

NONLINEARITIES IN MACROECONOMIC POLICY

A Ph.D. Dissertation

by
ÖMER FARUK AKBAL

Department of
Economics
İhsan Doğramacı Bilkent University
Ankara
August 2022

To my family and friends



NONLINEARITIES IN MACROECONOMIC POLICY

The Graduate School of Economics and Social Sciences
of
İhsan Doğramacı Bilkent University

by

ÖMER FARUK AKBAL

In Partial Fulfillment of the Requirements for the Degree of
DOCTOR OF PHILOSOPHY IN ECONOMICS

THE DEPARTMENT OF
ECONOMICS
İHSAN DOĞRAMACI BİLKENT UNIVERSITY
ANKARA

August 2022

NONLINEARITIES IN MACROECONOMIC POLICY

By Ömer Faruk Akbal

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Economics.

Sang Seok Lee (Advisor)

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Economics.

Levent Akdeniz

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Economics.

Cem Çakmaklı

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Economics.

Burçin Kısacıkoğlu

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy in Economics.

Seda Köymen Özer

Approval of the Graduate School of Economics and Social Sciences

Refet Soykan Gürkaynak
Director

ABSTRACT

NONLINEARITIES IN MACROECONOMIC POLICY

Akbal, Ömer Faruk

Ph.D., Department of Economics

Supervisor: Assist. Prof. Dr. Sang Seok Lee

August 2022

This thesis consists of four essays on nonlinear approaches in macroeconomic modeling. The first essay analyzes a structural vector autoregressive model to assess the contribution of inflation targeting monetary policy to economic stability. This study is conducted with the data of New Zealand, which was the first country to officially adopt inflation targeting. It is shown that the inflation targeting policy contributed to the overall macroeconomics stability. In the second essay, the expectation-maximization (EM) method, which is adapted to regime-switching settings, is derived. In the third essay, the performance of the EM method is tested by replicating well-known examples from the literature on regime-switching models. The method is shown to provide a good trade-off between accuracy and speed for a setting involving large data sets available in mixed frequencies. In the fourth essay, the performance of the EM method is tested using real-time US data. Here, a linear nowcasting model is used as a benchmark and compared to a modified framework

involving regime-switching. The nowcasting performance of both models, the benchmark and regime-switching, between 2007 and 2022 is evaluated. The regime-switching model is shown to outperform the linear model in terms of short-term prediction performance, especially during the COVID-19 pandemic.

Keywords: Regime-Switching, Monetary Policy, Dynamic Factor Models, Nowcasting, Big Data.



ÖZET

DOĞRUSAL OLMAYAN MAKROEKONOMİK POLİTİKALAR

Akbal, Ömer Faruk

Doktora, İktisat Bölümü

Tez Yöneticisi: Dr. Öğr. Üyesi Sang Seok Lee

Ağustos 2022

Bu tez, makroekonomik modellemede doğrusal olmayan yaklaşımlar üzerine dört makaleden oluşmaktadır. İlk makale, enflasyon hedeflemesi para politikasının ekonomik istikrara katkısını değerlendirmek için yapısal vektör otoregresif bir modeli analiz etmektedir. Bu çalışma, enflasyon hedeflemesini resmi olarak benimseyen ilk ülke olan Yeni Zelanda'nın verileriyle yapılmıştır. Enflasyon hedeflemesi politikasının genel makroekonomik istikrara katkı sağladığı gösterilmiştir. İkinci makalede, rejim değişimi için uyarlanmış beklenti maksimizasyon (BM) yöntemi türetilmiştir. Üçüncü makalede, BM yönteminin performansı, değişken rejim modelleri üzerine yazında iyi bilinen örnekler ile test edilmiştir. Yöntemin, farklı frekanslarda gözlem içeren büyük veri kümeleri için doğruluk ve hız arasında iyi bir denge sağladığı gösterilmiştir. Dördüncü denemede, EM yönteminin performansı gerçek zamanlı ABD verileri kullanılarak test edilmiştir. Burada, doğrusal bir angörü modeli kıyaslama amacıyla kul-

lanılmış ve rejim deęiřimi ieren bir angr modeli ile karřılařtırılmıştır. Her iki modelin de, 2007 ve 2022 yılları arasındaki performansları deęerlendirilmiřtir. Rejim deęiřimi modelinin, zellikle COVID-19 salgını sırasında kısa vadeli tahmin performansı aısından doęrusal modelden daha iyi performans gsterdięi gsterilmiřtir.

Anahtar Kelimeler: Degisken Rejim, Para Politikasi, Dinamik Faktor Modelleri, Angr, Byk Veri.



ACKNOWLEDGEMENTS

This thesis would not have been possible without the support and guidance of many people. Sincere thanks to all of them.

I would like to express my deepest gratitude to Sang Seok Lee for his invaluable guidance, support, and exceptional supervision throughout all stages of my study. It is no doubt that working with him made a difference in my understanding of economics.

I would also like to thank Levent Akdeniz and Burçin Kısacıkoğlu, who are the thesis monitoring committee members, for their continuous support and guidance. I wish to extend my thanks to Cem Çakmaklı and Seda Köymen Özer, who are the examining committee members, for their suggestions. I also wish to thank all of the professors in the Department of Economics for their support and guidance. I would like to extend my thanks to Emin Karagözoğlu and Hüseyin Çağrı Sağlam for their invaluable guidance. I need to mention Özlem Demiral and Nilgün Çorapcıoğlu for making their support available.

The financial support of TÜBİTAK during my studies is gratefully acknowledged.

It is important to thank my friends Mustafa Yeşilyurt, Yunus Emre Demirhan,

Ersin Yar, Mehmet Burak Kocamış, Yasin Soylu, and Hamdi Köyük for their sincere friendship.

Finally, my deep and sincere gratitude to my parents Dilfiruz and Zübeyir, my closest friend and lovely wife Gülserim, my sister Şeyda and my brother Ahmet Cihat for their encouragement and endless support without which I could not end what I have started. I dedicate this thesis to them.



TABLE OF CONTENTS

ABSTRACT	iii
--------------------	-----

ÖZET	v
----------------	---

TABLE OF CONTENTS	ix
-----------------------------	----

LIST OF TABLES	xii
--------------------------	-----

LIST OF FIGURES	xiii
---------------------------	------

CHAPTER 1: INTRODUCTION	1
-----------------------------------	---

CHAPTER 2: GOOD LUCK OR GOOD POLICY? A RECENT	
---	--

MACROECONOMIC HISTORY OF NEW ZEALAND	4
--------------------------------------	---

2.1 Introduction	4
----------------------------	---

2.2 Related Literature	7
----------------------------------	---

2.3 Model, Data and Methodology	11
---	----

2.3.1 Model	11
-----------------------	----

2.3.2 Data	17
----------------------	----

2.3.3 Estimation Methodology	18
--	----

2.4 Results and Discussions	20
---------------------------------------	----

2.5 Conclusion	25
--------------------------	----

**CHAPTER 3: THE ESTIMATION OF REGIME-SWITCHING
FACTOR MODELS WITH MIXED-FREQUENCY
AND MISSING DATA PROBLEMS USING THE
EXPECTATION-MAXIMIZATION ALGORITHM 27**

3.1	Introduction	27
3.2	Related Literature	29
3.3	Methodology	35
3.3.1	Definition and Notation	35
3.3.2	Estimation of Single-Regime DFMs	38
3.3.3	A Regime-Switching DFM	42
3.4	Conclusion	74

**CHAPTER 4: THE APPLICATION OF THE EXPECTATION-
MAXIMIZATION ALGORITHM TO REGIME-
SWITCHING DYNAMIC FACTOR MODELS:
REPLICATIONS AND MONTE CARLO SIM-
ULATIONS 78**

4.1	Introduction	78
4.2	Exercises	79
4.2.1	Kim and Nelson (2006)	79
4.2.2	Camacho et al. (2018)	87
4.2.3	Monte Carlo Simulation	92

CHAPTER 5: THE REGIME-SWITCHING NOWCASTING

MODEL: AN APPLICATION TO THE US ECON-

OMY 95

5.1 Introduction 95

5.2 Model and Methodology 96

5.3 Results and Discussion 99

5.4 Conclusion 105

REFERENCES 106

APPENDICES 114

A Appendix for Chapter 2 114

B Appendix for Chapter 3 152

LIST OF TABLES

2.1	Block Exogeneity Restrictions	12
2.2	Contemporaneous and Lag Structure of the Model	14
2.3	Marginal Data Density Comparison	19
2.4	Number of Lags	19
2.5	$A_0[s_t^2]$	21
2.6	Global Block	22
2.7	Domestic Block	24
4.1	Estimation Time for the Two Estimation Approaches	87
4.2	Estimation Time for the Two Estimation Approaches	92
5.1	Variables	98
2	Likelihood Ratio Test	142
3	Marginal Data Density Comparison between the Baseline and Modified Model	146

LIST OF FIGURES

2.1	CPI Inflation (YoY, % change)	8
2.2	Monetary Policy Regimes	20
2.3	Impulse Response Functions to Inflation Shock (Change in per- cent points)	22
2.4	Global Shock Regimes	23
2.5	Domestic Shock Regimes	24
4.1	Smoothed Recession Probabilities: the Original Model versus the Modified Model	85
4.2	Smoothed Recession Probabilities: the One-step Approach ver- sus the EM Approach	86
4.3	Smoothed Recession Probabilities: the Two-step Approach ver- sus the EM Approach	91
4.4	The True Regime Sequence and Three Model Estimates	94
5.1	Expanding-window RMSE and Real Time Recession Probabilities	100
5.2	Expanding-window RMSE for Each Data Release and Real Time Recession Probabilities	102
5.3	Real-time Recession Probabilities and NBER Recession Dates .	103
5.4	Backcasting of Recession Probabilities	104

5	New Zealand GDP	116
6	New Zealand CPI Inflation	116
7	New Zealand 90-day Bank Bill Rate	117
8	US GDP	117
9	Federal Funds Rate	118
10	Crude Oil Import Prices	118
11	New Zealand Currency to US Dollar	119
12	Potential Evapotranspiration Deficit	119
13	FAVAR Impulse Responses	144
14	Regime Probabilities under Baseline Model and the Robustness Check	147
15	The United States Data in the Baseline Model and Trade Weighted Counterpart	149
16	Regime Probabilities under Baseline Model and the Robustness Check	150

CHAPTER 1

INTRODUCTION

The quantity and quality of available data have substantially increased over the last decades, which brings about new opportunities as well as challenges to researchers and policymakers alike. An important task is to make the best use of these, which in many cases requires going beyond linear models. The literature has addressed these concerns by including time variations in the models, whether empirical or theoretical. This thesis focuses on regime-switching models with discrete time-variations, with the following three essays studying either policy changes or business cycle turning points.

The second chapter contributes to the literature by assessing the role of inflation-targeting monetary policy to the overall macroeconomic stabilization of New Zealand, the first country to formally adopt inflation targeting. The chapter addresses the question of whether it was good policy or good luck that achieved the stabilization through the lens of the regime-switching structural vector autoregressive (SVAR) model. The small open economy feature of New Zealand turns out to be an important advantage in assessing the relative contributions

of global conditions and domestic policy to the stabilization.

While VARs with regime-switching have been useful in empirical macroeconomics, there has been criticism of whether this class of models can sufficiently capture all relevant information within the model due to the well-known degrees of freedom issues. For this reason, in the third chapter, I study large-scale factor models which help overcome this limitation. Moreover, the chapter also demonstrates how to incorporate regime-switching into large scale factor models so that it is possible to consider both data reduction and nonlinearities jointly. In doing so, I propose to estimate regime-switching dynamic factor models (DFM) with the expectation-maximization (EM) approach which is demonstrated to provide a good trade-off between speed and accuracy.

In the fourth chapter, I test the performance of the EM approach in the regime-switching setting by replicating two well-known studies in the literature. The first example is a state-space model of ? and the second is a mixed-frequency DFM of [Camacho et al. \(2018\)](#) which extracts a real-time coincident indicator. The EM approach is used to estimate regime probabilities and these are compared to the regime probabilities provided by the aforementioned papers which are obtained using different estimation methods. I demonstrate that the EM based results are in good agreement with these even though the EM approach is shown to have nontrivial speed advantages. In addition to this, I conduct a Monte Carlo simulation study to test the validity of the EM approach.

In the fifth chapter, I propose to use the EM approach to nowcast real-time US GDP based on a mixed-frequency regime-switching DFM. The model used in

this chapter follows [Bok et al. \(2018\)](#) which documents the New York Fed Staff Nowcast. In this exercise, the GDP nowcast is estimated by the baseline model from [Bok et al. \(2018\)](#) and its regime-switching modification that I propose. I show that the regime-switching version of the model outperforms the single-regime baseline model in the context of nowcasting and also forecasting GDP at the onset of COVID-19 pandemic.

In summary, the thesis demonstrates the usefulness of the regime-switching models and the EM approach, and their potential for guiding economic policy.

CHAPTER 2

GOOD LUCK OR GOOD POLICY? A RECENT MACROECONOMIC HISTORY OF NEW ZEALAND¹

“ 2.1 Introduction

The Great Moderation was a global phenomenon that was marked by stable economic growth and inflation. From the early 1990s onward, both developed and developing countries experienced a significant reduction in economic volatility. However, there is still an ongoing debate on whether it was a change in policy or reduction in shock variances that contributed to the Great Moderation's success.

This paper addresses the question of whether it was good policy or good luck that contributed to this outcome through the lens of the regime-switching structural vector autoregression (SVAR) model. The regime-switching frame-

¹This chapter is accepted for publication at the Singapore Economic Review with DOI number 10.1142/S0217590822500436 to fulfill the publication requirement for the completion of the doctoral degree at Bilkent University. This is an edited version of the published manuscript.

work is an appropriate methodology to study this question because it can naturally accommodate these as competing explanations for the Great Moderation. To demonstrate this, I choose New Zealand, which was a pioneering country for inflation targeting.² The small open economy structure of New Zealand, which is modeled by block exogenous structural restrictions, is an advantage in assessing global conditions and domestic policy. In addition to its small open economy structure, New Zealand is an ideal laboratory to answer this question because it is the first country to adopt formal inflation targeting and systematically overcome macroeconomic stabilization challenges. Hence, it could have benefited from either policy changes or reduction in global shocks or both, which is a required precondition for the present analysis.

The literature has contested over relative contributions of shock variances and good policy to the success during the Great Moderation. While [Clarida et al. \(2000\)](#)³ highlight the role of good policy, [Sims and Zha \(2005\)](#), in their study of US monetary policy, argue that the Great Moderation was also characterized by the reduction in shock variances. Given that the US is a large country whose policy can affect the rest of the world, it is not clear to what extent the shock variance reduction was exogenous. On the other hand, New Zealand, due to being a small open economy, can circumvent this issue.

Indeed, the literature has examined this question beyond the US. [Bernanke](#)

²[Bernanke et al. \(1999\)](#) discuss that the Bundesbank had an implicit inflation targeting policy in practice and [Bollard and Karagedikli \(2006\)](#) discuss that the German experience was one to which the Reserve Bank of New Zealand looked in the early 1980s. However, the RBNZ Act of 1989 pioneered the formal and explicit adoption of inflation targeting. A detailed discussion and approximate adoption dates for inflation targeting can be found in [Svensson \(2010\)](#).

³Although this study does not explicitly focus on the Great Moderation, the time span covers the Great Moderation.

et al. (1999), in a cross-country analysis including advanced and emerging economies, show that both the realized and expected of inflation decreased considerably regardless of whether countries are inflation targeters or not. Buckle et al. (2002) and Wesselbaum (2014) are particularly pertinent to the analysis in this paper given their focus on New Zealand. This paper builds on their works using a richer model with regime-switching for structural identification. The model's main block can be taken to be a regime-switching extension to the seminal small open economy model of Kim and Roubini (2000). The model also considers contributions of climate conditions, which is an essential determinant of business cycles in New Zealand, as discussed in Buckle et al. (2007). The methodology used in this paper follows closely Sims et al. (2008). First, structural restrictions and block exogeneity are defined in the SVAR model. There are four blocks in the model: global block, domestic block, exchange rate block, and exogenous factors. International variable shocks (shocks to the global and exchange rate blocks), domestic variable shocks, and monetary policy coefficients are subject to regime shifts with independent Markov chains. Second, using a formal marginal data density (MDD) comparison procedure, the number of regimes for each independent Markov chain is determined. Then, the model is estimated using Bayesian techniques, and smoothed transition probabilities are mapped into historical events in order to discuss the contribution of reduction in shock variances and change in policy into macroeconomic stabilization.

The results suggest that both reductions in global shock variances and inflation-targeting policy played important roles in bringing about macroeconomic sta-

bilization to New Zealand during the Great Moderation. Moreover, it is shown that the Great Financial Contraction of 2008 manifested itself as global shocks to New Zealand and led to increased volatility. These demonstrate the usefulness of regime-switching empirical models for policy analysis, in particular, their abilities to deliver economic assessments in a unified framework.

The paper is organized as follows. Section 2.2 briefly discusses the related literature. Section 2.3 explains the model, data, and methodology. Section 2.4 provides results and discussions, and section 2.5 concludes.

2.2 Related Literature

The Great Inflation which took place between the late 1960s and late 1980s, was marked by unstable economic growth and high inflation. After fifteen years of the Great Inflation, the Great Moderation followed, which was characterized by stable growth and low inflation, globally. This was interrupted by the Great Recession, which began in the US. The financial crisis in the US spilled over to the world, and this increased economic volatility globally. Accordingly, central banks made an effort to bring international financial and macroeconomic volatilities to their pre-Great Recession levels. Figure 2.1 illustrates the consumer price inflation for selected countries from the Great Inflation to the post-Great Recession.

While most of the countries were successful in achieving low inflation levels starting from the early 1990s, the narrative and mechanism behind this success remained a popular debate among scholars. The macroeconomic stabilization,

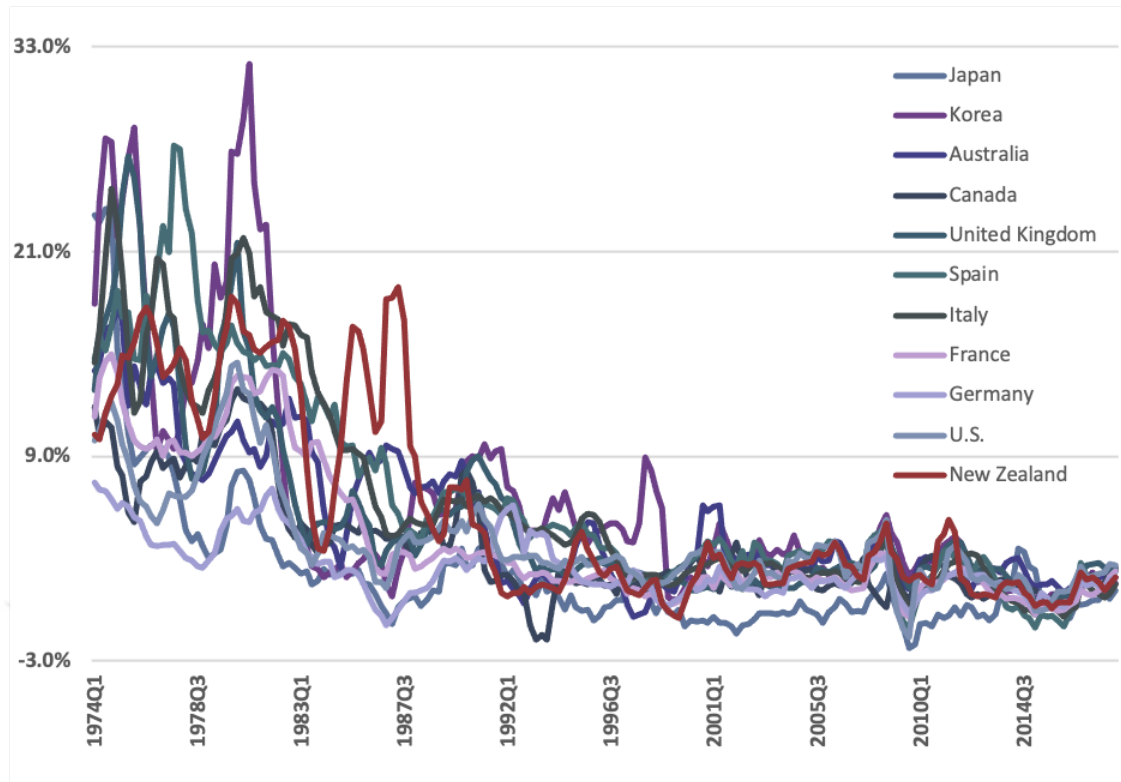


Figure 2.1: CPI Inflation (YoY, % change)

especially the reduction in inflation, has been widely examined. Some studies argued that overall progress is due to good policy, which is represented by a structural change in policy, while others sought convincing evidence in favor of good luck, i.e., reduction in shock variances.

On one hand, [Leeper and Zha \(2003\)](#); [Cogley and Sargent \(2005\)](#); [Kim and Nelson \(2006\)](#); [Boivin and Giannoni \(2006\)](#) document that the policy responses have changed in countries experiencing macroeconomic stabilization. Hence, the overall improvement is attributed to a good policy. On the other hand, [Sims and Zha \(2005\)](#); [Fernandez-Villaverde et al. \(2010\)](#); [Liu et al. \(2011\)](#) claim that the reduction in shock variance was the decisive factor. [Canova and Gambetti \(2004\)](#); [Primiceri \(2005\)](#); [Bianchi and Ilut \(2017\)](#) find evidence for both.

This paper is also related to another strand of the literature by focusing on the success of inflation targeting. This literature focuses on the question of whether inflation targeting improves inflation performance. There are a vast number of studies on inflation targeting and monetary policy, including, but not limited to, the works of [Bernanke et al. \(1999\)](#), [Mishkin and Schmidt-Hebbel \(2001\)](#), [Corbo et al. \(2001\)](#) and [Neumann and Hagen \(2002\)](#) that studied the collective progress of inflation targeting countries and concluded that adopting inflation targeting improves inflation performance. At the same time, [Cecchetti and Ehrmann \(2002\)](#) state that non-inflation-targeting countries also improved their inflation performance over the last three decades, which favors the argument involving the reduction in shock variances.

Studying the effectiveness of inflation targeting specifically in New Zealand, [Fischer \(1993\)](#) discusses the inflation performance of New Zealand and Canada, pointing to the descriptive statistics on realized and expected inflation levels. [McCallum \(1996\)](#) also discusses the limited evaluation of inflation targeting experience in New Zealand. Both of the early discussions positively commented on the overall performance. Observing the implementation of inflation targeting countries after a decade, [Kuttner \(2004\)](#) estimates a single equation forward-looking policy reaction function for New Zealand in 1997-2003 to show the interest rate rule is actually targeting the expected inflation level and the output gap in New Zealand. Similarly, [Huang et al. \(2001\)](#) estimate a time-invariant single policy rule equation for 1989-1998 and claimed that the policy rule aims at the inflation target. [Batini et al. \(2005\)](#) discuss, among other countries, medium-term macroeconomic and inflation performances improved after

adopting an inflation-targeting regime for New Zealand. They estimated the performance of variables using their past values as explanatory variables and the adoption of inflation targeting as a control variable. These studies provided a descriptive and statistical background for New Zealand inflation targeting performance. However, their approach does not evaluate the changes in the policy responsiveness over time. [Plantier and Scrimgeour \(2002\)](#) allow time-varying natural interest rate in a single equation policy rule setting. Their study showed that the policy response to inflation increased after the adoption of the inflation targeting regime. [Sekine and Teranishi \(2008\)](#) estimate a single equation interest rate rule that allows parameter and volatility changes for New Zealand between 1975 and 2005. They have observed a change in policy responsiveness in New Zealand prior to the adoption of New Zealand. On the other hand, [Baxa et al. \(2010\)](#) estimate a single interest rate rule in time-varying parameter setup from 1985 to 2007, and claimed the interest rate response to inflation didn't respond to the adoption of inflation targeting. Similarly, [Creel and Hubert \(2015\)](#) study a similar question under time-varying parameter models and regime-switching VAR models and concluded that there is no evidence of a change in the response of interest rate to inflation.

The literature is divided over the effectiveness in the effectiveness of the inflation targeting. This study suggests that the mixed results are due to a lack of identification concerning domestic and global shocks and policy equation. Therefore, it contributes to the literature by studying the policy rule changes in a structural model with open economy properties. It will be shown that a small open economy structure is essential to facilitate the identification of structural

shocks that spillover from the rest of the world and the policy changes.

2.3 Model, Data and Methodology

2.3.1 Model

The methodology follows [Sims et al. \(2008\)](#), where a structural vector autoregression (SVAR) model is allowed to have state-dependent parameters with independent Markov chains. The parameters that undergo regime-switching are structural equation coefficients and shock variances.

Consider the following SVAR(p) model:

$$Y_t A_0(s_t) = \sum_{i=1}^p Y_{t-i} A_i(s_t) + \varepsilon_t \Gamma(s_t) \quad (2.1)$$

where Y_t is the observable vector. Observables include global production (y_t^*), global interest rate (i_t^*), exchange rate (e_t), world oil price (wop_t), climate conditions (cl_t), domestic GDP (y_t), policy rate (i_t) and CPI inflation (π_t). $A_0(\cdot)$ and $A_i(\cdot)$ represent contemporaneous and lagged relationships among variables. $\Gamma(\cdot)$ is the shock variance, p is lag length and s_t is state indicator.

$$Y_t = [wop_t, cl_t, y_t^*, i_t^*, \pi_t, y_t, i_t, e_t]. \quad (2.2)$$

Zero Identifying Restrictions

This study relies on the block exogeneity assumption proposed by [Cushman and Zha \(1997\)](#)—and extended in [Zha \(1999\)](#), [Dungey and Pagan \(2000\)](#) and [Kim and Roubini \(2000\)](#)—which excludes recursive relations between domestic and exogenous variables. By imposing the block exogeneity, one can model a system of equations in which the domestic variables are allowed to be affected by exogenous and international variables; however, the reversed relation is restricted to zero.

Considering the small open economy characteristics, this paper models New Zealand in four blocks: exogenous factors, exchange block, domestic and global economies. Whereas the domestic economy is affected by exogenous factors, exchange rate, and global blocks, it has no effect either on the global economy or on exogenous factors. Table 2.1 visualizes block exogeneity restrictions where a checkmark indicates that the row entry depends on the column entry.

Table 2.1: Block Exogeneity Restrictions

	Explanatory:			
	GE	EF	EB	DE
Global Economy (GE):	✓			
Exogenous Factors (EF):		✓		
Exchange Block (EB):	✓		✓	✓
Domestic Economy (DE):	✓	✓	✓	✓

Global block includes y_t^* , i_t^* and wop_t ; domestic block consists of y_t , i_t and π_t , exchange block is e_t . In the model, the only exogenous factor is the climate variable cl_t , which is assumed to follow a persistent autoregressive process following [Buckle et al. \(2007\)](#). Although the exchange rate is an international variable and evolves strongly depending on the international factors, in terms

of block exogeneity, e_t is modeled under an independent block since e_t is affected by y_t^* , i_t^* and wop_t without having any effect on them. International characteristics of exchange rate will be taken into account in regime-switching restrictions, which will be presented in the next section.

The identifying restrictions of the SVAR model follow the works of [Buckle et al. \(2007\)](#) with minor modifications. The compact form of the model equations are shown in equation set (2.3) where $i \in [1, 2, \dots, p]$ is the lag length:⁴

$$\begin{aligned}
y_t &= k_{11}cl_t + \sum_{i=1}^p k_{12}^i y_{t-i}^* + \sum_{i=1}^p k_{13}^i cl_{t-i} + \sum_{i=1}^p k_{14}^i wop_{t-i} + \sum_{i=1}^p k_{15}^i y_{t-i} \\
&\quad + \sum_{i=1}^p k_{16}^i \pi_{t-i} + \sum_{i=1}^p k_{17}^i e_{t-i} + \sum_{i=1}^p k_{18}^i i_{t-i} + \varepsilon_{t,1} \\
\pi_t &= k_{21}wop_t + k_{22}e_t + \sum_{i=1}^p k_{23}^i cl_{t-i} + \sum_{i=1}^p k_{24}^i wop_{t-i} + \sum_{i=1}^p k_{25}^i y_{t-i} \\
&\quad + \sum_{i=1}^p k_{26}^i \pi_{t-i} + \sum_{i=1}^p k_{27}^i e_{t-i} + \sum_{i=1}^p k_{28}^i i_{t-i} + \varepsilon_{t,2} \\
i_t &= k_{31}\pi_t + k_{32}i_t^* + \sum_{i=1}^p k_{33}^i \pi_{t-i} + \sum_{i=1}^p k_{34}^i i_{t-i}^* + \sum_{i=1}^p k_{35}^i y_{t-i} \\
&\quad + \sum_{i=1}^p k_{36}^i i_{t-i} + \sum_{i=1}^p k_{37}^i e_{t-i} + \varepsilon_{t,3} \\
y_t^* &= \sum_{i=1}^p k_{41}^i y_{t-i}^* + \sum_{i=1}^p k_{42}^i i_{t-i}^* + \sum_{i=1}^p k_{43}^i wop_{t-i} + \varepsilon_{t,4} \\
i_t^* &= k_{51}y_t^* + k_{52}wop_t + \sum_{i=1}^p k_{53}^i wop_{t-i} + \sum_{i=1}^p k_{54}^i i_{t-i}^* + \sum_{i=1}^p k_{55}^i y_{t-i}^* + \varepsilon_{t,5} \\
wop_t &= k_{61}y_t^* + \sum_{i=1}^p k_{62}^i wop_{t-i} + \sum_{i=1}^p k_{63}^i y_{t-i}^* + \varepsilon_{t,6} \\
cl_t &= \sum_{i=1}^p k_{71}^i cl_{t-i} + \varepsilon_{t,7} \\
e_t &= \sum_{i=1}^p k_{81}^i y_{t,t-i} + \sum_{i=1}^p k_{82}^i \pi_{t,t-i} + \sum_{i=1}^p k_{83}^i i_{t,t-i} + \sum_{i=1}^p k_{84}^i wop_{t,t-i} \\
&\quad + \sum_{i=1}^p k_{85}^i y_{t,t-i}^* + \sum_{i=1}^p k_{86}^i i_{t,t-i}^* + \sum_{i=1}^p k_{87}^i e_{t-i} + \varepsilon_{t,8}
\end{aligned} \tag{2.3}$$

⁴For the detailed description of the model and identifying restrictions, please check Appendix A

Table 2.2 illustrates the contemporaneous and lagged zero restrictions. Each row represents an equation in which a shaded area indicates a contemporaneous relationship between the dependent and independent variables. A "✓" indicates a relation between the dependent and lagged independent variables. The empty cells indicate zero restriction. This isolates the changes in the global block from the changes in the domestic block of the model; therefore, it helps to separate changes in the global shock dynamics from the policy for identification purposes.

Table 2.2: Contemporaneous and Lag Structure of the Model

	Explanatory Variables:							
	y^*	i^*	cl	wop	y	π	e	i
y^*	✓	✓		✓				
i^*	✓	✓		✓				
cl			✓					
wop	✓			✓				
y	✓		✓	✓	✓	✓	✓	✓
π			✓	✓	✓	✓	✓	✓
e	✓	✓		✓	✓	✓	✓	✓
i		✓			✓	✓	✓	✓

Even though the contemporaneous and lagged restrictions are appropriate for New Zealand, one needs to check whether the model is properly identified. [Rubio-Ramirez et al. \(2010\)](#) show the necessary and sufficient conditions for global identification. Following their theorem of global identification, this study examines if the model is properly identified under given contemporaneous and lagged zero restrictions.

Proposition 1. *The model is identified under the given contemporaneous and lagged zero restrictions.*

Proof. See Appendix A. □

Regime Switching Restrictions

To estimate and map the realized regimes into historical events, this study allows variations in the coefficients and model blocks' variances. This property of the model, along with block exogeneity, allows separating contributions of global shocks and domestic policy.

Recall that both coefficients of the model equations and covariance matrices of structural shocks are subject to regime-switching. For identification, only the coefficients of the monetary policy equation (for i_t) are subject to change across different regimes to model the adoption of inflation targeting. However, all shock variances are permitted to change across regimes, and this enables assessing the reduction in shock variances hypothesis through the international block of the model. For this purpose, this study modifies the set of equations (2.3) and imposes independent Markov chains s_t^i 's as follows:

$$\begin{aligned}
y_t &= k_{11}cl_t + \sum_{i=1}^p k_{12}^i y_{t-i}^* + \sum_{i=1}^p k_{13}^i cl_{t-i} + \sum_{i=1}^p k_{14}^i wop_{t-i} + \sum_{i=1}^p k_{15}^i y_{t-i} \\
&\quad + \sum_{i=1}^p k_{16}^i \pi_{t-i} + \sum_{i=1}^p k_{17}^i e_{t-i} + \sum_{i=1}^p k_{18}^i i_{t-i} + \varepsilon_{t,1}[\mathbf{s}_t^1] \\
\pi_t &= k_{21}wop_t + k_{22}e_t + \sum_{i=1}^p k_{23}^i cl_{t-i} + \sum_{i=1}^p k_{24}^i wop_{t-i} + \sum_{i=1}^p k_{25}^i y_{t-i} \\
&\quad + \sum_{i=1}^p k_{26}^i \pi_{t-i} + \sum_{i=1}^p k_{27}^i e_{t-i} + \sum_{i=1}^p k_{28}^i i_{t-i} + \varepsilon_{t,2}[\mathbf{s}_t^1] \\
i_t &= k_{31}[\mathbf{s}_t^2]\pi_t + k_{32}[\mathbf{s}_t^2]i_t^* + \sum_{i=1}^p k_{33}^i[\mathbf{s}_t^2]\pi_{t-i} + \sum_{i=1}^p k_{34}^i[\mathbf{s}_t^2]i_{t-i}^* + \sum_{i=1}^p k_{35}^i[\mathbf{s}_t^2]y_{t-i} \\
&\quad + \sum_{i=1}^p k_{36}^i[\mathbf{s}_t^2]i_{t-i} + \sum_{i=1}^p k_{37}^i[\mathbf{s}_t^2]e_{t-i} + \varepsilon_{t,3}[\mathbf{s}_t^1] \\
y_t^* &= \sum_{i=1}^p k_{41}^i y_{t-i}^* + \sum_{i=1}^p k_{42}^i i_{t-i}^* + \sum_{i=1}^p k_{43}^i wop_{t-i} + \varepsilon_{t,4}[\mathbf{s}_t^3] \\
i_t^* &= k_{51}y_t^* + k_{52}wop_t + \sum_{i=1}^p k_{53}^i wop_{t-i} + \sum_{i=1}^p k_{54}^i i_{t-i}^* + \sum_{i=1}^p k_{55}^i y_{t-i}^* + \varepsilon_{t,5}[\mathbf{s}_t^3] \\
wop_t &= k_{61}y_t^* + \sum_{i=1}^p k_{62}^i wop_{t-i} + \sum_{i=1}^p k_{63}^i y_{t-i}^* + \varepsilon_{t,6}[\mathbf{s}_t^4] \\
cl_t &= \sum_{i=1}^p k_{71}^i cl_{t-i} + \varepsilon_{t,7}[\mathbf{s}_t^4] \\
e_t &= \sum_{i=1}^p k_{81}^i y_{t,t-i} + \sum_{i=1}^p k_{82}^i \pi_{t,t-i} + \sum_{i=1}^p k_{83}^i i_{t,t-i} + \sum_{i=1}^p k_{84}^i wop_{t,t-i} \\
&\quad + \sum_{i=1}^p k_{85}^i y_{t,t-i}^* + \sum_{i=1}^p k_{86}^i i_{t,t-i}^* + \sum_{i=1}^p k_{87}^i e_{t-i} + \varepsilon_{t,8}[\mathbf{s}_t^3]
\end{aligned} \tag{2.4}$$

Here, s_t^1 is the state indicator for variances of domestic variables y_t , i_t and π_t . s_t^2 is the state indicator for the monetary policy coefficients, s_t^3 is the state indicator for variances to international variables y_t^* , i_t^* and e_t , and s_t^4 is the state indicator for variances to oil price and climate.⁵ The exchange rate is defined in an exogenous block; however, it is appropriate for small open economies to assume that shock variances of exchange rates depend on global

⁵As a robustness check, I also conduct the analysis by allowing for an independent Markov Chain for the exchange rate. The results are presented in Appendix A.

shock regimes. Similarly, the world oil price is defined in the global block in terms of shock transmission characteristics; however, it is assumed that shock characteristics of world oil price are independent of other global variables.

Even though it is possible that using an independent Markov chain for each variable is a better approach, this study incorporates regime-switching only into the monetary policy coefficients due to the computational burden.

2.3.2 Data

The sample period is from 1974:Q2 to 2018:Q3. The beginning of the sample is determined by data availability. 1974:Q2 is beginning point of the sample for which all variables become available. The US serves as a proxy for the rest of the world, with the exchange rate and other global variables also coming from the US.⁶ The international real production (y_t^*) is measured by US real GDP, and the world nominal interest rate (i_t^*) is federal funds rate. Following the global block, the exchange rate (e_t) is New Zealand dollar per US dollar. The climate conditions (cl_t) are measured by potential evapotranspiration deficit (PED), which is a soil moisture index that represents total rain volume, and the world oil price (wop_t) as Brent crude oil import prices in US dollars. For the domestic block, this paper uses New Zealand's real GDP (y_t), CPI inflation (π_t), and 90-day bank bill rate (i_t). All series are deseasonalized and detrended using a one-sided HP filter.⁷

⁶For the data descriptions, sources, and data transformations, see Appendix A

⁷One-sided HP filter is used to detrend data since it is backward-looking. It assumes that only current and past states affect the current observations at each observation, which is in line with the model structure and assumptions.

