

ESSAYS ON MEASURING CREDIT AND PROPERTY PRICES GAPS

by

Nam Nguyen Tam Hoai

A Dissertation Submitted in
Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
in Economics

at

The University of Wisconsin-Milwaukee

August 2022

ABSTRACT

ESSAYS ON MEASURING CREDIT AND PROPERTY PRICES GAPS

by

Nam Nguyen Tam Hoai

The University of Wisconsin-Milwaukee, 2022
Under the Supervision of Professor Kundan Kishor

My dissertation studies credit expansion and its effect on house prices and financial stability. In the first chapter, I examine the idea that house prices and credit to households are jointly determined, affecting each other in the short and long run. I decompose the movements of the two variables of interest into permanent long-run and transitory short-run components using an unobserved components vector autoregressive model. The dynamic model shows findings to support the hypothesis that a short-run positive shock to house prices is associated with an increase in household credit above its long-run trend. Furthermore, by utilizing additional information generated by the unobserved component model, our multivariate model performs better than univariate models in capturing the dynamics of household credit and house prices over the last three decades, especially during the recent financial crisis. I also estimated the predictive ability of cyclical components of a variable on their counterparts from other variables by employing cross-correlation coefficients in the VAR model.

In the second chapter, I propose a new method to measure the credit gap - the deviation of the credit-to-GDP ratio from its long-run trend. Here, I utilize the idea proposed in Nelson (2008) that the deviation of a non-stationary variable from its long-run trend should predict future changes in the variable. Since different trend-cycle decomposition methods of credit-to-GDP ratio provide different credit gap measures, I handle the model uncertainty by assigning weights to these different credit gap measures based on their relative out-of-sample

predictive power based on Bates and Granger (1969) forecast combination method. I apply this approach to estimate the UK and the US credit gap using credit-to-GDP ratio data from 1960-2020.

My proposed credit gap measure dominates the alternate credit gaps, including the one provided by the Bank of International Settlements (BIS) regarding its relative out-of-sample predictive power. The proposed gap also has superior features in terms of early detection of turning points and relative insensitivity to the endpoint problem.

The third chapter of my dissertation overcomes model uncertainty in using the credit gap as an early warning indicator (EWI) of systemic financial crises in a binary outcome setting. I propose using model averaging of different credit gap measurements to achieve a better averaged model fit and out-of-sample prediction. I use binary logistic regression in a panel setup consisting of 40 countries. In this paper, I also propose a novel, superior criteria to judge the performance of an EWI than the one currently popularly used in the literature. The empirical results showed that our Bayesian averaged model could synthesize a single credit gap that outperforms other popularly studied credit gap measurements in terms of an early warning indicator.

© Copyright by Nam Nguyen Tam Hoai, 2022
All Rights Reserved

To
my parents,
and my teachers,

TABLE OF CONTENTS

Abstract	ii
List of Figures	vii
List of Tables	viii
1 Introduction	1
2 Credit and House Prices Cycles	3
2.1 Introduction	3
2.2 Literature Review	5
2.3 Data Description	11
2.4 Empirical Model	13
2.5 Empirical Results	19
2.6 Comparison with other Decomposition Methods	32
2.7 Conclusion	35
3 Measuring Credit Gap	37
3.1 Introduction	37
3.2 Related Literature	40
3.3 The Data	41
3.4 Empirical Methodology	42
3.5 Empirical Results	45
3.6 Conclusions	48
4 Identifying Unsustainable Credit Gap	52
4.1 Introduction	52
4.2 Literature Review	54
4.3 Data Description	57
4.4 Empirical Methodology	58
4.5 Empirical Results	76
4.6 Analyzing Time Series Graphs of Individual Countries	86
4.7 Conclusion	89
References	91
Appendix A: Chapter 2	99
Appendix B: Chapter 3	116
Appendix C: Chapter 4	119

LIST OF FIGURES

2.1	UK VAR(2)	25
2.2	UK VAR(2) 1 cross-lag	25
2.3	UK VAR(2) 2 cross-lags	26
2.4	US VAR(2)	26
2.5	US VAR(2) 1 cross-lag	27
2.6	US VAR(2) 2 cross-lags	27
2.7	Comparing Multivariate UC cycles with alternate decompositions: UK . . .	33
2.8	Comparing Multivariate UC cycles with alternate decompositions: US . . .	34
4.1	Procedure to standardize partial AUC	66
4.2	One sided weights graph	75
4.3	US time series graph	87
4.4	UK time series graph	88
A.1	UK VAR(2) - Posterior and Prior Distribution	101
A.2	UK VAR(2) 1 cross-lag - Posterior and Prior Distribution	102
A.3	UK VAR(2) 2 cross-lags - Posterior and Prior Distribution	103
A.4	US VAR(2) - Posterior and Prior Distribution	104
A.5	US VAR(2) 1 cross-lag - Posterior and Prior Distribution	105
A.6	US VAR(2) 2 cross-lags - Posterior and Prior Distribution	106
A.7	UK VAR(2) - Posterior chain	108
A.8	UK VAR(2) 1 cross-lag - Posterior chain	109
A.9	UK VAR(2) 2 cross-lags - Posterior chain	110
A.10	US VAR(2) - Posterior chain	111
A.11	US VAR(2) 1 cross-lag - Posterior chain	112
A.12	US VAR(2) 2 cross-lags - Posterior chain	113
A.13	Italy VAR(2) 1 cross-lag - Unobserved Components	114
A.14	Japan VAR(2) 1 cross-lag - Unobserved Components	115
B.1	Credit-to-GDP Ratios	116
B.2	Credit Gap Comparison (U.K.)	117
B.3	Credit Gap Comparison (U.S.)	118
C.1	Canada Graph	122
C.2	France Graph	122
C.3	Hong Kong (SAR) Graph	123
C.4	Japan Graph	123
C.5	South Korea Graph	124
C.6	Saudi Arabia Graph	124
C.7	Sweden Graph	125
C.8	Switzerland Graph	125
C.9	Thailand Graph	126

LIST OF TABLES

2.1	Descriptive statistics	12
2.2	Correlation matrix	13
2.3	Parameters description	19
2.4	UK Regression Results - All three models	20
2.5	US Regression Results - All three models	21
2.6	VAR(2) 1 cross-lag : UK and US regression results	22
2.7	VAR(2) 1 cross-lag : Cross countries estimates	30
3.1	Forecasting Performance of Credit Gap Models (U.K.)	50
3.2	Forecasting Performance of Credit Gap Models (U.S.)	51
4.1	Top 30 credit gap measurements ranked by psAUC	67
4.2	Top 25 credit gap measurements ranked by MC3 probability	68
4.3	Credit gaps performance as EWIs - Advanced Economies	77
4.4	Credit gaps performance as EWIs - Emerging Market Economies	78
4.5	Credit gaps performance as EWIs - Full sample	79
4.6	Credit gaps performance as EWIs - Out of sample prediction - AE	83
4.7	Credit gaps performance as EWIs - Out of sample prediction - EME	84
4.8	Credit gaps performance as EWIs - Out of sample prediction - Full sample	85

Chapter 1

Introduction

Financial market stability has been an increasingly important topic in recent years, especially after the great recession. Among major determinants of financial market booms and busts, credit supply expansion and property prices are popularly cited and studied. In my dissertation, I try to explore different aspects of the cycle of credit supply (deviation of credit-to-GDP ratio from its long-run trend), such as its forecasting ability, its dynamic relationship with property prices, and financial stability. My analysis focuses on the US and UK economies in the first two chapters. While the third chapter uses a panel data setting with 40 countries.

The first chapter of my dissertation exploits a model that allows for household credit and house prices to be jointly determined in both short run and long run. The methodology I use in this chapter allows for measuring the directions and magnitudes of the effects the two variables have on each other. More importantly, the timing of the effect can also be estimated using a state-space framework. The model empirical result suggests that, in the short run, a positive shock to house prices is associated with a future increase in household credit above its long-run trend. This result is robust in multiple structural specifications of the trend-cycle decomposition methods.

The second chapter of my dissertation proposes a new method to measure the credit gap - (deviation of credit-to-GDP ratio from its long-run trend) based on its out-of-sample forecast performance. I decompose the total credit-GDP ratio series into short-run deviations and long-run trends using many relevant and popularly practiced methods. I then combine the cycles to construct a singular cyclical series with the highest predictive power on out-of-sample forecasts. The empirical results suggest that by using forecast error weighted

average combination, we can produce a measurement of credit gaps that provides superior performance as an early indicator of turning points.

The third chapter of my dissertation overcomes model uncertainty in using the credit gap as an early warning indicator (EWI) of systemic financial crises in a binary outcome setting. I propose using model averaging of different credit gap measurements to achieve better averaged model fit and out-of-sample prediction. I also propose a novel, superior criteria to judge the performance of an EWI than the one currently popularly used in the literature. The empirical results showed that the Bayesian averaged model I proposed could synthesize a single credit gap that out-performs any other popularly studied credit gap measurements in terms of an early warning indicator.

Chapter 2

Credit and House Prices Cycles

2.1 Introduction

The Great Recession caused researchers to shift their focus on the narrative of credit, housing market and financial stability. However, the debate on whether house prices have been the main driving source of the credit cycle, or financial conditions (credit) are the main determinants of the house price cycle is still open. One strand of literature has argued that increase in credit supply played a major role in the boom and the subsequent bust in the housing market in the US. Another strand of literature has argued that credit supply itself cannot explain the big swings in house prices and has attributed beliefs and other unobserved characteristics as a major source of house price variations. At the same time, some researchers have argued that credit booms are preceded and sometimes driven by housing booms. The increase in collateral and relaxation of banks' funding constraints leads to an increase in the willingness of the banking sector to provide funding not only to the residential sector but also to commercial real estate as well as overall funding to the businesses.

Most of the work in the literature has considered the relationship between house prices and credit separately, that is, if house prices are affected by credit or changes in house prices affect credit. It is perfectly reasonable to assume that house prices and credit have a dynamic relationship, and the causality does not necessarily run from one variable to another. The novel contribution of this study is to develop a model to jointly examine the two variables of interest: household credit and house prices, and their interaction. In particular, I pay attention to the long-run and short-run movements in credit and house prices and model their joint dynamics. In doing so, I use data from the US and the UK. The methodology I use in this paper to extract transitory and permanent information from non-stationary

time series is a decomposition method called Unobserved Components model pioneered by Beveridge & Nelson (1981). The implementation details of the methodology are inspired by Morley (2007) and Huang & Kishor (2019). This method allows the permanent component to be shown as a random walk with a local growth rate component and the transitory component to be a stationary process with a mean zero. The stationary transitory component configuration is essential to infer meaningful structural linkage between the two variables of interest: household credit and house prices, as non-stationary transitory components do not offer meaningful inferences. This brings up the paper’s second novel contribution to the literature. Discovering causality in the structure of a dynamical system from observed time series is a traditional and important problem. By explicitly configuring cross-correlation coefficients on the cyclical components of the two variables, I will be able to examine the predictive ability of the cycles. This would produce a much desired inference for macro-prudential policy implications to stabilize the macroeconomy.

Using sample data from the US and the UK, I find interesting and meaningful results from the estimated multivariate correlated unobserved component model. The maximum-likelihood estimates of my proposed correlated multivariate UC model suggest that there is a strong positive correlation between the transitory shock to household credit and the transitory shock to the house prices index. This suggests that a temporary increase in household credit is associated with an increase in house prices above their long-run level. These results support the narrative evidence of the strong relationship between household credit and house prices. More importantly, I also find evidence that lags of the household credit cycle have the predictive ability in forecasting the magnitude of house prices gap above its long-run level by examining the cross-correlation coefficient on the cyclical components. I also find that the trend-cycle decomposition of the two variables of interest captures the recent boom and bust behavior and compares favorably to a univariate trend-cycle decomposition benchmark. In particular, I find that the house prices and household credit were significantly higher than their long-term trend before the financial crisis. Then there was an

overreaction during the crisis leading the house price and credit cycle to negative territory, implying that house prices and credit were below their long-run trend. The magnitude of the model's cyclical components for both house price and credit during this time of crisis is significantly higher than that of other univariate decomposition models' such as the HP filter and the VAR unobserved component model.

Finally, my sample data show that the correlation between house price and credit is much higher in the UK than in the US, as shown in Table 1. This help tests the robustness of my model. The plan of the paper is as follows. In section 2, I will discuss the relevant literature branches. In section 3, I will introduce the summary of data and their description. In section 4, I will describe the decomposition methodology using unobserved component model with vector auto-regression (VAR). Specifically, in subsection 4.2, I introduced my contribution to the numerical optimization process in the literature. In section 5, I will go over the model regression results and my interpretation. I will also show cross-countries evidence when using the data from other developed countries. In section 6, I will test the robustness of the model by comparing the results against some traditional methods of analyzing time series data. And lastly, in section 7, I will give my conclusion remark for the model.

2.2 Literature Review

There has been an increasing interest in the study of the interaction between credit, speculation, and house prices Mian & Sufi (2011); Mian & Sufi (2018) , Kishor (2020) , Guerrieri & Uhlig (2016), and Davis & Van Nieuwerburgh (2015) have detailed literature reviews on the dynamics between the housing market and credit conditions. I list the four literature branches that study the dynamics of the credit cycles, housing price cycles, and then the critical connection between boom-bust episodes in housing markets and boom-bust episodes in credit markets. There are two approaches to this interaction:(i) The house price cycles generates the credit cycles. (ii) The credit cycles generate the house price cycles.

One strand of the literature focuses on how credit cycles are generated. Kiyotaki & Moore (1997) modeled the fluctuation of credit conditions due to credit limits and asset prices. The model shows how exogenous shocks can create cyclical fluctuation in credit, asset prices, and real output. Myerson (2012) proposed a model of credit cycles generated by moral hazard in dynamic interactions among different generations of financial agents. Guerrieri & Uhlig (2016) used a catastrophe model for credit, in which multiple equilibria are possible due to adverse selection: as credit increases, the composition of borrowers worsens, this can generate a crash in the credit market. Boissay, Collard, & Smets (2016) studied the topic of endogenous boom and bust in the credit market using a dynamic stochastic general equilibrium (DSGE) model, in which moral hazard and asymmetric information may endogenously lead to sudden freezes and crises in the credit market. As for classifying periods of booms and bursts in credit conditions, Alessi & Detken (2018) used a random forest model to identify unsustainable credit growth periods.

The second branch of the literature focuses on the dynamics of house price changes. I first look at the generation of momentum in house price changes. Asset prices and valuations tend to vary when information about their performance is available. Barberis, Shleifer, & Vishny (2005) pointed out that securities with good performance records receive extremely high valuations, and those valuations will return to the mean on average. Hong & Stein (1999) suggested a model with information diffuses gradually across the population, and if agents implement simple univariate strategies, their attempts at arbitrage will lead to overreaction at long horizons. Capozza, Hendershott, & Mack (2004) analyzed dynamic properties of markets exhibiting serial correlation and mean reversion. These properties allow prices to overshoot equilibrium (cycles) and permanently diverge from equilibrium. Glaeser, Gyourko, & Saiz (2008) incorporated housing supply elasticity into the analysis of housing prices momentum and showed that the price run-ups of the 1980s were almost exclusively experienced in cities with more inelastic housing supply. Head, Lloyd-Ellis, & Sun (2014) showed that variation in the time it takes to sell houses induces transaction prices

to exhibit serially correlated growth. Glaeser & Nathanson (2017) modeled the leads house prices expectation approximation to display missing features in rational models: momentum at short-run horizon, mean reversion in the long-run horizon, and excess longer-term volatility relative to fundamentals. Kishor, Kumari, & Song (2015) studied the US housing market by using a combination of the Unobserved Component model and GARCH model to study the time-varying importance of permanent and transitory housing components in the US housing prices. Katharina Knoll, Schularick, & Steger (2017) constructed a house price index for 14 economies in over 140 years. They argue that real house prices have largely followed a “hockey stick” pattern: fairly consistent for an extended period, followed by a pronounced increase towards the second half of the century with substantial cross-country variation. Furthermore, they say that most of the price increase can be attributed to the increase in the price of land and that house prices have risen faster than income in recent decades. K. Knoll (2016) argued that a rise in house prices coincides with a rise in the price-rent ratio, a fundamental that shows the intrinsic value of housing.

Another branch of literature is the one that studies the hypothesis that house price cycles generate credit cycles. The dynamics of houses price on household credit can be viewed through the lens of the borrower balance sheet. Bernanke & Gertler (1989) developed a neoclassical model of the business cycle in which the condition of borrowers’ balance sheets is a source of output dynamics. The mechanism is that higher borrower net worth reduces the agency costs of financing real capital investments. The financial acceleration effects imply that stronger balance sheets due to higher asset prices will lead to a lower cost of borrowing to invest, suggesting that a housing price boom will lead to a boom in credit. Kiyotaki & Moore (1997) further incorporated this positive feedback through asset prices and the associated intertemporal multiplier process that affects borrowing capacity and output into their paper. An increase in home equity due to an increase in house prices will allow borrowers to borrow more to finance either personal consumption or more speculation housing investment. Mian & Sufi (2018) showed that the crash in the housing market and following credit crunch

showed the importance of housing prices for household balance sheets and banking sector balance sheets.

Some papers have studied the hypothesis that credit cycles generate house price cycles. Agnello & Schuknecht (2011) , Agnello, Castro, & Sousa (2018) examined different variables likely to create a bubble in housing markets. First is the effect of credit constraints on house prices. Stein (1995) is the first paper to explore the effects of down-payment requirements on house price volatility. The paper highlighted the self-reinforcing effect from house prices to down payments and housing demand back to house prices. If house prices decline, the value of households' collateral declines, depressing housing demand and pushing house prices further down. This multiplier effect can generate multiple equilibria and accounts for the house price boom-bust episodes. The self-reinforcing effect has the same spirit as the transmission mechanism by Kiyotaki & Moore (1997). In a recent related paper, Ortalo-Magne & Rady (2006) showed that income volatility of young households or relaxation of their credit constraints could explain excess volatility of house prices by identifying a powerful driver of the housing market: the ability of young households to afford the down payment on a starter home.

Another research branch has also explored the effect of financial innovation or financial liberalization on house prices. Kermani (2012) proposed a model to emphasize the importance of financial liberalization and its reversal to explain the housing boom and bust. He, Wright, & Zhu (2015) also proposed a model where housing collateralizes loans and house price boom and bust can be generated by financial innovation because the liquidity premium on housing is non-monotone in the loan-to-equity ratio. Huo & Ríos-Rull (2016) had a model with heterogenous households, housing, and credit constraints and also showed that financial shocks can generate large drops in housing prices. Favilukis, Kohn, Ludvigson, & Van Nieuwerburgh (2012) studied the impact of systemic changes in housing finance: changes in housing collateral requirements and the change in borrowing costs (the spread of mortgage rates over risk-free security) on how these factors affect risk premiums in housing markets,

and how those risk premiums, in turn, affect home prices. Favilukis, Ludvigson, & Van Nieuwerburgh (2017) developed a quantitative general equilibrium model with housing and collateral constraints to explore what drives fluctuations in house prices to rent ratio. They propose that relaxation of financing constraints leads to a significant boom in house prices. Furthermore, the boom in house prices is entirely the result of a decline in the housing risk premium. Mian & Sufi (2018) showed that speculation is a critical channel through which credit supply expansion affects the housing cycle.

After the financial crisis, there has also been an explosion of interest in the effect of credit expansion on house prices. Justiniano, Primiceri, & Tambalotti (2019) argued that loosening the collateral requirements alone cannot explain the recent housing boom in the US, but there must have been an expansion in the credit supply. The authors argued that house prices rose from 2000 to 2007 without an expansion of leverage. The cause of the housing boom before the recession was an increase in credit supply or available funds rather than an increase in leverage. The rates of mortgages to real estate remained constant. This contradicts the popular view that attributes the housing boom to looser borrowing constraints associated with lower collateral requirements, which would shift the demand for credit. In short, the increase in the supply of credit was the cause of the housing boom. Beyond 2007, the paper argues that there was an increase in collateralizing houses relative to available funds or that available funds for lending decreased, leading to a rise in mortgage rates and a collapse of house prices. More interestingly, in a series of papers, Jorda, Schularick, and Taylor have studied the interplay between credit, house price, and economic performance. Schularick & Taylor (2012) created new data sets for 14 developed countries over 140 years and showed how credit growth is a powerful predictor of financial crises. Jordà, Schularick, & Taylor (2016) claimed that mortgage lending booms were loosely related to financial crises before WWII but have become a more important predictor of financial fragility. The share of mortgages on banks' balance sheets doubled in the later half of the twentieth century, driven by household mortgage lending. Household debt to asset ratios have risen substantially in

the study’s many countries. Financial stability risks have been linked to real estate lending booms. Jordà, Schularick, & Taylor (2015) claims that there has been an increase in the mentality of “bets on the house” in the past century. Mortgage credit has risen dramatically as a share of banks’ balance sheets from about one-third at the beginning of the last century to about two-thirds nowadays. They use a novel IV local projection method to demonstrate that loose monetary conditions lead to booms in real estate and house price bubbles. These, in turn, lead to a higher risk of financial crises. Mortgage booms and house price bubbles have been closely associated with a higher likelihood of a financial crisis. Jordà, Schularick, & Taylor (2017) claimed that a century-long and stable ratio of credit to GDP gave way to rapid financialization and surging leverage in the last forty years. This coincides with a shift in foundational macroeconomic relationships. More financialized economies exhibit more tail risk and tighter real-real and real-financial correlations, including the credit and real estate correlation. The paper also shows that real house prices and mortgages in 17 sample countries display a “hockey stick” pattern. They both stay stable for a long time before ticking up drastically at the end of the sample. It can be shown that house price growth and mortgage growth generally co-move. Favara & Imbs (2015) showed an expansion in mortgage credit has significant effects on house prices using a spatial IV strategy with the US branching deregulation between 1994 and 2005 as an instrument for credit. The treated banks’ credit expansion led to increases in housing demand. Di Maggio & Kermani (2017) showed that a credit expansion could generate a boom and bust in house prices and real activity. The paper uses the exploitation of the same federal deregulation in preemption of local laws against predatory lending in 2004 to gauge the effectiveness of the supply of credit on the real economy.

However, the debate on whether house prices have been the main driving source of the credit cycle or financial conditions (credit) are the main driving force of the house price cycle is still open. In this paper, I will use a dynamic model to explain the relationship between these two variables in both the short-term and long-term.

2.3 Data Description

The sample periods include quarterly data from 1990:Q4 to 2021:Q3. The sample periods were chosen based on the nature of the change in the regulation of credit and housing markets beginning early 1990s. Also, many developed countries started recording their housing price index at the same time.¹ The main source of the data comes from the Bank of International Settlement (BIS). The real housing price index is based on the base index of 2010 as 100. The credit to household data is measured as a percentage of GDP. I take natural log of this series and use the log-transformed series in the model estimation.

