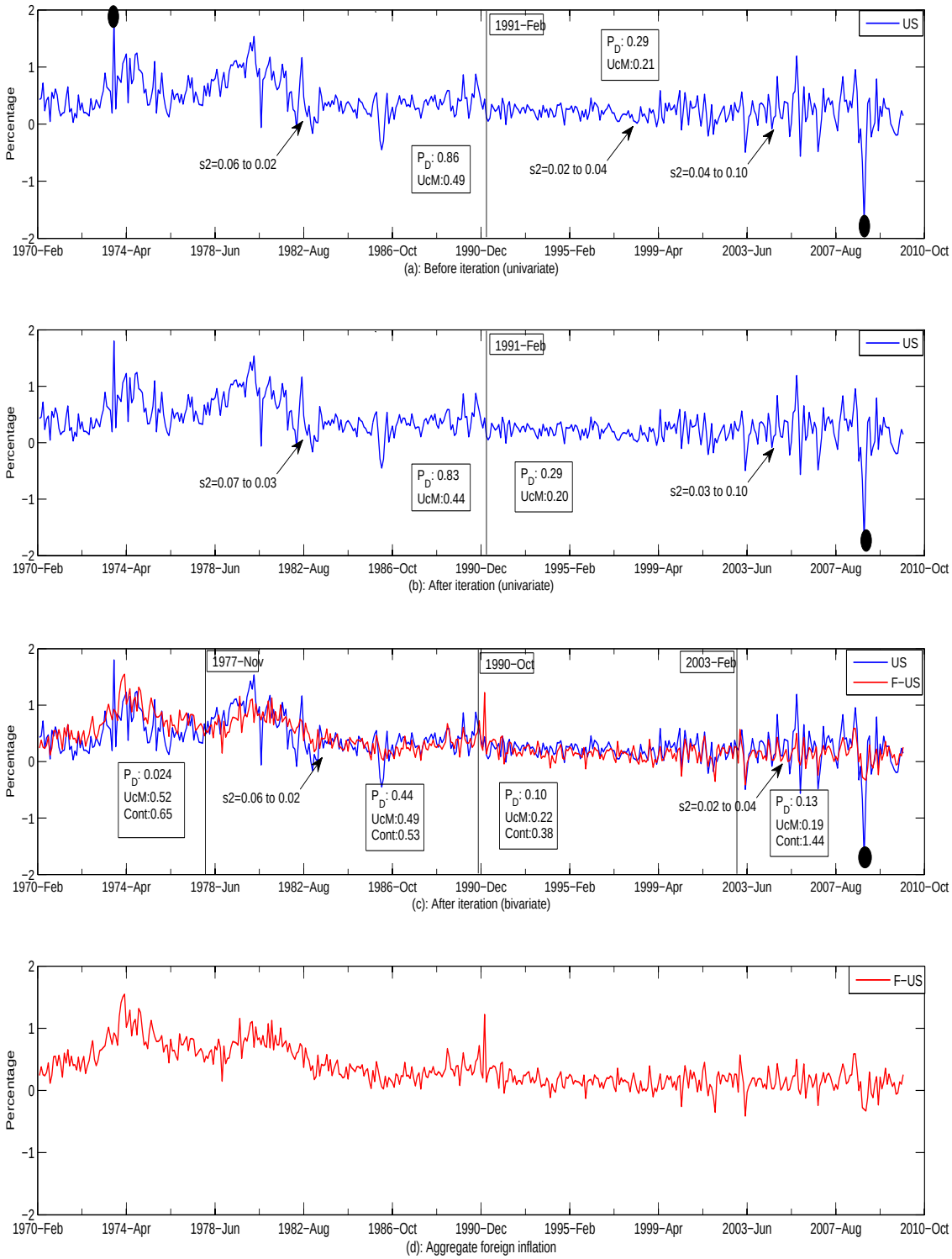
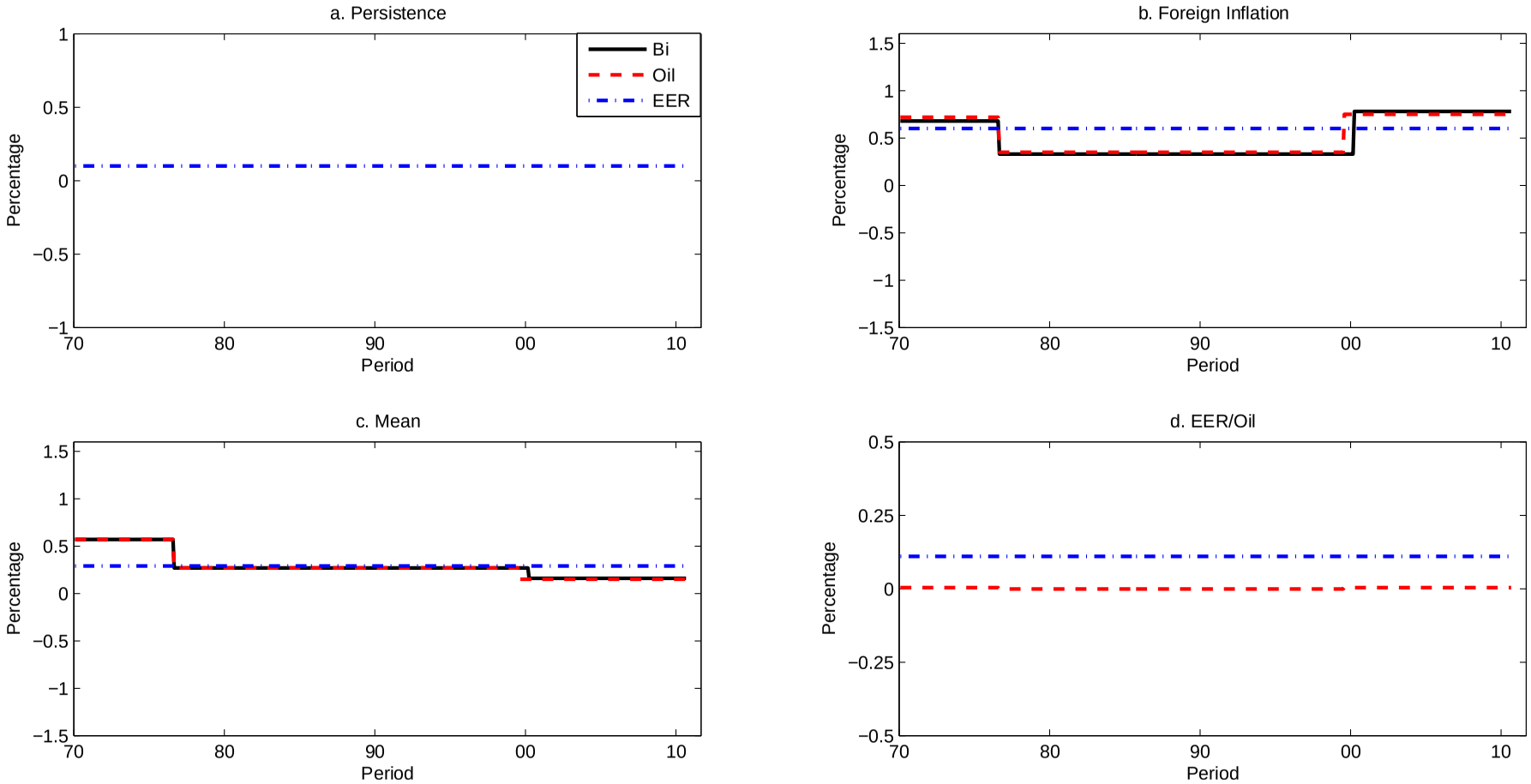


Figure 1.1.19: Inflation Dynamics: United States



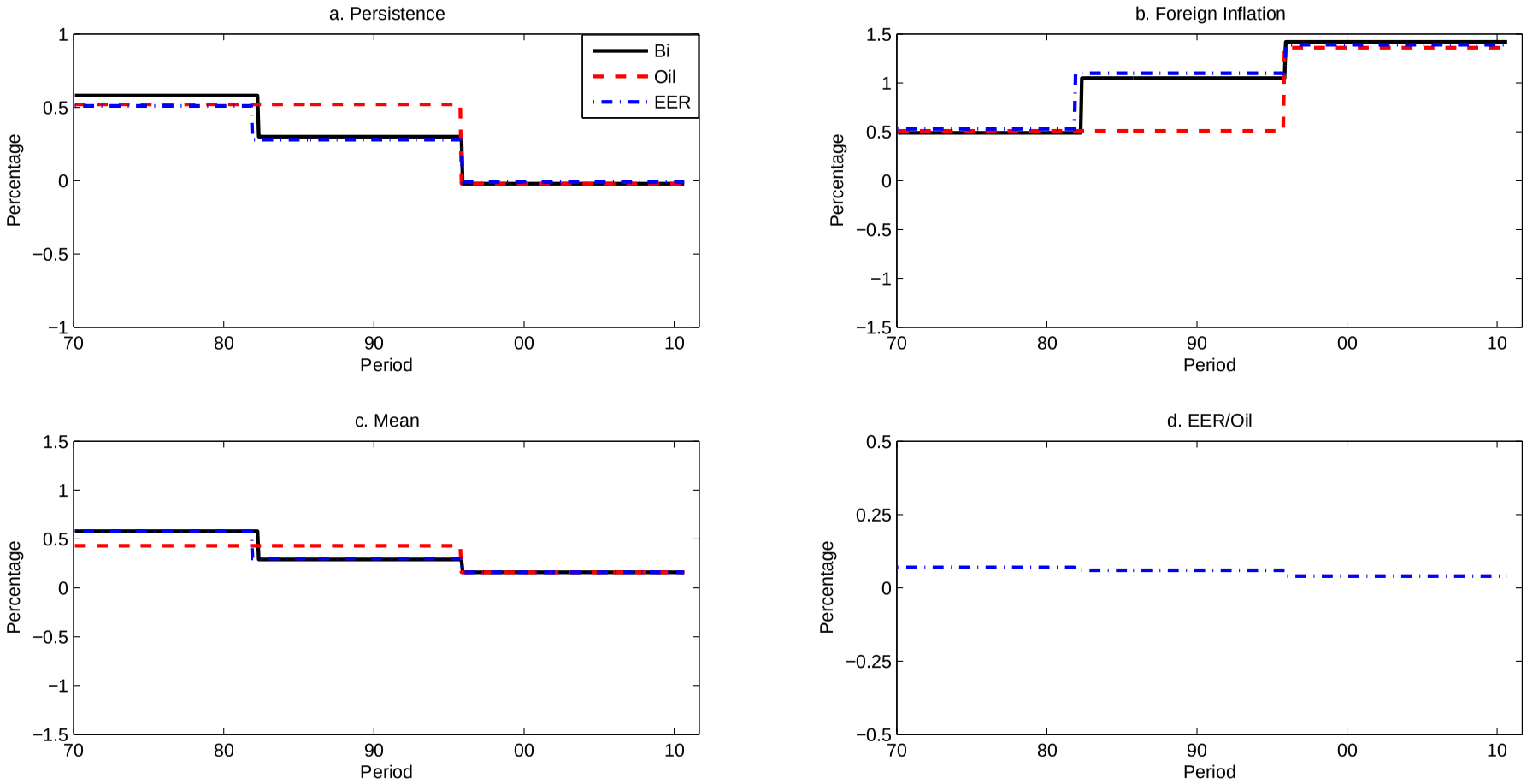
Figures 1.2: Coefficient changes

Figure 1.2.1: Coefficient changes: Austria



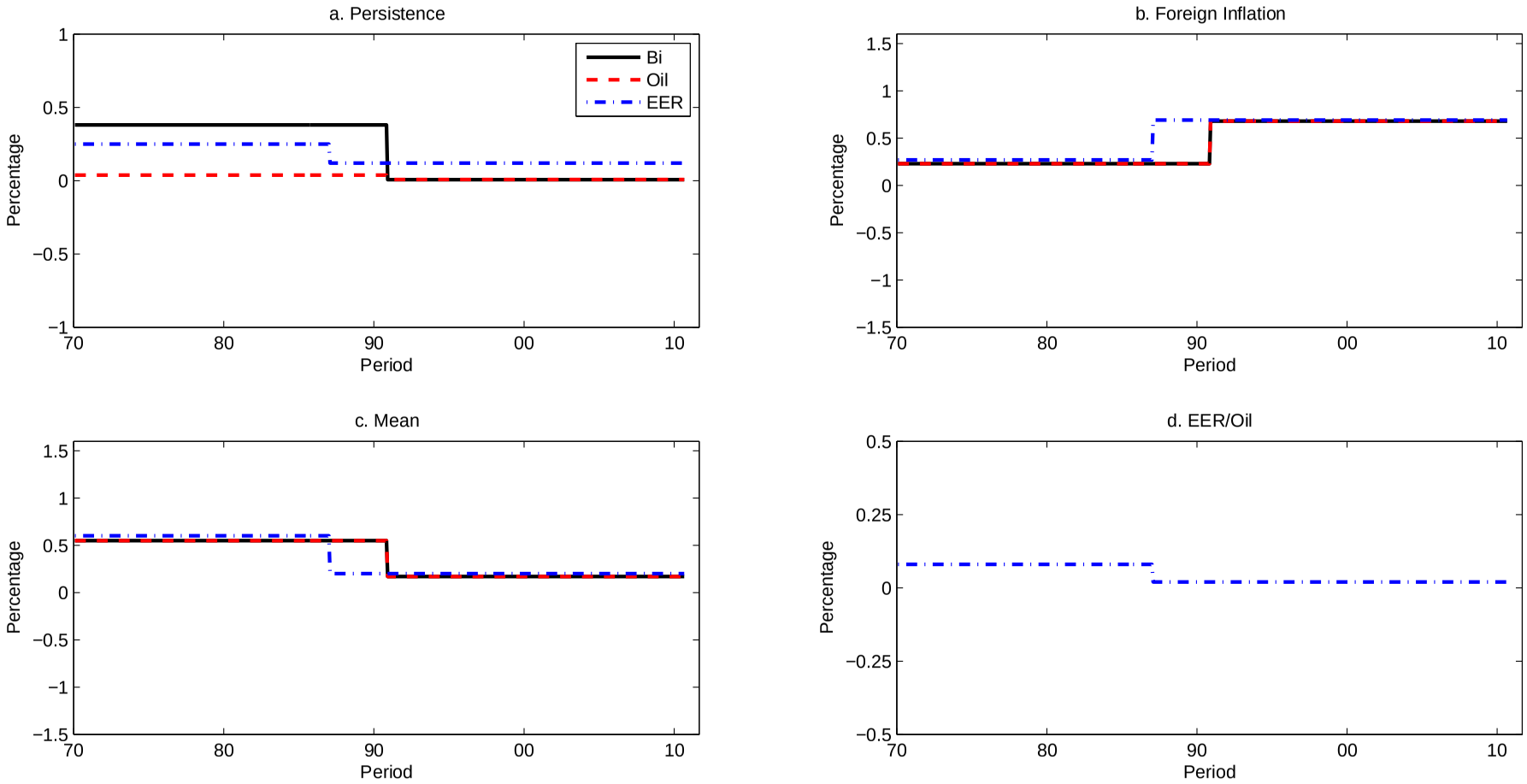
Notes: Subplot (a) depicts changes in persistence that are estimated using the bivariate model (in black line), the model with oil price inflation (in red line) and the model with EER (in blue line). Similarly, subplots (b) and (c) respectively, show changes in the contemporaneous foreign inflation coefficients and the subsample mean. Subplot (d) shows sum of estimated coefficients corresponding to the contemporaneous and lagged oil price inflation (in red line) and EER (in blue line). A missing line either in subplot (a) or (d) indicates the absence of the corresponding lags (and contemporaneous variable) in the model.

Figure 1.2.2: Coefficient changes: Belgium



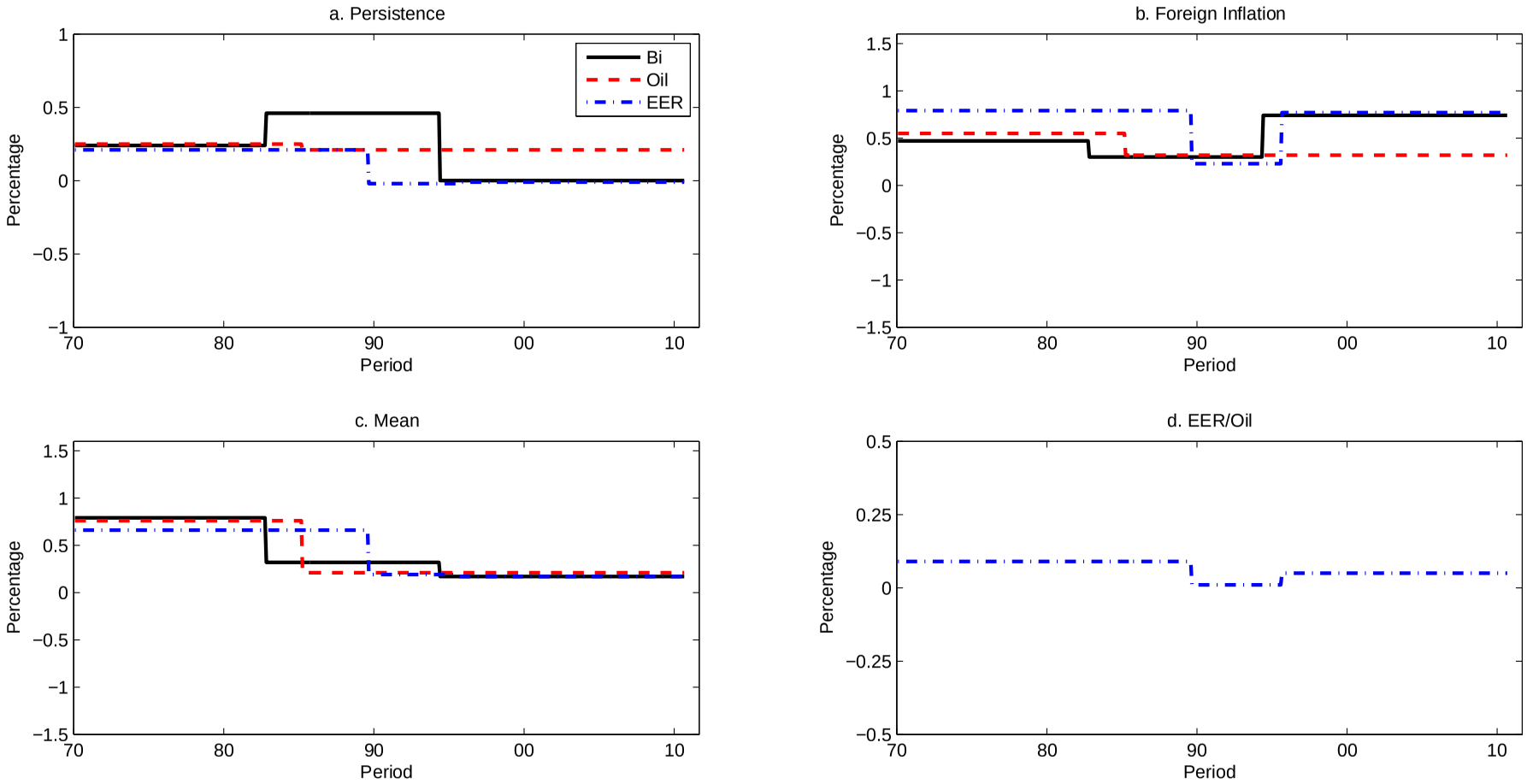
Notes: Same as figure 1.2.1.

Figure 1.2.3: Coefficient changes: Canada



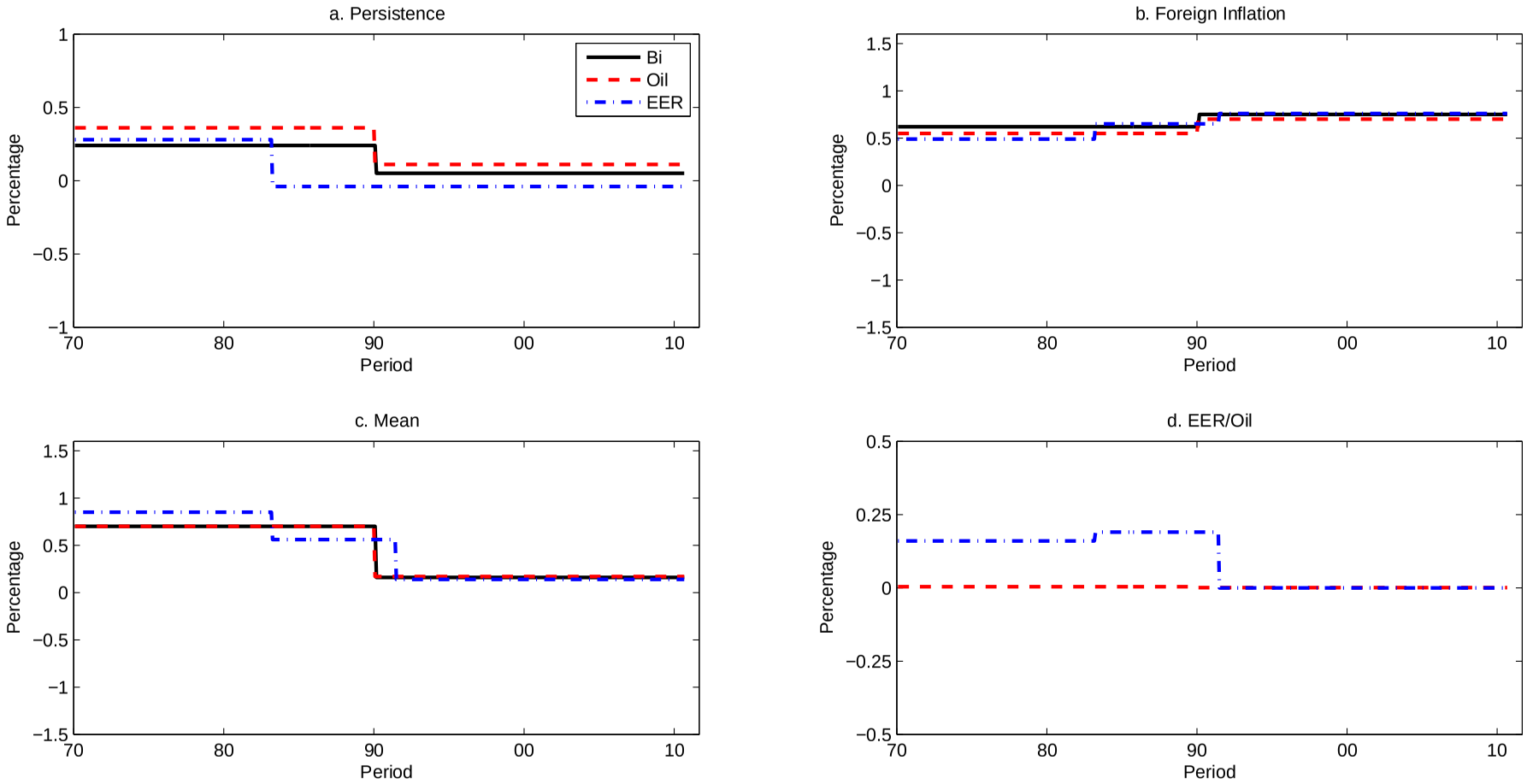
Notes: Same as figure 1.2.1.

Figure 1.2.4: Coefficient changes: Denmark



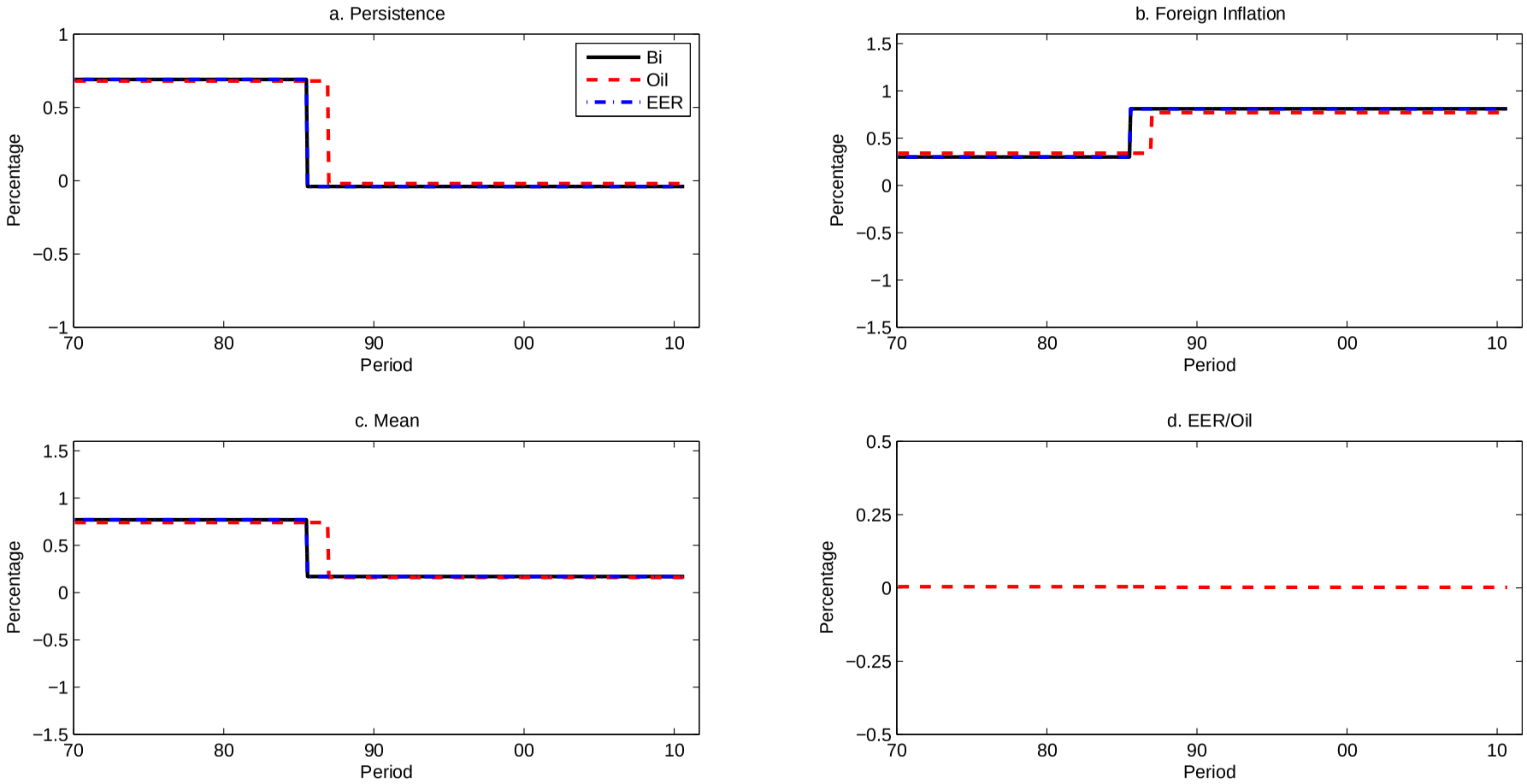
Notes: Same as figure 1.2.1.

Figure 1.2.5: Coefficient changes: Finland



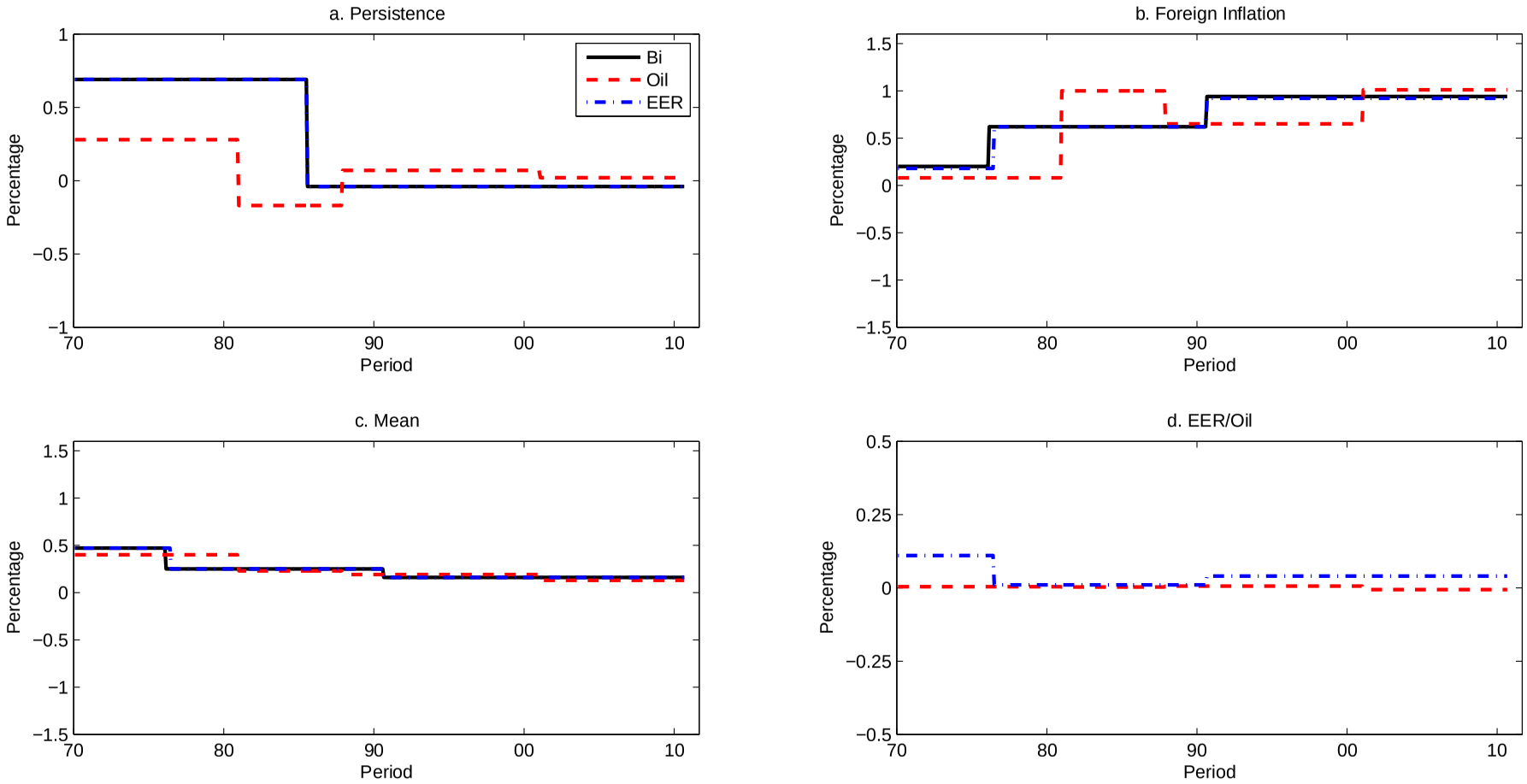
Notes: Same as figure 1.2.1.

Figure 1.2.6: Coefficient changes: France



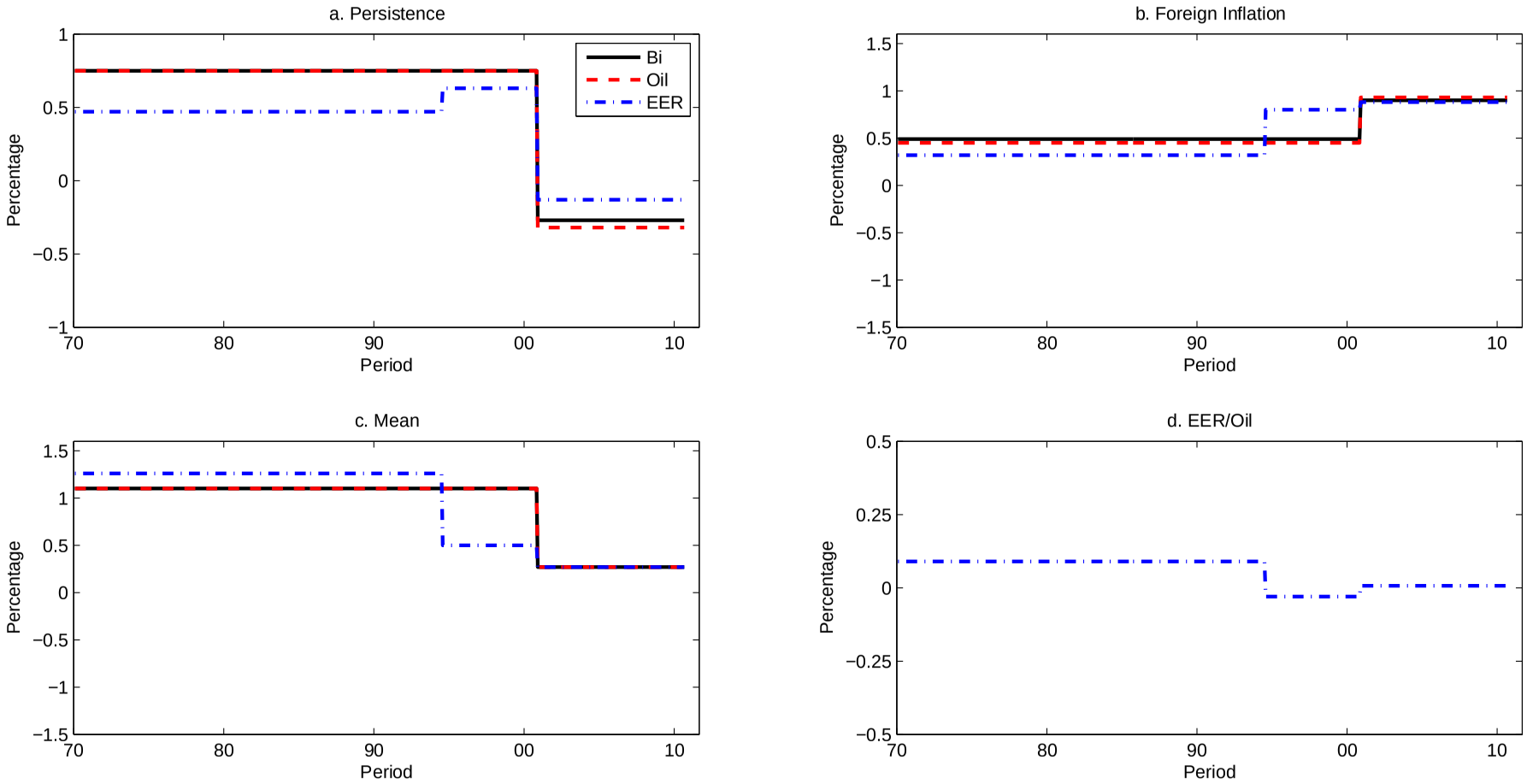
Notes: Same as figure 1.2.1.

Figure 1.2.7: Coefficient changes: Germany



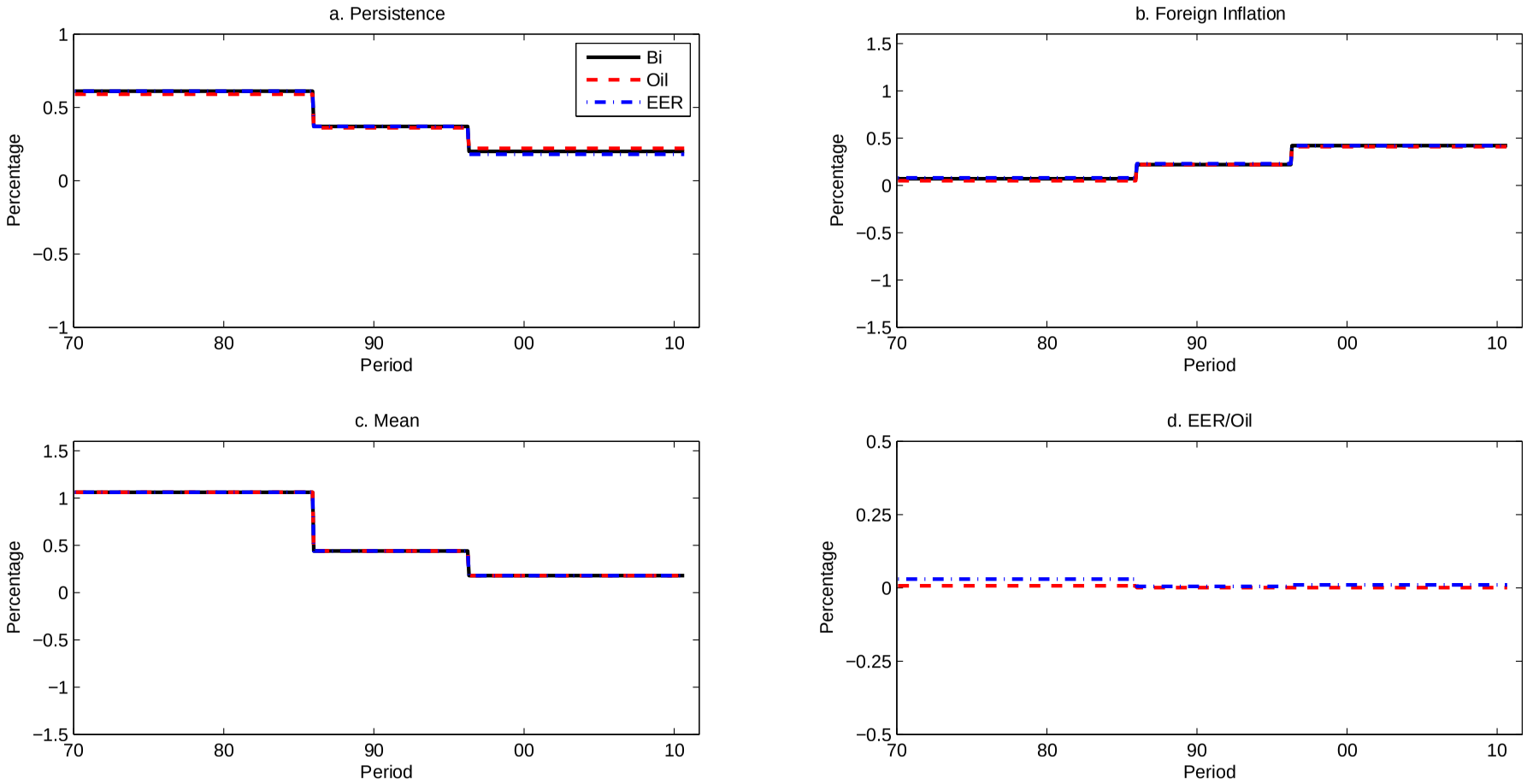
Notes: Same as figure 1.2.1.

Figure 1.2.8: Coefficient changes: Greece



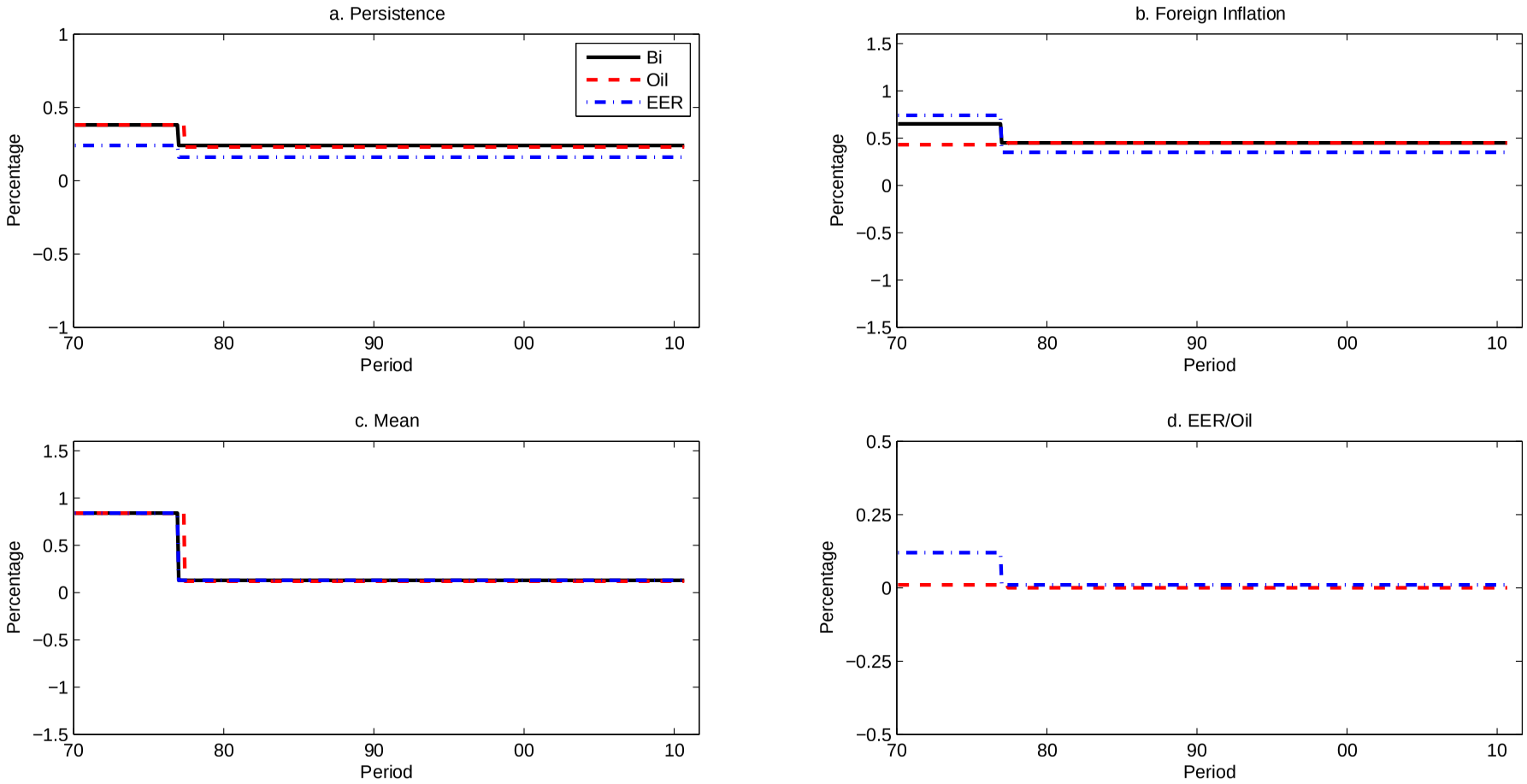
Notes: Same as figure 1.2.1.

Figure 1.2.9: Coefficient changes: Italy



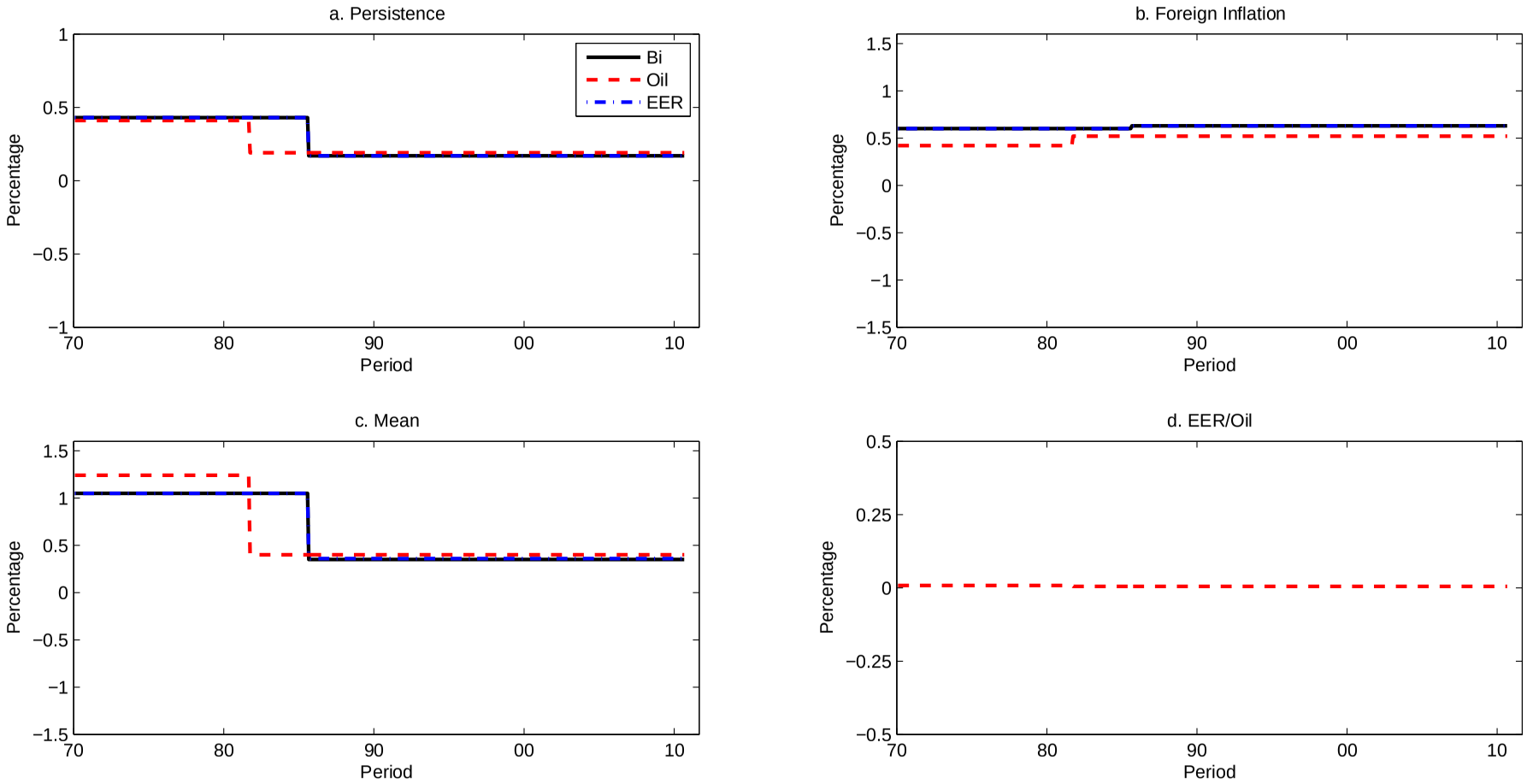
Notes: Same as figure 1.2.1.

Figure 1.2.10: Coefficient changes: Japan



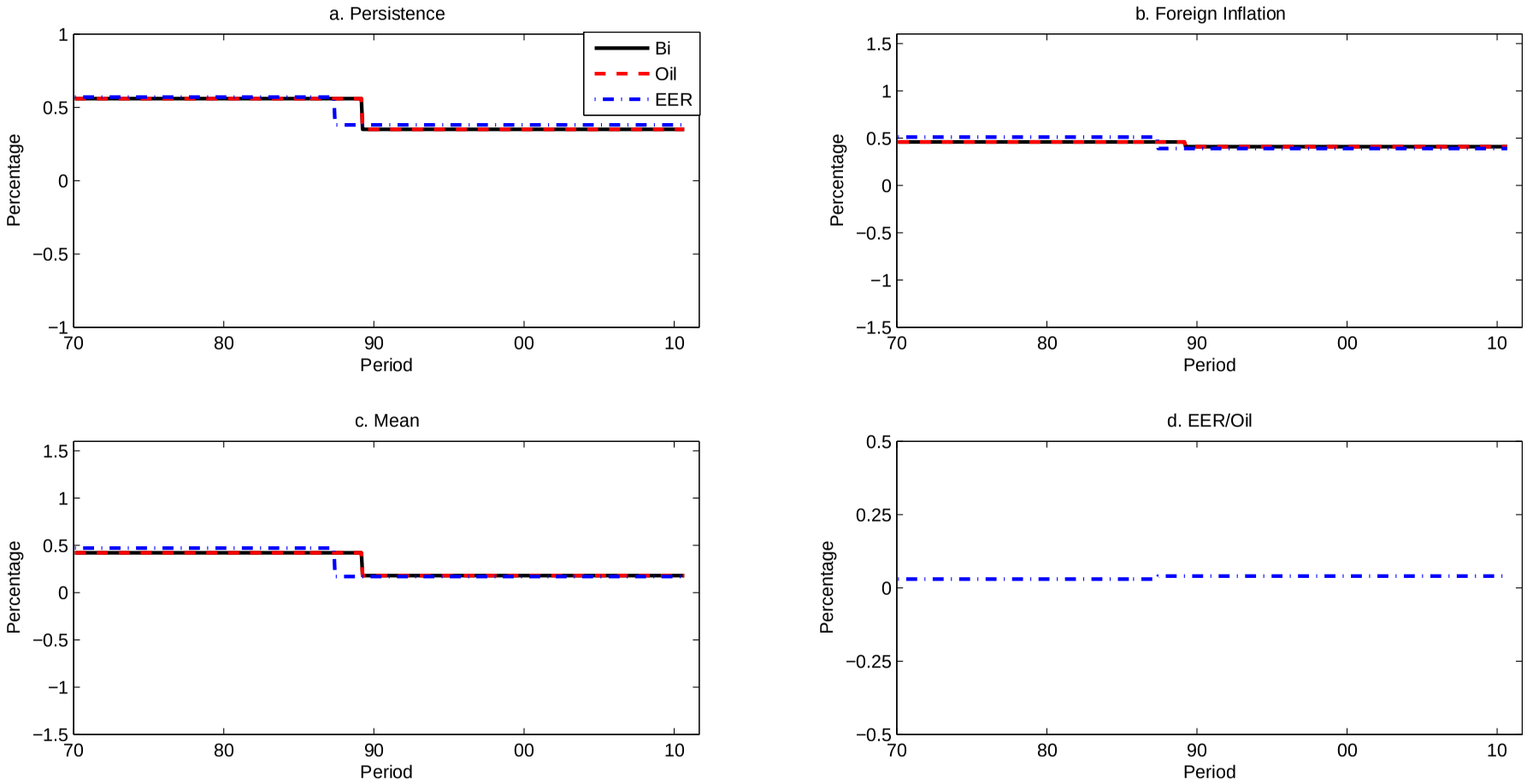
Notes: Same as figure 1.2.1.