2.3.3 Estimation Methodology

This paper modifies the analytical approach suggested by [Sims et al. \(2008\)](#) to a small open economy SVAR to highlight the New Zealand monetary policy experiment.

Even though the equations in (2.4) are based on four independent Markov chains, the number of regimes for each independent chain is still to be determined. For this purpose, this study follows [Chib \(1998\)](#) and uses a formal marginal data density (MDD) comparison. Different models with different posterior parameter modes can be compared based on the expected value of marginal log-likelihood in a Bayesian framework. Suppose that models M_j have a different number of regimes for each block, and model parameters are estimated as θ_j . MDD, or marginal log-likelihood, is the likelihood of the observable data over all model parameters:

$$p(Y|M_j) = \int p(Y|\theta_1)p([\theta_1, \theta_2, \dots, \theta_j])d_1, d_2, \dots, d_j \quad (2.5)$$

The integral is calculated by the modified harmonic mean method as in [Sims et al. \(2008\)](#). Table 2.3 reports the marginal data densities of the different number of regime choices around this study's assumptions. Note that the MDD and model performances are not necessarily monotone with the number of regimes. This measure compares the fit of data across different model specifications.

Increasing the number of regimes for monetary policy response, exogenous and

Table 2.3: Marginal Data Density Comparison

Number of Regimes	Bridge Sampling MDD's
$\beta = [2,2,3,2]$	3687.10
$\beta^* = [2,2,3,3]$	3702.29
$\beta = [2,2,3,4]$	3694.42
$\beta = [2,2,3,5]$	3683.22
$\beta = [2,2,3,6]$	3689.09

international blocks did not improve the results, and for the higher number of regimes in variances, an overfitting problem might become an issue. Consequently, this study uses two regimes for the monetary policy coefficients, three regimes for the domestic block shock variances, two regimes for the global and exchange rate block shock variances, and three regimes for the exogenous block and oil price shock variances.

Table 2.4: Number of Lags

		Information Criteria	
		AIC	BIC
Number of Lags	1	-7106.4	-6908.8
	2*	-7318.4	-7015.6
	3	-7294.8	-6886.8
	4	-7278.0	-6764.8

Both AIC and BIC suggest the number of lags p used in the SVAR model (2.4) to be two. Table 2.4 shows that both Akaike and Schwarz information criteria suggest the same lag length.⁸ $\Lambda = [0.57, 0.13, 0.1, 1.5, 5, 5]$. The prior is of the Dirichlet form in order to satisfy that persistence of regimes are symmetric to each other, and regime-switching probabilities are not affected by the distance of the regimes. The detailed discussion of the prior and posterior densities can be found in [Sims et al. \(2008\)](#).

⁸Both the estimation and MDD comparison procedures are performed using Dynare version 4.5.4. For replication purposes, the routine maximizes posterior in a Bayesian setup, with the prior hyperparameters

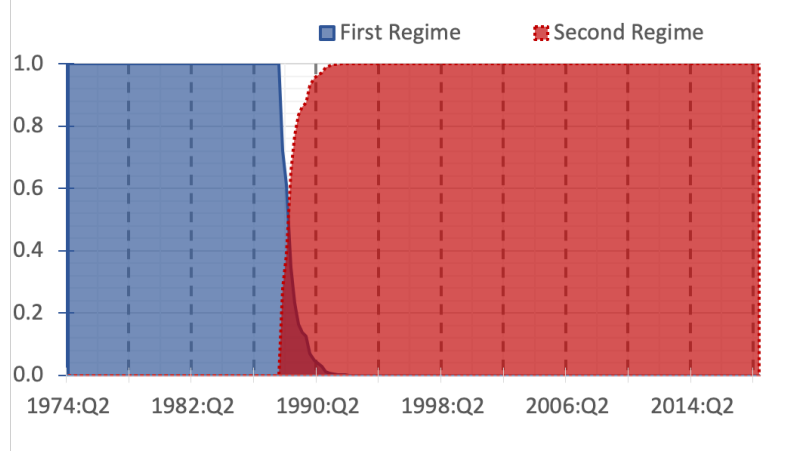


Figure 2.2: Monetary Policy Regimes

2.4 Results and Discussions

This section presents the main findings of the paper. Figure 2.2 shows the smoothed regime probabilities for monetary policy coefficients. The monetary policy equation starts to change in mid-1989, and the new regime dominates after the early 1990s.

The monetary policy response to an inflation shock, i.e., k_{31} in (2.4), increased from 10.92 to 14.82 while the model switched from the first regime to the second regime. The estimates of $A_0(s_t)$ are given in Table 2.5. The ratios of shaded cells give k_{31} coefficients for each regime. Although the values do not give any information on monetary policy's absolute tightness, they give information on relative tightness. In other words, the second regime responds aggressively to the changes in inflation compared to the first regime. Moreover, the regime switch's timing coincides with the timing of the adoption of inflation targeting in New Zealand, implying that the second regime represents the inflation-targeting monetary policy.

Table 2.5: $A_0[s_t^2]$

Contemporaneous coefficients								
	<i>wop</i>	<i>cl</i>	<i>y*</i>	<i>i*</i>	π	<i>y</i>	<i>i</i>	<i>e</i>
$A_0[1]:$	7.81	0.00	0.00	-0.52	-2.00	0.00	0.00	-0.54
	0.00	3.74	0.00	0.00	0.00	-0.15	0.00	0.00
	-6.82	0.00	106.31	-18.01	0.00	0.00	0.00	3.73
	0.00	0.00	0.00	78.63	0.00	0.00	3.96	-13.25
	0.00	0.00	0.00	0.00	54.48	0.00	-263.31	20.68
	0.00	0.00	0.00	0.00	0.00	104.07	0.00	0.32
	0.00	0.00	0.00	0.00	0.00	0.00	24.10	57.59
	0.00	0.00	0.00	0.00	-22.26	0.00	0.00	0.60
$A_0[2]:$	7.81	0.00	0.00	-0.52	-2.00	0.00	0.00	-0.54
	0.00	3.74	0.00	0.00	0.00	-0.15	0.00	0.00
	-6.82	0.00	106.31	-18.01	0.00	0.00	0.00	3.73
	0.00	0.00	0.00	78.63	0.00	0.00	36.19	-13.25
	0.00	0.00	0.00	0.00	54.48	0.00	-676.76	20.68
	0.00	0.00	0.00	0.00	0.00	104.07	0.00	0.32
	0.00	0.00	0.00	0.00	0.00	0.00	45.69	57.59
	0.00	0.00	0.00	0.00	-22.26	0.00	0.00	0.60

One important aspect of New Zealand's monetary policy conduct is the contract between the Finance Ministry and the Reserve Bank Governor, i.e., Policy Targets Agreements (PTAs). The contract, under the rubric of the RBNZ Act, has been revised several times. The revisions include the introduction of secondary policy objectives, changes to the target measure of inflation, and the time frame for inflation targeting. By design, the methodology in this paper is likely to pick up larger or more consequential changes as regimes, with these revisions contributing to estimating the features of the regimes they belong to. A detailed discussion on monetary policy governance in New Zealand can be found in [Buckle \(2020\)](#).

Figure 2.3 shows the impulse response functions (IRFs) of the monetary policy response to ten basis point inflation shocks under two regimes. Initially, the

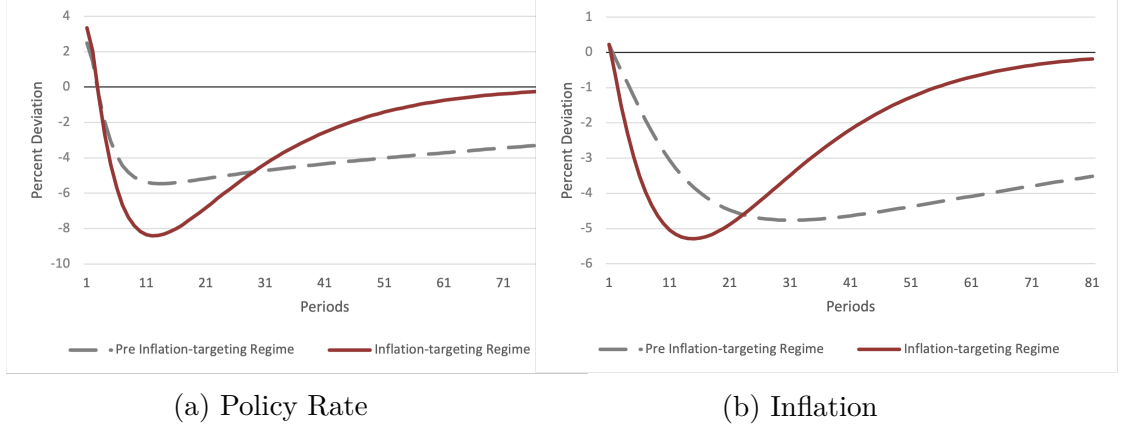


Figure 2.3: Impulse Response Functions to Inflation Shock (Change in percent points)

inflation targeting regime acts more aggressively, i.e., the initial response of the policy rate is larger. Afterward, in response to the large fall in inflation, the interest rate stays in the negative territory until the inflation goes back to its steady state. It is also clear that the pre-inflation-targeting regime takes a longer time to return to the steady state. Considering the timing of regime switch and IRFs together suggests that the adoption of inflation-targeting monetary policy has contributed to macroeconomic stabilization in New Zealand.

Table 2.6: Global Block

	Shock variances		
	y^*	i^*	e
$\Gamma[1]:$	4.39	0.00	0.00
	0.00	15.66	0.00
	0.00	0.00	9.53
$\Gamma[2]:$	1.00	0.00	0.00
	0.00	1.00	0.00
	0.00	0.00	1.00

Figure 2.4 and Table 2.6 provide smoothed transition probabilities and estimates for the global shock variances, respectively. For identification purposes,

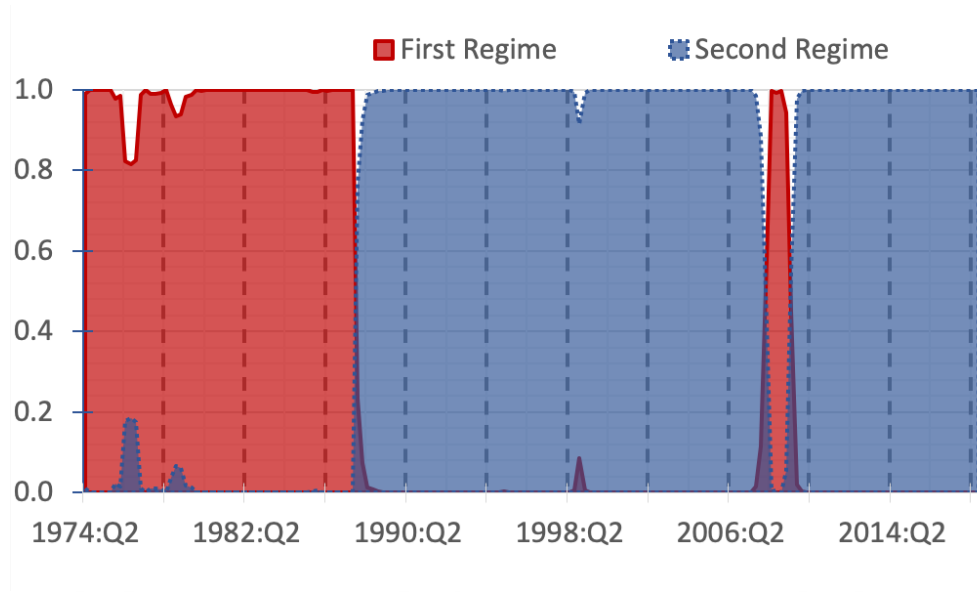


Figure 2.4: Global Shock Regimes

it is necessary to normalize one variance regime to unity ([Sims et al., 2008](#)). The second regime has lower global shock variance relative to the first regime. While the Great Moderation comes with a low variance in the global shocks, the Great Financial Contraction comes with a higher variance in global shocks.⁹

Figure 2.5 demonstrates the regime probabilities for the domestic shock variances whose estimates are given in Table 2.7. The first regime corresponds to a lower volatility period, and the second regime to a higher volatility regime. The third regime is interpreted as a transition period between the two.

The estimation results and smoothed transition probabilities highlight two main findings for the domestic block's shock variances. First, the decline in the global shock variances preceded the decline in the domestic shock variances, suggesting that New Zealand's success during the Great Moderation cannot be

⁹An interesting question to ponder on is whether the Great Recession was a simple interruption to the Great Moderation or the beginning of an entirely new regime in terms of international shock spillovers. This paper shows that the shock variances returned to those for the Great Moderation after the crisis ended. It would be possible to address this question more convincingly at a later time as more data become available.

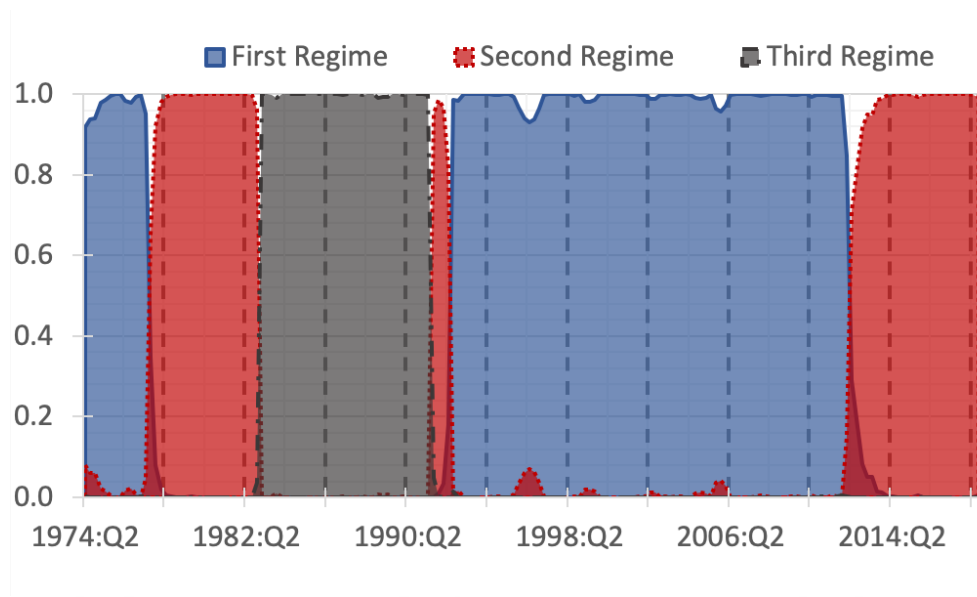


Figure 2.5: Domestic Shock Regimes

Table 2.7: Domestic Block

	Shock variances		
	π	y	i
$\Gamma[1]:$	0.70	0.00	0.00
	0.00	0.11	0.00
	0.00	0.00	0.33
$\Gamma[2]:$	2.05	0.00	0.00
	0.00	7.41	0.00
	0.00	0.00	4.70
$\Gamma[3]:$	1.00	0.00	0.00
	0.00	1.00	0.00
	0.00	0.00	1.00

attributed to inflation targeting alone. However, this does not invalidate the contribution of inflation targeting because the change in monetary policy coefficients also preceded the decline in the domestic shock variances, which are closer in terms of their timing. The evidence suggests that both reduction in global shock variances and adopting inflation-targeting monetary policy could have been important for macroeconomic stabilization during the Great Moderation because no single regime-change evidence dominates. Second, note

that commitment to inflation targeting was not affected by the Great Financial Contraction of 2008. As far as the New Zealand evidence is concerned, this was an episode in which the global shock variances increased first, and the domestic shock variances followed. The finding here corroborates that of [Fernandez-Villaverde et al. \(2010\)](#) for the US. They claim that, although the same policy stance holds, an increase in global shock variance can lead domestic conditions into a more volatile state. Therefore, the Global Financial Contraction influenced New Zealand's macroeconomic dynamics through a global increase in variances and then followed by a regime switch in the domestic shock variances.¹⁰

Whereas domestic shock variances had various regimes in the pre-inflation targeting period, global shocks were in the high variance regimes throughout this period. This suggests that a less aggressive policy together with a higher global shock variance led to a period of higher inflation.

2.5 Conclusion

I studied the macroeconomic dynamics of New Zealand to understand whether it was the reduction in shock variances or policy changes that have contributed to macroeconomic stabilization from the 1970s onward. A regime-switching SVAR model is estimated using the data from 1974 to 2018. The block exogeneity assumption due to the small open economy structure of New Zealand

¹⁰The results are robust to omitted variable bias according to a FAVAR setting. See Appendix A for robustness tests. An additional robustness check is provided by using trade-weighted international GDP, interest rate, and exchange rate instead of using the US as a proxy, as suggested by an anonymous referee. The results are provided in Appendix A.

makes it possible to separately identify the change in global shock variances and domestic policy.

The paper finds that both the adoption of inflation targeting and the reduction in global shock variances contributed to the macroeconomic stability in New Zealand during the Great Moderation. Moreover, the Great Financial Contraction is shown to have increased global shock variances without observing a change in monetary policy stance. Putting these evidences together, the paper concludes that macroeconomic stabilization was driven by both good policy and supportive global conditions.

Since Markov chains for domestic shock variances, global shock variances, and policy coefficients are defined to be independent, and the block exogeneity property of New Zealand restricts shock transmission from domestic block to global block, regime changes and transition probabilities of global shock variances and domestic policy can be identified. The contribution of global block shock variances and monetary policy regime changes to domestic block shock variances requires different identification strategies, which are left for future research.”

CHAPTER 3

THE ESTIMATION OF REGIME-SWITCHING FACTOR MODELS WITH MIXED-FREQUENCY AND MISSING DATA PROBLEMS USING THE EXPECTATION-MAXIMIZATION ALGORITHM

3.1 Introduction

The first chapter of this thesis examines the recent monetary policy and macroeconomic history of New Zealand under regime-switching. The main results claim that both the adoption of inflation targeting and the reduction in global shock variances contributed to the macroeconomic stability in New Zealand during the Great Moderation. An important criticism of the methodology used in the study, which has been addressed in Appendix A, is that the main results may be driven by a possible omitted variable bias.

The robustness test in Appendix A uses the factor-augmented vector autore-

gressive (FAVAR) model of [Bernanke et al. \(2005\)](#).¹ The idea is to include the factors obtained from other related indicators to the otherwise standard VAR. The early use of factor analysis by [Pearson \(1901\)](#), [Spearman \(1904\)](#), and [Hotelling \(1933\)](#) was driven by data dimension reduction purposes. Following the common component interpretation of [Geweke \(1977\)](#), the methodology integrated into the time series literature with factors as summary devices capturing the co-movement of multiple series. Among others, [Bernanke et al. \(2005\)](#) propose to test whether the factors obtained from additional data contribute to the analysis by comparing the impulse responses of the baseline VAR model and the factor-augmented model.

The methodology has demonstrated its usefulness in the business cycle analysis, in the tradition of [Burns and Mitchell \(1946\)](#). Recently, [Banbura et al. \(2010\)](#) show that the dynamic factor model (DFM) can be fruitfully used for the purpose of nowcasting where current economic activity, which is observed with delays, is estimated to aid policymaking in real-time. Moreover, the methodology can cope with challenging issues such as mixed frequency data and missing observations.

The aim of this chapter is to utilize large-dimensional datasets for the estimation of current economic activity based on a regime-switching DFM. Following the argument of [Hamilton \(1989\)](#), [Chauvet \(1998\)](#), and [Kim and Nelson \(1999\)](#) that using a multi-state factor structure can improve the assessment of economic activities, this chapter proposes an estimation methodology for

¹The robustness test proposed in Appendix A visualizes the impulse responses for the post inflation-targeting regime. The impulse responses are obtained using the Cholesky decomposition.

the DFM with regime-switching. The chapter is organized as follows. Section 3.2 summarizes the literature on DFMs and nowcasting, with the latter being the subject of the following chapters. This serves as the overall literature review for this and the following chapters of the thesis. Next, Section 3.3 first describes the estimation methodology for the single-regime DFM and its regime-switching modification. Then, the methodology is further modified to cope with additional challenges such as mixed frequency or missing data problems. Section 3.4 concludes.

3.2 Related Literature

GDP is arguably the most important macroeconomic indicator that researchers and policymakers closely follow as a measure of economic activity. Since the (preliminary) GDP is published with a delay, methodologies that investigate alternative ways of estimating the aggregate output have been widely studied in the literature. Among other approaches, exploiting information from higher frequency data via common components, i.e. factor model approaches, has become highly popular. Factor models are able to process a large number of data to estimate aggregate economic activity, circumventing the degrees of freedom issues in traditional empirical models.

[Pearson \(1901\)](#), [Spearman \(1904\)](#), and [Hotelling \(1933\)](#) introduced factor analysis as a data dimension reduction technique. [Geweke \(1977\)](#) and [Sargent et al. \(1977\)](#) provide early examples of factor models in economics by defining common components of time series data as coincident indicators of multiple time

series. [Forni et al. \(2000\)](#), [Stock and Watson \(2002a\)](#), and [Stock and Watson \(2002b\)](#) contribute to the literature by providing early examples of large-scale factor models in macroeconomic forecasting. [Giannone et al. \(2008\)](#) propose a systematic approach for using large-dimensional data to estimate the overall economic activity, which has become the platform for nowcasting at the New York FED (until its suspension during the COVID-19 pandemic; I will show that the regime-switching modification promises improved forecasting performance, alleviating the concern for the single-regime nowcasting model like the New York FED model.). [Banbura et al. \(2010\)](#) define nowcasting as the prediction of the present, the very near future, and the very recent past. Recently, using a similar systematic approach of [Giannone et al. \(2008\)](#), [Angelini et al. \(2011\)](#) estimate a short-term GDP forecast model for the euro area. [Barhoumi et al. \(2010\)](#), [Matheson \(2010\)](#), [Siliverstovs and Kholodilin \(2012\)](#), and [Modugno et al. \(2016\)](#) use similar nowcasting setups for other countries.

Models and techniques described above form an important part of the nowcasting literature. However, another essential reason for tracking economic activity is to determine the current state of the economy, i.e., the recession or the expansion. This strand of the literature is connected to the examination of turning points. For instance, the NBER business cycle dating committee announces ex-post turning points for the US economy. Since [Burns and Mitchell \(1946\)](#), many studies have focused on the determination of turning points of business cycles. [Harding and Pagan \(2002\)](#), [Stock and Watson \(2010\)](#), and [Stock and Watson \(2014\)](#) provide novel approaches to understand turning points. Using the factor models with asymmetries across recessions and expansions,

Hamilton (1989), Diebold and Rudebusch (1996), and Kim and Nelson (1999) propose to measure the economic state. In these models, regime changes occur in the intercept term of the underlying factors. Kholodilin et al. (2002b), and Bessec and Bouabdallah (2015) extend these models to a setting where the slope of the factors also are subject to regime-switching. Chauvet (1998) and Kholodilin et al. (2002a) study regime-switching on the volatility term of the factors. This chapter aims to contribute to the literature by proposing a regime-switching extension to the Bok et al. (2018)’s nowcasting setup. My main contribution is the synthesis of different ideas that have been explored in the literature, culminating to a toolkit that can be used to estimate large-scale regime-switching DFMs.² I will now discuss different estimation methods that have been considered by various researchers to explain how my contribution fits into the literature.

Because the factor analysis inherently operates through data reduction, it opens up the possibility of working with big datasets. Bai and Ng (2002), Stock and Watson (2002a), Stock and Watson (2002b), and Bernardo et al. (2003) are some of the early examples of the factor analysis with a large dimensional datasets. Even though increasing the size of the dataset has the benefit of working with additional information, it comes at the cost of increased computational burden as the number of parameters to be estimated also increases. For this reason, it is important to consider the trade-off between accuracy and speed in deciding the methodology to estimate factor models.

There are mainly two approaches for estimating factor models. The first is

²The toolkit will be available on a dedicated website in due time.

the frequentist approach based on the maximization of the likelihood. The second is the Bayesian approach which works with the posterior (of parameters) rather than the likelihood (where the posterior is proportional to the prior times the likelihood). However, some authors prefer to work with the principal component analysis instead, citing that both approaches are computationally burdensome. [Bai and Ng \(2002\)](#), [Stock and Watson \(2002a\)](#), and [Stock and Watson \(2002b\)](#) are some of the examples of this. As the computational power increases over time, this concern is likely to be less of a constraint.

For any estimation methodology, it is necessary to establish the consistency of the estimator. [Forni et al. \(2000\)](#), [Forni et al. \(2005\)](#), [Stock and Watson \(2002a\)](#), and [Stock and Watson \(2002b\)](#) study the consistency of the principal component estimators and show that they are the consistent estimators of the factors for large-scale cross-sectional data. [Forni et al. \(2009\)](#) and [Onatski \(2012\)](#) are some of well-known applications of the principal component methodology in the context of factor models.

The proponents of the maximum likelihood method argue that there are advantages to working with this over non-parametric methods such as the principal component analysis. For instance, the likelihood-based estimation of factor models has the advantage of deriving restrictions from economic theory. [Bernanke et al. \(2005\)](#), [Boivin and Giannoni \(2002\)](#), and [Reis and Watson \(2010\)](#) are some of the examples of this. Moreover, it is possible to estimate structural models (e.g., dynamic stochastic general equilibrium model, ?) using the maximum likelihood, which makes it a system based estimator.

Expectation-maximization (EM) is a popular iterative algorithm for maximizing the likelihood. An early contribution of this approach is [Dempster et al. \(1977\)](#). In the context of developments of nowcasting DFMs, [Doz et al. \(2011\)](#) propose a two-step approach to estimate the factor models with big datasets. This approach, in the first step, assumes that the principal component is the actual unobservable factors and estimate model parameters with OLS. In the second step, given the model parameters, the Kalman filter and smoother are used to estimate the factors. Building on this idea, [Bańbura and Modugno \(2014\)](#) and [Bok et al. \(2018\)](#) propose to use the EM where the two-steps above are repeated until convergence.

This chapter contributes to the literature by applying the EM approach to estimate the regime-switching version of the state-space model which can accommodate mixed frequency data as well as missing data problems. [Murphy \(1998\)](#) is a seminal application of the EM algorithm to Markov-switching linear state-space models. [Zhou and Shumway \(2008\)](#) and [Degras et al. \(2022\)](#) estimate regime-switching factor models where data frequency across variables is uniform.

[Hamilton \(1989\)](#) is the early influential work for the regime-switching literature. The work is based on VARs with an unobserved regime indicator. His work demonstrates how the EM approach can be used for estimating this class of models. Following this line of work, [Kim and Nelson \(1999\)](#) provide a one-step approach that combines the Kalman filter and the Hamiltonian filter to estimate the factors and the regime probabilities in a unified framework. The one-step approach proposed by [Kim and Nelson \(1999\)](#) numerically max-

imizes the approximate-likelihood (to be explained below) obtained from the Kalman filter. Since the one-step approach relies on numerical maximization, its computational cost becomes exorbitant as the model (parameter) dimension increases. To deal with this problem, [Diebold and Rudebusch \(1996\)](#), [Camacho et al. \(2015\)](#) and [Doz and Petronevich \(2016\)](#) propose a two-step approach, or a “short-cut” one-step approach. Here, the first step is to extract the factor by assuming that the model is linear, i.e., single-regime. Next, the model is estimated via the same methodology as in [Hamilton \(1989\)](#) since, given the factor, the dynamic factor model can be written as a VAR.³ [Doz and Petronevich \(2017\)](#) study the consistency of the two-step estimator and [Camacho et al. \(2018\)](#) apply the idea to a mixed-frequency nowcasting setup of [Bańbura and Modugno \(2014\)](#). For the Bayesian approach, see [Çakmakli et al. \(2021\)](#) who extend the mixed-frequency nowcasting model of [Bańbura and Modugno \(2014\)](#).

This chapter demonstrates how to apply the EM approach to a mixed-frequency regime-switching state-space model. In particular, I will focus on presenting the modified Kalman filter and smoother which are building blocks of the likelihood, the object to be maximized for maximum likelihood estimation. In the following chapter, I will demonstrate that this approach provides a good trade-off between accuracy and speed which makes it potentially useful for real-time applications of this methodology, especially for nowcasting involving large dimensional data.

³For the ease of reference, Appendix B provides a detailed mapping between DFMs and VARs. These materials build on [Stock and Watson \(2016\)](#) and [Doz and Fuleky \(2020\)](#). Refer to these papers for additional information.

3.3 Methodology

In this section, I explain the EM approach to estimate the regime-switching DFM in detail. In order to clarify the contribution of this chapter to the literature, I will first define the single-regime dynamic factor model and introduce the relevant notation. Next, I will briefly describe how the EM approach works in this setting. Then, I will provide the regime-switching extension to the baseline DFM and present the suitable EM approach for this modification. Finally, I will show how to incorporate mixed frequency data and missing observation problems into this setting.

3.3.1 Definition and Notation

This subsection introduces the DFM, building on important contributions of [Stock and Watson \(2016\)](#) and [Doz and Fuleky \(2020\)](#). Let Y_t be an $N \times 1$ vector of observables where each element, $y_{i,t}$, represents a different time series.

$$Y_t = \begin{bmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{N,t} \end{bmatrix}. \quad (3.1)$$

Define f_t as $q \times 1$ vector of unobservable factors. Let e_t be a “possibly serially correlated” innovations, and ε_t and v_t be normally distributed i.i.d shocks to the time series $t = 1, 2, \dots, T$. A dynamic form of the DFM is

$$Y_t = c_0 f_t + c_1 f_{t-1} + \dots + c_s f_{t-s} + e_t \quad (3.2)$$

$$f_t = a_1 f_{t-1} + a_2 f_{t-2} + \dots + a_h f_{t-h} + v_t \quad (3.3)$$

$$e_t = \delta_1 e_{t-1} + \delta_2 e_{t-2} + \dots + \delta_p e_{t-p} + \varepsilon_t \quad (3.4)$$

where equation (3.2) is called the observation or measurement equation which represents the observables as a linear combination of unobservable (common) factors, i.e., f_t . Here, the $N \times q$ coefficients c_i are called factor loadings which are essentially regression coefficients. Equation (3.3) is called the transition or state equation, which states that factors follow a VAR. This is what makes this factor model dynamic as today's factors depend on the lags of the factors. $q \times q$ coefficients a_i are vector autoregressive coefficients for f_t . $N \times N$ coefficients δ_i are persistence coefficients for innovations e_t when it follows an autoregressive process itself as in equation (3.4).

To convert the dynamic form of the DFM into a compact form, I define a vector with a lag length p :

$$F_t = \begin{bmatrix} f_t \\ f_{t-1} \\ \vdots \\ f_{t-p} \end{bmatrix}.$$

Then, (3.2)-(3.3) can be written as follows —which is called static, or stacked —representation:

$$Y_t = CF_t + e_t \quad (3.5)$$

$$F_t = AF_{t-1} + v_t \quad (3.6)$$

The definitions of exact and approximate DFM frameworks depend on the assumptions regarding to the innovation to the observation equation, i.e., e_t . [Stock and Watson \(2016\)](#) define the exact DFM as the model where e_t is uncorrelated cross-sectionally and serially. Equations (3.7) represent an exact static form DFM:

$$\begin{aligned} Y_t &= CF_t + e_t \\ F_t &= AF_{t-1} + v_t \\ e_t &\sim \mathcal{N}(0, R) \\ v_t &\sim \mathcal{N}(0, Q). \end{aligned} \quad (3.7)$$

However, the assumption of the exact DFM may be untenable for some applications. In this case, an approximate model which relaxes this assumption may be more suitable. Equations (3.8) represents an approximate static form DFM:

$$\begin{aligned}
Y_t &= CF_t + e_t \\
F_t &= AF_{t-1} + v_t \\
e_t &= \delta_1 e_{t-1} + \delta_2 e_{t-2} + \dots + \delta_p e_{t-p} + \varepsilon_t \\
\varepsilon_t &\sim \mathcal{N}(0, R) \\
v_t &\sim \mathcal{N}(0, Q).
\end{aligned} \tag{3.8}$$

This thesis is mainly concerned with an approximate DFM as represented in (3.8) given its focus on nowcasting and forecasting for which this is more relevant. However, it is often possible to recast this into an exact static form with an appropriate augmentation of state-space. For this reason, both representations will be invoked as I go along.

3.3.2 Estimation of Single-Regime DFMs

The DFM defined in equation (3.7), or similarly in equation (3.8), has two independent sets of parameters/latent factors to estimate. There are mainly two approaches for estimating factor models. The Bayesian approach treats the model parameters and unobservable factors as random variables. However, in the frequentist approach, the model parameters are fixed even though the factors are still treated as random variables. Before discussing the estimation of regime-switching DFMs with the EM approach, I will briefly discuss the single-regime model to introduce key concepts and motivate the regime-switching

counterpart.

Consider the static form DFM in (3.7). Discussing the computational challenges to estimate model parameters and unobservable factors, [Bai and Ng \(2002\)](#), [Stock and Watson \(2002a\)](#), and [Stock and Watson \(2002b\)](#) recommend using principal components to estimate factors. This approach is shown to give consistent estimators of factors for large datasets. However, the principal component estimator does not make use of economic theory in it being non-parametric. In order to make the factors more interpretable, a common practice is to use an orthogonal transformation of these as in [Gürkaynak et al. \(2005\)](#) where monetary policy surprises are separated into one referring to the current period ("target surprise") and another referring to beyond the current period ("path surprise").

The alternative estimation strategy is to use a parametric approach where it is possible to incorporate economic theory into structural estimation as in [Bernanke et al. \(2005\)](#), [Boivin and Giannoni \(2002\)](#), and [Reis and Watson \(2010\)](#). [Doz et al. \(2011\)](#) propose a two-step approach to estimate the dynamic factor models with large dimensional datasets as reviewed in Section 3.2. In what follows, I will use the iterative version of this which is referred to as the EM throughout the thesis.

Let P_t be the variance-covariance matrix of unobservable factor F_t , and $P = [P_1, \dots, P_t, \dots, P_T]$ and $F = [F_1, \dots, F_t, \dots, F_T]$ for time period $t = 1, 2, \dots, T$. And let $[P^k, F^k, C^k, R^k, A^k, Q^k]$ define the estimate at the k^{th} iteration given the initial values $[P^0, F^0, C^0, R^0, A^0, Q^0]$. The pseudo-code for the EM is

Algorithm 1: The Expectation-Maximization Approach

Initialization: Choose initial values for $F^0, P^0, C^0, R^0, A^0, Q^0$;

while *not Converged* **do**

 Expectation: Use the Kalman Filter and Smoother to estimate

F^{k+1}, P^{k+1} based on C^k, R^k, A^k, Q^k ;

 Maximization: Estimate $C^{k+1}, R^{k+1}, A^{k+1}, Q^{k+1}$ based on

F^{k+1}, P^{k+1} ;

end

The EM approach converges to the maximum likelihood estimate (MLE) independently of the initial values used under the linear Gaussian assumption. The EM approach can be summarized as follows. Let the expected values of the unobservable factors and their variance-covariance matrices for time period t' conditional on the information set at time t , i.e., Ω_t , be $E[F_{t'}|\Omega_t] = F_{t'|t}$ and $E[P_{t'}|\Omega_t] = P_{t'|t}$. Define N_{max} as the maximum number of iterations considered. Then, the mathematical expression for the EM algorithm is⁴

⁴I defer the discussion of these to Section 3.3.3 where these are compared to the regime-switching counterpart.

Algorithm 2: The Expectation-Maximization for Single-Regime

State-Space Model, see [Banbura et al. \(2010\)](#) for the detailed derivation.

The Expectation-Maximization Steps, $k \in [1, N_{max}]$:

The Expectation Step $t \in [1, T]$:

Kalman Filter:

Predict:

$$\begin{aligned} F_{t|t-1}^{k+1} &= C_t^k F_{t-1|t-1}^k \\ P_{t|t-1}^{k+1} &= C_t^k P_{t-1|t-1}^k (C_t^k)' + Q_t^k \end{aligned}$$

Update:

$$\begin{aligned} \eta_t^{k+1} &= Y_t - A_t^k F_{t|t-1}^{k+1} \\ S_t^{k+1} &= A_t^k P_{t|t-1}^{k+1} (A_t^k)' + R_t^k \\ K_t^{k+1} &= P_{t|t-1}^{k+1} (A_t^k)' S_t^{k+1-1} \\ F_{t|t}^{k+1} &= F_{t|t-1}^{k+1} + K_t^{k+1} \eta_t^{k+1} \\ P_{t|t}^{k+1} &= (I - K_t^{k+1} A_t^k) P_{t|t-1}^{k+1} \end{aligned}$$

Kalman Smoother:

$$\begin{aligned} J_{t+1|t}^{k+1} &= P_{t|t}^{k+1} C_t^{k+1} P_{t+1|t}^{k+1-1} \\ F_{t|T}^{k+1} &= F_{t|t}^{k+1} + J_{t+1|t}^{k+1} (F_{t+1|T}^{k+1} - C_t^{k+1} F_{t|t}^{k+1}) \\ P_{t|T}^{k+1} &= P_{t|t}^{k+1} + J_{t+1|t}^{k+1} (P_{t+1|T}^{k+1} - P_{t+1|t}^{k+1}) J_{t+1|t}^{k+1}{}' \\ P_{t,t-1|T}^{k+1} &= P_{t|T}^{k+1} J_{t+1|t}^{k+1}{}' \end{aligned}$$

The Maximization Step:

$$\begin{aligned} A^{k+1} &= \left(\sum_{t=1}^T P_{t,t-1|T}^{k+1} \right) \left(\sum_{t=1}^T P_{t-1|T}^{k+1} \right)^{-1} \\ C^{k+1} &= \left(\sum_{t=1}^T Y_t F_{t|T}^{k+1} \right) \left(\sum_{t=1}^T P_{t|T}^{k+1} \right)^{-1} \\ R^{k+1} &= \text{diag} \left(\frac{1}{T} \left(\sum_{t=1}^T Y_t Y_t' - C^{k+1} \sum_{t=1}^T F_{t|T}^{k+1} Y_t \right) \right) \\ Q^{k+1} &= \frac{1}{T} \left(\sum_{t=1}^T P_{t|T}^{k+1} - A^{k+1} \sum_{t=1}^T P_{t,t-1|T}^{k+1} \right) \end{aligned}$$

3.3.3 A Regime-Switching DFM

In this subsection, I present the EM algorithm for the DFM involving regime-switching, modifying the single-regime setting of the earlier Section 3.3.2 appropriately. The novelty here is the synthesis of different developments in the literature on regime-switching DFMs so far, providing a numerically tractable way to estimate these models when the data dimension is large.