Despite their importance, comparable cross-country data on residential property prices are hard to gather. The complicated nature of property transactions and property types, lack of standardization, and short time span of data available further complicate the compilation of a housing price index. To address this data gap, the BIS published a data set on residential property price statistics across the globe.² Combining with actual transaction prices and sources from appraisal and advertised prices, a comparable index of house prices of quarterly frequency is created for each country.

Even though there are other sources with the data regarding credit to households, such as the International Financial Statistics from IMF or the Federal Reserve Economic Data, I decided to use the credit to household data from the BIS for better compatibility and adjustments for breaks methodology in data collection.³ To achieve as long a period as possible for time series data on credit, the construction of the series combined data from institutional sector financial accounts, balance sheets of domestic banks, and international banking institutions.

¹The 17 countries with both credit and house price data available: Australia, Belgium, Canada, Finland, France, Germany, Hong Kong, Italy, Japan, Netherlands, New Zealand, Norway, South Korea, Spain, Sweden, United Kingdom, United States

²Housing price indices are available for 55 countries. https://www.bis.org/publ/qtrpdf/r_qt1409h.htm

³The BIS has constructed long series on credit to the private non-financial sector for 44 economies, both advanced and emerging. Credit is provided by domestic banks, all other sectors of the economy, and non-residents. https://www.bis.org/statistics/totcredit/credpriv_doc.pdf

In this study, I selected the US and UK as two representative countries to use because of the longevity and reliability of the time series data available.⁴ Table 2.1 describes the data used in this paper. House price data tends to fluctuate with greater magnitude than credit series. Furthermore, the housing prices in the UK increase at a faster rate than in the US. Table 2.2 shows the correlation of the series with its lag values of 1 and 2 quarters. The house price series in the UK is more closely correlated with its household credit series than in the US.

Table 2.1: Descriptive statistics

Country	Index	Mean	Max	Min	Frequency	Periods
UK	Δy_t	0.3802	2.6989	-1.7626	Quarterly	1990:Q4-2021:Q3
	Δh_t	0.5926	7.2322	-6.7250	Quarterly	1990:Q4-2021:Q3
US	Δy_t	0.1989	3.5088	-1.9427	Quarterly	1990:Q4-2021:Q3
	Δh_t	0.3004	3.4809	-6.7164	Quarterly	1990:Q4-2021:Q3

Δy_t is growth rate of credit to household series, Δh_t is growth rate of house prices index series. The measurements are in percentage.

⁴For other countries' analysis, I summarized results in subsection 2.5.4.

Table 2.2: Correlation matrix

Country		y_t	y_{t-1}	y_{t-2}	h_t	h_{t-1}	h_{t-2}
UK	y_t	1.0000	0.9991	0.9969		0.9442	0.9491
	h_t	0.9391	0.9314	0.9225	1.0000	0.9975	0.9925
US	y_t	1.0000	0.9983	0.9947		0.7232	0.7415
	h_t	0.7041	0.6891	0.6730	1.0000	0.9951	0.9817

y_t is credit to household series, h_t is housing price index series. Both are log transformed.

2.4 Empirical Model

2.4.1 Model Specification

My proposed model is a multivariate extension of the model used in (Morley, 2007). I use a bivariate unobserved component model to model the dynamics in credit to households as ratio to GDP (y_t) and house prices index (h_t). In my model, the credit and house prices series are log-transformed and are each sum of a trend and cycle components.

I begin with the notations of two series: (Credit) as credit to households as ratio to GDP and (HPI) as Housing Price Index.

$$\ln \frac{Credit}{GDP} = y_t = \tau_{yt} + c_{yt} \quad (2.1)$$

$$\ln HPI = h_t = \tau_{ht} + c_{ht} \quad (2.2)$$

Where τ_{yt} is the trend component of the credit series. c_{yt} is the cycle component of the credit series. Likewise, τ_{ht} and c_{ht} are trend and cycle components of the house prices series.

The trend components of the model have a random walk component and a time-varying

local growth rate component μ_t . This approach follows the structural time series model specification discussed in Beltran, Jahan-Parvar, & Paine (2021) and Campagnoli, Petrone, & Petris (2009):

$$\tau_{yt} = \mu_{yt-1} + \tau_{yt-1} + \eta_{yt}, \quad \eta_{yt} \sim iidN(0, \sigma_{\eta y}^2) \quad (2.3)$$

$$\mu_{yt} = \mu_{yt-1} + \eta_{\mu yt}, \quad \eta_{\mu yt} \sim iidN(0, 0.01) \quad (2.4)$$

$$\tau_{ht} = \mu_{ht-1} + \tau_{ht-1} + \eta_{ht}, \quad \eta_{ht} \sim iidN(0, \sigma_{\eta h}^2) \quad (2.5)$$

$$\mu_{ht} = \mu_{ht-1} + \eta_{\mu ht}, \quad \eta_{\mu ht} \sim iidN(0, 0.01) \quad (2.6)$$

The cycle components of the model follow a VAR(2) process:

$$c_{yt} = \phi_y^1 c_{yt-1} + \phi_y^2 c_{yt-2} + \phi_y^{x1} c_{ht-1} + \phi_y^{x2} c_{ht-2} + \varepsilon_{yt}, \quad \varepsilon_{yt} \sim iidN(0, \sigma_{\varepsilon y}^2) \quad (2.7)$$

$$c_{ht} = \phi_h^1 c_{ht-1} + \phi_h^2 c_{ht-2} + \phi_h^{x1} c_{yt-1} + \phi_h^{x2} c_{yt-2} + \varepsilon_{ht}, \quad \varepsilon_{ht} \sim iidN(0, \sigma_{\varepsilon h}^2) \quad (2.8)$$

Each series is decomposed into a stochastic trend with a local growth rate component ($\tau_{jt}, j = y, h$) and a cyclical component ($c_{jt}, j = y, h$) implying an $I(1)$ process for all the variables. The non-stationarity of these variables is confirmed by the unit root tests, where I do not reject the null hypothesis of unit root for all the variables.⁵ In contrast to (Morley, 2007), I do not impose a common trend restriction. The two variables have their own trend and cycle components, and these components are allowed to have a certain degree of correlation.

Secondly, I specify the dynamics of trend and cycle components. The cyclical component in each series is assumed to follow an AR(2) process, and in additional configurations, lags of the other series. This assumption captures the autocorrelation structures and provides rich dynamics in the data series to enable us to identify all the parameters under the state-space

⁵The detailed results are not reported here for brevity. They are available upon request

model framework.⁶ The trend components are assumed to follow a random walk process with a time-varying local growth rate component. As mentioned above, I do not impose a common trend between the two variables.

Thirdly, I assume the shocks to the trend and cyclical components follow a white noise process but allow for non-zero cross-correlation across series. The shocks to the trend components ($\eta_{jt}, j = y, h$) have a long-run effect on the trend because the trend is assumed to follow a random walk process. The shocks to the cyclical component ($\varepsilon_{jt}, j = y, h$) have a short-run effect on the cycles because the cycles follow a stationary autoregressive process with two lags. The shocks to each trend component are allowed to be correlated across each other, and so are the shocks to the cyclical components. However, I impose the zero correlation between the shocks to the trend components and the shocks to the cycle components within and between series. That is to say, I assume that the shocks that generate a long-run effect differ from those that generate a short-run effect. This assumption isolates the temporary shocks from permanent shocks.

The above dynamic equations can be represented in a state space form where the measurement equation is:

$$\tilde{y}_t = A + H\beta_t \tag{2.9}$$

⁶The cyclical dynamics, in theory, can also be modeled as VAR processes. The presence of cross-correlation among shocks and cross-cycle coefficients in our framework captures the cross-variable dynamics.

Where the measurement components are:

$$\begin{bmatrix} y_t \\ h_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \tau_{yt} \\ c_{yt} \\ c_{yt-1} \\ \tau_{ht} \\ c_{ht} \\ c_{ht-1} \\ \mu_{yt} \\ \mu_{ht} \end{bmatrix}$$

And the transition equations are:

$$\beta_t = F\beta_{t-1} + \tilde{v}_t \quad (2.10)$$

Where the transitory components are:

$$\begin{bmatrix} \tau_{yt} \\ c_{yt} \\ c_{yt-1} \\ \tau_{ht} \\ c_{ht} \\ c_{ht-1} \\ \mu_{yt} \\ \mu_{ht} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & \phi_y^1 & \phi_y^2 & 0 & \phi_y^{x1} & \phi_y^{x2} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & \phi_h^{x1} & \phi_h^{x2} & 0 & \phi_h^1 & \phi_h^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \tau_{yt-1} \\ c_{yt-1} \\ c_{yt-2} \\ \tau_{ht-1} \\ c_{ht-1} \\ c_{ht-2} \\ \mu_{yt-1} \\ \mu_{ht-1} \end{bmatrix} + \begin{bmatrix} \eta_{yt} \\ \varepsilon_{yt} \\ 0 \\ \eta_{ht} \\ \varepsilon_{ht} \\ 0 \\ \eta_{\mu yt} \\ \eta_{\mu ht} \end{bmatrix}$$

The covariance matrix for \tilde{v}_t , denoted Q , is:

$$Q = \begin{bmatrix} \sigma_{\eta y}^2 & 0 & 0 & \sigma_{\eta y \eta h} & 0 & 0 & 0 & 0 \\ 0 & \sigma_{\varepsilon y}^2 & 0 & 0 & \sigma_{\varepsilon y \varepsilon h} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \sigma_{\eta y \eta h} & 0 & 0 & \sigma_{\eta h}^2 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{\varepsilon y \varepsilon h} & 0 & 0 & \sigma_{\varepsilon h}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.01 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.01 \end{bmatrix} \quad (2.11)$$

Regarding variance and covariance estimates, it should be pointed out that, in the variance-covariance matrix of the shocks the the trend and cycle, $\sigma_{\eta y \eta h}$ is the covariance of the shocks to the trend of credit to household as percentage of GDP and house prices index, whereas $\sigma_{\varepsilon y \varepsilon h}$ is the covariance of the shocks to the cycles component of the two variables. The estimates of correlation coefficients, instead of covariances, will be reported in Table 2.4 - 2.6. Additionally, the variance of the local growth rate $\sigma_{\mu it}^2$ is fixed at 0.01 as I followed the steps in Campagnoli et al. (2009) and Beltran et al. (2021) of handling long cycle quarterly data and also for simplifying the computation problem.

The model estimation methodology I used is based on the classical maximum likelihood via the Kalman Filter.⁷ Additionally, my novel contribution to the literature on estimating trend-cycle component models is to overcome the common “curse of dimensionality” problem in estimating a complex unobserved component model, which resulted in perfect multicollinearity and problematic estimation results. To solve this problem, I implemented a Bayesian inference technique called random-walk Metropolis-Hastings sampler and referred to its implementation in Blake & Mumtaz (2017) to obtain draws from the posterior distribution. I use the last 1 million draws from the 1.5 million (Markov chain Monte Carlo) MCMC iterations to analyze the posterior distribution. The Bayesian method is a full information-based approach that uses all moments of the observations. The details of the implementation

⁷See (Kim & Nelson, 1999) and (Durbin & Koopman, 2012) for the details of the estimation procedure.

steps of the model estimation using the random-walk Metropolis-Hastings sampler are shown in appendix A.2.

2.4.2 Parameters Constraints

Another minor novel contribution of the paper is introducing a technique to constraint model parameters in feasible stationary regions by imposing penalties on magnitudes of stationary components. Configuring a feasible estimation procedure for the Unobserved Component model has been a difficult challenge of using the model.

The estimation of the unobserved component model uses a nonlinear log-likelihood function maximization in (Morley, 2007). Estimating this function requires a stationary constraint using numerical optimization. This method is prone to produce corner solutions that are not meaningful.

I did not put stationary constraints directly on the autoregressive parameters. Since such constraints on a VAR(2) system is complex to set up. However, to achieve feasible stationary transitory measurement, I implemented additional terms on the objective log-likelihood function:

$$l(\theta) = -0.5 \sum_{t=1}^T \ln[(2\pi)^2 |f_{t|t-1}|] - 0.5 \sum_{t=1}^T \eta'_{t|t-1} f_{t|t-1}^{-1} \eta_{t|t-1} - w_1 * \sum_{t=1}^T (c_{yt}^2) - w_2 * \sum_{t=1}^T (c_{ht}^2) \quad (2.12)$$

The last two terms in the objective function act as a penalty against too much transitory deviation from zero. Without this penalty, the trend would be linear, or transitory movements would match all the movements in the measured series. This paper will impose the additional penalty terms' weight w_i at 0.0005.

Regarding constraints on the covariance matrix, I applied the same constraints as in (Morley, 2007) to imply the positive-definite covariance matrix.

Table 2.3: Parameters description

Description	Parameter
Log-likelihood value	llv
Credit to household	
Credit to household 1st AR parameter	ϕ_y^1
Credit to household 2nd AR parameter	ϕ_y^2
Credit to household 1st cross cycle AR parameter	ϕ_y^{x1}
Credit to household 2nd cross cycle AR parameter	ϕ_y^{x2}
S.D. of permanent shocks to Credit to household	σ_{ny}
S.D. of transitory shocks to Credit to household	σ_{ey}
Housing Price Index	
Housing Price Index 1st AR parameter	ϕ_h^1
Housing Price Index 2nd AR parameter	ϕ_h^2
Housing Price Index 1st cross cycle AR parameter	ϕ_h^{x1}
Housing Price Index 2nd cross cycle AR parameter	ϕ_h^{x2}
S.D. of permanent shocks to Housing Price Index	σ_{nh}
S.D. of transitory shocks to Housing Price Index	σ_{eh}
Cross-series correlations	
Correlation: Permanent credit to household/Permanent Housing Price Index	ρ_{nynh}
Correlation: Transitory credit to household/Transitory Housing Price Index	ρ_{eyeh}

2.5 Empirical Results

In this section, I will apply the unobserved components state-space model to data from 2 countries: the US and UK. Since I use a Bayesian inference technique, instead of reporting the estimated expected mean and standard deviation as in the frequentist approach, I will report the estimate of the parameters using the median of the posterior distribution and its 10th and 90th percentile value. An estimate of a parameter with a negative 10th percentile and positive 90th percentile values would mean that we do not have evidence to reject the null hypothesis of said parameter equals zero.

Table 2.4: UK Regression Results - All three models

Parameters	VAR2		VAR2 1-cross lag		VAR2 2-cross lags	
	Median	[10%, 90%]	Median	[10%, 90%]	Median	[10%, 90%]
ϕ_y^1	1.9827	[1.9770, 1.9898]	1.4238	[1.3585, 1.4892]	1.4354	[1.3627, 1.5080]
ϕ_y^2	-1.0056	[-1.0126, -0.9985]	-0.4698	[-0.5305, -0.4090]	-0.4946	[-0.5599, -0.4301]
ϕ_y^{x1}			0.0238	[0.0154, 0.0319]	0.0023	[-0.0208, 0.0257]
ϕ_y^{x2}					0.0165	[-0.0075, 0.0399]
ϕ_h^1	1.4119	[1.3987, 1.4238]	1.3173	[1.2647, 1.3701]	1.2844	[1.2233, 1.3458]
ϕ_h^2	-0.4323	[-0.4464, -0.4227]	-0.3315	[-0.3885, -0.2746]	-0.3041	[-0.3686, -0.2409]
ϕ_h^{x1}			-0.0173	[-0.0464, 0.0062]	0.4847	[0.2707, 0.6894]
ϕ_h^{x2}					-0.4960	[-0.6698, -0.3198]
σ_{ny}	0.1055	[0.0896, 0.1254]	0.2714	[0.2150, 0.3155]	0.0737	[0.0463, 0.0987]
σ_{ey}	0.8113	[0.7957, 0.8259]	0.8021	[0.7699, 0.8376]	0.6336	[0.5803, 0.6925]
σ_{nh}	0.0062	[0.0055, 0.0072]	0.0789	[0.0742, 0.0845]	0.0062	[0.0055, 0.0071]
σ_{eh}	1.8647	[1.8332, 1.8845]	1.2242	[1.1886, 1.2613]	1.5020	[1.4080, 1.6063]
ρ_{nynh}	0.0589	[0.0418, 0.0808]	0.0189	[-0.3049, 0.3393]	0.0150	[-0.3101, 0.3306]
ρ_{eyeh}	0.3373	[0.2938, 0.3485]	0.2536	[0.1713, 0.3337]	0.2533	[0.1582, 0.3426]
llv	607.7600	[605.0700, 610.0600]	578.6200	[576.1600, 582.1500]	559.5500	[556.6400, 563.6200]

Note:

UK Bayesian method Metropolis-Hasting random walk posterior distribution estimates

Table 2.5: US Regression Results - All three models

Parameters	VAR2		VAR2 1-cross lag		VAR2 2-cross lags	
	Median	[10%, 90%]	Median	[10%, 90%]	Median	[10%, 90%]
ϕ_y^1	1.4826	[1.4216, 1.5446]	1.2074	[1.1374, 1.2785]	1.2004	[1.1227, 1.2753]
ϕ_y^2	-0.4887	[-0.5500, -0.4280]	-0.2483	[-0.3152, -0.1825]	-0.2554	[-0.3209, -0.1884]
ϕ_y^{x1}			0.0318	[0.0228, 0.0407]	0.0380	[0.0003, 0.0732]
ϕ_y^{x2}					-0.0088	[-0.0451, 0.0297]
ϕ_h^1	1.8594	[1.8276, 1.8915]	1.8038	[1.7700, 1.8363]	1.7999	[1.7658, 1.8345]
ϕ_h^2	-0.8728	[-0.9047, -0.8408]	-0.8261	[-0.8605, -0.7903]	-0.8316	[-0.8687, -0.7942]
ϕ_h^{x1}			0.0104	[0.0007, 0.0204]	0.3305	[0.2535, 0.4066]
ϕ_h^{x2}					-0.2882	[-0.3584, -0.2163]
σ_{ny}	0.0942	[0.0558, 0.1285]	0.2954	[0.2312, 0.3414]	0.0853	[0.0530, 0.1136]
σ_{ey}	0.8282	[0.7616, 0.9059]	0.8631	[0.8287, 0.9012]	0.7278	[0.6672, 0.7955]
σ_{nh}	0.0193	[0.0150, 0.0265]	0.1390	[0.1222, 0.1618]	0.0190	[0.0147, 0.0258]
σ_{eh}	0.8360	[0.7713, 0.9111]	0.8988	[0.8641, 0.9355]	0.8001	[0.7321, 0.8735]
ρ_{nynh}	0.0082	[-0.3118, 0.3230]	0.0082	[-0.3117, 0.3226]	0.0167	[-0.2998, 0.3328]
ρ_{eyeh}	0.1000	[-0.0181, 0.2185]	0.1537	[0.0399, 0.2619]	0.1642	[0.0460, 0.2764]
llv	197.7900	[195.5700, 201.0700]	204.9400	[202.4200, 208.4500]	187.7900	[184.8500, 192.1700]

Note:

US Bayesian method Metropolis-Hasting random walk posterior distribution estimates

Table 2.6: VAR(2) 1 cross-lag : UK and US regression results

Description	Para.	UK VAR2 1-cross lag		US VAR2 1-cross lag	
		Median	[10%, 90%]	Median	[10%, 90%]
Credit to household 1st AR parameter	ϕ_y^1	1.4238	[1.3585, 1.4892]	1.2074	[1.1374, 1.2785]
Credit to household 2nd AR parameter	ϕ_y^2	-0.4698	[-0.5305, -0.4090]	-0.2483	[-0.3152, -0.1825]
Credit to household 1st cross cycle AR parameter	ϕ_y^{x1}	0.0238	[0.0154, 0.0319]	0.0318	[0.0228, 0.0407]
Credit to household 2nd cross cycle AR parameter	ϕ_y^{x2}				
Housing Price Index 1st AR parameter	ϕ_h^1	1.3173	[1.2647, 1.3701]	1.8038	[1.7700, 1.8363]
Housing Price Index 2nd AR parameter	ϕ_h^2	-0.3315	[-0.3885, -0.2746]	-0.8261	[-0.8605, -0.7903]
Housing Price Index 1st cross cycle AR parameter	ϕ_h^{x1}	-0.0173	[-0.0464, 0.0062]	0.0104	[0.0007, 0.0204]
Housing Price Index 2nd cross cycle AR parameter	ϕ_h^{x2}				
S.D. of permanent shocks to Credit to household	σ_{ny}	0.2714	[0.2150, 0.3155]	0.2954	[0.2312, 0.3414]
S.D. of transitory shocks to Credit to household	σ_{ey}	0.8021	[0.7699, 0.8376]	0.8631	[0.8287, 0.9012]
S.D. of permanent shocks to Housing Price Index	σ_{nh}	0.0789	[0.0742, 0.0845]	0.1390	[0.1222, 0.1618]
S.D. of transitory shocks to Housing Price Index	σ_{eh}	1.2242	[1.1886, 1.2613]	0.8988	[0.8641, 0.9355]
Correlation: Permanent credit to household/Permanent HPI	ρ_{nynh}	0.0189	[-0.3049, 0.3393]	0.0082	[-0.3117, 0.3226]
Correlation: Transitory credit to household/Transitory HPI	ρ_{eyeh}	0.2536	[0.1713, 0.3337]	0.1537	[0.0399, 0.2619]
Log-likelihood value	llv	578.6200	[576.1600, 582.1500]	204.9400	[202.4200, 208.4500]

Note:

UK - US Bayesian method regression results

The tables 2.4 and 2.5 shows maximum-likelihood estimates of all three Unobserved Component VAR(2) models. The first model is a parsimonious UC VAR(2) model with no cross-cycle correlation terms (ϕ_y^x and ϕ_h^x are set to be zero). The next two models introduces one and two cross-cycle coefficients on the lags of cyclical component respectively. The model selection criteria, log-likelihood value, suggests that the best fit model is VAR(2) with 2 cross lags. But because this paper focuses on the estimate of the causal cross cycle correlation parameter, I select the second model - VAR(2) with only 1 cross lag in the cycle component as the one to limit the scope of analysis on.

2.5.1 Dynamic Relationship between Credit to Household and Housing Price

From Table 2.6, the VAR(2) - 1 cross lag model regression results suggest that transitory shocks dominate permanent shocks in terms of variation in both household credit and housing price variables. The standard deviation of the shocks in the cycle of credit σ_{ey} is 0.8021 in the UK and 0.8631 in the US, much higher than the standard deviation of the shocks to the trend of credit σ_{ny} in the UK of 0.2714 and the US of 0.2954. The same applies to housing prices, the standard deviation of the shocks in the cycle of housing price σ_{eh} is 1.2242 in the UK and 0.8988 in the US, higher than the standard deviation of the shocks to the trend of housing price σ_{nh} in the UK of 0.0789 and the US of 0.1390. This result also indicates that variations in the housing price cyclical components of the UK are bigger than in the US. In contrast, variations in other components of the UK do not differ from the US. Regarding the estimated parameters, the sum of AR parameters of the cyclical components in all three models is smaller, although close to one. This implies that shocks to the cycle are persistent but will eventually dissipate.