Figure 1.2.11: Coefficient changes: Korea



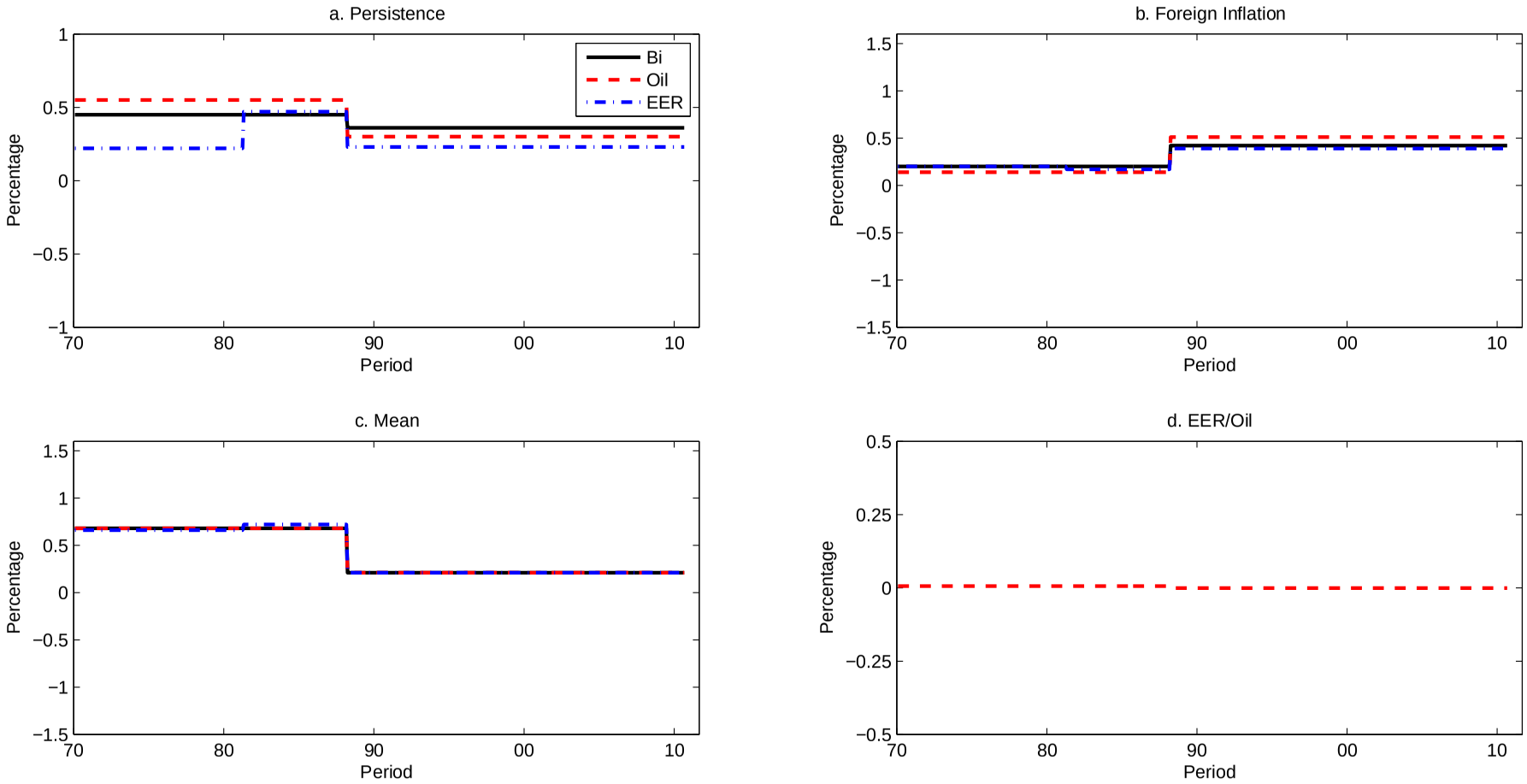
Notes: Same as figure 1.2.1.

Figure 1.2.12: Coefficient changes: Netherlands



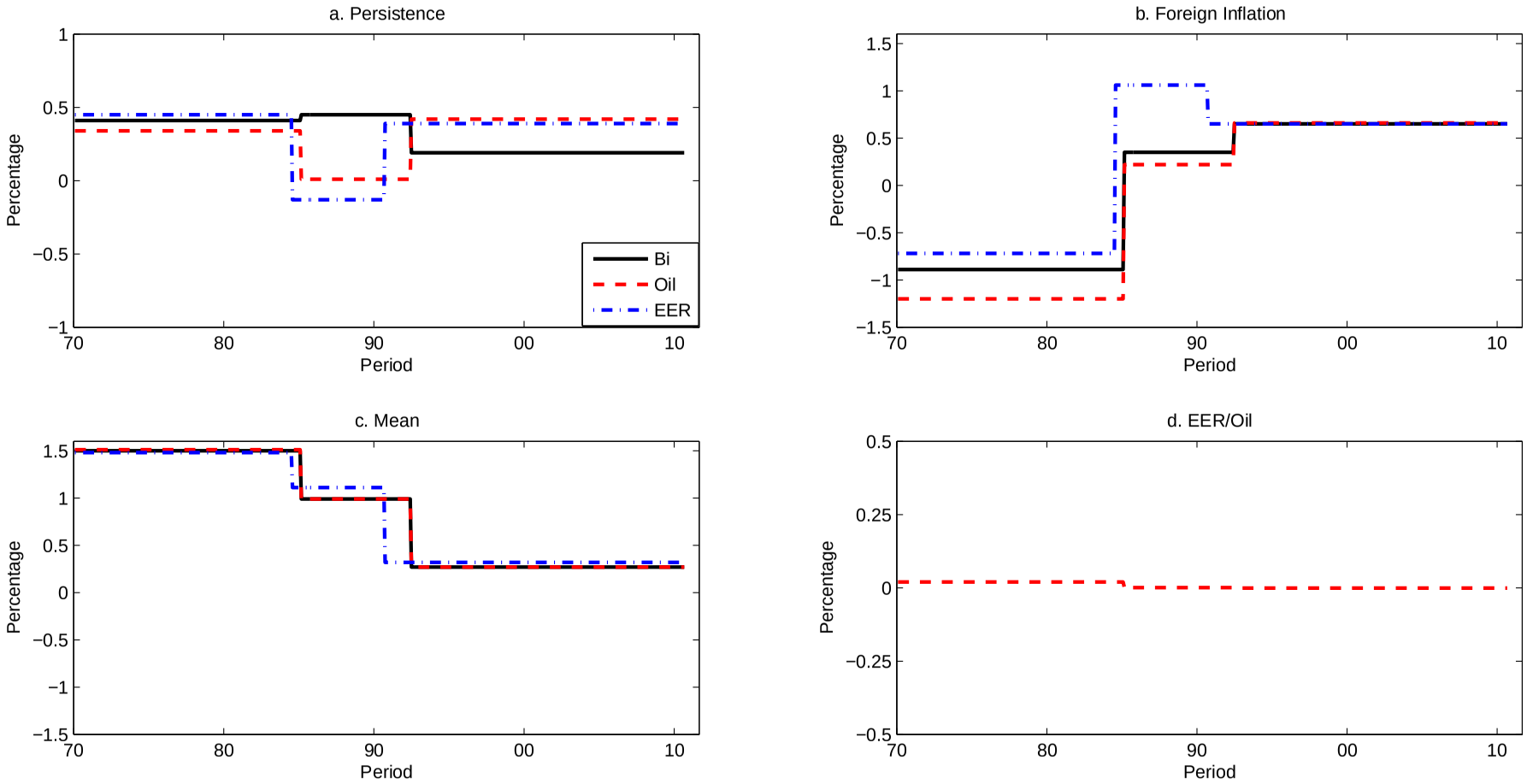
Notes: Same as figure 1.2.1.

Figure 1.2.13: Coefficient changes: Norway



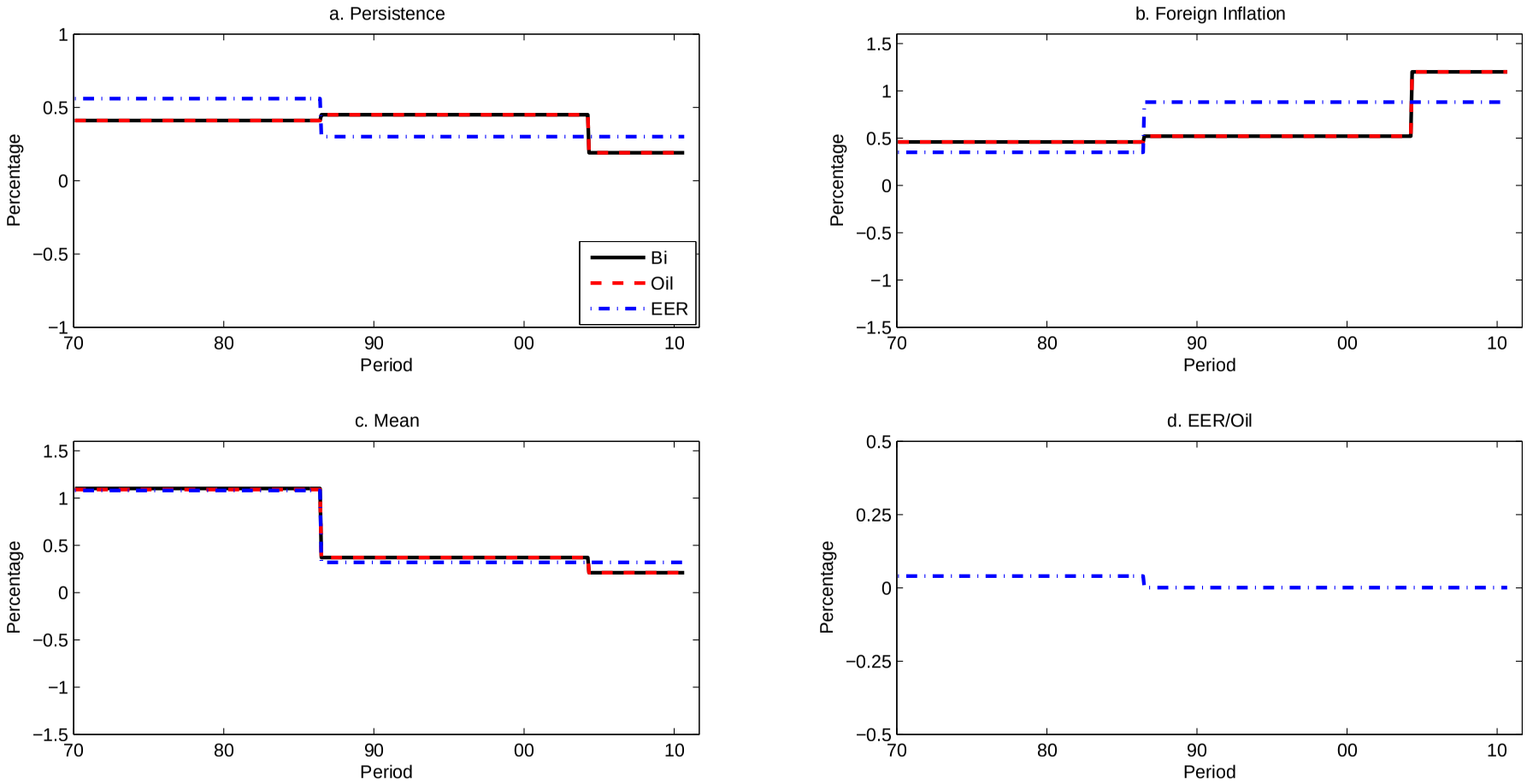
Notes: Same as figure 1.2.1.

Figure 1.2.14: Coefficient changes: Portugal



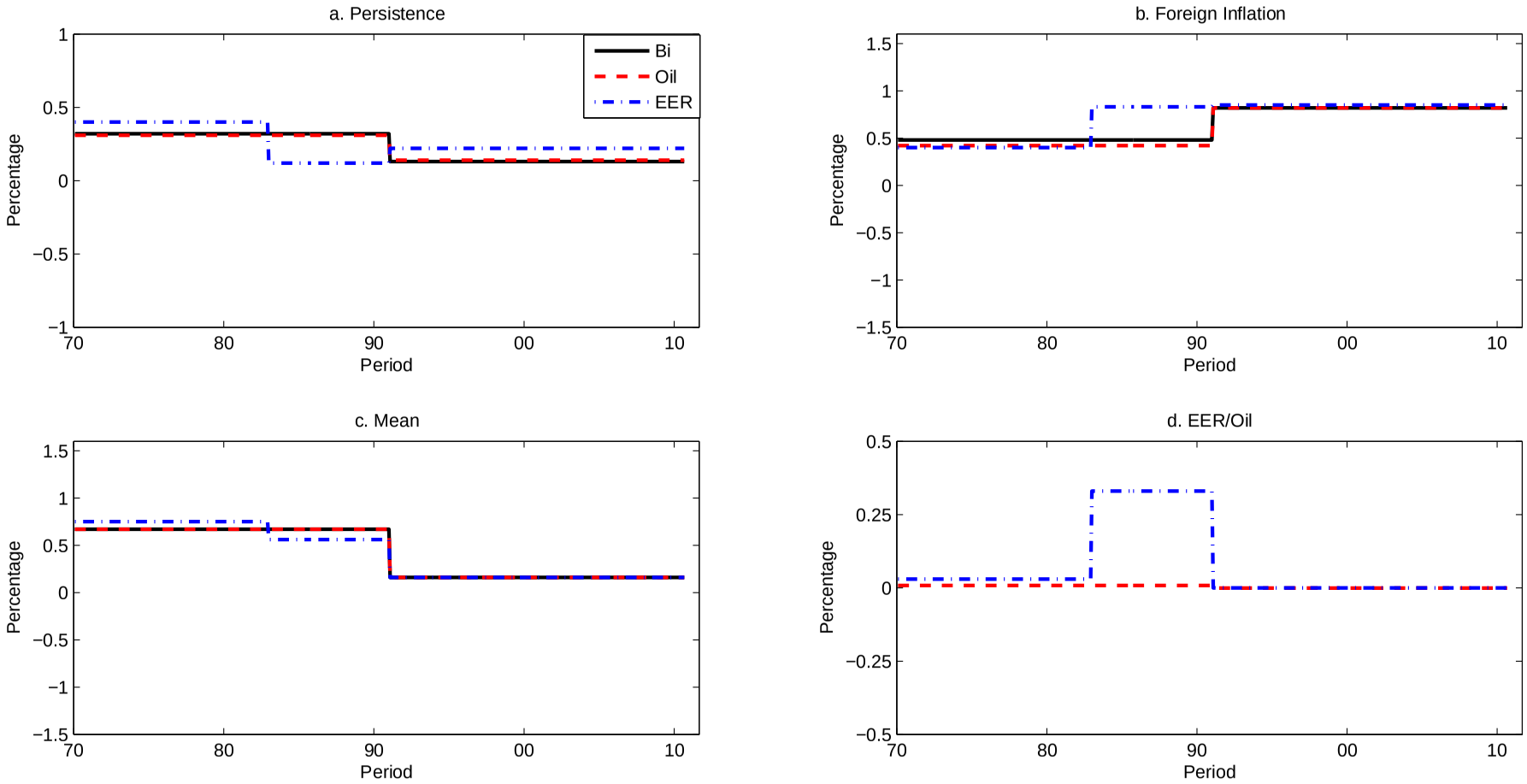
Notes: Same as figure 1.2.1.

Figure 1.2.15: Coefficient changes: Spain



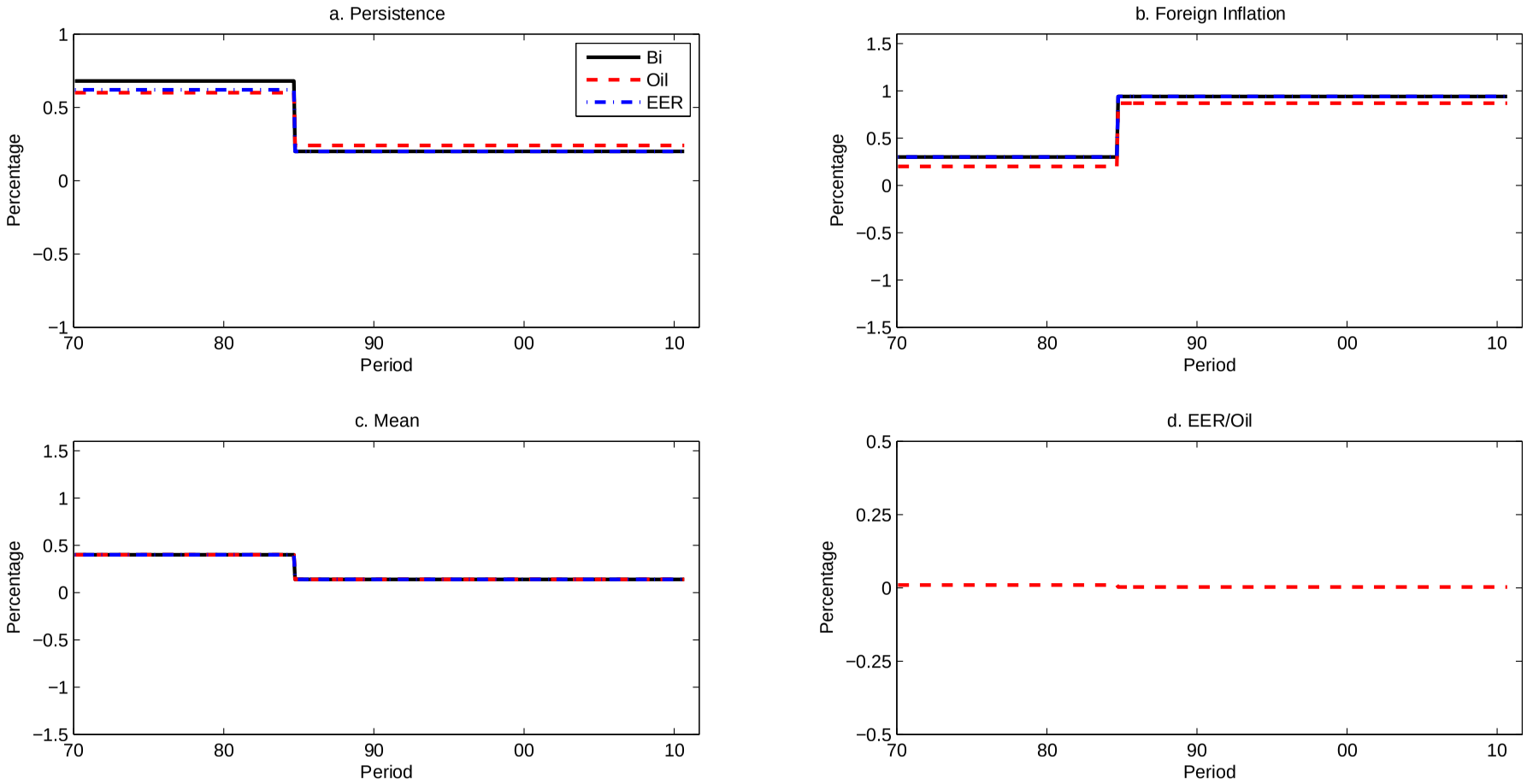
Notes: Same as figure 1.2.1.

Figure 1.2.16: Coefficient changes: Sweden



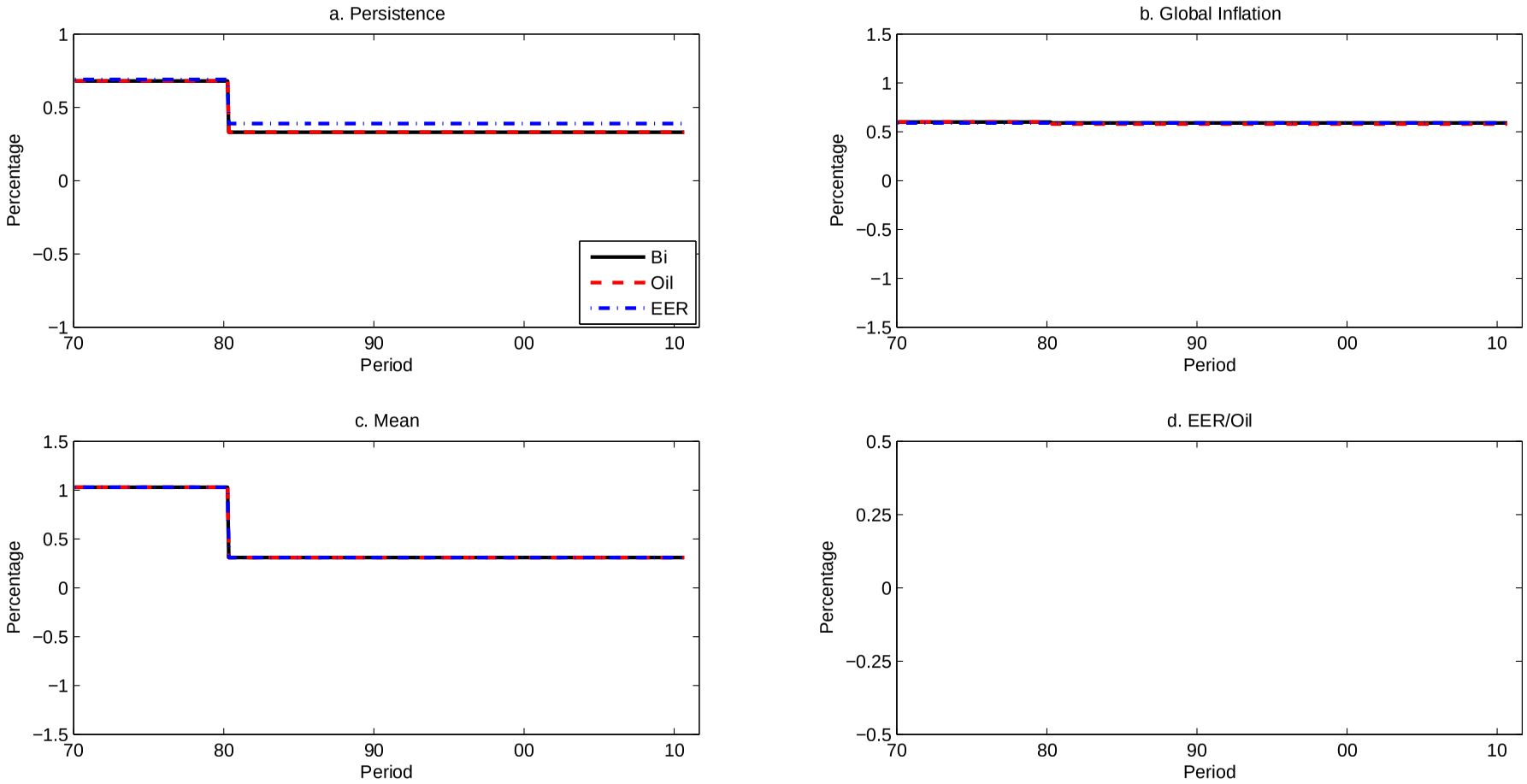
Notes: Same as figure 1.2.1.

Figure 1.2.17: Coefficient changes: Switzerland



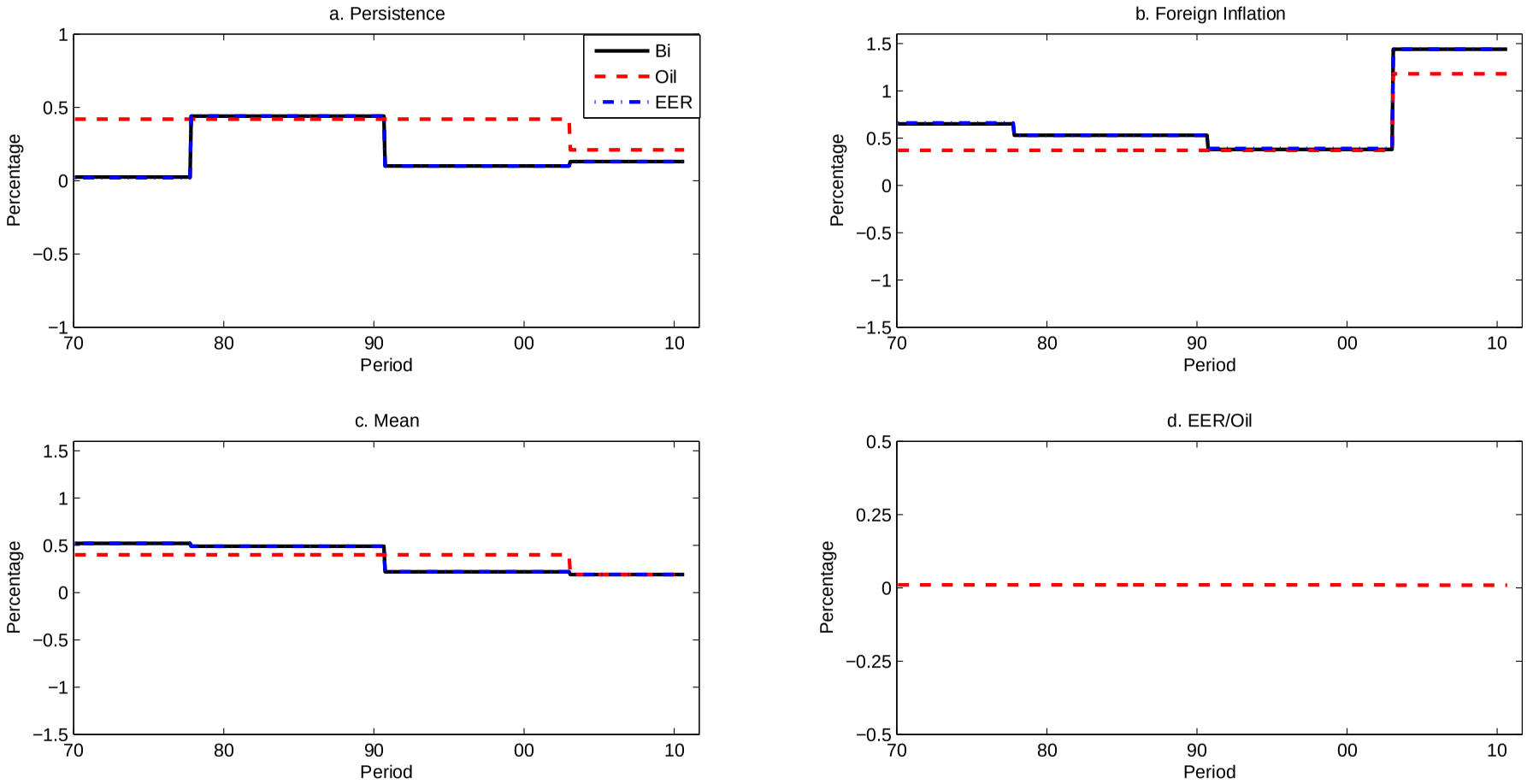
Notes: Same as figure 1.2.1.

Figure 1.2.18: Coefficient changes: United Kingdom



Notes: Same as figure 1.2.1.

Figure 1.2.19: Coefficient changes: United States



Notes: Same as figure 1.2.1.

Appendix A

A.1 Outliers and Breaks in The Exogenous Variable

Consider following data generating process which exhibits no break

$$\begin{aligned} y_t = \quad & 0.5 + 0.5y_{t-1} + 0.5z_t + e_t \\ & e_t \sim N(0,1) \end{aligned} \tag{A.1}$$

where y_{t-1} is one period lag of dependent variable and z_t is an exogenous independent variable with a random normal distribution. First, we undertook a simulation study to examine the performance of the Chow test with the explanatory variable z_t having moderate (observations between 90-100 are set equal to 50) and large (observations between 90-100 are set equal to 100) size outliers. Based on the 10000 replications, on average, the size of the test is unaffected at 10%, 5% and 1% significance levels.

Second, we analyze the size of the tests: WDmax and Sequential $SupF(l+1|l)$ rejecting the true null hypothesis of no break, when explanatory variable z_t is subject to a single break in the mean or variance. We impose a break in the mean by increasing it by 5, and also in the variance by multiplying it by 5, at the break point. A break occurs at $t = bT$ where $b = 0.5$ or $b = 0.75$. After 5000 replications, on average, the size of the WDmax and Sequential tests is well-sized at a 5% significance level.

Table A.1: Size of a coefficient break test: Outliers in the explanatory variable

Test	Method	Sample	Size of Outliers	Size*		
				α=0.10	α=0.05	α=0.01
Coefficient break test	Chow test	T=500	Moderate	0.100	0.051	0.010
		T=500	Large	0.099	0.051	0.010

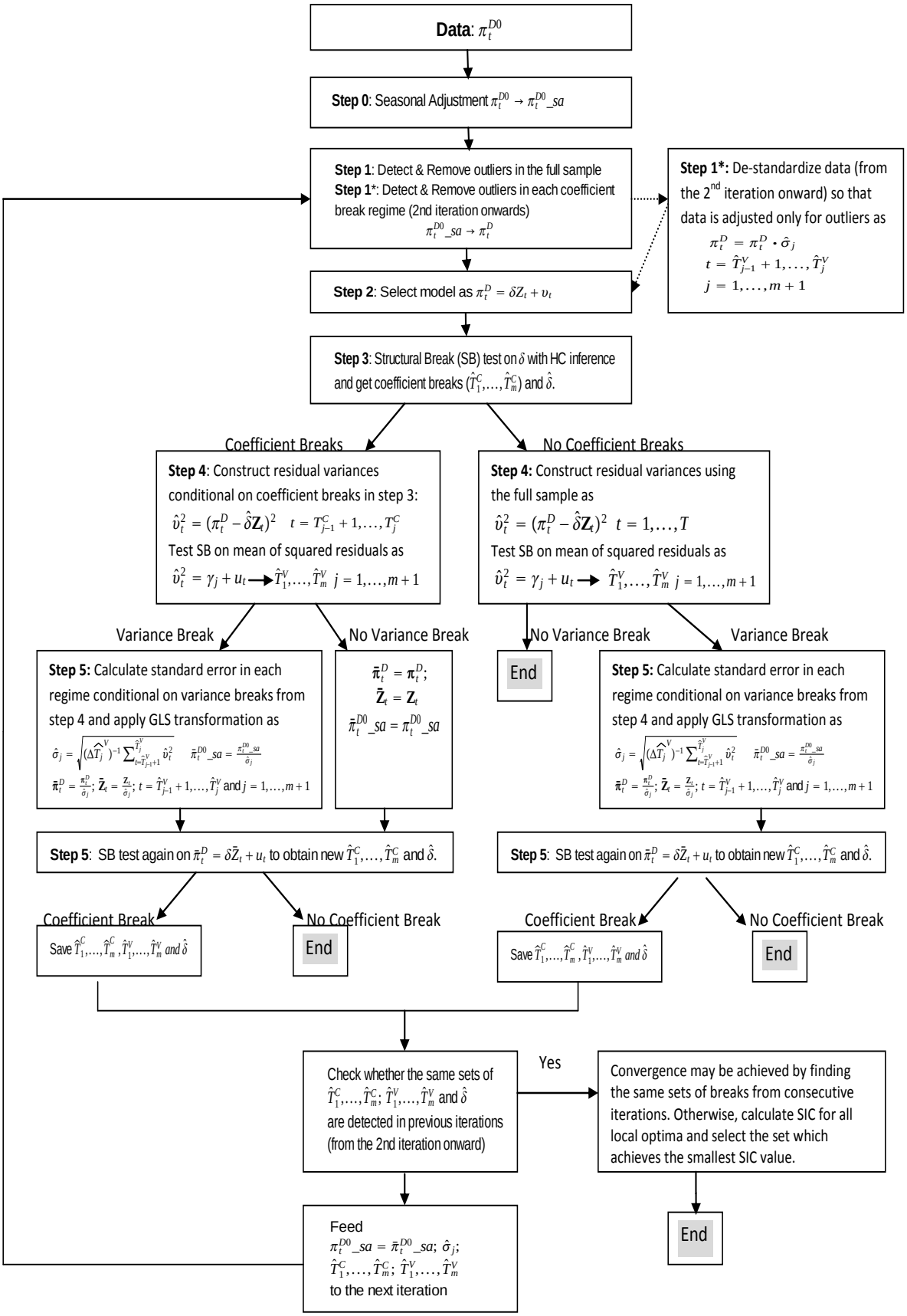
* After 10000 replications

Table A.2: Size of the tests: A break in the explanatory variable

Test	Method	Specification	Sample	Size*	
				b=0.5	b=0.75
A. A single break in the mean of explanatory variable					
Coefficient break test	SupF(l+1|l)	M=5, ϵ=0.15	T=500	0.046	0.047
	WDmax	M=5, ϵ=0.15	T=500	0.033	0.037
Variance break test	SupF(l+1|l)	M=5, ϵ=0.15	T=500	0.060	0.060
	WDmax	M=5, ϵ=0.15	T=500	0.075	0.073
B. A single break in the variance of explanatory variable					
Coefficient break test	SupF(l+1|l)	M=5, ϵ=0.15	T=500	0.020	0.020
	WDmax	M=5, ϵ=0.15	T=500	0.016	0.019
Variance break test	SupF(l+1|l)	M=5, ϵ=0.15	T=500	0.066	0.066
	WDmax	M=5, ϵ=0.15	T=500	0.079	0.077

*After 5000 replications

A.2 An algorithm for the iterative structural break testing method



A.3 Additional Tables

Table A.3: Test statistics and p values for the coefficient breaks tests for univariate models

	WDmax		SupF(l+1|l)									
Countries	SupF (K)	p-value	SupF(1)	p-value	SupF(2|1)	p-value	SupF(3|2)	p-value	SupF(4|3)	p-value	SupF(5|4)	p-value
Austria	18.198	0.236	18.198	0.167	9.636	0.988	9.049	1.000	3.446	1.000	0.000	1.000
Belgium	48.729	0.000	45.205	0.000	32.753	0.002	20.928	0.197	8.234	1.000	6.407	1.000
Canada	48.448	0.000	48.448	0.000	28.257	0.042	11.315	1.000	10.222	1.000	0.000	1.000
Denmark	39.426	0.000	39.426	0.000	23.749	0.028	6.850	1.000	3.494	1.000	0.000	1.000
Finland	37.183	0.000	37.183	0.000	19.378	0.129	12.907	0.810	5.832	1.000	5.607	1.000
France	51.215	0.000	51.215	0.000	25.891	0.026	20.184	0.244	4.026	1.000	0.000	1.000
Germany	29.735	0.002	29.735	0.003	13.075	0.795	7.019	1.000	15.084	0.838	0.000	1.000
Greece	30.923	0.001	30.923	0.001	18.123	0.190	17.610	0.312	9.537	1.000	10.196	1.000
Italy	28.333	0.002	23.943	0.013	22.452	0.045	13.722	0.723	4.920	1.000	0.000	1.000
Japan	45.990	0.000	45.990	0.000	33.466	0.007	15.356	0.939	12.524	1.000	0.000	1.000
Korea	37.303	0.000	37.303	0.000	21.641	0.060	17.580	0.315	11.808	0.962	0.000	1.000
Netherland	26.119	0.012	25.586	0.015	16.581	0.440	15.624	0.679	7.098	1.000	0.000	1.000
Norway	30.855	0.001	30.855	0.001	18.729	0.158	9.932	0.990	9.813	1.000	0.000	1.000
Portugal	35.826	0.000	34.097	0.000	25.067	0.007	7.952	0.997	6.996	1.000	0.000	1.000
Spain	38.338	0.000	29.791	0.006	26.405	0.043	22.126	0.229	14.156	0.980	6.628	1.000
Sweden	38.466	0.000	38.466	0.000	9.828	0.941	8.627	1.000	7.102	1.000	1.715	1.000
Switzerland	16.833	0.349	15.564	0.329	13.624	0.742	9.179	1.000	6.916	1.000	0.000	1.000
UK	36.344	0.000	35.546	0.000	22.595	0.020	9.139	0.971	21.582	0.058	0.000	1.000
US	38.544	0.000	38.544	0.000	20.139	0.100	12.787	0.822	11.027	0.988	11.111	0.997

Notes: Table provides WDmax and Sequential SupF(l+1|l) tests statistics, described in the text. Sequential test statistics are computed for up to 5 breaks regardless of non rejection of null hypothesis at certain points. Approximate p values are calculated corresponding to these statistics using Hall and Sakkas (2011). Consequently, the number of break dates are decided by the last rejection of null hypothesis by 5% siginificance level.

Table A.4: Test statistics and p values for the variance breaks tests for univariate models

Countries	WDmax		SupF(l+1|l)									
	SupF (K)	p-value	SupF(1)	p-value	SupF(2|1)	p-value	SupF(3|2)	p-value	SupF(4|3)	p-value	SupF(5|4)	p-value
Austria	16.251	0.001	16.251	0.001	3.071	0.805	1.588	1.000	1.170	1.000	0.000	1.000
Belgium	12.271	0.008	12.271	0.009	7.128	0.193	0.944	1.000	0.871	1.000	0.631	1.000
Canada	10.585	0.021	8.216	0.062	12.680	0.016	1.552	1.000	2.696	0.992	0.000	1.000
Denmark	43.583	0.000	43.583	0.000	24.920	0.000	8.217	0.178	2.581	0.996	0.000	1.000
Finland	33.848	0.000	33.848	0.000	22.771	0.000	1.050	1.000	4.086	0.846	0.000	1.000
France	8.190	0.077	8.190	0.063	10.128	0.052	2.778	0.959	0.000	1.000	0.000	1.000
Germany	7.886	0.090	6.146	0.156	10.991	0.035	2.671	0.969	1.344	1.000	0.150	1.000
Greece	49.472	0.000	49.472	0.000	29.727	0.000	3.976	0.774	1.316	1.000	0.261	1.000
Italy	69.104	0.000	69.104	0.000	34.201	0.000	1.855	1.000	5.500	0.600	0.000	1.000
Japan	75.835	0.000	75.835	0.000	32.424	0.000	1.562	1.000	1.421	1.000	1.321	1.000
Korea	103.018	0.000	103.018	0.000	7.922	0.138	1.909	1.000	6.050	0.507	0.000	1.000
Netherland	19.954	0.000	19.954	0.000	4.066	0.607	1.037	1.000	2.308	1.000	0.000	1.000
Norway	19.376	0.000	19.376	0.000	13.894	0.009	0.669	1.000	0.000	1.000	0.000	1.000
Portugal	142.665	0.000	142.665	0.000	101.644	0.000	43.285	0.000	2.237	1.000	1.210	1.000
Spain	30.594	0.000	30.594	0.000	37.966	0.000	8.976	0.129	11.615	0.051	0.000	1.000
Sweden	26.880	0.000	26.880	0.000	8.310	0.117	3.687	0.826	0.412	1.000	0.000	1.000
Switzerland	16.523	0.001	16.523	0.001	2.615	0.891	12.422	0.027	0.000	1.000	0.000	1.000
UK	53.670	0.000	53.670	0.000	4.756	0.483	7.642	0.226	0.000	1.000	0.000	1.000
US	15.082	0.002	12.828	0.007	14.230	0.008	6.008	0.419	17.193	0.003	0.000	1.000

Notes: Same as Table A.2

Table A.5: Test statistics and p values for the coefficient breaks tests for bivariate models

Countries	WDmax		SupF(l+1|l)									
	SupF (K)	p-value	SupF(1)	p-value	SupF(2|1)	p-value	SupF(3|2)	p-value	SupF(4|3)	p-value	SupF(5|4)	p-value
Austria	21.962	0.001	18.766	0.008	18.561	0.016	6.977	0.914	9.714	0.677	0.000	1.000
Belgium	71.323	0.000	71.323	0.000	26.662	0.001	7.885	0.963	7.320	0.999	0.000	1.000
Canada	59.903	0.000	59.903	0.000	17.381	0.236	9.638	0.994	17.403	0.409	6.266	1.000
Denmark	38.873	0.000	32.203	0.000	33.921	0.000	8.855	0.895	4.104	1.000	0.000	1.000
Finland	27.046	0.001	27.046	0.002	14.378	0.329	12.124	0.715	13.519	0.650	0.000	1.000
France	99.698	0.000	99.698	0.000	11.800	0.603	13.752	0.521	5.750	1.000	0.000	1.000
Germany	50.335	0.000	45.088	0.000	38.736	0.000	12.450	0.107	5.534	0.944	0.000	1.000
Greece	42.356	0.000	42.356	0.000	19.569	0.121	17.702	0.304	7.688	1.000	0.000	1.000
Italy	38.886	0.000	36.004	0.000	23.747	0.028	20.196	0.145	16.543	0.501	0.000	1.000
Japan	27.641	0.003	27.641	0.003	11.383	0.824	15.394	0.533	14.026	0.789	11.867	0.985
Korea	25.325	0.001	25.194	0.001	17.389	0.065	5.812	1.000	2.072	1.000	0.000	1.000
Netherland	26.862	0.001	21.381	0.016	18.394	0.095	11.092	0.829	4.645	1.000	0.000	1.000
Norway	34.002	0.000	34.002	0.000	13.193	0.632	12.764	0.824	13.502	0.842	0.000	1.000
Portugal	38.045	0.000	36.334	0.000	23.528	0.030	15.809	0.487	14.357	0.753	0.000	1.000
Spain	32.957	0.000	32.957	0.000	23.827	0.012	12.252	0.700	11.943	0.832	0.000	1.000
Sweden	21.033	0.020	21.033	0.019	11.028	0.695	9.085	0.973	5.964	1.000	0.000	1.000
Switzerland	53.086	0.000	53.086	0.000	15.742	0.223	14.818	0.404	9.416	0.990	0.000	1.000
UK	38.406	0.000	38.406	0.000	15.758	0.222	15.160	0.370	18.807	0.157	0.000	1.000
US	61.121	0.000	61.121	0.000	17.468	0.025	20.248	0.011	10.024	0.633	1.437	1.000

Notes: Same as Table A.2

Table A.6: Test statistics and p values for the variance breaks tests for bivariate models

Countries	WDmax		SupF(l+1|l)									
	SupF (K)	p-value	SupF(1)	p-value	SupF(2|1)	p-value	SupF(3|2)	p-value	SupF(4|3)	p-value	SupF(5|4)	p-value
Austria	21.271	0.000	21.271	0.000	7.851	0.142	1.503	1.000	2.436	0.998	0.000	1.000
Belgium	31.948	0.000	31.948	0.000	6.884	0.214	1.472	1.000	0.237	1.000	0.000	1.000
Canada	12.980	0.005	12.980	0.007	3.318	0.756	5.212	0.548	1.016	1.000	0.000	1.000
Denmark	39.623	0.000	39.623	0.000	26.253	0.000	1.528	1.000	3.984	0.862	0.000	1.000
Finland	48.049	0.000	48.049	0.000	10.476	0.044	2.743	0.962	1.295	1.000	0.000	1.000
France	20.012	0.000	20.012	0.000	5.819	0.326	4.225	0.728	2.078	1.000	1.067	1.000
Germany	11.302	0.014	11.302	0.015	2.719	0.872	0.456	1.000	0.433	1.000	0.622	1.000
Greece	53.351	0.000	53.351	0.000	29.896	0.000	3.041	0.927	0.612	1.000	0.000	1.000
Italy	59.358	0.000	59.358	0.000	20.789	0.000	5.812	0.449	2.753	0.989	0.000	1.000
Japan	95.393	0.000	95.393	0.000	42.919	0.000	11.147	0.049	1.118	1.000	0.000	1.000
Korea	125.302	0.000	125.302	0.000	12.387	0.018	1.747	1.000	8.143	0.236	0.000	1.000
Netherland	15.696	0.001	15.696	0.002	6.509	0.249	3.174	0.909	2.683	0.992	0.000	1.000
Norway	21.758	0.000	21.758	0.000	7.407	0.172	4.909	0.601	3.007	0.975	0.000	1.000
Portugal	107.337	0.000	107.337	0.000	94.088	0.000	29.350	0.000	4.963	0.696	1.127	1.000
Spain	36.326	0.000	36.326	0.000	31.839	0.000	19.132	0.001	7.196	0.341	0.000	1.000
Sweden	30.454	0.000	30.454	0.000	17.322	0.002	4.052	0.760	2.639	0.994	0.000	1.000
Switzerland	39.175	0.000	39.175	0.000	8.187	0.123	6.398	0.365	1.311	1.000	0.000	1.000
UK	12.396	0.008	12.396	0.009	6.376	0.262	2.069	1.000	0.889	1.000	0.000	1.000
US	24.508	0.000	24.508	0.000	16.960	0.002	3.601	0.841	10.536	0.084	0.073	1.000

Notes: Same as Table A.2

Chapter 2

What is the Globalization of Inflation

2.1 Introduction

In recent years policymakers and researchers have documented and discussed the globalization of inflation, namely the apparently strong international co-movement of inflation seen over the last two decades or more. Indeed, even in the context of the large economies of the US and Euro area, Bernanke (2007) and Trichet (2008), respectively, emphasize that their central banks need to monitor carefully international price developments and analyze their implications for the domestic economy. The strong link between domestic inflation and the international environment is also recognized in the models of Pesaran et al. (2004), Ciccarelli and Mojon (2010), Mumtaz and Surico (2012), Bataa et al. (2013a) and many others. Likewise, exploring the international link and examining its dynamics over time were also the focus of chapter 1 by which we document a positive and strengthening contemporaneous relationship between domestic and country specific foreign inflation.

Despite this now widespread recognition that inflation exhibits strong links across developed countries, limited progress has been made in uncovering the sources of these links. For example, the papers of Borio and Filardo (2007) on the one hand and Ihrig et al. (2010) on the other come to opposing views about the importance of the foreign output gap in Phillips curve models for a range of countries, while Ihrig et al. (2010) and Peacock and Baumann (2008) find little evidence that the role of import prices increases over time. Nevertheless, positive results are beginning to emerge, with Neely and Rapach (2011) finding that domestic inflation is more closely linked to foreign inflation in developed countries with greater central bank independence and more openness to trade,

while Mumtaz and Surico (2012) suggest that better monetary policy, and specifically the adoption of inflation targeting by a number of countries in the 1990s, has increased co-movement.

The above papers on the global aspects of inflation predominantly employ headline or aggregate inflation, as does the analysis in chapter 1. However, it is frequently argued (for example, Mishkin (2007)) that core inflation is the appropriate concept for monetary policy purposes, where the 'core' represents the underlying or trend inflation level. Although there is some debate about how this 'core' should be measured, it is most commonly represented as inflation in the consumer price index (CPI) omitting energy and food products; see Clark (2001) for a discussion on the concept and its measurement. Energy and food are seen as volatile components, with the former being subject to international demand and supply shocks and the latter to the vagaries of the weather, whose effects may not persist over time. Therefore, an analysis of core inflation may provide clearer evidence on whether monetary policy plays a role for inflation linkages and their changes over time. Further, while it is to be anticipated that energy inflation has a strong international dimension, it is unclear whether these have changed over time. On the other hand, there have been large changes in food supply for developed economies over the last forty years, in many cases moving from predominantly domestically produced to being largely imported, pointing to the possibility of increased international co-movement for food inflation.

The present chapter[1] asks what is the globalization of inflation by separating the core, energy and food components of aggregate CPI inflation, studying changes in the international linkages in these components since 1970 alongside a corresponding analysis of aggregate national CPI inflation for 13 OECD countries. In particular, we examine structural breaks in a dynamic model for domestic (aggregate or component) monthly inflation in relation to a corresponding country-specific foreign series, with the latter constructed as the bilateral trade-weighted average of inflation in the other countries of our sample. In addition, we dissect the globalization of aggregate inflation by examining whether breaks in international linkages at the aggregate level can be attributed to changes in the response of domestic inflation to foreign core, energy or food inflation. For this purpose, we employ an iterative structural break testing methodology proposed in chapter 1 in order to avoid the conflation of coefficient and volatility breaks.

[1]This chapter was presented at ESEM 2013 and is available on the conference website under the co-authorship of myself with Ralf Becker, George J Bratsiotis and Denise R Osborn. I undertook all the empirical work for the chapter and wrote the first draft of the text. The text was subsequently edited by my co-authors.

To preview our results, we find that relationships for core inflation are often distinctive from those of aggregate inflation, with the former providing less evidence of both changes in and the strength of international co-movement. Food inflation, on the other hand, displays a pattern of increasing co-movement, as (to a less clear extent) does energy. The most evident characteristic of the globalization of inflation for these countries is a convergence in the mean rates for each of aggregate, core and food inflation. The analysis of breaks in aggregate domestic inflation in terms of component foreign series shows an increased role for core inflation for a number of European countries, pointing to the importance of monetary policy in the context of the formation of the Euro area, while an increased linkage to movements in foreign energy inflation is important for the US.

The rest of the chapter is organized as follows. Section 2.2 describes our methodology, including our iterative procedure for structural break detection, with the data presented in section 2.3. Substantive results are reported and discussed in the following two sections, with the nature of change documented for individual (aggregate and component) series in section 2.4 and breaks in aggregate inflation decomposed in terms of foreign inflation series in section 2.5. Finally, section 2.6 concludes.

2.2 Methodology

The methodology outlined in chapter 1 is repeated in this chapter for convenience. However, it is written here explicitly for the bivariate domestic-foreign inflation model. Additionally, compared to the discussion of bivariate inflation model in subsection 1.3.3 of chapter 1, a more general specification of aggregate/component model is discussed in this section.

2.2.1 The model

Structural break analyses need to take account of possible breaks in both dynamics and volatility in order to avoid misleading inferences arising from model misspecification (Pitarakis, 2004). Therefore, our approach follows that of Bataa et al. (2013a,b) in iteratively taking account of breaks in both the coefficients and the innovation variance of inflation.