Model Representations

[Banbura et al. \(2010\)](#) discuss DFMs with and without serial correlation in the observation equation errors. Here, I follow their presentation by discussing these two cases, but this time featuring regime-switching additionally. The regime-switching DFM without the serial correlation is

$$\begin{aligned} Y_t &= C[s_t]F_t + e[s_t] \\ F_t &= A[s_t]F_{t-1} + v[s_t] \\ e[s_t] &\sim \mathcal{N}(0, R[s_t]) \\ v[s_t] &\sim \mathcal{N}(0, Q[s_t]). \end{aligned} \tag{3.9}$$

On the other hand, the regime-switching DFM with the serial correlation in the observation equation error is

$$\begin{aligned}
Y_t &= C[s_t]F_t + e_t \\
F_t &= A[s_t]F_{t-1} + v[s_t] \\
e_t &= \delta_1[s_t]e_{t-1} + \delta_2[s_t]e_{t-2} + \dots + \delta_p[s_t]e_{t-p} + \varepsilon_t \\
\varepsilon[s_t] &\sim \mathcal{N}(0, R[s_t]) \\
v[s_t] &\sim \mathcal{N}(0, Q[s_t]).
\end{aligned} \tag{3.10}$$

Here, s_t is the regime indicator, which is discrete. The regime transition probabilities are assumed to follow a first order Markov process where, for M regimes, the exogenous regime-switching probability from the i^{th} regime, i.e., $s_{t-1} = i$, to the j^{th} regime, i.e., $s_t = j$ is denoted by $\pi_{i,j}$. These form the transition matrix Z . Throughout the chapter, I will use $Pr(.)$ to denote probabilities.

$$Pr(s_t = j | s_{t-1} = i) = \pi_{i,j} \tag{3.11}$$

$$Z = \begin{bmatrix} \pi_{1,1} & \pi_{1,2} & \pi_{1,3} & \dots & \pi_{1,M} \\ \pi_{2,1} & \pi_{2,2} & \pi_{2,3} & \dots & \pi_{2,M} \\ \pi_{3,1} & \pi_{3,2} & \pi_{3,3} & \dots & \pi_{3,M} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \pi_{M,1} & \pi_{M,2} & \pi_{M,3} & \dots & \pi_{M,M} \end{bmatrix} \tag{3.12}$$

The main difference between the regime-switching model given in equation (3.9) and its single-regime counterpart in equation (3.7), and similarly between equation (3.8) and equation (3.10), is that depending on regime realization model parameters change, allowing economic structures being modeled to go through potentially abrupt transitions. This permits the analysis that cannot be considered in a single-regime setting, for instance showing how a recession is different from a normal time in its shocks transmission. It is important to recall that s_t is a latent variable that needs to be estimated just like factors.

Estimation of the Regime-Switching DFM Framework

The Algorithms 1 and 2 in Section 3.3.2 need to be modified to take into account regime-switching. Let $s = [s_1, \dots, s_t, \dots, s_T]$ for time period $t = 1, 2, \dots, T$, s^k be the estimate of these at k^{th} iteration, and let s^0 be the initial values of the vector. The pseudo-code for the EM is

Algorithm 3: The Expectation-Maximization Approach for Regime-Switching Model

Initialization: Choose initial values for $F^0, P^0, s^0, C^0, R^0, A^0, Q^0$;

while *not Converged* **do**

 Expectation: Use the Kalman Filter and Smoother to estimate

$F^{k+1}, P^{k+1}, s^{k+1}$ based on C^k, R^k, A^k, Q^k ;

 Maximization: Estimate $C^{k+1}, R^{k+1}, A^{k+1}, Q^{k+1}$ based on

$F^{k+1}, P^{k+1}, s^{k+1}$;

end

In what follows, I will describe the EM algorithm in detail. The expectation step follows the filtering and smoothing steps in [Kim and Nelson \(1999\)](#). The maximization step builds on the maximization steps in [Murphy \(1998\)](#) and the single-regime application in [Banbura et al. \(2010\)](#).

The Expectation Step

Consider the regime-switching DFM in equation (3.9). Once a dynamic model is written in state-space representation, the Kalman Filter can be applied to estimate the state for a given set of parameters (C, R, A, Q) . Recall that in a single-regime setting, the Kalman filter is

$$F_{t|t-1} = A_t F_{t-1|t-1} \quad (3.13)$$

$$P_{t|t-1} = A_t P_{t-1|t-1} A_t' + Q \quad (3.14)$$

$$\eta_{t|t-1} = Y_t - C_t F_{t|t-1} \quad (3.15)$$

$$f_{t|t-1} = C_t P_{t|t-1} C_t' + R \quad (3.16)$$

$$F_{t|t} = F_{t|t-1} + P_{t|t-1} C_t' f_{t|t-1}^{-1} \eta_{t|t-1} \quad (3.17)$$

$$P_{t|t} = (I - P_{t|t-1} C_t' f_{t|t-1}^{-1} C_t) P_{t|t-1} \quad (3.18)$$

$\eta_{t|t-1}$ is one-step ahead prediction error for the observable⁵ and $f_{t|t-1}$ is the variance-covariance matrix of this. As usual, ' stands for matrix transpose. The equations (3.13)-(3.14) are called the *prediction* step and (3.15)-(3.18) are called the *updating* step. In the regime-switching setting, these steps are

⁵Note that this makes use of information at time t . The notation features conditioning on $t - 1$ to emphasize that the conditional expectation of the observable uses information up to $t - 1$.

performed at each point in time for each transition from regime $s_{t-1} = i$ to $s_t = j$.⁶ The following equations present this:

$$F_{t|t-1}^{(i,j)} = A_t^j F_{t-1|t-1}^i \quad (3.19)$$

$$P_{t|t-1}^{(i,j)} = A_t^j P_{t-1|t-1}^i A_t^{j'} + Q_j \quad (3.20)$$

$$\eta_{t|t-1}^{(i,j)} = Y_t - C_j F_{t|t-1}^{(i,j)} \quad (3.21)$$

$$f_{t|t-1}^{(i,j)} = C_j P_{t|t-1}^{(i,j)} C_j' + R_j \quad (3.22)$$

$$F_{t|t}^{(i,j)} = F_{t|t-1}^{(i,j)} + P_{t|t-1}^{(i,j)} C_j' [f_{t|t-1}^{(i,j)}]^{-1} \eta_{t|t-1}^{(i,j)} \quad (3.23)$$

$$P_{t|t}^{(i,j)} = (I - P_{t|t-1}^{(i,j)} C_j' [f_{t|t-1}^{(i,j)}]^{-1} C_j) P_{t|t-1}^{(i,j)} \quad (3.24)$$

where

$$F_{t|t-1}^{(i,j)} = E[F_t | \Omega_{t-1}, s_{t-1} = i, s_t = j]$$

$$P_{t|t-1}^{(i,j)} = E[(F_t - F_{t|t-1})(F_t - F_{t|t-1})' | \Omega_{t-1}, s_{t-1} = i, s_t = j]$$

$$F_{t|t}^{(i,j)} = E[F_t | \Omega_t, s_{t-1} = i, s_t = j]$$

$$P_{t|t}^{(i,j)} = E[(F_t - E[F_t | \Omega_t, s_t = j])(F_t - E[F_t | \Omega_t, s_t = j])' | \Omega_t, s_{t-1} = i, s_t = j]$$

$$F_{t-1|t-1}^i = E[F_t | \Omega_{t-1}, s_{t-1} = i]$$

$$P_{t-1|t-1}^i = E[(F_{t-1} - F_{t-1|t-1}^i)(F_{t-1} - F_{t-1|t-1}^i)' | \Omega_{t-1}, s_{t-1} = i]$$

and Ω_t contains observables up to time t .

⁶While conceptually well-defined, this may not be computationally feasible. More on this below.

However, the curse of dimensionality does not allow for all possible paths of regimes s_t to be considered. For this reason, all algorithms used in the literature involve approximation of this. In order to overcome this issue, [Kim and Nelson \(1999\)](#) suggest the following:

$$\begin{aligned} F_{t|t}^j &= \sum_{i=1}^M \frac{Pr(s_t = j, s_{t-1} = i | \Omega_t)}{Pr(s_t = j | \Omega_t)} F_{t|t}^{(i,j)} \\ &= \sum_{i=1}^M Pr(s_{t-1} = i | s_t = j, \Omega_t) F_{t|t}^{(i,j)} \end{aligned} \quad (3.25)$$

$$\begin{aligned} P_{t|t}^j &= \sum_{i=1}^M \frac{Pr(s_t = j, s_{t-1} = i | \Omega_t)}{Pr(s_t = j | \Omega_t)} [P_{t|t}^{(i,j)} + (F_{t|t}^j - F_{t|t}^{(i,j)})(F_{t|t}^j - F_{t|t}^{(i,j)})'] \\ &= \sum_{i=1}^M Pr(s_{t-1} = i | s_t = j, \Omega_t) [P_{t|t}^{(i,j)} + (F_{t|t}^j - F_{t|t}^{(i,j)})(F_{t|t}^j - F_{t|t}^{(i,j)})'] \end{aligned} \quad (3.26)$$

Recall that M is the number of regimes. $Pr(s_t = j, s_{t-1} = i | \Omega_t)$ is transition probability from $s_{t-1} = i$ to $s_t = j$ conditional on the information set at time t . Similarly, $Pr(s_t = j | \Omega_t)$ is the probability that the current regime is j conditional on the information set at time t . These probabilities are calculated using the following equations:

$$\begin{aligned} Pr(s_t = j, s_{t-1} = i | \Omega_{t-1}) &= Pr(s_t = j | s_{t-1} = i, \Omega_{t-1}) Pr(s_{t-1} = i | \Omega_{t-1}) \\ &= Pr(s_t = j | s_{t-1} = i) Pr(s_{t-1} = i | \Omega_{t-1}) \end{aligned} \quad (3.27)$$

The second line of equation (3.27) follows because of conditional on s_{t-1}, Ω_{t-1} does not contain any additional information.

Also recall that the density function of a multivariate normal random variable

τ with the mean $\bar{\mu}$, the variance-covariance Σ , and the dimension d is

$$\mathcal{N}(\tau; \bar{\mu}, \Sigma) = \frac{1}{2\pi^{\frac{d}{2}}} |\Sigma|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\tau - \bar{\mu})' \Sigma^{-1} (\tau - \bar{\mu})\right\}$$

where $|\cdot|$ is the determinant of a square matrix. Then, [Kim and Nelson \(1999\)](#) define the conditional likelihood $Pr(Y_t|s_t = j, s_{t-1} = i, \Omega_{t-1})$ in terms of the conditional one-step ahead prediction error $\eta_{t|t-1}^{(i,j)}$, and the variance-covariance matrix $f_{t|t-1}^{(i,j)}$ as,

$$Pr(Y_t|s_t = j, s_{t-1} = i, \Omega_{t-1}) = \frac{1}{2\pi^{\frac{N}{2}}} |f_{t|t-1}^{(i,j)}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} \eta_{t|t-1}^{(i,j)'} f_{t|t-1}^{(i,j)-1} \eta_{t|t-1}^{(i,j)}\right\}. \quad (3.28)$$

Then,

$$Pr(Y_t|\Omega_{t-1}) = \sum_{j=1}^M \sum_{i=1}^M \left(Pr(Y_t|s_t = j, s_{t-1} = i, \Omega_{t-1}) \times Pr(s_t = j, s_{t-1} = i|\Omega_{t-1}) \right) \quad (3.29)$$

$$Pr(s_t = j, s_{t-1} = i|\Omega_t) = \left(Pr(Y_t|s_t = j, s_{t-1} = i, \Omega_{t-1}) \times \frac{Pr(s_t = j, s_{t-1} = i|\Omega_{t-1})}{Pr(Y_t|\Omega_{t-1})} \right) \quad (3.30)$$

$$Pr(s_t = i|\Omega_t) = \sum_{j=1}^M Pr(s_t = j, s_{t-1} = i|\Omega_t) \quad (3.31)$$

where the equation (3.31) makes use of the fact that Ω_t is Ω_{t-1} augmented with the time period t observables Y_t . The equations (3.19)-(3.24), (3.27)-(3.31), and (3.25)-(3.26) complete the filter from [Kim and Nelson \(2006\)](#).

In the smoothing step, I calculate the expected values of the latent factors and its variance-covariance using the whole sample information. This exercise

by definition is retrospective, inferring the state at a particular point in time using information beyond the period. And it makes use of the output from the filtering step as the input. Using the smoothing steps of [Kim and Nelson \(1999\)](#) for $t = (T - 1, T - 2, \dots, 1)$ (Here, time goes backwards starting from the end of the sample), I obtain⁷

$$J_{t+1|t}^{(j,k)} = P_{t|t}^j A_t^{k'} [P_{t+1|t}^{(j,k)}]^{-1} \quad (3.32)$$

$$F_{t|T}^{(j,k)} = F_{t|t}^j + J_{t+1|t}^{(j,k)} (F_{t+1|T}^k - A_t^k F_{t|t}^j) \quad (3.33)$$

$$P_{t|T}^{(j,k)} = P_{t|t}^j + J_{t+1|t}^{(j,k)} (P_{t+1|T}^k - P_{t+1|t}^{(j,k)}) J_{t+1|t}^{(j,k)'} \quad (3.34)$$

where

$$P_{t+1|t}^{(j,k)} = E[(F_{t+1} - F_{t+1|t})(F_{t+1} - F_{t+1|t})' | \Omega_t, s_t = j, s_{t+1} = k]$$

$$F_{t+1|T}^k = E[F_{t+1} | \Omega_T, s_{t+1} = k]$$

$$P_{t+1|T}^k = E[(F_{t+1} - F_{t+1|t}^k)(F_{t+1} - F_{t+1|t}^k)' | \Omega_T, s_{t+1} = k]$$

Similar to the regime-switching Kalman Filter, [Kim and Nelson \(1999\)](#) suggest

to approximate $F_{t|T}^{(j,k)}$ and $P_{t|T}^{(j,k)}$ to overcome the curse of dimensionality. Let

⁷The Kalman smoother without regime-switching is provided in Algorithm-2 above.

$$\begin{aligned}
F_{t|T}^j &= \sum_{k=1}^M \frac{Pr(s_t = j, s_{t+1} = k | \Omega_T)}{Pr(s_t = j | \Omega_T)} F_{t|T}^{(j,k)} \\
&= \sum_{k=1}^M Pr(s_{t+1} = k | s_t = j, \Omega_T) F_{t|T}^{(j,k)} \tag{3.35}
\end{aligned}$$

$$\begin{aligned}
P_{t|T}^j &= \sum_{k=1}^M \frac{Pr(s_t = j, s_{t+1} = k | \Omega_T)}{Pr(s_t = j | \Omega_T)} [P_{t|T}^{(j,k)} + (F_{t|T}^j - F_{t|T}^{(j,k)})(F_{t|T}^j - F_{t|T}^{(j,k)})'] \\
&= \sum_{k=1}^M Pr(s_{t+1} = k | s_t = j, \Omega_T) [P_{t|T}^{(j,k)} + (F_{t|T}^j - F_{t|T}^{(j,k)})(F_{t|T}^j - F_{t|T}^{(j,k)})'] \tag{3.36}
\end{aligned}$$

where

$$Pr(s_t = j, s_{t+1} = k | \Omega_T) = \frac{Pr(s_{t+1} = k | \Omega_T) Pr(s_{t+1} = k | s_t = j) Pr(s_t = j | \Omega_t)}{Pr(s_{t+1} = k | \Omega_t)} \tag{3.37}$$

$$Pr(s_t = j | \Omega_T) = \sum_{k=1}^M Pr(s_t = j, s_{t+1} = k | \Omega_T) \tag{3.38}$$

The equations (3.32)-(3.34), (3.37)-(3.38), and (3.35)-(3.36) complete the smoother of [Kim and Nelson \(2006\)](#).

Consider the general case of the model in (3.9) with the intercept term on the factor as follows:⁸

⁸Note that the equation for F_t now features the potentially regime-switching intercept term $\mu[s_t]$. This is for the expositional convenience. This equation can always be rewritten in the form of equation (3.9) where the vector of factors is augmented with the constant of one. Henceforth, both notations will be used without loss of generality.

$$\begin{aligned}
Y_t &= C[s_t]F_t + e[s_t] \\
F_t &= \mu[s_t] + A[s_t]F_{t-1} + v[s_t] \\
e[s_t] &\sim \mathcal{N}(0, R[s_t]) \\
v[s_t] &\sim \mathcal{N}(0, Q[s_t]).
\end{aligned} \tag{3.39}$$

In this section, I will provide a maximization solution to the model in (3.39).

Some new notations. For any parameter X , $X[s_t = j] = X_j$. Given the model parameter set of $\theta = \{(A_j, Q_j, C_j, R_j, \mu_j) : 1 < j < M; \pi_{i,j} \in Z\}$, let

- $E_\theta[\cdot | \Omega_T]$ be the expectations operator for the parameters $\hat{\theta}$ conditional on complete observable data.
- $\log L_c(\theta)$ be the log-likelihood function conditional on parameters θ .
- $Pr_\theta(\cdot)$ be the conditional probability $Pr(\cdot)$ on the parameters θ .

Recall that the density function for a multivariate random variable τ with mean $\bar{\mu}$, variance Σ and dimension d is

$$\mathcal{N}(\tau; \bar{\mu}, \Sigma) = \frac{1}{2\pi^{\frac{d}{2}}} |\Sigma|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\tau - \bar{\mu})' \Sigma^{-1} (\tau - \bar{\mu})\right\}. \tag{3.40}$$

Let the term $\mathbb{I}$ be the state indicator for $s_t = j$, i.e., $\mathbb{I}[s_t = j] = 1$. Following [Murphy \(1998\)](#) and [Shumway et al. \(2016\)](#), the log-likelihood function can be

written as

$$\begin{aligned}
\log L_c(\theta) &= -\frac{T(N+r)}{2} \log(2\pi) \\
&- \frac{1}{2} \sum_{t=1}^T \sum_{j=1}^M \mathbb{I}[s_t = j] \times \left[\log|R_j| + (Y_t - C_j F_t^j)' R_j^{-1} (Y_t - C_j F_t^j) \right. \\
&+ \log|Q_j| \\
&+ \left. \left((F_t^j - \mu_j - A_j F_{t-1}^j)' Q_j^{-1} (F_t^j - \mu_j - A_j F_{t-1}^j) \right) \right. \\
&+ \left. \sum_{i=1}^M \mathbb{I}[s_{t-1} = i] \times \left(\log \text{Pr}_\theta(s_t = j | s_{t-1} = i) \right) \right]
\end{aligned}$$

where $F_t^j = E[F_t | s_t = j]$. Note that $\mathbb{I}$ here is trivial in it being always equal to one. However, this representation allows the smoothed regime probabilities to be introduced when the expected value of $\log L_c(\theta)$ is evaluated below.

Next, I evaluate the expected value of $\log L_c(\theta)$ at the parameter value $\hat{\theta}$ which is obtained from solving

$$\hat{\theta} = \arg \max_{\theta} E_{\theta}[\log L_c(\theta) | \Omega_T] \quad (3.41)$$

where the expected value of the likelihood on the right hand side is based on the outcome from the previous iteration (the previous maximization step).

The resulting expression is

$$\begin{aligned}
Q &= E_{\hat{\theta}}[\log L_c(\theta)|\Omega_T] = -E_{\hat{\theta}}\left[\frac{T(N+r)}{2}\log(2\pi)\middle|\Omega_T\right] \\
&- \frac{1}{2}\sum_{t=1}^T\sum_{j=1}^M\left[\overbrace{E_{\hat{\theta}}\left[\mathbb{I}[s_t=j]\times(\log|R[s_t=j]|)\middle|\Omega_T\right]}^I\right. \\
&+ \overbrace{E_{\hat{\theta}}\left[\mathbb{I}[s_t=j]\times\left((Y_t-C[s_t=j]F_t^j)'R[s_t=j]^{-1}(Y_t-C[s_t=j]F_t^j)\right)\middle|\Omega_T\right]}^{II} \\
&+ \overbrace{E_{\hat{\theta}}\left[\mathbb{I}[s_t=j]\times(\log|Q[s_t=j]|)\middle|\Omega_T\right]}^{III} \\
&\left. + E_{\hat{\theta}}\left[\mathbb{I}[s_t=j]\times\left((F_t^j-\mu[s_t=j]-A[s_t=j]F_{t-1}^j)'\right.\right. \right. \\
&\quad \left.\left.\times Q[s_t=j]^{-1}(F_t^j-\mu[s_t=j]-A[s_t=j]F_{t-1}^j)\right)\middle|\Omega_T\right] \right\}^{IV} \\
&+ \sum_{i=1}^M\overbrace{E_{\hat{\theta}}\left[\mathbb{I}[s_t=j]\times\mathbb{I}[s_{t-1}=i]\times\left(\log Pr_{\hat{\theta}}(s_t=j|s_{t-1}=i)\right)\middle|\Omega_T\right]}^V\left. \right]
\end{aligned} \tag{3.42}$$

where the Roman numerals I to V refer to different parts of the expected likelihood Q .

Recall that since $\mathbb{I}[s_t=j]$ and $(Y_t-C_jF_t^j)'R_j^{-1}(Y_t-C_jF_t^j)$, and $\mathbb{I}[s_t=j]$ and $(F_t^j-\mu_j-A_jF_{t-1}^j)'Q_j^{-1}(F_t^j-\mu_j-A_jF_{t-1}^j)$ are mutually independent due to the exogeneity of the regime indicator, I have

$$\begin{aligned}
&E_{\hat{\theta}}[\mathbb{I}[s_t=j]\times\left((Y_t-C_jF_t^j)'R_j^{-1}(Y_t-C_jF_t^j)\right)|\Omega_T] \\
&= E_{\hat{\theta}}\left[\left((Y_t-C_jF_t^j)'R_j^{-1}(Y_t-C_jF_t^j)\right)|\Omega_T\right]\times E_{\hat{\theta}}[\mathbb{I}[s_t=j]|\Omega_T] \quad , \text{ and} \\
&E_{\hat{\theta}}[\mathbb{I}[s_t=j]\times\left((F_t^j-\mu_j-A_jF_{t-1}^j)'Q_j^{-1}(F_t^j-\mu_j-A_jF_{t-1}^j)\right)|\Omega_T] \\
&= E_{\hat{\theta}}\left[\left((F_t^j-\mu_j-A_jF_{t-1}^j)'Q_j^{-1}(F_t^j-\mu_j-A_jF_{t-1}^j)\right)|\Omega_T\right]\times E_{\hat{\theta}}[\mathbb{I}[s_t=j]|\Omega_T].
\end{aligned} \tag{3.43}$$

By definition,

$$\begin{aligned}
E_{\hat{\theta}}[\mathbb{I}[s_t = j]|\Omega_T] &= Pr_{\hat{\theta}}(s_t = j|\Omega_T) \\
E_{\hat{\theta}}[\mathbb{I}[s_{t-1} = i]\mathbb{I}[s_t = j]|\Omega_T] &= Pr_{\hat{\theta}}(s_t = j, s_{t-1} = i|\Omega_T)
\end{aligned} \tag{3.44}$$

For diagonal matrices Q_j and R_j , one can rewrite the quadratic forms above as

$$\begin{aligned}
& \left((Y_t - C_j F_t^j)' R_j^{-1} (Y_t - C_j F_t^j) \right) \\
&= tr \left((Y_t - C_j F_t^j)' R_j^{-1} (Y_t - C_j F_t^j) \right) \\
&= tr \left(R_j^{-1} (Y_t - C_j F_t^j) (Y_t - C_j F_t^j)' \right), \text{ and} \\
& \left((F_t^j - \mu_j - A_j F_{t-1}^j)' Q_j^{-1} (F_t^j - \mu_j - A_j F_{t-1}^j) \right) \\
&= tr \left((F_t^j - \mu_j - A_j F_{t-1}^j)' Q_j^{-1} (F_t^j - \mu_j - A_j F_{t-1}^j) \right) \\
&= tr \left(Q_j^{-1} (F_t^j - \mu_j - A_j F_{t-1}^j) (F_t^j - \mu_j - A_j F_{t-1}^j)' \right)
\end{aligned} \tag{3.45}$$

Next, I substitute the objects from the smoother into the Q function. To make this more transparent, I do this for each partition corresponding the Roman numerals I to V . With a slight abuse of notation, when the matrices are evaluated at parameter values $\hat{\theta}$, this is not explicitly stated to prevent the equations below from being over-cluttered.

For the first block I (again invoking the exogeneity of s_t),

$$\begin{aligned}
I &= E_{\hat{\theta}} \left[\mathbb{I}[s_t = j] \times (\log|R_j|) \middle| \Omega_T \right] \\
&= E_{\hat{\theta}} \left[(\log|R_j|) \middle| \Omega_T \right] \times E_{\hat{\theta}} \left[\mathbb{I}[s_t = j] \middle| \Omega_T \right] \\
&= \log|R_j| Pr_{\hat{\theta}}(s_t = j | \Omega_T)
\end{aligned} \tag{3.46}$$

For the second block II , substituting equations (3.43), (3.44) and (3.45) into this I obtain

$$\begin{aligned}
II &= E_{\hat{\theta}} \left[\mathbb{I}[s_t = j] \times \left((Y_t - C_j F_t^j)' R_j^{-1} (Y_t - C_j F_t^j) \right) \middle| \Omega_T \right] \\
&= E_{\hat{\theta}} \left[(Y_t - C_j F_t^j)' R_j^{-1} (Y_t - C_j F_t^j) \middle| \Omega_T \right] \times E_{\hat{\theta}} [\mathbb{I}[s_t = j] | \Omega_T] \\
&= E_{\hat{\theta}} \left[tr \left(R_j^{-1} (Y_t - C_j F_t^j) (Y_t - C_j F_t^j)' \right) \middle| \Omega_T \right] \times E_{\hat{\theta}} [\mathbb{I}[s_t = j] | \Omega_T] \\
&= E_{\hat{\theta}} \left[tr \left(R_j^{-1} (Y_t - C_j F_t^j) (Y_t - C_j F_t^j)' \right) \middle| \Omega_T \right] \times Pr_{\hat{\theta}}(s_t = j | \Omega_T) \\
&= tr \left(R_j^{-1} E_{\hat{\theta}} \left[(Y_t - C_j F_t^j) (Y_t - C_j F_t^j)' \middle| \Omega_T \right] \right) \times Pr_{\hat{\theta}}(s_t = j | \Omega_T) \\
&= tr \left(R_j^{-1} \times \left(E_{\hat{\theta}} [Y_t Y_t' | \Omega_T] - E_{\hat{\theta}} [Y_t (F_t^j)' C_j' | \Omega_T] \right. \right. \\
&\quad \left. \left. - E_{\hat{\theta}} [C_j F_t^j Y_t' | \Omega_T] + E_{\hat{\theta}} [C_j F_t^j (F_t^j)' C_j' | \Omega_T] \right) \right) \times Pr_{\hat{\theta}}(s_t = j | \Omega_T) \\
&= tr \left(R_j^{-1} \times \left(Y_t Y_t' - Y_t E_{\hat{\theta}} [(F_t^j)' | \Omega_T] C_j' \right. \right. \\
&\quad \left. \left. - C_j E_{\hat{\theta}} [F_t^j | \Omega_T] Y_t' + C_j E_{\hat{\theta}} [F_t^j (F_t^j)' | \Omega_T] C_j' \right) \right) \times Pr_{\hat{\theta}}(s_t = j | \Omega_T)
\end{aligned} \tag{3.47}$$

For the third block *III* (again invoking the exogeneity of s_t),

$$\begin{aligned}
III &= E_{\hat{\theta}} \left[\mathbb{I}[s_t = j] \times (\log|Q[s_t] = j|) \middle| \Omega_T \right] \\
&= E_{\hat{\theta}} \left[(\log|Q_j|) \middle| \Omega_T \right] \times E_{\hat{\theta}} \left[\mathbb{I}[s_t = j] \middle| \Omega_T \right] \\
&= \log|Q_j| Pr_{\hat{\theta}}(s_t = j | \Omega_T)
\end{aligned} \tag{3.48}$$

For the fourth block *IV*, substituting equations (3.43), (3.44) and (3.45) into this I obtain



$$\begin{aligned}
IV &= E_{\hat{\theta}} \left[\mathbb{I}[s_t = j] \times \left((F_t^j - \mu_j - A_j F_{t-1}^j)' Q_j^{-1} (F_t^j - \mu_j - A_j F_{t-1}^j) \right) \middle| \Omega_T \right] \\
&= E_{\hat{\theta}} \left[(F_t^j - \mu_j - A_j F_{t-1}^j)' Q_j^{-1} (F_t^j - \mu_j - A_j F_{t-1}^j) \middle| \Omega_T \right] \\
&\times E_{\hat{\theta}} \left[\mathbb{I}[s_t = j] \middle| \Omega_T \right] \\
&= E_{\hat{\theta}} \left[\text{tr} \left((F_t^j - \mu_j - A_j F_{t-1}^j)' Q_j^{-1} (F_t^j - \mu_j - A_j F_{t-1}^j) \right) \middle| \Omega_T \right] \\
&\times E_{\hat{\theta}} \left[\mathbb{I}[s_t = j] \middle| \Omega_T \right] \\
&= E_{\hat{\theta}} \left[\text{tr} \left((F_t^j - \mu_j - A_j F_{t-1}^j)' Q_j^{-1} (F_t^j - \mu_j - A_j F_{t-1}^j) \right) \middle| \Omega_T \right] \\
&\times \text{Pr}_{\hat{\theta}}(s_t = j | \Omega_T) \\
&= \text{tr} \left(Q_j^{-1} E_{\hat{\theta}} \left[(F_t^j - \mu_j - A_j F_{t-1}^j) (F_t^j - \mu_j - A_j F_{t-1}^j)' \middle| \Omega_T \right] \right) \\
&\times \text{Pr}_{\hat{\theta}}(s_t = j | \Omega_T) \\
&= \text{tr} \left(Q_j^{-1} \times \left(E_{\hat{\theta}}[F_t^j (F_t^j)' | \Omega_T] - E_{\hat{\theta}}[F_t^j \mu_j' | \Omega_T] - E_{\hat{\theta}}[F_t^j (F_{t-1}^j)' A_j' | \Omega_T] \right. \right. \\
&\quad - E_{\hat{\theta}}[\mu_j (F_t^j)' | \Omega_T] + E_{\hat{\theta}}[\mu_j \mu_j' | \Omega_T] + E_{\hat{\theta}}[\mu_j (F_{t-1}^j)' A_j' | \Omega_T] \\
&\quad - E_{\hat{\theta}}[A_j F_{t-1}^j (F_t^j)' | \Omega_T] + E_{\hat{\theta}}[A_j F_{t-1}^j \mu_j' | \Omega_T] \\
&\quad \left. \left. + E_{\hat{\theta}}[A_j F_{t-1}^j (F_{t-1}^j)' A_j' | \Omega_T] \right) \times \text{Pr}_{\hat{\theta}}(s_t = j | \Omega_T) \right) \\
&= \text{tr} \left(Q_j^{-1} \times \left(E_{\hat{\theta}}[F_t^j (F_t^j)' | \Omega_T] - E_{\hat{\theta}}[F_t^j | \Omega_T] \mu_j' - E_{\hat{\theta}}[F_t^j (F_{t-1}^j)' | \Omega_T] A_j' \right. \right. \\
&\quad - \mu_j E_{\hat{\theta}}[(F_t^j)' | \Omega_T] + \mu_j \mu_j' + \mu_j E_{\hat{\theta}}[(F_{t-1}^j)' | \Omega_T] A_j' \\
&\quad - A_j E_{\hat{\theta}}[F_{t-1}^j (F_t^j)' | \Omega_T] + A_j E_{\hat{\theta}}[F_{t-1}^j | \Omega_T] \mu_j' \\
&\quad \left. \left. + A_j E_{\hat{\theta}}[F_{t-1}^j (F_{t-1}^j)' | Y_1] A_j' \right) \times \text{Pr}_{\hat{\theta}}(s_t = j | \Omega_T) \right) \tag{3.49}
\end{aligned}$$

Finally for the fifth block V , substituting equations (3.43) and (3.44) gives

$$\begin{aligned}
V &= E_{\hat{\theta}} \left[\mathbb{I}[s_t = j] \mathbb{I}[s_{t-1} = i] \left(\log \Pr_{\hat{\theta}}(s_t = j | s_{t-1} = i) \right) \middle| \Omega_T \right] \\
&= E_{\hat{\theta}} \left[\mathbb{I}[s_t = j] \mathbb{I}[s_{t-1} = i] \middle| \Omega_T \right] \times E_{\hat{\theta}} \left[\left(\log \Pr_{\hat{\theta}}(s_t = j | s_{t-1} = i) \right) \middle| \Omega_T \right] \\
&= \Pr_{\hat{\theta}}(s_t = j, s_{t-1} = i | \Omega_T) \times \log \Pr_{\hat{\theta}}(s_t = j | s_{t-1} = i) \tag{3.50}
\end{aligned}$$

Given these, I rewrite the Q function (3.42) by substituting them into this.

$$\begin{aligned}
Q &= E_{\hat{\theta}} [\log L_c(\theta) | \Omega_T] = -\frac{T(N+r)}{2} \log(2\pi) \\
&- \frac{1}{2} \sum_{t=1}^T \sum_{j=1}^M \left[\log |R_j| \Pr_{\hat{\theta}}(\mathbf{s}_t = \mathbf{j} | \Omega_T) \right. \\
&+ \text{tr} \left(R_j^{-1} \left(Y_t Y_t' - Y_t \mathbf{E}_{\hat{\theta}}[(\mathbf{F}_t^j)' | \Omega_T] C_j' - C_j \mathbf{E}_{\hat{\theta}}[\mathbf{F}_t^j | \Omega_T] Y_t' \right. \right. \\
&+ \left. \left. C_j \mathbf{E}_{\hat{\theta}}[\mathbf{F}_t^j (\mathbf{F}_t^j)' | \Omega_T] C_j' \right) \right) \times \Pr_{\hat{\theta}}(\mathbf{s}_t = \mathbf{j} | \Omega_T) \\
&+ \log |Q_j| \Pr_{\hat{\theta}}(\mathbf{s}_t = \mathbf{j} | \Omega_T) + \text{tr} \left(Q_j^{-1} \left(\mathbf{E}_{\hat{\theta}}[\mathbf{F}_t^j (\mathbf{F}_t^j)' | \Omega_T] \right. \right. \\
&- \left. \left. \mathbf{E}_{\hat{\theta}}[\mathbf{F}_t^j | \Omega_T] \mu_j' - \mathbf{E}_{\hat{\theta}}[\mathbf{F}_t^j (\mathbf{F}_{t-1}^j)' | \Omega_T] A_j' - \mu_j \mathbf{E}_{\hat{\theta}}[(\mathbf{F}_t^j)' | \Omega_T] \right. \right. \\
&+ \left. \left. \mu_j \mu_j' + \mu_j \mathbf{E}_{\hat{\theta}}[(\mathbf{F}_{t-1}^j)' | \Omega_T] A_j' - A_j \mathbf{E}_{\hat{\theta}}[\mathbf{F}_{t-1}^j (\mathbf{F}_t^j)' | \Omega_T] \right. \right. \\
&+ \left. \left. A_j \mathbf{E}_{\hat{\theta}}[\mathbf{F}_{t-1}^j | \Omega_T] \mu_j' + A_j \mathbf{E}_{\hat{\theta}}[\mathbf{F}_{t-1}^j (\mathbf{F}_{t-1}^j)' | \Omega_T] A_j' \right) \right) \times \Pr_{\hat{\theta}}(\mathbf{s}_t = \mathbf{j} | \Omega_T) \\
&+ \left. \sum_{i=1}^M \Pr_{\hat{\theta}}(\mathbf{s}_t = \mathbf{j}, \mathbf{s}_{t-1} = \mathbf{i} | \Omega_T) \log \Pr_{\hat{\theta}}(s_t = j | s_{t-1} = i) \right] \tag{3.51}
\end{aligned}$$

where the bold text is used to emphasize that the object is an outcome from the smoothing.