The correlation analysis of the shocks to the cyclical components among the two variables suggests that cyclical variation among housing price and household credit is strongly positively correlated. The estimate ρ_{eyeh} at 0.2536 for the UK and 0.1537 for the US suggest

that transitory shock to housing credit is closely positively correlated to transitory shock in housing price. This implies that a transitory increase in household credit is correlated with an appreciation in housing prices above their long-run trend.

The correlation analysis of the shocks to the trends ρ_{nynh} among the two variables reveals no significant correlation between shocks to the trend components of household credit and housing price. The median estimate for the permanent component's correlation is much smaller than the correlation of the transitory components.

Lastly, I analyze the correlation value of the time-varying local growth rate components μ_{it} from both household credit and housing price index trends as specified in equations 2.3-2.6. After estimating the unobserved components, as in figures 2.2 and 2.5, we can calculate the correlation values between those two local growth rate components. The correlation value $\rho_{\mu yh}$ for the UK at 0.5603 is significantly higher than that for the US at -0.0574. This result supports the hypothesis that the underlying trend components of household credit and housing price index in the UK are more correlated than in the US.

Overall, the above analyses' results suggest that the two variables' short-run and long-run dynamics are very different. Therefore, there is a benefit in decomposing the series into trend and cyclical components.

2.5.2 Trend-Cycle Decomposition

The following graphs show the UC forecast series against the actual data series. As discussed in the previous subsection, I focus our analysis on the VAR(2) 1 cross-lag model in Figures 2.2 and 2.5.

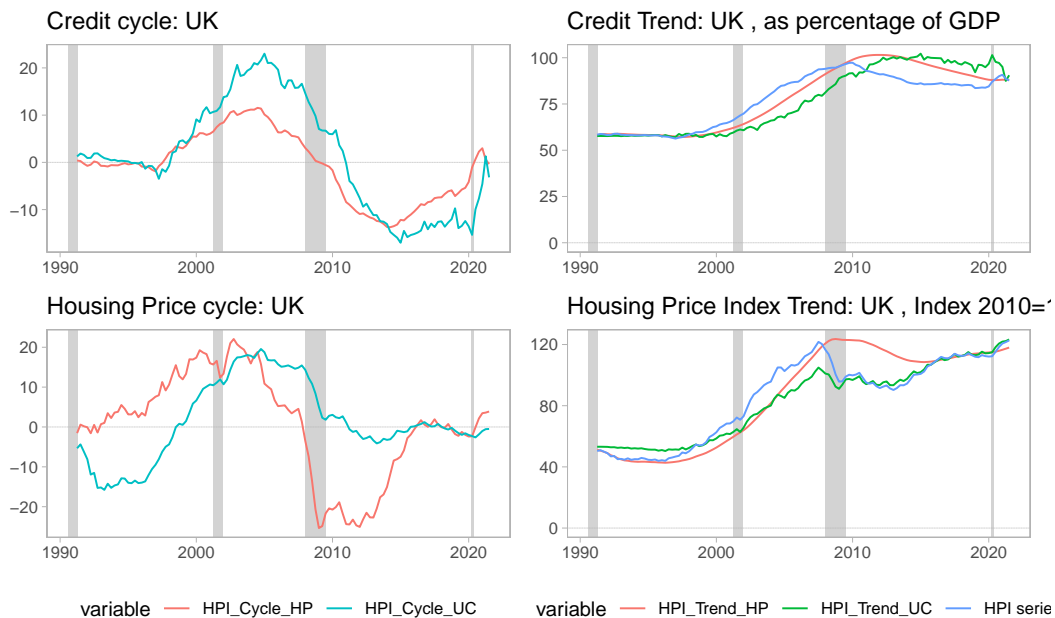


Figure 2.1: UK VAR(2)

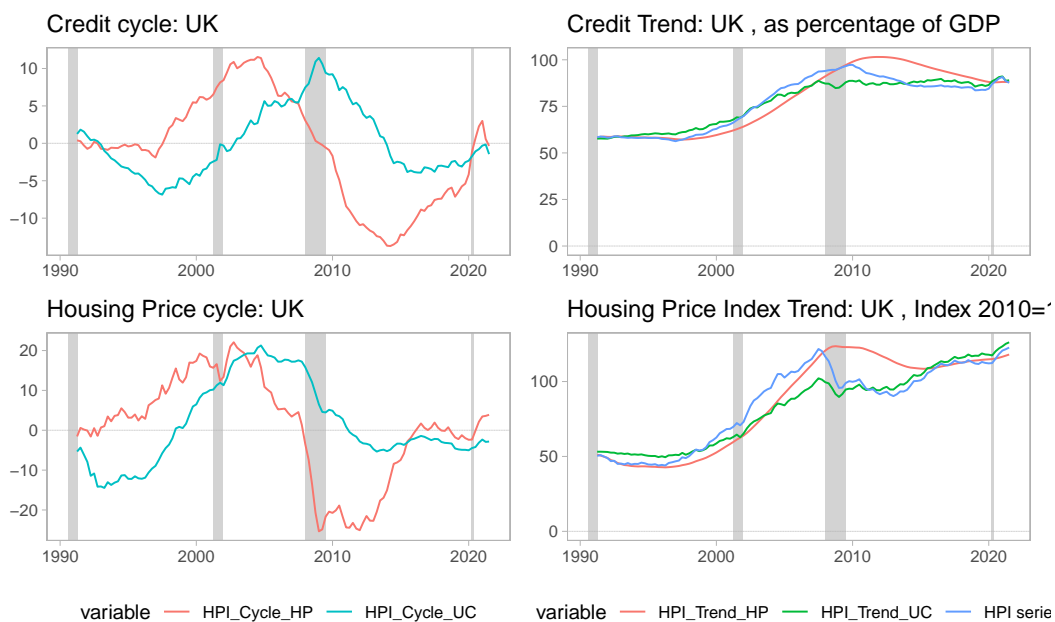


Figure 2.2: UK VAR(2) 1 cross-lag

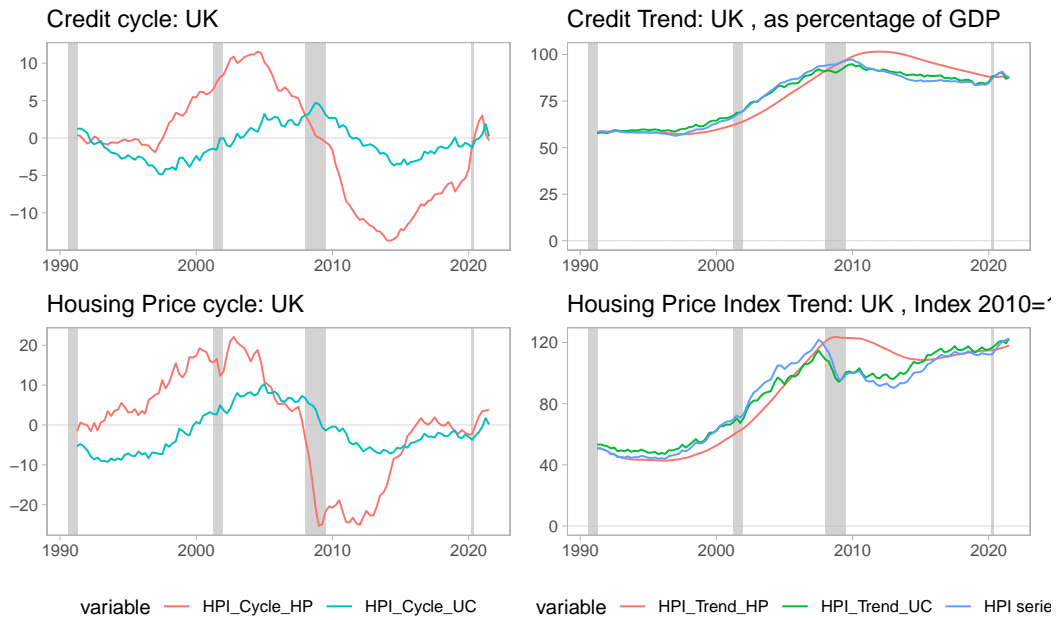


Figure 2.3: UK VAR(2) 2 cross-lags

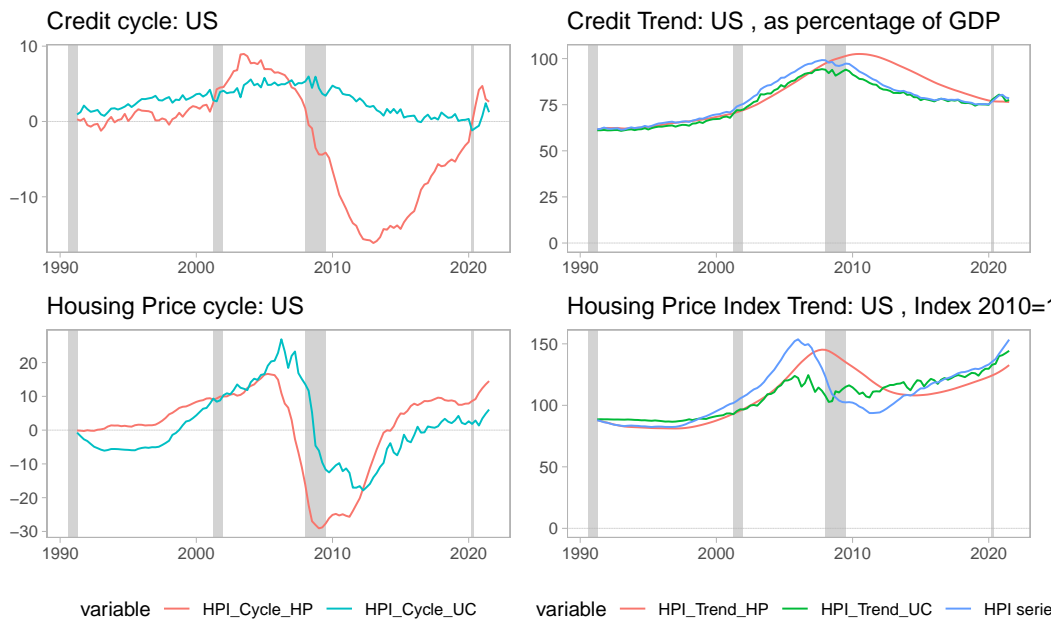


Figure 2.4: US VAR(2)

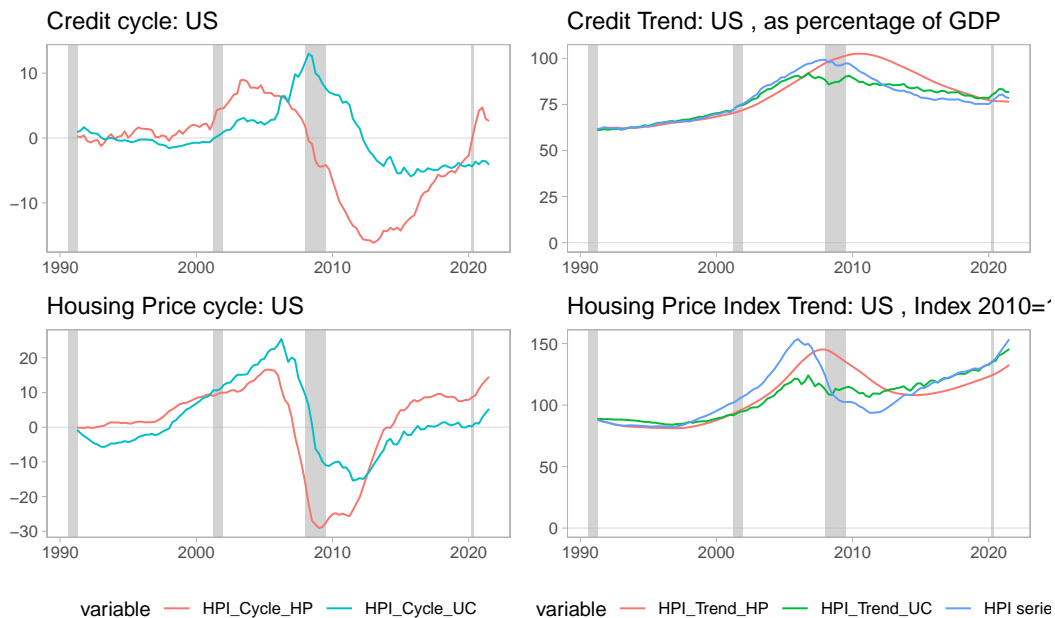


Figure 2.5: US VAR(2) 1 cross-lag

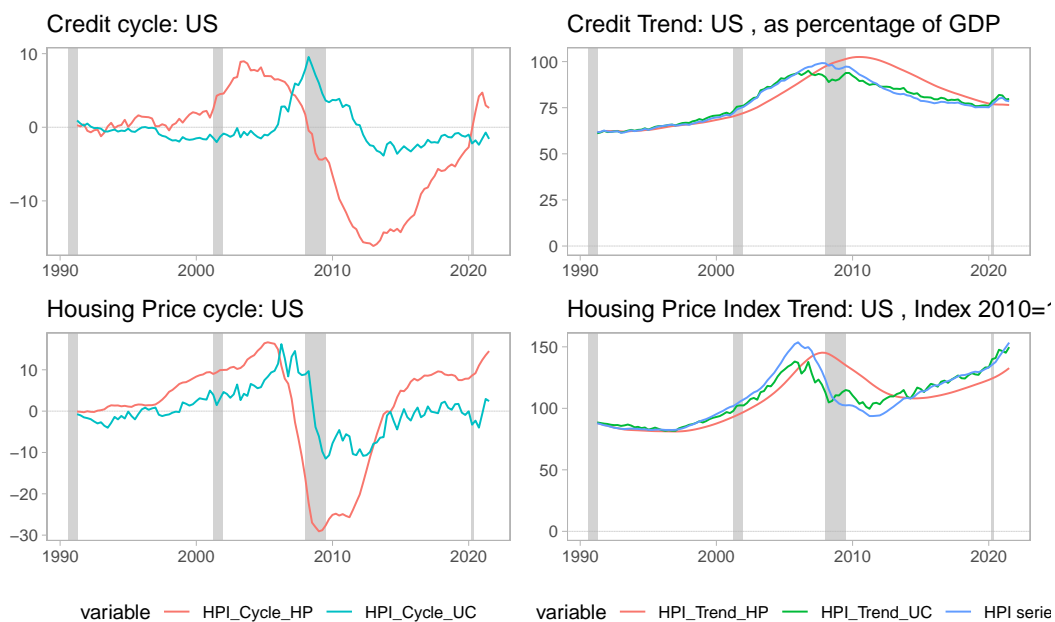


Figure 2.6: US VAR(2) 2 cross-lags

In this subsection, I decompose the trend and cycle of household credit and housing price using the correlated unobserved component model. The stochastic trend in the multivariate UC model captures the long-run evolution in household credit, housing price, and the effect of the recent global crisis. In the long run, there is an increasing trend in the housing price index. The household credit trend is also increasing, but since the series is credit to households as a ratio to GDP, the rate at which the household credit trend increases is lower than that of the housing price index. There was a downward movement of the trend components in both credit and housing prices after the crisis. However, the housing price index trends made a quicker recovery than household credit.

The cyclical components of the model capture the evolution of household credit, housing price, and their dynamic relationship. In Figures 2.1-2.6, we can see an increase in the credit transitory component before the crisis of 2008-2009 happened, and there is a negative shock to the transitory component of housing price after the recession is captured in the model as well. Specifically, in Figures 2.2 and 2.5, the VAR(2) model with 1 cross-lag coefficients identified the timing of temporary gaps increase in both credit and housing price before the crisis more accurately than the one-sided HP filter could.

It is also important to point out that our models capture a significantly bigger positive gap in transitory shock in both credit and house price than a Hodrick-Prescott (HP) filter would for the US in Figure 2.5. Our model utilizes additional information from decomposed long-run and short-run variables, which were extracted from a nonstochastic time series. Another approach in dealing with nonstochastic time series is to first-differencing the series, which loses much important information from a limited sample. Thus when dealing with a time series of low frequency and long-term assets such as housing prices, it is worthwhile to consider using the unobserved component model rather than simply applying an HP filter or first-differencing since it reveals more lower frequency information. The graphs indicate that the magnitude of transitory shocks the models capture is higher, and the movement frequency of the cycles is lower than that of other decomposition methods (HP filter).

2.5.3 Predictive Ability of Cyclical Components

A novel contribution of this paper is to introduce the cross-cycle parameter ϕ_h^{xt} and ϕ_y^{xt} in which it measures the effect of a change in previous periods' credit transitory component on the current housing price transitory component and vice versa. From Table 2.6, in both cross-cycle regressions in the UK and US, we can observe that there is a significant positive effect of last period house price cycle deviation on current household credit cycle component (ϕ_y^{x1}). While the coefficients of transitory household credit deviation on housing price index (ϕ_h^{x1}) are smaller, and even statistically insignificant in the case of the US. This holds true for the 2-crosscycle lags model also on the other regression results Tables 2.4 and 2.5; when we take the sum of the 2-crosscycle lags coefficients, they show similar results. This showed evidence that past transitory shocks to house price credit will cause a positive deviation in future transitory household credit. On the contrary, the effect in the opposite direction is much smaller.

2.5.4 Cross Countries Analysis

In this subsection, I will further discuss estimates of the causal coefficient ϕ_y^{xt} and ϕ_h^{xt} in cross-country settings. The window for sample data I selected starting from 1990:Q4 - 2021:Q3 coincides with the period when many developed countries started recording their housing price index. This created an opportunity for us to compare the estimates of the causal coefficient parameters. Specifically, I estimated our VAR(2) 1 cross-lag model using data from 17 developed countries, as shown in Table 2.7.

The overall result shows that there is cross-countries evidence to further support the hypothesis established in the previous subsection 2.5.3 that past transitory shocks to house price credit will cause a positive deviation in future transitory household credit; or that the ϕ_y^{x1} coefficient parameter estimates are positive and statistically significant.

Out of the 17 countries in our sample, 11 countries in Europe and North America, except

Table 2.7: VAR(2) 1 cross-lag : Cross countries estimates

Country	ϕ_y^{x1} HPI on Credit		ϕ_h^{x1} Credit on HPI	
	Median	[10%, 90%]	Median	[10%, 90%]
Australia	0.0157	[-0.0093, 0.0412]	0.0521	[0.0014, 0.1060]
Belgium	0.0279	[0.0013, 0.0559]	-0.0656	[-0.0980, -0.0339]
Canada	0.0191	[0.0032, 0.0332]	-0.0152	[-0.0343, 0.0025]
Finland	0.0080	[0.0017, 0.0156]	0.0085	[0.0021, 0.0156]
France	0.0298	[0.0185, 0.0411]	-0.0643	[-0.1098, -0.0241]
Germany	0.0728	[0.0500, 0.0917]	-0.0061	[-0.0282, 0.0052]
Hong Kong	-0.0031	[-0.0079, 0.0019]	-0.0629	[-0.0836, -0.0453]
Italy	0.1001	[0.0895, 0.1063]	-0.0027	[-0.0072, 0.0014]
Japan	-0.0088	[-0.0326, 0.0174]	0.1659	[0.1202, 0.2173]
Netherlands	0.0058	[-0.0039, 0.0166]	-0.0043	[-0.0156, 0.0070]
New Zealand	0.0078	[-0.0035, 0.0199]	-0.0139	[-0.0249, -0.0036]
Norway	0.0109	[0.0097, 0.0116]	0.0059	[0.0047, 0.0066]
South Korea	0.0106	[-0.0033, 0.0308]	0.0027	[-0.0251, 0.0369]
Spain	0.0144	[0.0003, 0.0331]	0.0051	[-0.0023, 0.0146]
Sweden	0.0159	[0.0071, 0.0252]	0.0400	[0.0218, 0.0617]
United Kingdom	0.0238	[0.0154, 0.0319]	-0.0173	[-0.0464, 0.0062]
United States	0.0318	[0.0228, 0.0407]	0.0104	[0.0007, 0.0204]

Note:

Cross Countries causal coefficients

for the Netherlands, have a positive and significant estimate for ϕ_y^{x1} (HPI on Credit causal coefficient). On the other hand, East Asian countries such as Japan, South Korea, and Hong Kong (SAR), along with Australia, and New Zealand, have insignificant estimates for ϕ_y^{x1} .

As for the other parameter, ϕ_h^{x1} (Credit on HPI causal coefficient), only six countries have positive and significant estimates: Australia, Finland, Japan, Norway, Sweden, and the US. Furthermore, three of those positive estimates are also smaller in magnitude compared to their counterpart ϕ_y^{x1} , except for Australia, Japan, and Sweden. Interestingly, there are also countries with negative and significant estimates for ϕ_h^{x1} , which are Belgium, France, Hong Kong, and New Zealand. Their negative values implied that a positive increase in transitory household credit would have a negative effect on future transitory household prices. Overall, from all 17 available countries' estimates, I found no strong evidence to support ϕ_h^{x1} (Credit on HPI causal effect) to be significant.

The two outliers in this analysis are Japan and Italy. Italy has a much higher ϕ_y^{x1} (HPI on Credit causal coefficient) than other countries, this could be a result of the country experiencing a boom in house prices before the global crisis. And after its bust in 2009, house prices have not been able to recover since, see graphs in Figure A.13. In contrast, Japan has a much higher ϕ_h^{x1} (Credit on HPI causal coefficient) than other countries. This could be due to Japan's economy's unique characteristics in which it has been experiencing three consecutive lost decades of stagnation trying to recover after the Japan Banking crisis in early 1990. Its household credit and house price index have declined ever since early 1990. However, there are signs of recovery in recent years for Japan, see graphs in Figure A.14.

2.6 Comparison with other Decomposition Methods

In this section, I check the robustness of the model results by comparing the estimated trend-cycle from my multivariate approach with univariate trend-cycle decomposition using different methods. In addition to estimating the correlation between shocks to the permanent and transitory component, the use of a multivariate model, in theory, should also provide us with a superior measurement of trend and cycle components compared to the univariate models. To test this hypothesis, I also perform trend-cycle decomposition using the univariate models (Figure 2.7 and 2.8). The univariate models include an HP filter model and a univariate VAR(2) UC model. The HP filter method uses an algorithm to smooth the original data series to estimate the trend component and the difference between them, which is the cyclical component. The parameter value λ is set at 125,000 as suggested by Hodrick and Prescott and Ravn & Uhlig (2002) for the quarterly data corresponding to a cyclical period of 15-17 years. The univariate UC model only uses a single series of either credit to household or house prices index to decompose a stochastic trend component and a cyclical component with the same specification as the multivariate UC model.