The present chapter focuses on the relationship between domestic inflation for country s ($s = 1, \ldots, N$) in month t ($\pi^D_{t,s}$) and a corresponding measure of foreign inflation ($\pi^F_{t,s}$); the inflation series under analysis can be either aggregate CPI inflation or a component (core, food or energy). In a dynamic

context, the relationship can be parsimoniously represented as[2]

$$\pi_{t,s}^{D} = \alpha_{0j} + \sum_{i=1}^{p} \alpha_{ij}\pi_{t-i,s}^{D} + \beta_{0j}\pi_{t,s}^{F} + \sum_{i=1}^{r} \beta_{ij}\pi_{t-i,s}^{F} + \varepsilon_{t,s}, \tag{2.1}$$

for $j = 1, \ldots, m+1$ and $t = T_{j-1}^{C} + 1, \ldots, T_{j}^{C}$, where $\pi_{t,s}^{F}$ is foreign inflation in relation to country s at time t, β_{0j} captures the contemporaneous co-movement between domestic and foreign inflation, while $(\beta_{1j}, ..., \beta_{rj})$ represents dynamic inflation spillovers, $(\alpha_{1j}, \ldots, \alpha_{pj})$ are the own inflation dynamics, and ε_t is a temporally uncorrelated (but possibly heteroskedastic) disturbance process. For estimation purposes, foreign inflation is treated as weakly exogenous for domestic inflation. The measurement of foreign inflation for country s is previously discussed in subsection 1.3.4 of chapter 1. The maximum lag order considered for both own and foreign inflation is $p = r = 12$, but the included lags are specified using a general to specific approach, as explained below. In line with the usual definition employed in a univariate context, $\rho_j^d = \sum_{i=1}^{p} \alpha_{ij}$ is referred to as inflation persistence, although it is measured in (2.1) conditionally on foreign inflation.

The coefficients of (2.1) are subject to change at the m break dates $(T_1^C, \ldots, T_m^C)$, with the convention that $T_0^C = 0$ and $T_{m+1}^C = T$, where T is the total sample size available for estimation (that is, after allowing for lags). Within each of $m+1$ coefficient regimes, $\alpha_j = (\alpha_{0j},\ \alpha_{1j},\ \ldots, \alpha_{pj})'$ and $\beta_j = (\beta_{0j},\ \beta_{1j},\ \ldots, \beta_{rj})'$ are time-invariant and all AR roots are assumed to lie strictly outside the unit circle. Similarly, the innovation variance $E[\varepsilon_t^2] = \sigma_k^2$ is constant and is assumed to be conditionally homoskedastic within each volatility regime $k = 1, \ldots, n+1$, but it is allowed to change at n volatility break dates. Both the numbers and dates of all breaks are unknown, with no restriction that coefficient and volatility breaks coincide, in either number or timing. Further, our procedure allows for the presence of additive outliers in π_t^D, which could be due to (say) changes in indirect taxes. A feature of our treatment of outliers is that these are measured after adjusting for volatility breaks.

Before turning to the iterative methodology used for structural break detection, some further discussion of (2.1) is warranted. It is noteworthy that, by construction, any common global influences are included in the foreign inflation series for country s, so that any increase over time in truly global inflation effects will be captured by the β_{ij} coefficients. This is not a problem for our analysis in that such global effects are, indeed, foreign in relation to country s. Dees et al. (2007) discuss related issues in the context of their GVAR

[2] All coefficients are country-specific, but the subscript s is omitted, and also from the disturbance in (2.1), for notational simplicity.

model, noting that it may imply that cross-section correlations exist across the $\varepsilon_{t,s}$; however, unlike those authors, we are not concerned with identifying country-specific shocks.

The assumption of weak exogeneity for $\pi^F_{t,s}$ may be questioned for the large economies of Germany and the US, both of which are included in our sample and (as noted in Section 2.3) attract large weight in our foreign inflation series for a number of other countries. Consequently, it is possible that an apparent increase in globalization for these countries may actually be due to an increased role for US or German inflation for other countries. However, allowing for simultaneity using the two-step methodology of Hall et al. (2012) is limited in the nature of breaks that can be examined[3]. An alternative model for aggregate inflation was therefore also examined for these countries, with $\pi^F_{t,s}$ omitted from the analysis. For Germany, the results are substantively unchanged from those reported below, whereas less evidence of globalization was uncovered in this modified specification for the US. We return to this point for the US in the final section of the chapter.

2.2.2 Iterative testing methodology

The iterative methodology proposed by Bataa et al. (2013b) employs structural break tests in conjunction with the outlier detection and removal procedure of Stock and Watson (2003) to examine structural breaks in each of the seasonal, mean, dynamic and volatility components of univariate inflation. While based on this approach, our procedure differs in three important respects. Firstly, we treat all elements of the coefficient vector $\delta_j = (\alpha'_j, \beta'_j)'$ of (2.1) together, rather than separating mean and dynamic breaks. In addition to mean breaks in this bivariate context requiring consideration of both domestic and foreign inflation, the simulation analysis in Bataa et al. (2013b, Table 1) shows that their procedure results in the mean break test being substantially oversized. This feature may be due to the initial tests for mean breaks applying HAC inference with un-modeled dynamics; Bai and Perron (2006) show that the use of HAC inference in this context can lead to badly oversized tests, whereas this is improved when the dynamics are modeled explicitly. Secondly, our procedure simplifies the iterations in respect to volatility breaks, since the results of Bataa et al. (2013a, Table 1) implies these are detected well without iteration. Finally, we take account of variance breaks when identifying outliers,

[3]More specifically, for a 'structural' equation such as (2.1), breaks are estimated first for the coefficients of the corresponding 'reduced form' equation. Possible breaks in the 'structural form' coefficients can then be examined only for the sub-intervals between the 'reduced form' breaks, rather than over the entire sample period.

whereas these are ignored in their procedure.

It is also important to note that seasonality is not a focus of interest in the present study, and hence we avoid the additional complications of the Bataa et al. (2013b) procedure in detecting breaks in a deterministic (dummy variable) representation of seasonality; in this context, it is reassuring that the robustness analysis in Bataa et al. (2013b) indicates that the detection of structural breaks in univariate dynamics is not substantively affected by the method of accounting for seasonality. As discussed in section 2.3, all series are seasonally adjusted prior to the application of the iterative procedure[4].

Step 1 - Outlier detection: In the initialization, outliers in the univariate $\pi^D_{t,s}$ series[5] are detected using the procedure by Stock and Watson (2003) over the full sample of data $t = 1, \ldots, T$. Outliers are defined as four times the interquartile range from the median[6], and are replaced by the median of the six neighboring non-outlier values.

Step 1* - Outlier detection for subsequent iterations: In subsequent iterations, the data are adjusted for volatility breaks (by standardizing $\pi^D_{t,s}$ using the residual standard deviation for the volatility regime corresponding to t), with outliers then examined separately within each coefficient regime. This is because, observations which were initially classified as outliers are reconsidered in every iteration as they may indeed part of the changed volatility and should not be classified as outliers any longer. Detected outliers are replaced by the median of the six neighboring non-outlier values and (except for the initialization) the data are rescaled to yield a series adjusted only for outliers.

Step 2 - Model selection: The model is specified using a simplified version of the general to specific multi path search algorithm proposed by Krolzig and Hendry (2001), in combination with the Schwartz Information Criterion (SIC). Since the intercept is always included, the model is initially evaluated with 25 individual (own and global) lags. Five starting points are then generated by initially eliminating the single variable that is the zth least significant ($z = 1, ..., 5$) in the general regression, calculating the corresponding SIC value in each case. From each starting point, the least significant variable is dropped sequentially one at a time, until only

[4]This is essentially the same algorithm proposed in chapter 1

[5]A small Monte Carlo analysis confirmed that the presence of aberrant observations in the explanatory variable did not affect the size of the structural break tests for (2.1).

[6]Based on a visual examination of results, using four times the interquartile range provides a balance between failing to detect 'obvious' outliers and apparently detecting too many.

the intercept remains. The selected model is that which achieves the smallest SIC across all 25 models and 5 paths[7].

Step 3 - Preliminary coefficient break test: After specifying the lags included in (2.1), the Bai and Perron (1998) multiple structural breaks test procedure is applied to the coefficients (intercept and all slope coefficients) employing heteroskedasticity consistent (HC) inference.

Step 4 - Variance break test: Using the residuals from the model with coefficient breaks as identified in step 3, variance breaks are examined through tests applied to the mean of the squared residuals.

Step 5 - Coefficient break test: Since HC inference can lead to oversized coefficient break tests (Bai and Perron, 2006), breaks in the coefficients are reconsidered conditional on the variance breaks from step 4. Following the proposal of Pitarakis (2004), this is achieved by applying homoskedastic inference in (2.1) after applying the feasible GLS transformation. If no volatility breaks are detected, coefficient tests are applied to the original data with a homoskedastic variance assumption.

The heart of the iterations just described is the multiple structural break testing procedure of Bai and Perron (1998) which is detailed in subsection 1.3.2 of chapter 1.

A single iteration comprises steps 1 to 5, inclusive, and the output of each iteration is two sets of break dates, namely identified coefficient and volatility break dates, and outliers. The maximum number of iterations is set to 10 and convergence may be achieved in two different ways. Firstly, identical dates may be output from two consecutive iterations; alternately, the iterations can cycle between (say) two or three sets of dates. In the latter case, we focus on coefficient breaks and choose the set which achieves the smallest SIC among those in the cycle. The version of SIC is that proposed by Yao (1988) for structural break inference, which is applied to the GLS transformed data and calculated for m breaks as

$$SIC(m) = \ln[T^{-1}S_T(\hat{T}_1^C, \ldots, \hat{T}_m^C)] + q^* \ln(T)/T \tag{2.2}$$

where $S_T(\hat{T}_1^c, \ldots, \hat{T}_m^c)$ is the sum of squared standardized residuals for $\pi_{t,s}^D$ computed over the $m+1$ coefficient regimes in (2.1) and $q^* = (m+1)q + m$ where q is the total number of coefficients (including the intercept) estimated

[7] A comparison of this SIC based procedure with a conventional testing down method led to the selection of very similar lags.

in the model. Note that, through q^*, the penalty term effectively treats each coefficient break date as an estimated parameter. The T sample observations for $\pi^D_{t,s}$ used in computing SIC are identical over all models in the comparison.

2.2.3 Decomposing foreign inflation

In order to shed light on whether individual breaks in international inflation linkages in (2.1) can be attributed to changed responses to a specific component of foreign inflation, we also study the following generalized version of this dynamic model:

$$\pi^D_t = \alpha_{0j} + \sum_{i=1}^{p} \alpha_{i,j}\pi^D_{t-i} + \sum_{l=1}^{3}\{\beta_{0l,j}\pi^F_{t,l} + \sum_{i=1}^{r} \beta_{il,j}\pi^F_{t-i,l}\} + \varepsilon_t \tag{2.3}$$

for $j = 1, \ldots, m+1$, and $t = T^C_{j-1}+1, \ldots, T^C_j$. where π^D_t is domestic aggregate inflation and $\pi^F_{t,l}$ $(l = 1, 2, 3)$ are the component foreign series relating to core, energy and food inflation in month t. Both the domestic and foreign inflation series are specific to country s, but this subscript is dropped from (2.3) for notational simplicity.

We employ two approaches to the identification of breaks in (2.3). One takes the coefficient and volatility breaks identified in (2.1) as also applying to (2.3). In this case, based on those dates, the more general model is estimated using the same dynamic specification as selected for (2.1), and also replacing the outliers as identified there for π^D_t, but replacing the aggregate series on the right-hand side of the model by the three component series. The second approach applies the iterative methodology of subsection 2.2.2 to the more general model of (2.3). Whereas the former provides a direct test of whether the given structural breaks in aggregate international inflation linkages can be associated with coefficient change for one or more specific components, the latter treats this as a distinct model and can be viewed as a check on the breaks identified from the aggregate. However, it is also relevant to note that the second approach is less parsimonious and hence may lack power for the detection of breaks compared to the first.

Due to the costs of searching in the more highly parameterized model of (2.3), model selection is modified when the iterative methodology of subsection 2.2.2 is applied in this context. To be specific, model selection uses the SIC-based procedure outlined there, but this is applied once, prior to the commencement of the iterations, so that the same dynamic specification is employed throughout.

The use of (2.3) in place of (2.1) effectively treats the aggregate foreign in-

flation series as a weighted sum of the component series with constant weights. However, both the weights we use to construct each foreign inflation series for country s from inflation data for other countries (see subsection 1.3.4 of chapter 1) and those implicitly used to construct an aggregate inflation series from its components within each country change over time. Hence it is only an approximation to consider aggregate foreign inflation as a fixed weighted sum of the corresponding component series. This is, indeed, an additional reason why it is appropriate to re-evaluate the existence and dates of breaks in international inflation linkages in the context of (2.3).

The primary focus of our analysis is the nature of changes in the international co-movement of inflation. Reflecting this, and for given (estimated) dates of structural breaks in (2.3) we apply a sequence of F-tests on the coefficients of this model[8]. Of particular interest is the test of the hypothesis of no change in the contemporaneous coefficient across regimes which is applied both separately for each components $l = 1, 2, 3$

$$H_0 : \beta_{0l,1} = \cdots = \beta_{0l,m+1} \tag{2.4}$$

and jointly across all three components

$$H_0 : \beta_{0l,1} = \cdots = \beta_{0l,m+1}, \quad all \quad l = 1, 2, 3. \tag{2.5}$$

The tests of (2.4) and (2.5) are applied to both sets of estimated break dates for (2.3), namely those based on the aggregate CPI model of (2.1) and those estimated directly from (2.3). In addition, the corresponding tests are applied to the lagged coefficients for foreign inflation, namely

$$H_0 : \beta_{il,1} = \cdots = \beta_{il,m+1}, \quad all \quad i = 1, ..., r \tag{2.6}$$

and

$$H_0 : \beta_{il,1} = \cdots = \beta_{il,m+1}, \quad all \quad i = 1, ..., r \quad and \quad l = 1, 2, 3. \tag{2.7}$$

Care needs to be taken when comparing the coefficients of foreign inflation components estimated in (2.3) with those for the separate models estimated using (2.1). To see this, consider the following special case of (2.1) for the

[8]Since Bai and Perron (1998) show that (with coefficient breaks of fixed magnitude) the estimated break fractions asymptotically converge to the true values at a rate of T, whereas the estimated coefficients in a model such as (2.3) converge at the usual rate of $T^{1/2}$. This implies that conventional hypothesis tests are asymptotically valid when applied to the coefficients conditional on the estimated break dates.

inflation component l ($l = 1, 2, 3$)

$$\pi_{tl}^D = \alpha_0 + \alpha_1 \pi_{t-1,l}^D + \beta_0 \pi_{tl}^F + \varepsilon_{tl}, \quad l = 1, 2, 3 \tag{2.8}$$

where no structural breaks apply and the country subscript s is omitted for simplicity. Note, in particular, that (2.8) assumes common coefficients across components. Further, the aggregate is a weighted sum of the component inflation series, and assuming constant weights both over time and across domestic and foreign inflation with $\pi_t^D = \sum_{l=1}^3 \omega_l \pi_{tl}^D$ and $\pi_t^F = \sum_{l=1}^3 \omega_l \pi_{tl}^F$, then aggregating (2.8) across components gives

$$\pi_t^D = \alpha_0 + \alpha_1 \pi_{t-1}^D + \sum_{l=1}^{3} (\omega_l \beta_0) \pi_{tl}^F + \varepsilon_t. \tag{2.9}$$

Consequently, the contemporaneous coefficient of the foreign component series l in (2.9) is not β_0, but rather $\beta_{0l} = \omega_l \beta_0$.

Therefore, consideration of this simple special case implies that coefficients estimated in the component models (2.1) should be scaled by the weights ω_l when compared with those estimated in (2.9). In the case of the US, for example, averaged over the period 1987 to 2012, goods and services that contribute to core inflation have a weight of 0.77 in aggregate inflation, with the weights of energy and food inflation being 0.09 and 0.14, respectively. Therefore, the estimated coefficients of foreign energy and food inflation can be anticipated to be substantially smaller in the context of (2.3) than in the separate models of (2.1), even when the same coefficients apply in the latter across the three components. Due to the larger role it plays in the aggregate, the reduction will be less marked for core inflation.

2.3 Data

The inflation data comprises monthly CPI aggregate inflation, together with the corresponding core, energy and food component inflation series, for OECD countries over January 1970 to September 2010. Our sample relates to 13 OECD countries, that is fewer than 19 OECD countries used in the aggregate analysis of chapter 1, due to the limitations of the availability of data at the sub-aggregate core, energy and food level and at the monthly frequency over an extended period. These include six countries that are currently members of the Euro area (Austria, Finland, France, Germany, Italy, Netherlands), four other European countries (Denmark, Sweden, Switzerland, UK) and three others (Canada, Japan, US). Although the sample of countries we use is dictated

purely by data availability, the inclusion of a number of Euro area countries sheds light on the impact of the formation of the Euro area on the nature of inflation in these countries.

All inflation series are calculated by differencing the logged monthly indices and multiplying by 100. The underlying monthly CPI values for both the aggregate and components are obtained from the OECD Main Economic Indicator database. To account for seasonality in the computed monthly inflation series, we undertake a prior seasonal adjustment for each inflation series using the widely applied X-12-ARIMA seasonal adjustment procedure[9].

For the construction of foreign inflation, the method described in subsection 1.3.4 of chapter 1 is employed based on the bilateral trade statistics of 13 OECD countries. Appendix Table B.1 shows bilateral trade weights for these countries which are averaged over the sample of approximately 40 years. Similar to bilateral trade weights computed for 19 OECD countries, in general, Germany is the most important trading partner for European countries, while the US is for the non-European countries of Japan and Canada. However, the UK does not have a dominant trading partner, although shares with respect to the US, Germany and France are relatively large compared to others. We should note, however, that these weights based on bilateral trade do not reflect effects of third-countries, such as the large emerging economies of China and India. However, the limitation of available data for those countries precludes their inclusion in our analysis.

2.4 Changes in Inflation Linkages

This section presents a summary of the results. Subsection 2.4.1 examines the break points which are detected in the coefficients and residual variances of aggregate inflation and its components. Subsection 2.4.2 studies the source(s) of changes by comparing estimates of the contemporaneous coefficients of foreign inflation across the aggregate and component analyses. Changes in the mean, persistence and volatility characteristics of inflation are then discussed in the remainder of the section.

As discussed in section 2.2, the dynamics employed in the general specification of (2.1) are selected within our iterative structural break testing procedure and Table 2.1 reports the resulting lags for each series (namely, aggregate, core, energy and food inflation) at the end of the iterations. Aggregate models pre-

[9]Official seasonally adjusted data is available for the US and our graphical comparison of this series with the comparable unadjusted series filtered using X-12-ARIMA showed these to have very similar properties. Hence we apply X-12-ARIMA seasonal adjustment to series for all countries.

sented in Table 2.1 are largely similar to those selected in chapter 1 (using more countries to construct the foreign inflation series) for the same countries. All later results using (2.1) are obtained using the models reported in this table; the domestic aggregate with foreign components model refers to (2.3) and is discussed in the next section.

It is noteworthy, in Table 2.1, that more domestic lags are selected in the aggregate and (particularly) core inflation models of (2.1) as compared to energy and food inflation. This is not surprising, due to the higher volatility of energy and food inflation those are less dependent on their past; indeed, for the majority of countries energy inflation has no persistence in the sense that the selected models include no domestic lags. Interestingly, more lags are selected for core than aggregate inflation. Another informative characteristic of Table 2.1 is that contemporaneous foreign inflation is always selected for the aggregate and also for both energy and food inflation, whereas in the clear majority of cases it would not be selected when the model is applied to core inflation. Indeed, for Japan and Switzerland, no foreign (contemporaneous or lagged) value would be selected at all for core inflation. These results suggest that short-run inflation linkages may apply primarily through energy and food inflation, rather than core inflation. Nevertheless, since our main focus is inflation linkages, the contemporaneous foreign value is always included when (2.1) is estimated and subject to structural breaks analysis.

2.4.1 Number and dates of breaks

Break dates detected in the coefficients and variances for (2.1) are shown in Table 2.2, with these presented for both (2.1) and (2.3); results for the latter are discussed in Section 2.5. The corresponding test statistics relating to (2.1) are provided in Appendix Tables B.2 to B.4. More specifically, Tables B.2 and B.3 provide detailed results for the model (2.1) applied to the aggregate inflation series, while (to conserve space) Table B.4 gives results for the $WDmax$ statistic applied to the core, energy and food inflation models; all p-values are obtained using the method of Hall and Sakkas (2013). All results shown here are at the end of the iterative procedure of Section 2.2[10] [11]. A glance at the

[10]For information, all sequential F-statistics are shown to $Sup(5|4)$, although the procedure stops when the relevant null hypothesis is not rejected at 5%.

[11]For all countries, with the exception of France, break dates detected in the coefficients and variances for aggregate specification of (2.1) converges to a unique set of coefficient and variance breaks in the iterative structural break testing methodology. For France, convergence is achieved by choosing the set which achieves the smallest SIC among other local optima, using (2.2). Similarly, in some cases break dates are selected based on (2.2) for core (Denmark, Germany, Netherlands, the UK), energy (Canada, Italy) and food (Canada, Finland) model of (2.1).

p-values of the $WDmax$ test statistics in these appendix tables indicates that there is usually strong evidence for breaks in both the coefficients and variances. This last statement applies particularly for the aggregate series, where the p-values of Table B.2 indicate strong evidence for breaks, with failure to reject at the 1 percent level only for the coefficients of the Austria model (p-value 0.019). However, these tables also indicate that the number of breaks is not always clearcut; for example, the choice of one rather than two breaks for aggregate inflation in Italy is based on the marginal p-value of 0.052 for the $SupF(2|1)$ statistic; see Table B.2. On the other hand, when core, energy and food inflation are considered separately, statistical evidence of breaks is not always uncovered (Table B.4).

Overall, coefficient and volatility breaks are found in aggregate inflation for all 13 countries, with coefficient breaks in 11, 12 and 10 countries for core, energy and food inflation, respectively (using a 5 percent significance level). Across the 13 countries, we find more breaks in total for the energy series (19 breaks) than others (16 for food, 15 each for core and aggregate inflation). Although a maximum of five breaks are allowed, the most uncovered in any series is three coefficient breaks and (with the exception only of aggregate inflation in Japan) two is the maximum number of volatility breaks. Hence, although we use four decades of monthly data, structural breaks occur relatively infrequently in either the coefficients or volatility of these international inflation relationships. However, our finding of fewer breaks overall than univariate studies that examine the mean level of inflation (such as Benati, 2008, Bataa et al., 2013b) can be attributed to our methodology that explicitly includes dynamics in the breaks analysis in order to avoid the oversizing of HAC methods, as discussed in Section 2.2 above.

Some clustering of coefficient break dates can be observed, with seven countries having estimated breaks in aggregate and/or core inflation between 1980 and 1982. In some cases (Denmark, Germany, the US) this is preceded by a break in the relationship energy inflation relationship; for example, Hooker (2002) documents a substantial role of oil prices in explaining US core inflation prior to 1981. More broadly, the clustering of these breaks in both the aggregate and energy inflation series around the first half of the 1980s point to the importance of energy prices for historical changes in inflation. On the other hand, breaks dated for Canada, Sweden and the UK in the early 1990s may be associated with the introduction of inflation targeting in those countries (in February 1991, January 1993 and October 1992 for Canada, Sweden and the UK, respectively), while a number of current members of the Euro area (specifically Finland, Germany, Italy, and the Netherlands) show coefficient breaks

in the decade leading to the launch of the common currency; these results are in line with Altissimo et al. (2006), Cecchetti and Debelle (2006), Bataa et al. (2013b) and others. The variance breaks in Table 2.2 cluster mainly in the decade from 1979 (32 in total) and far fewer occur at other periods, in contrast to the coefficient breaks where many occur in the early 1990s.

Corresponding to the regimes implied by the estimated break dates of Table 2.2, the corresponding coefficient estimates are summarized in Table 2.3. More specifically, this latter table gives the contemporaneous coefficient of foreign inflation $\widehat{\beta}_{0j}$ over coefficient regimes $j = 1, ..., m + 1$ and the corresponding sum of the lagged global coefficients $\sum_{i=1}^{r} \widehat{\beta}_{ij}$, together with the sample mean of the univariate inflation series and persistence measured by $\widehat{\rho}_j^d = \sum_{i=1}^{p} \widehat{\alpha}_{ij}$ for $j = 1, ..., m+1$. Finally, volatility is the residual variance obtained in (2.1) for each identified volatility regime $k = 1, ..., n + 1$. ; These results are discussed in the next three subsections.

2.4.2 International co-movement

Our findings on changes in international inflation linkages are summarized by Figures 2.1 and 2.2, which plot the estimated contemporaneous and (summed) lag coefficients, respectively over regimes for the 13 countries of our sample. In each figure, the values in these countries are presented over three panels: panel A (includes Austria, Canada, Denmark, Finland), panel B (includes France, Germany, Italy, Japan), and panel C (includes Netherlands, Sweden, Switzerland, United Kingdom, United States) for the readability of individual country's dynamics. More specifically, each panel consists of graphs a, b, c and d which correspond to aggregate, core, energy and food inflation, respectively, showing the break dates of Table 2.2 and the estimated coefficients of Table 2.3 for the model of (2.1). Note, however, that lagged foreign inflation is not always included in the models, and hence fewer than 13 countries are shown in each graph of Figure 2.2.

Graph a (in each panel) of Figure 2.1 shows a clear pattern of increasing contemporaneous co-movement for aggregate CPI inflation, with many countries showing an increased role for foreign inflation from the 1980s or early 1990s. This applies for most European countries, including France and Germany (though not the UK), and also for Canada. However, the substantial increase is dated to occur later for the US (in 2004). Therefore, our model applied to aggregate inflation reproduces the pattern of inflation globalisation documented in other studies (including Ciccarelli and Mojon, 2010, Neely and Rapach, 2011, Bataa et al., 2013a). This phenomenon apparently applies in

the very short-run, with graph a of Figure 2.2 showing no evident tendency for lagged foreign inflation to play a greater role over time. However, the pattern of increased contemporaneous co-movement for aggregate inflation is not reproduced when core inflation is examined in graph b of Figure 2.1. In particular, this shows the contemporaneous foreign coefficient to be relatively constant over time for core inflation and of smaller magnitude than that for aggregate inflation (graph a). Indeed, the estimated contemporaneous foreign coefficient for core inflation declines in a number of cases; for example, that for the UK changes from 0.32 to -0.13 after November 1990, around the time of the introduction of inflation targeting (see Tables 2.2 and 2.3). The corresponding coefficient for Canada increases to 0.25 with inflation targeting, which may reflect a closer alignment of its level to that of the US. The six Euro area countries present no substantial evidence that foreign core inflation (which Appendix Table B.1 shows to be influenced strongly by inflation in other European countries) plays a greater role with monetary integration, except that the relevant coefficient for Italy reverts from effectively zero to 0.48 from the mid-1990s. This conclusion is not substantively changed when the coefficients of lagged foreign core inflation are examined in Figure 2.2, albeit there are some individual cases (such as Sweden) where an increased role is indicated.

The implication is that non-core inflation elements largely drive the co-movement seen in aggregate inflation, and indeed graphs c and d of Figure 2.1 show similar characteristics in this respect to graph a. With the exceptions only of Japan (where foreign energy inflation apparently plays no role) and the UK, energy price inflation in all countries shows greater exposure to foreign movements over the last decade compared with the 1970s. The large estimated contemporaneous energy coefficient for the US from 1993 is particularly notable, with this being numerically very similar to the estimated foreign coefficients for the small countries of Finland and Switzerland in the latter part of the sample. With the exceptions only of Austria, Japan and Switzerland (where no breaks are uncovered), the foreign coefficients for food inflation are also larger at the end of the sample than in the early period. For instance, France and Germany experience an increased foreign food inflation role from around 1990, and this may be due to greater currency integration causing prices movements from other member states to be transmitted more fully to these large countries.

Our results based on (2.1) therefore suggest that short-term movements in national core inflation are primarily domestically determined and the role of international inflation has not increased overall during the last four decades.

This applies across countries, irrespective of whether they are members of the European monetary union or not. The increased co-movement seen in aggregate CPI inflation since the 1980s appears to be due, therefore, primarily to the components of energy and food, which do generally exhibit evidence of increased foreign effects; we will return to this issue through the model (2.3) in Section 2.5 below. Meanwhile, in subsection 2.4.3 we examine the impact of structural breaks on the mean and persistence of inflation in the context of (2.1).

2.4.3 Mean and persistence

The discussion of the previous subsection relates to co-movement across countries for monthly CPI inflation. However, breaks dated in (2.1) may refer to level shifts or changes in persistence, which are discussed here and depicted graphically in Figures 2.3 and 2.4. In common with earlier figures, all 13 countries are presented over panels A, B, C and graphs a to d in each panel show regime-specific values relating to aggregate, core, energy and food inflation. The mean inflation levels in Figure 2.3 are computed as the sample mean for the univariate inflation series over the indicated regimes, while persistence is measured as the sum of the AR coefficients in (2.1) and hence is conditional on the respective foreign inflation series; see also Table 2.3.

Any globalisation of inflation does not necessarily refer to month-on-month movements, but can be interpreted as longer term characteristics. With such an interpretation, graph a of Figure 2.3 presents strong evidence that all 13 OECD countries have effectively converged to a common inflation level of about 0.2 percent a month. Further, this common inflation level largely applies also to core inflation (graph b) since the late 1990s. The pattern is clear in both graphs, with each break being associated with a lowering of the respective inflation measure in each country, except for a temporary increase for core inflation in Italy between 1980 and 1986. In addition to the general decline of the early 1980s after the second world oil price shock, the declines for Canada and the UK in the early 1990s effectively coincide with the introduction of inflation targeting, while Euro area countries experience breaks in the run-up to full monetary integration (see also subsection 2.4.1). Italy is an interesting example of the last statement, where the inflation decline occurs in December 1995 and brings Italy's inflation in line with the requirements of the Maastricht Treaty[12].

[12]The formal Maastricht Treaty requirement is for annual inflation to be no more than 1.5 percent higher than the average of the three lowest inflation countries of the European Union.

As indicated by Table 2.3, with the single exception of the first break for core inflation in Italy, coefficient breaks in the relationships for aggregate and core inflation are associated with declines in the mean level, which is computed as the sample mean for the univariate inflation series over the indicated regimes. While the mean of energy inflation also declines in all countries from its initially high level of the 1970s, in some cases it increases again in the latter part of the sample period (June 2004 for Germany and Japan, and during the 1990s for Austria, Netherlands and Switzerland). The level of food inflation also shows a pattern of decline over time, with the corresponding breaks generally dated between the mid-1970s and mid-1980s. A summary measure of the extent of the decline in mean inflation is given by the ratio of the average values across countries of mean inflation for the first versus last coefficient regimes, yielding 3.5, 4, 3 and 4.5 for aggregate, core, energy and food inflation, respectively.

In line with results for univariate inflation models (including Benati, 2008, Bataa et al., 2013b) the persistence of aggregate inflation in (2.1) declines over time. As indicated in graph a of Figure 2.4 (see also Table 2.3), every country except German (which has very low persistence throughout) shares in the decline, with inflation persistence in all countries being 0.34 or less by the end of the sample. However, as for co-movement in Figure 2.1, the pattern for core inflation in graph b differs from the aggregate. More specifically, Denmark, Finland, Germany, the UK and Italy (from 1980) show persistence for core inflation to be effectively constant or to increase over time, with persistence typically higher for core than aggregate inflation in the final coefficient regimes.

Our models find little change over time in persistence for energy inflation, with this often being zero as no lagged dependent variable is required (Table 2.1). Similarly, the model for food inflation requires no lags in a number of cases, while persistence substantially declines in others (notably France, Italy and the UK). These findings support those of Altissimo et al. (2006), who note less persistence for non-processed foods and energy and higher persistence for industrial goods and services.

2.4.4 Volatility

An important but often overlooked feature of inflation is volatility, which is plotted in Figure 2.5 for each of the four series we analyse (but note the different scales used here for energy and food inflation versus aggregate and core) in the context of the model of (2.1). It is clear that energy inflation is more volatile than core or food inflation, with a variance of around 1.5-2

percent squared per month. Aggregate and core inflation are relatively smooth, with variances typically less than 0.5 percent squared variance per month from the early 1980s onward, while food inflation is intermediate, where volatility of less than 1 percent squared, with the exception of Italy and Japan.

The volatility reductions around the early 1980s in Figure 2.5 appear to be a manifestation of the international dimension of the so-called Great Moderation, which a number of studies link to improved monetary policy; see, for example, the discussion in Summers (2005). The volatility reductions apply particularly to aggregate and core inflation, with all countries except Austria and the Netherlands experiencing volatility declines for at least one of these. However, if these are due to improved monetary policy, it is surprising that the substantial policy changes of the 1990s in relation to the introduction of inflation targeting in Canada and the UK and the movement to the euro currency in Europe apparently had little impact on either aggregate or core inflation volatility, with the notable exception of the reduction in the volatility of core inflation in the UK in August 1992.

In general, energy inflation is volatile until the mid 1980s, with a somewhat mixed picture of declines and increases subsequently. On the other hand, the volatility of food inflation typically declines, but is generally stable from the mid-1990s onwards.

2.4.5 Discussion

The results just presented based on the model (2.1) shed light on the nature of the globalisation of inflation. More specifically, the cross-country convergence of mean rates of aggregate, core and food inflation in Figure 2.3 is striking. In contrast, energy inflation does not show such a pattern of mean convergence. Also, while Figure 2.1 indicates increased contemporaneous co-movement for aggregate CPI inflation, this is less clear for the components. In particular, core inflation provides less strong evidence of increased short-run co-movement than does the aggregate at the monthly frequency which we examine. Therefore, even though the model for aggregate inflation finds increased contemporaneous co-movement, the components models, and specifically that for core inflation, indicate the role played by the long-run downward shift in the mean level. Table 2.2 shows that estimated break dates may differ over components, allowing the possibility that long-run convergence in the sense of an effectively common mean core inflation level from the early 1990s could appear as short-run co-movement when the aggregate series is examined.

Previous authors, including Neely and Rapach (2011) and Mumtaz and

Surico (2012), suggest that monetary policy plays a role in explaining the globalisation of inflation, and our results support this in terms of the level of inflation. In particular, although the dates of structural change are estimated in terms of the domestic-foreign relationship of (2.1), those for core inflation in each case precede by a short period the introduction of inflation targeting for each of Canada, Sweden and the UK. While we cannot, therefore, necessarily attribute the break to this monetary policy change, nevertheless inflation targeting may help to keep inflation expectations at this new lower level and consequently make the downward shift permanent. It is also noteworthy that breaks for the Euro area countries of Finland, Italy and the Netherlands are dated in the run-up to the launch of the common currency. On the other hand, in line with the view that Euro area monetary policy largely follows that of Germany over the earlier period, Germany experiences no break in its core inflation relationship at this time.

On the other hand, the US does not have an explicit inflation target, but here also monetary policy is recognised as having changed after the appointment of Paul Volker as chairman of the US Federal Reserve in 1979; see, for example, Orphanides (2004). Although our model finds a break in the coefficients of the core inflation model for the US in 1980, with a variance break also in 1983, the only breaks in the aggregate series are dated in 2003/4. This suggests that aggregation, and specifically counteracting directions of change in the contemporaneous coefficients for foreign energy and food inflation in the early 1980s may have obscured the closer alignment of US core inflation with foreign values from the 1980s when the headline series is examined. Further, the increase in the volatility of US headline inflation in the early 2000s appears to be due to a substantial increase in the volatility of energy inflation.

Indeed, Table 2.3 suggests that energy and food inflation may play a role in explaining the globalisation of inflation. Recalling that no lagged foreign inflation is typically selected when (2.1) is specified for the energy series, it is striking that the contemporaneous foreign coefficient is higher at the end of the sample than at the beginning (in the 1970s) for every country we study, with the exceptions only of Japan, Sweden and the UK. Further, now with the exception of Austria, Japan and Switzerland, the role of contemporaneous foreign food inflation increases in that model.

The results of this section, therefore, support the proposition that the adoption of similar monetary policies (although not necessarily in the form of explicitly inflation targeting) across countries plays an important role in the globalisation of inflation, by effectively bringing the levels of core inflation into line across countries. This long-run aspect may also explain why Ciccarelli

and Mojon (2010) find that inclusion of a long run error-correction to foreign inflation improves the accuracy of domestic forecasting models. In terms of short-run movements, however, the individual models suggest that energy and food inflation may largely drive the increased synchronicity of monthly movements.

2.5 Decomposing International Co-Movements

Further insight into the nature and causes of changes in international inflation linkages are provided by the model of (2.3). More specifically, this focuses on CPI inflation at the aggregate level, since the globalisation or co-movement of inflation is documented using data at the aggregate level in the studies of Ciccarelli and Mojon (2010), Neely and Rapach (2011), Mumtaz and Surico (2012), Bataa et al. (2013a) and others. However, to shed light on the nature of this co-movement and how it changes over time, (country-specific) foreign inflation is decomposed into core, energy and food components.

As preliminary to the substantive results, consider the lags selected by our SIC procedure (with a maximum of 12 lags considered for both the lagged domestic inflation variable and each of the foreign components) in the context of (2.3) and shown in the final columns of Table 2.1. With the contemporaneous values treated in the same way as any individual lag, current foreign core inflation is selected in only three of the 13 cases, which may imply that this does not play a strong role overall in explaining movements in domestic CPI inflation. Foreign core inflation may nevertheless still play a role, with a (single) lagged value typically selected, although neither contemporaneous nor lagged foreign core inflation would be included in the models for either Japan or the US. The situation for foreign energy inflation stands in contrast to this, with the contemporaneous value always selected and a lagged value only for Germany. Finally, contemporaneous foreign food inflation is selected for a small majority of countries (7 of 13), while for Germany, Switzerland and the UK no contemporaneous or lagged value is selected. However, to facilitate comparison with the results discussed above for the model of (2.1), contemporaneous values of each foreign series are included in the specification of (2.3) when this is employed for structural break testing. Although there are some differences, the broad pattern of domestic lags selected for (2.3) is broadly similar to those for the model (2.1) estimated for aggregate CPI inflation.

As discussed in section 2.2 above, two sets of break dates are employed in this analysis decomposing foreign effects, namely those obtained from the model (2.1) applied to aggregate inflation and those breaks estimated from

application of our iterative procedure[13] directly to the model of (2.3); both of these are included in Table 2.2, with detailed structural break test results for the latter models included as Appendix Tables B.5 and B.6. It is notable that direct use of (2.3), including all contemporaneous foreign components, sometimes leads to the identification of two coefficient breaks rather than one revealed for aggregate inflation in (2.1). While there are differences, nevertheless a broad correspondence can generally be seen between these two sets of coefficient break dates. Not surprisingly, the coefficient break dates obtained directly from (2.3) also sometimes line up with those estimated for food or inflation in the context of (2.1). The variance breaks are quite clearcut, in the sense that the sets for aggregate inflation and for the aggregate-foreign components models in Table 2.2 are very similar overall.

Conditional on these break dates, Table 2.4 presents the estimated coefficients for foreign inflation in each regime together with p-values for the significance of changes across regimes, namely p-values for tests of the null hypotheses of (2.4) and (2.5) for contemporaneous coefficients, together with (2.6) and (2.7) for lag coefficients. To facilitate comparison, the first line of results in Table 2.4 for each country repeats the estimates from the aggregate inflation model of Table 2.3, but now including the corresponding p-values for the null hypothesis (2.4) or (2.6) applied in the aggregate model of (2.1). For example, the increased contemporaneous coefficient (from 0.23 to 0.66) in the aggregate model for Canada is significant at 1 percent. This can be attributable to the increases in responses of domestic inflation to foreign core, energy and food components, as their observed changes are significant at 1%, 1% and 10%, respectively. Consistently, the joint hypothesis of no change in the contemporaneous coefficient across components is also rejected at 1 percent significance level. As discussed in subsection 2.2.3, the estimated coefficients for foreign inflation components partly reflect the differing weights on components, so that those on energy and food are anticipated to be smaller than on core, with the latter reduced to a lesser extent from those of the aggregate equation, even when the same coefficients apply across the models of (2.1).

Seven countries have contemporaneous foreign aggregate inflation coefficients that exhibit significant change over regimes in (2.1), and the decomposition of (2.3) indicates the nature of change. Thus, the increases in the contemporaneous foreign inflation coefficients of the aggregate models for Italy, Sweden and Switzerland are associated with increases in the responses of do-

[13]For Finland and Italy only, our iterative methodology of structural break testing in the context of component model (2.3) identifies more than a single set of coefficient and variance breaks in which the iteration cycles among those, and thus the convergence is achieved based on the smallest SIC among them, using (2.2).

mestic inflation to foreign core inflation, with no significant changes in the responses to foreign energy or food inflation[14]; note, however, the overall null hypothesis of equal and unchanged contemporaneous coefficients (2.5) is not rejected for Italy in this specification. There is also some indication of an increase in the response to foreign core inflation for Germany in Table 2.4 when break dates are imposed from the aggregate specification results, but an increased response to energy inflation at the first break is apparently more important, while inflation in France exhibits increased responses to both foreign energy and food inflation. Finally, foreign core and energy inflation contribute to the increased foreign effects for Canada, while for the US it is attributed to an increased response to foreign energy, as the coefficient on foreign core inflation declines. It is noteworthy that, with the exception of Italy, the equality joint null hypothesis of (2.5) is also rejected for these seven countries in the context of the model of (2.3).

There is much less evidence of change when lagged foreign coefficients are examined, in the context of either model (2.1) or (2.3), when the break dates are imposed from the former. Indeed, where change is significant in the context of the latter, the foreign component coefficients generally increase.

The second set of results in Table 2.4 are based on break dates estimated in the context of the model of (2.3), and these are largely compatible with those obtained using breaks imposed from the aggregate inflation model of (2.1). Note, however, that although the quoted p-values are asymptotically justified by the analysis of Bai and Perron (1998), they may over-state significance in a finite sample context, especially for a joint test applied to (2.3) when the break dates are also endogenously determined within this model. Hence, it is unsurprising to find greater apparent significance of change compared to when the break dates are imposed from (2.1). With this caveat, the results nevertheless agree with the first set that changes relating to energy inflation are important for increased contemporaneous co-movement of inflation in Canada, France, Germany and the US, while those for core inflation play a role for some European countries, namely Germany, Italy and Sweden[15]. Only for France changes are relating to food inflation significant at 5 percent in both sets of results, with the overall null hypothesis of constant co-movement, (2.5), also rejected. Nevertheless, irrespective of whether the change is judged to be significant for an individual country or not, with endogenous dating of breaks in

[14]Care is required in interpreting the magnitudes of the coefficients for global energy and food inflation in the context of (2.3), since these make smaller contributions to aggregate inflation than does core inflation.

[15]While the contemporaneous coefficient on global core inflation declines for Denmark around 1990, it is noteworthy that this country is not a member of the Euro area.

the context of (2.5), Table 2.4 provides evidence that transmission of external food inflation to domestic aggregate inflation has become more important for all European countries since around the mid-1980s, except for Denmark and the Netherlands (the break for the latter is dated in 1977).

As noted above, contemporaneous core inflation would often not be selected in the context of the aggregate-foreign component model (2.3), so that changes in the corresponding lag coefficients may be informative. However, while such change is significant for Austria, Germany, Italy and the UK, the corresponding coefficients often decline.

The general pattern of persistence decline for aggregate inflation seen in Table 2.3 also applies in Table 2.5, where the latter is calculated for the decomposed foreign inflation model of (2.3) with breaks endogenously dated in the context of that model. There are, however, some interesting exceptions to the pattern of decline. Persistence changes significantly in seven countries, with the predominant direction being decline. However, it is effectively constant for Italy, the Netherlands and the UK , while that for the US is initially effectively zero in the early part of the 1970s, increases and then declines again to around 0.3. The overall decline in mean aggregate inflation seen in Figure 2.3 can also be associated with the significant declines in the intercept seen for a number of countries in Table 2.5.

2.6 Conclusions

This paper sheds new light on the nature of the globalisation of inflation by examining the separate roles of core, energy and food inflation in the increased synchronicity of monthly CPI inflation across OECD countries. To do so, we analyse changes in the linkages of domestic and country-specific foreign inflation (the latter constructed as a trade-weighted average) using an iterative methodology that allows for breaks in both coefficients and disturbance variances.

It is, perhaps, not surprising that our analysis reveals the importance of more than one feature. In terms of long-run movements, the apparent convergence in the mean levels of core inflation from the early 1990s indicates an important role for monetary policies that focus on inflation. This includes Euro area countries, such as Finland and Italy, as well as others that introduced inflation targeting during that decade (Canada, Sweden and the UK). There is, however, less strong evidence of globalisation in terms of short-run movements (contemporaneous relationships) of core than for aggregate inflation. Perhaps, these results support the proposition that the apparent globalisation in aggre-

gate inflation may due to the conduct of independent but similar monetary policy rules rather than the coordination across countries.

However, it is not easily distinguishable whether the convergence in the mean of core and aggregate inflation is caused mainly by the effectiveness of improved monetary policies or relatively stable global economic conditions such as reduced common shocks hitting the economies. It is also possible that these two explanations are jointly responsible for the (low and stable) mean convergence of inflation in a sense that a drop of exogenous shocks to the economies enables monetary policies to control inflation effectively (Ahmed et al., 2004, Summers, 2005). Nevertheless, short run movements of non-core elements (energy and food) suggest that they may largely drive the increased synchronicity of monthly movements, and thus their important roles in explaining globalisation of inflation.

Further examination based on the aggregate-foreign component models reveals an increased role for foreign energy inflation in explaining the apparent globalisation of domestic CPI inflation in both the US and Canada. This is a particularly interesting finding in the light of the discussion in Bernanke (2007) about the role of domestic monetary policy in the context of the globalisation of inflation. Indeed, domestic US monetary policy appears to be successful in the sense that our aggregate-foreign component model (2.3) finds no role for movements in foreign core inflation after the 1970s and no role for foreign food inflation from 2003. Therefore, the evidence indicates that US inflation is primarily domestically determined, except in respect of energy price movements. As a side-product, these findings also reduce the concern noted in subsection 2.2.1 about the possible endogeneity of foreign inflation for the US.

Tables

Table 2.1: Lags selected in domestic/foreign inflation models

Country	Aggregate Inflation		Core Inflation		Energy Inflation		Food inflation		Domestic Aggregate			
	Domestic	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic	Foreign	Domestic	Foreign Core	Foreign Energy	Foreign Food
Austria	11	0,4	10,11	0*,2	NA	0,1	NA	0	7	0,4	0	0
Canada	5,7,9	0,9	2,6,9,11	0*,1	10	0	NA	0,1	7	0*,1	0	0,9
Denmark	1,6,11	0	1,3,6,10	0	NA	0,2	NA	0,1	1,6	0*,10	0	0
Finland	2,9	0,4	2,3,7	0,5	NA	0	NA	0,4	2,9	0,6	0	0*,3
France	1,3,10	0	1,2,8	0,6	1	0	1,4,7,9	0,3	1,3,8,10	0	0	0
Germany	9	0	2,3,8	0*,1	NA	0,9	1	0	3,6	0*,1	0,9	0*
Italy	1,3,6	0,2	1,2,3,9	0*,2	NA	0,1	1	0	1,3,6	0*,2	0	0
Japan	3,5,9,11	0,1	2,3,4,5,6,8	0*	1,7	0,1	7,9,11	0,4	1,5,9,11	0*	0	0*,1
Netherland	4,6,8	0	4,5,6,9	0	NA	0,10	5,10	0	4,6,7,8	0*,1	0	0
Sweden	7,8,9	0	2,7,8,11	0*,2	NA	0	2,5,8	0	7,8,9	0*,5	0	0*,1
Switzerland	1,6,9	0	1,2,6,8,9	0*	NA	0	7	0	1,6,9,10	0*	0	0*
UK	1,2,3	0	1,2,5	0*,6	1,8	0	1,8	0,5	1,2,6	0*,7	0	0*
US	1,7	0	1,2,8	0*,4	1,11	0	1,9	0	1,6,11	0*	0	0,8

Notes: The aggregate, core, energy and food inflation models refer to (2.1), while the domestic aggregate with foreign components model is given by (2.3). Individual lags are selected to a maximum of $p = r = 12$. NA indicates that no lagged domestic values are selected. * indicates the contemporaneous value (lag 0)of foreign inflation is not selected, but is included in the estimated model.

Table 2.2: Estimated dates of structural breaks

		Breaks in coefficients				Breaks in variances			
Country	**Model**	**1970s**	**1980s**	**1990s**	**2000s**	**1970s**	**1980s**	**1990s**	**2000s**
Austria	Aggregate inflation	.	1981-May	.	.	1977-Jan	.	.	.
	Core inflation	.	.	.	.	1977-Jan	.	.	.
	Energy inflation	1976-Apr	.	1999-Aug	.	.	.	.	.
	Food inflation	.	.	.	.	.	1984-Oct	.	.
	Aggregate-foreign components	1976-Sep	1983-Jun	.	.	1976-Sep	.	.	.
Canada	Aggregate inflation	.	.	1990-Dec	.	1978-Nov	.	.	.
	Core inflation	.	1982-Nov	1991-Sep	.	.	1981-Feb	.	.
	Energy inflation	.	.	1991-Feb	.	.	.	1992-Jul	
	Food inflation	1976-Dec	1987-Jan	.	.	1978-Oct	1986-Sep	.	.
	Aggregate-foreign components	.	1982-Nov	1991-Jan	.	.	1981-Feb	.	.
Denmark	Aggregate inflation	.	1982-Nov	.	.	.	1980-Oct	1990-Jan	.
	Core inflation	.	1988-Mar	.	.	.	1987-Apr	.	.
	Energy inflation	.	1981-Jun	.	.	.	.	1990-Dec	.
	Food inflation	.	1985-Mar	.	2001-Sep	1976-Jul	1984-Mar	1991-Jan	.
	Aggregate-foreign components	1977-Nov	1989-May	.	.	.	1985-Apr	1990-Dec	.
Finland	Aggregate inflation	.	.	1990-Mar	.		1983-Jul	.	.
	Core inflation	.	.	1991-Jun	.	.	1983-Jul	1994-Feb	.
	Energy inflation	.	.	1991-Mar	.	.	.	.	.
	Food inflation	.	1985-Jun	.	.	.	1982-Oct	.	.
	Aggregate-foreign components	.	1989Nov	.	.	.	1983-Jul	.	.
France	Aggregate inflation	.	1985-Aug	.	.	.	1983-Jan	.	.
	Core inflation	.	1987-Mar	.	.	.	1983-May	.	.
	Energy inflation	.	1985-May	.	.	.	.	.	.
	Food inflation	.	1980-Mar 1989-Dec	.	.	.	.	1990-Mar	.
	Aggregate-foreign components	.	1987-Jan	.	.	.	1983-May	.	.
Germany	Aggregate inflation	.	1980-Nov	1991-Apr	.	.	1982-Jul	.	.
	Core inflation	1976-Oct	.	.	.	.	.	.	.
	Energy inflation	1979-Feb	.	1990-Oct	2004-Jun	.	.	1991-Nov	.
	Food inflation	1977-Oct	.	1991-Mar	.	1977-Feb	.	.	2000-Jan
	Aggregate-foreign components	.	1980-Feb	1991-Apr	.	.	1982-Jul	.	.
Italy	Aggregate inflation	.	.	1996-May	.	.	1981-Jan 1986-Jul	.	.
	Core inflation	.	1980-Feb 1986-Feb	1995-Dec	.	1976-Dec	1982-May	.	.
	Energy inflation		.	1995-Sep	.		1986-Jun	.	.
	Food inflation	1976-May	.	.	.	.	.	.	.
	Aggregate-foreign components	1976-Aug	1983-Feb	.	.	.	1982-Sep	.	.
Japan	Aggregate inflation	1977-Jan	.	.	.	1977-Jan	1985-Jan	1993-Jun	.
	Core inflation	.	.	.	.	.	1980-Feb	.	.
	Energy inflation	.	.	1997-Apr	2004-Jun	.	1987-Jun	.	2004-Feb
	Food inflation	.	.	.	.	1977-Jan	.	1998-Mar	.
	Aggregate-foreign components	1977-Jan	.	.	.	1977-Jan	1985-Dec	1993-Dec	.