Now the last couple of steps. First, the covariance decomposition allows rewriting the equation (3.51) as

$$\begin{aligned}
E_{\hat{\theta}}[F_t^j (F_t^j)' | \Omega_T] &= E_{\hat{\theta}}[F_t^j | \Omega_T] E_{\hat{\theta}}[(F_t^j)' | \Omega_T] + Cov_{\hat{\theta}}[F_t^j (F_t^j)' | \Omega_T] \\
E_{\hat{\theta}}[F_t^j (F_{t-1}^j)' | \Omega_T] &= E_{\hat{\theta}}[F_t^j | \Omega_T] E_{\hat{\theta}}[(F_{t-1}^j)' | \Omega_T] + Cov_{\hat{\theta}}[F_t^j (F_{t-1}^j)' | \Omega_T] \\
E_{\hat{\theta}}[F_{t-1}^j (F_t^j)' | \Omega_T] &= E_{\hat{\theta}}[F_{t-1}^j | \Omega_T] E_{\hat{\theta}}[(F_t^j)' | \Omega_T] + Cov_{\hat{\theta}}[F_{t-1}^j (F_t^j)' | \Omega_T] \\
E_{\hat{\theta}}[F_{t-1}^j (F_{t-1}^j)' | \Omega_T] &= E_{\hat{\theta}}[F_{t-1}^j | \Omega_T] E_{\hat{\theta}}[(F_{t-1}^j)' | \Omega_T] + Cov_{\hat{\theta}}[F_{t-1}^j (F_{t-1}^j)' | \Omega_T].
\end{aligned}$$

Second, the filter and smoother of [Kim and Nelson \(2006\)](#) are incomplete for the EM because their approach does not require estimating the covariance term $Cov_{\hat{\theta}}[F_t^j (F_{t-1}^j)' | \Omega_T]$. Here, I use the approach of [Degras et al. \(2022\)](#) to obtain this term. First define

$$P_{t,t+1|T}^{(j,k)} = E[(F_{t+1} - F_{t+1|t}^{(j,k)})(F_{t+1} - F_{t+1|t}^{(j,k)})' | \Omega_T, s_t = j, s_{t+1} = k]$$

[Degras et al. \(2022\)](#) extend the single-regime smoothing step of the covariance term in [Byron et al. \(2004\)](#) to the regime-switching framework using

$$P_{t,t+1|T}^{(j,k)} = P_{t+1|T}^k J_{t+1|t}^{(j,k)'} \quad (3.52)$$

where $(j, k) \in [1, \dots, M]$. Similar to the smoothing step of [Kim and Nelson \(2006\)](#), I calculate the covariance term by integrating out $P_{t,t+1|T}^{(j,k)}$ to $P_{t,t+1|T}^k$. [Murphy \(1998\)](#) shows that this can be done by

$$P_{t,t+1|T}^k = \sum_{j=1}^M \frac{Pr(s_t = j, s_{t+1} = k | \Omega_T)}{Pr(s_{t+1} = k | \Omega_T)} (P_{t,t+1|T}^{(j,k)} + (F_{t+1|T}^k - F_{t,t+1|T}^{(j,k)})(F_{t+1|T}^k - F_{t,t+1|T}^{(j,k)})'). \quad (3.53)$$

where

$$F_{t+1|T}^k = E[F_{t+1}|\Omega_T, s_{t+1} = k]$$

$$F_{t,t+1|T}^{(j,k)} = E[F_{t+1}|\Omega_T, s_t = j, s_{t+1} = k].$$

Finally, using the approximation suggested by [Degras et al. \(2022\)](#), $F_{t+1|T}^k \approx$

$F_{t,t+1|T}^{(j,k)}$. The expression reduces to

$$\begin{aligned} P_{t,t+1|T}^k &= \sum_{j=1}^M \frac{Pr(s_t = j, s_{t+1} = k|\Omega_T)}{Pr(s_{t+1} = k|\Omega_T)} P_{t,t+1|T}^{(j,k)} \\ &= \sum_{i=1}^M Pr(s_t = j|s_{t+1} = k, \Omega_T) P_{t,t+1|T}^{(j,k)} \end{aligned} \quad (3.54)$$

Then, I have

$$\begin{aligned}
Q &= E_{\hat{\theta}}[\log L_c(\theta)|\Omega_T] = -\frac{T(N+r)}{2}\log(2\pi) \\
&- \frac{1}{2} \sum_{t=1}^T \sum_{j=1}^M \left[\log|R_j| \Pr_{\hat{\theta}}(\mathbf{s}_t = \mathbf{j}|\Omega_T) \right. \\
&+ \text{tr} \left(R_j^{-1} (Y_t Y_t' - Y_t \mathbf{E}_{\hat{\theta}}[(\mathbf{F}_t^{\mathbf{j}})'|\Omega_T] C_j' - C_j \mathbf{E}_{\hat{\theta}}[\mathbf{F}_t^{\mathbf{j}}|\Omega_T] Y_t' \right. \\
&+ C_j (\mathbf{E}_{\hat{\theta}}[\mathbf{F}_t^{\mathbf{j}}|\Omega_T] \mathbf{E}_{\hat{\theta}}[(\mathbf{F}_t^{\mathbf{j}})'|\Omega_T] + \mathbf{Cov}_{\hat{\theta}}[\mathbf{F}_t^{\mathbf{j}}(\mathbf{F}_t^{\mathbf{j}})'|\Omega_T]) C_j' \left. \right) \\
&\times \Pr_{\hat{\theta}}(\mathbf{s}_t = \mathbf{j}|\Omega_T) + \log|Q_j| \Pr_{\hat{\theta}}(\mathbf{s}_t = \mathbf{j}|\Omega_T) \\
&+ \text{tr} \left(Q_j^{-1} ((\mathbf{E}_{\hat{\theta}}[\mathbf{F}_t^{\mathbf{j}}|\Omega_T] \mathbf{E}_{\hat{\theta}}[(\mathbf{F}_t^{\mathbf{j}})'|\Omega_T] + \mathbf{Cov}_{\hat{\theta}}[\mathbf{F}_t^{\mathbf{j}}(\mathbf{F}_t^{\mathbf{j}})'|\Omega_T]) \right. \\
&- \mathbf{E}_{\hat{\theta}}[\mathbf{F}_t^{\mathbf{j}}|\Omega_T] \mu_j' - (\mathbf{E}_{\hat{\theta}}[\mathbf{F}_t^{\mathbf{j}}|\Omega_T] \mathbf{E}_{\hat{\theta}}[(\mathbf{F}_{t-1}^{\mathbf{j}})'|\Omega_T] + \mathbf{Cov}_{\hat{\theta}}[\mathbf{F}_t^{\mathbf{j}}(\mathbf{F}_{t-1}^{\mathbf{j}})'|\Omega_T]) A_j' \\
&- \mu_j \mathbf{E}_{\hat{\theta}}[(\mathbf{F}_t^{\mathbf{j}})'|\Omega_T] + \mu_j \mu_j' + \mu_j \mathbf{E}_{\hat{\theta}}[(\mathbf{F}_{t-1}^{\mathbf{j}})'|\Omega_T] A_j' \\
&- A_j (\mathbf{E}_{\hat{\theta}}[\mathbf{F}_{t-1}^{\mathbf{j}}|\Omega_T] \mathbf{E}_{\hat{\theta}}[(\mathbf{F}_t^{\mathbf{j}})'|\Omega_T] + \mathbf{Cov}_{\hat{\theta}}[\mathbf{F}_{t-1}^{\mathbf{j}}(\mathbf{F}_t^{\mathbf{j}})'|\Omega_T]) \\
&+ A_j \mathbf{E}_{\hat{\theta}}[\mathbf{F}_{t-1}^{\mathbf{j}}|\Omega_T] \mu_j' + A_j (\mathbf{E}_{\hat{\theta}}[\mathbf{F}_{t-1}^{\mathbf{j}}|\Omega_T] \mathbf{E}_{\hat{\theta}}[(\mathbf{F}_{t-1}^{\mathbf{j}})'|\Omega_T] \\
&+ \mathbf{Cov}_{\hat{\theta}}[\mathbf{F}_{t-1}^{\mathbf{j}}(\mathbf{F}_{t-1}^{\mathbf{j}})'|\Omega_T]) A_j' \left. \right) \times \Pr_{\hat{\theta}}(\mathbf{s}_t = \mathbf{j}|\Omega_T) \\
&+ \sum_{i=1}^M \Pr_{\hat{\theta}}(\mathbf{s}_t = \mathbf{j}, \mathbf{s}_{t-1} = \mathbf{i}|\Omega_T) \log \Pr_{\hat{\theta}}(s_t = j | s_{t-1} = i) \left. \right] \quad (3.55)
\end{aligned}$$

where the expressions in bold are once again smoothed objects. To demonstrate its correctness, a useful test is to see whether the likelihood for the special case of single-regime is equivalent to that actually written for the single-regime case. This is the case.

Corollary 1. *For the number of regimes $M=1$, the intercept term $\mu = 0$ and $P(s_t = j, s_{t-1} = i|\Omega_T) = P(s_t = j|\Omega_T) = 1 \ \forall t \in [1, ..., T]$, the Q -function in (3.55) becomes*

$$\begin{aligned}
Q &= E_{\hat{\theta}}[\log L_c(\theta)|\Omega_T] = -\frac{T(N+r)}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^T \left[\log|R| \right. \\
&+ \text{tr}\left(R^{-1}(Y_t Y_t' - Y_t \mathbf{E}_{\hat{\theta}}[(\mathbf{F}_t)']|\Omega_T]C' - C \mathbf{E}_{\hat{\theta}}[\mathbf{F}_t|\Omega_T]Y_t' + C \mathbf{E}_{\hat{\theta}}[\mathbf{F}_t(\mathbf{F}_t)']|\Omega_T]C')\right) \\
&+ \log|Q| + \text{tr}\left(Q^{-1}(\mathbf{E}_{\hat{\theta}}[\mathbf{F}_t(\mathbf{F}_t)']|\Omega_T] - \mathbf{E}_{\hat{\theta}}[\mathbf{F}_t(\mathbf{F}_{t-1})']|\Omega_T]A' - A \mathbf{E}_{\hat{\theta}}[\mathbf{F}_{t-1}(\mathbf{F}_t)']|\Omega_T] \right. \\
&+ \left. A \mathbf{E}_{\hat{\theta}}[\mathbf{F}_{t-1}(\mathbf{F}_{t-1})']|\Omega_T]A')\right) \Big] \tag{3.56}
\end{aligned}$$

which is equivalent to the expected log-likelihood of [Bańbura and Modugno \(2014\)](#) for the single-regime DFM for an otherwise identical setting.

The Maximization Step

In this subsection, I estimate the parameters by maximizing $E_{\hat{\theta}}[\log L_c(\theta)|\Omega_T]$ with respect to $\hat{\theta} = \{(A_j, Q_j, C_j, R_j, \mu_j) : 1 < j < M; \pi \in Z\}$. Before deriving the first order conditions, I present useful standard matrix operations to be used repeatedly below:

$$\frac{\partial((Xa+b)'C(Xa+b))}{\partial X} = (C+C')(Xa+b)a' \tag{3.57}$$

$$\frac{\partial(a'Xb)}{\partial X} = ab' \tag{3.58}$$

$$\frac{\partial(a'F^{-1}b)}{\partial X} = -(X')^{-1}ab'(X')^{-1} \tag{3.59}$$

$$\frac{\partial \log|X|}{\partial X} = (X')^{-1} \tag{3.60}$$

I obtain the estimate $C[s_t = j] \forall j \in [1, \dots, M]$ by solving the first order conditions.

$$\frac{\partial E_{\hat{\theta}}[\log L_c(\theta)|\Omega_T]}{\partial C[s_t = j]} = 0 \tag{3.61}$$

Using (3.57), this becomes

$$\sum_{t=1}^T \sum_{j=1}^M \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}} \left[\{ 2R[s_t = j]^{-1} (Y_t - C[s_t = j] F_t^j) F_t^j \} | \Omega_T \right] \right] = 0.$$

Then, for each $s_t = j \in [1, \dots, M]$,

$$\begin{aligned} & 2R[s_t = j]^{-1} \sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}} \left[Y_t (F_t^j)' | \Omega_T \right] \right. \\ & \left. - C[s_t = j] Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}} \left[F_t^j (F_t^j)' | \Omega_T \right] \right] = 0. \end{aligned}$$

Rearranging the expression above, I obtain the estimate for $C[s_t = j]$

$$C[s_t = j] = \frac{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}} \left[Y_t (F_t^j)' | \Omega_T \right] \right]}{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}} \left[F_t^j (F_t^j)' | \Omega_T \right] \right]} \quad (3.62)$$

Similarly, for $A[s_t = j] \forall j \in [1, \dots, M]$, the first order condition is

$$\frac{\partial E_{\hat{\theta}}[\log L_c(\theta) | \Omega_T]}{\partial A[s_t = j]} = 0. \quad (3.63)$$

Using (3.57), this becomes

$$\begin{aligned} & \sum_{t=1}^T \sum_{j=1}^M \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}} \left[\{ 2Q[s_t = j]^{-1} \right. \right. \\ & \left. \left. \times (F_t^j - \mu[s_t = j] - A[s_t = j] F_{t-1}^j) (F_{t-1}^j)' \} | \Omega_T \right] \right] = 0. \end{aligned}$$

Then, for each $s_t = j \in [1, \dots, M]$,

$$2Q[s_t = j]^{-1} \sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) (E_{\hat{\theta}}[F_t^j (F_{t-1}^j)' | \Omega_T] - \mu[s_t = j] E_{\hat{\theta}}[(F_{t-1}^j)' | \Omega_T]) \right] \\ - 2Q[s_t = j]^{-1} \sum_{t=1}^T \left[A[s_t = j] Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_{t-1}^j (F_{t-1}^j)' | \Omega_T] \right] = 0.$$

Rearranging the expressions above, I obtain the estimate for $A[s_t = j]$

$$A[s_t = j] = \frac{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) (E_{\hat{\theta}}[F_t^j (F_{t-1}^j)' | \Omega_T]) \right]}{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_{t-1}^j (F_{t-1}^j)' | \Omega_T] \right]} \\ - \frac{\mu[s_t = j] \sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) (E_{\hat{\theta}}[(F_{t-1}^j)' | \Omega_T]) \right]}{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_{t-1}^j (F_{t-1}^j)' | \Omega_T] \right]}. \quad (3.64)$$

Now, I obtain the estimate of $\mu[s_t = j] \forall j \in [1, \dots, M]$ by solving the first order condition: (3.64).

$$\frac{\partial E_{\hat{\theta}}[\log L_c(\theta) | \Omega_T]}{\partial \mu[s_t = j]} = 0. \quad (3.65)$$

Using (3.57), this expression is

$$\sum_{t=1}^T \sum_{j=1}^M \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}} \left[\{ 2Q[s_t = j]^{-1} \right. \right. \\ \left. \left. \times (F_t^j - \mu[s_t = j] - A[s_t = j] F_{t-1}^j) \} | \Omega_T \right] \right] = 0$$

Then, for each $s_t = j \in [1, \dots, M]$,

$$2Q[s_t = j]^{-1} \sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_t^j | \Omega_T] \right] - 2Q[s_t = j]^{-1} (T-1) \mu[s_t = j] \\ - 2Q[s_t = j]^{-1} A[s_t = j] \sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_{t-1}^j | \Omega_T] \right] = 0$$

which can be rearranged as

$$\mu[s_t = j] = \frac{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_t^j | \Omega_T] \right]}{(T-1)} \\ - \frac{A[s_t = j] \sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_{t-1}^j | \Omega_T] \right]}{(T-1)}. \quad (3.66)$$

In order to obtain the analytical solutions for the estimates $A[s_t = j]$ and $\mu[s_t = j]$, I need to solve equations (3.64) and (3.66) simultaneously. For the purpose of tractability, I define four expressions:

$$\kappa_1 = \frac{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) \left(E_{\hat{\theta}}[F_t^j (F_{t-1}^j)' | \Omega_T] \right) \right]}{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_{t-1}^j (F_{t-1}^j)' | \Omega_T] \right]} \\ \kappa_2 = \frac{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) \left(E_{\hat{\theta}}[(F_{t-1}^j)' | \Omega_T] \right) \right]}{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_{t-1}^j (F_{t-1}^j)' | \Omega_T] \right]} \\ \kappa_3 = \frac{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_t^j | \Omega_T] \right]}{(T-1)} \\ \kappa_4 = \frac{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_{t-1}^j | \Omega_T] \right]}{(T-1)}.$$

I can now rewrite equations (3.64) and (3.66) as

$$A[s_t = j] = \kappa_1 - \mu[s_t = j]\kappa_2$$

$$\mu[s_t = j] = \kappa_3 - A[s_t = j]\kappa_4.$$

Then, the analytical solutions for $A[s_t = j]$ and $\mu[s_t = j]$ are

$$A[s_t = j] = \frac{\kappa_1 - \kappa_2\kappa_3}{(I - \kappa_2\kappa_4)} \quad (3.67)$$

and

$$\mu[s_t = j] = \frac{\kappa_3 - \kappa_1\kappa_4}{(I - \kappa_2\kappa_4)} \quad (3.68)$$

where under the restriction of $\mu[s_t] = 0, \forall t$, the estimate for $A[s_t = j]$ simplifies to κ_1 .

To find $R[s_t = j]$, I solve the first order condition $\forall j \in [1, \dots, M]$:

$$\frac{\partial E_{\hat{\theta}}[\log L_c(\theta) | \Omega_T]}{\partial R[s_t = j]} = 0. \quad (3.69)$$

Using (3.59) and (3.60), this becomes

$$\begin{aligned} & - \frac{1}{2} \sum_{t=1}^T \sum_{j=1}^M Pr_{\hat{\theta}}(s_t = j | \Omega_T) (R[s_t = j])^{-1} \\ & + \frac{1}{2} \sum_{t=1}^T \sum_{j=1}^M Pr_{\hat{\theta}}(s_t = j | \Omega_T) \times \left[E_{\hat{\theta}} \left[(R[s_t = j])^{-1} \right. \right. \\ & \quad \left. \left. \times (Y_t - C[s_t = j]F_t^j)(Y_t - C[s_t = j]F_t^j)'(R[s_t = j])^{-1} | \Omega_T \right] \right] = 0. \end{aligned}$$

Then, for each $s_t = j \in [1, \dots, M]$,

$$\begin{aligned} R[s_t = j]' &= \left(\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) \right)^{-1} \sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) \\ &\quad \times E_{\hat{\theta}} \left[(Y_t - C[s_t = j]F_t^j)(Y_t - C[s_t = j]F_t^j)' | \Omega_T \right] \end{aligned}$$

Rearranging this,

$$\begin{aligned} R[s_t = j]' &= \left(\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) \right)^{-1} \left[\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) Y_t Y_t' \right. \\ &\quad - \left(\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}} \left[Y_t (F_t^j)' | \Omega_T \right] \right) C[s_t = j]' \\ &\quad - C[s_t = j] \left(\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}} \left[(F_t^j) Y_t' | \Omega_T \right] \right) \\ &\quad \left. + C[s_t = j] \left(\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}} \left[(F_t^j)(F_t^j)' | \Omega_T \right] \right) C[s_t = j]' \right]. \end{aligned}$$

The inspection of equation (3.62) reveals that

$$\begin{aligned} &\left(\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}} \left[Y_t (F_t^j)' | \Omega_T \right] \right) C[s_t = j]' = \\ &C[s_t = j] \left(\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}} \left[(F_t^j)(F_t^j)' | \Omega_T \right] \right) C[s_t = j]'. \end{aligned}$$

I substitute this into the right hand side of the equation above to eliminate the second and the last term. This gives

$$\begin{aligned} R[s_t = j]' &= \left(\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) \right)^{-1} \times \left[\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) Y_t Y_t' \right. \\ &\quad \left. - C[s_t = j] \left(\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}} \left[(F_t^j) Y_t' | \Omega_T \right] \right) \right]. \quad (3.70) \end{aligned}$$

Finally, in order to find the estimate for $Q[s_t = j]$, I solve the first order

condition $\forall j \in [1, \dots, M]$:

$$\frac{\partial E_{\hat{\theta}}[\log L_c(\theta) | \Omega_T]}{\partial Q[s_t = j]} = 0. \quad (3.71)$$

Using (3.59) and (3.60), this becomes

$$\begin{aligned} & - \frac{1}{2} \sum_{t=1}^T \sum_{j=1}^M Pr_{\hat{\theta}}(s_t = j | \Omega_T) (Q[s_t = j]')^{-1} \\ & + \frac{1}{2} \sum_{t=1}^T \sum_{j=1}^M Pr_{\hat{\theta}}(s_t = j | \Omega_T) \left[E_{\hat{\theta}} \left[(Q[s_t = j]')^{-1} (F_t^j - \mu[s_t = j] - A[s_t = j] F_{t-1}^j) \right. \right. \\ & \times \left. \left. (F_t^j - \mu[s_t = j] - A[s_t = j] F_{t-1}^j)' (Q[s_t = j]')^{-1} | \Omega_T \right] \right] = 0. \end{aligned}$$

Then, for each $s_t = j \in [1, \dots, M]$,

$$\begin{aligned} Q[s_t = j]' &= \left(\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) \right)^{-1} \sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) \\ &\times E_{\hat{\theta}} \left[(F_t^j - \mu[s_t = j] - A[s_t = j] F_{t-1}^j) \right. \\ &\times \left. (F_t^j - \mu[s_t = j] - A[s_t = j] F_{t-1}^j)' | \Omega_T \right]. \end{aligned}$$

Rearranging this,

$$\begin{aligned}
Q[s_t = j]' &= \left(\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) \right)^{-1} \left[\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}} \left[(F_t^j)(F_t^j)' | \Omega_T \right] \right. \\
&\quad - \sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}} \left[(F_t^j) | \Omega_T \right] \mu'[s_t = j] \\
&\quad - \sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}} \left[(F_t^j)(F_{t-1}^j)' | \Omega_T \right] A'[s_t = j] \\
&\quad - \sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) \mu[s_t = j] E_{\hat{\theta}} \left[(F_t^j)' | \Omega_T \right] \\
&\quad + \sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) \mu[s_t = j] \mu'[s_t = j] \\
&\quad + \sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) \mu[s_t = j] E_{\hat{\theta}} \left[(F_{t-1}^j)' | \Omega_T \right] A'[s_t = j] \\
&\quad - \sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) A[s_t = j] E_{\hat{\theta}} \left[(F_{t-1}^j)(F_t^j)' | \Omega_T \right] \\
&\quad + \sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) A[s_t = j] E_{\hat{\theta}} \left[(F_{t-1}^j) | \Omega_T \right] \mu'[s_t = j] \\
&\quad \left. + \sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) A[s_t = j] E_{\hat{\theta}} \left[(F_{t-1}^j)(F_{t-1}^j)' | \Omega_T \right] A'[s_t = j] \right].
\end{aligned}$$

As before, it is possible to simplify this expression. The equation (3.63) suggests that for each $s_t = j \in [1, \dots, M]$

$$\begin{aligned}
2Q[s_t = j]^{-1} \sum_{t=1}^T &\left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) (E_{\hat{\theta}}[F_t^j(F_{t-1}^j)' | \Omega_T] - \mu[s_t = j] E_{\hat{\theta}}[(F_{t-1}^j)' | \Omega_T]) \right] \\
-2Q[s_t = j]^{-1} \sum_{t=1}^T &\left[A[s_t = j] Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_{t-1}^j(F_{t-1}^j)' | \Omega_T] \right] = 0.
\end{aligned}$$

This implies that

$$\begin{aligned}
&\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) \left(E_{\hat{\theta}}[F_t^j(F_{t-1}^j)' | \Omega_T] \right. \right. \\
&\quad \left. \left. - \mu[s_t = j] E_{\hat{\theta}}[(F_{t-1}^j)' | \Omega_T] - A[s_t = j] E_{\hat{\theta}}[F_{t-1}^j(F_{t-1}^j)' | \Omega_T] \right) \right] = 0.
\end{aligned}$$

Right-multiply the expression above by $A'[s_t = j]$ to get

$$\begin{aligned}
& \sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_t^j (F_{t-1}^j)' | \Omega_T] A'[s_t = j] \right] \\
&= \sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) \mu[s_t = j] E_{\hat{\theta}}[(F_{t-1}^j)' | \Omega_T] A'[s_t = j] \right] \\
&+ \sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) A[s_t = j] E_{\hat{\theta}}[F_{t-1}^j (F_{t-1}^j)' | \Omega_T] A'[s_t = j] \right]. \quad (3.72)
\end{aligned}$$

This eliminates the lines three, six, and nine on the right hand side of $Q[s_t = j]'$.

Similarly, (3.64) suggests that

$$\begin{aligned}
2Q[s_t = j]^{-1} \sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_t^j | \Omega_T] \right] - 2Q[s_t = j]^{-1} (T-1) \mu[s_t = j] \\
- 2Q[s_t = j]^{-1} A[s_t = j] \sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_{t-1}^j | \Omega_T] \right] = 0.
\end{aligned}$$

This implies that

$$\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) \left[E_{\hat{\theta}}[F_t^j | \Omega_T] - \mu[s_t = j] - A[s_t = j] E_{\hat{\theta}}[F_{t-1}^j | \Omega_T] \right] = 0.$$

Right-multiply the equation by $\mu'[s_t = j]$ to get

$$\begin{aligned}
& \sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) \left[E_{\hat{\theta}}[F_t^j | \Omega_T] \mu'[s_t = j] \right. \\
& \left. - \mu[s_t = j] \mu'[s_t = j] - A[s_t = j] E_{\hat{\theta}}[F_{t-1}^j | \Omega_T] \mu'[s_t = j] \right] = 0. \quad (3.73)
\end{aligned}$$

Substituting this into $Q[s_t = j]'$ additionally eliminates the lines two, five, and

eight, leading to the simplified expression

$$\begin{aligned}
Q[s_t = j]' &= \left(\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) \right)^{-1} \left[\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}} \left[(F_t^j)(F_t^j)' | \Omega_T \right] \right. \\
&\quad - \sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) \mu[s_t = j] E_{\hat{\theta}} \left[(F_t^j)' | \Omega_T \right] \\
&\quad \left. - \sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) A[s_t = j] E_{\hat{\theta}} \left[(F_{t-1}^j)(F_t^j)' | \Omega_T \right] \right] \quad (3.74)
\end{aligned}$$

This completes $\{(A_j, Q_j, C_j, R_j, \mu_j) : 1 < j < M\}$. The remaining object is the transition probabilities. As shown by [Hamilton \(1989\)](#), these are

$$\pi_{i,j} = \frac{\sum_{t=1}^T Pr_{\hat{\theta}}(s_{t-1} = i, s_t = j | \Omega_T)}{\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = i | \Omega_T)}. \quad (3.75)$$

To summarize, the equations (3.62), (3.67), (3.68), (3.70), (3.74), and (3.75) constitute the maximization step.

The Modifications for Mixed and Missing Data

In this subsection, I describe the standard approach for dealing with mixed-frequency and missing data problems. I follow [Bańbura and Modugno \(2014\)](#) and [Bok et al. \(2018\)](#) who use the approximation of [Mariano and Murasawa \(2003\)](#) for this purpose.

Let y_t be a quarterly time series data which is the geometric mean of the underlying monthly series $\hat{y}_t$, that is

$$y_t = \left(\hat{y}_t \times \hat{y}_{t-1} \times \hat{y}_{t-2} \right)^{\frac{1}{3}}. \quad (3.76)$$

Taking the logarithm of each side,

$$\log y_t = \frac{1}{3} \times \left(\log \hat{y}_t + \log \hat{y}_{t-1} + \log \hat{y}_{t-2} \right). \quad (3.77)$$

Define $\Delta_i x_t = \log x_t - \log x_{t-i}$ and take the quarterly difference to obtain

$$\begin{aligned} \Delta_3 y_t &= \log y_t - \log y_{t-3} \\ &= \frac{1}{3} \times \left(\log \hat{y}_t + \log \hat{y}_{t-1} + \log \hat{y}_{t-2} \right) - \frac{1}{3} \times \left(\log \hat{y}_{t-3} + \log \hat{y}_{t-4} + \log \hat{y}_{t-5} \right). \end{aligned} \quad (3.78)$$

For any variable x , it is possible to rewrite $(x_{t-i} - x_{t-i-3})$ as $(x_{t-i} - x_{t-i-1} + x_{t-i-1} - x_{t-i-2} + x_{t-i-2} - x_{t-i-3})$. Applying this to the quarterly difference above,

$$\begin{aligned} \Delta_3 y_t &= \log y_t - \log y_{t-3} \\ &= \frac{1}{3} \times \left(\log \hat{y}_t - \log \hat{y}_{t-3} \right) + \frac{1}{3} \times \left(\log \hat{y}_{t-1} - \log \hat{y}_{t-4} \right) \\ &\quad + \frac{1}{3} \times \left(\log \hat{y}_{t-2} - \log \hat{y}_{t-5} \right) \\ &= \frac{1}{3} \times \left(\log \hat{y}_t - \log \hat{y}_{t-1} + \log \hat{y}_{t-1} - \log \hat{y}_{t-2} + \log \hat{y}_{t-2} - \log \hat{y}_{t-3} \right) \\ &\quad + \frac{1}{3} \times \left(\log \hat{y}_{t-1} - \log \hat{y}_{t-2} + \log \hat{y}_{t-2} - \log \hat{y}_{t-3} + \log \hat{y}_{t-3} - \log \hat{y}_{t-4} \right) \\ &\quad + \frac{1}{3} \times \left(\log \hat{y}_{t-2} - \log \hat{y}_{t-3} + \log \hat{y}_{t-3} - \log \hat{y}_{t-4} + \log \hat{y}_{t-4} - \log \hat{y}_{t-5} \right) \\ &= \frac{1}{3} \times \left(\Delta_1 \hat{y}_t + 2\Delta_1 \hat{y}_{t-1} + 3\Delta_1 \hat{y}_{t-2} + 2\Delta_1 \hat{y}_{t-3} + \Delta_1 \hat{y}_{t-4} \right). \end{aligned} \quad (3.79)$$

Equation (3.79) defines a linear relation between the quarterly and monthly data. This is referred to as the Mariano-Murasawa approximation in the literature.

The mixed-frequency data problem is dealt with by incorporating the Mariano-Murasawa approximation into the DFM. This changes the observation equation and the state equation. Because it is possible to augment the DFM to include the Mariano-Murasawa approximation and still write the DFM in the form of equation (3.39), the EM algorithm above continues to apply. To see this, refer to [Banbura et al. \(2010\)](#).

To tackle the issue arising from the missing data problem, the expression from the maximization step

$$C[s_t = j] = \frac{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[Y_t(F_t^j)' | \Omega_T] \right]}{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_t^j(F_t^j)' | \Omega_T] \right]}$$

needs to be modified as

$$C[s_t = j] = \frac{\sum_{t \in t^*} \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[Y_t(F_t^j)' | \Omega_T] \right]}{\sum_{t \in t^*} \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_t^j(F_t^j)' | \Omega_T] \right]} \quad (3.80)$$

where t^* includes only the observed data. The estimation of $R[s_t = j]$ follows a similar modification. See [Bańbura and Modugno \(2014\)](#) for this.

3.4 Conclusion

This chapter presents a modified EM algorithm for estimating large-scale DFMs. This includes modified filtering and smoothing procedures which take the regime-switching into account. These are summarized by equations (3.19)-(3.24), (3.27)-(3.31), (3.25)-(3.26), (3.32)-(3.34), (3.37)-(3.38), (3.35)-(3.36), (3.62), (3.67), (3.68), (3.70), (3.74), and (3.75) which are reproduced here for the ease of reference:



$$\begin{aligned}
F_{t|t-1}^{(i,j)} &= A_t^j F_{t-1|t-1}^i \\
P_{t|t-1}^{(i,j)} &= A_t^j P_{t-1|t-1}^i A_t^{j'} + Q_j \\
\eta_{t|t-1}^{(i,j)} &= Y_t - C_j F_{t|t-1}^{(i,j)} \\
f_{t|t-1}^{(i,j)} &= C_j P_{t|t-1}^{(i,j)} C_j' + R_j \\
F_{t|t}^{(i,j)} &= F_{t|t-1}^{(i,j)} + P_{t|t-1}^{(i,j)} C_j' [f_{t|t-1}^{(i,j)}]^{-1} \eta_{t|t-1}^{(i,j)} \\
P_{t|t}^{(i,j)} &= (I - P_{t|t-1}^{(i,j)} C_j' [f_{t|t-1}^{(i,j)}]^{-1} C_j) P_{t|t-1}^{(i,j)} \\
Pr(s_t = j, s_{t-1} = i | \Omega_{t-1}) &= Pr(s_t = j | s_{t-1} = i) Pr(s_{t-1} = i | \Omega_{t-1}) \\
Pr(Y_t | s_t = j, s_{t-1} = i, \Omega_{t-1}) &= \frac{1}{2\pi^{\frac{N}{2}}} |f_{t|t-1}^{(i,j)}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} \eta_{t|t-1}^{(i,j)'} f_{t|t-1}^{(i,j)-1} \eta_{t|t-1}^{(i,j)}\right\} \\
Pr(Y_t | \Omega_{t-1}) &= \sum_{j=1}^M \sum_{i=1}^M \left(Pr(Y_t | s_t = j, s_{t-1} = i, \Omega_{t-1}) \right. \\
&\quad \left. \times Pr(s_t = j, s_{t-1} = i | \Omega_{t-1}) \right) \\
Pr(s_t = j, s_{t-1} = i | \Omega_t) &= \left(Pr(Y_t | s_t = j, s_{t-1} = i, \Omega_{t-1}) \right. \\
&\quad \left. \times \frac{Pr(s_t = j, s_{t-1} = i | \Omega_{t-1})}{Pr(Y_t | \Omega_{t-1})} \right) \\
Pr(s_t = i | \Omega_t) &= \sum_{i=1}^M Pr(s_t = j, s_{t-1} = i | \Omega_t) \\
F_{t|t}^j &= \sum_{i=1}^M Pr(s_{t-1} = i | s_t = j, \Omega_t) F_{t|t}^{(i,j)} \\
P_{t|t}^j &= \sum_{i=1}^M Pr(s_{t-1} = i | s_t = j, \Omega_t) \\
&\quad \times [P_{t|t}^{(i,j)} + (F_{t|t}^j - F_{t|t}^{(i,j)})(F_{t|t}^j - F_{t|t}^{(i,j)})'] \\
J_{t+1|t}^{(j,k)} &= P_{t|t}^j A_t^{k'} [P_{t+1|t}^{(j,k)}]^{-1} \\
F_{t|T}^{(j,k)} &= F_{t|t}^j + J_{t+1|t}^{(j,k)} (F_{t+1|T}^k - A_t^k F_{t|t}^j) \\
P_{t|T}^{(j,k)} &= P_{t|t}^j + J_{t+1|t}^{(j,k)} (P_{t+1|T}^k - P_{t+1|t}^{(j,k)}) J_{t+1|t}^{(j,k)'}
\end{aligned}$$

$$\begin{aligned}
Pr(s_t = j, s_{t+1} = k | \Omega_T) &= Pr(s_{t+1} = k | \Omega_T) \\
&\times \frac{Pr(s_{t+1} = k | s_t = j) Pr(s_t = j | \Omega_t)}{Pr(s_{t+1} = k | \Omega_t)} \\
Pr(s_t = j | \Omega_T) &= \sum_{k=1}^M Pr(s_t = j, s_{t+1} = k | \Omega_T) \\
F_{t|T}^j &= \sum_{k=1}^M Pr(s_{t+1} = k | s_t = j, \Omega_T) F_{t|T}^{(j,k)} \\
P_{t|T}^j &= \sum_{k=1}^M Pr(s_{t+1} = k | s_t = j, \Omega_T) \\
&\times [P_{t|T}^{(j,k)} + (F_{t|T}^j - F_{t|T}^{(j,k)})(F_{t|T}^j - F_{t|T}^{(j,k)})'] \\
C[s_t = j] &= \frac{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[Y_t(F_t^j)' | \Omega_T] \right]}{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_t^j(F_t^j)' | \Omega_T] \right]} \\
\kappa_1 &= \frac{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) (E_{\hat{\theta}}[F_t^j(F_{t-1}^j)' | \Omega_T]) \right]}{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_{t-1}^j(F_{t-1}^j)' | \Omega_T] \right]} \\
\kappa_2 &= \frac{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) (E_{\hat{\theta}}[(F_{t-1}^j)' | \Omega_T]) \right]}{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_{t-1}^j(F_{t-1}^j)' | \Omega_T] \right]} \\
\kappa_3 &= \frac{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_t^j | \Omega_T] \right]}{(T-1)} \\
\kappa_4 &= \frac{\sum_{t=1}^T \left[Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[F_{t-1}^j | \Omega_T] \right]}{(T-1)} \\
A[s_t = j] &= \frac{\kappa_1 - \kappa_2 \kappa_3}{(I - \kappa_2 \kappa_4)} \\
\mu[s_t = j] &= \frac{\kappa_3 - \kappa_1 \kappa_4}{(I - \kappa_2 \kappa_4)}
\end{aligned}$$

$$\begin{aligned}
R[s_t = j]' &= \left(\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) \right)^{-1} \times \left[\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) Y_t Y_t' \right. \\
&\quad \left. - C[s_t = j] \left(\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[(F_t^j) Y_t' | \Omega_T] \right) \right] \\
Q[s_t = j]' &= \left(\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) \right)^{-1} \left[\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) E_{\hat{\theta}}[(F_t^j)(F_t^j)' | \Omega_T] \right. \\
&\quad \left. - \sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) \mu[s_t = j] E_{\hat{\theta}}[(F_t^j)' | \Omega_T] \right. \\
&\quad \left. - \sum_{t=1}^T Pr_{\hat{\theta}}(s_t = j | \Omega_T) A[s_t = j] E_{\hat{\theta}}[(F_{t-1}^j)(F_t^j)' | \Omega_T] \right] \\
\pi_{i,j} &= \frac{\sum_{t=1}^T Pr_{\hat{\theta}}(s_{t-1} = i, s_t = j | \Omega_T)}{\sum_{t=1}^T Pr_{\hat{\theta}}(s_t = i | \Omega_T)}.
\end{aligned}$$

These are the expectation and maximization steps used in each iteration of the estimation procedure for the EM. Furthermore, the procedure can be modified to deal with mixed-frequency data as well as missing data problems as explained above.