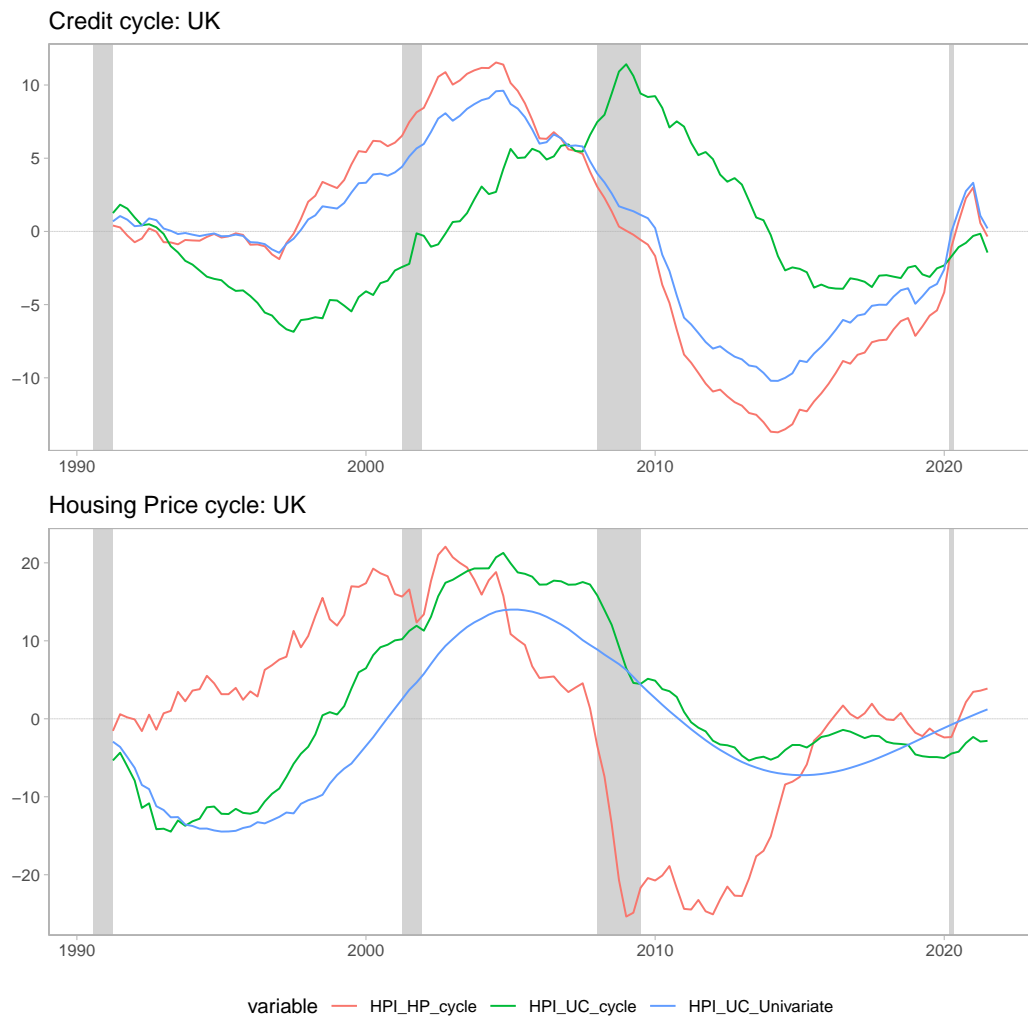


Figure 2.7: Comparing Multivariate UC cycles with alternate decompositions: UK

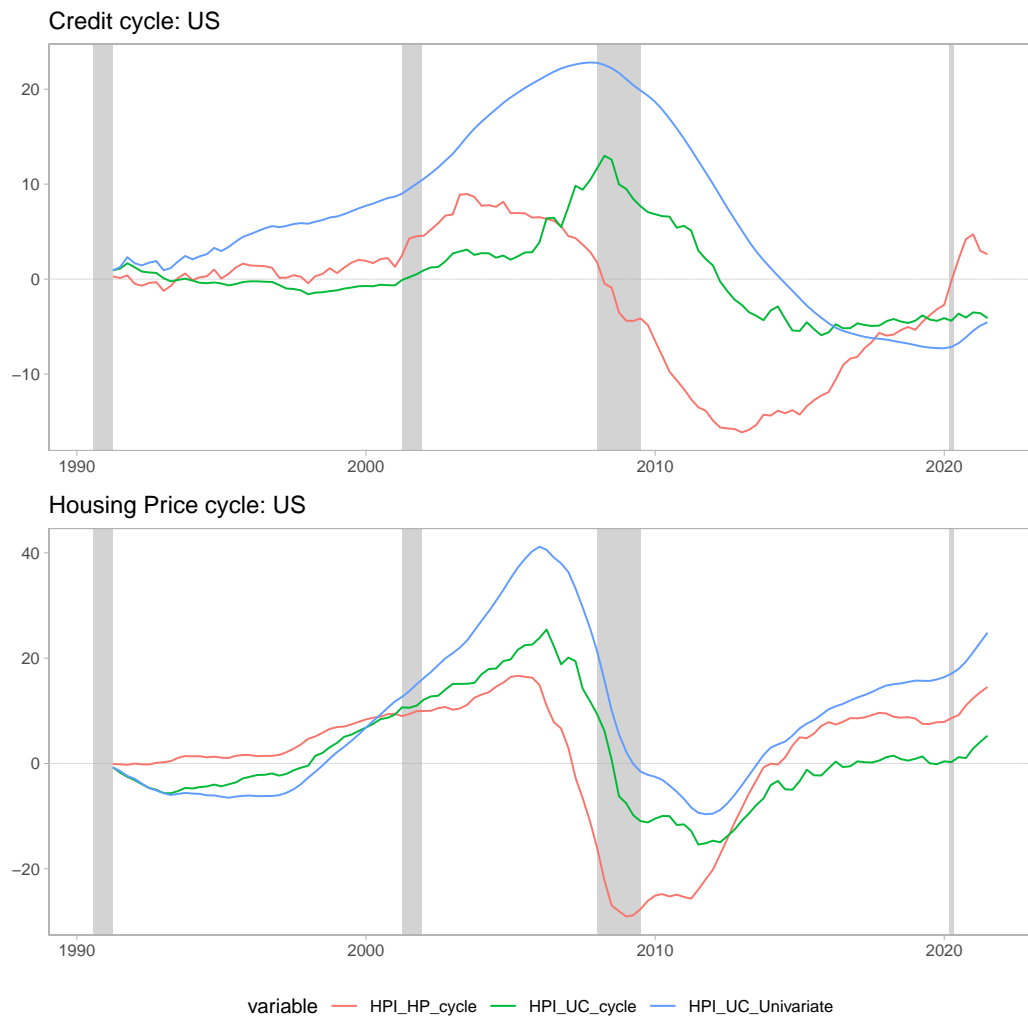


Figure 2.8: Comparing Multivariate UC cycles with alternate decompositions: US

The results from Figure 2.7 and Figure 2.8 suggest that the estimate of trends and cycles obtained from the multivariate UC model can capture the dynamics of the two variables during the sample period. The two univariate models fail to generate realistic trends and cycle series by ignoring the relationship between the two variables of interest. The HP cycle seems to do very well at remaining stationary, but by doing so, it missed out on capturing the boom of house prices in the US, leading to the Great Recession of 2009. The cycle from the univariate UC model is close to the multivariate counterpart but failed to fully indicate the magnitude of boom and bust in house prices in the UK before and after the crisis. Overall, it is clear from the analysis above that there is a valuable pay-off in utilizing information from extracting permanent and transitory components of credit to household and house prices index to study the dynamics of the two variables.

2.7 Conclusion

My study is based on the idea that house prices and credit are jointly determined, affecting each other in the short and long run. I decompose the movements of the two variables of interest into a permanent and transitory component. The correlations among the cyclical components support the idea that the rise of house prices is associated with an increase in household credit above its long-run trend. My multivariate model captures the dynamic features of the household credit and house prices series and performs better than univariate benchmarks in capturing the boom and bust during the last two decades. Additionally, employing cross-correlation effects on the transitory components of the two series allows me to test the predictive ability of the cyclical components and find evidence to support that a house prices gap can positively predict a household credit gap. These findings suggest macroprudential policy implications since house prices are increasingly becoming a more important topic.

Further development for this paper should include studying on policy implications of

credit and house price gaps with high magnitudes. More robust optimal constraints on parameters to ensure stability and model robustness rather than an ad-hoc approach to selecting weights. Lastly, it would be meaningful to link the information from this paper’s causal effect coefficients estimates to the bigger picture regarding macroprudential policy models that includes housing price and household credit, such as systemic financial stability model⁸ or house price growth-at-risk models⁹.

⁸Alessi & Detken (2018)

⁹Deghi, Katagiri, Shahid, & Valckx (2020)

Chapter 3

Measuring Credit Gap

3.1 Introduction

There is a consensus in macroeconomics and finance literature about the role of credit in overall macroeconomic activity. Not surprisingly, policymakers and practitioners pay a significant amount of attention to examine if credit growth in the economy is excessive. The problem, however, is that there is no unanimity on how to measure excessive credit. Among different measures of excessive credit, the credit-to-GDP gap measure published by the Bank of International Settlements (BIS) and proposed by C. E. V. Borio & Lowe (2002) has received widespread attention. This measure is estimated by the deviations of the credit-to-GDP ratio from a one-sided Hodrick & Prescott (1997) (HP) filter with a large smoothing parameter (400,000 for quarterly data). In subsequent work, C. E. V. Borio & Drehmann (2009), Drehmann, Borio, Gambacorta, Jiménez, & Trucharte (2010) , and Drehmann, Borio, & Tsatsaronis (2011) show that the credit-to-GDP gap was the best single Early Warning Indicator across different measures that they examined. Based on this work at the Bank of International Settlements (BIS), the Basel Committee for Banking Supervision (BCBS)(2010) has singled out the credit-to-GDP gap as a useful guide for setting countercyclical capital buffers (BCBS 2010b).

The credit gap ¹ measure proposed by the BIS is a measure of the cycle based on the HP trend-cycle decomposition of credit-to-GDP ratio, wherein the trend is the best estimate of where the variable will be in the long-run and the cycle is temporary fluctuations around the trend. Since there is no unique method of decomposing a series into a trend and a cycle, it is worth asking if the HP cycle measure used by the BIS is an appropriate measure

¹I use the credit gap and credit-to-GDP gap interchangeably throughout this paper.

of the credit gap. In fact, in a recent work Drehmann & Yetman (2021) make a similar argument and attempt to use the idea from Hamilton filter to use linear projections to get an estimate of the credit gap. They find that credit gaps based on linear projections in real time perform poorly when based on country-by-country estimation, and are subject to their own endpoint problems. But when they estimate as a panel, and impose the same coefficients on all economies, linear projections perform marginally better than the baseline credit-to-GDP gap, with somewhat larger improvements. This reinforces the argument that there is no unique method of decomposing a non-stationary series into a trend and a cycle, and that I need to take into account model uncertainty in estimating the credit gap.

I propose to use the idea in Nelson (2008) , where he argues that if measured cycle component is temporary then it predicts future growth rates of the opposite sign. For example, if the credit-to-GDP ratio is below trend, then a recovery in the ratio will require future growth in the credit-to-GDP ratio at an above-average rate. Conversely, if it is above trend, I can reasonably expect tepid growth in the coming time period. According to Nelson (2008) , predictability is the essence of ‘transitory’ variation, and is expected to be reversed in future periods. Predictability of the cycle provides us with a metric for measuring the effectiveness of alternative decompositions. This idea has also been emphasized by Cogley & Nason (1995) , Hodrick & Zhang (2003) , Orphanides & Norden (2002) , Rotemberg & Woodford (1996) and Wakerly, Scott, & Nason (2006) . The idea of predictability implies that a credit gap measure that encompasses all the information from other measures about future movements of the credit-to-GDP ratio is the appropriate credit gap measure. However, in practice, I do not have a measure of the credit gap that possesses this property. This is further complicated by instability in the predictive ability of different credit gap measures that are ubiquitous in the macro forecasting literature.

I combine the idea of predictability with model uncertainty in trend-cycle decomposition by combining credit gap measures obtained from different trend-cycle decomposition methods

using weights obtained from the out-of-sample forecasting exercise ². The weights are based on the Bates & Granger (1969) algorithm, where I perform a horse race among the most popular trend-cycle decomposition methods in an out-of-sample forecasting exercise ³. The relative weight on cycles from different trend-cycle decomposition methods is based on its forecast error variance in predicting out-of-sample credit-to-GDP ratio changes. In addition to taking into account model uncertainty by assigning weights on different models, this approach also handles the instability in the relative forecasting performance of different trend-cycle decomposition methods by assigning time-varying weights on different methods. These weights are time-varying since my method recalculates the weights based on the predictive ability of the model for each iteration in my recursive forecasting exercise.

I apply this approach to the credit-to-GDP ratios of the U.K. and the U.S. and estimate a credit gap measure for the 1994:Q1-2020:Q2 sample period. My results show that the weighted credit gap measure based on my approach dominates credit gaps from other trend-cycle decomposition methods in out-of-sample forecasting of changes in the credit-to-GDP ratio. My combined gap measure leads to an improvement of 13 percent for the U.S. and 9 percent for the U.K. for 1-4 quarter ahead forecast horizon over a benchmark AR(1) model. In contrast, the forecasting performance of the BIS gap is worse than the benchmark AR(1) model for both the U.S. and the U.K. In addition, the relative forecasting performance of different methods vary across two countries confirming model uncertainty in the estimation of credit gaps for different methods. My estimated combined credit gap measure for the U.S. and the U.K. exhibits smooth behavior with a smaller amplitude than the BIS gap. In addition, I observe a clear pattern in early detection of trough date by combined gap in both the countries in comparison to the BIS gap. Finally, the combined gap measure proposed in this paper does not suffer from the endpoint problem usually associated with HP filters.

²I use linear trend, quadratic trend, the HP filter, the Ravn & Uhlig (2002) modification of the HP filter, C. E. V. Borio & Lowe (2002) modification of the HP filter, Clark's unobserved component model (1987), Beveridge & Nelson (1981) Decomposition, Hamilton (2018) filter.

³My forecasts are based on 'quasi-real-time' data that uses revised data, but only the observations up to the historical date. When a decomposition requires estimation of parameters, they are re-estimated at each date before computing the cycle estimate.

The remainder of the paper is organized as follows: Section 2 provides brief literature review; Section 3 discusses the data; Section 4 presents my empirical methodology; Section 3 provides a description of the data used in my empirical analysis; Section 5 presents the empirical results; and Section 6 concludes.

3.2 Related Literature

The importance of the credit gap measured as proposed by C. E. V. Borio & Lowe (2002) can be gauged from the fact that Basel III suggests that policymakers use it as part of their countercyclical capital buffer frameworks. The baseline credit gap of C. E. V. Borio & Lowe (2002) is calculated as deviations of the credit-to-GDP ratio from a one-sided Hodrick-Prescott (HP) filter with large smoothing parameter-400,000 for quarterly data. The smoothing parameter is much larger than the one used for quarterly data in the business cycle literature. This choice of smoothing parameter is rationalized on the ground that credit cycles are on average about four times longer than standard business cycles and crises tend to occur once every 20–25 years (Drehmann et al., 2010) . Drehmann & Yetman (2021) use the critique of the HP filter outlined in Hamilton (2018) and examine different measures of the credit gap based on different horizons in local projection models. They find that credit gaps based on linear projections in real time perform poorly when based on country-by-country estimation, and are subject to their own endpoint problems. But when they estimate as a panel, and impose the same coefficients on all economies, linear projections perform marginally better than the baseline credit-to-GDP gap, with somewhat larger improvements concentrated in the post-2000 period and for emerging market economies. Several other papers have criticized the HP filter-based credit gap measure. These criticisms are based on the ground that HP filter-based cycle measures generate spurious dynamics and suffer from endpoint problems.

Several papers have tried to resolve this problem. Aikman, Haldane, & Nelson (2015)

have used band-pass filters to derive a measure of the credit gap. Galati, Hindrayanto, Koopman, & Vlekke (2016) estimate a financial cycle using a multivariate unobserved-components model on the credit-to-GDP ratio, total credit, and house prices for six economies, and find that the resulting medium-term cycles vary in terms of length and amplitude across countries and over time. Schüller, Hiebert, & Peltonen (2015) derive a financial cycle based on the common frequencies of credit and asset prices and find that this measure outperforms the credit-to-GDP gap in predicting systemic banking crises at horizons of one-to-three years. Aldasoro, Borio, & Drehmann (2018) show that combining various indicators of excessive debt with property prices can help to improve financial crisis prediction.

As is widely known, the HP filter is one measure of trend-cycle decomposition. Trend-cycle decomposition has a very rich history in business cycle literature. Yet there is no consensus on what constitutes appropriate decomposition methods. In this paper, I take an agnostic approach and utilize the idea that the cyclical component in the decomposition should have predictive power for future growth of the variable. This idea has been explored in Nelson (2008) to compare the forecasting performance of different trend-cycle decomposition methods for real GDP. I take the predictive ability argument for the cycle one step further and apply it in the context of the estimation of the credit gap. It is widely accepted that model uncertainty reigns supreme in applying a particular trend-cycle decomposition method. To take into account this model uncertainty, we weighted the cycles obtained from different decomposition methods using Bates and Granger forecast combination method.

3.3 The Data

My sample periods include quarterly data from 1960:Q1-2020:Q2. My variable of interest is the credit-to-GDP ratio for the U.S. and the U.K. The data has been sourced from the Bank of International Settlements (BIS). The measure of credit is total credit to the private non-financial sector, as published in the BIS database, capturing total borrowing from all

domestic and foreign sources. I do not include data for the recent pandemic period in my analysis. The credit-to-GDP ratios for these two countries are plotted in Figure B.1. As can be seen from the figure, these two series are clearly non-stationary. Unit root tests confirm this, with the null of unit root not being rejected at all conventional levels. This is robust to the use of different unit root tests.

3.4 Empirical Methodology

3.4.1 Trend-Cycle Decomposition

My combined measure of the credit gap is based on different trend-cycle decomposition methods for the credit-to-GDP ratio. In the literature, several methods have been proposed to decompose a non-stationary series into a trend and a cycle. Since there is no consensus on the true model of trend-cycle decomposition, I take an agnostic view in this paper. For my purposes, I use seven different measures of trend inflation. All these trend-cycle decomposition methods are based on the premise that a non-stationary series is the sum of a trend and a stationary cyclical component:

$$y_t = \tau_t + c_t \tag{3.1}$$

where y_t is an $I(1)$ process, and for my purposes, credit-to-GDP ratio. τ_t is trend component and c_t is cyclical component, and is stationary. Trend is usually modeled as random walk and cycle is modeled as following some $ARMA(p,q)$ process. In this paper, I use 8 different decomposition methods: linear trend, quadratic trend, the HP filter, the Ravn & Uhlig (2002) modification of the HP filter, C. E. V. Borio & Lowe (2002) modification of the HP filter, Clark's unobserved component model (Clark's unobserved component model (1987) , Beveridge & Nelson (1981) Decomposition, Hamilton (2018) filter. I do not use frequency-based filters because my approach uses the forecasting property of the cycle. The

observations at the end of the sample are either not available or highly volatile for frequency-based filters. The linear trend model is based on a deterministic time trend and assumes all variation in headline inflation is transitory, and hence due to cyclical components. The HP filter is an atheoretical smoothing method to obtain trend and cycle components of non-stationary series and is very popular in macroeconomics and finance literature. I follow the original prescription of Hodrick & Prescott (1997) and use smoothing parameter $\lambda=1600$ for the quarterly credit-to-GDP ratio data. Ravn & Uhlig (2002) have suggested modifying the smoothing parameter to account for the frequency of the data. Following their suggestion I use smoothing parameter $\lambda=3000$ for the other model. I call this model RU in my exercise. Higher λ yields a much smoother trend. My unobserved component model is based on the original Clark (1987) model. In particular, trend follows a random walk and cycle has ARMA (p,q) representation. Both trend shocks and cyclical shocks have time-varying volatility.

$$\tau_t = \tau_{t-1} + \eta_t, \eta_t \sim iid(0, \sigma_\eta^2) \quad (3.2)$$

$$c_t = \Phi(L)c_t + u_t, u_t \sim iid(0, \sigma_u^2) \quad (3.3)$$

Beveridge & Nelson (1981) methodology decomposes a non-stationary series into a random walk component and a stationary component which is the cycle of the non-stationary series. The BN decomposition of z_t has the following representation:

$$y_t = y_0 + \mu t + \Psi(1) \sum_{k=1}^t u_t + \tilde{u}_t - \tilde{u}_0 \quad (3.4)$$

where $\Psi(1) \sum_{k=1}^t u_t$ is the stochastic trend and $\tilde{u}_t - \tilde{u}_0$ represents the cycle. BN proposed that the long-run forecast is a measure of trend for time series such as GDP that do not follow a deterministic path in the long run. They showed that if the series is stationary in first differences, then the estimated trend is a random walk with drift that accounts for growth, and the cycle is stationary. In contrast to HP and UC decomposition, the BN decomposition attributes most variation in non-stationary series to trend shocks while the cycles are short

and brief.

In addition to these decomposition methods, I also use Hamilton (2018) approach to calculate credit gap. Hamilton (2018) argues that HP filter produces spurious dynamics that are not based on the underlying data-generating process and the dynamics at the ends of the sample differ from those in the middle. Hamilton (2018) proposes calculating cycle from the residual of the following linear projection model

$$y_{t+h} = \alpha + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \dots + \beta_p y_{t-p} + v_{t+h} \quad (3.5)$$

The estimated residual \hat{v}_{t+h} is the cyclical component from this approach. Hamilton suggests using $h=8$ for quarterly data and that is what I follow in my estimation.

3.4.2 Forecasting Model

Once the credit gap measures are obtained from different trend-cycle decomposition methods, I regress growth rate of credit-to-GDP ratio on its own lags and lags of credit gap. In particular, I estimate the following model:

$$\Delta y_t = \alpha + \beta(L)\Delta y_{t-1} + \gamma(L)GAP_{t-1} + v_t \quad (3.6)$$

The lags are chosen based on the Bayesian Information Criterion (BIC). We generate the forecasts from the above model for credit gaps generated from different trend-cycle decomposition approaches. My benchmark model is an AR(1) model. I compare the forecasts generated from the above model with this benchmark model to examine if the inclusion of lagged credit gap leads to an improvement in forecasting performance of growth rate of credit-to-GDP ratio.