Table 2.2 continued

Country	Model	Breaks in coefficients				Breaks in variances			
		1970s	1980s	1990s	2000s	1970s	1980s	1990s	2000s
Netherland	Aggregate inflation	.	1989-Apr	.	.	1979-Feb	.	.	.
	Core inflation	.	.	1998-Feb	.	1978-Aug	.	.	.
	Energy inflation	.	1983-Sep	1993-Oct	.	.	1988-Nov	.	.
	Food inflation	1976-Dec	.	1992-Jan 1998-Jan	.	.	.	.	.
	Aggregate-foreign components	1976-Dec	.	.	.	1977-May	.	.	.
Sweden	Aggregate inflation	.	.	1991-Feb	.	1977-Jul	.	1992-Apr	.
	Core inflation	.	.	1991-Feb	.	.	1983-Mar	.	.
	Energy inflation	.	.	.	.	.	1985-Dec	.	.
	Food inflation	.	.	1991-Feb	.	.	1987-Sep	.	.
	Aggregate-foreign components	.	.	1990-Aug	.	.	1985-Jun	1993-Jan	.
Switzerland	Aggregate inflation	.	1984-Oct	.	.	.	1983-Feb	.	.
	Core inflation	.	.	1993-Nov	.	1976-Nov	1985-Feb	.	.
	Energy inflation	.	1982-Feb	1991-Jun	.	.	1983-Feb	.	.
	Food inflation	.	.	.	.	1979-Aug	1985-Feb	.	.
	Aggregate-foreign components	.	.	1993-Sep	.	1976-Dec	1983-Apr	.	.
UK	Aggregate inflation	.	1980-May	1991-Jul	.	.	1982-Jul	.	.
	Core inflation	.	1980-Feb	1990-Nov	.	.	1980-Jun	1992-Aug	.
	Energy inflation	.	.	.	.	.	.	.	.
	Food inflation	.	1984-Jun	.	.	1976-Nov	.	.	.
	Aggregate-foreign components	.	1980-May	1990-Oct	.	.	1982-Jul	.	.
US	Aggregate inflation	.	.	.	2003-Feb	.	1983-May	.	2004-Mar
	Core inflation	.	1980-May	.	.	.	1983-Feb	.	.
	Energy inflation	.	1980-Jan 1986-Jul	1993-Jul	.	.	.	.	2000-Nov
	Food inflation	.	.	.	2004-Aug	1977-Mar	.	1994-Oct	.
	Aggregate-foreign components	1977-Nov	.	1990-Oct	.	.	1983-May	.	2004-Oct

Notes: Estimated dates of structural breaks are shown for the model of (2.1) applied separately to aggregate, core, energy and food inflation. In addition, the row labelled aggregate-foreign components shows estimated dates of structural breaks for the model of (2.3).

Table 2.3: Estimates within regimes for domestic/foreign inflation models

Country	Series	Contemporaneous				Lagged Foreign				Mean Inflation				Persistence				Volatility			
		R1	R2	R3	R4	R1	R2	R3	R4	R1	R2	R3	R4	R1	R2	R3	R4	R1	R2	R3	R4
Austria	Aggregate	0.55	0.48			0.38	0.11			0.50	0.21			0.19	0.05			0.06	0.03		
	Core	0.16				0.33				0.30				0.24				0.08	0.03		
	Energy	0.13	0.21	0.82		0.01	0.28	0.23		0.72	0.24	0.31		0.00	0.00	0.00		1.11			
	Food	0.57				NA				0.23				0.00				0.40	0.19		
Canada	Aggregate	0.23	0.66			0.30	0.08			0.56	0.17			0.38	0.01			0.09	0.05		
	Core	-0.09	-0.24	0.25		0.35	0.23	-0.03		0.61	0.43	0.15		0.56	-0.05	0.07		0.06	0.02		
	Energy	0.20	0.81			NA	NA			0.71	0.26			0.17	0.05			1.16	3.60		
	Food	0.46	0.99	0.86		0.35	0.46	0.01		0.61	0.70	0.18		0.00	0.00	0.00		0.75	0.45	0.20	
Denmark	Aggregate	0.35	0.46			NA	NA			0.79	0.25			0.41	0.18			0.19	0.07	0.02	
	Core	0.36	0.12			NA	NA			0.66	0.19			0.33	0.50			0.09	0.02		
	Energy	0.58	0.70			0.53	0.03			1.33	0.27			0.00	0.00			1.80	0.63		
	Food	0.51	0.26	0.82		0.30	-0.04	0.37		0.76	0.18	0.16		0.00	0.00	0.00		0.50	0.39	0.22	0.14
Finland	Aggregate	0.58	0.72			0.17	0.15			0.71	0.16			0.25	0.01			0.20	0.04		
	Core	0.36	0.08			0.39	-0.08			0.69	0.14			0.28	0.44			0.20	0.09	0.03	
	Energy	0.25	1.23			NA	NA			0.72	0.30			0.00	0.00			2.44			
	Food	0.17	0.64			0.40	0.09			0.85	0.11			0.00	0.00			0.68	0.23		
France	Aggregate	0.25	0.76			NA	NA			0.77	0.17			0.72	0.00			0.04	0.02		
	Core	0.28	0.25			0.14	0.13			0.70	0.16			0.60	0.34			0.04	0.01		
	Energy	0.24	0.81			NA	NA			0.99	0.13			0.21	0.10			1.14			
	Food	0.14	0.11	0.57		0.12	0.14	-0.02		0.77	0.60	0.14		0.36	0.77	0.16		0.05	0.12		
Germany	Aggregate	0.18	0.86	0.93		NA	NA	NA		0.40	0.23	0.16		0.08	-0.01	0.15		0.05	0.03		
	Core	-0.08	-0.03			0.13	0.27			0.46	0.20			0.40	0.35			0.04			
	Energy	0.17	0.82	1.16	0.66	0.03	0.14	0.31	0.01	0.64	0.29	0.27	0.39	0.00	0.00	0.00	0.00	1.14	0.68		
	Food	0.07	0.33	0.85		NA	NA	NA		0.41	0.18	0.11		0.01	0.14	0.13		0.12	0.05	0.10	
Italy	Aggregate	0.24	0.42			0.20	0.17			0.81	0.18			0.72	0.23			0.14	0.07	0.01	
	Core	0.48	-0.20	0.07	0.48	0.10	1.27	0.17	0.21	0.97	1.23	0.46	0.19	0.78	0.24	0.26	0.30	0.12	0.19	0.02	
	Energy	0.40	0.50			0.2	0.19			0.79	0.21			0.00	0.00			1.08	0.51		
	Food	0.42	0.63			NA	NA			0.72	0.55			0.65	0.17			3.02			

Table 2.3 continued

Country	Series	Contemporaneous				Lagged Foreign				Mean Inflation				Persistence				Volatility			
		R1	R2	R3	R4	R1	R2	R3	R4	R1	R2	R3	R4	R1	R2	R3	R4	R1	R2	R3	R4
Japan	Aggregate	0.33	0.31			0.65	0.19			0.84	0.13			0.32	0.25			0.41	0.14	0.07	0.03
	Core	0.13				NA				0.24				0.82				0.10	0.03		
	Energy	0.09	-0.04	0.08		0.05	0.09	0.15		0.27	-0.03	0.13		0.46	0.32	0.44		0.39	0.10	0.55	
	Food	0.31				0.34				0.23				0.36				1.84	1.00	0.41	
Netherlands	Aggregate	0.44	0.39			NA	NA			0.42	0.18			0.58	0.34			0.06	0.02		
	Core	0.31	-0.18			NA	NA			0.37	0.16			0.65	0.38			0.06	0.02		
	Energy	0.11	0.41	0.61		0.07	-0.06	0.16		0.87	-0.06	0.31		0.00	0.00	0.00		0.23	0.92		
	Food	0.44	0.54	1.94	0.76	NA	NA	NA	NA	0.54	0.19	0.10	0.12	0.39	0.02	0.12	0.38	0.17			
Sweden	Aggregate	0.41	0.78			NA	NA			0.67	0.16			0.34	0.13			0.28	0.16	0.03	
	Core	0.10	-0.07			0.17	1.03			0.64	0.14			0.48	0.10			0.19	0.07		
	Energy	0.70				NA				0.64				0.00				2.45	1.28		
	Food	0.29	0.90			NA	NA			0.73	0.08			0.28	0.08			0.44	0.17		
Switzerland	Aggregate	0.29	0.91			NA	NA			0.40	0.14			0.58	0.22			0.12	0.03		
	Core	0.07	-0.02			NA	NA			0.36	0.07			0.81	0.24			0.09	0.02		
	Energy	0.98	2.54	1.35		NA	NA			0.65	-0.17	0.21		0.00	0.00	0.00		3.59	1.74		
	Food	0.55				NA	NA			0.21				0.14							
UK	Aggregate	0.57	0.58	0.50		NA	NA	NA		1.03	0.53	0.19		0.71	0.31	0.25		0.11	0.03		
	Core	0.37	0.32	-0.13		0.83	-0.16	0.33		0.96	0.64	0.17		0.40	0.45	0.42		0.15	0.06	0.02	
	Energy	0.48				NA	NA	NA		0.59				0.34				1.07			
	Food	-0.01	0.59			0.34	0.24			1.01	0.25			0.61	0.14			0.63	0.21		
US	Aggregate	0.40	1.36			NA	NA			0.40	0.19			0.46	0.17			0.06	0.02	0.04	
	Core	-0.06	0.23			0.17	0.11			0.52	0.30			0.86	0.52			0.04	0.01		
	Energy	0.28	0.00	0.91	1.34	NA	NA	NA	NA	0.74	0.80	0.02	0.30	0.64	0.44	0.08	0.27	0.70	3.07		
	Food	0.44	0.66			NA	NA			0.37	0.22			0.30	0.20			0.46	0.20	0.09	

Notes: Estimated values within regimes (indicated by the prefix R) for the models of (2.1) estimated separately for aggregate, core, energy and food inflation. Values shown for lagged foreign inflation and persistence are shown as the sums of the relevant estimated coefficients. The mean is calculated as the sample mean of domestic inflation given the estimated break dates, while volatility refers to the estimated disturbance variance in (2.1). NA means not applicable, since no corresponding values are estimated

Table 2.4: Estimates of foreign inflation coefficients in aggregate and aggregate-foreign components models

Country	Coefficients	Estimated break dates from aggregate inflation model										Estimated break dates from aggregate-foreign components model									
		Contemporaneous coefficients			Equal-ity *p*-value	Equality joint *p*-value	Lag coefficients			Equal-ity *p*-value	Equality joint *p*-value	Contemporaneous coefficients			Equal-ity *p*-value	Equality joint *p*-value	Lag coefficients			Equal-ity *p*-value	Equality joint *p*-value
		R1	R2	R3			R1	R2	R3			R1	R2	R3			R1	R2	R3		
Austria	Aggregate model	0.55	0.48		0.52		0.38	0.11		0.08		NA	NA	NA	NA		NA	NA	NA	NA	
	Foreign core	0.37	0.17		0.23		0.20	0.18		0.89		0.41	0.43	0.20	0.31		0.08	0.58	0.17	0.02	
	Foreign energy	0.01	0.04		0.18	0.30	0.01	0.01		0.85	0.97	0.01	0.02	0.03	0.57	0.03	NA	NA	NA	NA	0.02
	Foreign food	0.16	0.14		0.85		NA	NA		NA		0.18	-0.04	0.17	0.00		NA	NA	NA	NA	
Canada	Aggregate model	0.23	0.66		0.00		0.30	0.08		0.05		NA	NA	NA	NA		NA	NA	NA	NA	
	Foreign core	-0.05	0.60		0.00		0.38	-0.02		0.06		-0.09	-0.44	-0.06	0.34		0.36	0.38	0.08	0.38	
	Foreign energy	0.00	0.06		0.00	0.00	NA	NA		NA	0.02	-0.01	0.00	0.06	0.00	0.00	NA	NA	NA	NA	0.07
	Foreign food	0.13	0.22		0.08		0.04	-0.08		0.03		0.06	0.12	0.14	0.26		0.12	0.08	-0.01	0.04	
Den-mark	Aggregate model	0.35	0.46		0.67		NA	NA		NA		NA	NA	NA	NA		NA	NA	NA	NA	
	Foreign core	0.15	-0.01		0.57		NA	NA		NA		0.24	0.45	-0.14	0.02		-0.31	0.04	0.04	0.18	
	Foreign energy	0.03	0.06		0.48	0.76	0.09	0.00		0.06	0.15	0.03	0.09	0.07	0.52	0.00	NA	NA	NA	NA	0.18
	Foreign food	0.00	0.14		0.42		0.05	-0.01		0.72		0.07	0.37	0.12	0.19		NA	NA	NA	NA	
Finland	Aggregate model	0.58	0.72		0.24		0.17	0.15		0.62		NA	NA		NA		NA	NA		NA	
	Foreign core	0.47	0.39		0.70		0.30	-0.07		0.05		0.42	0.28		0.50		0.35	0.15		0.26	
	Foreign energy	0.05	0.07		0.47	0.08	NA	NA		NA	0.14	0.03	0.07		0.12	0.05	NA	NA		NA	0.50
	Foreign food	-0.06	0.17		0.02		0.05	0.08		0.75		-0.04	0.16		0.04		0.11	0.15		0.65	
France	Aggregate model	0.25	0.76		0.00		NA	NA		NA		NA	NA		NA		NA	NA	NA	NA	
	Foreign core	0.16	0.29		0.32		0.12	0.08		0.78		0.22	0.32		0.39		NA	NA	NA	NA	
	Foreign energy	0.01	0.06		0.01	0.00	NA	NA		NA	0.02	0.03	0.07		0.02	0.00	NA	NA	NA	NA	NA
	Foreign food	0.05	0.16		0.01		0.05	-0.06		0.01		0.05	0.15		0.02		NA	NA	NA	NA	
Germany	Aggregate model	0.18	0.86	0.93	0.00		NA	NA	NA	NA		NA	NA	NA	NA		NA	NA	NA	NA	
	Foreign core	0.00	-0.09	0.54	0.04		0.21	0.55	0.69	0.10		0.02	-0.18	0.44	0.03		0.06	0.65	0.51	0.02	
	Foreign energy	0.00	0.12	0.06	0.00	0.00	-0.02	0.03	0.03	0.24	0.03	-0.01	0.12	0.07	0.00	0.00	0.00	0.01	0.03	0.31	0.01
	Foreign food	0.03	0.00	0.14	0.17		NA	NA	NA	NA		0.05	0.04	0.09	0.75		NA	NA	NA	NA	
Italy	Aggregate model	0.24	0.42		0.03		0.20	0.17		0.51		NA	NA	NA	NA		NA	NA	NA	NA	
	Foreign core	-0.17	0.17		0.04		0.41	0.14		0.09		-0.07	-1.08	0.15	0.00		0.41	1.39	0.31	0.00	
	Foreign energy	0.03	0.03		0.86	0.18	0.02	0.01		0.72	0.22	0.09	0.06	0.03	0.19	0.00	NA	NA	NA	NA	0.00
	Foreign food	0.14	0.09		0.21		NA	NA		NA		0.25	0.02	0.09	0.52		NA	NA	NA	NA	

Table 2.4 continued

Country	Coefficients	Estimated break dates from aggregate inflation model										Estimated break dates from aggregate-foreign components model									
		Contemporaneous coefficients			Equal-ity *p* -value	Equality joint *p* -value	Lag coefficients			Equal-ity *p* -value	Equality joint *p* -value	Contemporaneous coefficients			Equal-ity *p* -value	Equality joint *p* -value	Lag coefficients			Equal-ity *p* -value	Equality joint *p* -value
		R1	R2	R3			R1	R2	R3			R1	R2	R3			R1	R2	R3		
Japan	Aggregate model	0.33	0.31		0.49		0.65	0.19		0.22		NA	NA		NA		NA	NA		NA	
	Foreign core	0.07	0.43		0.49		NA	NA		NA		-0.16	0.45		0.25		NA	NA		NA	
	Foreign energy	0.01	0.02		0.89	0.56	0.43	0.02		0.00	0.01	0.29	0.03		0.01	0.05	NA	NA		NA	0.41
	Foreign food	0.18	0.04		0.24		0.09	0.06		0.80		0.05	0.04		0.92		0.15	0.05		0.41	
Nether-lands	Aggregate model	0.44	0.39		0.62		NA	NA		NA		NA	NA		NA		NA	NA		NA	
	Foreign core	0.27	0.13		0.27		NA	NA		NA		-0.09	0.08		0.43		0.45	0.09		0.10	
	Foreign energy	0.03	0.04		0.69	0.62	0.00	0.01		0.58	0.58	0.03	0.04		0.69	0.57	NA	NA		NA	0.10
	Foreign food	0.13	0.11		0.72		NA	NA		NA		0.04	0.12		0.40		NA	NA		NA	
Sweden	Aggregate model	0.41	0.78		0.00		NA	NA		NA		NA	NA		NA		NA	NA		NA	
	Foreign core	0.12	1.11		0.01		0.11	0.47		0.27		-0.30	0.30		0.04		0.07	0.46		0.12	
	Foreign energy	0.08	0.08		0.96	0.00	NA	NA		NA	0.27	0.12	0.07		0.10	0.04	NA	NA		NA	0.17
	Foreign food	-0.01	0.15		0.26		NA	NA		NA		0.04	0.17		0.30		0.17	0.03		0.22	
Switzer-land	Aggregate model	0.29	0.91		0.00		NA	NA		NA		NA	NA		NA		NA	NA		NA	
	Foreign core	-0.19	0.44		0.00		NA	NA		NA		-0.25	0.02		0.16		NA	NA		NA	
	Foreign energy	0.12	0.09		0.40	0.00	NA	NA		NA	NA	0.13	0.07		0.00	0.01	NA	NA		NA	NA
	Foreign food	0.07	0.13		0.47		NA	NA		NA		0.05	0.14		0.19		NA	NA		NA	
UK	Aggregate model	0.57	0.58	0.50	0.38		NA	NA	NA	NA		NA	NA	NA	NA		NA	NA	NA	NA	
	Foreign core	0.48	-0.10	0.04	0.42		0.97	0.16	0.25	0.14		0.36	0.28	0.09	0.59		0.97	-0.16	0.11	0.00	
	Foreign energy	0.08	0.08	0.04	0.27	0.49	NA	NA	NA	NA	0.07	0.07	0.08	0.04	0.25	0.41	NA	NA	NA	NA	0.00
	Foreign food	-0.05	0.13	0.09	0.55		0.04	-0.11	0.05	0.24		-0.07	0.06	0.09	0.32		NA	NA	NA	NA	
US	Aggregate model	0.40	1.36		0.00		NA	NA		NA		NA	NA	NA	NA		NA	NA	NA	NA	
	Foreign core	0.09	-0.42		0.03		0.11	0.09		0.90		0.41	0.01	0.06	0.12		NA	NA	NA	NA	
	Foreign energy	0.07	0.15		0.00	0.00	NA	NA		NA	0.90	0.07	0.05	0.11	0.01	0.00	NA	NA	NA	NA	0.04
	Foreign food	0.09	-0.01		0.08		NA	NA		NA		0.17	0.15	0.01	0.00		0.07	0.10	0.00	0.04	

Notes: The aggregate model is (2.1) estimated using aggregate inflation, with other results referring to the aggregate-foreign components model of (2.3). The equality *p* -value refers to a test of the null hypothesis of (2.4) or (2.6) for contemporaneous and lagged foreign coefficients, respectively, while the equality joint *p* -value refers to the hypothesis of (2.5) or (2.7) tested over foreign core, energy and food coefficients. All *p* -values are obtained using the conventional *F* -distribution, conditional on estimated aggregate or aggregate-foreign component model break dates shown in Table 2.2. The prefix R indicates regimes as implied by the respective break dates.

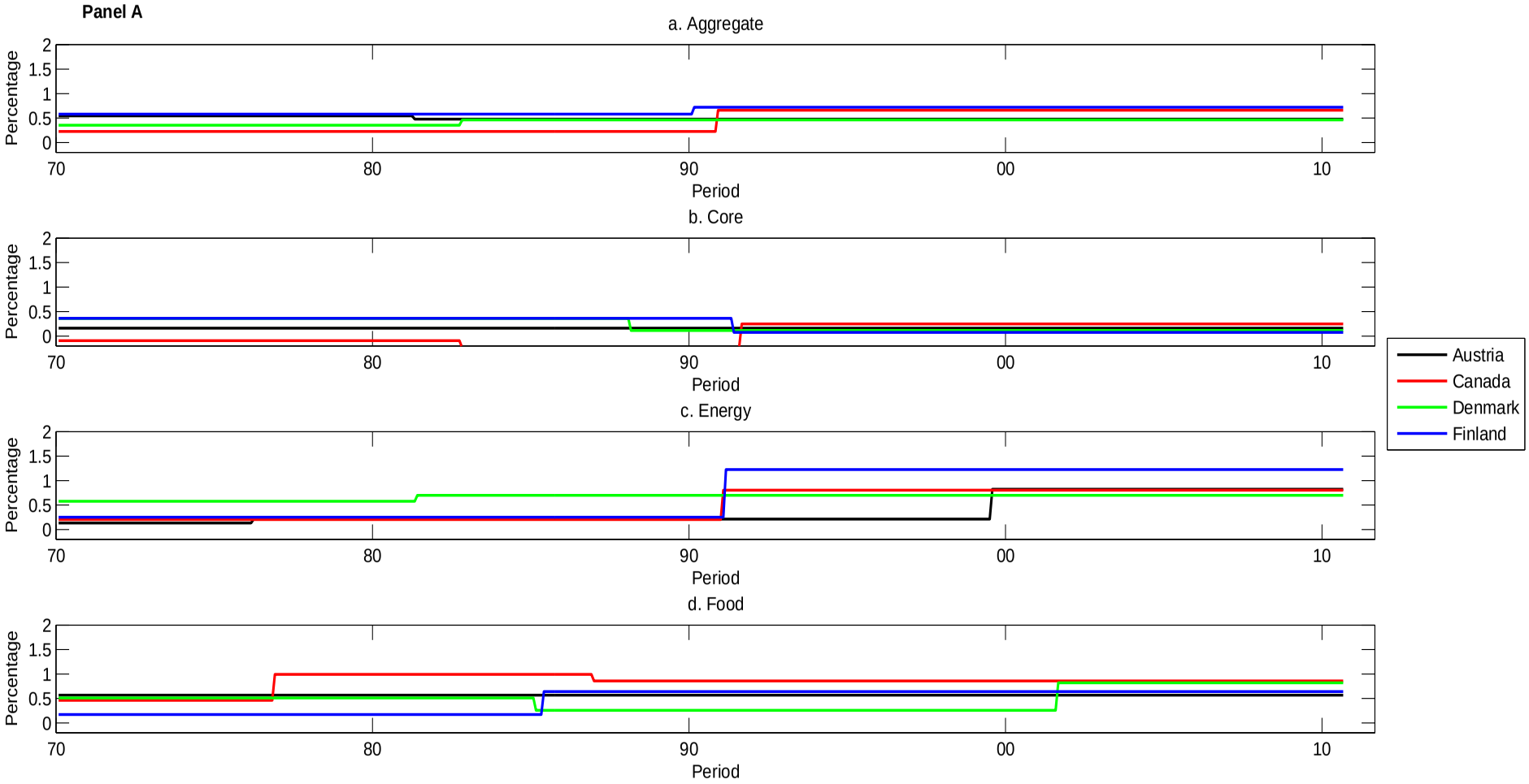
Table 2.5: Estimated persistence and intercepts in aggregate-foreign components model

Country	Persistence			Equality *p* - value	Intercepts			Equality *p* - value
	R1	R2	R3		R1	R2	R3	
Austria	0.02	0.12	0.11	0.76	0.17	-0.22	0.06	0.01
Canada	0.27	-0.04	0.03	0.00	0.26	0.34	0.10	0.03
Denmark	0.42	-0.01	-0.03	0.00	0.41	0.11	0.16	0.22
Finland	0.22	0.04		0.11	0.05	0.00		0.43
France	0.70	0.07		0.00	0.04	0.04		0.98
Germany	0.37	-0.12	0.06	0.01	0.18	-0.03	-0.05	0.06
Italy	0.66	0.64	0.67	0.97	-0.11	0.23	-0.01	0.39
Japan	0.46	0.24		0.41	0.22	-0.11		0.14
Netherland	0.34	0.35		0.94	0.19	0.05		0.34
Sweden	0.40	0.10		0.01	0.32	-0.05		0.00
Switzerland	0.52	-0.09		0.00	0.18	0.04		0.00
UK	0.38	0.38	0.35	0.95	-0.18	0.20	0.06	0.01
US	0.05	0.56	0.32	0.01	-0.03	0.09	0.11	0.22

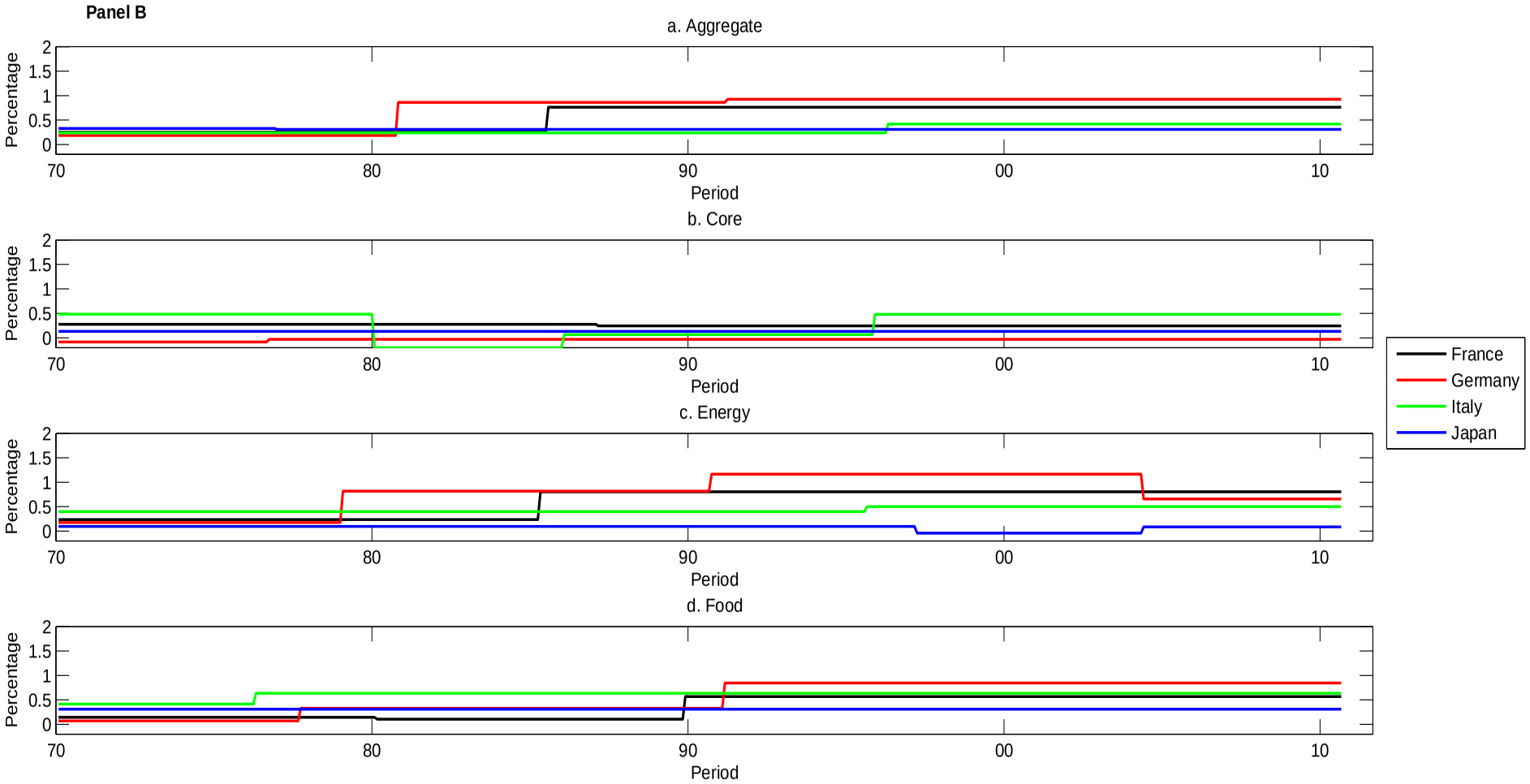
Notes: Breaks are tested in the context of (2.3), with break dates also estimated in this model. The prefix R indicates regime, while p-values are obtained using the conventional *F* -distribution.

Figures

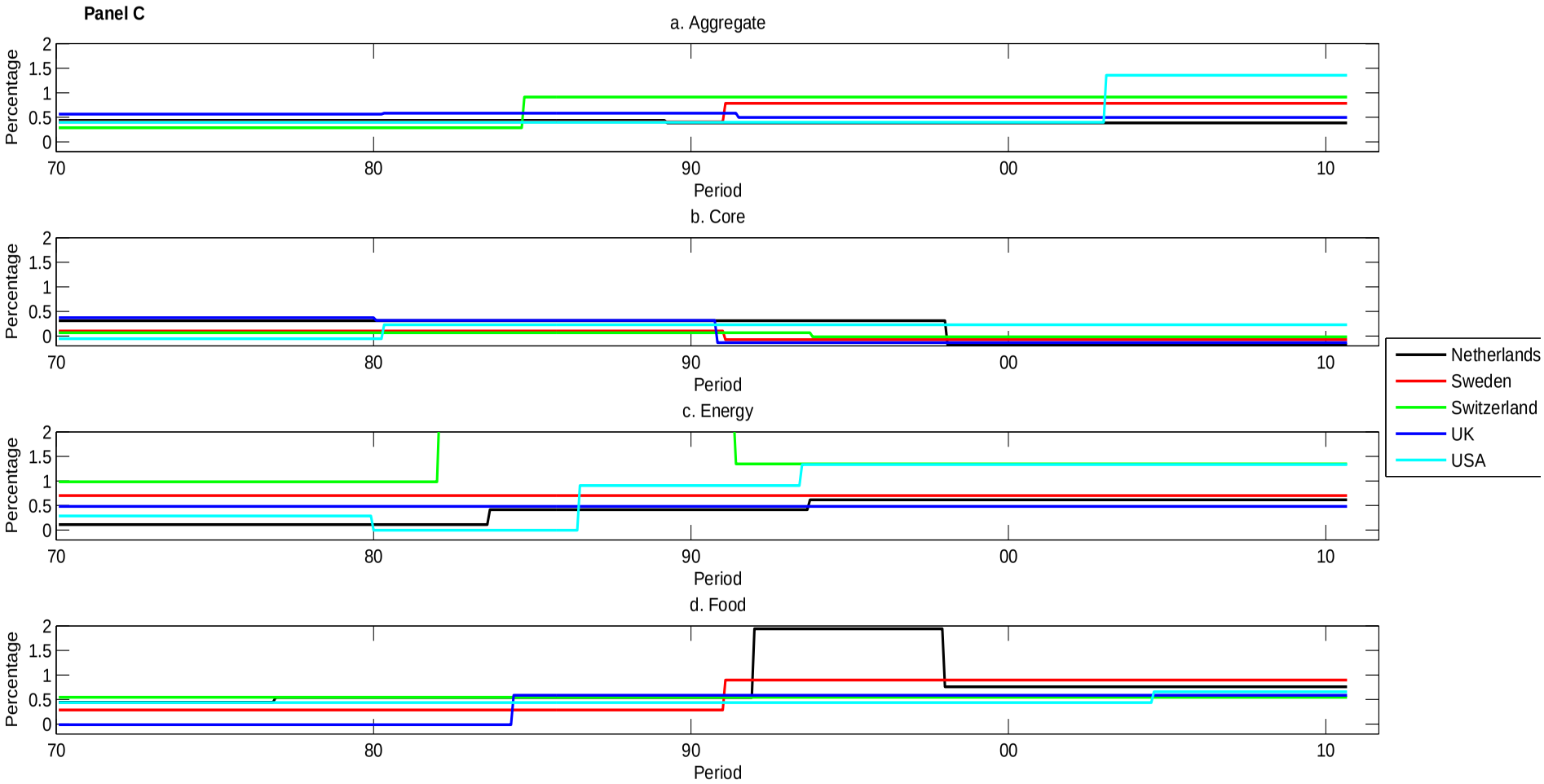
Figure 2.1: Contemporaneous foreign inflation coefficients



Notes: Coefficients are those of Table 2.3, over regimes as defined by the break dates of Table 2.2

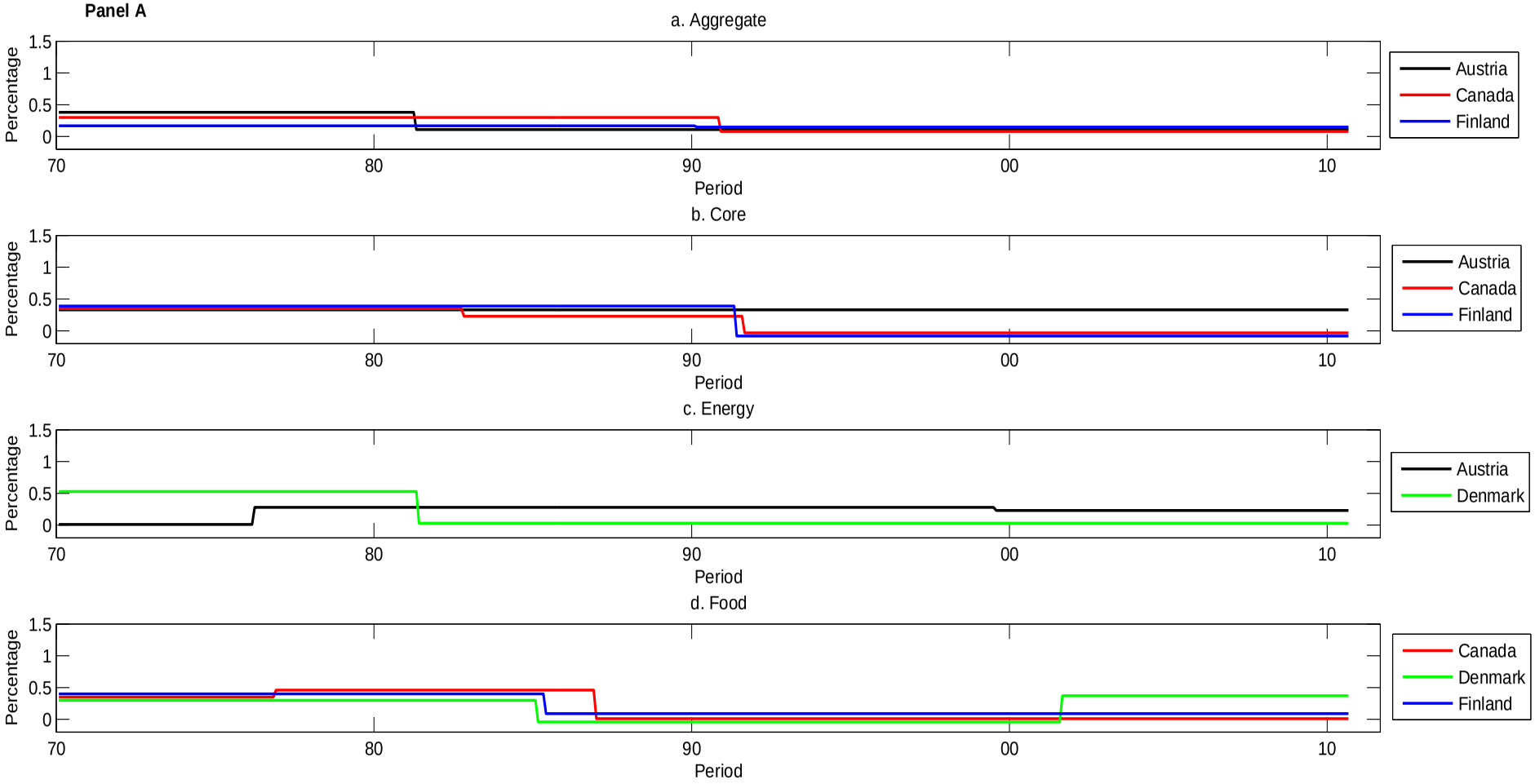


Notes: Coefficients are those of Table 2.3, over regimes as defined by the break dates of Table 2.2

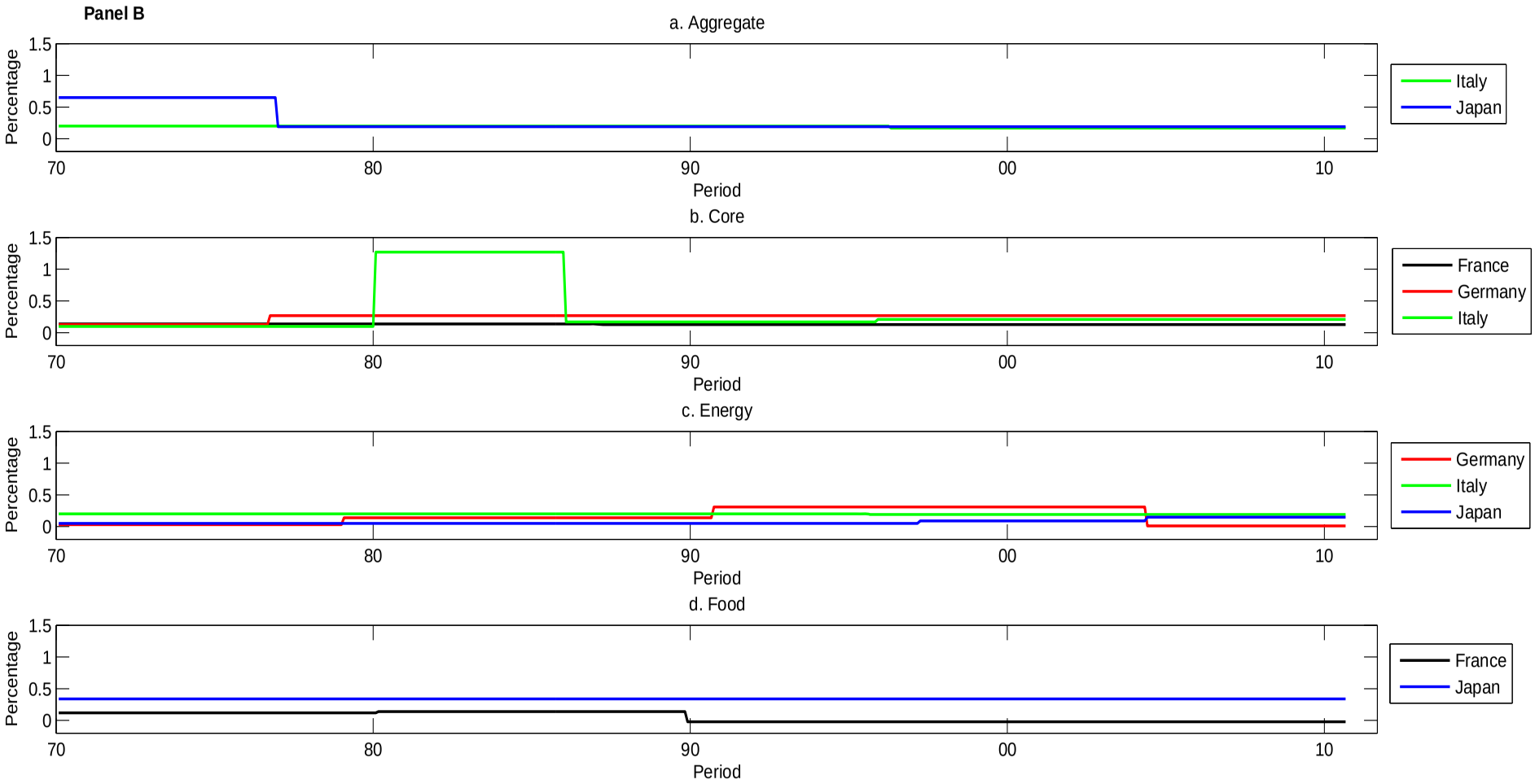


Notes: Coefficients are those of Table 2.3, over regimes as defined by the break dates of Table 2.2

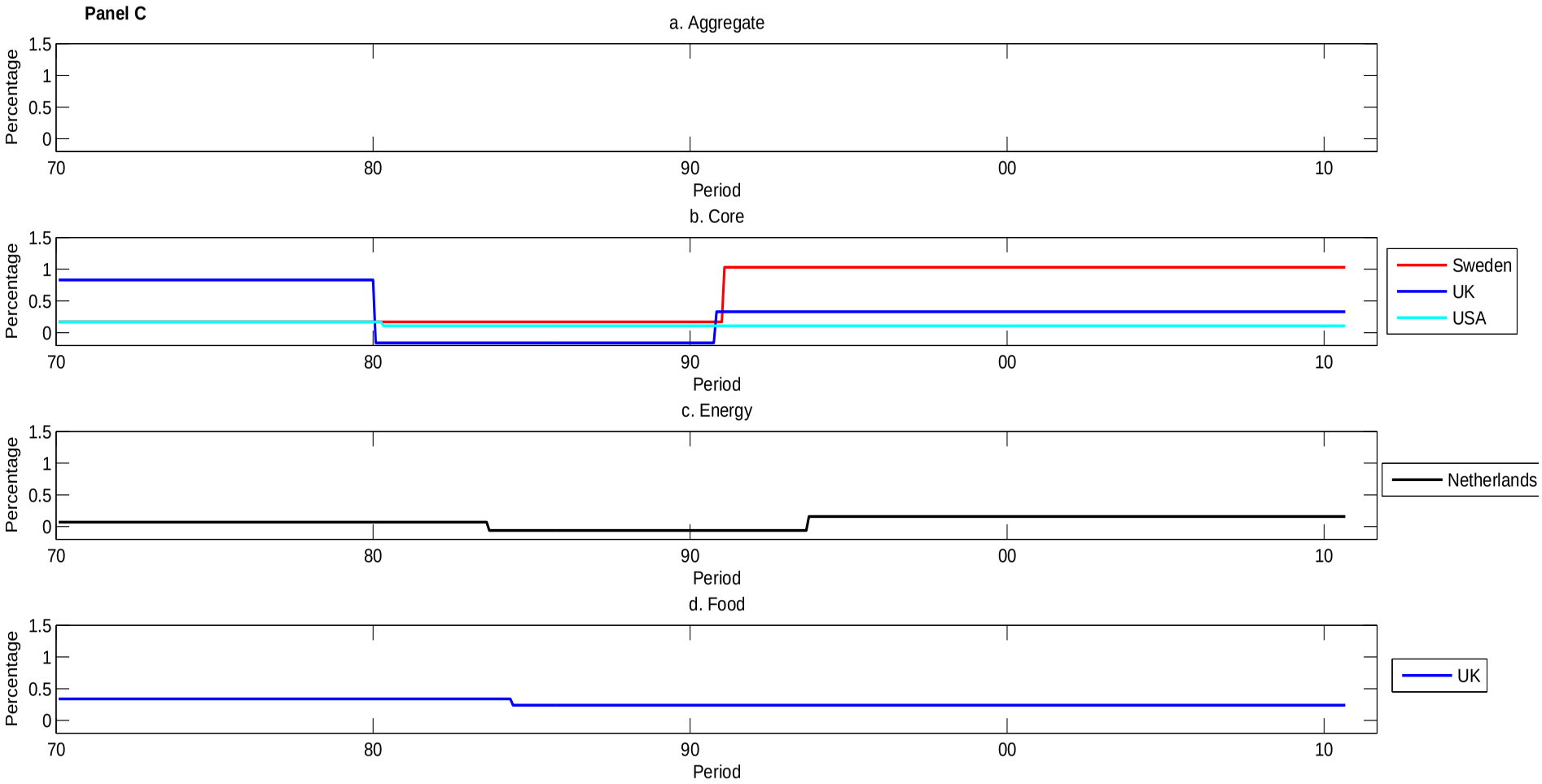
Figure 2.2: Lagged foreign inflation coefficients