CHAPTER 4

THE APPLICATION OF THE EXPECTATION-MAXIMIZATION ALGORITHM TO REGIME-SWITCHING DYNAMIC FACTOR MODELS: REPLICATIONS AND MONTE CARLO SIMULATIONS

4.1 Introduction

The previous chapter introduced the EM algorithm that can be applied to a regime-switching DFM when the dataset is available at mixed frequencies (e.g., monthly and quarterly) and features missing data problems. This chapter demonstrates the usefulness of this methodology by first replicating well-known examples in the literature. The first example is a monthly-frequency regime-switching DFM of [Kim and Nelson \(2006\)](#) (which was originally published in [Kim and Nelson \(1999\)](#)). The second example is a mixed-frequency regime-switching DFM of [Camacho et al. \(2018\)](#). This chapter is exclusively

concerned with the mixed-frequency data problem. The missing data problem will be dealt with in the following chapter. Finally, I will conduct Monte Carlo simulations to see whether the EM approach can correctly recover the regimes across models of different scales.

The chapter is organized as follows. Section 4.2.1 describes the model, the data, and the methodology in Chapter 5.7 of [Kim and Nelson \(2006\)](#). Using the same data and model structure, the model is first estimated by using the one-step approach described in [Kim and Nelson \(2006\)](#), and then by using the EM approach explained in Chapter 3 of this thesis. Then, the smoothed recession probabilities of two estimation procedures are compared. Section 4.2.2 carries out a similar exercise for a regime-switching mixed frequency DFM application of [Camacho et al. \(2018\)](#). In section 4.2.3, Monte Carlo simulations are conducted to evaluate the accuracy of the methodology in Chapter 3. The regime probabilities are estimated with mixed frequency data of different sizes.

4.2 Exercises

4.2.1 Kim and Nelson (2006)

Section 5.7 of [Kim and Nelson \(2006\)](#) propose a regime-switching DFM to study business cycles. The exercise involves growth rates of four variables, allowing for two regimes in the model, i.e., the economy is either in recession or in normal times. The variables considered are industrial production (IP), the total personal income less transfer payments in 1987 US dollars (GMYXPQ),

the total manufacturing and trade sales in 1987 US Dollars (MTQ), and employees on nonagricultural payrolls (LPNAG), all at monthly frequency. The sample covers from January 1959 to January 1995.

Let ΔY_t be a vector of the first difference of the logarithm of the aforementioned variables:

$$\Delta Y_t = \begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \\ \Delta y_{3t} \\ \Delta y_{4t} \end{bmatrix} = \begin{bmatrix} \log(IP_t) - \log(IP_{t-1}) \\ \log(GMYXPQ_t) - \log(GMYXPQ_{t-1}) \\ \log(MTQ_t) - \log(MTQ_{t-1}) \\ \log(LPNAG_t) - \log(LPNAG_{t-1}) \end{bmatrix}. \quad (4.1)$$

[Kim and Nelson \(2006\)](#) refer to the common component of these indicators as the coincident index, Δc_t . The relationship between the coincident index and the observed variables is given in the measurement equation:

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \\ \Delta y_{3t} \\ \Delta y_{4t} \end{bmatrix} = \begin{bmatrix} \gamma_{11} & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \gamma_{21} & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \gamma_{31} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \gamma_{41} & \gamma_{42} & \gamma_{43} & \gamma_{44} & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} \Delta c_t \\ \Delta c_{t-1} \\ \Delta c_{t-2} \\ \Delta c_{t-3} \\ e_{1,t} \\ e_{1,t-1} \\ e_{2,t} \\ e_{2,t-1} \\ e_{3,t} \\ e_{3,t-1} \\ e_{4,t} \\ e_{4,t-1} \end{bmatrix}. \quad (4.2)$$

The transition equation for the coincident index, which is an unobserved factor in the DFM language, is

$$\begin{bmatrix} \Delta c_t \\ \Delta c_{t-1} \\ \Delta c_{t-2} \\ \Delta c_{t-3} \\ e_{1t} \\ e_{1,t-1} \\ \vdots \\ e_{4,t} \\ e_{4,t-1} \end{bmatrix} = \begin{bmatrix} \phi_{st} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \phi_{1,1} & \phi_{1,2} & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \phi_{5,5} & \phi_{5,6} & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & \phi_{11,11} & \phi_{11,12} \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix} \times \begin{bmatrix} \Delta c_{t-1} \\ \Delta c_{t-2} \\ \Delta c_{t-3} \\ \Delta c_{t-4} \\ e_{1,t-1} \\ e_{1,t-2} \\ \vdots \\ e_{4,t-1} \\ e_{4,t-2} \end{bmatrix} + \begin{bmatrix} v_t \\ 0 \\ 0 \\ 0 \\ \epsilon_{1,t} \\ 0 \\ \vdots \\ \epsilon_{4,t} \\ 0. \end{bmatrix} \quad (4.3)$$

Equations (4.2) and (4.3) constitute a state-space model. In what follows, I will estimate the model using the original one-step approach of [Kim and Nelson \(2006\)](#), which maximizes the approximate log-likelihood in equation (3.29) of the previous chapter using the interior-point optimization (which is implemented by `fmincon` in Matlab), as well as the EM approach that this thesis propounds.

The initial values are as in [Kim and Nelson \(2006\)](#). The data are demeaned and normalized by their respective standard deviations. Recall that the model given by equations (3.39) is a DFM of the form

$$\begin{aligned} Y_t &= C[s_t]F_t + e[s_t] \\ F_t &= \mu[s_t] + A[s_t]F_{t-1} + v[s_t] \\ e[s_t] &\sim \mathcal{N}(0, R[s_t]) \\ v[s_t] &\sim \mathcal{N}(0, Q[s_t]). \end{aligned}$$

The model in (4.2)-(4.3) imposes the restrictions described by equation (3.39). First, the slope of the unobservable factor and the factor loadings of the observables are not subject to regime-switching. In other words, $C[s_t]$ and $A[s_t]$ are defined as:

$$C[s_t] = \bar{C} = \begin{bmatrix} \gamma_{11} & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \gamma_{21} & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \gamma_{31} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \gamma_{41} & \gamma_{42} & \gamma_{43} & \gamma_{44} & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

and

$$A[s_t] = \bar{A} = \begin{bmatrix} \phi_{1,1} & \phi_{1,2} & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \phi_{5,5} & \phi_{5,6} & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & \phi_{11,11} & \phi_{11,12} \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix}.$$

Furthermore, $R = \bar{R} = 0$, and $Q = \bar{Q}$ is restricted to be a nonzero only where appropriate. This is the original specification of [Kim and Nelson \(2006\)](#).

Below, I show that introducing a measurement error does not alter the results under the one-step estimation approach of [Kim and Nelson \(2006\)](#). For this reason, I will utilize the model with the measurement error when I compare the one-step approach to the EM approach.

Recall that there are four objects that are not regime-switching: $(\bar{C}, \bar{R}, \bar{A}, \bar{Q})$.

To apply the EM algorithm to these, I first estimate these regime by regime and then take the average across the regimes, setting the values in each regime to the average:

$$\bar{C} = \sum_{j=1}^M Pr(s_t = j | \Omega_T) C[s_t = j] \quad (4.4)$$

$$\bar{R} = \sum_{j=1}^M Pr(s_t = j | \Omega_T) R[s_t = j] \quad (4.5)$$

$$\bar{A} = \sum_{j=1}^M Pr(s_t = j | \Omega_T) A[s_t = j] \quad (4.6)$$

$$\bar{Q} = \sum_{j=1}^M Pr(s_t = j | \Omega_T) Q[s_t = j]. \quad (4.7)$$

This is done in each step of the iteration until convergence.

Figure 4.1 shows the smoothed recession probabilities for the original model of [Kim and Nelson \(2006\)](#) (blue solid line) without the measurement error, and the modified version of [Kim and Nelson \(2006\)](#) (orange dashed line) with the measurement error. Both are estimated (by maximizing the likelihood) using the one-step approach of [Kim and Nelson \(2006\)](#). The figure shows that the modification leads to negligible differences for the estimated recession probabilities. For this reason, I will use the modified model for the rest of this section because it facilitates computation for the EM approach.¹

Figure 4.2 compares the smoothed recession probabilities for the modified model obtained from the one-step approach (blue solid line) and the EM approach (red dashed line), respectively. It is clear that both approaches give more or less the same estimates.

Given that the two approaches give similar estimates for recession probabilities,

¹The EM approach is well-defined without the measurement error. However, it is a common practice to introduce the measurement error to facilitate convergence of the Kalman filter to reduce the incidence of the associated possible matrix singularity problem.

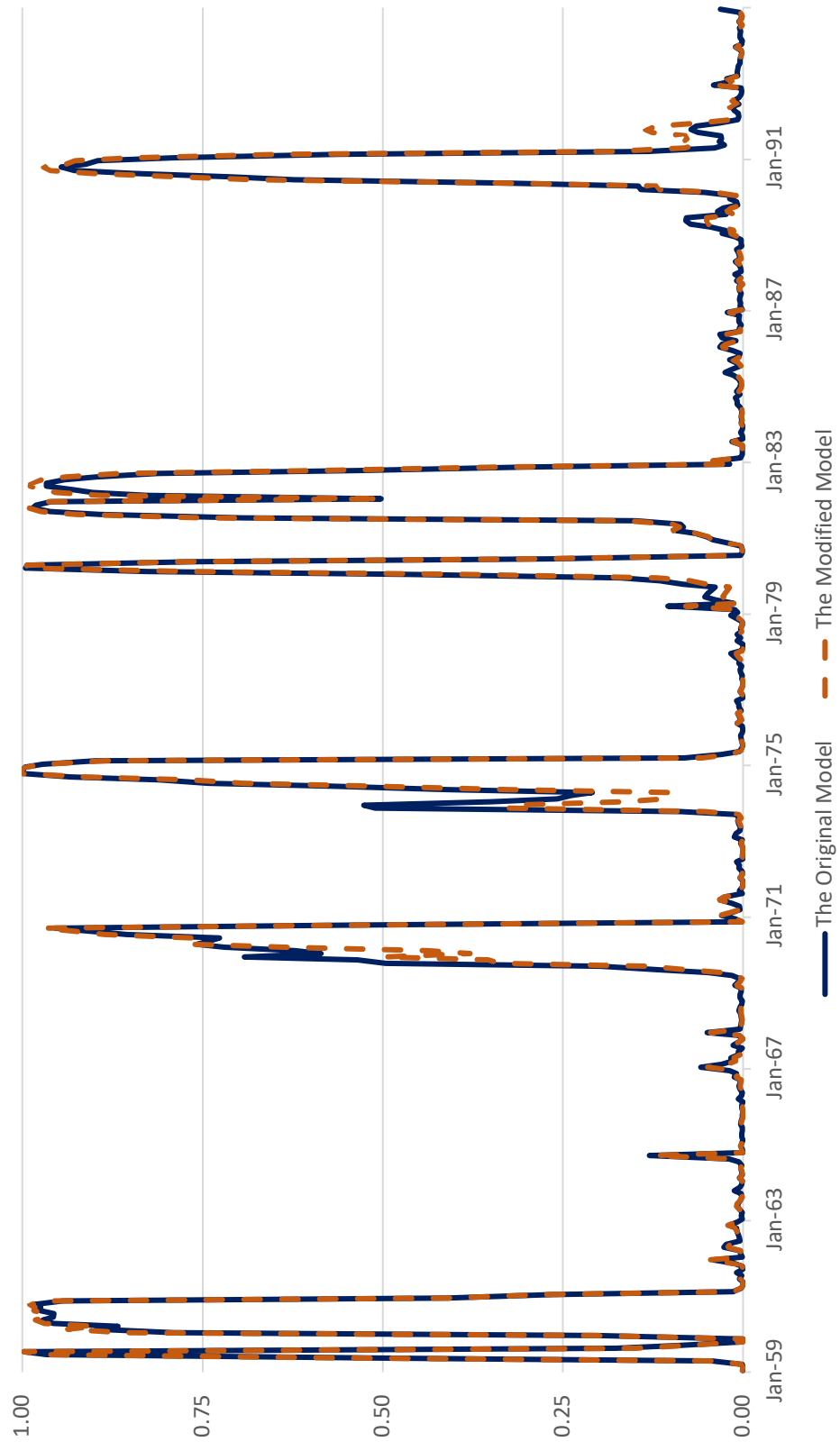


Figure 4.1: Smoothed Recession Probabilities: the Original Model versus the Modified Model

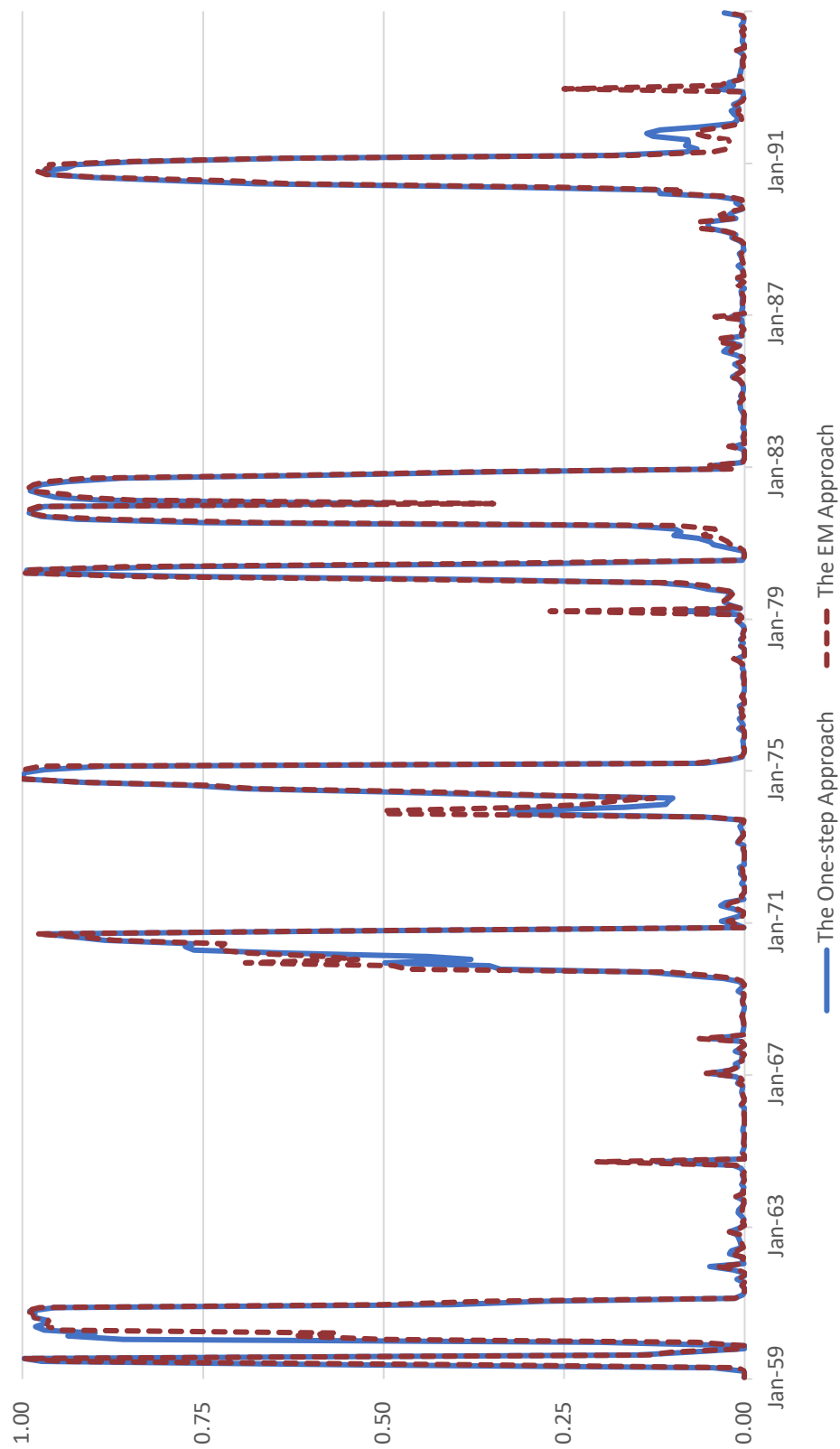


Figure 4.2: Smoothed Recession Probabilities: the One-step Approach versus the EM Approach

if there is any advantage of the EM approach it should be speed. This is indeed the case, as summarized in Table 4.2.² It will be shown that the computational advantage carries over to a setting with larger datasets in the next chapter.

Table 4.1: Estimation Time for the Two Estimation Approaches

Estimation Methodology	Total Computation Time (in Sec)
One-step Approach	56.1
EM Approach	41.0

4.2.2 Camacho et al. (2018)

This subsection replicates [Camacho et al. \(2018\)](#). The difference from the earlier replication exercise is that now I work with a mixed-frequency dataset consisting of monthly and quarterly series.

[Camacho et al. \(2015\)](#) and [Doz and Petronevich \(2016\)](#) propose an alternative two-step estimation method for regime-switching DFMs. The idea is to initialize the estimation with principal components (without imposing regime-switching) and then maximize the approximate likelihood treating the factors as observables (here with regime-switching). [Camacho et al. \(2018\)](#) extend this methodology to a setting with mixed-frequency data which I replicate in this subsection. Then, I compare this to the results from using the EM approach. As in the previous section, I consider a model with a measurement error because this facilitates computation for the EM approach without altering the results substantially. The comparison is based on the same initial values.

²This exercise, as well as other exercises in the chapter, is performed using Matlab2020b with i5 3.1 Ghz Processor and 8GB Ram. The data and the initial values for this particular exercise are from the replication package of [Kim and Nelson \(2006\)](#). See “<http://econ.korea.ac.kr/~cjkim/MARKOV/prgmlist.htm>” for the detail.

Let Y_t be an $N \times 1$ vector of observables

$$Y_t = \{y_{1,t}, y_{2,t}, \dots, y_{N,t}\}.$$

which enters the observation equation through

$$Y_t = CF_t + e_t$$

$$e_t = \delta_1 \varepsilon_t + \dots + \delta_m \varepsilon_{t-m} + \epsilon_t.$$

The unobservable common component f_t follows an autoregressive process

$$f_t = \mu_{st} + \phi_1 f_{t-1} + \dots + \phi_r f_{t-r} + v_t.$$

In the stacked DFM form,

$$y_t = CF_t + \varepsilon_t$$

$$F_t = \mu_{st} + AF_{t-1} + v_t$$

$$\varepsilon_t = \delta_1 \varepsilon_{t-1} + \dots + \delta_m \varepsilon_{t-m} + \epsilon_t$$

$$\epsilon_t \sim \mathcal{N}(0, R)$$

$$v_t \sim \mathcal{N}(0, Q).$$

In this exercise, the observable vector contains five elements: the industrial production index, non-farm payroll employment, personal income net of transfer payments, and real manufacturing and trade sales as the monthly indicators

and the real GDP as the quarterly indicator. The sample period is from April 1967 to March 2017. Following [Mariano and Murasawa \(2003\)](#), the quarterly data is approximated as in equation (3.79). Assuming the quarterly data is the first observable, i.e., $y_{1,t}$, the observation equation can be written as follows:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{5t} \end{bmatrix} = \begin{bmatrix} \frac{\gamma_{11}}{3} & \frac{2\gamma_{11}}{3} & \frac{3\gamma_{11}}{3} & \frac{2\gamma_{11}}{3} & \frac{\gamma_{11}}{3} & \frac{1}{3} & \frac{2}{3} & \frac{3}{3} & \frac{2}{3} & \frac{1}{3} & 0 & 0 \dots 0 & 0 \\ \gamma_{21} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \dots 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \gamma_{51} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \dots 1 & 0 \end{bmatrix} \begin{bmatrix} f_t \\ f_{t-1} \\ f_{t-2} \\ f_{t-3} \\ f_{t-4} \\ \varepsilon_{1,t} \\ \varepsilon_{1,t-1} \\ \varepsilon_{1,t-2} \\ \varepsilon_{1,t-3} \\ \varepsilon_{1,t-4} \\ \varepsilon_{2,t} \\ \varepsilon_{2,t-1} \\ \vdots \\ \varepsilon_{5,t} \\ \varepsilon_{5,t-1} \end{bmatrix}. \quad (4.8)$$

The transition equation is

$$\begin{bmatrix} f_t \\ f_{t-1} \\ f_{t-2} \\ f_{t-3} \\ f_{t-4} \\ \varepsilon_{1,t} \\ \varepsilon_{1,t-1} \\ \varepsilon_{1,t-2} \\ \varepsilon_{1,t-3} \\ \varepsilon_{1,t-4} \\ \varepsilon_{2,t} \\ \varepsilon_{2,t-1} \\ \vdots \\ \varepsilon_{5,t} \\ \varepsilon_{5,t-1} \end{bmatrix} = \begin{bmatrix} \mu_{st} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \phi_{1,1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \phi_{2,1} & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 \\ \vdots & 0 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & \phi_{5,1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{bmatrix} \begin{bmatrix} f_{t-1} \\ f_{t-2} \\ f_{t-3} \\ f_{t-4} \\ f_{t-5} \\ \varepsilon_{1,t-1} \\ \varepsilon_{1,t-2} \\ \varepsilon_{1,t-3} \\ \varepsilon_{1,t-4} \\ \varepsilon_{1,t-5} \\ \varepsilon_{2,t-1} \\ \varepsilon_{2,t-2} \\ \vdots \\ \varepsilon_{5,t-1} \\ \varepsilon_{5,t-2} \end{bmatrix} + \begin{bmatrix} v_{0,t} \\ 0 \\ 0 \\ 0 \\ 0 \\ \epsilon_{1,t} \\ 0 \\ 0 \\ 0 \\ 0 \\ \epsilon_{2,t} \\ 0 \\ \vdots \\ \epsilon_{5,t} \\ 0 \end{bmatrix} .(4.9)$$

Figure 4.3 plots the smoothed recession probabilities generated by the two approaches: the two-step approach of [Camacho et al. \(2018\)](#) (blue solid line) and the EM approach presented in Chapter 3 (red dashed line). I superimpose the NBER recession dates on top of these using the gray shaded areas.

The results show that two approaches mostly agree, which also closely resemble the NBER recession dates. Note that all of these are not real time estimates. In the next chapter, I will demonstrate how the regime-switching model, estimated with the EM approach, performs against the NBER dating in real time. As before, the EM approach comes with a speed advantage as shown in Table

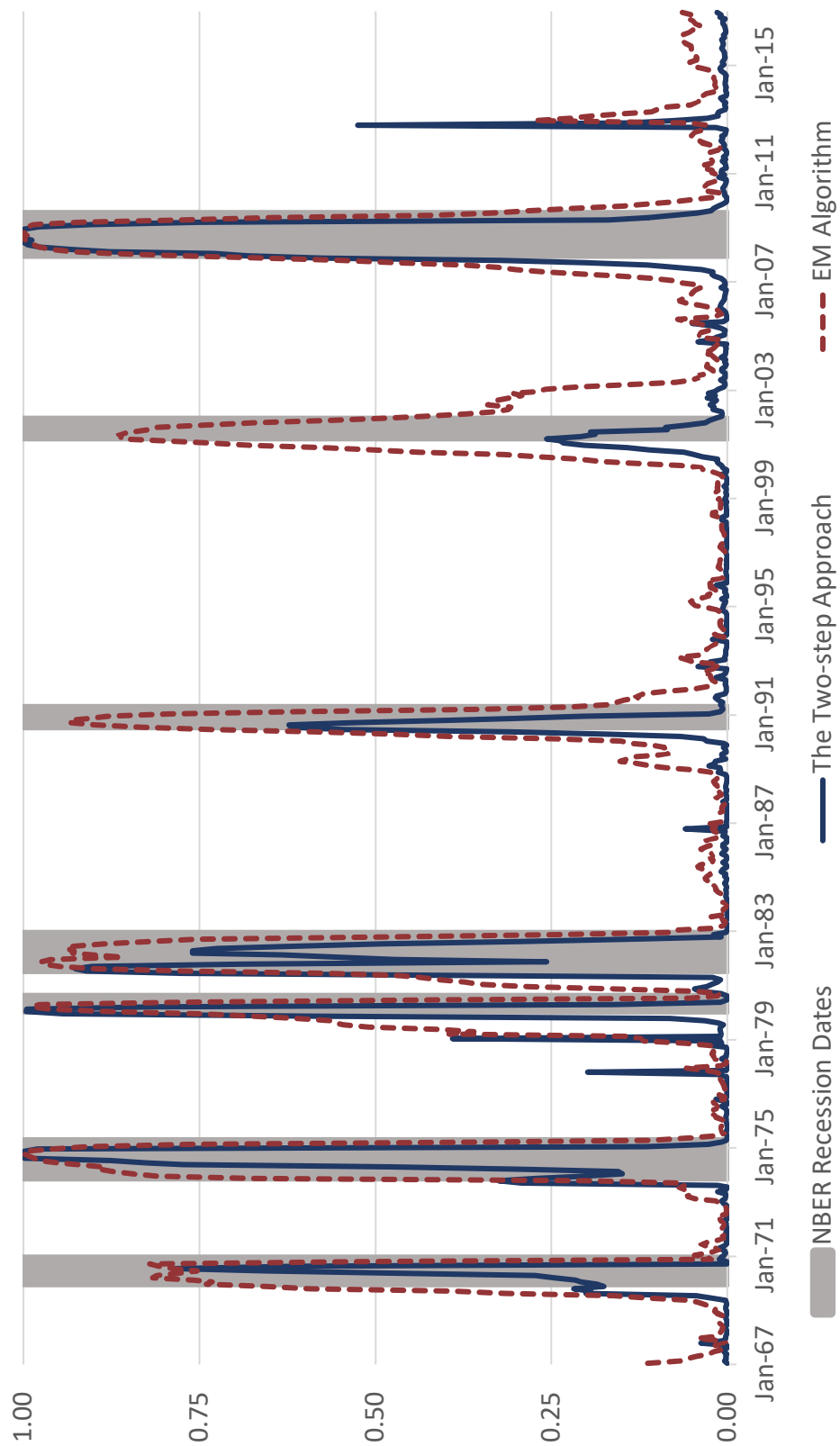


Figure 4.3: Smoothed Recession Probabilities: the Two-step Approach versus the EM Approach

Table 4.2: Estimation Time for the Two Estimation Approaches

Estimation Methodology	Total Computation Time (in Sec)
Two-step Approach	110.2
EM Approach	71.5

4.2.3 Monte Carlo Simulation

Sections 4.2.1 and 4.2.2 show that the EM approach presented in Chapter 3 performs similarly to other well-known approaches in the literature while promising speed advantages. In this subsection, I will conduct Monte Carlo simulations to test the accuracy of the approach.

The Monte Carlo experiment uses the following equations as the data generating mechanism:

$$y_t = CF_t + e_t$$

$$F_t = \mu[s_t] + AF_{t-1} + v_t.$$

To conduct the Monte Carlo experiment, I first generate a random sequence of regimes and then fix this. Then, I generate 1,000 simulations by drawing e_t and v_t from independent normal distributions, each simulation round having the sample length of 600. I do this for multiple sequences of regimes. In the interest of space, I present the results only for one specific sequence. However, in all other sequences I consider, I can successfully recover the true regime

sequences.

I consider three models of different scales:

- Small-scale Model: 4 monthly variables, 1 Quarterly variable
- Medium-scale Model: 8 monthly variables, 2 Quarterly variables
- Large-scale Model: 24 monthly variables, 4 Quarterly variables.

The observation equation for the small-scale model follows equation (4.10):

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{5t} \end{bmatrix} = \begin{bmatrix} \frac{\gamma_{11}}{3} & \frac{2\gamma_{11}}{3} & \frac{3\gamma_{11}}{3} & \frac{2\gamma_{11}}{3} & \frac{\gamma_{11}}{3} \\ \gamma_{21} & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \gamma_{51} & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} f_t \\ f_{t-1} \\ f_{t-2} \\ f_{t-3} \\ f_{t-4} \end{bmatrix} + \begin{bmatrix} e_{1,t} \\ e_{2,t} \\ \vdots \\ e_{5,t} \end{bmatrix}.$$

The medium-scale and large-scale models are modified accordingly. For the given values of Q and R matrices, the shocks are drawn from the following normal distributions:

$$e_t \sim \mathcal{N}(0, R)$$

$$v_t \sim \mathcal{N}(0, Q).$$

Figures 4.4 shows the smoothed regime probabilities, which are the simple average of simulations for each point in time, for the three models of different scales.

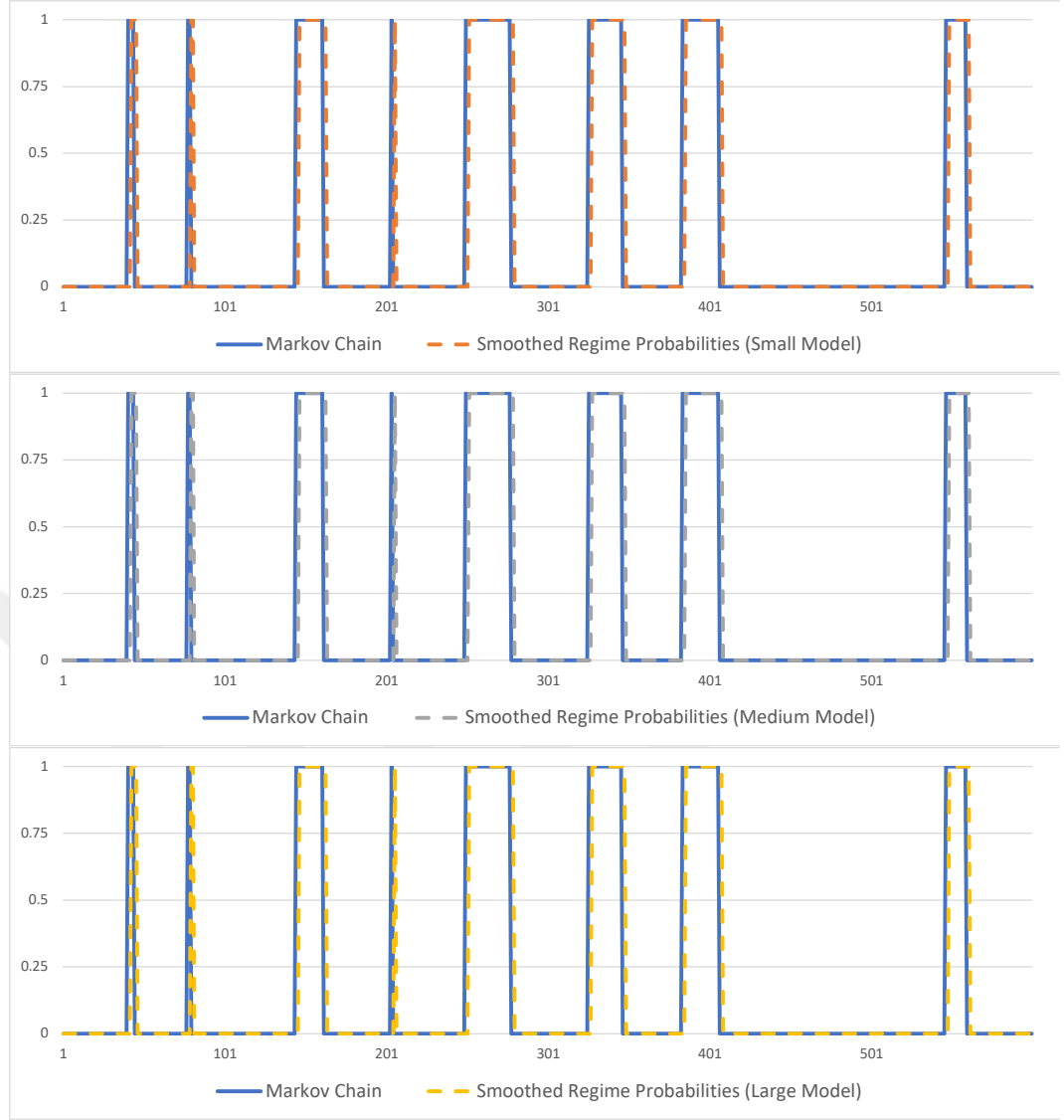


Figure 4.4: The True Regime Sequence and Three Model Estimates

The blue line gives the true regime sequence and the smoothed regime probabilities from the small-scale, medium-scale, and large-scale models are represented by the red, gray and yellow lines, respectively. The results indicate that the EM approach can successfully recover the true regime sequence.³

³Note that this exercise does not aim to recover true parameters. To be able to do so, I need to impose further identifying normalizations in the model. However, what is crucial for forecasting performance is the detection of true regimes. For this reason, this and next chapter exclusively focus on the regime detection.

CHAPTER 5

THE REGIME-SWITCHING NOWCASTING MODEL: AN APPLICATION TO THE US ECONOMY

5.1 Introduction

Academics and policymakers alike pay particular attention to the evolution of economic activities in real-time. This has led to a voluminous literature on nowcasting that was reviewed in Chapter 3. This chapter is concerned with nowcasting in a setting involving a potentially large dataset, which also comes in mixed frequencies and is subject to missing data problems. As discussed in Chapter 3, it is possible to apply the EM algorithm in this setting and doing so provides satisfactory performance relative to other estimation methods as shown in Chapter 4. One innovation relative to the existing nowcasting models, for instance those by [Bańbura and Modugno \(2014\)](#) and [Bok et al. \(2018\)](#), is that the model allows for regime-switching over time. [Çakmakli et al. \(2021\)](#) is a recent example of this class of models. Throughout this chapter, I will

focus on nowcasting real GDP and dating of recession in real-time for the US economy, but the application of this methodology is not limited to this particular dataset as well as this particular empirical exercise.

The chapter is organized as follows. In Section 5.2, I introduce the nowcasting model with regime-switching. In Section 5.3, I apply this to the US data, in real-time as the data releases happen, and demonstrate its performance relative to the NBER recession dating. Finally, I conclude in Section 5.4.

5.2 Model and Methodology

I first describe the New York FED nowcasting model which has produced popular nowcasts for the US economy until its suspension following the COVID-19 pandemic. I extend this model by allowing regime-switching, which makes it potentially applicable even in the sample including events involving large nonlinearities such as the recent pandemic. In the following subsection, I compare the performance of the single-regime New York FED model to my regime-switching model in predicting the current quarter US GDP. I also use this model to see whether it produces recession dating in real-time that is close to the NBER dating which becomes available with up to one year delay.

Recall that the single-regime DFM takes the form

$$\begin{aligned}
Y_t &= CF_t + e_t \\
F_t &= \mu + AF_{t-1} + v_t \\
e_t &= \delta_1 e_{t-1} + \delta_2 e_{t-2} + \dots + \delta_p e_{t-p} + \varepsilon_t \\
\varepsilon_t &\sim \mathcal{N}(0, R) \\
v_t &\sim \mathcal{N}(0, Q).
\end{aligned} \tag{5.1}$$

The regime-switching modification I consider allows regime-switching in the intercept of the transition equation for two regimes, either a recession or a normal time:¹

$$\begin{aligned}
Y_t &= CF_t + e_t \\
F_t &= \mu[s_t] + AF_{t-1} + v_t \\
e_t &= \delta_1 e_{t-1} + \delta_2 e_{t-2} + \dots + \delta_p e_{t-p} + \varepsilon_t \\
\varepsilon_t &\sim \mathcal{N}(0, R) \\
v_t &\sim \mathcal{N}(0, Q).
\end{aligned} \tag{5.2}$$

In the New York FED model, the common components are classified into four blocks, namely global, soft, real, and labor. The model on the New York FED website consists of 22 monthly indicators and three quarterly indicators.²

¹Because the data is typically standardized before being applied to the model, the intercept term in the single-regime case is restricted to be zero. In the case of a regime-switching model, it is the expected value of the intercept across regimes that is restricted to be zero.

²The model of [Bok et al. \(2018\)](#) uses 37 variables in which three of them are at quarterly frequency and 34 of them are at monthly frequency. Because some of the data are not publicly available, I only use the portion of the data that is publicly available. See

Every indicator loads on the global factor (hence the naming) and all other variables other than those related to the price level load on one additional factor. This is described in Table 5.1.