3.4.3 Forecast Combination

Once I have different measures of cycle, I use Bates and Bates & Granger (1969) method of forecast combination to assign weights to different cycles. The underlying idea is that the cyclical component should be able to predict changes in credit-to-GDP ratio. Bates-Granger weights are based on the following formula:

$$w_m = \frac{\hat{\sigma}_m^2}{\hat{\sigma}_1^2 + \hat{\sigma}_2^2 + \dots + \hat{\sigma}_M^2} \quad (3.7)$$

where $\hat{\sigma}_m^2$ is inverted out-of-sample forecast error variance of forecast M based on the cyclical component M. M is the number of forecasts. Weights are normalized by sum of inverted forecast error variances. Here the weight is assigned based on the predictive ability of different cycles.

3.5 Empirical Results

My analysis is performed in three steps: In the first step, I calculate the credit gap from 8 different trend cycle decomposition methods. These methods are: conventional HP filter (1997), Ravn-Uhlig's (2002) modification of the HP filter, HP filter with a high smoothing parameter (BIS gap), Beveridge-Nelson decomposition (1981), Hamilton filter (2018), Unobserved component model (Clark, 1987), linear trend, and quadratic trend. I perform these estimations recursively to preserve the 1-sided nature of the credit gap. My first estimation sample runs from 1960:Q1-1988:Q4 and saves the last estimate of the cycle. I keep adding one more observation to the estimation sample and keep saving the last observation of the cycle for different methods. This approach provides us with a 1-sided estimate of the credit gap from different methods. In the second step, I use these 1-sided credit gap measures to forecast changes in the credit-to-GDP ratio. The estimation sample for the first forecasts is 1989:Q2-1994:Q1. I then move ahead one quarter, re-estimate the forecasting model and

forecast 1995:Q2-1996:Q1, etc. My final set of forecasts, for 2019:Q3 2020:Q2, would have been prepared in 2019:Q2. I consider different quarterly horizon forecasts until $Q=4$. In addition to these quarterly forecasts, I also examine the average over the next four quarters. These averages are also used in the analysis to get around the noise associated with quarterly projections.

Tables 3.1 and 3.2 show the out-of-sample forecasting results for the above exercise for the U.K. and the U.S. The tables show the ratio of root mean squared errors in comparison to a benchmark AR(1) model. The ratio of less than unity implies that the model in question has a lower RMSEP (Root Mean Square Error of Prediction) than the benchmark AR (1) model. The results are reported for eight different methods of trend-cycle decomposition and two forecast combination methods. My preferred forecast combination method is Bates and Granger, though I also report the simple average of different forecasts. The results clearly suggest a significant improvement in forecasting performance using my combination approach. For the U.S., there is an improvement of around 13 percent in the RMSEP over the benchmark AR(1) model if I combine the forecasts based on the BG approach. Interestingly, for the U.S., the inclusion of the BIS gap in the forecasting equation (6) leads to a deterioration in forecasting performance. The conventional HP filter does a better job than Ravn-Uhlig and the BIS measure. Deterministic trend models also do not perform well. The credit gap measures obtained from the BN and the UC models perform better than the benchmark AR(1) model implying that inclusion of the lagged credit gap measure improves forecasting performance. For the U.K., the BG method yields an improvement of 9 percent in RMSEP over the benchmark AR(1) model over $h=1-4$ forecasting horizon. Even for the U.K., the BIS gap measure does not lead to an improvement in forecasting performance. The BN approach that improved the forecasting performance in the case of the U.S. performs poorly in the U.K. Overall, results for the forecasting exercise validate my argument that focusing on a single measure of trend-cycle decomposition to derive a measure of credit gap is fraught with severe limitations. The forecast combination that assigns weights based on

the predictive performance not only takes into account this model uncertainty, but in doing so improves the metric-forecasting performance- that should be the evaluation criteria for different trend-cycle decomposition methods.

Figures B.2 and B.3 plot the credit gaps obtained from my forecast combination approach along with the BIS credit gaps for the U.K. and the U.S. The sample period for the credit gap is 1994:02-2020:02. There are several interesting features that emerge from these graphs. First, I find that both the gap measures do move together. This is not surprising since the combined gap measure is a weighted measure of different cycle measures. Both the measures suggest that the credit-to-GDP ratio was above its long-run trend before the crisis for the U.K. and the U.S. These two measures became negative in the aftermath of the financial crisis and stayed negative for a significant time period. Although these credit gap measures tend to move together broadly, there are also significant differences within different subsamples. For the U.S., the combined credit gap became positive three years earlier than the BIS gap implying that the credit-to-GDP ratio was above its long-run trend starting in 1995:Q2. This was also the time period when the boom in the housing market started in the U.S. A similar pattern is observed in the U.K., where the combined gap became positive a year (in 1998:Q3) earlier than the BIS gap. Interestingly, there is no significant difference in the timeline when these gap measures turn negative.

Secondly, I find that the combined credit gap measure is significantly less volatile than BIS gap. The standard deviation of the combined gap is 6.84, whereas it is 11.57 for the BIS gap for the U.K. The corresponding numbers are 5.16 and 8.66 for the U.S. This becomes particularly evident during the post-financial crisis period where the trough of the cycle is significantly lower for the BIS gap than the combined gap for both the countries. The lower volatility of the combined gap is expected since it is derived from a weighted combination of different gap measures. This difference is even more stark for the U.K. than the U.S. Not only is there a significant differential between these two different credit gap measures after the financial crisis, these differences have persisted for a long period of time. Thirdly, I observe a

clear pattern in early detection of trough date by combined gap in both the countries. For the U.S., the trough of the cycle is 2012:Q3 according to the combined gap measure, whereas it occurred in 2013:Q1 according to the BIS measure. For the U.K., the combined gap suggests that the gap reached its bottom in 2015:Q1, whereas the corresponding date was 2015:Q2 for the BIS gap. For the peak, I do not find significant differences in the detection of turning points.

The comparison of credit gap measures across these two countries provides interesting insights. The peak in the credit-to-GDP ratio was reached in the U.K. much earlier than the beginning of the financial crisis. The peak for the U.K. as measured by these credit gap measures took place at the end of 2002 and the gap did not display any trend until the end of 2008. In the U.S., however, the peak occurred just before the financial crisis. Although both credit gap measures were positive before the crisis, the difference between them grew. Both the measures fell at a rapid pace for the U.S. during the crisis. In the U.K., however, the BIS gap measure fell at a much faster rate and the difference between these two measures started growing much earlier. The results from my forecasting exercise suggest that the rapid pace of fall in the credit gap for the BIS measure may have been overestimated. Finally, one of the criticisms of the HP filter is that it suffers from the endpoint problem. This can be observed in my figures where both BIS gap measures for the U.S. and the U.K. show a sudden shift at the end of the sample and display a ragged edge problem. My combined gap does not suffer from this problem.

3.6 Conclusions

This paper proposes an alternate measure of credit gap-deviation of credit-to-GDP ratio from long-run trend. My credit gap estimation approach is based on the premise that the deviation of a non-stationary time series from its trend should have predictive power for subsequent movements in the changes in the variables as noted by Nelson (2008). In addition,

I also take into account the model uncertainty by acknowledging that there is not a unique way to decompose a series into a trend and a cycle. For this purpose, I assign weights to different credit gaps derived from different trend-cycle decomposition methods. These weights are based on the out-of-sample forecast error variance as in Bates and Granger (1969). The model results show that this method of combining credit gaps yield a credit gap measure that dominates credit gaps from different trend-cycle decomposition methods including the one published by the BIS in terms of superior out-of-sample forecasting of changes in credit-to-GDP ratio. I apply this framework to the data from the U.S. for the recent time period. I also show that my proposed credit gap measure captures the trough in the credit cycle earlier than the BIS measure and also displays lower volatility.

Table 3.1: Forecasting Performance of Credit Gap Models (U.K.)

Horizon	HP	RU	BIS	Hamilton	Linear	Quadratic	BN	UC	Average	Bates-Granger
1	1.001	0.990	1.001	0.992	1.010	0.979	1.028	1.009	0.977	0.979
2	0.979	0.970	1.007	0.969	1.016	0.962	1.028	0.999	0.962	0.957
3	0.979	0.971	1.018	0.969	1.055	0.966	1.009	0.989	0.959	0.955
4	0.990	0.987	1.028	1.005	1.055	0.981	1.019	0.981	0.972	0.967
1 - 4	0.972	0.952	1.034	0.960	1.081	0.929	1.054	0.985	0.918	0.910

Notes:

The table shows the ratio of RMSEPs of different models in comparison to the benchmark AR(1) model. The first set of forecasts is for 1994:Q1-1994:Q4; the final set is for 2019:Q3-2020:Q2. Q=1-4 denotes averages over next 4-quarters. HP is Hodrick-Prescott, RU is Ravn-Uhlig, BIS is based on Borio and Lowe (2002), BN is Beveridge-Nelson, UC is Unobserved Component Model. The smallest ratio in each row is bolded.

Table 3.2: Forecasting Performance of Credit Gap Models (U.S.)

Horizon	HP	RU	BIS	Hamilton	Linear	Quadratic	BN	UC	Average	Bates-Granger
1	0.993	0.987	1.012	0.994	1.028	1.005	1.010	0.985	0.962	0.959
2	0.974	0.963	1.016	0.980	1.058	1.014	0.975	0.961	0.924	0.917
3	0.966	0.953	1.023	1.011	1.055	1.036	0.965	0.937	0.906	0.896
4	0.982	0.966	1.022	1.029	1.055	1.045	1.033	0.910	0.922	0.910
1 - 4	0.964	0.945	1.030	1.005	1.081	1.041	0.978	0.913	0.882	0.872

Notes:

The table shows the ratio of RMSEPs of different models in comparison to the benchmark AR(1) model. The first set of forecasts is for 1994:Q1-1994:Q4; the final set is for 2019:Q3-2020:Q2. Q=1-4 denotes averages over next 4-quarters. HP is Hodrick-Prescott, RU is Ravn-Uhlig, BIS is based on Borio and Lowe (2002), BN is Beveridge-Nelson, UC is Unobserved Component Model. The smallest ratio in each row is bolded.

Chapter 4

Identifying Unsustainable Credit Gap

4.1 Introduction

Credit-to-GDP gaps have been identified as the top candidate as an early warning indicator (EWI) for future financial crises. As the importance of the financial market to the real economy has increased, failing to predict a future systemic crisis and missing the correct timing to implement macroprudential instruments such as Countercyclical Capital Buffer (CCyB) can lead to significantly greater economic losses. As a result, there is a need to improve the performance of our early warning indicator models, specifically when using the credit gap in univariate EWI models.

In the EWI literature, The Bank of International Settlement (BIS) Basel credit gap using a quasi-real-time (one-sided) version of the Hodrick-Prescott filter (HP) has been established as a reference point for policymakers. The BIS Basel gap is easy to implement and communicate and has performed well as an early indicator. However, recent literature has found evidence that optimizing parameters of various decomposition filters can improve those filters' performance beyond the predictive power of the BIS Basel gap as an EWI. The details are discussed in Drehmann & Yetman (2021) and Beltran et al. (2021).

The development of embedding optimization of decomposition filter parameters into the process of building an EWI framework has significantly improved the performance of the credit gap filters. However, this also created another issue which is model uncertainty. Policymakers now face at least five decomposition methods, each with a different set of feature parameters to consider optimizing for, such as smoothing parameters, robustness, and rolling window. Furthermore, changes in the sample window would change the performance of those filters with optimized parameters. Selecting a robust and best-performing decomposition fil-

ter and optimizing its feature parameters can be a paradox of choice for policymakers.

There is a need for a robust methodology to create a univariate EWI that adapts well to changes in dynamics of the underlying financial market but at the same time uses only the credit gap as an explanatory variable. This paper attempts to solve the model uncertainty issue by using Bayesian model selection and average methods. I first create a set of possible decomposition filters with rich features from the recent literature findings. Out of all possible initial credit gaps created, I then implement a series of model selection methods using an additional novel metric psAUC (partial standardized area under the curve of receiver operating characteristic), discussed in subsection 4.4.4. Subsequently, from a finalized subset of variables that passed the performance filters, I perform model averaging of all possible combinations of the variables. Raftery (1995) proposed the Bayesian model average method that averages its estimated parameters of interest across all possible combinations of variables. The model that uses the average estimate has proven to be more robust and performs better in out-of-sample forecasts than any other individual model in the selected subset of model combinations.

Using data from 1970:Q4 to 2017:Q4 across 43 countries, I created 90 credit gaps with varying features found in the literature. I implemented the model selection and average methods described in the paragraph above. I then found evidence that the new EWI performance metric psAUC, which constraints analysis to the region where Type II error $< 1/3$, is a robust single value criterion. I also successfully incorporated the new metric into my model selection and averaging method. Lastly, to ease policy implications, I propose to create a single credit gap measurement from weighted-averaging other popularly studied credit gap measurements. The single weighted average credit gap is tested as more robust in resampling bootstrapping estimates. It also performs better in out-of-sample forecasts k-fold cross-validation than any other individual model in the filtered subset of model combinations.

The paper is structured as follows. In the next section, I discuss relevant branches of literature. In Section 3, I describe the data used in this paper. Section 4 outlines the method-

ology I use for model selection, model average, and a weighted combination of variables of interest, as well as a novel metric for testing the performance of the selected variable and averaged model. Section 5 describes the empirical results of the selected variables in the horse race. Section 6 analyzes time series graphs of individual countries. Section 7 summarizes the main conclusions.

4.2 Literature Review

In this section, I discuss the branches of literature that my study is relevant and contributes to. C. Borio & Lowe (2002) first documented credit gaps’ property as a useful early warning indicator (EWI) for banking crises. Earlier literature on excessive credit leading to crises includes Minsky (1977), Manias (1978), and Kaminsky & Reinhart (1999). C. E. V. Borio & Drehmann (2009) assessed different smoothing parameters for the one-sided HP filter and later determined $\lambda = 400,000$ to be the ideal value for measuring credit gap based on its long smooth trend property and performance as an early warning indicator. Banking Supervision (BCBS) (2010) later established this particular HP filter as the Basel credit gap (I call it BIS Basel gap in this paper), and it became the reference for all EWI models using credit gap filters. Drehmann & Tsatsaronis (2014) defended the Basel gap’s performance as an EWI against criticism it received, such as its failure to perform well in emerging market economies.

Drehmann & Juselius (2014) set out a list of criteria for an EWI. It needs to have the “right timing” of being able to detect future financial crises from 20-6 quarters ahead of a future systemic crisis; a “stable signal” that does not decrease as the periods approach the actual crises date; “robustness” that fits well with different data set; and “interpretability” which is ideal for a univariate model such as the credit gap to determine an optimized threshold of classifying excessive credit. Lastly, EWI S_i outperforms EWI S_j for horizon h if $AUC(S_{i,h}) > AUC(S_{j,h})$. This last criterion, however, receives criticism regarding the

lower left region of the ROC curve that represents low predictive power (True Positive Rate) of a model, which the AUC metric also measures.¹ Additionally, C. Borio (2014) reported stylized features of the credit cycles. These cycles have a much lower frequency than business cycles. It also pointed out that credit cycle peaks are closely associated with the onset of financial crises. Because of that property, they can detect crises in real-time within the macroprudential policies implementation window.

Hamilton (2018) argues against using HP filter for macroeconomics series as it introduces spurious dynamics. Drehmann & Yetman (2018) defended the HP filter again but will later admit other filters with additional features, such as a Hamilton filter estimated in a panel setting, outperformed the Basel gap in Drehmann & Yetman (2021). Furthermore, features for the decomposition methods include rolling sample windows of 15 and 20 years as proposed by Galán (2019) to help improve the performance of the decomposition filter. Beltran et al. (2021) embedded the optimization of features parameters into the construction of an EWI model and constraint the search range for the loss function to the region where True Positive rates $\geq 2/3$ or Type II error rates $< 2/3$ following suggestions in C. E. V. Borio & Drehmann (2009), the paper also introduced using Structural Time Series model (STM) gap, Moving average (MA) gap as possible candidates for the EWI model.

Galán (2019) also proposed adding other features to the Basel gap that individual countries implement which are out of the scope of this paper's implementation. Such adjustments include adjusting for GDP decrease, combining one and two-sided filters for robustness, adjusting by projection as in Gerdrup, Kvinlog, & Schaanning (2013), and adjusting for currency fluctuations.

However, there is no consensus on which decomposition with which features would be the best performing and most robust EWI. In this paper, I aim to construct a methodology to find the best combination of decomposition filters and features to be used as an EWI and propose an average model that performs as well as the best combination, if not better. The

¹the details are discussed in subsection 4.4.4

methodology I use for the model selection, and average steps are discussed in the literature branch below.

Babecký et al. (2014) used a model selection method called Markov Chain Monte Carlo Model Comparison (MC^3) developed by Madigan, York, & Allard (1995). This method allows for a feasible search for best-fitted combinations out of a large set of variables with similar features, which would not be efficiently estimated using other model selection methods such as RIDGE, LASSO, and Elastic Net. Moreover, Furnival & Wilson (2000) proposed a “leap and bound” algorithm to find the best-fitted subsets without examining all possible subsets of variables.

The Bayesian model averaging method is formally described in Raftery (1995). It aims to overcome model uncertainty by averaging the performance of a set of models. The averaged model performs better than any selected individual models on out-of-sample forecasts. The model average method could be applied to multivariate settings, but that will cost us the ability to synthesize a combined weighted credit gap. Following the motivation from Chapter 3, I also attempt to create a combined weighted credit gap in this chapter that performs well in a horse race with other credit gap measurements.

Multivariate early warning system (EWS) is also a significant branch of early warning literature and is worth mentioning. The variables selected in Babecký et al. (2014) that perform well in EWS are then used in the following papers. Drehmann & Juselius (2014) established that with the credit gap, the Debt-to-Service ratio is another EWI that performs well within 1-4 quarters of the future systemic financial crisis. Aldasoro et al. (2018) is an extension of Drehmann & Juselius (2014). Furthermore, Alessi & Detken (2018) used a Random Forest model to achieve a higher EWI performance metric than traditional generalized linear regression. Recently, Holopainen & Sarlin (2017) implemented ensemble methods to combine various machine learning model predictions and found significant improvement in the ensembled multivariate early warning system performance. Even though multivariate models can achieve superior model fit and prediction accuracy than univariate models, they

also require richer data availability and limit the scope of the model implementation. Certain emerging market economies with insufficient data would benefit more from a well-performing univariate model.

Lastly, I touch on the literature on partially estimating the Area Under the Curve of Receiver Operating Characteristic (psAUC). Detken et al. (2014) first introduced the use of partial standardized AUC in the EWI setting. The partial standardized AUC is popularly studied in the medical statistics literature, e.g., McClish (1989), but is relatively new in the EWI literature. I will use this novel (psAUC) metric, inspired by both Detken et al. (2014) and the policy function restriction in Beltran et al. (2021), to improve my model selection and averaging process in terms of satisfying policy implication requirements (e.g. $\text{TPR} \geq 2/3$). The method for partial estimation of the ROC curve is documented in Robin et al. (2011).

4.3 Data Description

The model sample periods include quarterly data from 1970:Q4 to 2017:Q4 across 43 countries.² The sample periods were chosen based on the availability of systemic crisis data and total credit to GDP ratio data. The primary source of the credit data comes from the Bank of International Settlement (BIS). The credit variable is measured as total credit to the private non-financial sector as a percentage of GDP.³

Regarding data for dating systemic crises, there is no single consensus source of database. I follow the literature and use the crisis dates reported in the European Systemic Risk Board crisis data set (Lo Duca et al., 2017) for countries in the EU and (Laeven & Valencia, 2013) for the rest of the world. The systemic crisis dates are available from after WWII. However,

²The list of countries with data available are: Argentina, Austria, Australia, Belgium, Brazil, Canada, Switzerland, Chile, China, Colombia, Czech Republic, Germany, Denmark, Spain, Finland, France, United Kingdom, Greece, Hong Kong SAR, Hungary, Indonesia, Ireland, Israel, India, Italy, Japan, Korea, Luxembourg, Mexico, Malaysia, Netherlands, Norway, New Zealand, Poland, Portugal, Russia, Saudi Arabia, Sweden, Singapore, Thailand, Turkey, United States, and South Africa.

³Source of the data could be accessed here: https://www.bis.org/statistics/totcredit/credpriv_doc.pdf

the first systemic crisis did not happen until the early 1970s.

There are three types of crises: banking, currency, and debt crises, all of which could lead to a systemic crisis. A crisis can be labeled as systemic when certain conditions are met, such as significant signs of financial distress in the banking system, bank runs, and significant policy intervention measurements are imposed. The authors also consulted with subject matter experts in each country to determine the dates and details of the systemic crises.

While the credit data is available until 2021:Q3, my systemic crisis data stopped at 2017:Q4. Therefore, I will limit my analysis till 2017:Q4. I will also omit periods of countries with shorter measurements of credit.

4.4 Empirical Methodology

4.4.1 Overview

The total credit series is nonstationary. In order to extract a useful stationary credit signal to use as an early warning indicator for future financial crises, I first decompose the series into a nonstationary trend and a stationary cycle using various popularly studied filters in subsection 4.4.2.

The idea behind using the credit gap as an early warning indicator is that a prolonged period of excessive credit growth could be linked to future financial crises. Therefore, I can use logistic regression to identify pre-crisis periods with unsustainable high credit gap values over a certain threshold.

In subsection 4.4.3, I discuss strategies for identifying the dependent variable (pre-crisis periods) and the logistic regression empirical method. In subsection 4.4.4 I then discuss the metrics to determine the performance of a credit gap as an early warning indicator EWI model such as BIC, AIC values, Area Under Curve (AUC) of receiver operating characteristic, how I determine a policy loss function, optimize credit gap threshold and estimate the

minimized Type I and Type II error rates. I will then propose a novel metric - partial standardized AUC (psAUC) that, when used in conjunction with other metrics, will provide more desirable properties for policy implication and overcome some of the criticism that the traditional AUC metric receives.

This paper aims to overcome model uncertainty when using credit gap filters as an early warning indicator for future financial crises. Since each decomposition filter of the total credit series provides a stationary series that performs qualitatively differently from other filters as an EWI, I propose selecting the best gap candidates and performing model averaging to achieve model stability and improved out-of-sample prediction as proposed in (Raftery, 1995).

In subsection 4.4.5, I will use the metrics discussed in subsection 4.4.3 to select credit gap filters that perform well individually and in combination with other credit gaps. Then in subsection 4.4.6, I discuss the theory behind the Bayesian model average method, my model customization to fit the policy implication need, and the new metric (psAUC).

Lastly, in subsection 4.4.7, for ease of policy implication, I propose the creation of a crisis-weighted credit gap that would inherit the information and features from the averaged model of 30 other credit gaps, and would perform as well as the averaged model prediction does.