Notes: As for Figure 2.1

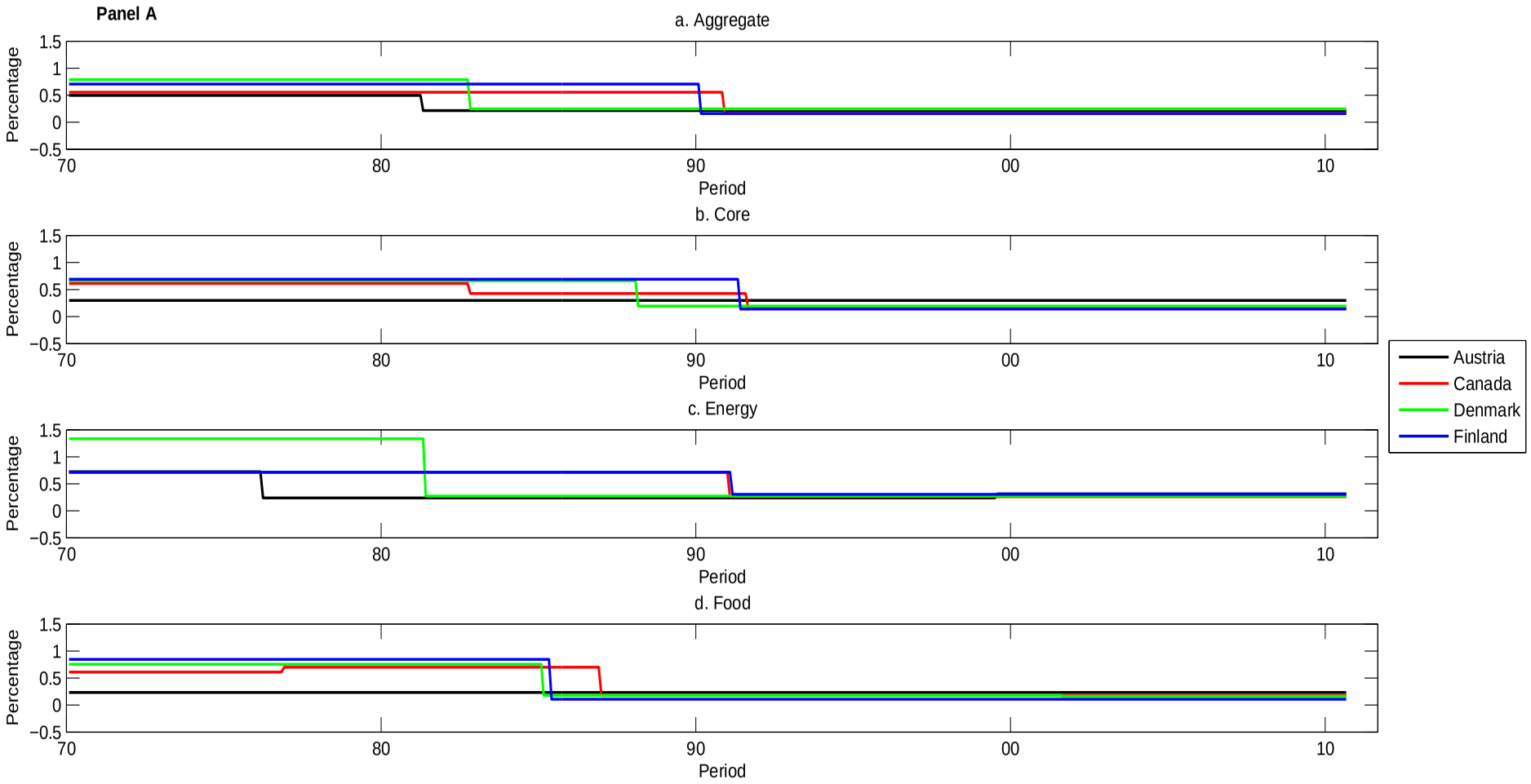


Notes: As for Figure 2.1

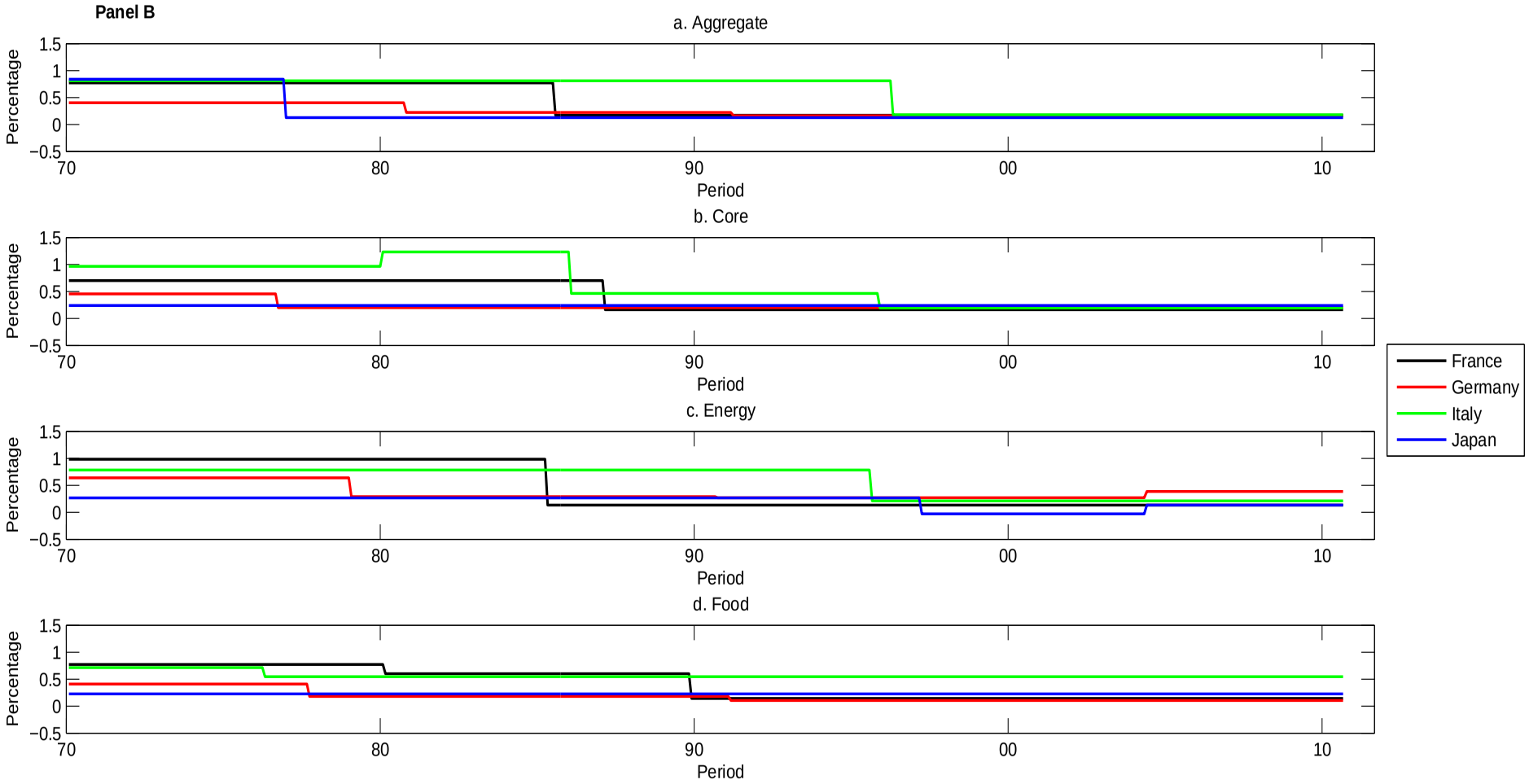


Notes: As for Figure 2.1

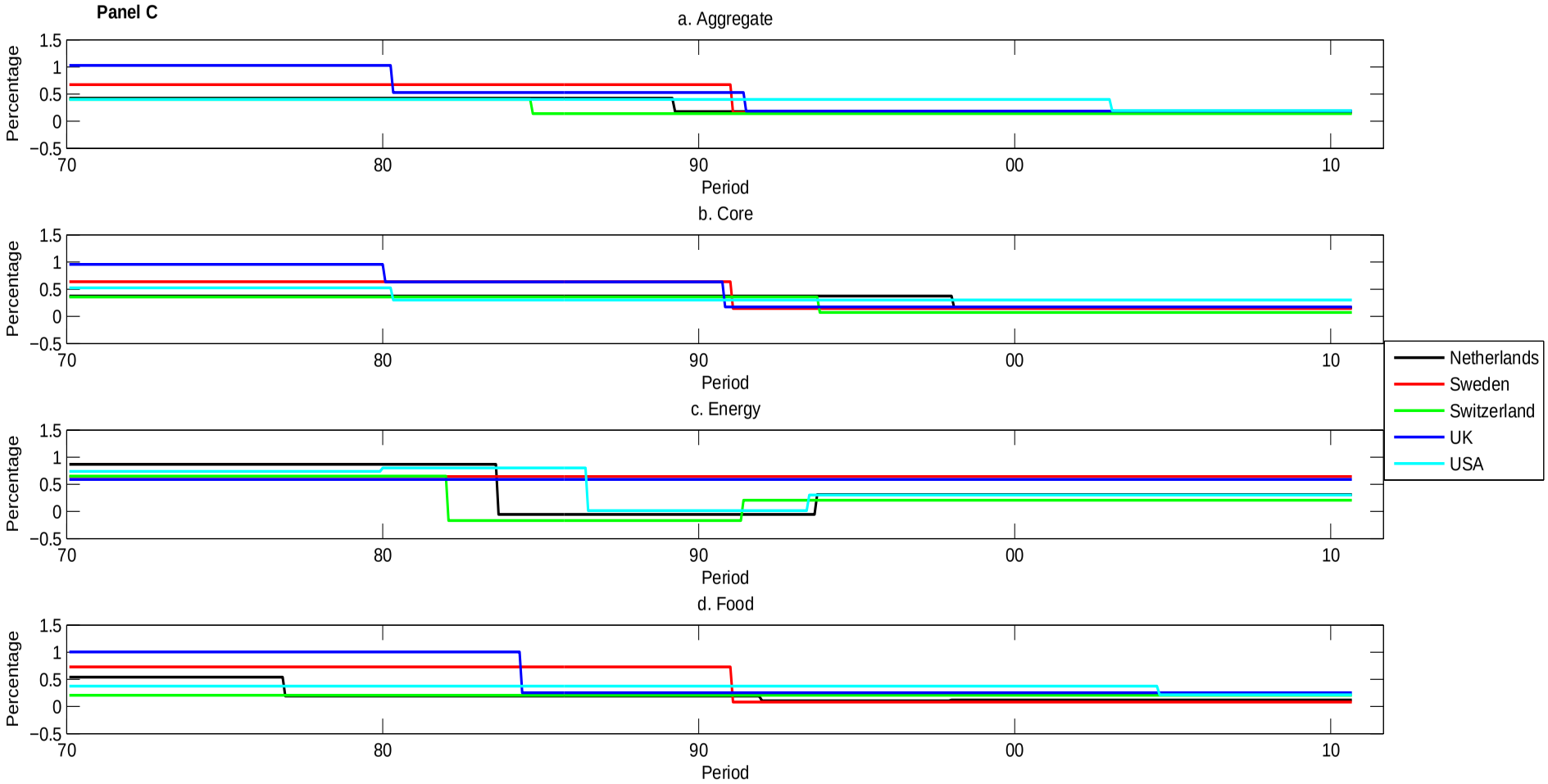
Figure 2.3: Mean of domestic inflation



Notes: As for Figure 2.1

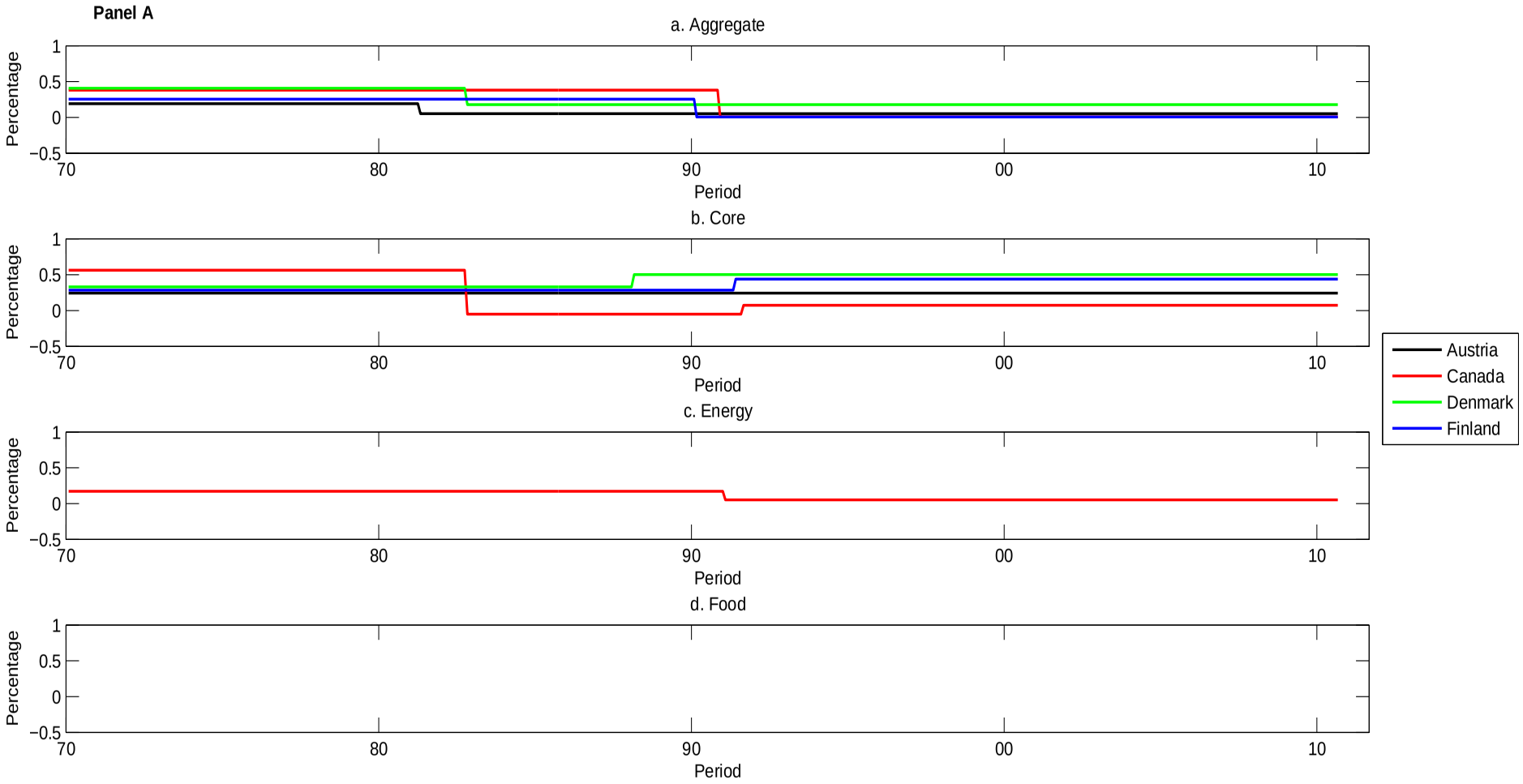


Notes: As for Figure 2.1

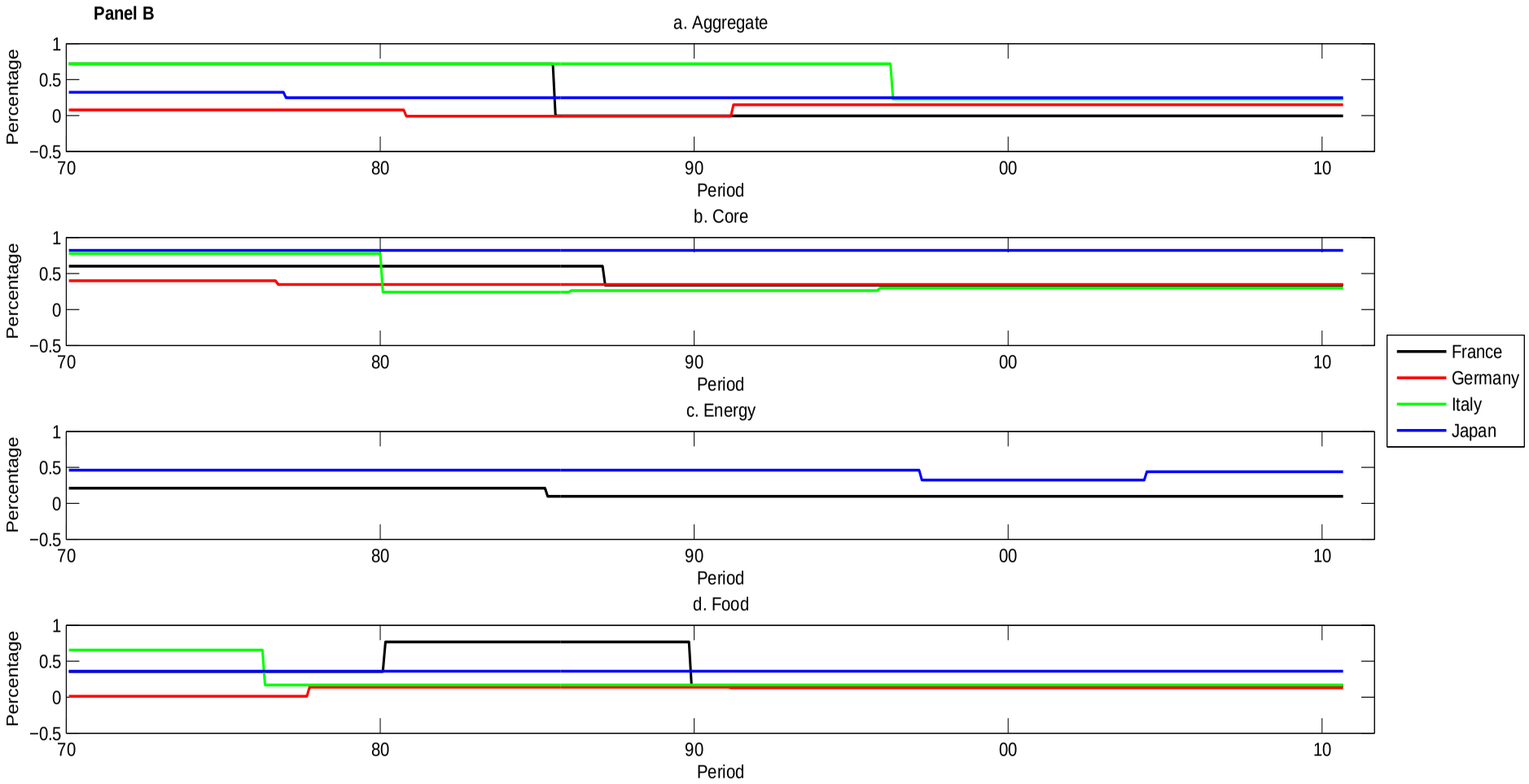


Notes: As for Figure 2.1

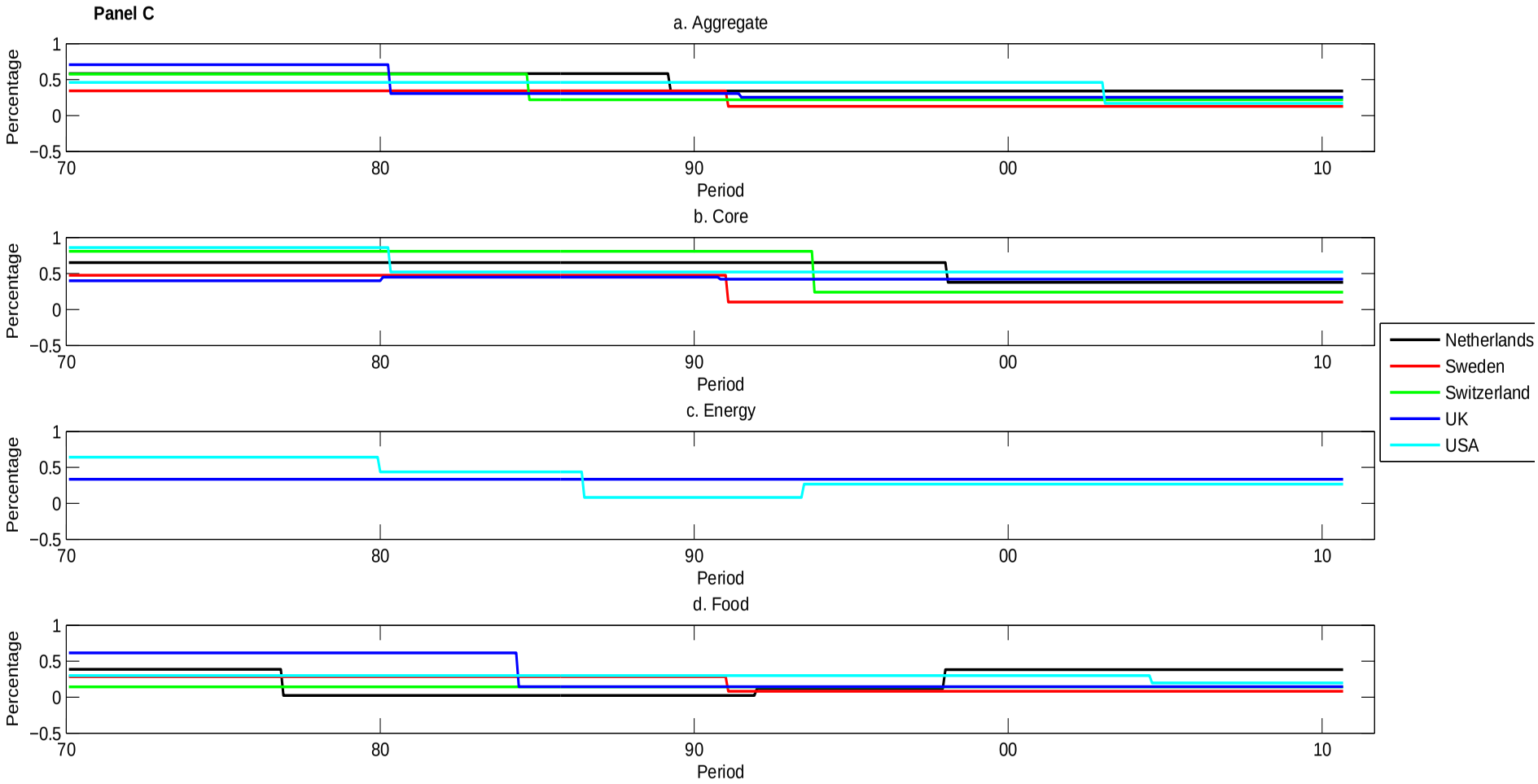
Figure 2.4: Persistence



Notes: As for Figure 2.1

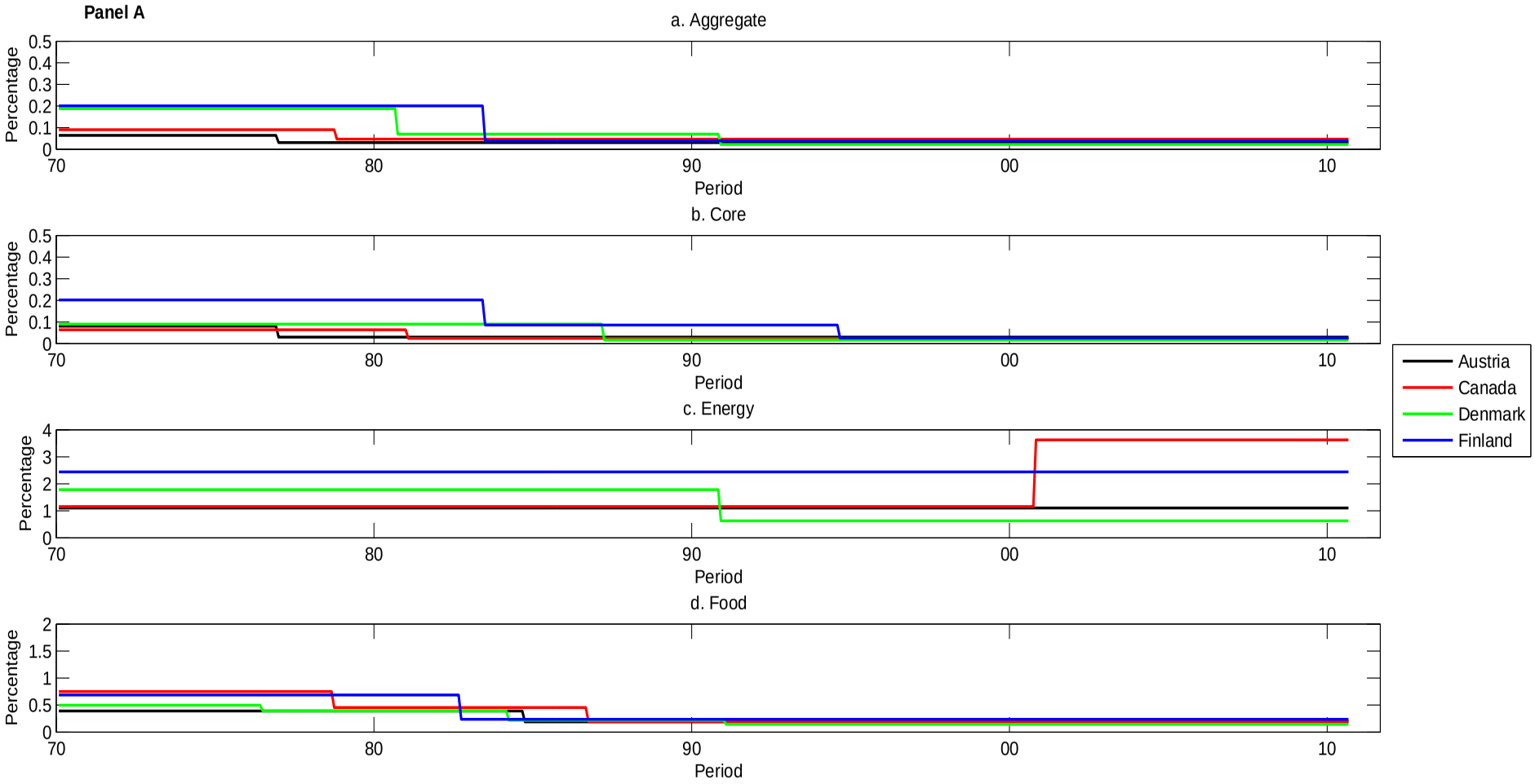


Notes: As for Figure 2.1

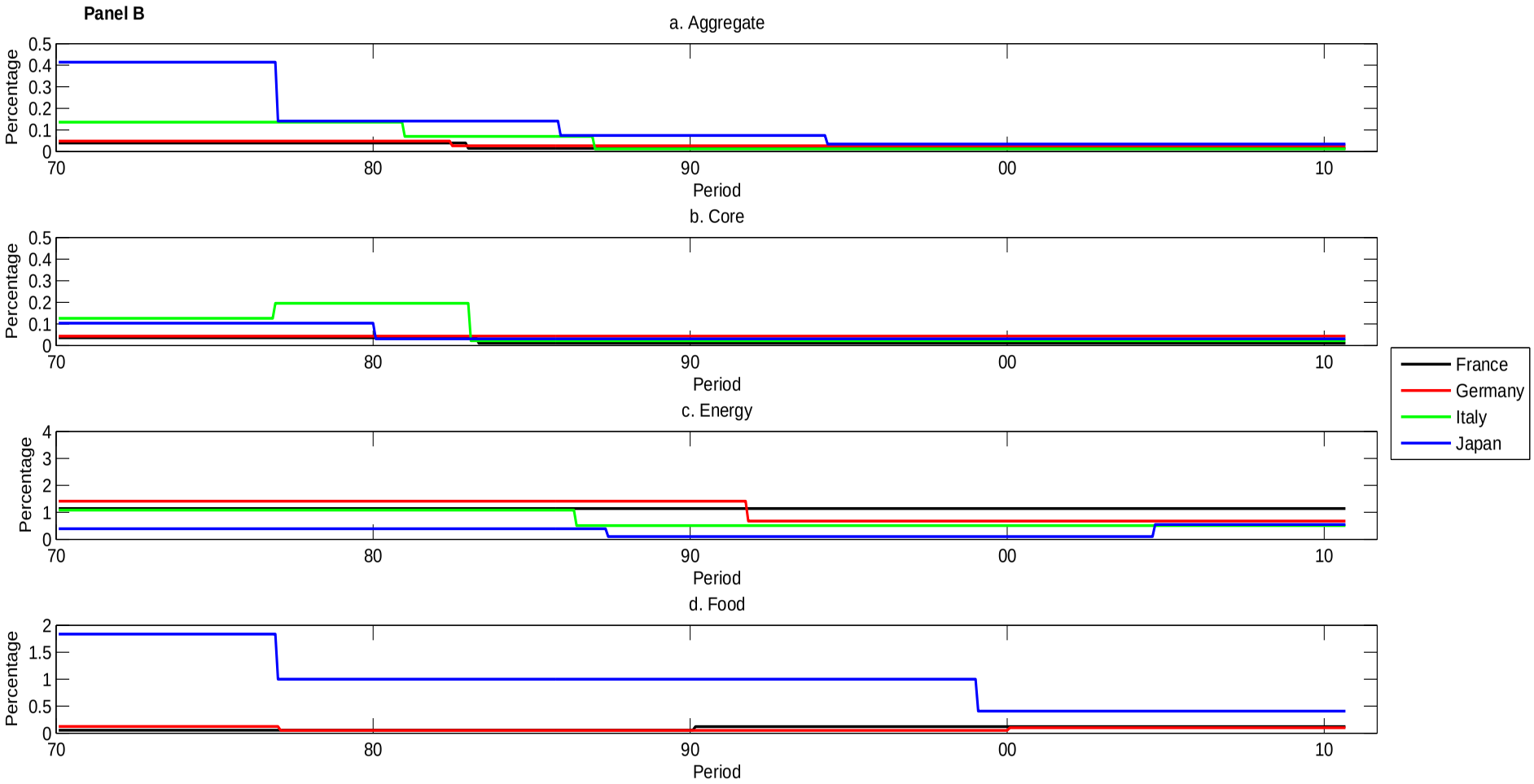


Notes: As for Figure 2.1

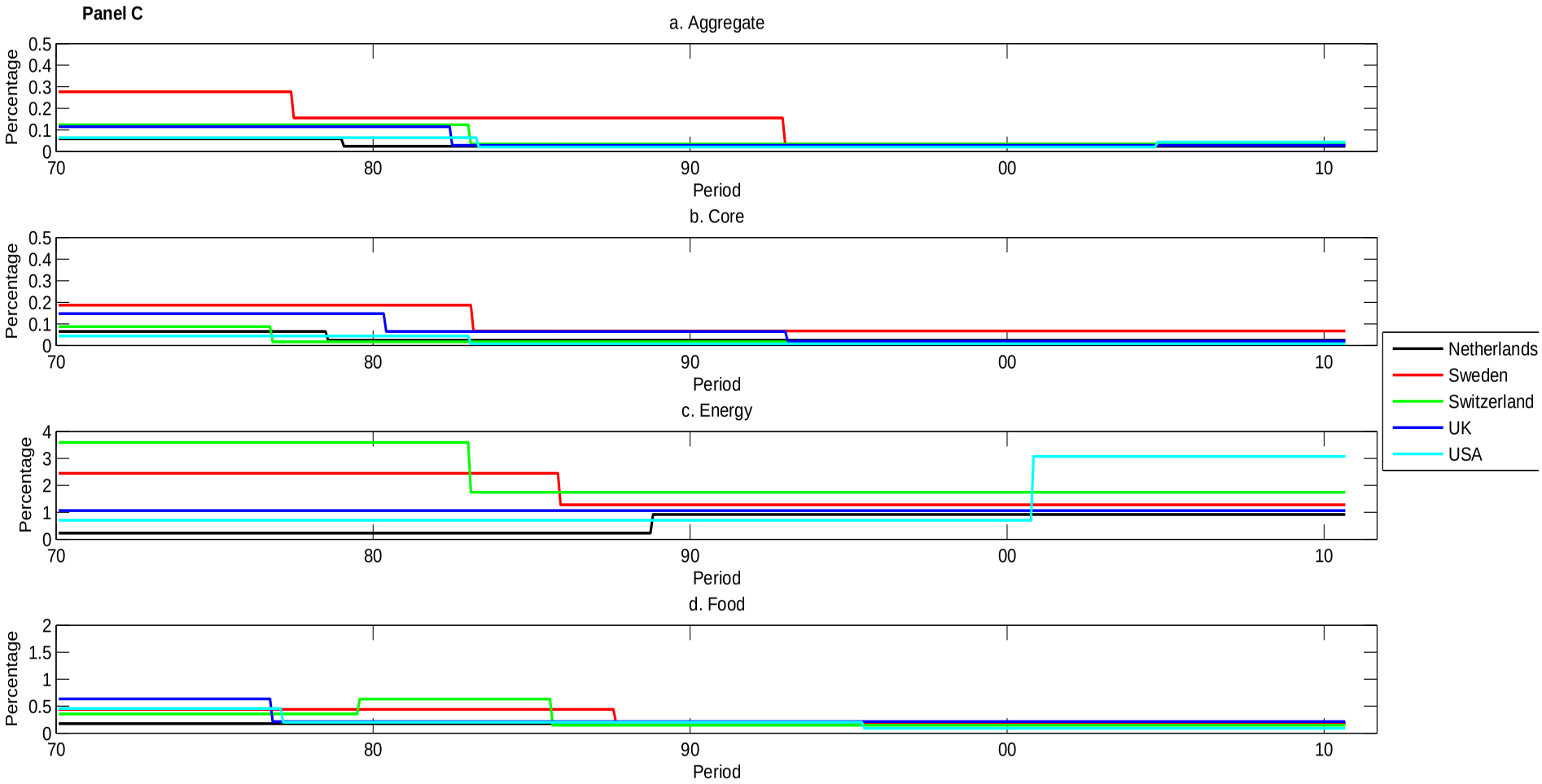
Figure 2.5: Residual variances



Notes: As for Figure 2.1



Notes: As for Figure 2.1



Notes: As for Figure 2.1

Appendix B

B.1 Additional Tables

Table B.1: Average trade weights by partner countries

	Aus	Can	Den	Fin	Fra	Ger	Ita	Jap	Net	Swe	Swi	Uk	US	SUM
Aus	0.000	0.008	0.013	0.010	0.056	0.523	0.119	0.027	0.038	0.026	0.077	0.051	0.053	1.000
Can	0.002	0.000	0.002	0.002	0.013	0.022	0.011	0.055	0.008	0.004	0.005	0.037	0.840	1.000
Den	0.015	0.009	0.000	0.038	0.067	0.279	0.054	0.036	0.073	0.181	0.024	0.144	0.080	1.000
Fin	0.017	0.012	0.051	0.000	0.067	0.218	0.049	0.046	0.064	0.206	0.026	0.150	0.095	1.000
Fra	0.014	0.014	0.013	0.010	0.000	0.311	0.177	0.037	0.090	0.025	0.055	0.130	0.125	1.000
Ger	0.075	0.013	0.030	0.016	0.188	0.000	0.130	0.046	0.153	0.039	0.072	0.110	0.129	1.000
Ita	0.038	0.015	0.014	0.008	0.225	0.305	0.000	0.028	0.073	0.020	0.065	0.092	0.118	1.000
Jap	0.006	0.067	0.009	0.006	0.039	0.096	0.027	0.000	0.034	0.013	0.023	0.062	0.618	1.000
Net	0.015	0.008	0.021	0.013	0.137	0.412	0.072	0.028	0.000	0.032	0.023	0.137	0.102	1.000
Swe	0.020	0.014	0.114	0.090	0.076	0.229	0.051	0.037	0.078	0.000	0.027	0.151	0.114	1.000
Swi	0.054	0.010	0.014	0.009	0.137	0.346	0.122	0.044	0.049	0.024	0.000	0.083	0.107	1.000
Uk	0.013	0.039	0.031	0.021	0.140	0.203	0.074	0.051	0.122	0.051	0.048	0.000	0.204	1.000
US	0.005	0.396	0.007	0.005	0.052	0.100	0.042	0.231	0.039	0.015	0.022	0.086	0.000	1.000

Note: The shares for each country with respect to partner countries are given in the rows of the table. The values shown are computed as monthly averages of trade shares over 40 years of data.

Table B.2: Detailed results for coefficient break tests applied to aggregate inflation

	WDmax test		**Sequential *F-tests***							
Country	**Statistic**	**p-value**	**SupF(2|1)**	**p-value**	**SupF(3|2)**	**p-value**	**SupF(4|3)**	**p-value**	**SupF(5|4)**	**p-value**
Austria	18.737	0.019	13.466	0.252	9.906	0.784	0.000	1.000	0.000	1.000
Canada	60.343	0.000	17.559	0.224	9.556	0.995	16.234	0.536	6.437	1.000
Denmark	37.280	0.000	18.874	0.080	10.627	0.874	0.000	1.000	0.000	1.000
Finland	32.589	0.000	15.843	0.216	11.362	0.801	13.734	0.624	0.000	1.000
France	100.975	0.000	18.828	0.081	13.213	0.585	4.979	1.000	0.000	1.000
Germany	44.393	0.000	30.301	0.000	4.360	1.000	1.253	1.000	0.000	1.000
Italy	38.430	0.000	22.024	0.052	19.442	0.184	19.099	0.259	0.000	1.000
Japan	29.250	0.003	11.506	0.917	17.160	0.513	15.089	0.838	0.000	1.000
Netherland	26.783	0.001	18.650	0.086	10.999	0.839	2.855	1.000	0.000	1.000
Sweden	27.319	0.001	11.100	0.686	7.936	0.997	5.488	1.000	0.000	1.000
Switzerland	45.171	0.000	16.489	0.177	14.011	0.491	11.348	0.888	0.000	1.000
UK	28.182	0.001	21.688	0.028	10.127	0.915	19.331	0.131	0.000	1.000
US	63.893	0.000	17.450	0.063	19.207	0.048	19.335	0.059	5.742	1.000

Notes: Results are shown for structural break test statistics applied to the coefficients of (2.1) for aggregate inflation using the iterative procedure of subsection 2.2.2; p -values are obtained using the method of Hall and Sakkas (2012).

Table B.3: Detailed results for variance break tests applied to aggregate inflation

	WDmax test		Sequential *F-tests*							
Country	**Statistic**	**p-value**	**SupF(2|1)**	**p-value**	**SupF(3|2)**	**p-value**	**SupF(4|3)**	**p-value**	**SupF(5|4)**	**p-value**
Austria	16.728	0.001	8.946	0.088	1.538	1.000	2.042	1.000	0.000	1.000
Canada	12.724	0.006	3.414	0.737	5.355	0.523	1.076	1.000	0.000	1.000
Denmark	31.125	0.000	31.283	0.000	2.573	0.978	0.307	1.000	0.000	1.000
Finland	44.920	0.000	6.176	0.284	2.471	0.985	1.312	1.000	0.000	1.000
France	17.832	0.000	6.621	0.238	4.485	0.679	1.936	1.000	0.998	1.000
Germany	13.781	0.003	3.521	0.715	0.812	1.000	0.513	1.000	0.000	1.000
Italy	62.459	0.000	13.164	0.013	2.835	0.953	0.366	1.000	0.093	1.000
Japan	84.407	0.000	39.520	0.000	12.080	0.032	1.471	1.000	0.000	1.000
Netherlands	15.844	0.001	5.276	0.401	2.892	0.946	2.648	0.994	0.000	1.000
Sweden	31.275	0.000	15.744	0.004	4.702	0.639	2.666	0.993	0.000	1.000
Switzerland	34.942	0.000	6.793	0.222	7.004	0.290	1.240	1.000	0.000	1.000
UK	58.954	0.000	5.007	0.442	8.787	0.140	2.266	1.000	0.004	1.000
US	22.579	0.000	17.624	0.001	3.025	0.930	7.534	0.300	0.307	1.000

Notes: As for Table B.2, except that tests are applied to the variance.

Table B.4: WDmax tests for no breaks in component domestic/foreign inflation models

	Coefficient breaks						Variance breaks					
	Core inflation		Energy inflation		Food inflation		Core inflation		Energy inflation		Food inflation	
Country	**WDmax**	**p-value**	**WDmax**	**p-value**	**WDmax**	**p-value**	**WDmax**	**p-value**	**WDmax**	**p-value**	**WDmax**	**p-value**
Austria	18.828	0.050	117.995	0.000	11.115	0.081	29.079	0.000	6.917	0.148	15.856	0.001
Canada	37.572	0.000	85.488	0.000	27.135	0.000	38.713	0.000	28.653	0.000	40.135	0.000
Denmark	50.291	0.000	25.167	0.000	24.471	0.000	30.502	0.000	25.672	0.000	24.448	0.000
Finland	30.375	0.001	148.391	0.000	26.366	0.000	40.091	0.000	3.923	0.572	31.846	0.000
France	41.912	0.000	63.192	0.000	75.692	0.000	36.891	0.000	6.564	0.176	12.776	0.006
Germany	28.938	0.001	37.024	0.000	63.299	0.000	3.804	0.598	14.848	0.002	12.917	0.006
Italy	30.063	0.002	18.470	0.008	19.690	0.005	34.864	0.000	13.574	0.004	5.234	0.330
Japan	20.130	0.216	25.977	0.002	20.041	0.068	34.068	0.000	18.569	0.000	23.687	0.000
Netherlands	25.695	0.007	79.503	0.000	32.630	0.000	24.371	0.000	18.986	0.000	8.541	0.064
Sweden	36.663	0.000	11.327	0.074	35.793	0.000	20.180	0.000	11.052	0.016	15.687	0.001
Switzerland	26.216	0.012	54.693	0.000	12.076	0.145	61.604	0.000	19.675	0.000	43.072	0.000
UK	30.030	0.001	17.172	0.040	36.268	0.000	16.027	0.001	3.988	0.558	26.347	0.000
US	37.040	0.000	171.177	0.000	17.750	0.031	35.776	0.000	41.531	0.000	46.242	0.000

Notes: As for Tables B.2 and B.3, except that the models are estimated using core, energy or food inflation series.

Table B.5: Detailed results for coefficient break tests in aggregate-foreign components model

	WDmax test		Sequential *F-tests*							
Country	**Statistic**	**p-value**	**SupF(2|1)**	**p-value**	**SupF(3|2)**	**p-value**	**SupF(4|3)**	**p-value**	**SupF(5|4)**	**p-value**
Austria	28.629	0.002	24.808	0.019	15.242	0.550	13.360	0.855	0.000	1.000
Canada	84.525	0.000	26.721	0.019	13.787	0.859	10.048	1.000	13.891	0.969
Denmark	90.912	0.000	24.522	0.042	11.665	0.977	6.040	1.000	0.000	1.000
Finland	33.981	0.001	23.350	0.115	16.082	0.786	13.302	0.995	9.987	1.000
France	111.103	0.000	20.300	0.270	24.482	0.115	11.737	1.000	0.000	1.000
Germany	45.427	0.000	38.995	0.000	24.312	0.121	3.836	1.000	0.000	1.000
Italy	40.716	0.000	30.451	0.010	21.868	0.245	17.408	0.764	0.000	1.000
Japan	35.323	0.001	18.683	0.538	18.939	0.647	12.161	1.000	0.000	1.000
Netherlanc	31.717	0.005	22.370	0.241	24.869	0.168	12.329	1.000	4.410	1.000
Sweden	42.046	0.000	15.503	0.831	16.783	0.847	12.218	1.000	0.000	1.000
Switzerlanc	60.017	0.000	10.743	0.991	12.731	0.986	8.702	1.000	0.000	1.000
UK	42.967	0.000	34.968	0.002	12.453	0.991	19.475	0.547	0.000	1.000
US	43.430	0.000	27.685	0.028	25.235	0.091	19.537	0.540	15.333	0.964

Notes: As for Table B.2, except that results refer to the model of (2.3).

Table B.6: Detailed results for variance break tests applied to aggregate-foreign components model

	WDmax test		Sequential *F-tests*							
Country	**Statistic**	**p-value**	**SupF(2|1)**	**p-value**	**SupF(3|2)**	**p-value**	**SupF(4|3)**	**p-value**	**SupF(5|4)**	**p-value**
Austria	21.107	0.000	6.517	0.248	1.492	1.000	1.518	1.000	0.000	1.000
Canada	16.264	0.001	6.730	0.227	1.325	1.000	0.072	1.000	0.000	1.000
Denmark	57.259	0.000	29.217	0.000	3.757	0.814	0.725	1.000	0.331	1.000
Finland	32.273	0.000	6.456	0.254	1.767	1.000	1.594	1.000	0.000	1.000
France	23.662	0.000	5.192	0.413	4.373	0.700	0.857	1.000	0.493	1.000
Germany	18.383	0.000	9.902	0.058	2.265	0.995	0.582	1.000	0.000	1.000
Italy	63.614	0.000	8.564	0.105	0.992	1.000	0.448	1.000	0.112	1.000
Japan	104.269	0.000	41.768	0.000	11.428	0.043	0.601	1.000	0.000	1.000
Netherlands	11.168	0.015	2.448	0.919	1.260	1.000	1.959	1.000	0.000	1.000
Sweden	28.789	0.000	12.796	0.015	0.647	1.000	0.410	1.000	0.000	1.000
Switzerland	35.553	0.000	10.473	0.044	3.981	0.773	3.776	0.892	2.577	1.000
UK	42.874	0.000	5.991	0.305	7.429	0.246	1.194	1.000	0.000	1.000
US	24.733	0.000	18.323	0.001	4.869	0.609	7.712	0.280	0.426	1.000

Notes: As for Table B.5, except that tests are applied to the variance.

Chapter 3

Forecasting Time Series in the Presence of Structural Breaks

3.1 Introduction

Structural changes which characterize many macroeconomics time series are often unanticipated and they are a major cause of forecast failure (Hendry and Clements, 2003). If structural changes are defined as permanent shifts in the data generating process, a previously best in-sample model generally no longer applies after the parameters of the data generating process change. Since forecasts are generated relying on parameter estimates of the model, unaccounted breaks may lead to substantial deterioration in forecasting performance.

The existing literature includes various forecasting methods which differ in their handling of structural breaks. These methods can be grouped depending on whether their forecasts are based on a carefully selected single estimation window, average across multiple estimation windows or apply different weights to observations in the full sample. A detailed review of methods in each category is presented in the next section. In brief, many of these single estimation window and observational weighting methods exploit information on breaks (described as non-robust methods). Whereas, window averaging methods are often robust to structural breaks of unknown break dates and sizes (described as robust methods).

More specifically, when estimating the parameters of the forecasting model using single window methods, forecasters attempt to detect structural breaks that may have occurred in the past and eliminate observations which are generated from different regimes from that which currently applies. The rationale for this procedure is related to the potential bias in the parameter estimates when such observations are included in the estimation sample. The common

example is the post break window method which only employs observations after the most recent break. However, if the length of the post break window is short, variances of the parameter estimates and resulting forecast errors increase. Pesaran and Timmermann (2007) analytically shows that such variances can be reduced by employing a longer estimation window which includes pre-break observations despite the bias of parameter estimators using such a window. This trade-off method exploits an optimal estimation window by trading off bias against variances of the OLS estimators.

Rather than defining an estimation window within which all observations are weighted equally, information on breaks can also be used to derive observational weights in which all observations in the sample are utilized but weighted differently when estimating the parameters of the forecast model. For example, Pesaran et al. (2013) propose observational weights that are optimal in the sense that the resulting mean squared forecast error is minimized by exploiting information on the dates and sizes of breaks.

Nevertheless, forecast accuracies of these methods heavily rely on how well the true break date is estimated. However, in practice, estimates of break dates can be imprecise and this rules out achieving the full efficiency of these methods. Indeed, precise identification of break dates depends on a number of parameters that researchers have to assume but often do not have prior knowledge about, including the size, the number of breaks, and frequency of breaks. Alternatively, some forecasters advocate robust methods to structural breaks, for example, forecast combination methods, to avoid inaccurate estimates of breaks.

Specifically, forecasts from multiple windows of different sizes can be combined using various weighting schemes, which is found to be superior compared to a single estimation window method (Pesaran and Pick, 2011, Tian and Anderson, 2011, Eklund et al., 2013, and many others). In particular, forecast averaging across estimation windows is shown to work well when the breaks are small or recurring or only occur in the variance. However, these methods do not perform well in the presence of large breaks (Pesaran and Timmermann, 2004, 2007, Eklund et al., 2013). This suggests that sizes of breaks plays an important role in determining the success of this method. In practice, both the date and size of breaks need to be estimated. The latter can be more difficult to be estimated with high accuracy as the number of available observations in each break segment can be small (Pesaran and Pick, 2011).

This paper proposes a forecast method that is effectively a mixture of the two types of approaches: namely, non-robust and robust, mentioned above. Specifically, we employ a confidence interval for the estimated break date for

choosing the range of windows to be averaged. Each date in the confidence interval is treated as one of a sequence of choices for the potential break date, and the corresponding post break window forecasts are averaged. To some extent, the size of breaks is incorporated in the width of the interval – for example, when the break is large, the corresponding interval is narrow and hence windows that use less relevant data are excluded from the forecast combination process. Therefore, our approach can be seen as an improvement on existing methods that combine forecasts from all possible windows, many of which may yield large forecast errors, consequently leading to distortions in overall forecast accuracy.

The performance of the proposed confidence interval method is assessed against other related forecast methods, using Monte Carlo simulations and an empirical application to inflation. We report an overall good performance of our method in the presence of large and small breaks that occur in the coefficients of the forecast model. In particular, confidence interval methods can outperform forecast combination methods in coefficient break experiments regardless of the locations of breaks. Furthermore, when variance breaks are present in the data generating process, we employ an iterative structural break testing methodology proposed in the preceding chapters (subsections 1.3.1 and 2.2.2). The iteration based confidence interval which accounts breaks in the both coefficients and residual variances improves on single window forecast methods that ignore the possibility of changing variances.

A second contribution of this paper is to improve forecastability when there are multiple breaks which have the form of reverting coefficients. Most forecast methods that break date estimates are well suited in the presence of a single break. The prevailing way to adapt these methods in a multiple break environment is to use the most recent break only. However, time series may exhibit a regime-switching process, in which the detected structural breaks capture switches between two or more distinct (but recurring) regimes. In this respect, using only the most recent break ignores the fact that data prior to any previous break can be informative with regard to a forecast value. To account for this, we propose a procedure of re-ordering data segments associated with estimated coefficient breaks based on the relative closeness of the estimated parameters compared with those in the segment after the last identified break. Monte Carlo simulations show that the procedure of re-ordering data segments substantially improves forecast accuracies of all methods assessed when coefficient reversion applies.

This paper proceeds as follows. In section 3.2, we review the existing methods in the literature. Section 3.3 describes available procedures to estimate a

break date confidence interval and details the proposed confidence interval forecast. This section also outlines the procedure of re-ordering data segments in the presence of multiple estimated breaks. Section 3.4 sets up the Monte Carlo simulations and the simulation results are presented in section 3.5. Section 3.6 examines the performances of forecast methods for observed data for inflation in the G7 countries. Section 3.7 concludes and some remaining details are presented in the appendix.

3.2 Review of current methods

In order to examine the issues, consider the multivariate regression model

$$y_t = \beta_t' \mathbf{x}_{t-1} + \sigma_t \varepsilon_t \qquad \varepsilon_t \sim IID(0,1) \tag{3.1}$$

where $\mathbf{x}_{t-1}$ is a $k \times 1$ vector of stationary lagged regressors, β_t is the $k \times 1$ coefficient vector which is subject to a single discrete break at time T_1. ε_t is a serially uncorrelated error term with mean zero and variance of σ_t^2, which is assumed to be independently distributed to $\mathbf{x}_{t-1}$. As in previous chapters, we allow the variance of ε_t to also be subject to structural breaks.

Suppose that data $t = 1, 2..., T$ are available to make a forecast for period $T + 1$. In the absence of structural breaks in the coefficients or disturbance variance, all available data are utilized for parameter estimation of the forecast model and this is called the full sample forecast. This is often used as a benchmark to assess the performances of forecast methods which allow the possibility of one or more structural breaks. In the presence of structural breaks, several types of approaches are proposed in the literature, which we group through their treatment of the estimation window or of observational weights in the subsections which follow.

3.2.1 Single estimation window

The forecast can be generated using a carefully selected single window which includes a subset of the available data. The main question in relation to this strategy is how much data should be used to estimate the parameters of the forecast model in order to minimize forecast errors according to a measure such as the Mean Squared Forecast Error (MSFE). A standard solution is to use a window that only includes observations after the most recent coefficient break, as it is anticipated these will be more informative with regard to future values. Specifically, given the estimate of the break date $\hat{T}_1$, β_t is estimated using data $[\hat{T}_1 + 1 : T]$, yielding $\hat{\beta}_{\hat{T}_1+1:T}$ and the forecast for $T + 1$ is computed as

$\hat{y}_{T+1} = \hat{\beta}'_{\hat{T}_1+1:T}\,\mathbf{x}_T$. This is referred to as the post break forecast methodology in the literature and it involves pre-testing for a break and estimating the break date using one of the conventional tests such as Andrews (1993), or Bai and Perron (1998, 2003a).

Alternatively, a sequence of papers by Pesaran and Timmermann (2004, 2005, 2007) proves analytically and empirically that the optimal estimation window can include some, but not necessarily all, pre-break observations. The rationale is that, although it increases the bias of OLS, including pre-break observations can reduce the variances of the parameter estimates. For this reason, the optimal window can be selected by trading off bias against reduction of forecast error variance. For practical implementation, Pesaran and Timmermann (2007) propose the trade-off function

$$f(v_1) = \lambda^2(\mu' \sum\nolimits_{v_1} \sum\nolimits_{v}^{-1} \mathbf{x}_T)^2 + \frac{1}{v}(\mathbf{x}'_T \sum\nolimits_{v}^{-1} \mathbf{x}_T)^2 + \frac{\lambda\psi}{v}(\mathbf{x}'_T \sum\nolimits_{v}^{-1} \sum\nolimits_{v_1} \sum\nolimits_{v}^{-1} \mathbf{x}_T)$$

where $\mu = (\hat{\beta}_2 - \hat{\beta}_1)/\hat{\sigma}_2$, $\psi = (\hat{\sigma}_1^2 - \hat{\sigma}_2^2)/\hat{\sigma}_2^2$, $\lambda = v_1/v$ and $v = v_1 + v_2$, with v_1 and v_2 the numbers of pre and post break observations respectively used in estimation. $\hat{\beta}_1$ and $\hat{\sigma}_1^2$ are estimated using $v_1 = \hat{T}_1 - m + 1$ for $1 \leq m < \hat{T}_1 < T$ where m is the first observation used in the estimation window, while $\hat{\beta}_2$ and $\hat{\sigma}_2^2$ are estimated over $v_2 = T - \hat{T}_1$, and

$$\sum\nolimits_{v_1} = v_1^{-1} \sum_{t=m}^{\hat{T}_1} \mathbf{x}'_{t-1}\mathbf{x}_{t-1}, \quad \sum\nolimits_{v_2} = v_2^{-1} \sum_{\hat{T}_1+1}^{T} \mathbf{x}'_{t-1}\mathbf{x}_{t-1}, \quad \sum\nolimits_{v} = \lambda \sum\nolimits_{v_1} + (1-\lambda) \sum\nolimits_{v_2}$$

Selecting v_1 to minimize this function, β_t is estimated over the sample of $[\hat{T}_1 - v_1 + 1 : T]$ yielding in our notation $\hat{\beta}_{\hat{T}_1 - v_1 + 1:T}$, and the forecast is calculated as $\hat{y}_{T+1} = \hat{\beta}'_{\hat{T}_1 - v_1 + 1:T}\,\mathbf{x}_T$. Here, a single break date is assumed, which can apply to coefficients as well as to variances.