Table 5.1: Variables

Series Name	Frequency	Global Block	Soft Block	Real Block	Labor Block	Transformation	Category
Payroll Employment	m	1	0	0	1	chg	Labor
Job Openings	m	1	0	0	1	chg	Labor
Real GDP	q	1	0	1	0	pca	National Accounts
Consumer Price Index	m	1	0	0	0	pch	Prices
Durable Goods Orders	m	1	0	1	0	pch	Manufacturing
Retail Sales	m	1	0	1	0	pch	Retail and Consumption
Unemployment Rate	m	1	0	0	1	chg	Labor
Housing Starts	m	1	0	1	0	pch	Housing and Construction
Industrial Production	m	1	0	1	0	pch	Manufacturing
Personal Income	m	1	0	1	0	pch	National Accounts
Exports	m	1	0	1	0	pch	International Trade
Imports	m	1	0	1	0	pch	International Trade
Construction Spending	m	1	0	1	0	pch	Housing and Construction
Import Price Index	m	1	0	0	0	pch	International Trade
Core Consumer Price Index	m	1	0	0	0	pch	Prices
Core PCE Price Index	m	1	0	0	0	pch	Prices
PCE Price Index	m	1	0	0	0	pch	Prices
Building Permits	m	1	0	1	0	chg	Housing and Construction
Capacity Utilization Rate	m	1	0	1	0	chg	Manufacturing
Business Inventories	m	1	0	1	0	pch	International Trade
Unit Labor Cost	q	1	0	0	1	pca	Labor
Export Price Index	m	1	0	0	0	pch	International trade
Empire State Mfg Index	m	1	1	0	0	lin	Surveys
Real Consumption Spending	m	1	0	1	0	pch	Retail and Consumption
Philadelphia Fed Mfg Index	m	1	1	0	0	lin	Surveys

“<https://github.com/FRBNY-TimeSeriesAnalysis/Nowcasting>” for the detail.

I extended the dataset from the New York FED website both forward and backward by using real-time data from the ALFRED, i.e., the archival database website of St. Louis FED, with the total span of data starting in January 2000 and ending in March 2022.

I use the first seven years of the data to train the nowcasting models. For this reason, the nowcasting exercise starts from January 2007. In what follows, all nowcasts will be in real-time, i.e., for each data release.

5.3 Results and Discussion

The forecast target for the nowcasting exercise is the first flash estimate of the US real GDP. I evaluate the relative nowcasting performance of the single-regime New York FED model to my regime-switching model using the relative root mean square error (RMSE):

$$RMSE_{\text{relative}} = \frac{\sqrt{\sum_1^{\bar{T}} \left(G\hat{D}P_{\text{Regime Switching}} - GDP_{\text{First Estimate}} \right)^2}}{\sqrt{\sum_1^{\bar{T}} \left(G\hat{D}P_{\text{Bok et al. (2018)}} - GDP_{\text{First Estimate}} \right)^2}} \quad (5.3)$$

where $\bar{T}$ is the number of data releases in the sample.

Figure 5.1 gives the relative RMSE (red line) which is calculated based on the expanding-window subsample. The shaded areas are the real-time (smoothed) recession probabilities from the regime-switching model. The solid horizontal line is the benchmark level of one which means two models perform equally well, and the dashed horizontal line is the mean of the relative RMSE. Note that each point in the horizontal axis is a data release which is not as regular

as the calendar time.

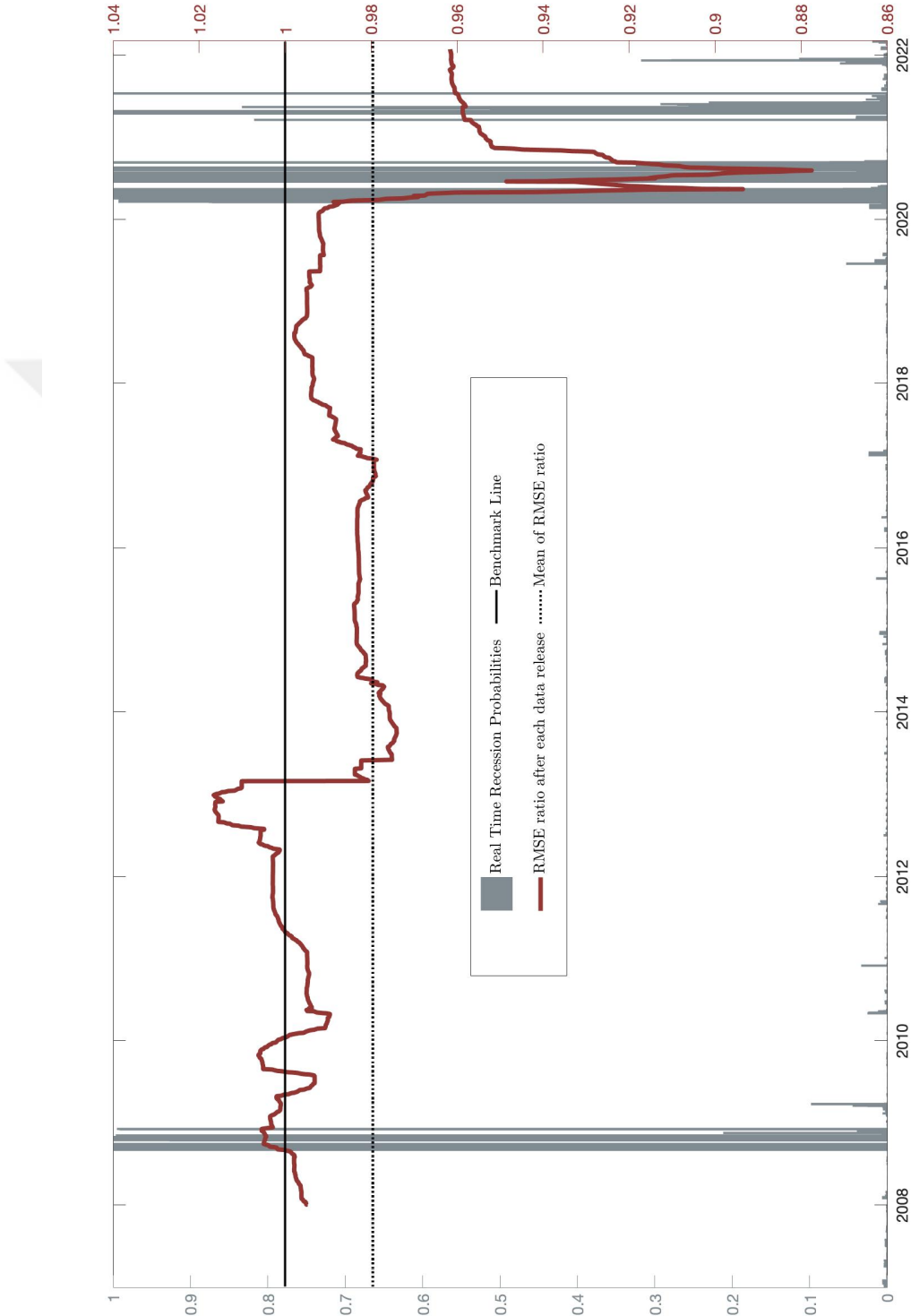


Figure 5.1: Expanding-window RMSE and Real Time Recession Probabilities

Recall that there are only two recessions in this sample, the first being the global financial crisis of 2008 and the second recent COVID-19 pandemic. Because the regime-switching model trains itself using the early observations, its forecasting performance is not substantially better relative to the single-regime New York FED model. However, one can see better relative forecasting performance starting 2013, which comes out very clearly when the COVID-19 pandemic starts with the RMSE dropping substantially.

Next, I break down Figure 5.1 release by release. Figure 5.2 shows that the regime-switching model has clear superior forecasting performance with the following indicators: nonfarm payroll employment, consumer price index, durable good orders, and unemployment rate. These are some of the most important indicators for monitoring the state of the economy and this is why it is not surprising to see that they are particularly informative for nowcasting real GDP.

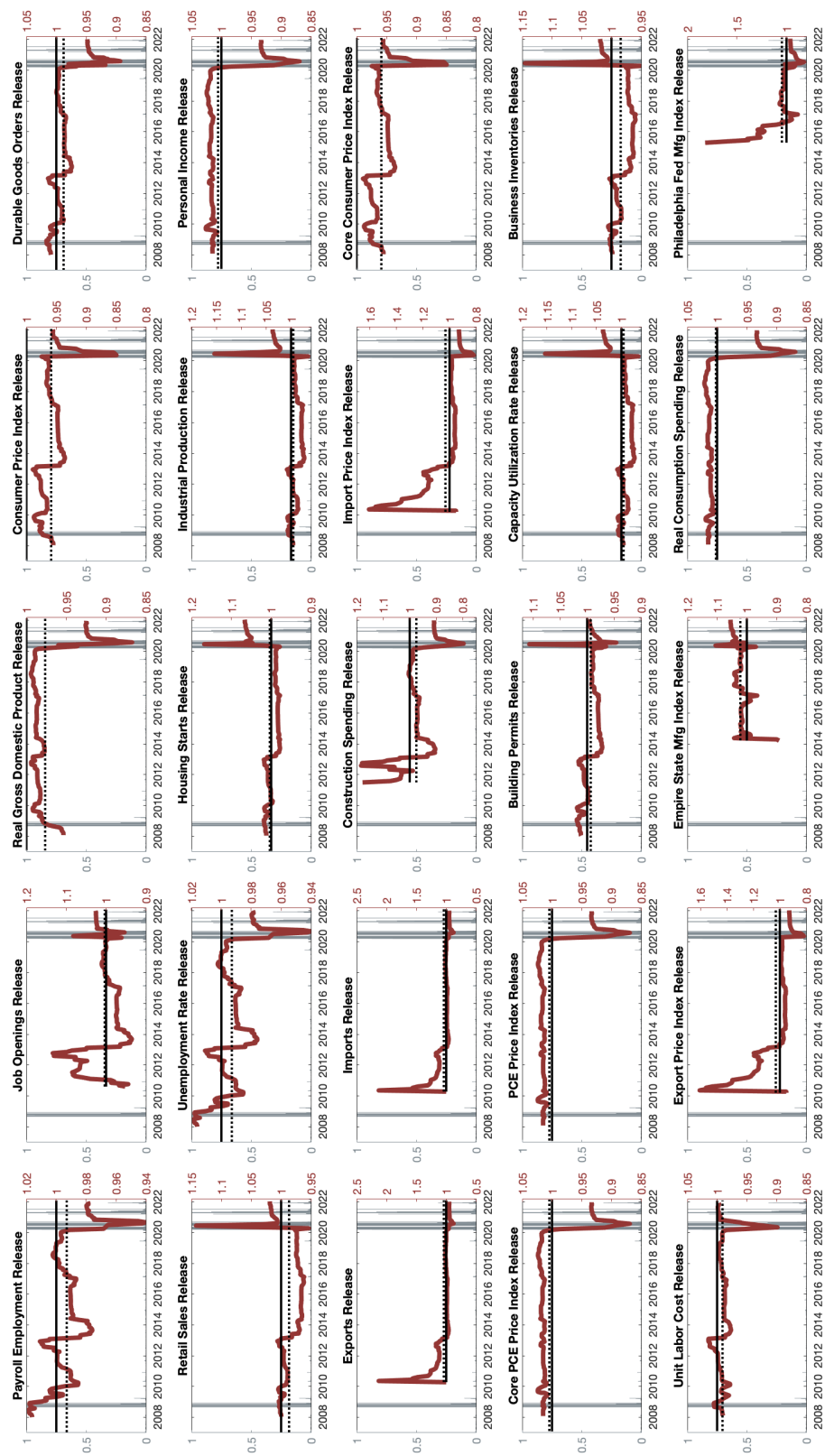


Figure 5.2: Expanding-window RMSE for Each Data Release and Real Time Recession Probabilities

Now, I compare the recession probabilities out of the regime-switching model to the recession dating of the NBER. The former is a real-time object based on the latest release of the data which by definition is based on less information than the latter because it becomes available with delays of a quarter to a year. For the both recessions in the US data, these two datings sufficiently overlap despite the information disadvantage of my model. This demonstrates the usefulness of the regime-switching nowcasting model for practical purposes.

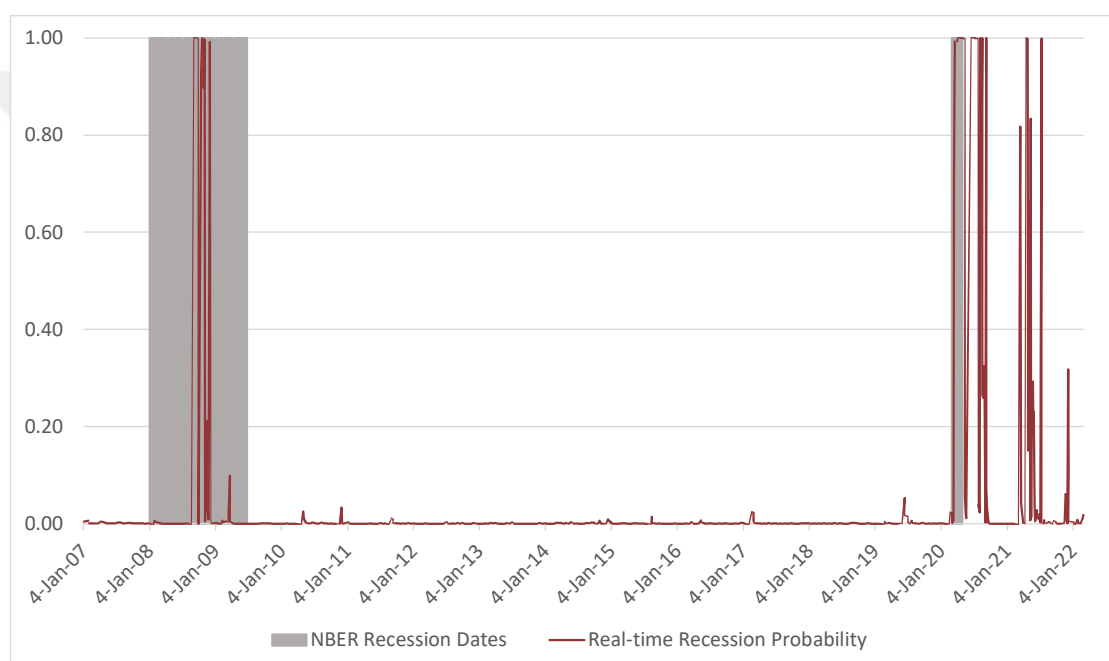


Figure 5.3: Real-time Recession Probabilities and NBER Recession Dates

Finally, for a fair comparison, I use the nowcasting model to backcast recessions. I allow using the information up to one year out relative to the time period being dated. Figure 5.4 shows that for the COVID-19 period, the model-based dating and the NBER dating of the recession become closer as more information becomes available. This indicates that the regime-switching nowcasting model holds a good promise as an alternative automated recession dating tool. The future research will extend the sample backwards and ex-

plore alternative model specifications to further improve the performance of the model. This defines my research agenda for years to come.

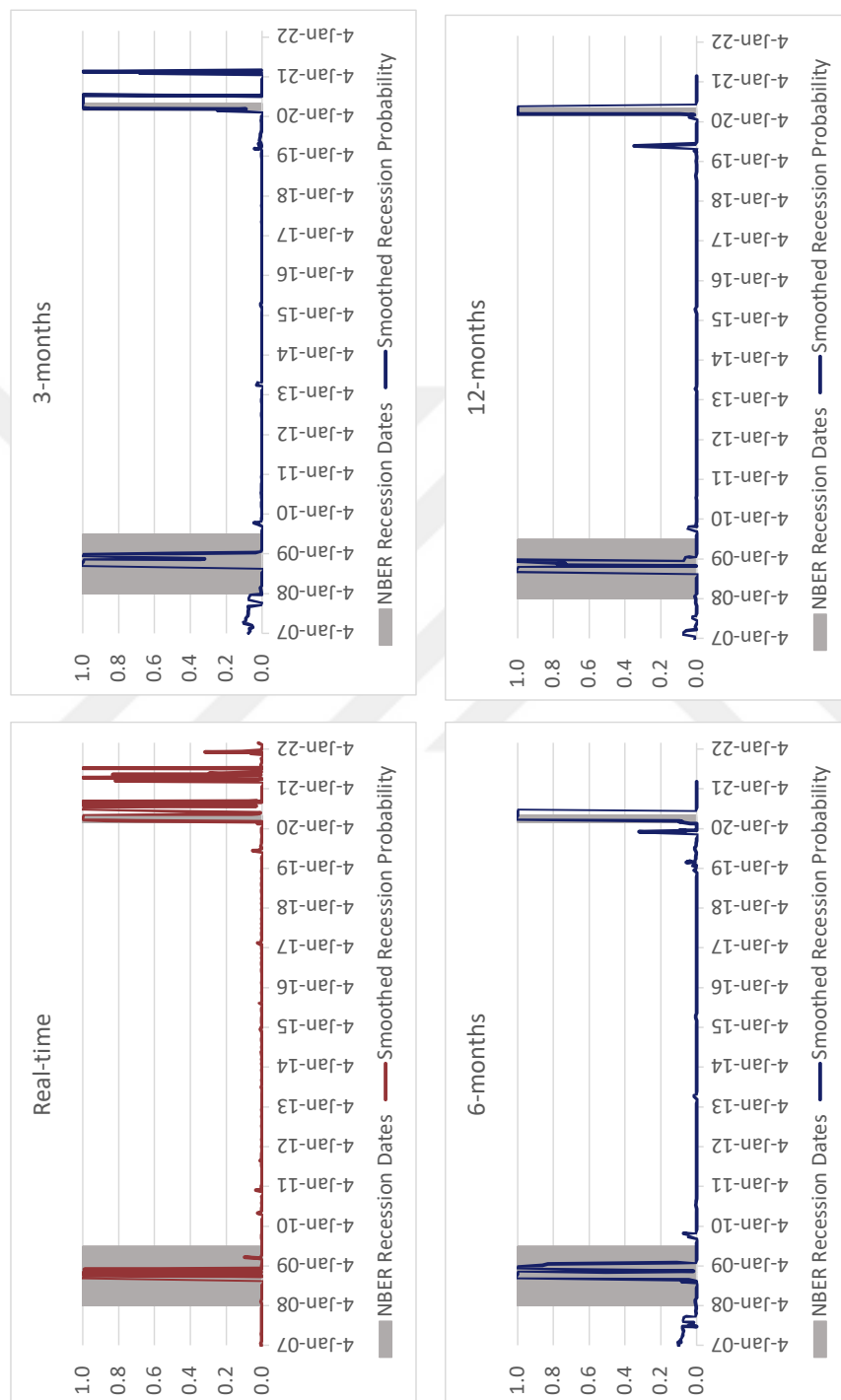


Figure 5.4: Backcasting of Recession Probabilities

5.4 Conclusion

This chapter puts together all the developments of the last two chapters in one place, which culminates to a regime-switching nowcasting model. The model is applied to nowcast the US real GDP in real-time, which as a by-product also provides real-time probabilistic recession dating. I demonstrate that the model outperforms the single-regime New York FED nowcasting model in relative terms and match the NBER recession dating closely. The model is flexible enough to be of use for other applications, for instance nowcasting GDP for other countries or other policy-relevant variables such as unemployment which is a key input for policymaking. These are the applications I will consider for future studies.

REFERENCES

- Angelini, E., Camba-Mendez, G., Giannone, D., Reichlin, L. and Rünstler, G. (2011), ‘Short-term forecasts of euro area gdp growth’.
- Bai, J. and Ng, S. (2002), ‘Determining the number of factors in approximate factor models’, *Econometrica* **70**(1), 191–221.
- Banbura, M., Giannone, D. and Reichlin, L. (2010), ‘Nowcasting’.
- Barhoumi, K., Darné, O. and Ferrara, L. (2010), ‘Are disaggregate data useful for factor analysis in forecasting french gdp?’, *Journal of Forecasting* **29**(1-2), 132–144.
- Batini, N., Kuttner, K. and Laxton, D. (2005), Does Inflation Targeting Work in Emerging Markets, World Economic Outlook (WEO) 4.
- Baxa, J., Horváth, R. and Vasíček, B. (2010), How Does Monetary Policy Change? Evidence on Inflation Targeting Countries, Working Papers wpdea1007, Department of Applied Economics at Universitat Autònoma of Barcelona.
URL: <https://ideas.repec.org/p/uab/wprdea/wpdea1007.html>
- Bañbura, M. and Modugno, M. (2014), ‘Maximum Likelihood Estimation Of Factor Models On Datasets With Arbitrary Pattern Of Missing Data’, *Journal of Applied Econometrics* **29**(1), 133–160.
URL: <https://ideas.repec.org/a/wly/japmet/v29y2014i1p133-160.html>
- Bernanke, B. S. and Boivin, J. (2003), ‘Monetary policy in a data-rich environment’, *Journal of Monetary Economics* **50**(3), 525–546.
URL: <https://ideas.repec.org/a/eee/moneco/v50y2003i3p525-546.html>
- Bernanke, B. S., Boivin, J. and Elias, P. (2005), ‘Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach’, *The Quarterly Journal of Economics* **120**(1), 387–422.
URL: <https://dx.doi.org/10.1162/0033555053327452>
- Bernanke, B. S., Laubach, T., Mishkin, F. S. and Posen, A. S. (1999), *Inflation Targeting: Lessons from the International Experience*, Princeton University Press.
URL: <http://www.jstor.org/stable/j.ctv301gdr>

- Bernardo, J., Bayarri, M., Berger, J., Dawid, A., Heckerman, D., Smith, A. and West, M. (2003), ‘Bayesian factor regression models in the “large p, small n” paradigm’, *Bayesian statistics* **7**, 733–742.
- Bessec, M. and Bouabdallah, O. (2015), ‘Forecasting gdp over the business cycle in a multi-frequency and data-rich environment’, *Oxford Bulletin of Economics and Statistics* **77**(3), 360–384.
- Bianchi, F. and Ilut, C. (2017), ‘Monetary/Fiscal Policy Mix and Agent’s Beliefs’, *Review of Economic Dynamics* **26**, 113–139.
- Boivin, J. and Giannoni, M. (2002), ‘Assessing changes in the monetary transmission mechanism: a VAR approach’, *Economic Policy Review* (May), 97–111.
URL: <https://ideas.repec.org/a/fip/fednep/y2002imayp97-111nv.8no.1.html>
- Boivin, J. and Giannoni, M. (2006), ‘Has monetary policy become more effective?’, *The Review of Economics and Statistics* **88**(3), 445–462.
URL: <https://EconPapers.repec.org/RePEc:tpr:restat:v:88:y:2006:i:3:p:445-462>
- Bok, B., Caratelli, D., Giannone, D., Sbordone, A. M. and Tambalotti, A. (2018), ‘Macroeconomic nowcasting and forecasting with big data’, *Annual Review of Economics* **10**(1), 615–643.
URL: <https://doi.org/10.1146/annurev-economics-080217-053214>
- Bollard, A. and Karagedikli, O. (2006), Inflation Targeting: The New Zealand Experience and Some Lessons, Inflation targeting: Performance and challenges” conference, hosted by the central bank of the republic of turkey.
- Buckle, R. A. (2020), Monetary policy governance and inflation targeting in new zealand, in ‘Public Policy and Governance Frontiers in New Zealand’, Emerald Publishing Limited.
- Buckle, R. A., Haugh, D. and Thomson, P. (2002), Growth and volatility regime switching models for New Zealand GDP data, Treasury Working Paper Series 02/08, New Zealand Treasury.
URL: <https://ideas.repec.org/p/nzt/nztwps/02-08.html>
- Buckle, R. A., Kim, K., Kirkham, H., McLellan, N. and Sharma, J. (2007), ‘A structural VAR business cycle model for a volatile small open economy’, *Economic Modelling* **24**(6), 990–1017.
URL: <https://ideas.repec.org/a/eee/ecmode/v24y2007i6p990-1017.html>
- Burns, A. F. and Mitchell, W. C. (1946), *Measuring business cycles*, number burn46-1, National bureau of economic research.
- Byron, M. Y., Shenoy, K. V. and Sahani, M. (2004), Derivation of kalman filtering and smoothing equations, in ‘Technical report’, Stanford University.

- Çakmakli, C., Demircani, H. and Altug, S. (2021), ‘Modelling of economic and financial conditions for real-time prediction of recessions’, *Oxford Bulletin of Economics and Statistics* **83**(3), 663–685.
- Camacho, M., Perez-Quiros, G. and Poncela, P. (2015), ‘Extracting nonlinear signals from several economic indicators’, *Journal of Applied Econometrics* **30**(7), 1073–1089.
- Camacho, M., Perez-Quiros, G. and Poncela, P. (2018), ‘Markov-switching dynamic factor models in real time’, *International Journal of Forecasting* **34**(4), 598–611.
URL: <https://ideas.repec.org/a/eee/intfor/v34y2018i4p598-611.html>
- Canova, F. and Gambetti, L. (2004), On the time variations of us monetary policy: Who is right?, Money macro and finance (mmf) research group conference 2004, Money Macro and Finance Research Group.
URL: <https://EconPapers.repec.org/RePEc:mmf:mmfc04:96>
- Cecchetti, S. G. and Ehrmann, M. (2002), Does Inflation Targeting Increase Output Volatility?: An International Comparison of Policymakers’ Preferences and Outcomes, in N. Loayza, K. Schmidt-Hebbel, N. L. S. Editor) and K. S.-H. (Series, eds, ‘Monetary Policy: Rules and Transmission Mechanisms’, Vol. 4 of *Central Banking, Analysis, and Economic Policies Book Series*, Central Bank of Chile, chapter 9, pp. 247–274.
URL: <https://ideas.repec.org/h/chb/bcchsb/v04c09pp247-274.html>
- Chauvet, M. (1998), ‘An econometric characterization of business cycle dynamics with factor structure and regime switching’, *International Economic Review* **39**(4), 969–996.
URL: <http://www.jstor.org/stable/2527348>
- Chib, S. (1998), ‘Estimation and comparison of multiple change-point models’, *Journal of Econometrics* **86**(2), 221–241.
URL: <https://ideas.repec.org/a/eee/econom/v86y1998i2p221-241.html>
- Clarida, R., Galí, J. and Gertler, M. (2000), ‘Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory’, *The Quarterly Journal of Economics* **115**(1), 147–180.
URL: <https://ideas.repec.org/a/oup/qjecon/v115y2000i1p147-180..html>
- Cogley, T. and Sargent, T. J. (2005), ‘Drift and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S’, *Review of Economic Dynamics* **8**(2), 262–302.
URL: <https://ideas.repec.org/a/red/issued/v8y2005i2p262-302.html>
- Corbo, V., Landerretche, O. and Schmidt-Hebbel, K. (2001), Assessing Inflation Targeting after a Decade of World Experience, Working Papers 51, Oesterreichische Nationalbank (Austrian Central Bank).
URL: <https://ideas.repec.org/p/onb/oenbwp/51.html>
- Creel, J. and Hubert, P. (2015), ‘Has inflation targeting changed the conduct of monetary policy?’, *Macroeconomic Dynamics* **19**(1), 1–21.

- Cushman, D. O. and Zha, T. (1997), ‘Identifying monetary policy in a small open economy under flexible exchange rates’, *Journal of Monetary Economics* **39**(3), 433–448.
- Degras, D., Ting, C.-M. and Ombao, H. (2022), ‘Markov-switching state-space models with applications to neuroimaging’, *Computational Statistics & Data Analysis* **174**, 107525.
- Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977), ‘Maximum likelihood from incomplete data via the em algorithm’, *Journal of the Royal Statistical Society: Series B (Methodological)* **39**(1), 1–22.
- Diebold, F. X. and Rudebusch, G. D. (1996), ‘Measuring Business Cycles: A Modern Perspective’, *The Review of Economics and Statistics* **78**(1), 67–77.
- Doz, C. and Fuleky, P. (2020), ‘Dynamic factor models’, *Macroeconomic forecasting in the era of big data* pp. 27–64.
- Doz, C., Giannone, D. and Reichlin, L. (2011), ‘A two-step estimator for large approximate dynamic factor models based on kalman filtering’, *Journal of Econometrics* **164**(1), 188–205.
- Doz, C. and Petronevich, A. (2016), Dating business cycle turning points for the french economy: An ms-dfm approach, *in* ‘Dynamic Factor Models’, Emerald Group Publishing Limited.
- Doz, C. and Petronevich, A. (2017), ‘On the consistency of the two-step estimates of the ms-dfm: A monte carlo study’.
- Dungey, M. and Fry, R. (2009), ‘The identification of fiscal and monetary policy in a structural VAR’, *Economic Modelling* **26**(6), 1147–1160.
URL: <https://ideas.repec.org/a/eee/ecmode/v26y2009i6p1147-1160.html>
- Dungey, M. and Pagan, A. (2000), ‘A Structural VAR Model of the Australian Economy’, *The Economic Record* **76**(235), 321–342.
URL: <https://ideas.repec.org/a/bla/ecorec/v76y2000i235p321-42.html>
- Fernandez-Villaverde, J., Guerron, P. and Rubio-Ramirez, J. F. (2010), Fortune or virtue: Time-variant volatilities versus parameter drifting in u.s. data, CEPR Discussion Papers 7813, C.E.P.R. Discussion Papers.
URL: <https://EconPapers.repec.org/RePEc:cpr:ceprdp:7813>
- Fischer, A. M. (1993), Inflation Targeting : The New Zealand and Canadian Cases, *Cato Journal* **13** 13.
- Forni, M., Giannone, D., Lippi, M. and Reichlin, L. (2009), ‘Opening the black box: Structural factor models with large cross sections’, *Econometric Theory* **25**(5), 1319–1347.
- Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2000), ‘The generalized dynamic-factor model: Identification and estimation’, *Review of Economics and statistics* **82**(4), 540–554.

- Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2005), 'The generalized dynamic factor model: one-sided estimation and forecasting', *Journal of the American statistical association* **100**(471), 830–840.
- Ftiti, Z. (2011), 'Taylor rule and inflation targeting: Evidence from new zealand', *International Business and Economics Research Journal (IBER)* **7**.
- Geweke, J. (1977), 'The dynamic factor analysis of economic time series', *Latent variables in socio-economic models*.
- Giannone, D., Reichlin, L. and Small, D. (2008), 'Nowcasting: The real-time informational content of macroeconomic data', *Journal of Monetary Economics* **55**(4), 665–676.
URL: <https://www.sciencedirect.com/science/article/pii/S0304393208000652>
- Gürkaynak, R., Sack, B. and Swanson, E. (2005), 'Do actions speak louder than words? the response of asset prices to monetary policy actions and statements', *International Journal of Central Banking* **1**(1), 55–93.
- Hamilton, J. D. (1989), 'A new approach to the economic analysis of nonstationary time series and the business cycle', *Econometrica: Journal of the econometric society* pp. 357–384.
- Harding, D. and Pagan, A. (2002), 'Dissecting the cycle: a methodological investigation', *Journal of monetary economics* **49**(2), 365–381.
- Hotelling, H. (1933), 'Analysis of a complex of statistical variables into principal components.', *Journal of educational psychology* **24**(6), 417.
- Huang, A., Margaritis, D. and Mayes, D. (2001), 'Monetary Policy Rules in Practice: Evidence from New Zealand', *Multinational Finance Journal* **5**(3), 175–200.
URL: <https://ideas.repec.org/a/mfj/journal/v5y2001i3p175-200.html>
- Kholodilin, K. A. et al. (2002a), 'Some evidence of decreasing volatility of the us coincident economic indicator', *Economics Bulletin* **3**(20), 1–20.
- Kholodilin, K. A. et al. (2002b), 'Two alternative approaches to modelling the nonlinear dynamics of the composite economic indicator', *Economics Bulletin* **3**(26), 1–18.
- Kim, C.-J. and Nelson, C. R. (1999), *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*, Vol. 1 of *MIT Press Books*, The MIT Press.
URL: <https://ideas.repec.org/b/mtp/titles/0262112388.html>
- Kim, C.-J. and Nelson, C. R. (2006), 'Estimation of a forward-looking monetary policy rule: A time-varying parameter model using ex post data', *Journal of Monetary Economics* **53**(8), 1949–1966.
URL: <https://ideas.repec.org/a/eee/moneco/v53y2006i8p1949-1966.html>

- Kim, S. and Roubini, N. (2000), ‘Exchange rate anomalies in the industrial countries: A solution with a structural VAR approach’, *Journal of Monetary Economics* **45**(3), 561–586.
URL: <https://ideas.repec.org/a/eee/moneco/v45y2000i3p561-586.html>
- Kuttner, K. N. (2004), A Snapshot of Inflation Targeting in its Adolescence, in C. Kent and S. Guttman, eds, ‘The Future of Inflation Targeting’, RBA Annual Conference Volume (Discontinued), Reserve Bank of Australia.
URL: <https://ideas.repec.org/h/rba/rbaacv/acv2004-02.html>
- Leeper, E. and Zha, T. (2003), ‘Modest policy interventions’, *Journal of Monetary Economics* **50**(8), 1673–1700.
URL: <https://EconPapers.repec.org/RePEc:eee:moneco:v:50:y:2003:i:8:p:1673-1700>
- Liu, Z., Waggoner, D. and Zha, T. (2011), ‘Sources of macroeconomic fluctuations: A regime-switching dsge approach’, *Quantitative Economics* **2**(2), 251–301.
URL: <https://EconPapers.repec.org/RePEc:ecm:quante:v:2:y:2011:i:2:p:251-301>
- Mariano, R. S. and Murasawa, Y. (2003), ‘A new coincident index of business cycles based on monthly and quarterly series’, *Journal of Applied Econometrics* **18**(4), 427–443.
- Matheson, T. D. (2010), ‘An analysis of the informational content of new zealand data releases: The importance of business opinion surveys’, *Economic Modelling* **27**(1), 304–314.
- McCallum, B. T. (1996), Inflation Targeting in Canada, New Zealand, Sweden, the United Kingdom, and in General, NBER Working Papers 5579, National Bureau of Economic Research, Inc.
URL: <https://ideas.repec.org/p/nbr/nberwo/5579.html>
- Mishkin, F. S. and Schmidt-Hebbel, K. (2001), One Decade of Inflation Targeting in the World: What Do We Know and What Do We Need to Know?, NBER Working Papers 8397, National Bureau of Economic Research, Inc.
URL: <https://ideas.repec.org/p/nbr/nberwo/8397.html>
- Modugno, M., Soybilgen, B. and Yazgan, E. (2016), ‘Nowcasting turkish gdp and news decomposition’, *International Journal of Forecasting* **32**(4), 1369–1384.
- Murphy, K. P. (1998), ‘Switching kalman filters’.
- Neumann, M. J. M. and Hagen, J. V. (2002), ‘Does inflation targeting matter?’, *Review* (Jul), 127–148.
URL: <https://ideas.repec.org/a/fip/fedlrv/y2002ijulp127-148nv.84no.4.html>
- New Zealand Drought Index and Drought Monitor Framework* (2017).

- Onatski, A. (2012), ‘Asymptotics of the principal components estimator of large factor models with weakly influential factors’, *Journal of Econometrics* **168**(2), 244–258.
- Pearson, K. (1901), ‘Principal components analysis’, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **6**(2), 559.
- Plantier, L. C. and Scrimgeour, D. (2002), Estimating a Taylor Rule for New Zealand with a time-varying neutral real rate, Reserve Bank of New Zealand Discussion Paper Series DP2002/06, Reserve Bank of New Zealand.
URL: <https://ideas.repec.org/p/nzb/nzbdps/2002-06.html>
- Primiceri, G. E. (2005), ‘Time Varying Structural Vector Autoregressions and Monetary Policy’, *Review of Economic Studies* **72**(3), 821–852.
URL: <https://ideas.repec.org/a/oup/restud/v72y2005i3p821-852.html>
- Reis, R. and Watson, M. W. (2010), ‘Relative goods’ prices, pure inflation, and the phillips correlation’, *American Economic Journal: Macroeconomics* **2**(3), 128–57.
- Rubio-Ramirez, J. F., Waggoner, D. and Zha, T. (2010), ‘Structural vector autoregressions: Theory of identification and algorithms for inference’, *Review of Economic Studies* **77**(2), 665–696.
URL: <https://EconPapers.repec.org/RePEc:oup:restud:v:77:y:2010:i:2:p:665-696>
- Sargent, T. J., Sims, C. A. et al. (1977), ‘Business cycle modeling without pretending to have too much a priori economic theory’, *New methods in business cycle research* **1**, 145–168.
- Sekine, T. and Teranishi, Y. (2008), Inflation Targeting and Monetary Policy Activism, IMES Discussion Paper Series 08-E-13, Institute for Monetary and Economic Studies, Bank of Japan.
URL: <https://ideas.repec.org/p/ime/imedps/08-e-13.html>
- Shumway, R. H., Stoffer, D. S. and Stoffer, D. S. (2016), *Time series analysis and its applications: with R examples Edition 4*, Vol. 3, Springer.
- Siliverstovs, B. and Kholodilin, K. A. (2012), ‘Assessing the real-time informational content of macroeconomic data releases for now-/forecasting gdp: Evidence for switzerland’, *Jahrbücher für Nationalökonomie und Statistik* **232**(4), 429–444.
- Sims, C. A., Waggoner, D. F. and Zha, T. (2008), ‘Methods for inference in large multiple-equation Markov-switching models’, *Journal of Econometrics* **146**(2), 255–274.
URL: <https://ideas.repec.org/a/eee/econom/v146y2008i2p255-274.html>
- Sims, C. A. and Zha, T. (2005), Were There Regime Switches in U.S. Monetary Policy?, Working Papers 92, Princeton University, Department of Economics, Center for Economic Policy Studies.
URL: <https://ideas.repec.org/p/pri/cepsud/110sims.pdf.html>

- Spearman, C. (1904), ‘General intelligence objectively determined and measured’, *American Journal of Psychology* **15**, 201–293.
- Stock, J. H. and Watson, M. W. (2002a), ‘Forecasting using principal components from a large number of predictors’, *Journal of the American statistical association* **97**(460), 1167–1179.
- Stock, J. H. and Watson, M. W. (2002b), ‘Macroeconomic forecasting using diffusion indexes’, *Journal of Business & Economic Statistics* **20**(2), 147–162.
- Stock, J. H. and Watson, M. W. (2010), ‘Indicators for dating business cycles: Cross-history selection and comparisons’, *American Economic Review* **100**(2), 16–19.
- Stock, J. H. and Watson, M. W. (2014), ‘Estimating turning points using large data sets’, *Journal of Econometrics* **178**, 368–381.
- Stock, J. H. and Watson, M. W. (2016), Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics, in ‘Handbook of macroeconomics’, Vol. 2, Elsevier, pp. 415–525.
- Svensson, L. E. (2010), Chapter 22 - inflation targeting, Vol. 3 of *Handbook of Monetary Economics*, Elsevier, pp. 1237–1302.
URL: <https://www.sciencedirect.com/science/article/pii/B9780444534545000104>
- Wesselbaum, D. (2014), Fiscal and monetary policy interactions in new zealand, Mpra paper, University Library of Munich, Germany.
URL: <https://EconPapers.repec.org/RePEc:pra:mpapa:58763>
- Zha, T. (1999), ‘Block recursion and structural vector autoregressions’, *Journal of Econometrics* **90**(2), 291–316.
URL: <https://ideas.repec.org/a/eee/econom/v90y1999i2p291-316.html>
- Zhou, T. and Shumway, R. (2008), ‘One-step approximations for detecting regime changes in the state space model with application to the influenza data’, *Computational Statistics Data Analysis* **52**(5), 2277–2291.
URL: <https://www.sciencedirect.com/science/article/pii/S0167947307003325>

APPENDICES

“

A Appendix for Chapter 2

The second chapter, which is accepted for publication at the Singapore Economic Review, satisfies the publication requirement for completion of the degree. This appendix is the online appendix for the published chapter that provides the information about the dataset, methodology, and robustness test that are mentioned but not included in the main text. In the first subsection, the data sources and transformation are discussed in detail. In the following subsection, the model structure is presented in detail. Next subsection proves Proposition 1 using Theorem 1 of [Rubio-Ramirez et al. \(2010\)](#). The fourth subsection gives robustness tests for potential omitted variable biases and also tests whether exogenous climate indicators provide useful additional information to the model.