4.4.2 Credit Gap Decompositions

$$100 * \frac{Credit}{GDP} = y_it = \tau_{itj} + c_{itj} \quad (4.1)$$

I started my variable selection process by creating 90 candidate one-sided credit gap measurements based on the literature.⁴ The nonstationary Credit to GDP ratio series is

⁴List of filters created is can be viewed in Appendix C.1

decomposed into a nonstationary trend τ_{yit} and a stationary cycle credit gap c_{itj} components. Once a country has more than 15 years of credit measurement available, I start storing its one-sided credit gap values onward. My decomposition filter methods include one-sided Hodrick-Prescott, Hamilton (panel and non-panel setting), Moving Average, Structural Time-series model, Beveridge-Nelson, linear, quadratic, and polynomial decompositions. All are decomposed with full sample, rolling 15 years and 20 years rolling window features when possible.

Apart from the filters that I already introduced in Chapter 3 - “Measuring Credit Gap” (Equations 3.1 - 3.5), the two additional filters I will use in this chapter are Moving Average, Structural Time-Series model filters that Beltran et al. (2021) introduced.⁵

4.4.3 Early Warning Indicator - GLM Logistic Regression

I begin testing the fitness of the EWI model using credit gap measurements by implementing binary logistic regression. I identify the dependent variable as the periods of 5-12 quarters ahead of a systemic financial crisis. The underlying classification model assumes that the credit to GDP ratio gap will increase sharply as a response to an excessive credit increase due to expansionary monetary policy or a sharp decline in GDP, which is the ratio’s denominator. When the credit gap is above a certain threshold for a prolonged period, it signals that the economy is at risk of a future financial crisis. The periods with excessive credit before financial crises will be labeled “pre-crisis” periods.

It is also worth pointing out that there is no consensus on setting the periods to label as pre-crisis. The literature labeled pre-crisis periods with various ranges from 1-16 quarters before a crisis. I decided to use a labeling strategy with ideal policy implication properties. It is recommended that macroprudential policy takes more than four quarters to be effective because of policy implementation lag. As a result, a signal telling that the economy will experience a financial crisis in 1-4 quarters has little policy implication. Therefore, I start

⁵Refer to their equation representations here in Appendix C.1.1

labeling the pre-crisis periods in the 5th quarter before a systemic crisis. I stop labeling in the 12th quarter before a crisis because implementing counter-cyclical policy too early will be ineffective in lowering the risk of a future crisis and would only exacerbate economic conditions by unnecessary credit tightening.

Equation (4.2) represents the logistic regression:

$$pre.crisis_{it} \sim c_{itj} \quad (4.2)$$

In which i is the country indicator, and j is the credit gap filter type. The binary dependent variable $pre.crisis_{it}$ takes values of 1 when the economy is between 5-12 quarters before a systemic financial crisis, labeling the periods as “pre-crisis” positive. I discard measurements between 1-4 quarters before a crisis, during, and post-crisis management periods identified in Lo Duca et al. (2017) and Laeven and Valencia (2018). The credit measurements in these periods are discarded to avoid biased estimation as credit gaps behave erratically during a crisis.

The indicator is set to 0 at other periods. Additionally, pre-crisis periods of imported crises identified in the dataset are also set to 0 since the model aims to capture pre-crisis periods created by a country’s excess credit growth, not from exogenous shocks. However, for labeling conformity, I still discard measurements of periods during and post-crisis of these imported crises, as countries experiencing imported crises could also implement reactionary monetary policies.

4.4.4 EWI Performance Metrics

Likelihood Values and BIC

The logistic regression model parameters are estimated by maximizing the cross-entropy log-likelihood values. The relationship between the BIC (Bayesian Information Criterion)

and likelihood value is expressed in Equation (4.7). The model selection criterion here is to choose the filter method with a lower BIC value associated with a higher likelihood value and fitness of the general linearized model or a higher degree of freedom by using fewer explanatory variables.

Area Under the Curve (AUC) of Receiver Operating Characteristic (ROC)

A receiver operating characteristic (ROC) curve in the EWI literature setting represents True Positive Rate (TPR) and False Positive Rate (FPR) Cartesian coordinates of different credit gap thresholds, which are used as indicators for identifying pre-crisis periods.⁶ The credit gap thresholds are determined by the logistic regression model predicted response values which range from (0,1) and map increasingly with the credit gap magnitude values.

A logistic model is strictly preferred over another comparable model if all of the coordinates of its ROC curve are closer to the top-left region of the graph, which minimizes FPR and maximizes TPR. However, it is not always the case when I have such a clear-cut difference. Models can have overlapping ROC curves. To simplify the process of model selection in such cases, the area under the curve (AUC) of ROC metric is used.

$$AUC = \int_0^1 TPR d(FPR)$$

Each logistic regression with a different gap measurement yields a value of Area Under the Curve (AUC) of receiver operating characteristic. There is an underlying assumption that the higher the AUC value is, the better the overall performance of a credit gap is as an EWI; on average, its ROC curve has coordinates with higher TPR and lower FPR.

True Positive Rate represents a model's predictive power (P) to identify pre-crisis periods correctly (positive case). $TPR = TP/(TP + FN)$. True Negative Rate represents a model's predictive power to identify normal periods correctly (negative case). $TNR = TN/(TN +$

⁶Refer to Figure 4.1 for illustration

FP). False Positive Rate represents the Type I error rate that a model misidentifies a calm, normal period (negative case) as a pre-crisis period (positive case). $FPR = 1 - TNR = FP/(FP + TN)$. Lastly, I discuss the False Negative Rate or the Type II error rate at which a model misidentifies a pre-crisis period (positive case) as a normal calm period (negative case). $FNR = 1 - TPR = FN/(TP + FN)$.⁷

Central planners working with classification problems have dual objectives to minimize Type I and Type II errors. To simplify this problem, a policy loss function with the input of the two error types are used:

$$L_{\theta} = \alpha TypeI(\theta) + (1 - \alpha)TypeII(\theta) \quad (4.3)$$

With $\alpha = 0.5$ by default if policymakers are indifferent about the two types of errors. Here, the model selection criteria would be to select minimized policy loss function with the smallest combination of Type I and Type II error rates.

A binary logistic regression model with a substantially higher AUC value will have smaller sets of Type I and Type II metrics values. Therefore, its optimized credit gap threshold value that minimizes the policy loss function would represent superior Type I and Type II error rates.

However, because the AUC value is an aggregate value representing information of the whole ROC curve, the area on its lower left corner, where the predictive power of the threshold (TPR) is low, and type II error rates are high, is not relevant for policy implications discussion. Before the Great Financial Crisis of 2008, central planners tended to weigh Type I errors more than Type II errors since the cost of misspecifying a financial crisis was not apparent, and a Type I error meant costing the economy unnecessary credit constraints by implementing reactionary policies. But since 2009, there has been a shift in weighing Type II errors more as the cost of misspecifying a systemic financial crisis became significantly heavier than a

⁷Our notation of Type I and Type II error follows Beltran (2021) which deviated from previous literature.

Type I error.

Partial Standardised AUC

To overcome the issue of unnecessary information included in the full AUC, an approach to estimate only a partial relevant area under the ROC was proposed in McClish (1989) and later implemented in the EWI setting by Detken et al. (2014). I will also discuss the standardization of the partial AUC metric (psAUC) in Equation (4.5).

Detken et al. (2014) gave their remarks on the usage of partial standardized AUC (AUROC and AUC terms are used interchangeably in the literature):

“Instead of considering only the full AUROC (e.g. Drehmann and Juselius, 2014), this paper also presents a partial standardized AUROC (psAUROC) that cuts off the area associated with a preference parameter of $\theta < 0.5$ ”

“While the psAUROC has been used extensively in the area of medical statistics to assess the performance of a classifier only in specific regions of the ROC curve (e.g., McClish, 1989 and Jiang et al., 1996), it is a new approach in the literature evaluating EWMs.”

“The results reported in this paper show that the psAUROC can reveal useful additional information as long as the partial area does not become too restricted.”

Detken et al. (2014) proposed cutting off the area representing the low values of preference response parameter in a multivariate regression setting. However, this application does not directly translate to a univariate credit gap model. Using a univariate credit gap EWI model, (Beltran et al., 2021) constrained the policy loss function to regions where $\text{TPR} \geq 2/3$ or Type II error rate $< 1/3$. They then estimated the policy loss function value at different points on the partial ROC curve by assigning different policy preference parameter values α .

To merge these two ideas above and to solve the uncertainty issue of having to estimate policy loss function at different policy preference parameter values to determine model performance, I propose to restrict the consideration of the ROC curve to $TPR \geq 2/3$, then estimate the partial standardize psAUC of the restricted ROC curve region instead. Note in Equation (4.4), I estimated the partial area under the curve that represents TPR and TNR (instead of TPR and FNR) and took integral along the TPR axis.

$$pAUC = \int_{\frac{2}{3}}^1 TNR d(TPR) = \int_{\frac{2}{3}}^1 specificity d(sensitivity) \quad (4.4)$$

Because I am limiting my analysis only to the portion of the ROC curve that satisfies $TPR \geq 2/3$, the policy loss function (4.3) used in previous literature tends to have a corner solution of optimized thresholds where $TPR = 2/3$. I propose using an alternative policy loss function that priorities points on the ROC closest to the top left regions where both Type I and II errors are the lowest:

$$L_{\theta} = TypeI(\theta)^2 + TypeII(\theta)^2 = (1 - sensitivity)^2 + (1 - specificity)^2$$

The last step is to standardize the partial AUC value:

$$psAUC = \frac{1}{2} \left[1 + \frac{pAUC - min}{max - min} \right] \quad (4.5)$$

The standardization step helps with the comparison of EWI models' performance. Traditional AUC values range from 0 to 1, with 0.5 being the value for a null hypothesis model or an information-less guess. When only estimated partially, the partial AUC value will not retain the same range making it harder to interpret their meaning. Standardizing the partial

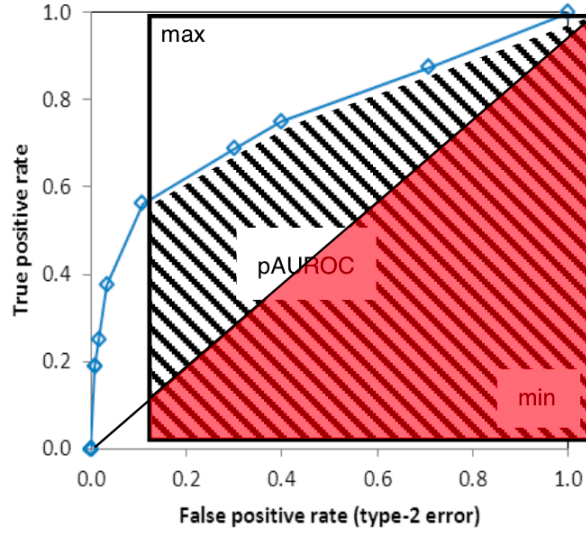


Figure 4.1: Procedure to standardize partial AUC

AUC values restores the ranges of possible values to 0 and 1.

In section 4.5, I will discuss the relevance and performance of the psAUC as an EWI metric and the other metrics introduced in this subsection.

4.4.5 Variable Selection

Our overall methodology goal is to implement model averaging of top candidate variables. In order to achieve that, I need to select top EWI candidates using the EWI performance criteria discussed in the previous section. I also search for models with combinations of variables that, when combined using positive weights, would have a better model fit than a univariate model would. I aim to select 29 credit gap measurements based on these two criteria.

Firstly, I compared performances of individual credit gaps using partial area under the curve (psAUC) values. Table 4.1 ranked decomposition filters by psAUC.

Secondly, I test for performances of multivariate models with combinations of different credit gaps.

Table 4.1: Top 30 credit gap measurements ranked by psAUC

Cycles	BIC	AIC	AUC	psAUC	c.Threshold	Type.I	Type.II	Policy.Loss.Function
null	0.0000	0.0000	0.5000	0.5000		1.0000	0.0000	1.0000
c.bn6.r20	-108.0679	-114.4506	0.7048	0.6379	0.6581	0.3962	0.3019	0.2481
c.hamilton28.panel	-149.8518	-156.2346	0.7107	0.6359	9.7674	0.3912	0.3066	0.2470
c.hamilton13.panellr20	-150.2442	-156.6269	0.7036	0.6333	5.9895	0.4261	0.2547	0.2464
c.hamilton24.panel	-134.4093	-140.7920	0.6991	0.6322	7.1794	0.4383	0.2689	0.2644
c.hamilton20.panellr20	-151.5617	-157.9445	0.7048	0.6313	7.9350	0.4321	0.3066	0.2807
c.ma	-120.8108	-127.1936	0.6922	0.6313	5.7813	0.3989	0.3160	0.2590
c.hamilton20.panellr15	-135.3713	-141.7540	0.6985	0.6312	7.5244	0.4616	0.2689	0.2854
c.hamilton13.panellr15	-126.2968	-132.6796	0.6924	0.6311	6.5289	0.4297	0.2830	0.2647
c.hamilton28.panellr20	-164.6015	-170.9842	0.7158	0.6302	10.8558	0.3948	0.2925	0.2414
c.hamilton24.panellr20	-155.8638	-162.2466	0.7096	0.6301	9.1672	0.4251	0.2830	0.2608
c.hamilton24.panellr15	-143.2235	-149.6062	0.7033	0.6299	10.4963	0.3984	0.3160	0.2586
c.hamilton20.panel	-126.8625	-133.2452	0.6907	0.6288	5.6212	0.4686	0.2830	0.2997
c.hamilton28.panellr15	-154.4533	-160.8361	0.7091	0.6270	11.5510	0.3854	0.2972	0.2369
c.hamilton13.panel	-133.9347	-140.3175	0.6922	0.6250	4.9769	0.4285	0.2877	0.2664
c.bn2.r20	-109.3128	-115.6955	0.6963	0.6218	0.2776	0.4080	0.3255	0.2724
c.linear	-135.4069	-141.7896	0.6879	0.6204	3.9989	0.4616	0.2925	0.2986
c.bn2	-135.9914	-142.3741	0.6842	0.6165	0.1864	0.4530	0.3113	0.3021
c.bn6	-132.7915	-139.1742	0.6835	0.6113	0.4710	0.4371	0.2830	0.2712
c.bn6.r15	-54.9953	-61.3781	0.6756	0.6070	0.5680	0.4179	0.3255	0.2806
c.bn2.r15	-83.9469	-90.3297	0.6749	0.6047	0.1349	0.4761	0.3302	0.3357
c.poly4.r20	3.5738	-2.8090	0.5772	0.6011	0.1651	0.4980	0.3302	0.3570
BIS Basel gap	-121.5910	-127.9738	0.6733	0.5960	3.0578	0.4441	0.3255	0.3032
c.bn4	-169.1186	-175.5014	0.6892	0.5943	1.2840	0.3837	0.3255	0.2532
c.bn4.r15	-89.6147	-95.9975	0.6669	0.5929	0.4435	0.4792	0.2925	0.3152
c.bn5.r20	-99.7674	-106.1501	0.6744	0.5928	0.5016	0.4234	0.3302	0.2883
c.stm.r15	-79.5531	-85.9358	0.6575	0.5924	2.0027	0.4778	0.3160	0.3281
c.hp125k	-92.2897	-98.6725	0.6562	0.5924	2.5216	0.4547	0.3302	0.3158
c.hp221k	-106.8842	-113.2670	0.6656	0.5921	2.6641	0.4561	0.3160	0.3079
c.linear.r15	-66.3472	-72.7300	0.6474	0.5916	2.6344	0.4626	0.3302	0.3230
c.hp400k.r15	-67.1228	-73.5055	0.6472	0.5912	2.6223	0.4592	0.3255	0.3168

Table 4.2: Top 25 credit gap measurements ranked by MC3 probability

Variable	Pr(B!=0)	Variable	Pr(B!=0)
Intercept	1	c.bn2	0.3721
c.hamilton28_panel	0.999968	c.bn3_r10	0.2478
c.bn7_r15	0.93097125	c.hp400k	0.2309
c.poly3	0.915792	c.hamilton24_r20	0.2175
c.hamilton13_panel	0.733247	c.hamilton13_r20	0.1958
c.hamilton13_r10	0.70601225	c.bn2_r10	0.1916
c.hamilton28	0.65141375	c.stm	0.1894
c.hamilton13	0.6438865	c.hamilton28_r20	0.1799
c.bn3_r15	0.514728	c.hp_r20	0.1794
c.hp3k	0.4888055	c.hp	0.1761
c.hp3k_r20	0.4887615	c.bn8_r20	0.1561
c.poly6	0.39826475	c.linear_r20	0.1526
		c.hp221k	0.1440

Note:

Pr(B!=0) is the posterior probability of the variable being selected

$$Model_k : pre.crisis_{ti} \sim \sum_j \beta_j * credit.gap_{tij}$$

I implement the gaps combination model selection using Markov Chain Monte Carlo Model Comparison (MC^3) method developed by Madigan et al. (1995). The method assigns a posterior probability for different credit gaps being selected in the most likely models/combinations with the lowest BIC values. With 90 variables, there are $2^{90} = 10^{26}$ possible subsets of combinations to choose from. Babecký et al. (2014) used this MC^3 method to identify potential variables in multivariate EWS models. For more efficient search implementation, each variable is given increasing prior weights depending on their univariate EWI performance. I then estimated the posterior probability of each variable included in the most likely models using 4,000,000 MCMC iterations. The results are reported in Table 4.2.

Using the two results table above, I selected the candidate gaps not only by their psAUC

values but also by their other metrics and types of decomposition methods and features to ensure the robustness of the model averaging method. The 29 selected decomposition filters are reported in Figure 4.2. 29 is also the number of recommended variables to start the Model Averaging method in the following subsection.

4.4.6 Model Averaging

The Bayesian Model Average method is formalized in Raftery (1995) to account for model uncertainty and achieve better prediction results in out-of-sample forecasts. Equation (4.6) represents a Model k 's posterior probability or weight in the averaging process:

$$P(M_k|D) = \frac{P(D|M_k)P(M_k)}{\sum_{l=1}^K P(D|M_l)P(M_l)} \approx \frac{\exp(-\frac{1}{2}BIC_k)}{\sum_{l=1}^K \exp(-\frac{1}{2}BIC_l)} \quad (4.6)$$

Where $P(M_k)$ is model prior probability and can be ignored if all models are assumed to have equal prior weights. $P(D|M_k)$ is the marginal likelihood. And $P(D|M_k) \propto \exp(-\frac{1}{2}BIC_k)$.

In which:

$$BIC_k = 2\log(\text{Bayesfactor}_{sk}) = \chi_{sk}^2 - df_k \log(n) \quad (4.7)$$

The subscript s indicates the saturated model. χ_{sk}^2 is the deviance of model K from the saturated model. $\chi_{sk}^2 = 2(\mathcal{U}(Ms) - \mathcal{U}(Mk))$. And $\mathcal{U}(Mk)$ is the log-likelihood of model Mk given data D . A model with better fitness will have a higher log-likelihood value, hence smaller deviance and smaller BIC value. df_k is the degree of freedom of model K , and n is the number of observations.

Alternative BIC Measurement

I propose using psAUC in addition to the log-likelihood in the measurement of deviance. Hence, an alternative BIC value can be estimated at:

$$BIC_{alt,k} = 2\log(Bayesfactor_{alt,sk}) \quad (4.8)$$

$$= 2(1000 * (psAUC_s - psAUC_k)) - df_k \log(n) \quad (4.9)$$

I scaled the psAUC value by 1000 since $0 < psAUC < 1$. Also, by design, the saturated model has $psAUC_s = 1$.

Final Real-time Search for Best Models Combinations

All of the following search steps described below are done in quasi-real time. As additional quarterly data are added, I will perform the search steps again to find the best possible combinations of models that fit the updated data.

From the set of 29 candidate filter gaps selected in the previous subsection, I add an intercept parameter to make a list of 30 possible variables from which to choose a subset of combinations. The number of possible subsets of combinations is still significantly high at $2^{30} = 10^9$ possible combinations subset, which will cost us computation time and resources. I still need to limit the number of possible subsets further to consider. Furnival & Wilson (2000) proposed an algorithm for computing the residual sums of squares for all possible regressions with minimum computing power requirement. With a “leap and bound” technique, the paper found it possible to find the best subsets without examining all possible subsets. This feature reduced the number of operations required to find the best subsets by several orders of magnitude.

From the list of reduced subsets using “leap and bound” algorithm, I further filter the number of possible gaps combinations using a method called Occam’s razor window described in Raftery (1995). The Bayesian Model Average (BMA) R package, which incorporated “leap and bound” algorithm, also has an inbuilt function to filter out less likely models using Occam’s razor window method. The package also allowed for a multiple-pass filter through

a custom function.

In this variable selection step using Occam's Razor, I will first select the models with BIC values within an Occam's Razor window of the best-fitted model with the lowest BIC value. Secondly, I repeat the same step above but regarding the alternative BIC value derived from psAUC as in Equation (4.8).

Lastly, I filter out models with negative weight combinations for model averaging stability. In order to reach feasible results in these three-pass filter steps, I relax the Occam's razor (OC) value from the default value of 20 to 2000, corresponding to a change in the actual search window $(2 * \ln(OC))$ that searches for models with 6 times less likelihood value to 15 times.

Posterior Distribution of Coefficients of Interest

With all possible candidate models of credit gap combinations selected, I can start implementing the model averaging steps. First, I estimate each model k separately, then store their maximum likelihood estimated parameters $\beta_j(k)$ and their standard deviations. I use the notation $\beta_j(k)$ as the coefficient of credit gap j (c_j) in a logistic regression model k against pre-crisis indicator.

I am interested in the averaged estimate of the parameter $\beta_j(K)$ across all models K s, this is the estimate of the parameter of interest β_j in an averaged model. When considering a particular β_1 , the parameter has this probability distribution in a Bayesian averaged model setting:

$$p(\beta_1|D, \beta_1 \neq 0) = \sum_{A_1} p(\beta_1|D, M_k) p'(M_k|D) \quad (4.10)$$

Where $p'(M_k|D) = p(M_k|D)/pr[\beta_1 \neq 0|D]$. And $pr[\beta_1 \neq 0|D] = \sum_{A_1} p(M_k|D)$.

$p(M_k|D)$ is the posterior probability discussed in Equation (4.6). While $pr[\beta_1 \neq 0|D]$ is

the probability that β_1 is in the averaged model; and $A_1 = \{M_k : k = 1, \dots, K; \beta_1 \neq 0\}$ is the set of models that includes β_1 .

Estimation of Coefficients of Interest

From Equation (4.10), an expected value of the parameter of interest β_1 could be estimated at:

$$\hat{\beta}_1 = E[\beta_1 | D, \beta_1 \neq 0] = \sum_{A_1} \hat{\beta}_1(k) p'(M_k | D) \quad (4.11)$$

$$SD^2[\beta_1 | D, \beta_1 \neq 0] = [\sum_{A_1} [se_1^2(k) + \hat{\beta}_1(k)] p'(M_k | D) - E[\beta_1 | D, \beta_1 \neq 0]^2] \quad (4.12)$$

Where $\hat{\beta}_1(k)$ and $se_1^2(k)$ are respectively the MLE and standard error of β_1 under the model M_k .