Clearly, both post break and trade off methods may be prone to imprecise estimation of the break point location. In practice, the difficulty of estimating the break date is well known, especially when size of the break is small in magnitude (Elliott, 2005).

Another method that carefully selects an estimation window is the cross validation approach of Pesaran and Timmermann (2007). This considers all possible estimation windows with different lengths of observations and chooses a single window which achieves the smallest pseudo out of sample forecast error. Then the selected window is used for parameter estimation of the forecast model. Specifically, for each possible starting point of the estimation window,

m, the recursive pseudo out of sample MSFE value is calculated as

$$MSFE(m|T,\tilde{w}) = \tilde{w}^{-1} \sum_{\tau=T-\tilde{w}}^{T-1} (y_{\tau+1} - \mathbf{x}'_\tau \hat{\beta}_{m:\tau})^2$$

where $\hat{\beta}_{m:\tau}$ is the OLS estimate based on the observation window $[m:\tau]$ and $m \in 1, ..., \min(\hat{T}_1+1, T-\tilde{w}-\underline{w})$, having a minimum estimation window length $\underline{w}$ and reserving the last $\tilde{w}$ observations of the data for an out of sample evaluation. Out of all possible windows, the one which generates the smallest $MSFE$ is selected as

$$m^*(T,\hat{T}_1,\tilde{w},\underline{w}) = \arg \min_{m=1,\ldots,\min(\hat{T}_1+1,T-\tilde{w}-\underline{w})} \left\{ \tilde{w}^{-1} \sum_{\tau=T-\tilde{w}}^{T-1} (y_{\tau+1} - \mathbf{x}'_\tau \hat{\beta}_{m:\tau})^2 \right\}$$

Then, the parameters of the forecast model are estimated on the sample $[m^* : T]$ and the forecast is $\hat{y}_{T+1} = \hat{\beta}'_{m^*:T}\, \mathbf{x}_T$. This method eases the dependence on potentially poor estimation of the break point as it indirectly use information on break point location to restrict the number of estimation windows considered. It also can be adapted for use without any information of break point location. In particular, assuming unknown break date, this method searches m^* along all available $m = 1, \ldots, T - \tilde{w} - \underline{w}$.

A common method that is unconditional on the location of the break point is the rolling window forecast which employs a single estimation window with a constant number of recent observations to estimate the parameters of the forecast model. A difficulty also arises, however, in determining the appropriate length of the rolling window.

Despite the problems corresponding to the above mentioned methods, they generally outperform the natural benchmark of the full sample forecast method which assumes no break when breaks indeed occur. The details of previous simulation results concerning their performances will be discussed in subsection 3.2.5.

3.2.2 Multiple estimation windows

There is a stream of methods that combine forecasts from multiple windows instead of attempting to exploit a single best estimation window. Specifically, forecasts using the same model estimated over different sizes of windows are averaged to generate a single forecast for $T + 1$. In relation to combining forecasts from different windows, various weighting schemes are proposed.

Forecasts based on different windows with sizes spanning $[\underline{w}, T]$ and aver-

aged using equal weights (Pesaran and Timmermann, 2007, Pesaran and Pick, 2011, among others) are given by

$$\hat{y}_{T+1}(T, \underline{w}) = (T - \underline{w})^{-1} \sum_{m=1}^{T-\underline{w}} (\mathbf{x}_T' \hat{\beta}_{m:T})$$

This is often referred as a pooled forecast in the literature and can also be applied when the estimate of the break date is available by setting $m = 1, ..., \min(\hat{T}_1 + 1, T - \underline{w})$, as values of m greater than $\hat{T}_1 + 1$ lead to inefficient estimators as all data after the most recent break are not used.

Alternatively, unequal weights can be employed when pooling forecasts from different windows. For example, weights proportional to the inverses of the associated pseudo out of sample MSFE values (Pesaran and Timmermann, 2007) result in

$$\hat{y}_{T+1}(T, \hat{T}_1, \tilde{w}, \underline{w}) = \frac{\sum_{m=1}^{\min(\hat{T}_1+1, T-\tilde{w}-\underline{w})} (\mathbf{x}_T' \hat{\beta}_{m:T}) \; MSFE^{-1}(m|T, \tilde{w})}{\sum_{m=1}^{\min(\hat{T}_1+1, T-\tilde{w}-\underline{w})} MSFE^{-1}(m|T, \tilde{w})} \tag{3.2}$$

Also, weights based on the values of reversed ordered Cusum (ROC) structural break test statistics, and a location weight which assigns heavier weights to forecasts based on recent observations, are proposed in Tian and Anderson (2011). Their simulation results show an important role for the location weight rather than the ROC weight in window averaging forecasts. In effect, therefore, this procedure gives relatively large weights to the most recent observations. Since we employ other similar types of observational weighting methods, such as exponential smoothing weights, we do not replicate the location weight and thus the details are not provided in this chapter.

These forecast combination methods just mentioned involve all possible windows with different sizes. A version averages forecasts from only two windows: the post break window and the full sample window (Eklund et al., 2013). Specifically, in the context of a recent break, they carry out repeated tests for a break using a Cusum based test, and once a break is detected the forecast is obtained by combining forecasts from using the full sample and the post break window. However, this chapter focuses on accounting for historical breaks, rather than recent breaks and therefore the details of this method are not included in this chapter. Clark and McCracken (2009) also investigate the effectiveness of forecast combination based on recursive and rolling window forecasts: the first uses all available observations in the estimation sample and the latter employs a rolling window of the most recent observations after the break. Linear convex combining weights for these two forecasts are analytically

derived based on the bias-variance trade off characterization. They report the superior performance of this combination method compared to the forecasts either using recursive window only or rolling window only.

In general, combination methods are suited to reduce the bias from imprecise estimation of the break date, as they mainly assume this is unknown. Also, combining forecasts from different subsamples may deliver a better forecast than a single window by canceling upward and downward biased forecasts arising from different windows (Hendry and Clements, 2003).

3.2.3 Observational weight

Averaging forecasts from different windows implicitly assigns decaying weights to older observations. However, some forecast methods, including rolling window, exponential smoothing, optimal and robust observational weights, explicitly weight past observations in parameter estimation. Implementation of the first two methods do not require estimates of the date and size of the break, but the down-weighting parameter should be pre-specified. In the literature, the down-weighing parameter is often chosen arbitrarily and forecasting performance is sensitive to the choice of this parameter (Pesaran et al., 2013).

The recent paper by Giraitis et al. (2013) proposes a data-dependent method to define the down-weighting rate in the context of ongoing structural change. Specifically, their weighting strategy is based on a cross validation method by numerically minimizing the mean squared forecast error of the in-sample forecasts to obtain the down-weighting parameter. Furthermore, their theoretical analysis explores the properties of the new forecasts using a simple model of $y_t = \beta_t + u_t$ where β_t follows a variety of processes including stationary, unit root, deterministic trend and a structural break in the mean, and show the validity of their method in the presence of different forms of structural breaks (see Giraitis et al., 2013, for details). Although interesting, this approach is not replicated in our simulation analysis due to time constraints. Instead, as a representation of this type of method and as it is also found to work quite well in the simulation study by Giraitis et al. (2013), the exponential smoothing method with a fixed down-weighting parameter is implemented in our simulation study. Specifically, exponentially decreasing weights are

$$w_t = \frac{1-\gamma}{1-\gamma^T}\gamma^{T-t} \quad for \quad t = 1, 2, ..., T$$

where the γ is a down-weighting parameter which is subject to forecasters' choice.

Pesaran et al. (2013) proposes optimal and robust observational weighting

schemes, which are more suitable for discrete breaks. These are optimal in the sense that the resulting forecasts minimize the MSFE conditioning on the break size and location and the weights follow a step function that are constant within a regime but different weights across regimes. Specifically, with known parameters, including break size and location, the weights are defined as

$$w_t = \begin{cases} \frac{1}{T}\frac{1}{b+(1-b)(q^2+Tb\phi^2)} & for \quad t \leq \hat{T}_1 \\ \frac{1}{T}\frac{q^2+Tb\phi^2}{b+(1-b)(q^2+Tb\phi^2)} & for \quad t > \hat{T}_1 \end{cases}$$

where $\phi = \frac{\mathbf{x}_T'\theta}{(\mathbf{x}_T'\Omega_{xx}^{-1}\mathbf{x}_T)^{1/2}}$, $E(\mathbf{x}_t\mathbf{x}_t') = \Omega_{xx}$ and $\theta = (\beta_1 - \beta_2)/\sigma_2$, the size of the break relative to the disturbance standard deviation after break. Also, the weights depend on the pre-break sample fraction, $b = \hat{T}_1/T$ and a ratio of standard deviations of the error term before and after break, $q = \sigma_1/\sigma_2$. In practice, these parameters are replaced by their estimates.

An extension of optimal weights to the case of multiple breaks is provided by the authors. We describe their optimal weights in the presence of two breaks in Appendix C.1, as it is appropriate within our framework of Monte Carlo studies in section 3.4 where data generating processes with two breaks are considered. Its performance is then evaluated, in subsection 3.5.2, together with other forecast methods described in subsection 3.3.2.

A robust weighting scheme which is also proposed by Pesaran et al. (2013) does not use information regarding the break date, but uses a range over which the potential break date is assumed to be uniformly distributed. There are two types of robust weights: restricted and unrestricted depending on assumptions regarding the distribution. Assuming that the fraction of break point to full sample, b, is uniformly distributed over the range $\underline{b}$ and $\bar{b}$ with $0 < \underline{b} < \bar{b} < 1$, weights for each point in time can be described as

$$w_t = \begin{cases} 0 & for \quad a < \underline{b} \\ \frac{-1}{T(\bar{b}-\underline{b})}\log(\frac{1-a}{1-\underline{b}}) & for \quad \underline{b} \leq a \leq \bar{b} \\ \frac{-1}{T(\bar{b}-\underline{b})}\log(\frac{1-\bar{b}}{1-\underline{b}}) & for \quad a > \bar{b} \end{cases} \tag{3.3}$$

where $a = t/T$. This is a restricted version of a robust weight. An unrestricted robust weighting scheme is defined when b is allowed to occur any time in $[0, 1]$ as

$$w_t = \begin{cases} \frac{-\log(1-a)}{T-1} & for \quad t = 1, 2, ..., T-1 \\ \frac{\log(T)}{T-1} & for \quad t = T \end{cases}$$

where $a = t/T$. These unrestricted robust weights do not sum to unity as a

discrete time approximation is applied, they are scaled as

$$w_t^* = \frac{w_t}{\sum_{t=1}^{T} w_t}, \quad for\ t = 1, 2, ..., T.$$

Common to all methods based on weighted observations, the parameters of the forecast model are estimated over the estimation period $1, ..., T$ using Weighted Least Squares as

$$\hat{\beta}_T = \left(\sum_{t=1}^{T} w_t \mathbf{x}_{t-1}' \mathbf{x}_{t-1} \right)^{-1} \sum_{t=1}^{T} w_t \mathbf{x}_{t-1}' y_t$$

where $\sum_{t=1}^{T} w_t = 1$ and the forecast for $T+1$ is computed as,

$$\hat{y}_{T+1} = \hat{\beta}_T' \ \mathbf{x}_T.$$

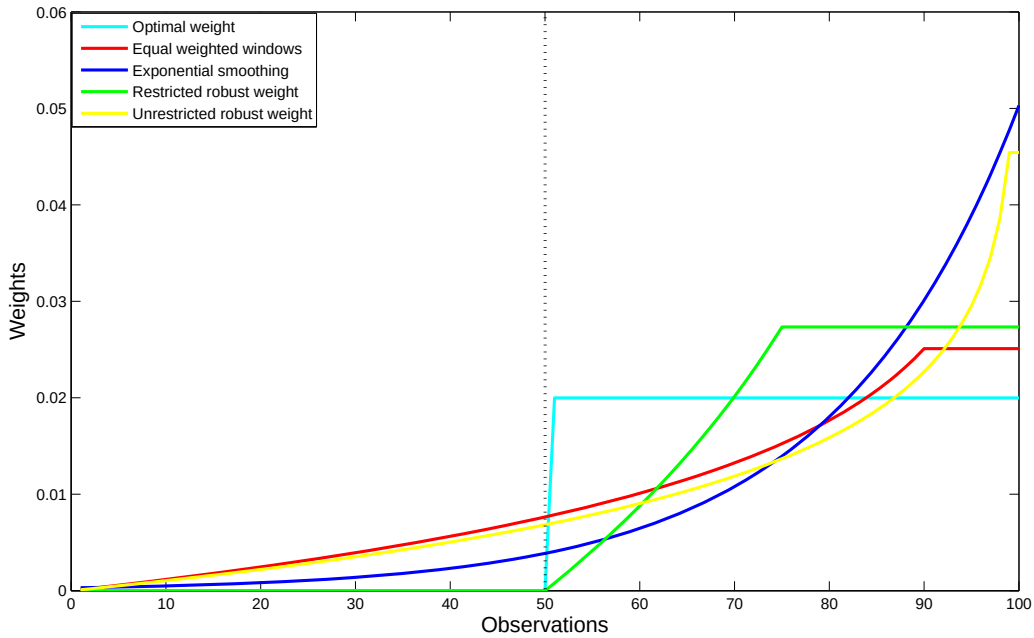
3.2.4 Approximated Weights

Understanding the relationships between window based forecasts in subsections 3.2.1 and 3.2.2, and observational weighted forecasts in 3.2.3 is important to identify the underlying key factors for the relative performance of different forecast methods under a variety of situations. Perhaps, the easiest way is to translate weights on the forecasts from different windows into weights on the observations. For instance, window averaging with equal weights implies smoothly decaying weights on the observations (Pesaran et al., 2013). This is because each forecast is generated using an estimation window which includes at least the last $\underline{w}$ observations and expands backward. Overall, more recent observations are used in every estimation window and hence they carry higher weights compared to earlier observations.

Using the same logic, a given window weighting scheme can be approximated in terms of observational weights using the relative importance of each observation for the implicit $\hat{\beta}_T$. In general, this approximation does not precisely replicate the forecasts from schemes such as window averaging or single window methods. However, the equivalence does hold for mean only forecast models, namely, $y_t = \alpha_t + \varepsilon_t$ (Pesaran and Pick, 2011). The details of approximated observational weights are outlined in Appendix C.2. To provide a visual intuition, figure 3.1 plots some selected observational weights for the case of homoskedastic disturbances.

In brief, an equal weighted window averaging forecast assigns smoothly increasing weights throughout the sample compared to the unrestricted robust and exponential smoothing weights which get steeper toward the end of

Figure 3.1: Weights for observations



Notes: The vertical dotted line indicates the location of the break used in the optimal weighting scheme. The restricted robust weight is plotted based on the assumption that the break date is uniformly distributed in the observation range of [50:75]. All other forecast methods are unconditional on the estimate of break date. The exponential smoothing parameter in this figure is $\gamma = 0.95$. A minimum estimation window of 10 observations is employed for the equal weighted window averaging method.

sample, allocating relatively higher weights to the most recent observations. The optimal weighting scheme appears to be similar to the implied weights for post break[1] window method when there is a small single break in the autoregressive coefficient. However, these could be substantially different when the variance of the disturbances is subject to a change (details of these weights are in Pesaran et al. (2013)). Finally, the restricted robust weights assign increasing weights within a range, with constant and the highest weights on recent observations.

3.2.5 A summary of previous findings

Based on simulation results in the literature, the following conclusions emerge. Firstly, the relative performance of forecasting methods depends on the size of the break. When the break is small or only occurs in the disturbance variance, forecast combination methods deliver good forecasts (Pesaran and

[1]The shape of weights corresponding to the single window forecasts, including post break, cross validation and trade-off, are similar to the shape of the optimal weighting scheme. Differences lie in defining the origin of the estimation sample but all observations within the estimation sample receive equal weights.

Timmermann, 2004, 2007, and others). However, when the break in coefficients is large, the forecast errors of estimation windows that include more pre-break observations are high and including these deteriorates overall forecast accuracy. This is because the correlation between forecast and realization weakens for windows that include both pre and post break observations when coefficient differences are large (Pesaran and Timmermann, 2004). On the other hand, methods involving pre-testing break dates for example, cross validation, trade-off and post break window methods, perform quite well when the break in coefficients is large (Pesaran and Timmermann, 2007). It should be noted however, that the gain from pre-testing for breaks is only marginal even if their size is large when breaks appear towards the end of the estimation sample (Eklund et al., 2013). This is due to the small number of observations available after the break for estimating the parameters of the forecast model.

Secondly, forecast accuracy crucially depends on how precisely the timing of the break is estimated (Pesaran and Timmermann, 2007, Pesaran and Pick, 2011, Elliott, 2005, Pesaran et al., 2013, among others). This issue is especially relevant for methods which directly hinge on estimated break dates. In practice, estimating the timing of breaks is notoriously difficult when breaks are small in magnitude. The widely used methods based on least squares can perform poorly for estimating the break date. As shown by the Monte Carlo experiment in Elliott (2005), when the break is small, the distribution of break dates is spread evenly across all possible dates, except for peaks around the true break date and smaller peaks at the two ends of the sample. The bias of the forecast then depends on how far the break is identified from the true break date: the larger the distance the higher the bias. On the other hand, the presence of a larger break increases the chance that its date is estimated well. However, Elliott (2005) also shows that structural break testing procedures based on least squares are highly asymmetric in estimating the break date. In particular, most breaks are estimated to appear after the true break date, which increases the variances of parameter estimation leading to larger mean squared forecast errors.

Thirdly, breaks in variance appear to have a negligible effect on forecast accuracy. None of above approaches tests for the location and size of variance breaks or handles them explicitly in the forecast. However, the optimal weighting scheme in Pesaran et al. (2013) uses the size of variance break, which is calculated using available observations before and after the estimated coefficient break date assuming that variance changes at the same time as the change in coefficients. If the break occurs at any other time, the size of the break may be estimated inaccurately. Some studies indirectly use informa-

tion regarding the direction of variance changes for selecting window size or observational weights. Pesaran and Timmermann (2007) shows analytically that if the variance increases after the break, including more pre-break observations in the estimation window is optimal provided that the size of the coefficient break and the number of post break observations are small. Also Pesaran et al. (2013) propose robust weights on observations which monotonically rise with time when the pre-break variance is less than the post break variance. However, these may not be sufficient methods to deal with variance breaks. Unaccounted variance breaks complicate the identification of coefficient breaks in the data generating process so that overall forecast accuracy could be reduced.

Finally, forecast methods which involve estimation of the break date are well suited in the presence of a single break. The prevailing way to apply them to multiple breaks is to employ the most recent break only (Pesaran and Timmermann, 2002, 2005, 2007, and others). Specifically, a multiple[2] structural break testing methodology such as Bai and Perron (1998) is applied, and then information from the last break date is used when more than a single break is estimated. This is due to the assumption of higher relevance of most recent observations for forecasting purpose. However, information from earlier regimes could be relevant, especially when breaks have the form of the 'reverting' coefficients in which the detected breaks capture switches between two or more distinct but recurring regimes.

3.2.6 Other relevant literatures

This chapter focuses on forecast issues related to structural break(s) in the dynamics[3] of a single equation forecast model, and breaks are assumed to occur in the estimation sample. Additionally, this chapter does not investigate more complex models for example, models with multiple explanatory variables. Instead, we employ small scale VAR models with a single exogenous variable for the Monte Carlo simulation and univariate autoregressive models for the empirical study. However, for completeness, this subsection briefly overviews other forecast methods that deal with deterministic breaks, model uncertainties, and breaks in the forecast horizon which are not investigated further in this chapter.

[2]Pesaran and Timmermann (2002) also used the reversed order Cusum test as an alternative to estimating multiple breaks. Specifically, as the last break is assumed to be empirically relevant, a single break testing Cusum procedure is applied to the reverse ordered data in which the first identified break date is the last break in the non reversed time series.

[3]See section 3.4 for the details of a Monte Carlo simulation where long run means of the series are set up to remain unchanged.

In addition to the methods examined in this chapter, other tools are available to deal with uncertainties of structural breaks including differencing and intercept correction. Hendry and Clements (2003), and Castle and Hendry (2008) claim that breaks in deterministic terms are the major source of forecast failure in practice and suggest using intercept corrections in which the intercept is adjusted using realized equation errors when such a break is present. Moreover, in the presence of deterministic breaks, differencing (or double differencing) some or all variables in forecasting models works well because it removes intercepts and linear trends (Clark and McCracken, 2008, Hendry and Clements, 2003).

Uncertainties associated with model selection are well studied in the forecast literature. It is almost impossible to replicate true data generating process, as a model can be misspecified in many different ways and a poorly selected single model can be anticipated to yield poor forecasts. In this respect, the Bayesian model averaging technique is widely used, whereby different models use the same set of data to produce forecasts which are averaged (Clark and McCracken, 2008, and others). Also, as an alternative to the Bayesian approach to specification of the weights, information-theoretic model averaging method is proposed by Kapetanios et al. (2008) in which various models are weighted based on their relative model likelihoods. However, model averaging techniques often make an implicit assumption of model stability while at least some of models may be subject to breaks. To overcome this problem, Pesaran et al. (2009) propose a double averaging method that averages forecasts over different models and different observation windows. Specifically, it employs Bayesian model averaging in which coefficient changes in each model are accounted for by averaging over different estimation windows. They note that double averaged forecasts outperform methods based on model averages only and window averages only.

Furthermore, most studies employ simple linear models with a single exogenous variable or univariate autoregressive models. Clark and McCracken (2008) assess forecasting performance using small scale VAR models, each comprises of different measures of output, inflation and short run interest rates. They employ a total of 86 different forecasting methods incorporating choice of lags, estimation windows, level and difference, intercept correction, time varying parameters, break dating, discounted least square, Bayesian shrinkage, de-trending and model averaging methods. Aggregating all models, horizons and variables, model averaging and Bayesian shrinkage methods perform best, while fixed width rolling window and its less extreme version – discounted least squares method appear to yield the worst forecasts.

The forecasting performance of the Global VAR (GVAR) specification that take into account the increasing interdependencies across countries and markets, is examined by Pesaran et al. (2009). They find improved forecasting accuracy from the GVAR model employed with the double averaging forecast combination method, compared to its competitors, namely univariate autoregressive and random walk models. Moreover, Banerjee et al. (2008) attempt to improve forecasting accuracy in a short sample with structural change by using a diffusion index (common factor). Specifically, a model with a diffusion index, which is extracted using a large amount of data, is compared with autoregressive and VAR models with robust forecasting devices such as differencing and intercept correction. They find good performance of the diffusion index, especially when its factor loading is time varying, compared to other candidate forecasting models.

The above mentioned studies assume single or multiple breaks occurring in the estimation period. Another interesting framework which is not studied in this chapter is the possibility of a break occurring in the forecast horizon. Pesaran et al. (2006) propose a new approach that addresses the problem of breaks occurring not only in the estimation period but also in the forecast horizon. They allow random breaks in a forecast horizon using a Hierarchical Markov Chain method. Specifically, they employ the Bayesian approach to choose the number of breaks in which the marginal likelihood estimate of each model with n breaks is calculated and ranked using a Bayes factor. Then, given the number of breaks, probabilities of all possible break dates are computed using the Hierarchical Markov Chain where parameters are drawn from some common meta distribution. They show an improved forecasting performance over a range of methods including those which assume no break in the forecast horizons and time varying parameter models which assume breaks in each period.

Overall, these studies suggest that use of advanced methods rather than simple constant parameter autoregressive models could help to extrapolate more accurate forecasts. This chapter pursues this in a single equation context.

3.3 Methodology

This section describes our proposed forecasting methodologies. Recall equation (3.1) where we assume up to n breaks occur in a set of coefficients β_t at times $T_1, T_2, ..., T_n$. We also consider a scenario where the variance of disturbances σ_t^2 may be subject to structural changes, but these changes do not necessarily

coincide with coefficient breaks. Our interest lies in forecasting y_{T+1} given available observations of $\Gamma_t = \{y_t, \mathbf{x}_t : t = 1, 2, ..., T\}$, recognizing that break dates may not be well estimated.

3.3.1 Single break confidence interval forecast

Suppose that we do not know the true break date but estimate it as $\hat{T}_1$, using an appropriate method. Instead of a point estimate $\hat{T}_1$, we employ a confidence interval to reduce loss associated with a poor single break date estimate, since the confidence interval provides a collection of potential break dates. For convenience of exposition, assume that the dates within the confidence interval are contiguous and denote $[\hat{T}_{1L}, \hat{T}_{1U}]$ as the lower and upper bounds of the confidence interval for T_1, which we will describe how to obtain in subsection 3.3.4. Then treating each date in the interval as one of a sequence of choices for the potential break date, the corresponding post break window forecasts are averaged with equal weights as

$$\hat{y}_{T+1,CIE} = \frac{1}{\hat{T}_{1U} - \hat{T}_{1L} + 1} \sum_{t=\hat{T}_{1L}+1}^{\hat{T}_{1U}+1} \hat{\beta}'_{t:T}\, \mathbf{x}_T. \tag{3.4}$$

Furthermore, to ensure that a minimum $\underline{w}$ observations are used in each estimation, $\hat{T}_{1U} + 1$ in the summation is replaced by $\min(\hat{T}_{1U} + 1, T - \underline{w})$. If $T - \underline{w} < \hat{T}_{1U} + 1$, the denominator of (3.4) is then also adjusted. In an analogous manner to the discussion in subsection 3.2.4, approximated observational weights for the confidence interval approach are given by

$$w_t = \begin{cases} 0 & for \quad t \leq \hat{T}_{1L} \\ \frac{1}{(\hat{T}_{1U} - \hat{T}_{1L} + 1)} \sum_{m=\hat{T}_{1L}+1}^{t} \frac{1}{T-m+1} & for \quad \hat{T}_{1L} + 1 \leq t \leq \hat{T}_{1U} + 1 \\ \frac{1}{(\hat{T}_{1U} - \hat{T}_{1L} + 1)} \sum_{m=\hat{T}_{1L}+1}^{\hat{T}_{1U}+1} \frac{1}{T-m+1} & for \quad \hat{T}_{1U} + 1 < t \leq T \end{cases}$$

and the weights sum to one.

We also employ cross-validation weights proposed by Pesaran and Timmermann (2007) within a confidence interval. Specifically, we adopt the framework of cross validation by reserving $\tilde{w}$ observations for the pseudo out of sample evaluation. Then, in a similar way to equation (3.2), estimation windows with starting points $m \in \hat{T}_{1L} + 1, ..., \min(\hat{T}_{1U} + 1, T - \tilde{w} - \underline{w})$ are evaluated in $\tilde{w}$, and the corresponding $MSFEs$ are used to weight the forecasts. The forecast

for $T+1$ period is

$$\hat{y}_{T+1,CIW}(T,\hat{T}_{1L},\hat{T}_{1U},\tilde{w},\underline{w}) = \frac{\sum_{m=\hat{T}_{1L}+1}^{\min(\hat{T}_{1U}+1,T-\tilde{w}-\underline{w})}(\mathbf{x}_T'\hat{\beta}_{m:T})\ MSFE(m|T,\tilde{w})}{\sum_{m=\hat{T}_{1L}+1}^{\min(\hat{T}_{1U}+1,T-\tilde{w}-\underline{w})} MSFE(m|T,\tilde{w})} \tag{3.5}$$

and the corresponding approximate observational weights are

$$w_t = \begin{cases} 0 & for \quad t \leq \hat{T}_{1L} \\ \sum_{m=\hat{T}_{1L}+1}^{t} \frac{\mu_m}{T-m+1} & for \quad \hat{T}_{1L}+1 \leq t \leq \hat{T}_{1U}+1 \\ \sum_{m=\hat{T}_{1L}+1}^{\hat{T}_{1U}+1} \frac{\mu_m}{T-m+1} & for \quad \hat{T}_{1U}+1 < t \leq T \end{cases}$$

where $\mu_m = MSFE(m|T,\tilde{w})/\sum_{m=\hat{T}_{1L}+1}^{\hat{T}_{1U}+1} MSFE(m|T,\tilde{w})$. Further, when $\tilde{w}$ is large relative to the full sample and the estimated break date is close to the end of the sample, both the lower and upper bounds exceed $T-\tilde{w}-\underline{w}$. In that case, we estimate the parameters of the model on the sample $[T-\tilde{w}-\underline{w}:T]$.

We further pursue the idea of using a confidence interval for the estimated break date to develop the robust weighting scheme proposed by Pesaran et al. (2013). In order to derive the restricted robust weighting scheme, they make an assumption that the break date fraction of the full sample, b is uniformly distributed over the range $\underline{b}$, $\overline{b}$ with $0 < \underline{b} < \overline{b} < 1$. The range of b is selected arbitrarily, yet it may yield a poor forecast when the true break date is out of the pre-specified range. One way to develop this method is to set the lower and upper bounds of the contiguous confidence interval as $\underline{b}$ and $\overline{b}$, respectively. In this way, the true break date is likely to be included within the range. Using restricted robust weights in equation (3.3), the weights are described as

$$w_t = \begin{cases} 0 & for \quad a < \hat{T}_{1L}/T \\ \frac{-1}{(\hat{T}_{1U}-\hat{T}_{1L})}\log(\frac{1-a}{1-(\hat{T}_{1L}/T)}) & for \quad \hat{T}_{1L}/T \leq a \leq \hat{T}_{1U}/T \\ \frac{-1}{(\hat{T}_{1U}-\hat{T}_{1L})}\log(\frac{1-(\hat{T}_{1U}/T)}{1-(\hat{T}_{1L}/T)}) & for \quad a > \hat{T}_{1U}/T \end{cases} \tag{3.6}$$

where $a = t/T$.

3.3.2 Multiple breaks confidence interval method

So far we have assumed the presence of a single break when describing various forecast methods. However, in practice time series may be subject to multiple structural breaks. In relation to multiple breaks, we consider two scenarios in the data generating process (DGP). First, changes in a set of coefficients are in the same direction after each break by either increasing or decreasing. Second, breaks have the form of reverting coefficients, in which the values shift

after the first break but revert to the original values after the second break. The latter case is interesting in this context, as it implies that earlier and later observations can be informative, but observations between the first and second breaks have the least information regarding the DGP relevant for forecasting.

We propose a method designed to improve forecast accuracy when there are multiple breaks which lead to reversion in the coefficients. First, we test for multiple breaks in equation (3.1) using the test procedure of Bai and Perron (1998). Suppose that n breaks are estimated at $\hat{T}_1, \hat{T}_2, ..., \hat{T}_n$ and those breaks divide the full sample into $n+1$ segments of observations. By re-ordering such segments based on the relative closeness of their estimated coefficients, we aim to exploit more information and improve forecasting performance. The process of re-ordering segments is described in the following.

Firstly, we estimate the set of coefficients in each segment conditional on the estimated break dates:

$$\beta'_t = \begin{cases} \beta'_1 & for \quad 1 \leq t \leq \hat{T}_1 \\ \beta'_2 & for \quad \hat{T}_1 < t \leq \hat{T}_2 \\ \vdots & \vdots \qquad \vdots \\ \beta'_{n+1} & for \quad \hat{T}_n < t \leq T \end{cases}$$

Treating the break dates as known we then employ the Chow (1960) test to examine whether a set of estimated coefficients in each segment is statistically different from the coefficients in the final data segment. Specifically, the hypotheses are

$$\begin{cases} H_0 : \beta'_1 = \beta'_{n+1} & vs \quad H_A : \beta'_1 \neq \beta'_{n+1} \\ H_0 : \beta'_2 = \beta'_{n+1} & vs \quad H_A : \beta'_2 \neq \beta'_{n+1} \\ \vdots & \vdots \qquad \vdots \\ H_0 : \beta'_n = \beta'_{n+1} & vs \quad H_A : \beta'_n \neq \beta'_{n+1} \end{cases} \tag{3.7}$$

After conducting this sequence of coefficient equality tests, the corresponding p-values are saved. Based on the 5% significance level, any p-values higher than 0.05 are taken to justify the union of the corresponding segments with the last $(n+1^{th})$ segment as they are judged to be not statistically different from those of the final segment. If more than one segment needs to be combined with the $n+1^{th}$ segment, their corresponding p-values are used to decide the order of segments. That is, the smaller the p-value, the higher the rejection level and thus the corresponding segment should be located further from the $n+1^{th}$ segment. On the other hand, any p-values smaller than 0.05 are taken to indicate differences in the estimated coefficients and such segments are not

Figure 3.2: Using the most recent break

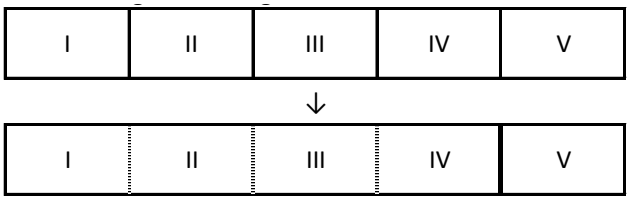
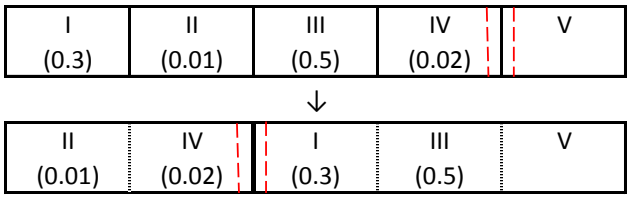


Figure 3.3: Data Re-ordering



combined with the last segment of observations. Although not combined with the last segment, these segments are also re-ordered based on their p-value ranks, so that the "most similar" (in p-value terms) are located closest to the final segment. We should note however, that the Chow test is based on known break dates, whereas our break dates are estimated, so p-values may appear more significant than they are.

For example, say, we find 4 breaks which divide the full sample into 5 segments. As mentioned previously, the majority of studies recognize only the most recent break in forecasting when multiple breaks are estimated. This is described in figure 3.2 by the solid bold line and the remaining estimated breaks represented by dotted lines are ignored. Figure 3.3 sketches an example to describe our process of re-ordering the original data. Hypothetical p-values from testing coefficient equality of each segment against the last segment are given in brackets. Then, data are re-ordered [4] as shown in figure 3.3 where we also use the most recent break in forecasting, indicated by the solid bold line. However, due to re-ordering data, the location of the last break moves to reflect the pooled segments. In relation to this, bounds of the corresponding confidence interval also should move. To keep things simple, we apply the width of the original confidence interval to the new break on both sides. This is described by hypothetical lower and upper bounds for the confidence interval using red dashed lines.

In terms of the tests of (3.7), we anticipate that the last null hypothesis $H_0 : \beta'_n = \beta'_{n+1}$ is rejected, as the parameter differences in last two segments are implied by the last break date in the original sample. In this sense, the last two segments are never joined and thus the last break date is always maintained

[4] In the procedure, not only the value of dependent variable is re-ordered but also all regressor values are re-ordered in the same way of dependent variable.

while other breaks may disappear once the segments are re-ordered. However, unless account is taken of possible variance breaks, it is also possible that the apparent significance may be due to a variance break rather than a coefficient break.

The Chow (1960) test is only valid when disturbance variances are stable over time. To allow for the possibility of changing variances, we assume that any change in the disturbance variances occur at the same time as the coefficient breaks and perform a GLS transformation to the original data when estimating the coefficients in each regime and testing the hypotheses of (3.7). However, in the iterative structural break testing methodology in which variance breaks are tested as in Chapters 1 and 2, coefficient tests are applied to variance break adjusted data without making any prior assumption on the dates of variance breaks.

Finally, we apply all single break forecast methods described in subsection 3.3.1 as well as some of the relevant methods in the literature on both re-ordered and original data by exploiting information on the last break, and assess any value added by our re-ordering data procedure.

3.3.3 Structural break testing methodology

The methodology we employ for structural break detection uses, firstly, a heteroskedastic version of Bai and Perron (1998) testing procedure to test for a break in the set of coefficients in equation (3.1). Later, we employ an iterative structural break testing methodology in order to identify breaks in both coefficients and variances which do not necessarily occur at the same time.

Variance breaks are often overlooked in the forecast literature due to the dominant effect of coefficient breaks. However, testing for coefficient breaks requires consideration be given to potential breaks in the variance because the presence of variance breaks can affect inferences on coefficient breaks (Detailed discussion is provided in subsection 1.3.1 in chapter 1 and subsection 2.2.2 in chapter 2). For example, the testing procedure on coefficients could wrongly identify a variance break as a coefficient break while indeed there is none. In this case, a spurious coefficient break will lead to increased MSFE, since estimation is less efficient than using all observations.

Steps for the iterative procedure are outlined in the following. Although this is similar to that used in the previous chapters, the different procedure in testing for variance breaks should be noted, namely this is based on the absolute value, rather than the squares, of the residuals. The absolute value specification is robust to non-normality of residuals compared to the conven-

tional variance estimator, a mean of squared residuals (Davidian and Carroll, 1987, McConnell and Perez-Quiros, 2000). Additionally, in our simulation, the absolute value specification is found to work well in identifying the timing of the break compared to employing the mean of squared residuals.

Step 1 - Preliminary coefficient break test The Bai and Perron (1998) multiple structural breaks testing procedure is applied to the coefficients, β, employing heteroskedasticity consistent (HC) inference.

Step 2 - Variance break test Using residuals from the model after allowing for coefficient breaks identified in step 1, variance breaks are tested using the absolute value specification. Specifically, the test regression is

$$\sqrt{\frac{\pi}{2}}|\hat{\varepsilon}_t| = \alpha + \epsilon_t \tag{3.8}$$

to which the homoskedastic version of the Bai and Perron (1998) multiple break testing methodology is applied. The resulting $\hat{\alpha}_t$ yields the estimated disturbance standard deviation for each variance regime.

Step 3 - Coefficient break test Since HC inference can lead to oversized coefficient break tests (Bai and Perron, 2006), breaks in the coefficients are reconsidered conditional on the variance breaks from step 2. Following the proposal of Pitarakis (2004), this is achieved by applying homoskedastic inference after applying the feasible GLS transformation. If no variance breaks are detected, coefficient tests are applied to the original data with a homoskedastic variance assumption.

3.3.4 Estimating confidence intervals

We employ two procedures that have been proposed in the literature for estimating confidence intervals for break dates.

3.3.4.1 Confidence interval I

As described in the previous subsection, Bai and Perron (1998) testing procedure is used to test for breaks and, along with the coefficient break dates, we obtain their corresponding 95% confidence intervals. The confidence intervals are constructed using the asymptotic framework of the break dates. Recall equation (3.1), where we assume n breaks occur in a set of coefficients β_t. For $j = 1, ..., n$, let $\hat{\triangle}_j = \hat{\beta}_{j+1} - \hat{\beta}_j$ be the difference between estimated coefficients in consecutive data segments, and $\hat{\sigma}_j^2$ be the estimated variance of disturbances that are serially uncorrelated and assumed to be homoskedastic within each

coefficient segment which can be estimated using $\hat{\sigma}_j^2 = (\triangle\hat{T}_j)^{-1}\sum_{t=\hat{T}_{j-1}+1}^{\hat{T}_j}\hat{\varepsilon}_t^2$ with $\triangle\hat{T}_j = \hat{T}_j - \hat{T}_{j-1}$. As the regressors are assumed to be identically distributed across regimes, define $\hat{Q} = T^{-1}\sum_{t=1}^{T}\mathbf{x}_{t-1}'\mathbf{x}_{t-1}$ and then the confidence intervals can be constructed using the following approximation

$$\frac{(\hat{\triangle}_j'\hat{Q}\hat{\triangle}_j)}{\hat{\sigma}_j^2}(\hat{T}_j - T_j^0) \Longrightarrow \arg\max_s\{W^{(j)}(s) - |s|/2\} \tag{3.9}$$

where T_j^0 indicates a true break date and $W^{(j)}(s)$ denotes a two-sided Brownian motion; see Bai and Perron (1998).

3.3.4.2 Confidence interval II

Despite the popularity of the Bai and Perron (1998) procedure, the coverage rate for the associated confidence intervals are far below the nominal rate for small breaks, as shown by Elliott and Muller (2007) and by the simulation results in Bai and Perron (2006). Therefore, we also use the confidence interval proposed by Elliott and Muller (2007) to assess the robustness of forecast performances based on the Bai and Perron (1998) confidence interval. However, the Elliott and Muller (2007) confidence interval is constructed based on the assumption that a single break occurs in time series, hence it is compared with Bai and Perron (1998) confidence interval only for a single break cases.

The confidence interval is constructed based on series of hypothesis tests for the maintained break at times $\tau_m = 1, ..., T$. Specifically, the null hypothesis that time τ_m is the true break date τ_0 is tested against the alternative that the break occurs at some other time, namely

$$H_0 : \tau_0 = \tau_m \; against \; H_1 : \tau_0 \neq \tau_m$$

using the statistic

$$\begin{aligned} U_T(\tau_m) &= \tau_m^{-2}\sum_{t=1}^{\tau_m}(\sum_{s=1}^{t}\mathbf{u_s})'\hat{\Omega}_1^{-1}(\sum_{s=1}^{t}\mathbf{u_s}) + \\ &\quad (T-\tau_m)^{-2}\sum_{t=\tau_m+1}^{T}(\sum_{s=\tau_m+1}^{t}\mathbf{u_s})'\hat{\Omega}_2^{-1}(\sum_{s=\tau_m+1}^{t}\mathbf{u_s}) \end{aligned} \tag{3.10}$$

where $\{\mathbf{u_t}\}_{t=1}^{T} = \{\mathbf{x}_{t-1}\hat{\varepsilon}_t\}_{t=1}^{T}$ and $\hat{\varepsilon}_t$ are the estimated residuals from equation (3.1); $\hat{\Omega}_1$ and $\hat{\Omega}_2$ are the long run variance estimators of $\{\mathbf{u_t}\}_{t=1}^{\tau_m}$ and $\{\mathbf{u_t}\}_{t=\tau_m+1}^{T}$, respectively. The computed test statistics are compared with the critical values tabulated by Elliott and Muller (2007). When $\hat{U}_T(\tau_m) < CV$, include τ_m in the confidence interval as it is not rejected by the test, and ex-

clude it otherwise. The testing procedure runs for each point in time, and the resulting confidence interval need not be a contiguous set.

3.4 Monte carlo simulations

We conduct Monte carlo simulations to evaluate our proposed forecast methodologies, following the simulation setup in Pesaran and Timmermann (2007). A similar setting is also adopted in Clark and McCracken (2005) and Tian and Anderson (2011).

Consider the following bivariate VAR(1) DGP,

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} \mu_{yt} \\ \mu_{xt} \end{pmatrix} + \mathbf{A}_t \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{xt} \end{pmatrix} \tag{3.11}$$

where the pre-break unconditional mean is

$$\mu_0 = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$$

and the intercept vector

$$\begin{pmatrix} \mu_{yt} \\ \mu_{xt} \end{pmatrix} = (I - A_t)^{-1}\mu_0$$

is such that the long run means of the series remain unchanged. The VAR coefficient matrix is

$$\mathbf{A}_t = \begin{pmatrix} \beta_{11t} & \beta_{12t} \\ 0 & \beta_{22t} \end{pmatrix}$$

with $\beta_{21} = 0$ as x Granger causes y and not vice versa, with the parameters in $\mathbf{A}_t$ subject to changes. Error terms are normally distributed with variance covariance matrix

$$\begin{pmatrix} \sigma^2_{\varepsilon_{yt}} & E(\varepsilon_{yt}\varepsilon_{xt}) \\ E(\varepsilon_{xt}\varepsilon_{yt}) & \sigma^2_{\varepsilon_{xt}} \end{pmatrix} = \begin{pmatrix} \sigma^2_{yt} & 0 \\ 0 & 1 \end{pmatrix}$$

where σ^2_{yt} may also be subject to changes. Details of the DGPs with breaks in $\mathbf{A}_t$ and σ^2_t are given in Table 3.1 and Table 3.2 for single and multiple break cases, respectively. In the testing procedure, we examine breaks in the first equation of (3.11).

Experiment 1 in Table 3.1 corresponds to the case that the parameters of the model are unaffected by a break in which we anticipate the best fore-

casting performance to be the benchmark which ignores a break and uses all observations. In experiments 2 and 3, a break occurs in the autoregressive coefficient, β_{11t}, which leads to a small (declined by 0.2) and large (declined by 0.4) changes after the break, respectively. Similarly, the effects of a small (increased by 0.5) and large (increased by 1) changes to the coefficient of the lagged exogenous variable on y_t, β_{12t}, are considered in experiments 4 and 5, respectively. Experiment 6 is the combination of experiments 2 and 5, where a single break affects both β_{11t} and β_{12t} coefficients simultaneously. The effects of increasing and decreasing variances are studied in experiments 7 and 8, with the post-break variance increasing by a factor of four and declining by a factor of two, respectively. In addition to the experiments in Pesaran and Timmermann (2007), we consider situations in which all parameters of the model are affected by the break. Specifically, experiment 9 introduces simultaneous changes in coefficients and variances by combining experiments 6 and 7, and similarly for experiment 10 which is the combination of experiments 6 and 8.

In the single break DGPs, we set a single break occurring at 1/4 (quarter), 1/2 (half) and 3/4 (three-quarters) of the full sample size in order to examine the sensitivity of the break point location for forecasting performance. We employ the Bai and Perron (1998) procedure to test for breaks in which we allow a maximum of one break with trimming $\epsilon = 0.10$ (10% of the full sample).

We also employ a DGP with two breaks in the Monte Carlo experiments, which occur at 1/3 (one-third) and 2/3 (two-thirds) of the full sample. However, we allow up to three breaks in the break point testing procedure with trimming $\epsilon = 0.10$ (10% of the full sample). More specifically, first the WDmax test is employed to check the presence of at least one break. If the WDmax test indicates the presence of any breaks, then the sequential $Sup\ F(l+1|l)$ procedure is applied to estimate the number of breaks starting from the test $Sup\ F(2|1)$. If WDmax fails to reject the null hypothesis of no break, the procedure stops and concludes the presence of no break. All tests are conducted at the 5% significance level.

We consider two patterns of coefficients in DGPs with two breaks, with details provided in Table 3.2. A scenario where the coefficients either increase or decrease in the same direction after each break is investigated in experiments 12, 14 and 16; changes occur in β_{11t} for the first two and in β_{12t} for the latter. We do not expect substantial differences in terms of forecast performances using original and re-ordered data for these cases. More interesting scenarios are in experiments 11, 13, 15 where reversion in coefficients is considered, with coefficients equal in the first and third segments, where data re-ordering may

yield some forecast benefits. Additionally, experiments 17 and 18 cover cases of decreasing and increasing variances after the break, in which there are no coefficient breaks and a single variance break occurs at 2/3 of the sample.

Furthermore, to assess the impact of the sample size on forecasting performance, sample sizes $T = 100, 200$ and 500 are employed in the simulations. For each DGP, as an initialization, we set the first value of $y_{t-1} = \mu_{yt}$ and simulate $T_0 + T + 1$ observations for the corresponding DGP with $T_0 = 100$. Then the first T_0 is thrown away and observations $T_0 + 1, \ldots, T$ are used for the parameter estimation to generate a forecast for the observation $T + 1$, and repeat the simulation 5000 times[5].

The forecast accuracy of the methods described in section 3.3 as well as methods in section 3.2 are assessed in the Monte Carlo experiments based on sample Mean Squared Forecast Error (MSFE), the average squared difference between forecasted and realized values as

$$MSFE = 1/S \sum_{i=1}^{S} (y_{T+1} - \hat{y}_{T+1})^2$$

where S denotes the number of simulations. In the results reported, the computed MSFE for each method is divided by the MSFE of the benchmark model which uses the full sample in the parameter estimation regardless of the presence of breaks. Ratios lower than 1 indicate better performances of the corresponding methods than the benchmark, and higher than 1 points to worse performances compared to the benchmark model.

Although the evaluation of point forecasts is widely based on the smallest MSFEs or its alternative MAEs (Mean Absolute Errors) across different forecast methods, more advanced tests for evaluating the accuracy of a method relative to another method are available in the literature. An extensive review of the relevant literature including recent developments in the forecast evaluation is provided by Clark and McCracken (2013). For example, tests of equal predictive accuracy – the MSE-t and MSE-F, can be applied to forecasts from nested models where the tests have non-standard asymptotic distributions and require the use of bootstrapped critical values (see Clark and McCracken, 2013, for details). Alternatively, equal predictive accuracy of forecasts from nested models can be tested based on the comparison of MSFEs after adjusting for the upward bias in the MSFE of the larger model, proposed by Clark and West (2007). However, these tests are not appropriate for our forecast method comparison and more applicable test to forecasts from non-nested models by

[5]The initial seed is set for each DGP so that all forecasting methods are evaluated based on exactly the same sample data.

Diebold and Mariano (2002) is employed in our empirical analysis in section 3.6.

Furthermore, researchers often make inferences about the relative forecast performance of competing models based on (not limited to) above mentioned tests which crucially rely on the size of the estimation window. On the other hand, the choice of the estimation window affects the size and power of tests in assessing the null hypothesis of equal predictive ability (Rossi and Inoue, 2012). Rossi and Inoue (2012) proposes a method that is robust to the choice of the estimation window size. Specifically, they evaluate the performance of forecast models using different window sizes and then taking summary statistics, and this method is applicable to many of existing predictive ability tests to forecasts from non-nested and nested models. However, due to time constraints, this methodology is not adapted in this chapter.