New Zealand's exchange rate regime changed from a crawling-pegged exchange rate system to a completely floating exchange rate system. Therefore, the fifth subsection checks the main results and discusses an alternative setup where the shock variance regime of exchange rate is defined by an independent Markov chain. In addition to complementing the idea proposed in the fifth subsection, the last subsection checks whether the results change if a trade-weighted proxy for the rest of the world is used instead of the United States to model the rest of the world. The results are the same.

Data sources, description and transformations

This subsection first defines and describes the data used in the estimation of the model. Then, the data source and data transformations are stated. Finally, for each time-series, the raw and transformed data are visualized. Recall that the observable vector Y_t includes the domestic block, the international block, the exchange rate block, and the climate block that is assumed to be exogenous. Observables include domestic GDP (y_t), policy rate (i_t) and CPI inflation (π_t) in the domestic block, global production (y_t^*), global interest rate (i_t^*) and the world oil price (wop_t) in the international block, exchange rate (e_t) in the exchange rate block and climate conditions (cl_t) in the exogenous block.

The data used in this study cover the period between 1974:Q1 and 2018:Q3 at quarterly frequency. Unless otherwise noted, the data are downloaded from the St. Louis FED's database FRED. I consider transitory deviations from the trend as cycles which is a common practice in the literature. For this reason, all time series except the climate indicator are detrended from their long-run trends, which are estimated by the Hodrick-Prescott filter with the smoothing parameter of $\lambda=1600$. The climate indicator is demeaned relative to the sample average. The all data except for the climate indicator are log-transformed.

The domestic block consists of New Zealand's GDP (y_t), CPI inflation (π_t), interest rate (i_t), and exchange rate (e_t). New Zealand's GDP is measured by annual total gross domestic product at constant prices with the base year of 2015. The real GDP values are log-transformed and seasonally adjusted (using an ARIMA model) before the detrending procedure. The CPI inflation of New Zealand measures the price level changes with the base year of 2015. CPI inflation is taken as a year-to-year inflation at quarterly frequency. The data is seasonally adjusted and detrended before being used in the model. The policy tool that the Reserve Bank of New Zealand (RBNZ) uses is the official cash rate (OCR). The OCR is announced eight times a year since 1999 and

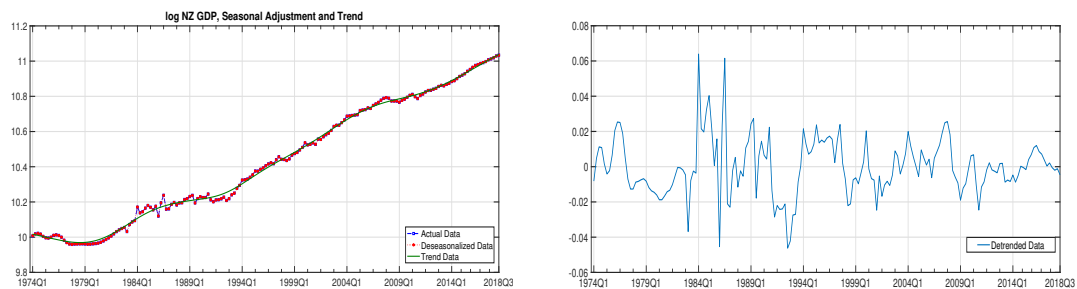


Figure 5: New Zealand GDP

seven times a year since 2016. However, the actual interest rate that the RBNZ targets is the 90-day bank bill rate, which is a better proxy for the policy.

I use the GDP index LORSGPORNZQ661S with the base year 2015. Figure 5 visualizes raw data, seasonal adjustment, trend, and detrended data.

The New Zealand CPI inflation series is NZLCPIALLQINMEI. Figure 6 visualizes raw data, seasonal adjustment, trend, and detrended data.

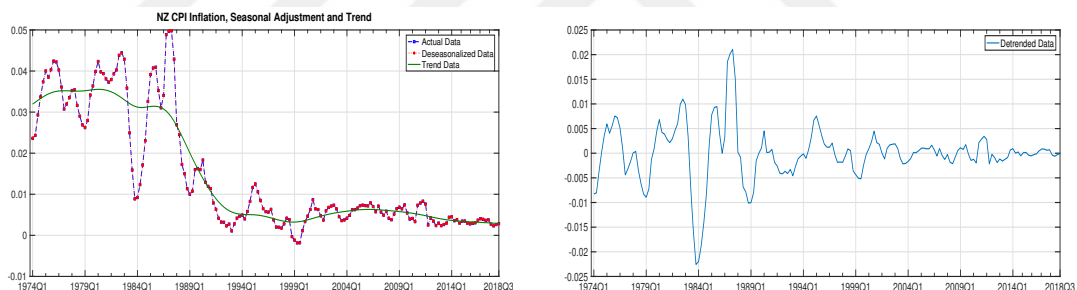


Figure 6: New Zealand CPI Inflation

The monetary policy response is measured using the 90-day bank bill rate. The data series is IR3TBB01NZQ156N. Figure 7 visualizes raw data, seasonal adjustment, trend, and detrended data.

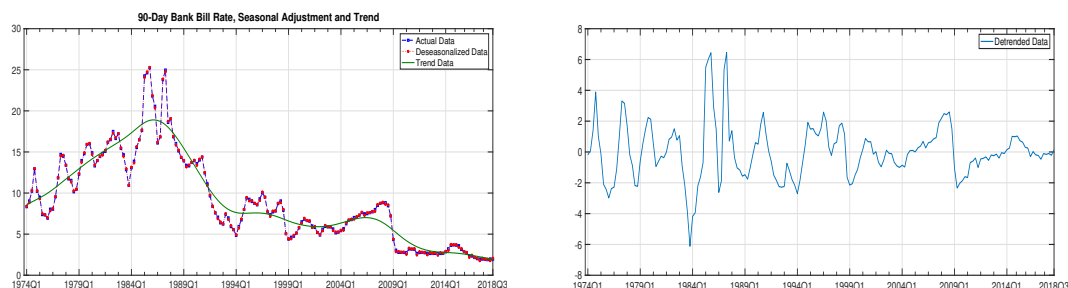


Figure 7: New Zealand 90-day Bank Bill Rate

The international block is modeled with the US data where the international demand (y_t^*) is proxied by US real output, and the international interest rate (i_t^*) is the federal funds rate. Considering US demand and interest rates as international indicators for New Zealand is both adequate in terms of New Zealand macroeconomic conditions and complementary to the exchange rate indicator defined above. The federal funds rate (i_t^*) is a three-month rate averaged within a quarter. The real US GDP (y_t^*) is the annualized quarterly constant price gross domestic product with the base year of 2012. The series code is GDPC1. Figure 8 visualizes raw data, seasonal adjustment, trend and detrended data.

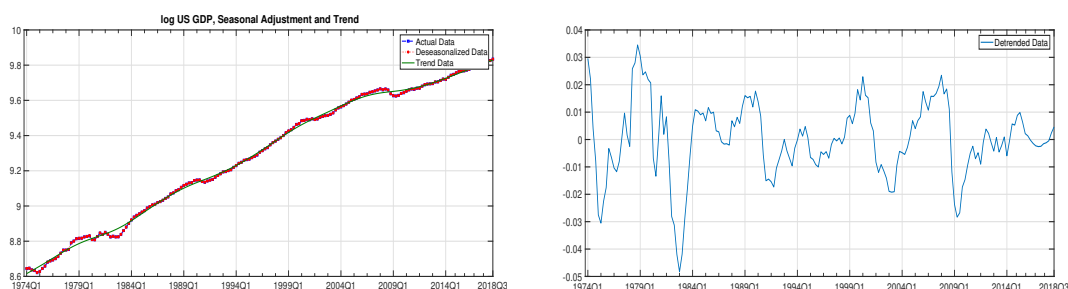


Figure 8: US GDP

The federal funds rate series is FEDFUNDS. Figure 9 visualizes the raw data, seasonal adjustment, trend and detrended data.

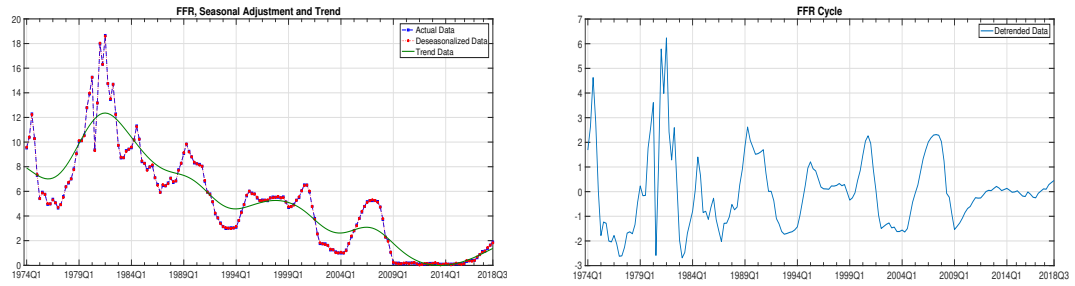


Figure 9: Federal Funds Rate

The world oil price (wop_t) is the Brent crude oil import prices in US dollars. Crude oil import prices are monthly series taken from the OECD database and averaged within a quarter. Figure 10 visualizes raw data, seasonal adjustment, trend, and detrended data.

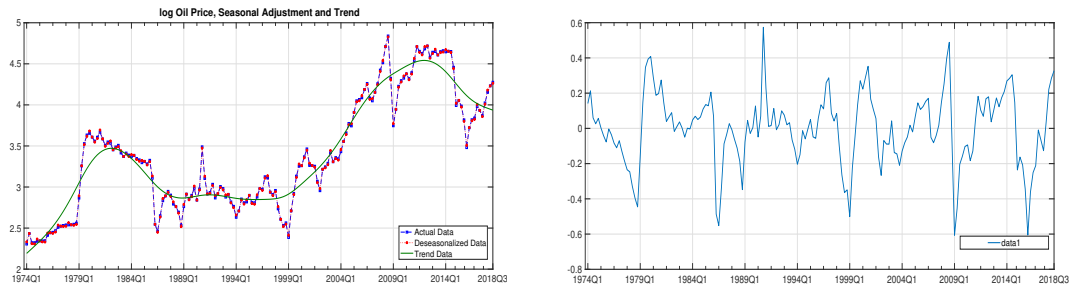


Figure 10: Crude Oil Import Prices

The exchange rate (e_t) is the national currency to US dollar exchange rate value at the end of the quarter. The exchange rate series is CCUSMA02NZQ618N. Figure 11 visualizes raw data, seasonal adjustment, trend and detrended data.

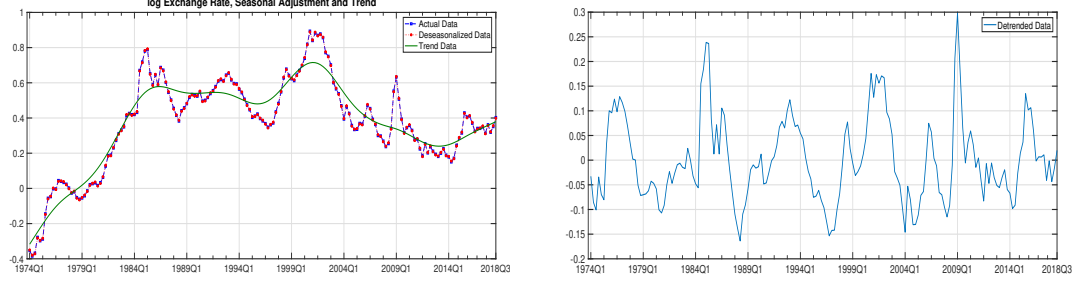


Figure 11: New Zealand Currency to US Dollar

The climate factor (cl_t) is soil moisture where potential evapotranspiration deficit (PED) serves as a measure of soil moisture.³ The climate moisture indicator PDE data are taken from *New Zealand Drought Index and Drought Monitor Framework* (2017), and the long-run trend is modeled by the sample mean. Figure 12 visualizes the raw data, seasonal adjustment, trend, and detrended data.

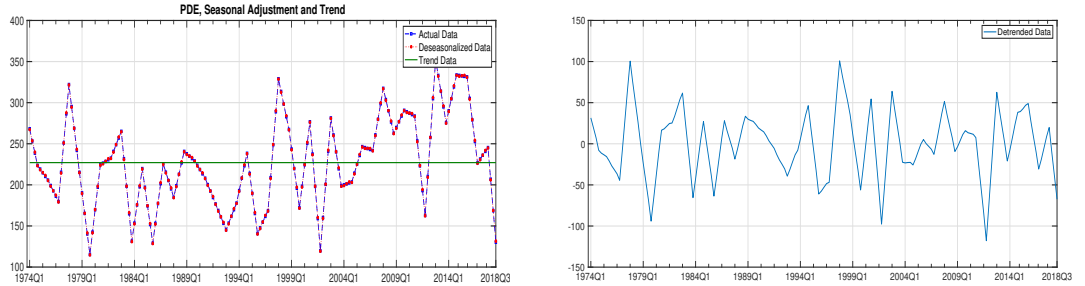


Figure 12: Potential Evapotranspiration Deficit

Model Description

The equations in (2.4) describe the structure of the model and the assumptions for the structural identification. This subsection describes the economic

³A clear explanation of the relation between the actual soil moisture deficit (SMD) and potential evapotranspiration deficit (PED) is provided in the *New Zealand Drought Index and Drought Monitor Framework* (2017) report prepared by National Institute of Water and Atmospheric Research of New Zealand.

intuition for these.

New Zealand's real GDP (y_t) is assumed to be both contemporaneously and with lags affected by climate conditions (cl_t). Therefore, any external shock to the climate conditions will immediately affect the level of production. The effects of international production (y_t^*), domestic inflation (π_t), world oil price (wop_t), exchange rate (e_t) and interest rate (i_t) appear with lag. In the original model of [Buckle et al. \(2007\)](#), the effects of interest rate and price changes work through a demand channel. This model does not feature a standalone demand channel and assumes that i_t and π_t affect the output with lags. The effect of climate on New Zealand business cycle dynamics has been widely studied in the literature. [Buckle et al. \(2007\)](#) and [Dungey and Fry \(2009\)](#) explain the importance of soil moisture in agricultural production as well as energy production. [Kim and Roubini \(2000\)](#) includes supply shocks in the model economy through world oil prices, where small open economies are directly affected by the cyclical behavior of the Brent crude. Equation (4) shows how all of these affect the dynamics of New Zealand's real GDP.

$$y_t = k_{11}cl_t + k_{12}y_{t-i}^* + k_{13}cl_{t-i} + k_{14}wop_{t-i} + k_{15}y_{t-i} + \dots \quad (4)$$

$$\dots + k_{16}\pi_{t-i} + k_{17}e_{t-i} + k_{18}i_{t-i} + \varepsilon_{t,1}$$

The inflation level in New Zealand (π_t) is both contemporaneously and with lags effected by world oil price (wop_t) and the exchange rate (e_t). Climate conditions (cl_t), domestic demand (y_t) and the response of the monetary policy (i_t) are assumed to be working with a delay. The effect of monetary policy on inflation is observed with lag through the simplified demand channel. These assumptions are realistic to the extent that monetary authority does not observe current period data. Equation (5) describes the dynamics of New

Zealand's CPI inflation.

$$\begin{aligned}\pi_t = & k_{21}wop_t + k_{22}e_t + k_{23}cl_{t-i} + k_{24}wop_{t-i} + k_{25}y_{t-i} + \dots \\ & \dots + k_{26}\pi_{t-i} + k_{27}e_{t-i} + k_{28}i_{t-i} + \varepsilon_{t,2}\end{aligned}\quad (5)$$

The monetary policy reaction function captures the mandate of the central bank. Following the early works of [Huang et al. \(2001\)](#), [Plantier and Scrimgeour \(2002\)](#), and [Ftiti \(2011\)](#), a Taylor-type rule is assumed to approximate New Zealand's policy reaction function. Monetary policy authority responds to the inflation (π_t) and the federal funds rate (i_t^*) contemporaneously and with lags. The policymaker also considers the real domestic production (y_t), the exchange rate (e_t) and inflation (π_t) with lags. Thus, the policy reaction function that responds to the price level, the output, the policy rate, and international conditions. Equation (6) describes the monetary policy rule.

$$i_t = k_{31}\pi_t + k_{32}i_t^* + k_{33}\pi_{t-i} + k_{34}i_{t-i}^* + k_{35}y_{t-i} + k_{36}i_{t-i} + k_{37}e_{t-i} + \varepsilon_{t,3} \quad (6)$$

The international block consists of US GDP (y_t^*), the federal funds rate (i_t^*) and the world oil price (wop_t). New Zealand's domestic economy is assumed to have no effect on the international block. The US GDP is assumed to be affected by lagged US GDP, world oil price, and international interest rate values. Equation (7) describes the dynamics of the US GDP.

$$y_t^* = k_{41}y_{t-i}^* + k_{42}i_{t-i}^* + k_{43}wop_{t-i} + \varepsilon_{t,4} \quad (7)$$

The world nominal interest rate (i_t^*), proxied by the federal funds rate, is assumed to be affected by US GDP (y_t^*) and world oil price (wop_t) contemporaneously and with lags. The effect of the federal funds rate (i_t^*) is observed with lags. Equation (8) describes the world interest rate.

$$i_t^* = k_{51}y_t^* + k_{52}wop_t + k_{53}wop_{t-i} + k_{54}i_{t-i}^* + k_{55}y_{t-i}^* + \varepsilon_{t,5} \quad (8)$$

The world oil price is assumed to be contemporaneously affected by the rest of the world's production (y_t^*) and its lags. It also depends on the lagged oil prices. Equation (9) describes the world oil price.

$$wop_t = k_{61}y_t^* + k_{62}wop_{t-i} + k_{63}y_{t-i}^* + \varepsilon_{t,6} \quad (9)$$

The exchange rate (e_t) is assumed to be affected by all variables, both contemporaneously and with lags, except for the climate conditions (cl_t). The domestic climate (cl_t) is assumed to evolve independently of the other two blocks, i.e., the domestic and international blocks, with an autoregressive process.

The proof of Proposition 1

This subsection builds on *Theorem 1* of [Rubio-Ramirez et al. \(2010\)](#) to prove Proposition 1. [Rubio-Ramirez et al. \(2010\)](#) accommodates both recursive and non-recursive identifications which makes their framework more generally applicable. The first part of the proof describes the model parameters and their mapping to *Theorem 1* of [Rubio-Ramirez et al. \(2010\)](#). The second part completes the proof by demonstrating that the relevant rank conditions hold for each variable. The idea is to show that conditional on the regime, structural shocks are identified. Since the regime-switching process is independent of other shocks, the regime by regime identification is sufficient for my purpose.

The general class of SVARs can be represented as

$$y_t A_0 = x_t A_+ + \varepsilon_t'$$

where y_t is the vector of observables, A_0 is the contemporaneous relation matrix, and

$$x_t = \begin{bmatrix} y_{t-1} \\ y_{t-2} \end{bmatrix}$$

is the stacked form of the lagged observations with the lag length $p = 2$ which is the selected lag length for the analysis. Define

$$A_+ = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$$

as the stacked form of the coefficient matrices for the lagged observable vectors x_t . Finally, the number of variables considered here is $n = 8$. For the given zero restrictions on the contemporaneous and lagged matrices, the matrices A_0 and A_+ are given in the main text as follows:

$$A_0 = \begin{bmatrix} a_{1,1} & 0 & 0 & a_{1,4} & a_{1,5} & 0 & 0 & a_{1,8} \\ 0 & a_{2,2} & 0 & 0 & 0 & a_{2,6} & 0 & 0 \\ a_{3,1} & 0 & a_{3,3} & a_{3,4} & 0 & 0 & 0 & a_{3,8} \\ 0 & 0 & 0 & a_{4,4} & 0 & 0 & a_{4,7} & a_{4,8} \\ 0 & 0 & 0 & 0 & a_{5,5} & 0 & a_{5,7} & a_{5,8} \\ 0 & 0 & 0 & 0 & 0 & a_{6,6} & 0 & a_{6,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & a_{7,7} & a_{7,8} \\ 0 & 0 & 0 & 0 & a_{8,5} & 0 & 0 & a_{8,8} \end{bmatrix} \quad (10)$$

$$A_+ = \begin{bmatrix} b_{1,1} & 0 & b_{1,3} & b_{1,4} & b_{1,5} & b_{1,6} & 0 & b_{1,8} \\ 0 & b_{2,2} & 0 & 0 & b_{2,5} & b_{2,6} & 0 & 0 \\ b_{3,1} & 0 & b_{3,3} & b_{3,4} & 0 & b_{3,6} & 0 & b_{3,8} \\ 0 & 0 & b_{4,3} & b_{4,4} & 0 & 0 & b_{4,7} & b_{4,8} \\ 0 & 0 & 0 & 0 & b_{5,5} & b_{5,6} & b_{5,7} & b_{5,8} \\ 0 & 0 & 0 & 0 & b_{6,5} & b_{6,6} & b_{6,7} & b_{6,8} \\ 0 & 0 & 0 & 0 & b_{7,5} & b_{7,6} & b_{7,7} & b_{7,8} \\ 0 & 0 & 0 & 0 & b_{8,5} & b_{8,6} & b_{8,7} & b_{8,8} \\ c_{1,1} & 0 & c_{1,3} & c_{1,4} & c_{1,5} & c_{1,6} & 0 & c_{1,8} \\ 0 & c_{2,2} & 0 & 0 & c_{2,5} & c_{2,6} & 0 & 0 \\ c_{3,1} & 0 & c_{3,3} & c_{3,4} & 0 & c_{3,6} & 0 & c_{3,8} \\ 0 & 0 & c_{4,3} & c_{4,4} & 0 & 0 & c_{4,7} & c_{4,8} \\ 0 & 0 & 0 & 0 & c_{5,5} & c_{5,6} & c_{5,7} & c_{5,8} \\ 0 & 0 & 0 & 0 & c_{6,5} & c_{6,6} & c_{6,7} & c_{6,8} \\ 0 & 0 & 0 & 0 & c_{7,5} & c_{7,6} & c_{7,7} & c_{7,8} \\ 0 & 0 & 0 & 0 & c_{8,5} & c_{8,6} & c_{8,7} & c_{8,8} \end{bmatrix} \quad (11)$$

Let $k \times n$ matrix ($k = 3 \times n$ by construction)

$$f(A_0, A_+) = \begin{bmatrix} A_0 \\ A_+ \end{bmatrix}$$

be the function that stacks A_0 and A_+ . Let e_j be the j^{th} column of the $n \times n$ identity matrix. Define $k \times k$ matrix Q_j as the linear restrictions basis matrix where

$$Q_j f(A_0, A_+) e_j = 0, \forall j \in 1 \leq j \leq n$$

Proof. According to *Theorem 1* of [Rubio-Ramirez et al. \(2010\)](#), the global and exact identification must satisfy the following sufficiency conditions

“Consider an SVAR with admissible restrictions represented by R . If $(A_0, A_+) \in R$ and $M_j(f(A_0, A_+))$ is of rank $n \forall j \in 1 \leq j \leq n$, then the SVAR is globally identified at the parameter point (A_0, A_+) .”

For $1 \leq j \leq n$, linear restrictions basis matrix Q_j , and any $k \times n$ matrix $\mathbf{X}$, [Rubio-Ramirez et al. \(2010\)](#) defines $M_j(X)$ by

$$M_j(X) = \begin{bmatrix} [Q_j X]_{kn} \\ \left[\begin{array}{cc} I_{j \times j} & 0_{j \times (n-j)} \end{array} \right] \end{bmatrix} \quad (12)$$

where $I_{j \times j}$ is the $j \times j$ identity matrix. For the number of equations $n = 8$ and the lag length of $p=2$, I have $k = 24$. For each $M_j(X)$ where $\forall j \in 1 \leq i \leq 8$, showing that all $M_j(X)$'s are of rank 8 will complete the proof.

For $j = 1$, Q_1 is

$$Q_1 =$$

, and the matrix $M_1(f(A_0, A_+))$ is

$$M_1(f(A_0, A_+)) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{2,2} & 0 & 0 & 0 & a_{2,6} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{4,4} & 0 & 0 & a_{4,7} & a_{4,8} \\ 0 & 0 & 0 & 0 & a_{5,5} & 0 & a_{5,7} & a_{5,8} \\ 0 & 0 & 0 & 0 & 0 & a_{6,6} & 0 & a_{6,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & a_{7,7} & a_{7,8} \\ 0 & 0 & 0 & 0 & a_{8,5} & 0 & 0 & a_{8,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b_{2,2} & 0 & 0 & b_{2,5} & b_{2,6} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & b_{4,3} & b_{4,4} & 0 & 0 & b_{4,7} & b_{4,8} \\ 0 & 0 & 0 & 0 & b_{5,5} & b_{5,6} & b_{5,7} & b_{5,8} \\ 0 & 0 & 0 & 0 & b_{6,5} & b_{6,6} & b_{6,7} & b_{6,8} \\ 0 & 0 & 0 & 0 & b_{7,5} & b_{7,6} & b_{7,7} & b_{7,8} \\ 0 & 0 & 0 & 0 & b_{8,5} & b_{8,6} & b_{8,7} & b_{8,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & c_{2,2} & 0 & 0 & c_{2,5} & c_{2,6} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & c_{4,3} & c_{4,4} & 0 & 0 & c_{4,7} & c_{4,8} \\ 0 & 0 & 0 & 0 & c_{5,5} & c_{5,6} & c_{5,7} & c_{5,8} \\ 0 & 0 & 0 & 0 & c_{6,5} & c_{6,6} & c_{6,7} & c_{6,8} \\ 0 & 0 & 0 & 0 & c_{7,5} & c_{7,6} & c_{7,7} & c_{7,8} \\ 0 & 0 & 0 & 0 & c_{8,5} & c_{8,6} & c_{8,7} & c_{8,8} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (14)$$

whose rank is 8.

For $j = 2$, Q_2 is

[illegible]

, and the matrix $M_2(f(A_0, A_+))$ is

$$M_2(f(A_0, A_+)) = \begin{bmatrix} a_{1,1} & 0 & 0 & a_{1,4} & a_{1,5} & 0 & 0 & a_{1,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{3,1} & 0 & a_{3,3} & a_{3,4} & 0 & 0 & 0 & a_{3,8} \\ 0 & 0 & 0 & a_{4,4} & 0 & 0 & a_{4,7} & a_{4,8} \\ 0 & 0 & 0 & 0 & a_{5,5} & 0 & a_{5,7} & a_{5,8} \\ 0 & 0 & 0 & 0 & 0 & a_{6,6} & 0 & a_{6,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & a_{7,7} & a_{7,8} \\ 0 & 0 & 0 & 0 & a_{8,5} & 0 & 0 & a_{8,8} \\ b_{1,1} & 0 & b_{1,3} & b_{1,4} & b_{1,5} & b_{1,6} & 0 & b_{1,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ b_{3,1} & 0 & b_{3,3} & b_{3,4} & 0 & b_{3,6} & 0 & b_{3,8} \\ 0 & 0 & b_{4,3} & b_{4,4} & 0 & 0 & b_{4,7} & b_{4,8} \\ 0 & 0 & 0 & 0 & b_{5,5} & b_{5,6} & b_{5,7} & b_{5,8} \\ 0 & 0 & 0 & 0 & b_{6,5} & b_{6,6} & b_{6,7} & b_{6,8} \\ 0 & 0 & 0 & 0 & b_{7,5} & b_{7,6} & b_{7,7} & b_{7,8} \\ 0 & 0 & 0 & 0 & b_{8,5} & b_{8,6} & b_{8,7} & b_{8,8} \\ c_{1,1} & 0 & c_{1,3} & c_{1,4} & c_{1,5} & c_{1,6} & 0 & c_{1,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ c_{3,1} & 0 & c_{3,3} & c_{3,4} & 0 & c_{3,6} & 0 & c_{3,8} \\ 0 & 0 & c_{4,3} & c_{4,4} & 0 & 0 & c_{4,7} & c_{4,8} \\ 0 & 0 & 0 & 0 & c_{5,5} & c_{5,6} & c_{5,7} & c_{5,8} \\ 0 & 0 & 0 & 0 & c_{6,5} & c_{6,6} & c_{6,7} & c_{6,8} \\ 0 & 0 & 0 & 0 & c_{7,5} & c_{7,6} & c_{7,7} & c_{7,8} \\ 0 & 0 & 0 & 0 & c_{8,5} & c_{8,6} & c_{8,7} & c_{8,8} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (16)$$

whose rank is 8.

For $j = 3$, Q_3 is

[illegible]

, and the matrix $M_3(f(A_0, A_+))$ is

$$M_3(f(A_0, A_+)) =$$

$$\begin{bmatrix} a_{1,1} & 0 & 0 & a_{1,4} & a_{1,5} & 0 & 0 & a_{1,8} \\ 0 & a_{2,2} & 0 & 0 & 0 & a_{2,6} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{4,4} & 0 & 0 & a_{4,7} & a_{4,8} \\ 0 & 0 & 0 & 0 & a_{5,5} & 0 & a_{5,7} & a_{5,8} \\ 0 & 0 & 0 & 0 & 0 & a_{6,6} & 0 & a_{6,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & a_{7,7} & a_{7,8} \\ 0 & 0 & 0 & 0 & a_{8,5} & 0 & 0 & a_{8,8} \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b_{2,2} & 0 & 0 & b_{2,5} & b_{2,6} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & b_{5,5} & b_{5,6} & b_{5,7} & b_{5,8} \\ 0 & 0 & 0 & 0 & b_{6,5} & b_{6,6} & b_{6,7} & b_{6,8} \\ 0 & 0 & 0 & 0 & b_{7,5} & b_{7,6} & b_{7,7} & b_{7,8} \\ 0 & 0 & 0 & 0 & b_{8,5} & b_{8,6} & b_{8,7} & b_{8,8} \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & c_{2,2} & 0 & 0 & c_{2,5} & c_{2,6} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{5,5} & c_{5,6} & c_{5,7} & c_{5,8} \\ 0 & 0 & 0 & 0 & c_{6,5} & c_{6,6} & c_{6,7} & c_{6,8} \\ 0 & 0 & 0 & 0 & c_{7,5} & c_{7,6} & c_{7,7} & c_{7,8} \\ 0 & 0 & 0 & 0 & c_{8,5} & c_{8,6} & c_{8,7} & c_{8,8} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(18)

whose rank is 8.

For $j = 4$, Q_4 is

[illegible]

, and the matrix $M_4(f(A_0, A_+))$ is

$$M_4(f(A_0, A_+)) =$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{2,2} & 0 & 0 & 0 & a_{2,6} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{4,4} & 0 & 0 & a_{4,7} & a_{4,8} \\ 0 & 0 & 0 & 0 & a_{5,5} & 0 & a_{5,7} & a_{5,8} \\ 0 & 0 & 0 & 0 & 0 & a_{6,6} & 0 & a_{6,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & a_{7,7} & a_{7,8} \\ 0 & 0 & 0 & 0 & a_{8,5} & 0 & 0 & a_{8,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b_{2,2} & 0 & 0 & b_{2,5} & b_{2,6} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & b_{4,3} & b_{4,4} & 0 & 0 & b_{4,7} & b_{4,8} \\ 0 & 0 & 0 & 0 & b_{5,5} & b_{5,6} & b_{5,7} & b_{5,8} \\ 0 & 0 & 0 & 0 & b_{6,5} & b_{6,6} & b_{6,7} & b_{6,8} \\ 0 & 0 & 0 & 0 & b_{7,5} & b_{7,6} & b_{7,7} & b_{7,8} \\ 0 & 0 & 0 & 0 & b_{8,5} & b_{8,6} & b_{8,7} & b_{8,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & c_{2,2} & 0 & 0 & c_{2,5} & c_{2,6} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & c_{4,3} & c_{4,4} & 0 & 0 & c_{4,7} & c_{4,8} \\ 0 & 0 & 0 & 0 & c_{5,5} & c_{5,6} & c_{5,7} & c_{5,8} \\ 0 & 0 & 0 & 0 & c_{6,5} & c_{6,6} & c_{6,7} & c_{6,8} \\ 0 & 0 & 0 & 0 & c_{7,5} & c_{7,6} & c_{7,7} & c_{7,8} \\ 0 & 0 & 0 & 0 & c_{8,5} & c_{8,6} & c_{8,7} & c_{8,8} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(20)

whose rank is 8.

For $j = 5$, Q_5 is

[illegible]

, and the matrix $M_5(f(A_0, A_+))$ is

$$M_5(f(A_0, A_+)) =$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{2,2} & 0 & 0 & 0 & a_{2,6} & 0 & 0 \\ a_{3,1} & 0 & a_{3,3} & a_{3,4} & 0 & 0 & 0 & a_{3,8} \\ 0 & 0 & 0 & a_{4,4} & 0 & 0 & a_{4,7} & a_{4,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_{6,6} & 0 & a_{6,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & a_{7,7} & a_{7,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ b_{3,1} & 0 & b_{3,3} & b_{3,4} & 0 & b_{3,6} & 0 & b_{3,8} \\ 0 & 0 & b_{4,3} & b_{4,4} & 0 & 0 & b_{4,7} & b_{4,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ c_{3,1} & 0 & c_{3,3} & c_{3,4} & 0 & c_{3,6} & 0 & c_{3,8} \\ 0 & 0 & c_{4,3} & c_{4,4} & 0 & 0 & c_{4,7} & c_{4,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

(22)

whose rank is 8.

For $j = 6$, Q_6 is

[illegible]

, and the matrix $M_6(f(A_0, A_+))$ is

$$M_6(f(A_0, A_+)) = \begin{bmatrix} a_{1,1} & 0 & 0 & a_{1,4} & a_{1,5} & 0 & 0 & a_{1,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{3,1} & 0 & a_{3,3} & a_{3,4} & 0 & 0 & 0 & a_{3,8} \\ 0 & 0 & 0 & a_{4,4} & 0 & 0 & a_{4,7} & a_{4,8} \\ 0 & 0 & 0 & 0 & a_{5,5} & 0 & a_{5,7} & a_{5,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & a_{7,7} & a_{7,8} \\ 0 & 0 & 0 & 0 & a_{8,5} & 0 & 0 & a_{8,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & b_{4,3} & b_{4,4} & 0 & 0 & b_{4,7} & b_{4,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & c_{4,3} & c_{4,4} & 0 & 0 & c_{4,7} & c_{4,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \quad (24)$$

whose rank is 8.

For $j = 7$, Q_7 is

[illegible]

, and the matrix $M_7(f(A_0, A_+))$ is

$$M_7(f(A_0, A_+)) = \begin{bmatrix} a_{1,1} & 0 & 0 & a_{1,4} & a_{1,5} & 0 & 0 & a_{1,8} \\ 0 & a_{2,2} & 0 & 0 & 0 & a_{2,6} & 0 & 0 \\ a_{3,1} & 0 & a_{3,3} & a_{3,4} & 0 & 0 & 0 & a_{3,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_{6,6} & 0 & a_{6,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & a_{8,5} & 0 & 0 & a_{8,8} \\ b_{1,1} & 0 & b_{1,3} & b_{1,4} & b_{1,5} & b_{1,6} & 0 & b_{1,8} \\ 0 & b_{2,2} & 0 & 0 & b_{2,5} & b_{2,6} & 0 & 0 \\ b_{3,1} & 0 & b_{3,3} & b_{3,4} & 0 & b_{3,6} & 0 & b_{3,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ c_{1,1} & 0 & c_{1,3} & c_{1,4} & c_{1,5} & c_{1,6} & 0 & c_{1,8} \\ 0 & c_{2,2} & 0 & 0 & c_{2,5} & c_{2,6} & 0 & 0 \\ c_{3,1} & 0 & c_{3,3} & c_{3,4} & 0 & c_{3,6} & 0 & c_{3,8} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad (26)$$

whose rank is 8.