So far, I have established averaging estimated parameters using the Bayesian Model Average framework. The response values predicted by the generalized linear model regression can perform better than any individual model in out-of-sample prediction. However, this approach only solves half of the model uncertainty problem. Since I use a multivariate model, deducting an optimized credit gap threshold is still impossible through minimizing the policy loss function. I propose a solution to this issue in the following subsection.

4.4.7 Weighted Credit Gap Creation

Our motivation to create an averaged weighted gap is to find a solution to the issue of not being able to determine an optimized credit threshold discussed above. Since all credit gap variables measure one set of information: how much in absolute value the total credit series deviated from its long-run trend, I can strategically assign weights to them to create a weighted credit gap.

Our objective is to assign weights to the credit gaps to combine them and create a single credit gap that performs as well as the multivariate Bayesian averaged model in the previous subsection 4.4.6. After creating the weighted cycle, I can perform an optimized credit gap threshold analysis, assign a policy loss function value, and compare its EWI performance with other credit gap filters.

Multivariate GLM binary logistic predicted response values:

$$\widehat{pre.crisis}_{ti} = \widehat{probability}_{ti} = \frac{1}{1 + \exp(-(a + \sum_j \hat{\beta}_j c_{tij}))} \quad (4.13)$$

From the previous subsection 4.4.6, in a Bayesian averaged model: $\hat{\beta}_j = E[\beta_j | D, \beta_j \neq 0] = \sum_{A_j} \hat{\beta}_j(k) p'(M_k | D)$.

Using the information in Equation (4.13), I propose creating a single weighted credit gap \hat{c}_{ti} that satisfies:

$$\frac{1}{1 + \exp(-(a + \hat{\beta} \hat{c}_{ti}))} = \frac{1}{1 + \exp(-(a + \sum_j \hat{\beta}_j c_{tij}))}$$

Or

$$\hat{\beta} \hat{c}_{ti} = \sum_j \hat{\beta}_j c_{tij} \quad (4.14)$$

This condition allows us to create a weighted gap \hat{c}_{ti} that will have similar predicted response values and averaged characteristics of the averaged decomposition filters selected. Additionally, to further simplify construction of the weighted gap, I then propose $\hat{\beta} = \sum_j \hat{\beta}_j$.

Therefore,

$$\hat{c}_{ti} = \frac{\sum_j (\hat{\beta}_j c_{tij})}{\hat{\beta}} = \frac{\sum_j (\hat{\beta}_j c_{tij})}{\sum_j \hat{\beta}_j} = \sum_j w_j c_{tij} \quad (4.15)$$

The weight of each candidate credit gap j is $w_j = \frac{\hat{\beta}_j}{\sum_j \hat{\beta}_j}$, which is the fraction of the estimate of coefficient β_j in the averaged model over the sum of the estimated coefficients of all selected credit gaps.

One-sided Crisis-weighted Averaged Credit Gap

I save the weights w_j at every incremental period t of available data to create a one-sided weight vector w_{tj} . In the next section, Figure 4.2 illustrates time series of changes in the weights of credit gaps. I estimate the first 15 years of data available for stable model implementation as a full sample regression and assign the weights as constant during that period.

To create one-sided crisis weighted averaged credit gap for each country i (\hat{c}_{ti}), I compute:

$$\hat{c}_{ti,one-sided} = \sum_j w_{tj} * c_{tij} \quad (4.16)$$

At this point, I have established a framework to create a one-sided crisis weighted averaged credit gap. In the next section, Empirical Results, I will compare the performance as an EWI of this one-sided crisis weighted credit gap (one-sided weighted gap) with other top candidate credit gaps.

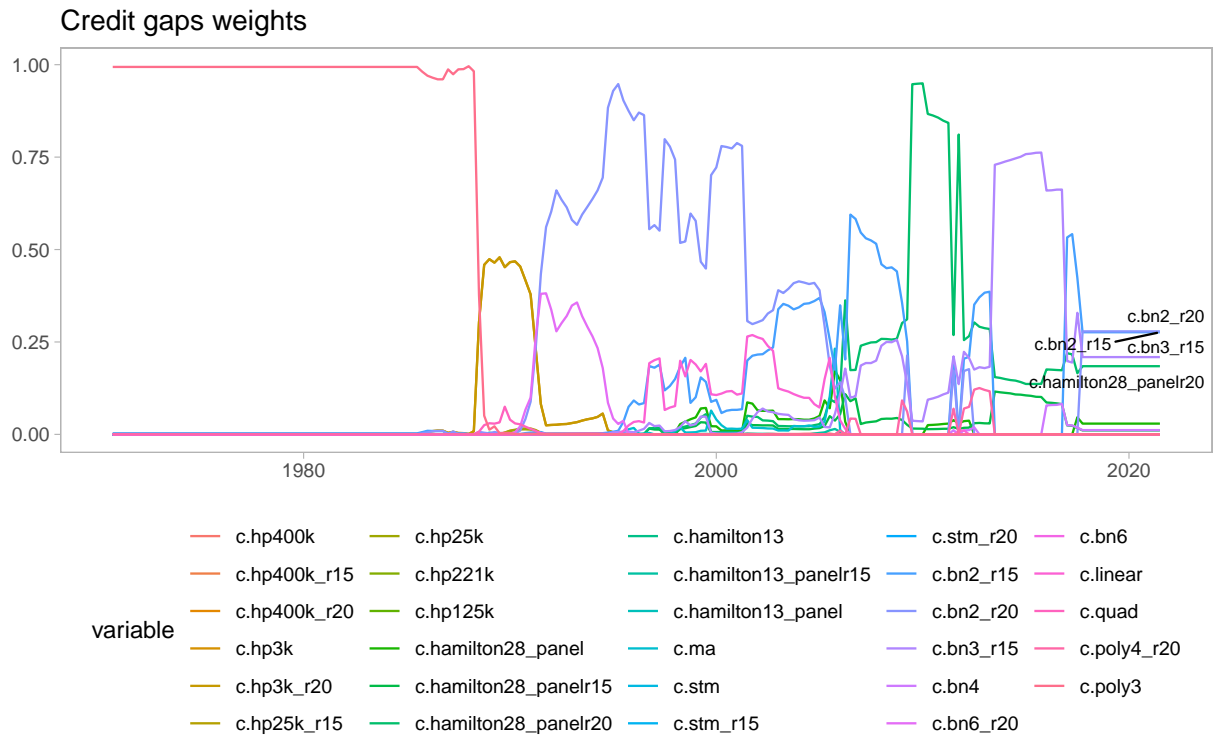
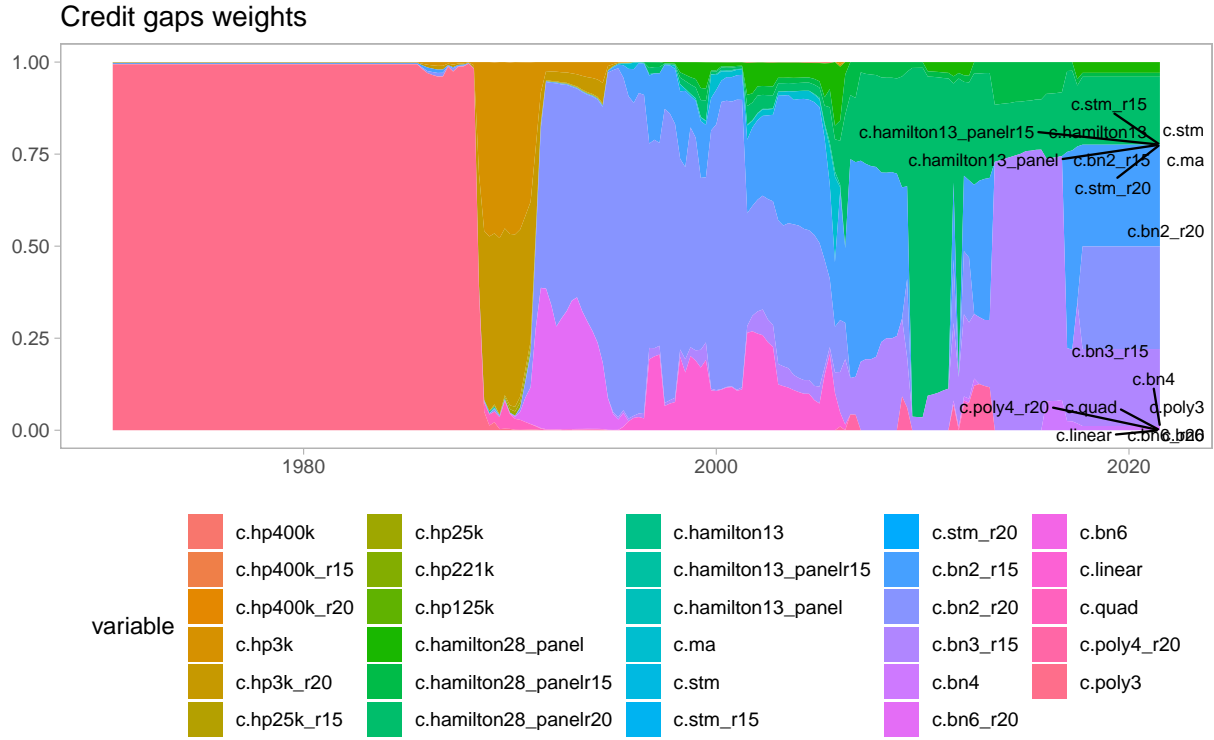


Figure 4.2: One sided weights graph

4.5 Empirical Results

4.5.1 One-sided Weights Time Series Graph

In Figure 4.2, we see that for the first 15-18 years of the data sample, the parsimony (c.poly3) 3rd degree polynomial filter's weight dominated all the other filters. However, after about 20 years, the (c.poly3) filter lost all of its weight. This could be explained by the lack of initial data, leading to poorer performance of other credit gap measurements.

However, as more data were updated, the dynamics of the weights changed significantly. During the early 1990s, Hodrick-Prescott filters' weights overshadowed others, then the Beveridge-Nelson, and the Structural Time-Series model did. As the Great Financial Crisis happened, the Hamilton filter in a panel setting gained weight. And lastly, at the end of the crisis data period 2017:Q4, we have 3 Beveridge-Nelson filters (c.bn2_r20, c.bn2_r15 and c.bn3_r15) and 1 Hamilton filter (c.hamilton28_panelr20) sharing nearly equal weights. To apply the findings in the limited sample and make predictions using latest available credit data, I fixed the weights for 2017:Q4 at constant and extrapolated the weights to 2021:Q3, the end period of this paper's credit data availability.

4.5.2 Model Fitness Results

The standard deviation for each estimate is deducted from a 95% confidence interval of the bootstrapping results using 2000 stratified bootstrap replicates.⁸

Performance Metrics

Along with the results from Table 4.3 - 4.5, I discuss psAUC as a performance measurement criteria of a credit gap series as an EWI. Firstly, the psAUC metric overall gives a consistent signal about the quality of the credit gap decomposition filters. A filter with a

⁸I estimate partial ROC and confidence interval using the R package "pROC" by Robin et al. (2011) whose methodology is based on Carpenter & Bithell (2000)

Table 4.3: Credit gaps performance as EWIs - Advanced Economies

Cycle	BIC	AIC	AUC	psAUC	c.Threshold	Type I	Type II	Policy Loss Function
null	0.0000	0.0000	0.5000	0.5000	NA	1.0000	0.0000	1.0000
1.sided weighted.cycle	-128.8749	-134.9430	0.7545 (0.0177)	0.6889 (0.0195)	3.5486	0.3177 (0.0088)	0.2841 (0.0348)	0.1817 (0.0254)
c.hamilton13.panelr15	-116.9260	-122.9941	0.7282 (0.0174)	0.6846 (0.0178)	8.8000	0.3469 (0.0087)	0.3011 (0.0348)	0.2110 (0.0270)
c.hamilton28.panelr15	-141.5772	-147.6453	0.7491 (0.0184)	0.6825 (0.0198)	11.5510	0.3834 (0.0088)	0.2216 (0.0319)	0.1961 (0.0212)
c.hamilton28.panelr20	-147.4322	-153.5003	0.7514 (0.0185)	0.6811 (0.0195)	13.1072	0.3489 (0.0086)	0.2898 (0.0348)	0.2057 (0.0262)
c.bn6.r20	-90.1022	-96.1703	0.7338 (0.0183)	0.6751 (0.0171)	0.6581	0.4126 (0.0094)	0.2273 (0.0319)	0.2219 (0.0222)
c.hamilton28.panel	-129.1730	-135.2411	0.7382 (0.0187)	0.6740 (0.0189)	9.7674	0.3977 (0.0090)	0.2386 (0.0319)	0.2151 (0.0224)
c.hamilton13.panel	-122.0054	-128.0735	0.7208 (0.0182)	0.6684 (0.0174)	7.2247	0.3509 (0.0090)	0.3239 (0.0362)	0.2280 (0.0296)
c.ma	-97.4868	-103.5549	0.7133 (0.0181)	0.6677 (0.0174)	6.0375	0.4073 (0.0087)	0.2727 (0.0333)	0.2403 (0.0255)
c.bn6	-104.1404	-110.2085	0.7042 (0.0198)	0.6426 (0.0184)	0.8256	0.4040 (0.0089)	0.2955 (0.0348)	0.2505 (0.0277)
c.hp125k	-92.8331	-98.9012	0.6977 (0.0195)	0.6397 (0.0185)	3.6972	0.3788 (0.0089)	0.3239 (0.0348)	0.2484 (0.0293)
BIS Basel gap	-113.9826	-120.0507	0.7124 (0.0200)	0.6390 (0.0197)	3.9706	0.3847 (0.0085)	0.3011 (0.0348)	0.2387 (0.0275)
c.hp221k	-104.1190	-110.1871	0.7070 (0.0198)	0.6388 (0.0197)	3.8869	0.3755 (0.0086)	0.3011 (0.0362)	0.2317 (0.0285)
c.stm.r15	-86.6063	-92.6744	0.6989 (0.0195)	0.6382 (0.0204)	4.0784	0.3486 (0.0090)	0.3295 (0.0348)	0.2301 (0.0292)
c.stm	-89.9073	-95.9754	0.6929 (0.0197)	0.6357 (0.0193)	3.3071	0.3924 (0.0091)	0.3295 (0.0348)	0.2626 (0.0300)
c.hp400k.r15	-73.5902	-79.6582	0.6869 (0.0194)	0.6357 (0.0201)	4.2355	0.3496 (0.0085)	0.3125 (0.0348)	0.2199 (0.0277)
c.linear	-111.9570	-118.0251	0.7069 (0.0206)	0.6356 (0.0177)	3.9964	0.4620 (0.0092)	0.2443 (0.0319)	0.2732 (0.0241)
c.hp400k.r20	-93.8576	-99.9257	0.6991 (0.0197)	0.6340 (0.0197)	4.3393	0.3622 (0.0090)	0.3295 (0.0362)	0.2398 (0.0306)
c.bn2.r20	-84.0800	-90.1481	0.6955 (0.0201)	0.6297 (0.0168)	0.2773	0.4216 (0.0090)	0.3182 (0.0348)	0.2790 (0.0301)
c.stm.r20	-87.6992	-93.7673	0.6879 (0.0200)	0.6282 (0.0178)	3.1055	0.4073 (0.0089)	0.3239 (0.0348)	0.2708 (0.0298)

The standard deviation for each estimates are deducted from 95% confidence interval of the bootstrapping results using 2000 stratified bootstrap replicates.

Table 4.4: Credit gaps performance as EWIs - Emerging Market Economies

Cycle	BIC	AIC	AUC	psAUC	c.Threshold	Type I	Type II	Policy Loss Function
null	0.0000	0.0000	0.5000	0.5000	NA	1.0000	0.0000	1.0000
c.bn3.r15	-46.2774	-51.3507	0.7365 (0.0471)	0.6308 (0.0451)	0.6244	0.3059 (0.0134)	0.3333 (0.0779)	0.2047 (0.0623)
c.poly3	5.3862	0.3129	0.5737 (0.0358)	0.6046 (0.0264)	1.8089	0.5280 (0.0152)	0.3056 (0.0709)	0.3721 (0.0593)
c.bn2.r15	-13.2062	-18.2795	0.6879 (0.0505)	0.5879 (0.0508)	0.2952	0.3566 (0.0140)	0.3333 (0.0779)	0.2383 (0.0642)
c.poly4.r20	7.0732	1.9999	0.5040 (0.0339)	0.5816 (0.0160)	-0.9609	0.5962 (0.0147)	0.3333 (0.0779)	0.4665 (0.0717)
1.sided weighted.cycle	6.2094	1.1361	0.5325 (0.0402)	0.5811 (0.0183)	-1.0639	0.6827 (0.0136)	0.1111 (0.0496)	0.4784 (0.0310)
c.linear	-9.3676	-14.4409	0.5787 (0.0471)	0.5774 (0.0208)	-0.9783	0.6294 (0.0143)	0.2222 (0.0709)	0.4455 (0.0495)
c.bn2.r20	-16.6411	-21.7144	0.6760 (0.0519)	0.5751 (0.0462)	0.1470	0.4510 (0.0154)	0.3056 (0.0709)	0.2968 (0.0572)
c.hamilton13	6.5749	1.5016	0.5468 (0.0386)	0.5710 (0.0229)	3.6354	0.6206 (0.0143)	0.3056 (0.0709)	0.4785 (0.0610)
c.ma	-11.6401	-16.7133	0.5572 (0.0546)	0.5457 (0.0164)	-0.2250	0.7220 (0.0134)	0.1667 (0.0567)	0.5491 (0.0382)
c.hamilton28.panel	-7.8687	-12.9420	0.5392 (0.0522)	0.5384 (0.0170)	-1.7750	0.6958 (0.0132)	0.2778 (0.0709)	0.5613 (0.0577)
c.quad	6.1997	1.1264	0.4654 (0.0474)	0.5334 (0.0140)	-6.4882	0.7456 (0.0127)	0.1944 (0.0638)	0.5938 (0.0455)
c.hamilton13.panelr15	-1.6660	-6.7393	0.5087 (0.0550)	0.5274 (0.0130)	-1.1064	0.7002 (0.0136)	0.3333 (0.0779)	0.6014 (0.0732)
c.hp25k.r15	3.6420	-1.4313	0.5018 (0.0558)	0.5265 (0.0101)	-3.5672	0.7850 (0.0125)	0.1111 (0.0496)	0.6285 (0.0320)
c.hp25k	3.9466	-1.1267	0.4975 (0.0560)	0.5247 (0.0099)	-3.7339	0.7893 (0.0118)	0.1111 (0.0496)	0.6354 (0.0310)
c.hp3k	3.3678	-1.7054	0.5276 (0.0556)	0.5235 (0.0146)	-1.1119	0.7019 (0.0136)	0.3333 (0.0850)	0.6038 (0.0758)
c.hp3k.r20	3.3703	-1.7030	0.5276 (0.0556)	0.5235 (0.0149)	-1.1125	0.7028 (0.0132)	0.3333 (0.0779)	0.6050 (0.0726)
c.hamilton13.panel	-1.6294	-6.7027	0.5166 (0.0554)	0.5222 (0.0134)	-2.9398	0.7500 (0.0132)	0.2778 (0.0709)	0.6397 (0.0591)
BIS Basel gap	-0.7015	-5.7748	0.4928 (0.0544)	0.5217 (0.0100)	-5.3969	0.7920 (0.0118)	0.1389 (0.0638)	0.6465 (0.0382)
c.hamilton28.panelr20	-4.9986	-10.0719	0.5123 (0.0544)	0.5213 (0.0147)	-1.5578	0.6932 (0.0136)	0.3333 (0.0710)	0.5916 (0.0663)

The standard deviation for each estimates are deducted from 95% confidence interval of the bootstrapping results using 2000 stratified bootstrap replicates.

Table 4.5: Credit gaps performance as EWIs - Full sample

Cycle	BIC	AIC	AUC	psAUC	c.Threshold	Type I	Type II	Policy Loss Function
null	0.0000	0.0000	0.5000	0.5000	NA	1.0000	0.0000	1.0000
1.sided weighted.cycle	-127.5308	-133.9135	0.7182 (0.0174)	0.6454 (0.0165)	2.8892	0.3532 (0.0076)	0.3255 (0.0313)	0.2307 (0.0257)
c.bn6.r20	-108.0679	-114.4506	0.7048 (0.0184)	0.6379 (0.0176)	0.6581	0.3962 (0.0077)	0.3019 (0.0313)	0.2481 (0.0250)
c.hamilton28.panel	-149.8518	-156.2346	0.7107 (0.0185)	0.6359 (0.0169)	9.7674	0.3912 (0.0075)	0.3066 (0.0313)	0.2470 (0.0251)
c.ma	-120.8108	-127.1936	0.6922 (0.0181)	0.6313 (0.0170)	5.7813	0.3989 (0.0077)	0.3160 (0.0313)	0.2590 (0.0259)
c.hamilton13.panelr15	-126.2968	-132.6796	0.6924 (0.0180)	0.6311 (0.0175)	6.5289	0.4297 (0.0079)	0.2830 (0.0313)	0.2647 (0.0245)
c.hamilton28.panelr20	-164.6015	-170.9842	0.7158 (0.0189)	0.6302 (0.0181)	10.8558	0.3948 (0.0074)	0.2925 (0.0313)	0.2414 (0.0241)
c.hamilton28.panelr15	-154.4533	-160.8361	0.7091 (0.0190)	0.6270 (0.0184)	11.5510	0.3854 (0.0077)	0.2972 (0.0325)	0.2369 (0.0254)
c.hamilton13.panel	-133.9347	-140.3175	0.6922 (0.0184)	0.6250 (0.0169)	4.9769	0.4285 (0.0075)	0.2877 (0.0301)	0.2664 (0.0239)
c.bn2.r20	-109.3128	-115.6955	0.6963 (0.0188)	0.6218 (0.0170)	0.2776	0.4080 (0.0079)	0.3255 (0.0313)	0.2724 (0.0268)
c.linear	-135.4069	-141.7896	0.6879 (0.0191)	0.6204 (0.0145)	3.9989	0.4616 (0.0079)	0.2925 (0.0325)	0.2986 (0.0264)
c.bn6	-132.7915	-139.1742	0.6835 (0.0194)	0.6113 (0.0188)	0.4710	0.4371 (0.0075)	0.2830 (0.0313)	0.2712 (0.0242)
c.bn2.r15	-83.9469	-90.3297	0.6749 (0.0192)	0.6047 (0.0149)	0.1349	0.4761 (0.0074)	0.3302 (0.0313)	0.3357 (0.0277)
c.poly4.r20	3.5738	-2.8090	0.5772 (0.0154)	0.6011 (0.0144)	0.1651	0.4980 (0.0077)	0.3302 (0.0313)	0.3570 (0.0284)
BIS Basel gap	-121.5910	-127.9738	0.6733 (0.0197)	0.5960 (0.0164)	3.0578	0.4441 (0.0077)	0.3255 (0.0325)	0.3032 (0.0281)
c.bn4	-169.1186	-175.5014	0.6892 (0.0204)	0.5943 (0.0206)	1.2840	0.3837 (0.0075)	0.3255 (0.0325)	0.2532 (0.0268)
c.stm.r15	-79.5531	-85.9358	0.6575 (0.0193)	0.5924 (0.0167)	2.0027	0.4778 (0.0075)	0.3160 (0.0325)	0.3281 (0.0279)
c.hp125k	-92.2897	-98.6725	0.6562 (0.0194)	0.5924 (0.0164)	2.5216	0.4547 (0.0077)	0.3302 (0.0337)	0.3158 (0.0293)
c.hp221k	-106.8842	-113.2670	0.6656 (0.0196)	0.5921 (0.0163)	2.6641	0.4561 (0.0077)	0.3160 (0.0325)	0.3079 (0.0274)
c.hp400k.r15	-67.1228	-73.5055	0.6472 (0.0190)	0.5912 (0.0164)	2.6223	0.4592 (0.0075)	0.3255 (0.0313)	0.3168 (0.0272)

The standard deviation for each estimates are deducted from 95% confidence interval of the bootstrapping results using 2000 stratified bootstrap replicates.

higher psAUC value, in general, will have a higher quality set of possible policy loss function values to optimize. The standard deviation values of the weighted cycle metrics are also comparable to the other gap measurements.