Finally, in the forecast methods, we allow a minimum estimation window $\underline{w}$ equals to 10% and the length of the pseudo out of sample forecast period used in cross-validation, $\tilde{w}$, equals to 25% of the full sample data. For the exponential smoothing method, we set $\gamma = 0.95$ and $\gamma = 0.98$, and for the restricted robust weighting scheme, we assume that a break occur in the range of observations [75:95] following Pesaran et al. (2013).

3.5 Simulation Results

3.5.1 Single break

This subsection summarizes the simulation results of different forecast methods for the DGPs with a single break. Table 3.3-Table 3.5 report relative MSFEs of forecast methods to the benchmark when a single break occurs at the 1/4, 1/2, and 3/4 of the full sample, respectively, with a sample of length of 100 observations. Specifically, in these tables, DGPs with a single break refer to the experiments specified in Table 3.1. Each row of the table then shows the results for the selected forecasting methods, which are grouped as single window, multiple window, observational weights and our proposed confidence interval methods, as panel A, B, C and D respectively. The forecast methods in panels A, B and C are described in subsections 3.2.1, 3.2.2 and 3.2.3, respectively. Panel D includes our proposed methods which are introduced under subsection 3.3.1.

Specifically, the first and second rows of panel D in each case show the results for the non-iterated (denoted as 'CI') and the iterated versions (denoted as 'iterated'), respectively, of confidence interval methods which employ

equation (3.4). More specifically, when detecting the break date and obtaining its confidence interval, the homoskedastic Bai and Perron (1998) procedure is used for the non-iterated version, while an iterative coefficient/variance testing methodology (described in subsection 3.3.3) is employed for the iterated version of the confidence interval method. Also, cross validation weights are employed within the iterated confidence interval as in equation (3.5) and the results are reported in the third row of panel D (denoted as 'CV weight'). The EM confidence interval in panel D refers to forecasts that are generated using the confidence interval proposed by Elliott and Muller (2007), as outlined in subsection 3.3.4.2. Finally, the iterated confidence interval is used to define the bounds for the restricted robust weighting scheme as in equation (3.6) and the resulting forecasts are presented in the last row.

This grouping facilitates us to find reliable procedures not only among all of the forecast methods, but also among each group. First, we will discuss the overall performances of methods across groups, and later we turn to the detailed discussion for each group of methods.

As one would expect, in experiment 1 in each table where no break applies, the best forecasting method is full sample OLS (benchmark). This is due to the efficiency of using all available data when a break is absent. Consistent with this, forecast methods which use longer estimation windows such as exponential smoothing with $\gamma = 0.98$ work well with less than 1% MSFE accuracy loss. Worst is the restricted robust weighting method which assumes a break occurring between the 75th and 95th observations and cuts off all other observations, leading to a short estimation window. However, in the presence of any size of coefficient break (Exp 1-6), all forecast methods yield smaller MSFE than the benchmark, especially the reduction in a relative MSFE is substantial when the break is large. Furthermore, the largest gain over the benchmark for all forecast methods is observed in experiment 10 which is associated with a break that changes coefficients and leads to a decreased variance. In experiment 9 where the same coefficient break occurs but the variance is increasing, most forecast methods perform worse than the benchmark. Similar patterns are seen for experiments 7 and 8 where the break affects only the variance of the series.

In general, across all DGPs with a coefficient break excluding experiment 9, our confidence interval methods consistently generates good forecasts along with the single window methods (post break, trade-off and cross validation). The best performance in terms of the smallest MSFE interchangeably appear either in single window methods or in confidence interval methods, although the differences are small in magnitude. This applies regardless of the location

of break, that is across Tables 3.3-3.5. Essentially, these two groups of forecast methods condition on break date information and their corresponding MSFEs are much smaller than the MSFEs of other methods when the break is large (see for example, Exp 3 and 5). However, the differences are small when the break is small in magnitude (see for example Exp 2 and 4). This is related to the increased accuracy of the break point estimation procedure with a large or moderate size of break.

Moreover, confidence interval methods (except the confidence interval using Elliott and Muller (2007)) outperform forecast combination (averaging multiple windows) methods and most of observational weighting methods in coefficient break experiments regardless of the locations of breaks, with the exception of experiment 9.

For experiments related to the changing variances (Exp 7, 8, 9 but not 10), the most accurate performances are cross validated window average and exponential smoothing with $\gamma = 0.98$ when the break occurs in the early or middle part of the sample. The iterated confidence interval method combined with cross validation weights also yields relatively good accuracy. When the break occurs in the later part of the sample (see Table 3.5), the pooled window averaging methods in Panel B tend to yield good forecasts in addition to the previously accurate methods mentioned above for the variance break cases. Although experiment 10 associates with a variance break case, forecast methods behave similar to the coefficient break experiments.

Overall, across Table 3.3 to 3.5, the gain from employing a confidence interval over using an estimated post break window is marginal except in experiment 4 and the increasing variance experiments of 7 and 9. More important benefits come from employing a confidence interval based on the iterated coefficient/variance break testing procedure, which substantially reduces MSFEs in all variance break cases compared to both the post break window and confidence interval methods that assume homoskedastic errors. Moreover, there is no loss associated with the iterative confidence interval in the absence of a changing variance, hence it is a useful method when forecasting different types of DGPs.

Finally, as discussed in subsection 3.3.4, an alternative confidence interval to that of Bai and Perron (1998) and proposed by Elliott and Muller (2007) is implemented and this is referred as the EM confidence interval in the tables. We note that the EM interval does not outperform Bai and Perron's confidence interval for our purpose except in the case of an increased post break variance. However, even in this case, once we employ iteration, the forecasts from the EM confidence interval are dominated by the forecasts from Bai and Perron's

confidence interval. In the following, we discuss forecast methods within each group.

Single window: The post break method is the most accurate among single window forecasts in panel A for most coefficient break experiments (see Exp 2, 3, 5, 6, 10 in Tables 3.3-3.5). However, the trade off and the cross validation methods are not far off, especially when the break is small. The explanation for this is simple. If the break is large, it is estimated with a high precision and the full efficiency of the post break method can be exploited. On the other hand, small breaks are hard to detect and may result in either efficiency loss due to late or bias due to early detection of the break. The trade-off and cross validation approaches allow pre-break observations to be used which introduces potential inefficiency and bias. For experiments 4 and 9, the cross validation method performs with the highest accuracy of this group. In the latter case of increasing post break variance, the cross validation method may be suited well by introducing pre-break observations and choosing the forecast which achieves the smallest MSFE. To support this, experiment 7 with increasing variance DGP also favours the cross validation method regardless of the location of the break (see Exp 7 and 9 in Tables 3.3-3.5).

However, when DGPs have a single break only in the variance (Exp 7 and 8), the post break and trade-off methods perform poorly, even worse than the benchmark. This is due to the poor performance of the homoskedastic Bai and Perron (1998) testing procedure in the presence of a variance break. The deterioration is even higher when the location of a variance break is later in the sample. We observe that the testing procedure tends to erroneously identify a coefficient break around the timing of the volatility break. This falsely cuts off available observations, leaving a relatively small estimation sample. Cross validation works better in this case because it does not directly use an estimate of the break date.

Multiple windows: The pooled scheme which does not use any break date information consistently performs well when the break is in the coefficients (with exception of Exp 9). In particular, the relative MSFEs compared to the benchmark are smaller when the break is large. This is because forecast errors associated with the full sample estimation are large in the presence of a large break. Additionally, using an estimate of the break date does not improve upon the equal weighting scheme as it is not always precisely estimated. In the presence of a variance break only, the cross validated weighting scheme performs well, implying that weighting forecasts by their corresponding MSFEs helps to reduce overall forecast errors by assigning smaller weights to those with high MSFEs. These patterns do not vary with different break locations.

Observational weights: Among this group, the most accurate method in the coefficient break experiments overall is the optimal weighting scheme which is based on the location, magnitude of the coefficient break and changes in the standard deviations. Although optimal weights can be prone to an imprecise break date estimate, exploiting this information adds value compared to other observational weighting schemes. Exponential smoothing with $\gamma = 0.95$ also does well, often having a MSFE closer to that of the optimal weights than other methods. However, this deteriorates when $\gamma = 0.98$. The performance of the robust weighting scheme is poor when the break is early in the sample, because the true break date is out of the pre-assumed range $[75 : 95]$ and important observations are given zero weights in the estimation. The performance, however, improves when the true break date falls in the range such as when the break is at 75th observation. In this case, the robust weighting scheme generally yields the smallest MSFE in panel C. The optimal weights continue to perform well in most coefficient break experiments regardless of the different break locations. However, it should be noted that experiment 9 is an exception to these observed patterns as it behaves similarly to the DGPs with only changing post break variance.

The ranking of most accurate methods changes substantially in the variance break cases. The most accurate method both for an increase and decrease of the post break variance is exponential smoothing with $\gamma = 0.98$. This is reasonable, since this implies a longer estimation window and should work well with no coefficient break. Now the robust weighting scheme is the worst performer even when the break is late in the sample (see Table 3.5) – approximately 30% loss in accuracy in increasing and 25% loss in decreasing post variance experiments, compared to the benchmark. This is related to the large data loss in the absence of a coefficient break.

Confidence interval methods: The non-iterated confidence interval method with equal weights performs well in all DGPs except those with a variance break (Exp 7 to 9 in Tables 3.3-3.5). This is not surprising as the confidence interval is associated with a coefficient break and any variance break is ignored. To account for a variance break, we also employ our iterated procedure when estimating the coefficient break date and confidence interval. The simulation results show forecast improvements with the iterated procedure when there is a variance break. As expected, in the DGPs with only coefficient breaks, there is no difference between iteration and non-iteration. This implies no accuracy loss is associated with the iteration procedure for forecasting.

Furthermore, we employ cross validated weights in the iterated confidence

interval which yields further improvement in the small break experiments 2 and 4 when the break occurs either early or in the middle of the sample (see Tables 3.3 and 3.4). However, its performance deteriorates when the break is late in the sample (see Table 3.5), which is related to the small sample available after reserving the required data for the evaluation period. Cross validated weights in the iterated confidence intervals also improves forecast accuracy of the iterated confidence interval method for experiments 7, 8 and 9 that include a variance break, regardless of the different break locations. Therefore, in panel D, this method is the most accurate across all DGPs when the break occurs in the early or middle part of the sample.

The usefulness of using a confidence interval to give the bounds for the restricted robust weighting scheme is found to be non-negligible. The improvement over the restricted robust weighting scheme of Pesaran et al. (2013) is evident for all DGPs and all break locations (see Exp 1-10 in Tables 3.3-3.5). Differences in terms of their MSFEs are large when the true break is out of the pre-assumed interval for the robust weight. The explanation for this is clear – using the estimated confidence interval increases the probability of including the true break date in the range when employing a robust weighting scheme. Even when the true break date falls in the assumed interval, replacing it by the estimated interval improves the forecast performance as can be seen in Table 3.5. It is among the best methods in coefficient break cases, having approximately equal MSFEs with the simple confidence interval forecasts.

Large sample performance: We repeat all simulations with a larger sample of 200 observations. The results are reported in appendix Table C.1-Table C.3. The observed patterns and rankings of the forecast methods remain largely unchanged from those discussed for T=100. However, the relative MSFEs for all methods are reduced compared to the simulations with 100 observations. Probably, this is due to an increased accuracy of the estimated break date with a larger sample as well as the availability of more post break observations for parameter estimation. For instance, in experiment 7 with 100 observations, our iteration largely improves on the non-iterated confidence interval, even though iteration still yields some loss over the benchmark. In particular, with 200 sample observations, the loss decreases from 4% to 1% when the break is early, 6% to 2% when the break is in the middle or later in the sample. We also simulated experiments 7 and 8 with 500 observations to see whether the efficiency loss in our iterated confidence interval over the benchmark disappears. The results are presented in appendix Table C.4. Indeed, we find an improvement in the MSFE sense – having an accuracy loss less than 1% in experiment 7 and accuracy gains of 3%-7% in experiment 8

with different break dates.

3.5.2 Multiple breaks

As discussed in subsection 3.2.5, to date methods use information only on the most recent break date. In addition to evaluating methods based on that approach, Table 3.6 also employs our re-ordering procedure described in subsection 3.3.2. Therefore, for each method, two sets of results are shown, the first of which does not apply re-ordering whereas the second does. A sample size of $T = 100$ observations is employed.

Despite similar conclusions often applying when employing original and re-ordered data, the resulting MSFEs are distinct in the parameter reversion cases of experiments 11, 13, 15. To be more specific, the procedure of re-ordering data segments reduces MSFEs, often substantially, for all forecast methods in the coefficient reversion cases. The improvements are reported in Table 3.7, in terms of the MSFE differences between the two approaches for the corresponding experiments. The improvements are larger in experiments 11 and 15, where persistence is high (around 0.9) compared to experiment 13 in which autoregressive coefficients are relatively small (0.3 to 0.5 and back to 0.3). This is reasonable, because the gain from using a longer data sample for estimation will be greater for a more persistent process.

For the trended coefficient and variance break cases, the differences using original and re-ordered data segments are negligible. The implication is that the re-ordering procedure does not change the ordering of the original data when there is no coefficient reversion. Therefore, it assures no loss by using the re-ordering procedure for all DGPs when forecasting a time series that is subject to multiple coefficient breaks.

When reversion occurs in the coefficients, exponential smoothing with $\gamma = 0.95$, optimal weights, unrestricted robust weights and pooled average methods tend to generate good forecasts compared to others. Their good performances relate to the long data coverage that includes observations prior to the first break which are informative with regard to the forecast value. When the coefficients are either increasing or decreasing after each break (Exp 12, 16), the most accurate methods are the single window forecasts and our confidence interval forecast methods. This is the same conclusion that arises from the experiments with a single break occurring in the set of coefficients. The implication is that using the most recent break information is appropriate when the parameters are changing in the same direction after each break. Although experiment 14 is a DGP with increasing coefficients, somewhat similar findings

to coefficient reversion cases appear in terms of the favoured forecast methods. We also note that our confidence interval method with cross validated weights generally performs well for both reverting and trending coefficient DGPs compared to other methods.

For DGPs with variance breaks, cross validation weighted window averaging, exponential smoothing with $\gamma = 0.98$ and confidence interval with cross validated weights methods are the most accurate forecasters which are similar to the findings for single break cases.

Table 3.6 also includes the forecasting results using the optimal weights under multiple break information proposed by Pesaran et al. (2013). This method is an extension of their optimal weighting scheme in a single break, when multiple breaks are assumed and we described the case with two breaks in Appendix C.1. We find that it does not perform well compared to our re-ordering method or the optimal weighting scheme for a single break where the estimate of the last break point is used.

When the sample size increases from 100 to 200, the MSFE ratios for all methods decrease substantially. The results are presented in appendix Table C.5. The general conclusions based on a sample size of 100 largely carry over. Due to the increased sample, the structural break testing method yields more accurate estimates of break dates. Consequently, the performances of single window forecasts and confidence interval methods which use estimates of break dates improve when using the original data for the coefficient reverting experiments of 11 and 15. Finally, Table C.6 in the appendix shows differences between the forecasts using original and re-ordered data in the MSFE sense for the parameter reversion experiments with $T = 200$. Compared with Table 3.7, we observe that the gain from employing re-ordered data is even larger for the majority of cases when the sample size increases.

3.6 Application to G7 inflation

3.6.1 The empirical model

In order to find out how well our proposed methods work in practice, we forecast aggregate inflation series of the G7 countries, namely Canada, France, Germany, Italy, Japan, UK and US. The exercise uses one step ahead pseudo out of sample forecasts which are based on simple autoregressive models (AR) with a fixed lag order of 4 as,

$$\hat{\pi}_{t+1|t} = \hat{\beta}_0 + \sum_{i=1}^{4} \hat{\beta}_i \pi_t + \varepsilon_t \tag{3.12}$$

where $\hat{\beta}_0$ and $\hat{\beta}_i$ are the OLS estimates using a rolling window of 400 observations up to period t, to generate a forecast for period $t+1$. All inflation data are monthly and cover the period between February 1970 and September 2010. Having 4 lags in the forecast model, the remaining sample is divided into an initial estimation sample (June 1970–September 2003) in which 400 observations are available and a forecast sample (October 2003–September 2010) with a length of 84 observations. For instance, the initial estimation sample is used to produce a one-step ahead forecast for October 2003, and then the estimation window moves forward by one month (but excludes the oldest one observation from the estimation window) and a forecast is produced for November 2003, and so on.

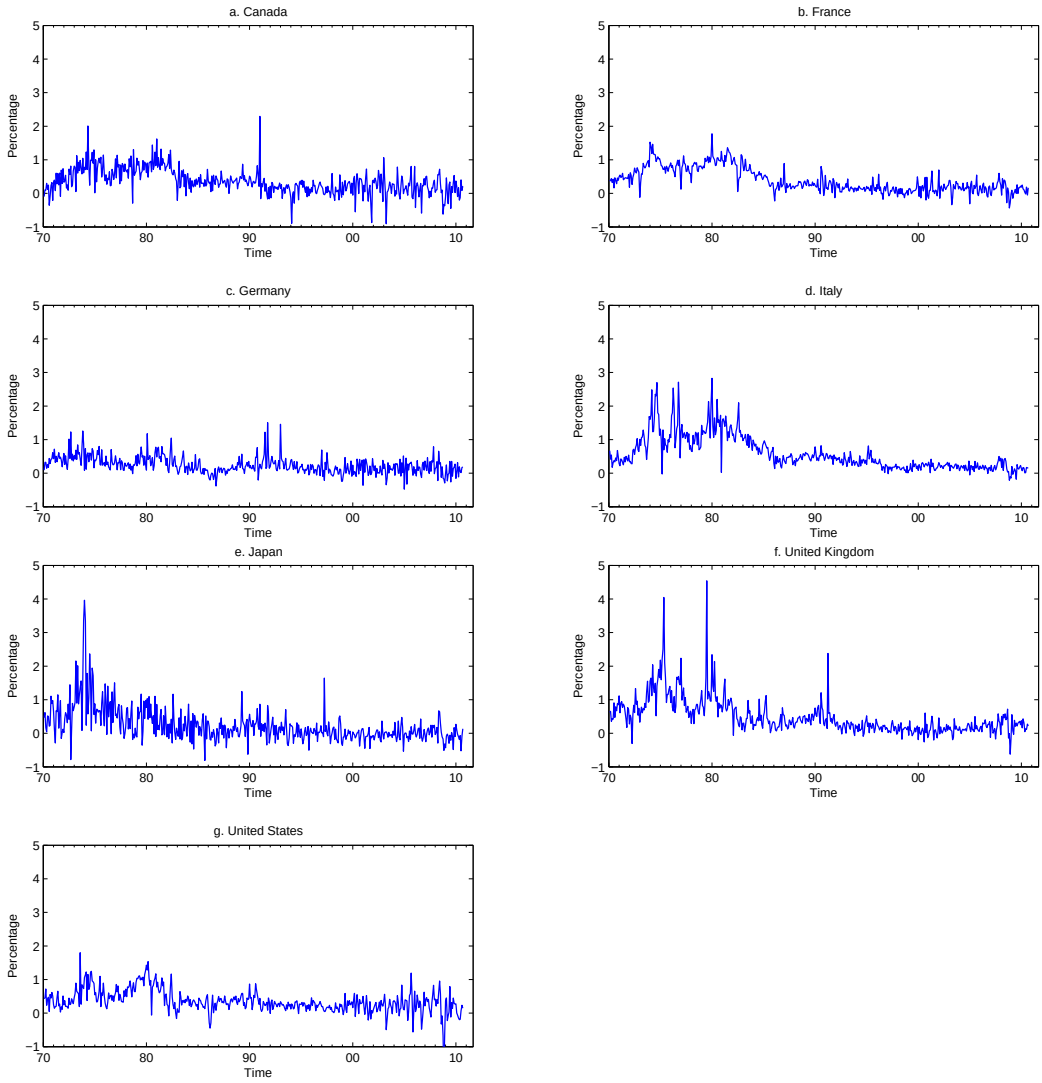
For each rolling window, multiple structural break testing procedure runs to assess the existence of breaks which, if any detected, are taken into account to generate one step ahead forecast. We acknowledge that the repeated application of multiple break testing procedure to each rolling window may raise an issue that the probability of rejecting a true null hypothesis of no change approaches one as the number of application grows. This is discussed in Robbins (1970), Chu et al. (1996) in the context of expanding estimation window. However, in our application of repeated tests, the same conclusions are generally found as to the existence and non-existence of breaks with the estimates of breaks remaining at the same temporal locations.

The inflation series are constructed by differencing the logged monthly CPI indices[6] and multiplying by 100. After computing the inflation series, we apply X-12-ARIMA seasonal adjustment procedure to account for any seasonal effects in each inflation series and forecast the seasonally adjusted values.

The current study is not focused on the presence of breaks in the forecast horizon and it is consequently desirable to rule out this possibility when evaluating forecast methods. For an initial inspection, figure 3.4 plots the inflation series for the full sample, which does not suggest any obvious breaks in the forecast horizon except a slight increase in the volatility of Canada and US inflation. Further, we test for breaks in the inflation series of G7 countries using the iterative break point testing procedure outlined in subsection 3.3.3 on the full sample. The results are reported in Table C.7 in the appendix and they are reassuring as no break is uncovered in the forecast horizon either in

[6]We take the underlying CPI index values from the OECD Main Economic Indicator Database.

Figure 3.4: G7 Inflation Dynamics



the coefficients or variance of inflation for these countries.

In parameter estimation for the purpose of forecasting, we test for multiple breaks using the Bai and Perron (1998) procedure allowing a maximum of 3[7] breaks with trimming of 10% and set the significance level to 5%. The number of breaks is estimated as described in section 3.4 and all forecast methods considered in the simulation experiments are re-evaluated on each rolling data window. If no break is detected, the full sample ($T = 400$) estimation is used for the coefficients of the forecast model. In the presence of any number of breaks, forecast methods are applied based on the most recent break both on the original (as in subsection 3.3.1) and re-ordered data (as in subsection 3.3.2). Additionally, in the application we set a minimum estimation sample $\underline{w}$ equals to 10% of the rolling sample. In cross validation based methods,

[7] As indicated by the results of chapter 1 for the study of 19 OECD countries, allowing a maximum of 3 breaks is appropriate.

we allow the out of sample evaluation period $\tilde{w}$ equals to 20% of the rolling estimation sample. For the exponential smoothing method, we set $\gamma = 0.95$ and $\gamma = 0.98$.

For each forecast method, we again report the out of sample MSFE relative to the full sample benchmark as our measure of forecast accuracy. We also employ the Diebold and Mariano (2002)[8] test of equal predictive accuracy for two competing methods. Specifically, this asymptotic test is based on the loss differentials series as given by the difference between MSFE values. To save space, only a few of the best performing forecasts based on their relative MSFEs are selected and tested against each other. The null hypothesis is equal accuracy of the forecast methods compared. Rejection of the null hypothesis indicates that lower MSFE is delivered by the alternative method (2nd method) if the test value is positive. Conversely, if the test value is negative, the test favours the null forecast method (1st method) as it implies smaller MSFE compared to the alternative method.

3.6.2 Empirical Results

Table 3.8 reports the estimated break dates and the corresponding changes in mean, persistence and standard deviation of disturbance of each country's model as obtained from the iterative procedure outlined in subsection 3.3.3, using the initial estimation period of 400 observations. As the estimation sample moves forward by 1 observation at a time, in general, the estimates of breaks remain at these temporal locations. Conditional on coefficient breaks, in each regime the mean is calculated as the sample average and persistence is computed as the sum of autoregressive coefficients. Changes in absolute values of residuals are also reported based on breaks in the mean of absolute errors. These measures are indicative with regards to the magnitude and nature of breaks, and helps us understand why some methods perform well for some countries and not so well for other countries.

To measure forecast accuracy, out of sample MSFEs relative to the full sample benchmark are reported in Table 3.9 where same grouping of methods is used as in Tables 3.3. MSFEs in this table are computed based on the most recent break using data that has not been re-ordered. We also report MSFEs calculated using re-ordered data in Table C.8 in the appendix. Since there are no coefficient reversion cases for these series, the values in Table 3.9 and Table

[8]Clark and West (2007) propose a test for equal predictive accuracy of forecasts which is appropriate for nested models. However, the test compares a parsimonious null model to a larger model that collapses to the null model when certain parameters are set to zero, and this set up is not applicable to our forecast method comparisons.

C.8 are almost identical. Hence, the following analysis is focused on Table 3.9.

In general, iterated confidence interval methods consistently perform well, ranking amongst the most accurate forecast methods for all countries with the exception of France. Depending on the nature of breaks in different countries, some other methods also produce good forecasts. Brief summaries of the results in Table 3.8 and Table 3.9 for each country are as follows.

Canada: A single break in the coefficients leads to substantial declines both in mean and persistence, as shown in Table 3.8. It is anticipated that this large break will be estimated with good precision and thus forecast methods which use the location of a break are expected to work well. Consistent with this, single estimation window forecasts and confidence interval methods which directly hinge on the break date estimate perform well compared to the methods in panels B and C. Although the mean of the residual absolute values increases slightly after the second variance break, the less volatile sample is included in the relevant coefficient estimation window which reduces variances of estimated coefficients in the forecasting model.

France: Averaging forecasts from multiple windows, such as equally weighted average with break date information and cross validation weighted average, yield the most accurate forecasts for France inflation. According to the simulation results, these methods work well when the coefficient break is small and/or for DGPs with changing variances which are the case for France. Indeed, the most recent break is relevant in forecasting and the second break estimated in the coefficients around June 1992 results in small reductions in the mean and persistence compared to the period before first break around July 1985. Additionally, we document that France inflation has undergone two variance breaks and the second break leads to an increase in variance.

Germany: A break in July 1982 seems to affect only the mean of German inflation which is reduced to less than half its previous value. Considering the initial relatively low level of inflation, the size of the break is not large. In this respect, methods that do not use an estimate of the break date may yield higher accuracy of forecasting. Indeed, equally weighted window averaging, which does not use the estimate of break date, unrestricted robust weight and exponential smoothing with $\gamma = 0.98$ methods are the most accurate forecasters for German inflation. However, unlike methods based on an estimate of the break date, our iterated version of confidence interval methods perform almost as well as the most accurate methods just mentioned. It is a little surprising that the iterated testing methodology works better than the non-iterated homoskedastic version of Bai and Perron (1998) in the absence of variance breaks. Presumably this implies that a variance break is found in some other

estimation windows (beyond the initial one for which results are presented in Table 3.8), and using this information helps accuracy. Nevertheless the lower MSFEs yielded by use of the iterated confidence interval is evident compared to their non-iteration based alternatives.

Italy, UK: There are similarities between UK and Italy inflation series in terms of changing patterns in the coefficients and variances. Each break in the coefficients (only one break for Italy) results in substantial declines both in mean and persistence. For both countries, the residuals variances increase after the first variance break but decline to even lower levels after the second break. Under these changes, the iterated confidence interval method, with both equally and cross validation weighted, and the robust weighting method combined with confidence interval bounds yield the smallest MSFEs among all forecast methods. This result is consistent with the simulation results where we find the accuracy of confidence interval methods increase when the coefficient break is large and volatility declines towards the end of the sample. With large breaks, single window forecasts are also expected to work well. However, estimation of the coefficient break date with the homoskedastic version of the Bai and Perron (1998) procedure may not be accurate when the variance is changing. Perhaps, for this reason, single window methods and non-iterated confidence interval methods do not perform as well as the iterated confidence interval method for Italy and UK inflation series.

Japan: The mean level of inflation and residual variances decline after each break for the Japan inflation series. However, the relatively high negative persistence after the second coefficient break should be noted. In this case, it may be appropriate to assign higher weights to the more recent observations as they are more informative with regards to a forecast value. Consistent with the expectation, optimal weight, unrestricted robust weight and confidence interval methods which put heavier weights on recent past observations yield small MSFEs compared to other methods. Additionally, the persistence (albeit here negative) appears to increase substantially in magnitude after the second coefficient break and models with persistent parameters benefit from a longer coverage in the estimation sample. Perhaps, good performances of optimal weight and unrestricted robust weight methods may be related to not only assigning heavier weights to the recent observations but also employing a longer estimation sample compared to other methods (see figure 3.1 for observational weights).

To support this view, the accuracy of exponential smoothing with $\gamma = 0.98$ is not far from these best performing methods. Furthermore, the late occurrence of a coefficient break around late 1997 shortens the effective estimation

sample. As supported by the simulation analysis and the empirical study, our confidence interval methods work well when the break is close to the end of sample. Besides, as the window moves through time, more post-break observations become available in the estimation sample assuming that there are no further breaks. In this sense, confidence interval methods are expected to work well as they are essentially the average of post break window forecasts and this is the case for Japan.

US: The most accurate methods to forecast US inflation are cross validation, cross validation weighted average, equally weighted average with break date information and iterated confidence interval methods. These methods use an estimate of the break date and judging by their good performance break date information adds value for forecasting US inflation. However, this excludes post break, trade-off and non-iterated confidence interval methods which do not tackle changes in variance, whereas (according to Table 3.8)) the residual variance has changed twice, the first change leads a decline and the second leads an increase. Although cross validation based approaches and equal weighted average methods do not directly handle variance breaks, the simulation analysis reveals these work well in DGPs with changing variance. This is shown again in the empirical analysis and they perform as well as iterated confidence interval methods which are designed to account for changing variances.

Table 3.10 provides Diebold and Mariano (2002) statistics and statistical significance at 5% and 10% levels are indicated by double and single stars respectively. Each panel corresponds to tests for a chosen null model against other forecast methods for the G7 countries. For example, in panel A, the full sample method is tested against nine other forecast methods (selected as representative of the various approaches) and the statistics shows that cross validation, pooled average and iterated confidence interval methods are significantly better than the full sample benchmark method for most G7 countries' inflation. The iterated confidence interval method appears to be significantly better than other methods that do or do not recognize breaks for only a small number of countries. In other panels, the majority of tests statistics are not statistically significant and this may due to a relatively short out of sample period as the asymptotic distribution of the test statistics rely on a large sample. To this extent, the relative performance of these methods in Table 3.9 should be treated with some caution.

3.7 Conclusion

This chapter considers the problem of forecasting in the presence of a single or multiple structural breaks in the data generating process. We take an approach that attempts to identify the locations of breaks and then uses this information in the context of forecast combination methods. However, we do not rely on a single estimate of the break date as it is often difficult to estimate accurately. Instead, we examine the usefulness of employing a confidence interval of the estimated break date in improving forecastability of time series.

Monte Carlo and empirical studies undertaken here show an overall good performance for the confidence interval methods. In particular, confidence interval methods can outperform forecast combination methods in coefficient break experiments regardless of the locations of breaks, but result in a higher MSFE when a variance break presents in the DGP. However, the confidence interval obtained using the iterative procedure takes account of breaks both in the set of coefficients and residual variances. Employing the iterative confidence interval, especially when it is incorporated with cross validation weights, reduces forecast errors such that MSFE is lower than that obtained using other forecast combination approaches in most experiments.

However, we should note that the improvement over single window methods is marginal in the coefficient break experiments. Nevertheless, the iteration based confidence interval improves other methods that hinge on a point estimate of break dates when variance breaks are present in the DGP. Overall, our proposed methods show no loss compared to benchmark and the majority of other competing methods for coefficient break only cases.

We also propose a method to improve forecastability when there are multiple breaks which lead to reversion in the coefficients. This involves re-ordering data segments associated with estimated coefficient breaks based on the relative closeness of the estimated parameters. Monte Carlo simulations show that the procedure of re-ordering data segments substantially reduces MSFEs of all forecast methods in the parameter reversion cases. Further, the gain from employing such a procedure is even larger when the sample size increases, suggesting this is related to the increased accuracy of the estimated break dates and corresponding coefficients with a larger sample size.

Tables

Table 3.1: Simulation setup: single break

	Parameters	β_{11}		β_{12}		σ_{yt}	
Exp	**Comments/Regimes**	**R1**	**R2**	**R1**	**R2**	**R1**	**R2**
EXP 1	No Break	0.9	0.9	1	1	1	1
EXP 2	Small break in β_{11}	0.9	0.7	1	1	1	1
EXP 3	Large break in β_{11}	0.9	0.5	1	1	1	1
EXP 4	Small break in β_{12}	0.9	0.9	1	1.5	1	1
EXP 5	Large break in β_{12}	0.9	0.9	1	2	1	1
EXP 6	Break in β_{11} and β_{12}	0.9	0.7	1	2	1	1
EXP 7	Increase in σ_{yt}	0.9	0.9	1	1	1	4
EXP 8	Decrease in σ_{yt}	0.9	0.9	1	1	1	0.5
EXP 9	Break in $\beta_{11,}$ β_{12}+Increase in σ_{yt}	0.9	0.7	1	2	1	4
EXP 10	Break in $\beta_{11,}$ β_{12}+Decrease in σ_{yt}	0.9	0.7	1	2	1	0.5

Note: β_{21}=0; β_{22}=0.9. The values of coefficients before and after the break are given in R1 and R2 columns, respectively.

Table 3.2: Simulation setup: multiple breaks

	Parameters	β_{11}			β_{12}			σ_{yt}		
Exp	**Comments/Regimes**	**R1**	**R2**	**R3**	**R1**	**R2**	**R3**	**R1**	**R2**	**R3**
EXP 11	Coefficient reversion in β_{11}	0.9	0.7	0.9	1	1	1	1	1	1
EXP 12	Decline in β_{11}	0.9	0.7	0.5	1	1	1	1	1	1
EXP 13	Coefficient reversion in β_{11}	0.3	0.5	0.3	1	1	1	1	1	1
EXP 14	Increase in β_{11}	0.3	0.5	0.7	1	1	1	1	1	1
EXP 15	Coefficient reversion in β_{12}	0.9	0.9	0.9	0	1	0	1	1	1
EXP 16	Increase in β_{12}	0.9	0.9	0.9	0	1	2	1	1	1
EXP 17	Increase in σ_{yt}	0.9	0.9	0.9	1	1	1	1	1	4
EXP 18	Decline in σ_{yt}	0.9	0.9	0.9	1	1	1	1	1	0.5

Note: β_{21}=0; β_{22}=0.9. The values of coefficients before and after the breaks are given in R1, R2 and R3 columns.

Table 3.3: MSFE ratios (T=100, $T_1 = 25$)

Sample size: T=100 Break location: T_1=25	No Break	Small break in β_{11}	Large break in β_{11}	Small break in β_{12}	Large break in β_{12}	Break in β_{11} and β_{12}	Increase in σ_{yt}	Decrease in σ_{yt}	EXP6 + EXP 7	EXP6 + EXP 8
	EXP 1	EXP 2	EXP 3	EXP 4	EXP 5	EXP 6	EXP 7	EXP 8	EXP 9	EXP 10
A. Single estimation windows										
Post break	1.027	0.721	0.552	0.917	0.664	0.703	1.112	1.001	1.220	0.380
Trade-off	1.017	0.726	0.577	0.908	0.667	0.704	1.071	0.998	1.135	0.385
Cross Validation (CV)	1.036	0.727	0.562	0.910	0.678	0.717	1.044	1.005	1.013	0.385
B. Averaging multiple windows										
Pooled (without break info)	1.042	0.739	0.580	0.916	0.694	0.723	1.046	1.016	1.019	0.388
Pooled (with break info)	1.036	0.840	0.784	0.925	0.786	0.806	1.038	1.006	1.001	0.606
Cross validated weights	1.012	0.821	0.746	0.916	0.760	0.786	1.020	0.989	0.990	0.514
C. Observational weight										
Optimal weight	1.010	0.753	0.622	0.923	0.693	0.717	1.041	0.998	1.088	0.397
Unrestricted Robust weight	1.034	0.773	0.700	0.912	0.697	0.732	1.038	1.010	1.017	0.410
Restricted Robust weight (75-95)	1.326	0.908	0.696	1.157	0.865	0.887	1.278	1.294	1.255	0.456
Exponential (Y=0.95)	1.056	0.760	0.649	0.927	0.698	0.732	1.060	1.031	1.038	0.387
Exponential (Y=0.98)	1.009	0.850	0.833	0.920	0.758	0.794	1.014	0.990	0.997	0.565
D. Confidence interval methods										
Confidence Interval (CI)	1.021	0.723	0.555	0.910	0.664	0.703	1.093	0.999	1.169	0.378
Confidence Interval (iterated)	1.030	0.723	0.555	0.914	0.664	0.703	1.043	1.017	0.998	0.371
Confidence interval (CV weight)	1.011	0.719	0.555	0.906	0.664	0.703	1.014	0.994	0.988	0.371
EM confidence interval	1.087	0.801	0.688	0.935	0.721	0.775	1.095	1.049	1.049	0.479
Restricted Robust -CI bounds	1.062	0.724	0.554	0.909	0.664	0.703	1.072	1.040	1.004	0.371

Note: The details of forecast methods presented in panel A, B, C and D can be found in subsections 3.2.1, 3.2.2, 3.2.3 and 3.3.1, respectively. DGPs in the first row of the table are from table 3.1. The values in the table are the ratios of MSFEcandidate/MSFEbenchmark. Larger than 1 indicates worse and smaller than 1 indicates better performance than the benchmark model.

Table 3.4: MSFE ratios (T=100, $T_1 = 50$)

Sample size: T=100 Break location: T_1=50	No Break	Small break in β_{11}	Large break in β_{11}	Small break in β_{12}	Large break in β_{12}	Break in β_{11} and β_{12}	Increase in σ_{yt}	Decrease in σ_{yt}	EXP6 + EXP 7	EXP6 + EXP 8
	EXP 1	EXP 2	EXP 3	EXP 4	EXP 5	EXP 6	EXP 7	EXP 8	EXP 9	EXP 10
A. Single estimation windows										
Post break	1.027	0.610	0.455	0.813	0.446	0.509	1.311	1.008	1.165	0.201
Trade-off	1.017	0.613	0.477	0.806	0.450	0.514	1.193	1.002	1.086	0.214
Cross Validation (CV)	1.036	0.616	0.465	0.797	0.456	0.521	1.043	1.004	1.003	0.205
B. Averaging multiple windows										
Pooled (without break info)	1.042	0.663	0.552	0.806	0.525	0.577	1.053	1.000	0.982	0.315
Pooled (with break info)	1.036	0.813	0.786	0.849	0.675	0.729	1.040	0.997	0.968	0.584
Cross validated weights	1.012	0.785	0.735	0.842	0.619	0.684	1.023	0.982	0.965	0.431
C. Observational weight										
Optimal weight	1.010	0.652	0.538	0.820	0.492	0.526	1.107	1.003	1.048	0.215
Unrestricted Robust weight	1.034	0.757	0.743	0.809	0.548	0.629	1.046	0.994	0.988	0.397
Restricted Robust weight (75-95)	1.326	0.746	0.561	0.969	0.561	0.627	1.288	1.265	1.201	0.248
Exponential (Y=0.95)	1.056	0.698	0.665	0.794	0.490	0.560	1.069	1.010	0.997	0.279
Exponential (Y=0.98)	1.009	0.843	0.835	0.852	0.671	0.745	1.019	0.979	0.978	0.596
D. Confidence interval methods										
Confidence Interval (CI)	1.021	0.609	0.456	0.806	0.447	0.510	1.264	1.003	1.119	0.203
Confidence Interval (iterated)	1.030	0.610	0.457	0.808	0.447	0.510	1.057	1.001	1.083	0.202
Confidence interval (CV weight)	1.011	0.607	0.456	0.801	0.447	0.509	1.021	0.987	1.046	0.202
EM confidence interval	1.087	0.706	0.613	0.826	0.549	0.621	1.097	1.038	1.028	0.360
Restricted Robust -CI bounds	1.062	0.610	0.457	0.792	0.447	0.510	1.074	1.025	1.082	0.202

Note: Same as Table 3.3

Table 3.5: MSFE ratios (T=100, $T_1 = 75$)

Sample size: T=100 Break location: T_1=75	No Break	Small break in β_{11}	Large break in β_{11}	Small break in β_{12}	Large break in β_{12}	Break in β_{11} and β_{12}	Increase in σ_{yt}	Decrease in σ_{yt}	EXP6 + EXP 7	EXP6 + EXP 8
	EXP 1	EXP 2	EXP 3	EXP 4	EXP 5	EXP 6	EXP 7	EXP 8	EXP 9	EXP 10
A. Single estimation windows										
Post break	1.027	0.616	0.427	0.847	0.367	0.459	1.548	1.014	1.181	0.157
Trade-off	1.017	0.609	0.453	0.819	0.369	0.462	1.314	1.007	1.084	0.172
Cross Validation (CV)	1.036	0.813	0.781	0.806	0.541	0.670	1.030	1.012	1.005	0.474
B. Averaging multiple windows										
Pooled (without break info)	1.042	0.727	0.671	0.775	0.546	0.640	1.053	1.003	0.969	0.479
Pooled (with break info)	1.036	0.825	0.809	0.810	0.645	0.746	1.042	0.999	0.972	0.636
Cross validated weights	1.012	0.872	0.866	0.855	0.702	0.813	1.019	0.995	0.984	0.715
C. Observational weight										
Optimal weight	1.010	0.648	0.531	0.821	0.415	0.471	1.163	1.006	1.036	0.181
Unrestricted Robust weight	1.034	0.788	0.763	0.777	0.569	0.705	1.050	0.991	0.986	0.568
Restricted Robust weight (75-95)	1.326	0.638	0.467	0.812	0.390	0.485	1.292	1.243	1.126	0.173
Exponential (Y=0.95)	1.056	0.746	0.719	0.737	0.476	0.605	1.073	1.004	0.979	0.425
Exponential (Y=0.98)	1.009	0.866	0.848	0.849	0.723	0.822	1.021	0.979	0.980	0.746
D. Confidence interval methods										
Confidence Interval (CI)	1.021	0.615	0.428	0.832	0.366	0.457	1.465	1.005	1.137	0.159
Confidence Interval (iterated)	1.030	0.615	0.428	0.832	0.366	0.458	1.065	0.999	1.095	0.158
Confidence interval (CV weight)	1.011	0.805	0.768	0.830	0.514	0.629	1.007	0.993	1.033	0.439
EM confidence interval	1.087	0.704	0.626	0.803	0.511	0.605	1.099	1.042	1.011	0.405
Restricted Robust -CI bounds	1.062	0.616	0.432	0.793	0.366	0.458	1.072	1.018	1.092	0.157

Note: Same as Table 3.3

Table 3.6: MSFE ratios under multiple breaks (T=100)

Sample size: T=100 Break locations: T_1=33, T_2=66	Coeff reversion in β_{11}	Decline in β_{11}	Coeff reversion in β_{11}	Increase in β_{11}	Coeff reversion in β_{12}	Increase in β_{12}	Increase in σ_{yt}	Decline in σ_{yt}
	EXP 11	**EXP 12**	**EXP 13**	**EXP 14**	**EXP 15**	**EXP 16**	**EXP 17**	**EXP 18**
A. Single estimation windows								
Post break	1.004	0.576	1.048	0.809	0.847	0.222	1.495	1.011
Post break	0.967	0.576	1.045	0.809	0.814	0.222	1.388	1.011
Trade-off	0.970	0.576	1.026	0.787	0.834	0.222	1.293	1.006
Trade-off	0.934	0.578	1.025	0.787	0.794	0.223	1.227	1.005
Cross Validation (CV)	0.950	0.564	1.005	0.752	0.840	0.232	1.034	1.012
Cross Validation (CV)	0.900	0.568	1.005	0.758	0.773	0.228	1.047	1.012
B. Averaging multiple windows								
Pooled (without break info)	0.951	0.573	0.985	0.753	0.893	0.331	1.055	0.998
Pooled (without break info)	0.896	0.578	0.982	0.756	0.766	0.334	1.060	0.998
Pooled (with break info)	0.991	0.721	0.991	0.810	1.010	0.440	1.043	0.995
Pooled (with break info)	0.912	0.728	0.987	0.813	0.804	0.445	1.041	0.996
Cross validated weights	0.998	0.698	1.003	0.809	1.006	0.371	1.022	0.987
Cross validated weights	0.915	0.703	0.997	0.811	0.794	0.370	1.024	0.987
C. Observational weight								
Optimal weight	0.945	0.634	1.018	0.785	0.800	0.252	1.153	1.004
Optimal weight	0.928	0.632	1.019	0.789	0.781	0.251	1.153	1.005
Optimal weight (multiple break)	0.933	0.722	1.013	0.788	0.843	0.588	1.079	1.005
Unrestricted Robust weight	0.939	0.716	1.000	0.751	0.884	0.369	1.049	0.989
Unrestricted Robust weight	0.887	0.721	0.996	0.754	0.767	0.371	1.055	0.990
Restricted Robust weight (75-95)	1.074	0.606	1.166	0.848	0.877	0.249	1.290	1.252
Restricted Robust weight (75-95)	1.072	0.607	1.166	0.849	0.880	0.250	1.291	1.252
Exponential (Y=0.95)	0.919	0.645	0.997	0.733	0.832	0.279	1.073	1.003
Exponential (Y=0.95)	0.889	0.651	0.994	0.735	0.763	0.281	1.080	1.003
Exponential (Y=0.98)	0.950	0.826	0.991	0.806	0.920	0.566	1.021	0.977
Exponential (Y=0.98)	0.898	0.828	0.986	0.808	0.802	0.568	1.025	0.977
D. Confidence interval methods								
Confidence Interval (CI)	0.983	0.574	1.033	0.789	0.843	0.221	1.433	1.006
Confidence Interval (CI)	0.950	0.576	1.033	0.793	0.810	0.221	1.360	1.006
Confidence Interval (iterated)	0.981	0.573	1.034	0.791	0.846	0.221	1.084	1.008
Confidence Interval (iterated)	0.949	0.575	1.033	0.794	0.810	0.221	1.085	1.009
Confidence interval (CV weight)	0.941	0.556	1.007	0.755	0.826	0.224	1.026	0.993
Confidence interval (CV weight)	0.919	0.561	1.007	0.760	0.787	0.222	1.027	0.993
Restricted Robust -CI bounds	0.983	0.574	1.040	0.784	0.849	0.222	1.095	1.032
Restricted Robust -CI bounds	0.945	0.580	1.031	0.791	0.806	0.222	1.084	1.006

Note: The details of forecast methods presented in panel A, B, C and D can be found in subsections 3.2.1, 3.2.2, 3.2.3 and 3.3.1, respectively. DGPs in the first row of the table are from table 3.2. For each method, the first row reports the relative MSFEs obtained using the original sample, while the second row presents the relative MSFEs from using the re-ordering procedure described in subsection 3.3.2. The third row for the optimal weights is calculated using the weights derived under multiple break information, proposed by Pesaran et al (2013).

Table 3.7: Changes in MSFE ratios in the coefficient reverting experiments

Sample size: T=100 **Break locations: T_1=33, T_2=66**	**Coeff reversion in β_{11}**	**Coeff reversion in β_{11}**	**Coeff reversion in β_{12}**
	EXP 11	**EXP 13**	**EXP 15**
Post break	-0.037	-0.002	-0.033
Trade-off	-0.036	-0.002	-0.040
Cross Validation (CV)	-0.050	0.000	-0.067
Pooled (without break info)	-0.054	-0.003	-0.127
Pooled (with break info)	-0.079	-0.004	-0.206
Cross validated weights	-0.083	-0.006	-0.211
Optimal weight	-0.016	0.001	-0.019
Unrestricted Robust weight	-0.052	-0.004	-0.117
Restricted Robust weight (75-95)	-0.002	0.000	0.003
Exponential (ϒ=0.95)	-0.030	-0.003	-0.069
Exponential (ϒ=0.98)	-0.052	-0.004	-0.119
Confidence Interval (CI)	-0.033	-0.001	-0.033
Confidence Interval (iterated)	-0.032	-0.001	-0.036
Confidence interval (CV weight)	-0.022	0.000	-0.040
Restricted Robust -CI bounds	-0.038	-0.009	-0.043

Note: The values in the table are differences between MSFEs of the first and second rows for each forecast method for experiments 11, 13, 15 in Table 3.6. Negative values indicate gains from employing re-ordering data method in MSFE sense compared to using the original data. Similarly, positive values indicate the loss associated with a re-ordering data method.

Table 3.8: Structural breaks and corresponding parameter changes in inflation models of G7 countries

Countries	**Structural Breaks**		**Mean**			**Persistence**			**Residual Variance**		
	Coefficients	**Variance**	**R1**	**R2**	**R3**	**R1**	**R2**	**R3**	**R1**	**R2**	**R3**
Canada	Feb-1991	Dec-1983 Feb-2000	0.555	0.170		0.772	0.048		0.309	0.205	0.408
France	Jul-1985 Jun-1992	Jun-1983 Dec-1999	0.770	0.257	0.132	0.813	-0.051	0.080	0.176	0.125	0.226
Germany	Jul-1982	NA	0.420	0.173		0.357	0.460		0.209		
Italy	Jul-1995	Nov-1973 Dec-1982	0.821	0.221		0.868	0.496		0.231	0.456	0.118
Japan	Dec-1980 Nov-1997	Dec-1974 Feb-1986	0.713	0.141	-0.037	0.483	0.014	-0.636	0.848	0.443	0.250
United Kingdom	Jan-1982 Dec-1991	Oct-1973 May-1980	1.048	0.448	0.154	0.542	0.238	0.131	0.330	0.595	0.201
United States	Oct-1981 Feb-1991	May-1983 Feb-2000	0.636	0.348	0.213	0.754	0.382	0.117	0.271	0.138	0.225

Notes: Reported break dates are estimated using an estimation window which includes first 400 observations. Prefix R indicates a regime. In each regime, the mean is calculated as the sample average of aggregate inflation and the persistence is computed as a sum of estimated autoregressive coefficients conditional on the estimated coefficient breaks. The residual variance refers to the mean of absolute values of residuals in (3.8) in each variance break regime. These errors are used in the GLS transformation. NA means not applicable, since no break is detected.