For $j = 8$, Q_8 is

[illegible]

, and the matrix $M_8(f(A_0, A_+))$ is

Robustness Tests on Omitted Variables

The study so far identifies global (i.e., the rest of the world) shocks and domestic policy changes under a small open economy structure. Orthogonal to this, there are at least two pertinent questions that can affect model dynamics. First, is it the case that there is an omitted variable in the monetary policy rule potentially affecting the domestic shocks? Second, is the climate condition actually crucial to the model?

To address the questions above, business confidence survey (BCS_t) and climate conditions (cl_t) are included and excluded from the original model in the text respectively, and the resulting models are compared to the original model using a standard likelihood ratio test. The results are provided in Table 2. Here, $h = 1$ indicates that the unrestricted model should be chosen instead of the restricted model, similarly $h = 0$ indicates that the unrestricted model should be rejected in favor of the restricted model. For both cases, the restricted model is the one with less variables.

Table 2: Likelihood Ratio Test

	Likelihood Values		Restrictions (q)	h-value
	Unrestricted	Restricted		
Model with cl_t :	3921.82	3851.23	8	1
Model with BCS_t :	3196.33	3921.82	24	0

The likelihood ratio test shows that including the climate variable cl_t is favored, while including the business confidence survey BCS_t is not. This suggests that the original model is well-specified.

A more powerful approach to test the omitted variable bias involves factor-augmented vector autoregressions (FAVARs) ([Bernanke et al., 2005](#)). However, this is not directly applicable based on VARs with regime-switching, which is investigated in the following chapters based on dynamic factor models. Here, I consider only the post-inflation targeting period without further regime-switches.

Since [Bernanke and Boivin \(2003\)](#) it is well-known in the literature that small-scale structural VAR models are unlikely to cover all relevant information sets of policymakers and relevant structures that capture business cycles. Although the addition of new variables might solve this issue, this process severely decreases the degree of freedom in structural models. [Bernanke et al. \(2005\)](#) suggests a way to overcome this issue by including factors directly into the VARs, leading to FAVARs. However, the FAVAR setup is robust to only small parameter changes. The following exercise forms factors (up to two) from 19 additional variables and includes it in the VAR as an observable.

The additional variables used in order to create factors are commodity prices, export, import, stock market index, business confidence survey, credit card balances, current account balances, monetary conditions expectation survey, hourly earnings, agricultural employment, working-age population, young age unemployment, unemployment, government consumption, total industry production excluding construction, car registrations, PPI, total non-financial sector credit and real residential property prices.

Using the factor-augmented VAR model of [Bernanke et al. \(2005\)](#),

$$\begin{bmatrix} F_t \\ Y_t \end{bmatrix} = \Phi(L) \begin{bmatrix} F_{t-1} \\ Y_{t-1} \end{bmatrix} + \varepsilon_t. \quad (29)$$

Here, the model is reduced form with the matrix $\Phi(L)$ representing the lagged coefficients. I impose the Cholesky decomposition so that shocks have structural interpretations. The ordering of variables in Y_t is climate, world oil prices, federal funds rate, US GDP, domestic inflation, real domestic production, monetary policy rule and the exchange rate.

The data is at quarterly frequency from 1990:Q1 to 2018:Q3. The impulse response functions of domestic interest rate, domestic production, and domestic inflation with no additional factor, one additional factor, and two additional factors are illustrated in Figure 13.

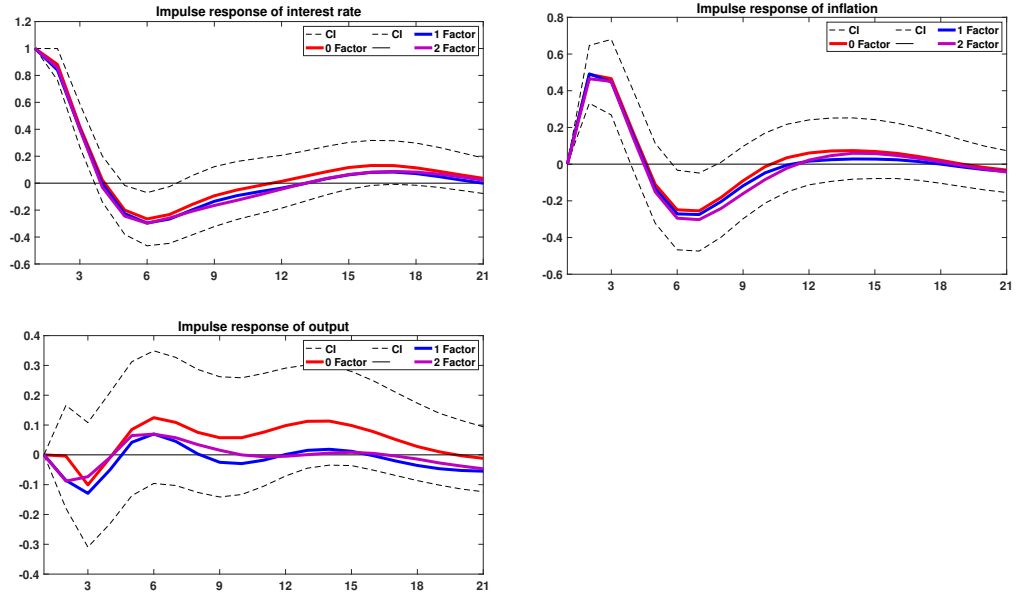


Figure 13: FAVAR Impulse Responses

The figure in the upper left is the impulse response of the domestic interest rate (i_t) with and without factors. The figure in the top right is the impulse

response of the domestic inflation (π_t), and the figure in the bottom is for the domestic production (y_t).

Impulse responses show that the additional variables do not create a significant difference for the model, with the figure plotting the 90% confidence intervals for the original model only when the deviations are statistically significant. The results suggest that the model dynamics are broadly similar at least qualitatively.

Robustness Checks on Exchange Rate Regimes

The New Zealand's exchange rate regime has seen several changes in the sample period of this chapter. From a fixed exchange rate regime under the Smithsonian Agreement of the early 1970s, New Zealand switched to a crawling pegged exchange regime in 1979. After 1985, New Zealand switched to a floating exchange rate regime against all currencies. One possible question regarding regime switching is whether the changes in the exchange rate regime are confounding the regime changes in the international block of the model. To answer this question, I separate the regime-switch in these two blocks of the model using distinct independent Markov chains.

Thus, the regime-switching equations in (2.4) are modified so that the shock variance regimes for the exchange rate is defined with and independent Markov chain s_t^5 . The set of equations in (30) represents the new setting where the exchange rate shock variance regimes are independent of other blocks of the model.

$$\begin{aligned}
y_t &= k_{11}cl_t + \sum_{i=1}^p k_{12}^i y_{t-i}^* + \sum_{i=1}^p k_{13}^i cl_{t-i} + \sum_{i=1}^p k_{14}^i wop_{t-i} + \sum_{i=1}^p k_{15}^i y_{t-i} \\
&\quad + \sum_{i=1}^p k_{16}^i \pi_{t-i} + \sum_{i=1}^p k_{17}^i e_{t-i} + \sum_{i=1}^p k_{18}^i i_{t-i} + \varepsilon_{t,1}[\mathbf{s}_t^1] \\
\pi_t &= k_{21}wop_t + k_{22}e_t + \sum_{i=1}^p k_{23}^i cl_{t-i} + \sum_{i=1}^p k_{24}^i wop_{t-i} + \sum_{i=1}^p k_{25}^i y_{t-i} \\
&\quad + \sum_{i=1}^p k_{26}^i \pi_{t-i} + \sum_{i=1}^p k_{27}^i e_{t-i} + \sum_{i=1}^p k_{28}^i i_{t-i} + \varepsilon_{t,2}[\mathbf{s}_t^1] \\
i_t &= k_{31}[\mathbf{s}_t^2]\pi_t + k_{32}[\mathbf{s}_t^2]i_t^* + \sum_{i=1}^p k_{33}^i[\mathbf{s}_t^2]\pi_{t-i} + \sum_{i=1}^p k_{34}^i[\mathbf{s}_t^2]i_{t-i}^* \\
&\quad + \sum_{i=1}^p k_{35}^i[\mathbf{s}_t^2]y_{t-i} + \sum_{i=1}^p k_{36}^i[\mathbf{s}_t^2]i_{t-i} + \sum_{i=1}^p k_{37}^i[\mathbf{s}_t^2]e_{t-i} + \varepsilon_{t,3}[\mathbf{s}_t^1] \\
y_t^* &= \sum_{i=1}^p k_{41}^i y_{t-i}^* + \sum_{i=1}^p k_{42}^i i_{t-i}^* + \sum_{i=1}^p k_{43}^i wop_{t-i} + \varepsilon_{t,4}[\mathbf{s}_t^3] \\
i_t^* &= k_{51}y_t^* + k_{52}wop_t + \sum_{i=1}^p k_{53}^i wop_{t-i} + \sum_{i=1}^p k_{54}^i i_{t-i}^* + \sum_{i=1}^p k_{55}^i y_{t-i}^* + \varepsilon_{t,5}[\mathbf{s}_t^3] \\
wop_t &= k_{61}y_t^* + \sum_{i=1}^p k_{62}^i wop_{t-i} + \sum_{i=1}^p k_{63}^i y_{t-i}^* + \varepsilon_{t,6}[\mathbf{s}_t^4] \\
cl_t &= \sum_{i=1}^p k_{71}^i cl_{t-i} + \varepsilon_{t,7}[\mathbf{s}_t^4] \\
e_t &= \sum_{i=1}^p k_{81}^i y_{t,t-i} + \sum_{i=1}^p k_{82}^i \pi_{t,t-i} + \sum_{i=1}^p k_{83}^i i_{t,t-i} + \sum_{i=1}^p k_{84}^i wop_{t,t-i} \\
&\quad + \sum_{i=1}^p k_{85}^i y_{t,t-i}^* + \sum_{i=1}^p k_{86}^i i_{t,t-i}^* + \sum_{i=1}^p k_{87}^i e_{t-i} + \varepsilon_{t,8}[\mathbf{s}_t^5]
\end{aligned}$$

Although the results here are similar to those of the original model, the formal marginal data density (MDD) comparison procedure still favors the baseline model.

Table 3: Marginal Data Density Comparison between the Baseline and Modified Model

Number of Regimes	Bridge Sampling MDD's
$\beta^* = [2,2,3,3,1]$	3702.29
$\beta = [2,2,3,3,2]$	3695.82

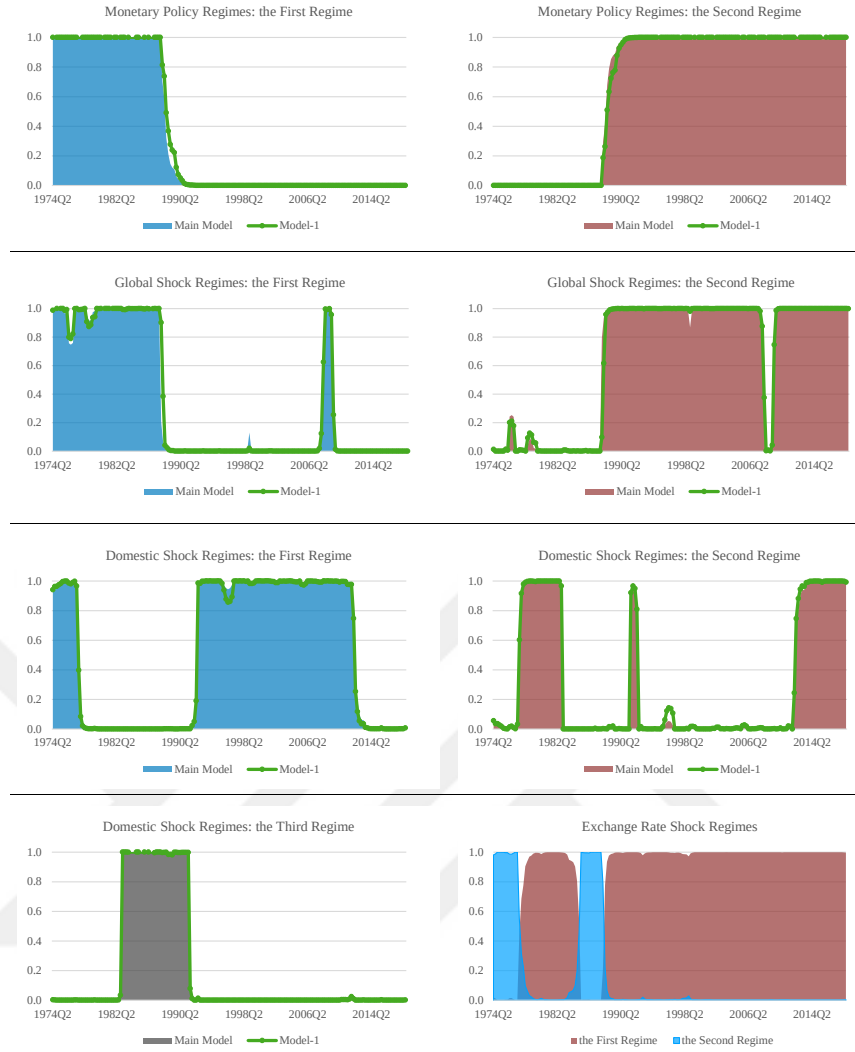


Figure 14: Regime Probabilities under Baseline Model and the Robustness Check

Figure 14 shows the smoothed regime probabilities for the monetary policy regimes, the global shock regimes, and the domestic shock variance regimes. Model-1 is the modified model described in this subsection. Estimated regime probabilities show that the main estimation results are robust to considering independent exchange rate regime switches. The probabilities at the bottom right visualize the smoothed regime probabilities for the exchange rate shock variances.

Robustness Checks on the International Block

Here, I check whether the results in the main text are sensitive to using the US as the proxy for the rest of world. To do so, I construct the trade-weighted exchange rate, the trade-weighted global GDP, and a trade partner's GDP-weighted interest rate as alternatives to the variables belonging to the exchange rate block and the international block, as in [Buckle et al. \(2007\)](#).

In the early 1980s, New Zealand's dominant trade partners were the United States and Australia. Over the last decade (after 2010), China emerged as an important trading partner. Due to the data availability issues, the trade weights between 1974 and 1984 are calculated using the data for the following 10 years. This practice is likely to be benign given the persistence of the trade shares in the early part of the sample.

For this robustness analysis, the top seven trade partners of New Zealand are considered. The countries are Australia, China, the Euro Area, Japan, Korea, the United Kingdom, and the United States. Following [Buckle et al. \(2007\)](#), and normalizing the total trade shares of the seven countries to 100%, a trade-weighted exchange rate series is constructed. Similarly, a trade-weighted international GDP and an international interest rate are also constructed. Following [Buckle et al. \(2007\)](#), the international GDP is the trade-weighted real GDP index, and the international interest rate is the weighted average of the 90-day bank bill rates of countries weighted by their nominal GDP in US dollars. When the trade share data is not available for a country, it is dropped from the calculation and the total of the rest of the countries is normalized to 100%. The resulting data and the US counterparts are shown in the Figure 15, which demonstrates that the United States is a reasonable proxy for the rest of the world.

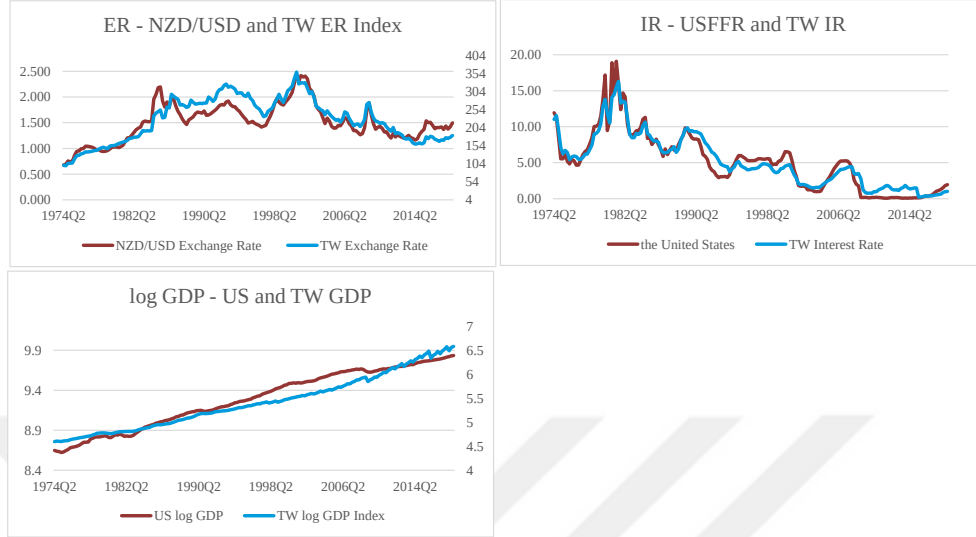


Figure 15: The United States Data in the Baseline Model and Trade Weighted Counterpart

Using the series above, I estimate two additional models: (i) a regime-switching SVAR model replacing the NZD/USD exchange rate with the trade-weighted exchange rate (Model-2), (ii) a regime-switching SVAR model replacing US GDP with trade-weighted GDP, FFR by GDP-weighted interest rate and NZD/USD by the trade-weighted exchange rate (Model-3). The results are presented in Figure 16. According to Figure 16, as far as the working of the inflation targeting policy regime and global shocks are considered, the results are robust.

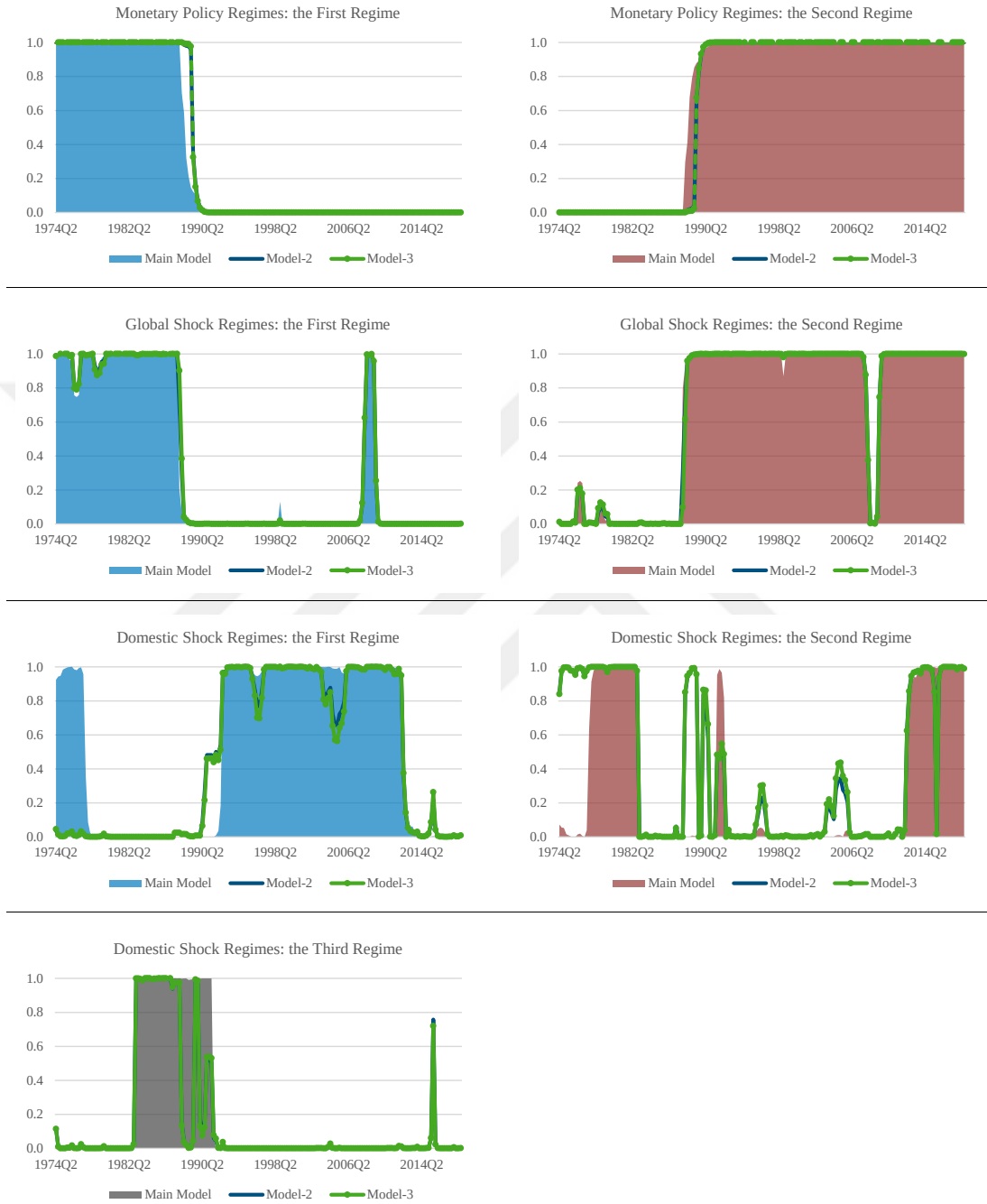


Figure 16: Regime Probabilities under Baseline Model and the Robustness Check

In terms of domestic shock variances, the regime transitions are different from the baseline model in the late 1970s and early 1980s. However, the main results of the paper are preserved, and the following discussions still hold for domestic shock variances as in the baseline analysis: First, the decline in the global shock variances preceded the decline in the domestic shock variances, and second, although the commitment to inflation targeting was not affected by the Great Financial Contraction of 2008, this was an episode in which the global shock variances increased first, and the domestic shock variances followed.

It is important to note that, in the current setup, the working of the rest of the world block has two dimensions. First, the rest of the world should have an effect on global demand, global monetary policy, and oil prices. Second, the rest of the world should affect New Zealand through the exchange rate block. It turns out the United States is a reasonable proxy which allows capturing these concerns. Although trade-weighted exchange rate, trade-weighted global GDP, and a trade partner's GDP-weighted interest rate are natural alternatives to consider for modeling the international block, data availability limits their usage.”

B Appendix for Chapter 3

This appendix provides additional information for Chapter 3, in particular the relationship between VARs and DFMs, building on the work of [Stock and Watson \(2016\)](#) and [Doz and Fuleky \(2020\)](#). The notation is consistent with that in Chapter 3. This aims to serve as background materials for those who are interested in the topic.

Let Y_t be an $N \times 1$ vector of observables, f_t be a $q \times 1$ vector of factors,⁴ e_t be a vector of possibly serially correlated innovations, and v_t be i.i.d shocks to the factors for time period $t = 1, 2, \dots, T$. A dynamic factor model (DFM) is represented by the following equations:

$$Y_t = c_0 f_t + c_1 f_{t-1} + \dots + c_s f_{t-s} + e_t \quad (31)$$

$$f_t = a_1 f_{t-1} + a_2 f_{t-2} + \dots + a_h f_{t-h} + v_t \quad (32)$$

$$e_t = \delta_1 e_{t-1} + \delta_2 e_{t-2} + \dots + \delta_p e_{t-p} + \varepsilon_t. \quad (33)$$

If there is no serial correlation for e_t , $\delta_i = 0, \forall i \in [1, \dots, p]$. $N \times q$ coefficients c_i are called factor loadings, $q \times q$ coefficients a_i are autoregressive coefficients for the unobserved factors f_t , and $N \times N$ coefficients δ_i are persistence parameters for the innovations e_t . Equation (31) is called the measurement equation which states the linear relation between the unobservable factors and the observables. Equation (32) is the transition equation which defines the dynamic relationship for the unobservable factors which gives the name DFM to the modelling framework.

Define for the case where $h \leq s$

⁴Note that factors need not be unobserved even though in this thesis I focus on the case where they are.

$$C = \begin{bmatrix} c_0 & c_1 & \dots & c_s \end{bmatrix}, F_t = \begin{bmatrix} f_t \\ f_{t-1} \\ \vdots \\ f_{t-s} \end{bmatrix} \text{ and } A = \begin{bmatrix} a_1 & a_2 & \dots & a_h & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 1 & 0 \end{bmatrix}$$

or for the case where $s < h$

$$C = \begin{bmatrix} c_0 & c_1 & \dots & c_s & 0 & \dots & 0 \end{bmatrix}, F_t = \begin{bmatrix} f_t \\ f_{t-1} \\ \vdots \\ f_{t-s} \\ f_{t+1-s} \\ \vdots \\ f_{t-h} \end{bmatrix} \text{ and } A = \begin{bmatrix} a_1 & a_2 & \dots & a_h & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}.$$

$$\text{Also define } G = \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Then, equations (31)-(31) can be stated as

$$Y_t = c_0 f_t + c_1 f_{t-1} + \dots + c_s f_{t-s} + e_t \quad (34)$$

$$f_t = a_1 f_{t-1} + a_2 f_{t-2} + \dots + a_h f_{t-h} + G v_t \quad (35)$$

$$e_t = \delta_1 e_{t-1} + \delta_2 e_{t-2} + \dots + \delta_p e_{t-p} + \varepsilon_t \quad (36)$$

which in the vector form is

$$Y_t = C F_t + e_t \quad (37)$$

$$F_t = A F_{t-1} + G v_t \quad (38)$$

$$e_t = \delta_1 e_{t-1} + \delta_2 e_{t-2} + \dots + \delta_p e_{t-p} + \varepsilon_t. \quad (39)$$

Stock and Watson (2016) define equations (34)-(36) the dynamic representation of DFM, while equations (37)-(39) the static (or stacked) representation of DFM. I will show the relationship between standard VAR models and DFMs using the DFM in equations (34)-(36) (or (37)-(39) in the matrix form).

A Mapping from VAR Model to DFM

Recall that Y_t is a vector of N observables $y_{i,t}$ for time period $t = 1, 2, \dots, T$. λ_i is an $N \times N$ matrix of coefficients for the i^{th} lagged observables Y_{t-i} , p is the overall lag length, and η_t is an $N \times 1$ vector of i.i.d shocks to the observables. Then, a standard VAR model is represented as

$$Y_t = \sum_{i=1}^p \lambda_i Y_{t-i} + \eta_t, \quad Y_t = \begin{bmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{N,t} \end{bmatrix}. \quad (40)$$

This can be rewritten as a DFM by taking the following steps. First, I can rewrite the observable variables $y_{i,t}$'s as in the following equation where c_0 is an $N \times N$ identity matrix, c_i for $i \neq 0$ are $N \times N$ zero matrices, and e_t is an $N \times 1$ vector of zeros $\forall t$:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{N,t} \end{bmatrix} = c_0 \begin{bmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{N,t} \end{bmatrix} + c_1 \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \vdots \\ y_{N,t-1} \end{bmatrix} + \dots + c_s \begin{bmatrix} y_{1,t-s} \\ y_{2,t-s} \\ \vdots \\ y_{N,t-s} \end{bmatrix} + e_t. \quad (41)$$

This is an identity mapping. Second, let the transition equation for the VAR model be:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{N,t} \end{bmatrix} = a_1 \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \vdots \\ y_{N,t-1} \end{bmatrix} + a_2 \begin{bmatrix} y_{1,t-2} \\ y_{2,t-2} \\ \vdots \\ y_{N,t-2} \end{bmatrix} + \dots + a_p \begin{bmatrix} y_{1,t-p} \\ y_{2,t-p} \\ \vdots \\ y_{N,t-p} \end{bmatrix} + \eta_t. \quad (42)$$

This is the original VAR, but now interpreted as the equation describing the dynamic relationship among the observed factors. Because

$$Y_t = f_t, \quad N = q$$

equation (41) reduces to

$$\begin{aligned} y_t = \begin{bmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{N,t} \end{bmatrix} &= c_0 \begin{bmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{N,t} \end{bmatrix} + c_1 \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \vdots \\ y_{N,t-1} \end{bmatrix} + \dots + c_s \begin{bmatrix} y_{1,t-s} \\ y_{2,t-s} \\ \vdots \\ y_{N,t-s} \end{bmatrix} + e_t \\ &= c_0 f_t + c_1 f_{t-1} + \dots + c_s f_{t-s} + e_t \end{aligned}$$

and equation (42) reduces to

$$f_t = \begin{bmatrix} Y_{1t} \\ Y_{2t} \\ \vdots \\ Y_{Nt} \end{bmatrix} = a_1 \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \vdots \\ y_{N,t-1} \end{bmatrix} + a_2 \begin{bmatrix} y_{1,t-2} \\ y_{2,t-2} \\ \vdots \\ y_{N,t-2} \end{bmatrix} + \dots + a_p \begin{bmatrix} y_{1,t-p} \\ y_{2,t-p} \\ \vdots \\ y_{N,t-p} \end{bmatrix} + \eta_t$$

$$= a_1 f_{t-1} + a_2 f_{t-2} + \dots + a_h f_{t-h} + \eta_t.$$

Then, a standard VAR model defined above becomes a special case of DFM where

- The estimated factors are observable, that is, $f_t = Y_t$
- The autoregressive coefficients a_i of the DFM in equation (35) are equal to the autoregressive coefficients of the VAR model, that is, $a_i = \lambda_i$
- a_0 of the DFM in equation (35) is an identity matrix and a_i for $i \neq 0$ is a vector of zeros.
- The shock to the observation equation e_t of the DFM in equation (35) is a vector of zeros.
- The shock to the transition equation of the DFM in equation (35) is the shock to the VAR model, that is, $v_t = \eta_t$.

A Mapping from SVAR Model to DFM

Now, I turn to SVAR where further restrictions are imposed on VAR to be able to interpret shocks as structural. Suppose Y_t is an $N \times 1$ matrix of observables for time period $t = 1, 2, \dots, T$, $\tilde{\lambda}_i$ is an $N \times N$ coefficients for the i^{th} lagged observables Y_{t-i} , $\tilde{\lambda}_0$ is a non-singular contemporaneous relation matrix (a matrix including restrictions on the current period Y_t), p is the overall lag length, and $\tilde{\eta}_t$ is an $N \times 1$ vector of i.i.d shocks to Y_t . Then, following [Stock and Watson \(2016\)](#), a standard SVAR model is represented as

$$\tilde{\lambda}_0 Y_t = \sum_{i=1}^p \tilde{\lambda}_i Y_{t-i} + \tilde{\eta}_t, \quad Y_t = \begin{bmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{N,t} \end{bmatrix}. \quad (43)$$

Now define

$$\lambda_i = \tilde{\lambda}_0^{-1} \tilde{\lambda}_i$$

$$\eta_t = \tilde{\lambda}_0^{-1} \tilde{\eta}_t.$$

Then, the reduced-form representation of the structural VAR model is

$$Y_t = \sum_{i=1}^p \lambda_i Y_{t-i} + \eta_t, \quad Y_t = \begin{bmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{N,t} \end{bmatrix}.$$

The rest are identical to the previous subsection.

A Mapping from DFM to VAR Model

Consider the measurement equation in equation (34), but for a single observation:

$$y_t = c_0 f_t + c_1 f_{t-1} + c_2 f_{t-2} + \dots + c_s f_{t-s} + e_t.$$

Lagging this equation,

$$y_{t-1} = c_0 f_{t-1} + c_1 f_{t-2} + \dots + c_s f_{t-1-s} + e_{t-1}$$

$$y_{t-2} = c_0 f_{t-2} + c_1 f_{t-3} + \dots + c_s f_{t-2-s} + e_{t-2}$$

$$\vdots$$

$$y_{t-p} = c_0 f_{t-p} + c_1 f_{t-p-1} + \dots + c_s f_{t-p-s} + e_{t-p}.$$

Multiplying each observable y_{t-i} by $(-\delta_i) \forall i \in [1, 2, \dots, p]$,

$$\begin{aligned}
y_t &= c_0 f_t + c_1 f_{t-1} + c_2 f_{t-2} + \dots + c_s f_{t-s} + e_t \\
-\delta_1 y_{t-1} &= -\delta_1 c_0 f_{t-1} + -\delta_1 c_1 f_{t-2} + \dots + -\delta_1 c_s f_{t-1-s} + -\delta_1 e_{t-1} \\
-\delta_2 y_{t-2} &= -\delta_2 c_0 f_{t-2} + -\delta_2 c_1 f_{t-3} + \dots + -\delta_2 c_s f_{t-2-s} + -\delta_2 e_{t-2} \\
-\delta_3 y_{t-3} &= -\delta_3 c_0 f_{t-3} + -\delta_3 c_1 f_{t-4} + \dots + -\delta_3 c_s f_{t-3-s} + -\delta_3 e_{t-3} \\
&\vdots \\
-\delta_p y_{t-p} &= -\delta_p c_0 f_{t-p} + -\delta_p c_1 f_{t-p-1} + \dots + -\delta_p c_s f_{t-p-s} + -\delta_p e_{t-p}
\end{aligned}$$

where the symbols in red are used to make the equations more readable. Substitute f_t from equation (35) into the above and add up the equations to get

$$\begin{aligned}
y_t &= \delta_1 y_{t-1} + \delta_2 y_{t-2} + \delta_3 y_{t-3} + \dots + \delta_p y_{t-p} \\
&+ c_0 a_1 f_{t-1} + c_0 a_2 f_{t-2} + \dots + c_0 a_h f_{t-h} + c_0 \varepsilon_t \\
&+ c_1 f_{t-1} + c_2 f_{t-2} + \dots + c_s f_{t-s} \\
&- \delta_1 c_0 f_{t-1} - \delta_1 c_1 f_{t-2} - \dots - \delta_1 c_s f_{t-1-s} \\
&- \delta_2 c_0 f_{t-2} - \delta_2 c_1 f_{t-3} - \dots - \delta_2 c_s f_{t-2-s} \\
&- \delta_3 c_0 f_{t-3} - \delta_3 c_1 f_{t-4} - \dots - \delta_3 c_s f_{t-3-s} \\
&\vdots \\
&- \delta_p c_0 f_{t-p} - \delta_p c_1 f_{t-p-1} - \dots - \delta_p c_s f_{t-p-s} + e_t.
\end{aligned}$$

Rearrange the equation above to get

$$\begin{aligned}
y_t = & \delta_1 y_{t-1} + \delta_2 y_{t-2} + \delta_3 y_{t-3} + \dots + \delta_p y_{t-p} \\
& + (\Gamma_1 c_0 a_1 + \Theta_1 c_1 - \delta_1 c_0) f_{t-1} \\
& + (\Gamma_2 c_0 a_2 + \Theta_2 c_2 - \delta_1 c_1 - \delta_2 c_0) f_{t-2} \\
& + (\Gamma_3 c_0 a_3 + \Theta_3 c_3 - \delta_1 c_2 - \delta_2 c_1 - \delta_3 c_0) f_{t-3} \\
& \vdots \\
& + (\Gamma_i c_0 a_i + \Theta_i c_i - \sum_{k=1}^i \delta_k c_{i-k} \xi_k \mu_{i,k}) f_{t-i} \\
& \vdots \\
& + (-\delta_p c_s) f_{t-p-s} \\
& c_0 \varepsilon_t + e_t
\end{aligned} \tag{44}$$

where the terms $(\Gamma_i, \Theta_i, \xi_i, \mu_{i,k})$ are indicators defined as

$$\begin{aligned}
\Gamma_i &= 1 \quad \text{for } i \leq h, \\
\Theta_i &= 1 \quad \text{for } i \leq s, \\
\xi_i &= 1 \quad \text{for } i \leq p, \\
\mu_{i,k} &= 1 \quad \text{for } i - k < s,
\end{aligned}$$

and otherwise zero.

Now, I define $j = \max(h, s)$, and the vector Y_t where

$$Y_t = \begin{bmatrix} y_t \\ f_t \\ f_{t-1} \\ \vdots \\ f_{t-j} \end{bmatrix}.$$

Then, equation (44) can be rewritten in the following form:

$$Y_t = \begin{bmatrix} y_t \\ f_t \\ f_{t-1} \\ \vdots \\ f_{t-j} \end{bmatrix} = \lambda_1 \begin{bmatrix} y_{t-1} \\ f_{t-1} \\ f_{t-2} \\ \vdots \\ f_{t-j-1} \end{bmatrix} + \lambda_2 \begin{bmatrix} y_{t-2} \\ f_{t-2} \\ f_{t-3} \\ \vdots \\ f_{t-j-2} \end{bmatrix} + \dots + \lambda_p \begin{bmatrix} y_{t-p} \\ f_{t-p} \\ f_{t-p-1} \\ \vdots \\ f_{t-j-p} \end{bmatrix} + c_0 \varepsilon_t + e_t \quad (45)$$

where λ_i is a $(j+1) \times (j+1)$ autoregressive coefficient matrix in which the first element is δ_i . Let $\eta_t = c_0 \varepsilon_t + e_t$ where the e_t and ε_t are by definition independent. Then, one can rewrite (45) as a VAR.⁵

$$Y_t = \sum_{i=1}^p \lambda_i Y_{t-i} + \eta_t, \quad Y_t = \begin{bmatrix} y_t \\ f_t \\ f_{t-1} \\ \vdots \\ f_{t-j} \end{bmatrix}. \quad (46)$$

This completes the mapping from the DFM to the VAR. One can impose additional restrictions to show the mapping to the SVAR.

⁵It is important to note that without further restrictions on λ_i 's, the mapping from the DFM to the VAR is not unique.