Table 3.9: MSFEs for G7 inflation (based on the last break)

Countries	Can	Fra	Ger	Ita	Jap	UK	US
A. Single estimation windows							
Post break	0.829	1.032	0.965	0.881	0.913	0.938	0.983
Trade-off	0.838	1.003	0.947	0.882	0.903	0.939	0.951
Cross Validation (CV)	0.829	0.954	0.951	0.878	0.939	0.962	0.888
B. Averaging multiple windows							
Pooled (without break info)	0.863	0.920	0.910	0.862	0.891	0.951	0.914
Pooled (with break info)	0.909	0.907	0.954	0.908	0.924	0.956	0.904
Cross validated weights	0.912	0.909	0.956	0.912	0.925	0.957	0.903
C. Observational weight							
Optimal weight	0.846	0.933	0.943	0.867	0.846	0.958	0.939
Unrestricted Robust weight	0.894	0.960	0.925	0.893	0.892	1.010	0.980
Restricted Robust weight (75-95)	0.859	1.016	0.936	0.918	0.957	1.131	1.001
Exponential (ϒ=0.95)	0.869	1.057	1.005	0.979	1.015	1.332	1.172
Exponential (ϒ=0.98)	0.849	0.957	0.912	0.878	0.900	1.059	0.996
D. Confidence interval methods							
Confidence Interval (CI)	0.828	1.021	0.962	0.879	0.903	0.948	0.966
Confidence Interval (iterated)	0.837	1.011	0.917	0.830	0.895	0.927	0.908
Confidence interval (CV weight)	0.838	0.986	0.918	0.844	0.884	0.927	0.908
Restricted Robust -CI bounds	0.842	1.010	0.917	0.829	0.890	0.921	0.906

Table 3.10: Diebold and Mariano test statistics

Methods	Canada	France	Germany	Italy	Japan	UK	US
A. Full sample vs							
Post break	2.00**	-0.23	0.39	1.78**	0.98	0.62	0.11
Trade off	2.00**	0.39	0.60	1.92**	0.85	0.48	1.01
Cross Validation	2.68**	1.81**	1.86**	2.78**	1.64**	1.23	1.49*
Pooled Average	2.67**	1.72**	1.84**	2.78**	1.66**	1.23	1.37*
CV weighted average	2.20**	-0.15	0.44	2.02**	1.18	0.56	0.25
Optimal Weight	1.68**	-1.75**	0.51	-0.22	1.00	-1.27	0.24
Exp smoothing (Υ=0.98)	1.33*	-0.43	-0.04	0.17	-0.10	-1.96**	-0.82
Confidence interval (iterated)	2.25**	0.92	1.73**	2.02**	1.54*	0.74	0.91
Restricted Robust-CI bounds	2.15**	0.11	0.83	1.52*	1.26	0.69	0.83
B. Post break vs							
Trade off	-0.05	2.01**	1.11	0.14	-0.72	-0.81	1.43*
Cross Validation	-1.51*	1.28*	0.14	-0.70	-0.23	-0.26	0.85
Pooled Average	-1.48*	1.36*	0.16	-0.61	-0.20	-0.25	0.89
CV weighted average	0.09	1.83**	0.67	0.23	1.22	-0.69	0.97
Optimal Weight	-0.69	-0.26	0.39	-1.85**	0.37	-1.46*	0.10
Exp smoothing (Υ=0.98)	-0.72	-0.25	-0.45	-0.97	-0.96	-2.24**	-1.99**
Confidence interval (iterated)	-1.12	1.85**	1.24	0.53	0.64	-0.33	1.00
Restricted Robust-CI bounds	-0.39	1.94**	1.30*	0.68	0.92	0.66	0.91
C. Trade-off vs							
Cross Validation	-1.51*	0.62	-0.08	-0.93	0.36	0.11	-0.27
Pooled Average	-1.48*	0.69	-0.06	-0.85	0.41	0.14	-0.30
CV weighted average	0.10	-1.93**	-1.05	-0.05	1.10	0.65	-1.48*
Optimal Weight	-0.68	-0.93	0.17	-2.14**	0.62	-1.42*	-1.04
Exp smoothing (Υ=0.98)	-0.72	-1.36*	-0.64	-1.04	-0.59	-2.28**	-1.94**
Confidence interval (iterated)	-1.12	0.92	1.13	0.53	1.02	0.72	-0.80
Restricted Robust-CI bounds	-0.39	-1.04	0.88	0.61	0.99	0.96	-0.40
D. Cross validation vs							
Pooled Average	2.43**	0.54	1.15	2.54**	1.17	1.31*	-0.15
CV weighted average	1.79**	-1.22	-0.10	0.96	0.49	0.14	-0.81
Optimal Weight	1.09	-2.45**	0.15	-2.11**	0.38	-1.56*	-0.84
Exp smoothing (Υ=0.98)	0.58	-1.53*	-0.46	-0.63	-0.68	-2.26**	-1.63*
Confidence interval (iterated)	1.69**	-0.27	1.44*	1.20	0.89	0.16	-0.34
Restricted Robust-CI bounds	1.63*	-0.86	0.47	0.89	0.72	0.38	-0.08
E. Pooled average vs							
CV weighted average	1.76**	-1.30*	-0.12	0.86	0.46	0.13	-0.85
Optimal Weight	1.05	-2.38**	0.14	-2.20**	0.36	-1.57*	-0.83
Exp smoothing (Υ=0.98)	0.54	-1.60*	-0.47	-0.68	-0.70	-2.27**	-1.69**
Confidence interval (iterated)	1.65**	-0.37	1.43*	1.13	0.86	0.13	-0.34
Restricted Robust-CI bounds	1.59*	-0.93	0.46	0.85	0.69	0.37	-0.07

Table 3.10 continues

Methods	Canada	France	Germany	Italy	Japan	UK	US
F. CV weighted average vs							
Optimal Weight	-0.66	-0.34	0.33	-2.08**	0.13	-1.42*	-0.06
Exp smoothing (ϒ=0.98)	-0.75	-0.38	-0.48	-1.02	-1.04	-2.23**	-1.95**
Confidence interval (iterated)	-1.56*	1.80**	1.24	0.57	0.43	-0.09	0.99
Restricted Robust-CI bounds	-0.62	1.60*	1.14	0.64	0.65	1.06	0.83
G. Optimal weight vs							
Exp smoothing (ϒ=0.98)	-0.48	0.10	-1.14	0.27	-1.24	-1.03	-1.25
Confidence interval (iterated)	-0.54	1.56*	0.43	2.17**	0.18	1.59*	0.69
Restricted Robust-CI bounds	0.17	0.60	0.62	1.72**	0.68	1.50*	0.60
H. Exponential smoothing vs							
Confidence interval (iterated)	0.11	1.71**	1.14	1.47*	1.20	2.44**	1.75**
Restricted Robust-CI bounds	0.52	0.76	1.29*	1.76**	1.28*	2.22**	1.57*
I. Confidence interval (iterated)							
Restricted Robust-CI bounds	0.89	-1.27	-0.16	0.50	0.29	0.51	0.15

Notes: The table reports Diebold and Mariano (2002) test statistics. The null hypothesis is that a null forecast method and an alternative method have equal predictibility. The rejection of the null hypothesis is indicated by a double star at the 5% level of significance, and a single star at the 10% level of significance. If the rejection occurs with a positive test value, the alternative method is better than the null method, while with a negative test value, the null method is better than the alternative method. Finally, due to a symmetry of the test, we only report the test statistics from testing the accuracy of a first method against a second method, but not the second against a first method. For this reason, the latter panels shrink as previous panels include the information for all other possible test combinations.

Appendix C

C.1 Optimal weights in the presence of 2 breaks

As proposed by Pesaran et al. (2013) optimal weights to the observations in each coefficient break segment for the case of 2 breaks (3 segments) are given by

$$w_{(1)} = \frac{1}{T}\frac{1 + T(b_2 - b_1)\phi_{(2)}^2 - T(b_2 - b_1)\phi_{(1)}\phi_{(2)}}{a_{a,2}}$$

$$w_{(2)} = \frac{1}{T}\frac{1 + Tb_1\phi_{(1)}^2 - Tb_1\phi_{(1)}\phi_{(2)}}{a_{a,2}}$$

$$w_{(3)} = \frac{1}{T}\frac{1 + Tb_1\phi_{(1)}^2 - T(b_2 - b_1)\phi_{(2)}^2}{a_{a,2}}$$

where $a_{a,2} = 1+T(1-b_2)b_1\phi_{(1)}^2+T(b_2-b_1)(1-b_2)\phi_{(2)}^2+Tb_1(b_2-b_1)(\phi_{(1)}-\phi_{(2)})^2$ and $\phi_{(i)} = \frac{\mathbf{x}_T'\theta_{(i)}}{(\mathbf{x}_T'\Omega_{xx}^{-1}\mathbf{x}_T)^{1/2}}$ with $E(\mathbf{x}_t\mathbf{x}_t') = \Omega_{xx}$ and $\theta_{(i)} = (\beta_i - \beta_3)/\sigma$ for $i = 1, 2$. The fraction of pre-break sample for each break is $b_i = \hat{T}_i/T$. Here the authors assume a constant residual variance and focus only on breaks in the coefficients.

C.2 Approximated observational Weights

Weights on sample observations can be approximated for the following window forecast methods based on their relative contributions to the sample used for estimation different windows.

1. Post-break window

   Given the time of the break $\hat{T}_1$ is either known or estimated, the weight on observation at time t is defined as

$$w_t = \begin{cases} 0 & for \quad t \leq \hat{T}_1 \\ \frac{1}{(T-\hat{T}_1)} & for \quad t > \hat{T}_1 \end{cases}$$

2. Cross - validation

   Assuming that cross validation approach finds the optimal starting point of the estimation sample as m^*, then the estimation period has a sample of $[m^* : T]$ and weights on observations can be approximated as

$$w_t = \begin{cases} 0 & for \;\; t < m^* \\ \frac{1}{(T-m^*-1)} & for \;\; t \geq m^* \end{cases}$$

3. Averaging windows using equal weights

   After reserving a minimum estimation sample $\underline{w}$, starting point of the estimation window is defined as $m = 1, 2, ..., T - \underline{w}$ and window size spans $[\underline{w}, T]$. By adding up the weight given to each observation in each window, we can approximate weights as,

$$w_t = \begin{cases} \frac{1}{(T-\underline{w})} \sum_{m=1}^{t} \frac{1}{T-m+1} & for \;\; t \leq T - \underline{w} \\ \frac{1}{(T-\underline{w})} \sum_{m=1}^{T-\underline{w}} \frac{1}{T-m+1} & for \;\; T - \underline{w} < t \leq T \end{cases}$$

   Since the last $\underline{w}$ observations are used in all estimation windows, they receive the highest and equal weights.

4. Averaging windows using cross validation weights

   Using a similar method as equal weighting scheme, we can approximate a weight on observation at time t as,

$$w_t = \begin{cases} \sum_{m=1}^{t} \frac{\mu_m}{T-m+1} & for \;\; t \leq T - \tilde{w} - \underline{w} \\ \sum_{m=1}^{T-\tilde{w}-\underline{w}} \frac{\mu_m}{T-m+1} & for \;\; T - \tilde{w} - \underline{w} < t \leq T \end{cases}$$

   where $\mu_m = MSFE(m|T, \tilde{w}) / \sum_{m=1}^{T-\tilde{w}-\underline{w}} MSFE(m|T, \tilde{w})$, is a cross validation weight given to each window starting from $m = 1, ..., T - \tilde{w} - \underline{w}$ with $\tilde{w}$ is being a pseudo out of sample estimation sample.

C.3 Additional Tables

Table C.1: MSFE ratios (T=200, $T_1 = 50$)

Sample size: T=200 Break location: T_1=50	No Break	Small break in β_{11}	Large break in β_{11}	Small break in β_{12}	Large break in β_{12}	Break in β_{11} and β_{12}	Increase in σ_{yt}	Decrease in σ_{yt}	EXP6 + EXP 7	EXP6 + EXP 8
	EXP 1	EXP 2	EXP 3	EXP 4	EXP 5	EXP 6	EXP 7	EXP 8	EXP 9	EXP 10
A. Single estimation windows										
Post break	1.014	0.650	0.500	0.916	0.713	0.659	1.056	0.998	1.020	0.343
Trade-off	1.010	0.656	0.517	0.916	0.714	0.662	1.040	0.997	1.002	0.347
Cross Validation (CV)	1.017	0.659	0.507	0.927	0.725	0.669	1.019	1.009	0.976	0.348
B. Averaging multiple windows										
Pooled (without break info)	1.016	0.673	0.529	0.925	0.729	0.680	1.016	1.007	0.975	0.366
Pooled (with break info)	1.016	0.818	0.779	0.943	0.817	0.795	1.013	1.007	0.975	0.601
Cross validated weights	1.008	0.796	0.737	0.942	0.799	0.774	1.008	1.001	0.974	0.507
C. Observational weight										
Optimal weight	1.005	0.678	0.549	0.922	0.729	0.666	1.022	0.997	0.996	0.350
Unrestricted Robust weight	1.020	0.722	0.670	0.929	0.734	0.694	1.019	1.010	0.981	0.387
Restricted Robust weight (75-95)	1.120	0.718	0.551	1.014	0.791	0.728	1.112	1.109	1.065	0.373
Exponential (Y=0.95)	1.072	0.692	0.533	0.972	0.759	0.702	1.065	1.063	1.024	0.360
Exponential (Y=0.98)	1.020	0.718	0.666	0.928	0.732	0.691	1.019	1.010	0.980	0.382
D. Confidence interval methods										
Confidence Interval (CI)	1.011	0.652	0.501	0.917	0.713	0.661	1.044	0.996	1.004	0.344
Confidence Interval (iterated)	1.013	0.652	0.501	0.917	0.713	0.660	1.015	1.000	0.954	0.341
Confidence interval (CV weight)	1.002	0.652	0.501	0.916	0.713	0.660	0.998	0.998	0.954	0.341
EM confidence interval	1.045	0.759	0.661	0.953	0.781	0.753	1.036	1.032	0.996	0.493
Restricted Robust -CI bounds	1.030	0.651	0.499	0.916	0.713	0.660	1.031	1.016	0.954	0.341

Note: Same as Table 3.3

Table C.2: MSFE ratios (T=200, $T_1 = 100$)

Sample size: T=200 Break location: T_1=100	No Break	Small break in β_{11}	Large break in β_{11}	Small break in β_{12}	Large break in β_{12}	Break in β_{11} and β_{12}	Increase in σ_{yt}	Decrease in σ_{yt}	EXP6 + EXP 7	EXP6 + EXP 8
	EXP 1	EXP 2	EXP 3	EXP 4	EXP 5	EXP 6	EXP 7	EXP 8	EXP 9	EXP 10
A. Single estimation windows										
Post break	1.014	0.535	0.411	0.774	0.453	0.465	1.123	1.003	0.960	0.184
Trade-off	1.010	0.542	0.429	0.774	0.454	0.468	1.082	1.002	0.947	0.189
Cross Validation (CV)	1.017	0.540	0.415	0.782	0.458	0.470	1.014	1.009	0.936	0.186
B. Averaging multiple windows										
Pooled (without break info)	1.016	0.617	0.528	0.804	0.528	0.548	1.018	0.999	0.936	0.307
Pooled (with break info)	1.016	0.796	0.785	0.867	0.687	0.719	1.015	1.001	0.948	0.578
Cross validated weights	1.008	0.762	0.733	0.858	0.631	0.668	1.010	0.996	0.948	0.417
C. Observational weight										
Optimal weight	1.005	0.563	0.461	0.790	0.476	0.473	1.046	1.001	0.945	0.188
Unrestricted Robust weight	1.020	0.725	0.738	0.810	0.540	0.595	1.022	1.002	0.949	0.375
Restricted Robust weight (75-95)	1.120	0.581	0.448	0.845	0.495	0.506	1.114	1.098	1.009	0.200
Exponential (Y=0.95)	1.072	0.564	0.455	0.811	0.475	0.489	1.068	1.053	0.970	0.194
Exponential (Y=0.98)	1.020	0.698	0.711	0.798	0.512	0.557	1.022	1.002	0.940	0.320
D. Confidence interval methods										
Confidence Interval (CI)	1.011	0.537	0.412	0.775	0.453	0.467	1.091	1.000	0.949	0.185
Confidence Interval (iterated)	1.013	0.537	0.412	0.775	0.453	0.467	1.021	0.998	0.947	0.184
Confidence interval (CV weight)	1.002	0.537	0.412	0.774	0.453	0.467	1.002	0.995	0.939	0.184
EM confidence interval	1.045	0.695	0.625	0.833	0.596	0.603	1.038	1.026	0.955	0.396
Restricted Robust -CI bounds	1.030	0.537	0.412	0.774	0.453	0.466	1.030	1.015	0.947	0.184

Note: Same as Table 3.3

Table C.3: MSFE ratios (T=200, $T_1 = 150$)

Sample size: T=200 Break location: T_1=150	No Break	Small break in β_{11}	Large break in β_{11}	Small break in β_{12}	Large break in β_{12}	Break in β_{11} and β_{12}	Increase in σ_{yt}	Decrease in σ_{yt}	EXP6 + EXP 7	EXP6 + EXP 8
	EXP 1	EXP 2	EXP 3	EXP 4	EXP 5	EXP 6	EXP 7	EXP 8	EXP 9	EXP 10
A. Single estimation windows										
Post break	1.014	0.479	0.351	0.678	0.312	0.379	1.234	1.002	0.960	0.135
Trade-off	1.010	0.486	0.374	0.675	0.316	0.383	1.146	1.001	0.934	0.141
Cross Validation (CV)	1.017	0.704	0.685	0.763	0.504	0.587	1.014	1.011	0.942	0.411
B. Averaging multiple windows										
Pooled (without break info)	1.016	0.697	0.650	0.770	0.560	0.632	1.019	1.000	0.928	0.492
Pooled (with break info)	1.016	0.806	0.790	0.829	0.675	0.749	1.015	1.001	0.952	0.653
Cross validated weights	1.008	0.852	0.844	0.868	0.728	0.812	1.008	1.004	0.969	0.727
C. Observational weight										
Optimal weight	1.005	0.511	0.409	0.692	0.348	0.389	1.080	1.001	0.923	0.143
Unrestricted Robust weight	1.020	0.756	0.734	0.768	0.560	0.688	1.024	1.001	0.951	0.563
Restricted Robust weight (75-95)	1.120	0.495	0.368	0.683	0.324	0.395	1.119	1.085	0.942	0.141
Exponential (ϒ=0.95)	1.072	0.564	0.536	0.672	0.340	0.422	1.073	1.043	0.918	0.190
Exponential (ϒ=0.98)	1.020	0.743	0.722	0.750	0.527	0.646	1.023	1.000	0.937	0.509
D. Confidence interval methods										
Confidence Interval (CI)	1.011	0.479	0.353	0.674	0.313	0.380	1.180	1.001	0.939	0.136
Confidence Interval (iterated)	1.013	0.480	0.353	0.676	0.313	0.380	1.024	0.992	0.981	0.135
Confidence interval (CV weight)	1.002	0.692	0.670	0.753	0.497	0.572	1.002	0.991	0.966	0.404
EM confidence interval	1.045	0.658	0.599	0.771	0.507	0.583	1.052	1.025	0.954	0.396
Restricted Robust -CI bounds	1.030	0.479	0.354	0.674	0.313	0.380	1.029	1.015	0.974	0.135

Note: Same as Table 3.3

Table C.4: MSFE ratios (T=500)

Sample size: T=500	Increase in σ_{yt}	Decrease in σ_{yt}	Increase in σ_{yt}	Decrease in σ_{yt}	Increase in σ_{yt}	Decrease in σ_{yt}
	T_1=125		T_1=250		T_1=375	
	EXP 7	EXP 8	EXP 7	EXP 8	EXP 7	EXP 8
A. Single estimation windows						
Post break	1.019	0.999	1.047	0.996	1.085	0.999
Trade-off	1.015	0.999	1.035	0.996	1.059	0.999
Cross Validation (CV)	1.004	0.999	1.004	0.997	1.004	0.996
B. Averaging multiple windows						
Pooled (without break info)	1.005	0.997	1.006	0.994	1.007	0.994
Pooled (with break info)	1.004	0.997	1.005	0.996	1.005	0.995
Cross validated weights	1.002	0.995	1.003	0.994	1.003	0.997
C. Observational weight						
Optimal weight	1.009	0.998	1.023	0.997	1.039	0.999
Unrestricted Robust weight	1.008	1.001	1.010	0.997	1.011	0.996
Restricted Robust weight (75-95)	1.044	1.040	1.045	1.035	1.047	1.031
Exponential (ϒ=0.95)	1.084	1.089	1.086	1.084	1.087	1.079
Exponential (ϒ=0.98)	1.029	1.023	1.030	1.018	1.033	1.014
D. Confidence interval methods						
Confidence Interval (CI)	1.014	0.999	1.031	0.997	1.057	0.999
Confidence Interval (iterated)	1.005	0.997	1.007	0.994	1.009	0.993
Confidence interval (CV weight)	1.000	0.995	1.001	0.992	1.001	0.993
EM confidence interval	1.021	1.019	1.028	1.016	1.030	1.017
Restricted Robust -CI bounds	1.009	1.004	1.008	1.001	1.009	1.003

Note: Same as Table 3.3

Table C.5: MSFE ratios under multiple breaks (T=200)

Sample size: T=200 Break locations: T_1=66, T_2=132	Coeff reversion in β_{11}	Decline in β_{11}	Coeff reversion in β_{11}	Increase in β_{11}	Coeff reversion in β_{12}	Increase in β_{12}	Increase in σ_{yt}	Decline in σ_{yt}
	EXP 11	**EXP 12**	**EXP 13**	**EXP 14**	**EXP 15**	**EXP 16**	**EXP 17**	**EXP 18**
A. Single estimation windows								
Post break	0.914	0.470	0.998	0.740	0.774	0.202	1.199	1.003
Post break	0.881	0.470	0.988	0.741	0.761	0.202	1.154	1.003
Trade-off	0.907	0.469	0.991	0.737	0.771	0.203	1.129	1.002
Trade-off	0.875	0.470	0.982	0.738	0.754	0.203	1.100	1.002
Cross Validation (CV)	0.913	0.473	0.966	0.734	0.788	0.216	1.011	1.013
Cross Validation (CV)	0.866	0.471	0.953	0.736	0.748	0.212	1.016	1.013
B. Averaging multiple windows								
Pooled (without break info)	0.950	0.545	0.966	0.757	0.902	0.341	1.019	0.998
Pooled (without break info)	0.871	0.546	0.951	0.758	0.749	0.340	1.021	0.997
Pooled (with break info)	1.012	0.647	0.991	0.812	1.030	0.454	1.015	0.998
Pooled (with break info)	0.901	0.648	0.968	0.813	0.801	0.454	1.014	0.998
Cross validated weights	1.015	0.617	1.005	0.808	1.018	0.376	1.009	0.999
Cross validated weights	0.900	0.616	0.980	0.808	0.791	0.371	1.008	0.998
C. Observational weight								
Optimal weight	0.898	0.510	0.981	0.740	0.763	0.223	1.070	1.002
Optimal weight	0.872	0.509	0.977	0.741	0.752	0.222	1.062	1.002
Optimal weight (multiple break)	0.893	0.763	0.982	0.766	0.795	0.589	1.034	1.003
Unrestricted Robust weight	0.946	0.696	0.980	0.755	0.883	0.362	1.024	1.000
Unrestricted Robust weight	0.873	0.697	0.961	0.756	0.762	0.362	1.025	0.999
Restricted Robust weight (75-95)	0.939	0.484	1.001	0.761	0.795	0.213	1.118	1.089
Restricted Robust weight (75-95)	0.941	0.484	1.000	0.763	0.796	0.213	1.125	1.089
Exponential (Y=0.95)	0.910	0.475	0.969	0.737	0.784	0.210	1.071	1.045
Exponential (Y=0.95)	0.902	0.475	0.971	0.738	0.763	0.210	1.076	1.045
Exponential (Y=0.98)	0.931	0.676	0.967	0.745	0.851	0.326	1.023	0.998
Exponential (Y=0.98)	0.870	0.677	0.952	0.746	0.756	0.326	1.025	0.998
D. Confidence interval methods								
Confidence Interval (CI)	0.907	0.468	0.990	0.735	0.772	0.202	1.159	1.001
Confidence Interval (CI)	0.877	0.469	0.981	0.737	0.759	0.202	1.127	1.002
Confidence Interval (iterated)	0.906	0.469	0.991	0.736	0.771	0.202	1.027	0.993
Confidence Interval (iterated)	0.876	0.469	0.982	0.737	0.757	0.202	1.028	0.994
Confidence interval (CV weight)	0.907	0.467	0.980	0.732	0.775	0.215	1.003	0.990
Confidence interval (CV weight)	0.866	0.467	0.973	0.734	0.745	0.211	1.003	0.991
Restricted Robust -CI bounds	0.906	0.469	0.986	0.736	0.772	0.202	1.033	1.016
Restricted Robust -CI bounds	0.876	0.470	0.982	0.737	0.756	0.202	1.026	0.994

Note: Same as Table 3.6

Table C.6: Changes in MSFE ratios in the coefficient reverting experiments

Sample size: T=200 **Break locations: T_1=66, T_2=132**	**Coeff reversion in β_{11}**	**Coeff reversion in β_{11}**	**Coeff reversion in β_{12}**
	EXP 11	**EXP 13**	**EXP 15**
Post break	-0.033	-0.010	-0.013
Trade-off	-0.032	-0.009	-0.017
Cross Validation (CV)	-0.048	-0.013	-0.040
Pooled (without break info)	-0.079	-0.015	-0.153
Pooled (with break info)	-0.111	-0.023	-0.228
Cross validated weights	-0.115	-0.026	-0.227
Optimal weight	-0.027	-0.004	-0.011
Unrestricted Robust weight	-0.073	-0.019	-0.122
Restricted Robust weight (75-95)	0.002	-0.001	0.001
Exponential (ϒ=0.95)	-0.007	0.002	-0.021
Exponential (ϒ=0.98)	-0.061	-0.014	-0.094
Confidence Interval (CI)	-0.030	-0.009	-0.013
Confidence Interval (iterated)	-0.030	-0.009	-0.014
Confidence interval (CV weight)	-0.041	-0.007	-0.030
Restricted Robust -CI bounds	-0.030	-0.003	-0.016

Note: The values in the table are differences between MSFEs of the first and second rows for each forecast method for experiments 11, 13, 15 in Table C.5. Negative values indicate gains from employing re-ordering data method in MSFE sense compared to using the original data. Similarly, positive values indicate the loss associated with a re-ordering data method.

Table C.7: Structural breaks and corresponding parameter changes in inflation models of G7 countries (based on the full sample)

Countries	**Structural Breaks**		**Mean**			**Persistence**			**Residual variance**		
	Coefficients	**Variance**	**R1**	**R2**	**R3**	**R1**	**R2**	**R3**	**R1**	**R2**	**R3**
Canada	Jan-1991	Nov-1983 Dec-1999	0.555	0.162		0.772	0.032		0.309	0.205	0.313
France	Jun-1985 Nov-1991	NA	0.770	0.263	0.134	0.797	-0.025	0.193			
Germany	Jul-1975 Aug-1994	NA	0.485	0.274	0.122	-0.011	0.595	-0.052			
Italy	Jun-1985 Jun-1995	May-1974 Nov-1982	1.075	0.443	0.193	0.799	0.447	0.580	0.294	0.457	0.122
Japan	Nov-1980 Aug-1993	Nov-1974 Oct-1986	0.713	0.167	-0.003	0.467	-0.034	-0.002	0.893	0.427	0.224
United Kingdom	Apr-1991	May-1982	0.770	0.187		0.754	0.370		0.454	0.188	
United States	Sep-1981 Jan-1991	Apr-1983 Mar-1999	0.636	0.348	0.207	0.754	0.380	0.221	0.269	0.144	0.223

Notes: Reported break dates are estimated using the full sample which includes 484 observations. Prefix R indicates a regime. In each regime, the mean is calculated as the sample average of aggregate inflation and the persistence is computed as a sum of estimated autoregressive coefficients conditional on the estimated coefficient breaks. The residual variance refers to the mean of absolute values of residuals in (3.8) in each variance break regime. These errors are used in the GLS transformation. NA means not applicable, since no break is detected.

Table C.8: MSFEs for G7 inflation (based on data using the re-ordering procedure)

Countries	Can	Fra	Ger	Ita	Jap	UK	US
A. Single estimation windows							
Post break	0.829	1.032	0.965	0.852	0.913	0.938	0.983
Trade-off	0.838	1.003	0.947	0.849	0.898	0.939	0.953
Cross Validation (CV)	0.829	0.954	0.951	0.845	0.924	0.962	0.886
B. Averaging multiple windows							
Pooled (without break info)	0.863	0.920	0.910	0.858	0.891	0.951	0.915
Pooled (with break info)	0.909	0.907	0.954	0.903	0.919	0.959	0.909
Cross validated weights	0.912	0.909	0.956	0.906	0.921	0.960	0.904
C. Observational weight							
Optimal weight	0.846	0.933	0.943	0.838	0.844	0.958	0.931
Unrestricted Robust weight	0.894	0.960	0.925	0.886	0.900	1.012	0.982
Restricted Robust weight (75-95)	0.859	1.016	0.936	0.918	0.956	1.131	1.005
Exponential (Υ=0.95)	0.869	1.057	1.005	0.978	1.009	1.332	1.177
Exponential (Υ=0.98)	0.849	0.957	0.912	0.875	0.899	1.059	1.002
D. Confidence interval methods							
Confidence Interval (CI)	0.828	1.021	0.962	0.846	0.901	0.948	0.966
Confidence Interval (iterated)	0.837	1.011	0.917	0.831	0.895	0.927	0.908
Confidence interval (CV weight)	0.838	0.986	0.918	0.846	0.884	0.927	0.908
Restricted Robust -CI bounds	0.842	1.010	0.917	0.830	0.890	0.921	0.906

Conclusion

This thesis has addressed several issues which have previously arisen in the literature of inflation dynamics and forecasting. In this conclusion, the main results of the thesis are summarized and some potential directions for further research are offered.

In the course of this thesis, an iterative structural break testing procedure proposed in chapter 1 plays a key role for the all contributions made. The method is motivated by the theoretical result that identification of breaks in one component (either mean or variance) requires consideration be given to the presence of breaks in another component (Pitarakis, 2004, Sensier and van Dijk, 2004). To avoid the potential misspecification (omitting changes in either mean or variance of a time series), the iterative procedure identifies distinct breaks in conditional mean and variance parameters by iterating tests between them, with also outliers identified in relation to conditional mean and variance break regimes.

The use of this procedure enables chapter 1 to pin down the nature and dates of change in international inflation linkages and consequently, a strong and increasing co-movement of inflation is uncovered. Additionally, as prominent outcomes from the iterative algorithm, clusters of variance breaks which reflect substantial declines in the volatility of inflation are documented across many industrialized countries, casting doubt on the common claim in the literature that changes of inflation have been mainly in the mean.

Further research will develop the results of this chapter in several aspects. Firstly, formal assessments on the performance of the iterative procedure are warranted both in terms of the size and power of the test. For instance, one of the differences of our iterative procedure from the similar procedure by Bataa et al. (2013b) is a joint test of mean and dynamic coefficients to avoid substantial oversized results in the mean break test, revealed in their Monte Carlo analysis. The improvements will be explored through simulation analysis. Secondly, due to the availability of inflation and bilateral trade data at the monthly frequency over the extended period, most countries in the set examined are currently members of the Euro area. However, this analysis does

not cover the use of third-countries' data such as large emerging economies of China and India. This would provide a clearer answer to the question of whether the observed inflation dynamics is truly global in a sense that it is dictated by globally common shocks. Thirdly, an inclusion of more countries in the analysis would help to ensure the weak exogeneity assumption on contemporaneous foreign inflation which is a relevant issue for the large economies of Germany and the US those included in our data set.

The second chapter of thesis closely relates to the first chapter in terms of both the subject of study, the globalization of inflation, and the methodology used, except it examines the separate roles of core, energy and food components in addition to the headline inflation. In this respect, the above suggested research extensions are applicable also to the analysis in chapter 2. However, the aim of the second chapter differs from the first as it attempts to shed light on the nature of the inflation globalization while the objective was to measure the degree of co-movement in chapter 1. The results reveal some important features of components, including the apparent convergences in the mean levels of aggregate and each component (with the largest extent seen in core and lesser extent in energy inflation) in the long run. Specially, the convergence in core is notable from the 1990s in countries that introduced inflation targeting and as well as in Euro area countries, indicating the importance of monetary policy in explaining the co-movements. Moreover, the short run dynamics of food and energy components implies their important roles for co-movements in aggregate inflation.

In addition to the previously suggested extensions, more efforts may be directed to the modeling aspects of inflation in chapter 2. On the one hand, the models that link domestic and foreign inflation could be extended by explicitly including monetary variables in order to assess the relative importance of policy effects in explaining the dynamics of domestic inflation compared to a proxy for the international environment. Similarly, on the other hand, the supporting results for globalization may suggest, to some extent, that inflation should be modeled as a global rather than domestic phenomenon. Then, the conventional inflation models such as the Phillips equation can be augmented by a global variable which may yield some interesting insights.

The empirical examination of aggregate and component inflation in chapter 1 and chapter 2 show that structural breaks either in the conditional mean or variance parameters of inflation are a common feature. This is an important source of forecast failure if unaccounted for, as forecasts are generated relying on parameter estimates of a particular model. Therefore, the third chapter of the thesis focuses on the problem of forecasting in the presence of structural

breaks. In particular, two contributions are made to the literature on forecasting. Firstly, a forecast method that deals with structural break uncertainty is proposed. Specifically, the method involves averaging forecasts from different estimation windows where the range of estimation windows is selected using a confidence interval for the estimated break date. This method adds value to the forecast combination literature in the presence of both large and small breaks occurring in the coefficients of the forecast model, shown in the Monte Carlo simulations and empirical analysis to univariate inflation models.

The second contribution of chapter 3 relates to the case where coefficient changes are reversed with multiple breaks. In order to exploit additional information from similar but distinct coefficient regimes, a data re-ordering procedure is proposed in which the similarity of coefficients is tested and observations in the corresponding segments are re-ordered based on their p-value ranks. Monte Carlo simulation shows improvement in forecast accuracy for all methods, often substantially for the experiments with multiple breaks which have the form of reverting coefficients.

This chapter could be developed in several ways. Firstly, in the re-ordering procedure, not only the orders of coefficient break segments are rearranged but also some of the segments can be combined if the estimated coefficients in those segments are judged to be not statistically different and the corresponding breaks are removed. However, when two or more segments are combined and re-ordered based on their p-value ranks, it is possible that two segments that were previously apart from each other come together. In this case, the possibility of a break between these regimes should be re-assessed. Secondly, at the end of the re-ordering procedure, only the most recent break is acknowledged as a break although the location of such break can move to reflect the pooled segments. Then, due to time constraints and for sake of simplicity, the width of the original confidence interval is applied to the newly located break. This is not the statistically correct way of handling confidence interval. This should be re-estimated and the performance of this procedure re-assessed through Monte Carlo simulation.

Finally, this thesis can be further developed by incorporating the results from the first and second chapters within the forecasting methodology proposed in the third chapter to improve forecasting performance for inflation series. Specifically, chapter 1 and 2 suggest that domestic inflation, at some extent, should be modeled with global variables to reflect ongoing inflation globalization and use of such information may be anticipated to increase the predictability of inflation. Nevertheless, the empirical analysis using univariate inflation for G7 countries in chapter 3 yields overall good forecast performance

for our proposed methodology compared to most other forecast methods. Potentially, the application of our proposed methodology to inflation models augmented by global variables may offer further improvements in forecasting accuracy.

Bibliography

Ahmed, S., Levin, A., and Wilson, B. A. (2004): "Recent U.S. Macroeconomic Stability: Good Policies, Good Practices, or Good Luck?" *Review of Economics and Statistics*, 86, 824–832.

Altissimo, F., Bilke, L., Levin, A., Matha, T., and Mojon, B. (2006): "Sectoral and Aggregate Inflation Dynamics in the Euro Area," *Journal of the European Economic Association*, 4, 585–593.

Andrews, D. W. K. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858.

——— (1993): "Tests for Parameter Instability and Structural Change With Unknown Change Point," *Econometrica*, 61, 821–856.

Andrews, D. W. K. and Chen, H.-Y. (1994): "Approximately Median-Unbiased Estimation of Autoregressive Models," *Journal of Business & Economic Statistics*, 12, 187–204.

Bai, J. (1997): "Estimating Multiple Breaks One at a Time," *Econometric Theory*, 13, 315–352.

Bai, J. and Perron, P. (1998): "Estimating and Testing Linear Models with Multiple Structural Changes," *Econometrica*, 66, 47–78.

——— (2003a): "Computation and analysis of multiple structural change models," *Journal of Applied Econometrics*, 18, 1–22.

——— (2003b): "Critical values for multiple structural change tests," *Econometrics Journal*, 6, 72–78.

——— (2006): "Multiple Structural Change Models: A Simulation Analysis," In Econometric Theory and Practice: Frontiers of Analysis and Applied Research, D. Corbea, S. Durlauf and B. E. Hansen (eds.), Cambridge University Press, 212–237.

Ball, L. M. (2006): "Has Globalization Changed Inflation?" Working Paper 12687, National Bureau of Economic Research.

Banerjee, A., Marcellino, M., and Masten, I. (2008): "Chapter 4 Forecasting Macroeconomic Variables Using Diffusion Indexes in Short Samples with Structural Change," *Frontiers of Economics and Globalization*, 3, 149–194.

Bataa, E., Osborn, D. R., Sensier, M., and van Dijk, D. (2013a): "Structural Breaks in the International Dynamics of Inflation," *Review of Economics and Statistics*, 95, 646–659.

Bataa, E., Osborn, D. R., Sensier, M., and Dijk, D. v. (2013b): "Identifying Changes in Mean, Seasonality, Persistence and Volatility for G7 and Euro Area Inflation," *Oxford Bulletin of Economics and Statistics*, forthcoming.

Benati, L. (2008): "Investigating Inflation Persistence Across Monetary Regimes," *The Quarterly Journal of Economics*, 123, 1005–1060.

Benati, L. and Kapetanios, G. (2002): "Structural Breaks in Inflation Dynamics," manuscript, Bank of England.

Bernanke, B. S. (2007): "Globalization and Monetary Policy," Speech presented at the Fourth Economic Summit, Stanford Institute for Economic Policy Research, 2 March. Downloaded from www.federalreserve.gov/newsevents/.

Borio, C. and Filardo, A. (2007): "Globalisation and inflation: New cross-country evidence on the global determinants of domestic inflation," Working Paper 227, Monetary and Economic Department, Bank for International Settlements, Basel, Switzerland.

Calza, A. (2009): "Globalization, Domestic Inflation and Global Output Gaps: Evidence from the Euro Area," *International Finance*, 12, 301–320.

Canova, F. and Ferroni, F. (2012): "The dynamics of US inflation: Can monetary policy explain the changes?" *Journal of Econometrics*, 167, 47–60.

Castle, J. L. and Hendry, D. F. (2008): "Chapter 2 Forecasting UK Inflation: The Roles of Structural Breaks and Time Disaggregation," *Frontiers of Economics and Globalization*, 3, 41–92.

Cecchetti, S. G. and Debelle, G. (2006): "Has the inflation process changed?" *Economic Policy*, 21, 311–352.

Census Bureau, U. (2011): "X-12-ARIMA reference Manual."

Chen, C. and Liu, L.-M. (1993): "Joint Estimation of Model Parameters and Outlier Effects in Time Series," *Journal of the American Statistical Association*, 88, 284–297.

Chow, G. C. (1960): "Tests of Equality Between Sets of Coefficients in Two Linear Regressions," *Econometrica*, 28, 591–605.

Chu, C.-S. J., Stinchcombe, M., and White, H. (1996): "Monitoring Structural Change," *Econometrica*, 64, 1045–1065.

Ciccarelli, M. and Mojon, B. (2010): "Global Inflation," *Review of Economics and Statistics*, 92, 524–535.

Clarida, R., Gali, J., and Gertler, M. (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *The Quarterly Journal of Economics*, 115, 147–180.

Clark, T. E. (2001): "Comparing measures of core inflation," *Economic Review-Federal Reserve Bank of Kansas City*, 86, 5–32.

——— (2006): "Disaggregate evidence on the persistence of consumer price inflation," *Journal of Applied Econometrics*, 21, 563–587.

Clark, T. E. and McCracken, M. W. (2005): "The power of tests of predictive ability in the presence of structural breaks," *Journal of Econometrics*, 124, 1–31.

——— (2008): "Chapter 3 Forecasting with Small Macroeconomic VARs in the Presence of Instabilities," *Frontiers of Economics and Globalization*, 3, 93–147.

——— (2009): "Improving Forecast Accuracy by Combining Recursive and Rolling Forecasts," *International Economic Review*, 50, 363–395.

——— (2013): "Advances in Forecast Evaluation," In Handbook of Economic Forecasting SET 2A-2B, Elliot, G., and Timmermann, A. (eds.), Elsevier.

Clark, T. E. and West, K. D. (2007): "Approximately normal tests for equal predictive accuracy in nested models," *Journal of Econometrics*, 138, 291–311.

Corvoisier, S. and Mojon, B. (2005): "Breaks in the Mean of Inflation: How they Happen and What to do With Them," Working Paper No. 451, European Central Bank, Rochester, NY.

Davidian, M. and Carroll, R. J. (1987): "Variance Function Estimation," *Journal of the American Statistical Association*, 82, 1079–1091.

Dees, S., Mauro, F. d., Pesaran, M. H., and Smith, L. V. (2007): "Exploring the international linkages of the euro area: a global VAR analysis," *Journal of Applied Econometrics*, 22, 1–38.

Diebold, F. X. and Chen, C. (1996): "Testing structural stability with endogenous breakpoint A size comparison of analytic and bootstrap procedures," *Journal of Econometrics*, 70, 221–241.

Diebold, F. X. and Mariano, R. S. (2002): "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, 20, 134–144.

Eklund, J., Kapetanios, G., and Price, S. (2013): "Robust Forecast Methods and Monitoring during Structural Change," *The Manchester School*, forthcoming.

Elliott, G. (2005): "Forecasting when there is a single break," manuscript, University of California at San Diego.

Elliott, G. and Muller, U. K. (2007): "Confidence sets for the date of a single break in linear time series regressions," *Journal of Econometrics*, 141, 1196–1218.

EVIEWS. (2009): *software version 7 (EVIEWS 7)*, IHS Global Inc., United States.

Gadzinski, G. and Orlandi, F. (2004): "Inflation Persistence in the European Union, the Euro area, and the United States," Working Paper No. 414, European Central Bank, Rochester, NY.

Giordani, P., Kohn, R., and van Dijk, D. (2007): "A unified approach to nonlinearity, structural change, and outliers," *Journal of Econometrics*, 137, 112–133.

Giraitis, L., Kapetanios, G., and Price, S. (2013): "Adaptive forecasting in the presence of recent and ongoing structural change," *Journal of Econometrics*, forthcoming.

Hall, A. R., Han, S., and Boldea, O. (2012): "Inference regarding multiple structural changes in linear models with endogenous regressors," *Journal of Econometrics*, 170, 281–302.

Hall, A. and Sakkas, N. (2013): "Approximate p-values of certain tests involving hypotheses about multiple breaks," *Journal of Econometric Methods*, forthcoming.

Hansen, B. E. (2000): "Testing for structural change in conditional models," *Journal of Econometrics*, 97, 93–115.

Hendry, D. F. and Clements, M. P. (2003): "Economic forecasting: some lessons from recent research," *Economic Modelling*, 20, 301–329.

Hooker, M. A. (2002): "Are Oil Shocks Inflationary?: Asymmetric and Nonlinear Specifications versus Changes in Regime," *Journal of Money, Credit, and Banking*, 34, 540–561.

Ihrig, J., Kamin, S. B., Lindner, D., and Marquez, J. (2010): "Some Simple Tests of the Globalization and Inflation Hypothesis," *International Finance*, 13, 343–375.

Kapetanios, G., Labhard, V., and Price, S. (2008): "Forecasting Using Bayesian and Information-Theoretic Model Averaging," *Journal of Business & Economic Statistics*, 26, 33–41.

Krolzig, H.-M. and Hendry, D. F. (2001): "Computer automation of general-to-specific model selection procedures," *Journal of Economic Dynamics and Control*, 25, 831–866.

Levin, A. T. and Piger, J. (2003): "Is Inflation Persistence Intrinsic in Industrial Economies?" Working Paper No. 2002-023A, Federal Reserve Bank of St. Louis, Rochester, NY.

MATLAB. (2011): *software version 7.12.0.635 (R2011a)*, The MatWorks Inc., Natick, Massachusetts, United States.

McConnell, M. M. and Perez-Quiros, G. (2000): "Output Fluctuations in the United States: What Has Changed Since the Early 1980's?" *The American Economic Review*, 90, 1464–1476.

Mishkin, F. S. (2007): "Headline versus core inflation in the conduct of monetary policy," Speech presented at the Business Cycles, International Transmission and Macroeconomic Policies Conference, HEC Montreal, Canada, 20 October. Downloaded from http://www.drduru.com/money/SavedFiles/071020_FRB- Speech-Mishkin-Ination.htm.

——— (2009): "Globalization, Macroeconomic Performance, and Monetary Policy," *Journal of Money, Credit and Banking*, 41, 187–196.

Monacelli, T. and Sala, L. (2009): "The International Dimension of Inflation: Evidence from Disaggregated Consumer Price Data," *Journal of Money, Credit and Banking*, 41, 101–120.

Mumtaz, H., Simonelli, S., and Surico, P. (2011): "International comovements, business cycle and inflation: A historical perspective," *Review of Economic Dynamics*, 14, 176–198.

Mumtaz, H. and Surico, P. (2012): "Evolving International Inflation Dynamics: World and Country-Specific Factors," *Journal of the European Economic Association*, 10, 716–734.

Neely, C. J. and Rapach, D. E. (2011): "International comovements in inflation rates and country characteristics," *Journal of International Money and Finance*, 30, 1471–1490.

O'Reilly, G. and Whelan, K. (2005): "Has Euro-Area Inflation Persistence Changed Over Time?" *Review of Economics and Statistics*, 87, 709–720.

Orphanides, A. (2004): "Monetary Policy Rules, Macroeconomic Stability, and Inflation: A View from the Trenches," *Journal of Money, Credit, and Banking*, 36, 151–175.

Osborn, D. R. and Ghysels, E. (2001): *The Econometric Analysis of Seasonal Time Series*, Cambridge University Press.

Peacock, C. and Baumann, U. (2008): "Globalisation, import prices and inflation dynamics," Working Paper No. 359, Bank of England.

Pesaran, M. H., Pettenuzzo, D., and Timmermann, A. (2006): "Forecasting Time Series Subject to Multiple Structural Breaks," *The Review of Economic Studies*, 73, 1057–1084.

Pesaran, M. H. and Pick, A. (2011): "Forecast Combination Across Estimation Windows," *Journal of Business & Economic Statistics*, 29, 307–318.

Pesaran, M. H., Pick, A., and Pranovich, M. (2013): "Optimal forecasts in the presence of structural breaks," *Journal of Econometrics*, forthcoming.

Pesaran, M. H., Schuermann, T., and Smith, L. V. (2009): "Forecasting economic and financial variables with global VARs," *International Journal of Forecasting*, 25, 642–675.

Pesaran, M. H., Schuermann, T., and Weiner, S. M. (2004): "Modeling Regional Interdependencies Using a Global Error-Correcting Macroeconometric Model," *Journal of Business & Economic Statistics*, 22, 129–162.

Pesaran, M. H. and Timmermann, A. (2005): "Small sample properties of forecasts from autoregressive models under structural breaks," *Journal of Econometrics*, 129, 183–217.

——— (2007): "Selection of estimation window in the presence of breaks," *Journal of Econometrics*, 137, 134–161.

Pesaran, M. and Timmermann, A. (2002): "Market timing and return prediction under model instability," *Journal of Empirical Finance*, 9, 495–510.

——— (2004): "How costly is it to ignore breaks when forecasting the direction of a time series?" *International Journal of Forecasting*, 20, 411–425.

Pitarakis, J.-Y. (2004): "Least squares estimation and tests of breaks in mean and variance under misspecification," *Econometrics Journal*, 7, 32–54.

Pivetta, F. and Reis, R. (2007): "The persistence of inflation in the United States," *Journal of Economic Dynamics and Control*, 31, 1326–1358.

Qu, Z. and Perron, P. (2007): "Estimating and Testing Structural Changes in Multivariate Regressions," *Econometrica*, 75, 459–502.

Robbins, H. (1970): "Statistical Methods Related to the Law of the Iterated Logarithm," *The Annals of Mathematical Statistics*, 41, 1397–1409.

Rossi, B. and Inoue, A. (2012): "Out-of-Sample Forecast Tests Robust to the Choice of Window Size," *Journal of Business & Economic Statistics*, 30, 432–453.

Sensier, M. and van Dijk, D. (2004): "Testing for Volatility Changes in U.S. Macroeconomic Time Series," *Review of Economics and Statistics*, 86, 833–839.

Stock, J. H. and Watson, M. W. (1996): "Evidence on Structural Instability in Macroeconomic Time Series Relations," *Journal of Business & Economic Statistics*, 14, 11–30.

——— (2003): "Forecasting Output and Inflation: The Role of Asset Prices," *Journal of Economic Literature*, 41, 788–829.

Summers, P. M. (2005): "What caused the Great Moderation? Some cross-country evidence," *Economic Review-Federal Reserve Bank of Kansas City*, 90, p. 5.

Tian, J. and Anderson, H. M. (2011): "Forecasting Under Structural Break Uncertainty," working paper, Monash University, Department of Econometrics and Business Statistics.

Trichet, J.-C. (2008): "Globalisation, inflation and the ECB monetary policy," Lecture at the Barcelona Graduate School of Economics, 14 February. Downloaded from www.ecb.int/press/.

Wang, P. and Wen, Y. (2007): "Inflation dynamics: A cross-country investigation," *Journal of Monetary Economics*, 54, 2004–2031.

Yao, Y.-C. (1988): "Estimating the number of change-points via Schwarz' criterion," *Statistics & Probability Letters*, 6, 181–189.

Yohei Yamamoto. (2012): *m-Break*, [MATLAB code]. Downloaded from http://people.bu.edu/perron/.