In some scenarios, psAUC can be a deciding factor if policymakers are indifferent about the AUC value of a filter. In Table 4.5, the Moving Average (c.ma) and Hamilton 13th lag using panel setting (c.hamilton13.panel) filters have the same AUC values. While the BIC and AIC values, the overall linearized binary logistic regression fitness criteria, would favor (c.hamilton13.panel) filter. The psAUC indicates otherwise. If we only focus on the region where FPR or the optimized threshold's predictive power is at least $2/3$, the minimized policy loss function would favor c.ma.

However, there are also instances where psAUC failed to be a single criterion for determining the performance of a filter. In Table 4.5, the 4th-degree polynomial filter with a rolling sample of 20 years (c.poly4.r20) have a higher psAUC value than the BIS Basel gap. However, all of (c.poly4.r20) filter's other metrics have less preferred values than the BIS Basel gap. As a result, we deem (c.poly4.r20) less preferred than the BIS Basel gap.

In conclusion, even though psAUC is an overall good measurement for a specific case of binary regression model, where Type II errors are costlier than Type I errors, and we focus the optimized threshold on the region where Type II errors are less than $1/3$; We still have to use it in conjunction with other criteria such as BIC, AUC, and loss function to determine the performance of a filter. Using it alone will give biased results of the model performance.

Weighted Gap and other Gaps Performance Comparison

In this part, I will discuss the performance of the weighted credit gap series as an EWI and compare it with other selected series.

From Table 4.3 and Table 4.5 regarding Advanced Economies and full sample results, the weighted credit cycle outperformed other cycles in AUC, psAUC, and Policy Loss function values. Specifically, while maintaining approximately the same optimized threshold as the

BIS Basel gap, the weighted credit cycle can achieve lower Type I and II error rates than the BIS Basel gap in all samples.

The findings in Advanced Economies and full sample results also align with the literature findings. Hamilton and Moving Average filters with optimized parameters perform well, which agrees with the findings in Beltran et al. (2021). The findings also agree with Drehmann & Yetman (2021) proposal of using Hamilton filter in a panel setting with fixed sloping coefficients across countries, and with Galán (2019) who found rolling sample with 15 and 20 years window helps with improvement of the early indicator performance.

On the other hand, across the three tables, optimizing and changing parameters on the one-sided HP filter does not improve its performance over the BIS Basel gap. This is contrary to the finding in Beltran et al. (2021) and Galán (2019). In (Galán, 2019), the author did not use a full sample analysis but by individual countries. The structural time series filter proposed in Beltran et al. (2021) also did not perform better than the BIS Basel gap.

Unsurprisingly, parsimony filters such as linear and polynomial does not perform well in AE and full sample. However, they can outperform other well-studied gaps in EMEs. This confirmed my finding in the previous section 4.5.1 that when data are limited, and the market economies are not well established, parsimony filters outperformed other complex filters.

I want to reiterate the importance of avoiding the pitfall of using a single criterion for selecting models, even for the psAUC metric. In Table 4.4. The (c.poly3) filter has a higher psAUC value than its next filter on the table (c.bn2.r15). However, (c.poly3)'s other metrics have lower quality than its counterpart. The policy loss function, Type I and II errors rates are higher for (c.poly3). Therefore, c.bn2.r15 is preferred over c.poly3.

Last but not least, the emerging market economies (EMEs) have proven challenging to predict crises using the credit gap as an EWI because of their more limited sample sizes and less established financial systems. Both Beltran et al. (2021) and Drehmann & Yetman (2021) found that credit gap-based EWIs do not perform well in emerging market economies.

In this paper, I found evidence that adding Beveridge-Nelson decomposition filters to the model selection process improved the model performance for the EMEs, specifically the (c.bn3.r15) Beveridge-Nelson filter with a smoothing parameter of 3 lags and 15 years rolling sample. Later graphical analysis in this paper on EMEs will include (c.bn3.r15) filter as another reference. Another Beveridge-Nelson filter with specific features such as (c.bn6.r20) also performs well in both AE and full samples.

4.5.3 Out-of-Sample Forecast

I followed Alessi & Detken (2018) in implementing k-fold leave-one-out cross-validation to perform an out-of-sample forecast analysis. To assess the out-of-sample predictive power of the credit gaps, I perform a 3-fold cross-validation, in which the data is randomly split into three partitions. For each partition, the model is estimated using the other two groups. The resulting estimated parameter is then used to predict the dependent variable in the unused group. I then compute psAUCs and other metrics values and obtain the average metrics and their standard deviations from 10 repeats.

The overall results of the out-of-sample prediction exercise show similar findings as in the previous subsection 4.5.2 that used bootstrapping method. The ranking of the filters does not change significantly. These results overall affirm the robustness of using credit gaps as EWI models.

On Table 4.6 regarding the AEs, the weighted cycle has a lower value than the (c.hamilton.panelr15) filter. However, the weighted cycle has a lower loss function value and lower overall Type I and Type II error rates.

Expectedly, for EMEs, on Table 4.7 the parsimony (c.poly3) 3th power polynomial filter dropped its psAUC in out-of-sample prediction context compared to its bootstrapped value on Table 4.4. This could be predicted by its poor metrics such as BIC and Loss Function on the model fitness results in the previous subsection.

Through the two subsections 4.5.2 and 4.5.3, the model proved to perform well in identi-

Table 4.6: Credit gaps performance as EWIs - Out of sample prediction - AE

Cycle	BIC	AUC	psAUC	c.Threshold	Type I	Type II	Policy Loss Function
c.null	0.0000 (1.4201)	0.4729 (0.0099)	0.4906 (0.0032)	0.0000	1.0000	0.0000	1.0000
c.hamilton13.panelr15	-92.1757 (3.1133)	0.7247 (0.0050)	0.6801 (0.0053)	8.0229 (1.3409)	0.3694 (0.0218)	0.2909 (0.0245)	0.2220 (0.0085)
1.sided weighted.cycle	-83.3823 (5.0485)	0.7469 (0.0084)	0.6776 (0.0127)	3.7972 (0.8036)	0.3261 (0.0200)	0.2915 (0.0253)	0.1923 (0.0158)
c.hamilton28.panelr15	-102.2043 (3.1634)	0.7442 (0.0036)	0.6755 (0.0059)	12.1456 (1.7511)	0.3595 (0.0202)	0.2744 (0.0300)	0.2058 (0.0070)
c.hamilton28.panelr20	-106.5651 (3.8537)	0.7469 (0.0040)	0.6748 (0.0061)	12.2954 (1.8840)	0.3625 (0.0151)	0.2722 (0.0213)	0.2061 (0.0037)
c.hamilton28.panel	-89.7510 (2.9903)	0.7354 (0.0024)	0.6705 (0.0036)	10.0439 (1.9885)	0.3930 (0.0157)	0.2585 (0.0226)	0.2220 (0.0041)
c.hamilton13.panel	-101.4619 (1.6108)	0.7184 (0.0022)	0.6665 (0.0022)	6.7644 (1.0174)	0.3714 (0.0260)	0.2983 (0.0355)	0.2287 (0.0039)
c.ma	-69.1570 (3.2513)	0.7091 (0.0036)	0.6630 (0.0051)	6.0750 (0.7204)	0.3994 (0.0244)	0.2818 (0.0405)	0.2410 (0.0065)
c.bn6.r20	-72.3756 (2.7401)	0.7062 (0.0122)	0.6353 (0.0158)	0.6105 (1.4701)	0.4187 (0.0333)	0.3028 (0.0349)	0.2691 (0.0324)
c.hp125k	-71.8112 (5.1241)	0.6930 (0.0039)	0.6351 (0.0030)	3.5090 (1.0440)	0.4030 (0.0250)	0.3057 (0.0217)	0.2568 (0.0098)
BIS Basel gap	-87.6905 (2.4274)	0.7089 (0.0029)	0.6350 (0.0049)	4.4951 (0.9743)	0.3818 (0.0255)	0.3131 (0.0283)	0.2451 (0.0059)
c.stm.r15	-65.4667 (2.8230)	0.6950 (0.0018)	0.6349 (0.0017)	3.7851 (0.7083)	0.3723 (0.0170)	0.3216 (0.0108)	0.2424 (0.0102)
c.hp400k.r15	-58.7622 (3.6325)	0.6830 (0.0028)	0.6330 (0.0037)	4.2265 (0.9549)	0.3756 (0.0276)	0.3051 (0.0191)	0.2352 (0.0108)
c.linear	-87.1078 (2.9343)	0.7045 (0.0024)	0.6328 (0.0035)	4.8208 (0.8031)	0.4265 (0.0104)	0.3040 (0.0184)	0.2747 (0.0091)
c.hp221k	-79.5029 (4.4132)	0.7001 (0.0051)	0.6311 (0.0070)	3.6423 (1.1691)	0.3779 (0.0149)	0.3165 (0.0154)	0.2434 (0.0116)
c.stm	-71.3061 (2.9283)	0.6881 (0.0028)	0.6298 (0.0044)	2.9453 (0.8258)	0.4040 (0.0143)	0.3153 (0.0140)	0.2631 (0.0071)
c.hp400k.r20	-74.2327 (2.0353)	0.6940 (0.0050)	0.6286 (0.0067)	3.7081 (1.2202)	0.3932 (0.0258)	0.3165 (0.0085)	0.2554 (0.0208)
c.bn6	-76.6398 (2.4656)	0.6914 (0.0100)	0.6243 (0.0141)	1.1094 (0.6884)	0.4392 (0.0417)	0.2960 (0.0324)	0.2830 (0.0268)
c.stm.r20	-70.6443 (2.8287)	0.6818 (0.0049)	0.6219 (0.0051)	2.9958 (1.0397)	0.4207 (0.0069)	0.3227 (0.0128)	0.2813 (0.0047)
c.bn4	-94.6995 (1.7677)	0.6964 (0.0072)	0.6174 (0.0103)	1.8741 (0.8777)	0.3953 (0.0306)	0.3125 (0.0149)	0.2549 (0.0286)

Note:

3-fold cross-validation results. Standard errors are reported in parentheses.

Table 4.7: Credit gaps performance as EWIs - Out of sample prediction - EME

Cycle	BIC	AUC	psAUC	c.Threshold	Type I	Type II	Policy Loss Function
c.null	0.0000 (2.1053)	0.4445 (0.0294)	0.4828 (0.0081)	0.0000	1.0000	0.0000	1.0000
c.bn3.r15	-42.6016 (2.2377)	0.7211 (0.0293)	0.6224 (0.0301)	0.3484 (0.4116)	0.3928 (0.0764)	0.3111 (0.0255)	0.2569 (0.0709)
c.linear	-18.9590 (4.6826)	0.5665 (0.0110)	0.5675 (0.0123)	0.7792 (2.2166)	0.6094 (0.0510)	0.2889 (0.0559)	0.4599 (0.0407)
c.bn2.r20	-7.2882 (2.2840)	0.6564 (0.0162)	0.5579 (0.0164)	0.1180 (0.3002)	0.4755 (0.0576)	0.3167 (0.0268)	0.3300 (0.0459)
c.bn2.r15	-3.5368 (1.8649)	0.6418 (0.0392)	0.5570 (0.0247)	0.0252 (0.4690)	0.4911 (0.0757)	0.3306 (0.0088)	0.3557 (0.0745)
c.poly3	1.7710 (0.9026)	0.5143 (0.0404)	0.5499 (0.0367)	3.0514 (7.6000)	0.6313 (0.0705)	0.2917 (0.0458)	0.4900 (0.1062)
c.hamilton28.panel	-23.8796 (4.1713)	0.5309 (0.0074)	0.5267 (0.0097)	-0.8265 (4.2248)	0.6839 (0.0330)	0.3028 (0.0357)	0.5615 (0.0352)
c.ma	-26.0114 (7.1586)	0.5466 (0.0068)	0.5244 (0.0168)	1.2194 (1.5300)	0.6885 (0.0302)	0.3000 (0.0450)	0.5666 (0.0401)
1.sided weighted.cycle	1.6191 (0.8146)	0.4787 (0.0301)	0.5219 (0.0306)	1.5348 (4.9140)	0.6621 (0.0877)	0.3056 (0.0393)	0.5400 (0.1028)
c.hamilton13	1.6095 (0.8780)	0.4816 (0.0404)	0.5197 (0.0347)	1.7494 (3.5850)	0.6581 (0.0923)	0.3111 (0.0410)	0.5391 (0.1137)
c.quad	1.6542 (0.9705)	0.4498 (0.0291)	0.5040 (0.0267)	-8.3825 (9.5714)	0.7454 (0.0623)	0.2806 (0.0357)	0.6389 (0.0948)
c.hamilton13.panel	-7.5078 (7.0689)	0.4985 (0.0153)	0.5037 (0.0107)	-0.4802 (2.9623)	0.7158 (0.0277)	0.3139 (0.0187)	0.6119 (0.0342)
c.hamilton28.panellr20	-18.0060 (5.4148)	0.5006 (0.0075)	0.5034 (0.0176)	-5.4259 (2.7504)	0.7463 (0.0456)	0.2833 (0.0410)	0.6407 (0.0542)
BIS Basel gap	-6.5423 (5.8202)	0.4791 (0.0128)	0.5023 (0.0138)	-4.3181 (2.1378)	0.7816 (0.0374)	0.2417 (0.0764)	0.6759 (0.0408)
c.hamilton28.panellr15	-16.2657 (6.9303)	0.4841 (0.0190)	0.5022 (0.0214)	-2.1619 (3.4804)	0.7701 (0.0471)	0.2583 (0.0508)	0.6641 (0.0625)
c.hp3k	1.0131 (0.9964)	0.5102 (0.0157)	0.5017 (0.0232)	-0.2747 (1.0254)	0.7488 (0.0234)	0.2861 (0.0322)	0.6440 (0.0445)
c.hamilton13.panellr15	-11.6105 (6.0026)	0.4688 (0.0612)	0.4976 (0.0259)	1.2963 (6.5172)	0.7628 (0.0461)	0.2861 (0.0541)	0.6682 (0.0660)
c.hp25k.r15	1.0960 (1.5266)	0.4604 (0.0283)	0.4957 (0.0135)	-1.7970 (2.1943)	0.7691 (0.0419)	0.2917 (0.0419)	0.6797 (0.0437)
c.hp221k	-3.6880 (3.7753)	0.4692 (0.0089)	0.4955 (0.0139)	-3.9673 (1.9937)	0.7913 (0.0415)	0.2306 (0.0859)	0.6874 (0.0354)
c.bn6.r20	-6.2230 (4.0505)	0.5007 (0.0238)	0.4948 (0.0125)	-0.1197 (0.6508)	0.6855 (0.0541)	0.3028 (0.0306)	0.5650 (0.0699)

Note:

3-fold cross-validation results. Standard errors are reported in parentheses.

Table 4.8: Credit gaps performance as EWIs - Out of sample prediction - Full sample

Cycle	BIC	AUC	psAUC	c.Threshold	Type I	Type II	Policy Loss Function
c.null	0.0000 (1.8172)	0.4767 (0.0128)	0.4907 (0.0049)	0.0000	1.0000	0.0000	1.0000
1.sided weighted.cycle	-80.9638 (5.0654)	0.7134 (0.0033)	0.6404 (0.0037)	2.9939 (0.6513)	0.3707 (0.0282)	0.3208 (0.0092)	0.2411 (0.0164)
c.hamilton28.panel	-121.3928 (1.7471)	0.7078 (0.0025)	0.6318 (0.0032)	9.6472 (0.3081)	0.4043 (0.0195)	0.3071 (0.0155)	0.2583 (0.0113)
c.hamilton13.panehr15	-120.5975 (2.5998)	0.6903 (0.0016)	0.6291 (0.0019)	6.6028 (0.5314)	0.4120 (0.0158)	0.3071 (0.0187)	0.2646 (0.0056)
c.ma	-98.5015 (7.0341)	0.6892 (0.0018)	0.6276 (0.0030)	5.5675 (0.6929)	0.4045 (0.0146)	0.3113 (0.0201)	0.2611 (0.0059)
c.hamilton28.panehr20	-139.0052 (3.0642)	0.7134 (0.0018)	0.6272 (0.0027)	10.5537 (1.1148)	0.3910 (0.0119)	0.3042 (0.0165)	0.2458 (0.0070)
c.hamilton13.panel	-123.8627 (4.9205)	0.6894 (0.0021)	0.6220 (0.0034)	5.1997 (0.6877)	0.4206 (0.0194)	0.3132 (0.0199)	0.2757 (0.0092)
c.hamilton28.panehr15	-132.7462 (5.5529)	0.7052 (0.0041)	0.6217 (0.0053)	11.6613 (0.5448)	0.3874 (0.0245)	0.3127 (0.0135)	0.2486 (0.0141)
c.linear	-106.5647 (4.8347)	0.6846 (0.0017)	0.6158 (0.0028)	4.0160 (1.5341)	0.4621 (0.0110)	0.3113 (0.0156)	0.3108 (0.0080)
c.bn2.r20	-76.7863 (2.1205)	0.6908 (0.0043)	0.6137 (0.0066)	0.2458 (0.2248)	0.4327 (0.0239)	0.3208 (0.0116)	0.2907 (0.0149)
c.bn6.r20	-86.4636 (2.8925)	0.6757 (0.0160)	0.5987 (0.0205)	0.2144 (1.1466)	0.4713 (0.0471)	0.3231 (0.0114)	0.3286 (0.0445)
c.bn6	-101.2737 (4.0553)	0.6693 (0.0130)	0.5935 (0.0176)	0.3956 (0.6125)	0.4611 (0.0472)	0.3250 (0.0061)	0.3203 (0.0459)
BIS Basel gap	-110.9026 (4.6247)	0.6707 (0.0018)	0.5932 (0.0025)	3.3012 (0.7148)	0.4490 (0.0080)	0.3250 (0.0052)	0.3073 (0.0059)
c.hp221k	-100.8887 (2.0973)	0.6632 (0.0020)	0.5908 (0.0015)	3.4199 (0.5762)	0.4530 (0.0103)	0.3292 (0.0020)	0.3137 (0.0098)
c.hp125k	-88.2934 (2.9343)	0.6540 (0.0012)	0.5905 (0.0019)	2.4639 (0.7551)	0.4691 (0.0079)	0.3236 (0.0067)	0.3248 (0.0089)
c.bn2.r15	-47.5774 (2.8755)	0.6637 (0.0110)	0.5901 (0.0152)	0.1127 (0.2168)	0.5010 (0.0270)	0.3212 (0.0090)	0.3550 (0.0294)
c.bn4	-143.6399 (1.8203)	0.6835 (0.0043)	0.5889 (0.0061)	0.9797 (0.7268)	0.4146 (0.0233)	0.3236 (0.0116)	0.2772 (0.0186)
c.stm.r15	-66.3874 (2.6907)	0.6530 (0.0024)	0.5879 (0.0025)	2.4833 (0.8012)	0.4742 (0.0311)	0.3179 (0.0190)	0.3272 (0.0208)
c.stm	-84.2480 (3.8697)	0.6500 (0.0021)	0.5879 (0.0026)	2.3616 (0.5334)	0.4767 (0.0063)	0.3250 (0.0068)	0.3330 (0.0056)
c.hp400k.r15	-61.6848 (2.2021)	0.6439 (0.0014)	0.5876 (0.0034)	2.6897 (0.8146)	0.4627 (0.0125)	0.3226 (0.0102)	0.3184 (0.0075)

Note:

3-fold cross-validation results. Standard errors are reported in parentheses.

fyng at-risk periods before a crisis and normal periods of countries in the sample, especially regarding its criteria for selecting robust EWI early warning indicators and its method of weighted-averaging the top candidates for improved model certainty and out-of-sample forecast. With this evidence, I will move our discussion to specific individual countries' analysis in the next section.

4.6 Analyzing Time Series Graphs of Individual Countries

This section will discuss the BIS Basel gap and weighted gap as EWIs for the US and UK. Based on the estimated optimized gap threshold values for the weighted gap and BIS Basel gap reported in Table 4.5 and 4.8 for full sample results, I will use an optimized credit gap threshold of 3.00 for ease of comparison. The black shaded area on Figures 4.3 and 4.4 ranges from 2018:Q1 to 2021:Q3, representing periods where we do not have sufficient crisis data, but credit data are available.

For the US, the weighted gap and BIS Basel gap performed similarly. They were able to predict the late 80s - early 90s financial crisis. They raised above the threshold starting early 2000s, which would result in misclassification of pre-crisis periods (Type I error) since my model aims to only identify periods of 5-12 months before a systemic financial crisis that would not happen until 2008. Nevertheless, the two gap series showed a build-up of an unsustainable credit gap for a prolonged period before the Great Financial Crisis (GFC) in the US. It is also worth pointing out that the weighted gap dipped below the threshold again before climbing again to identify the pre-crisis period of the GFC correctly.

The graph for the UK tells quite a similar story to the US. Both the credit gaps were able to detect the pre-crisis periods in the early 1990s and the GFC on 2008 for the UK. However, regarding the 90s crisis, the BIS Basel gap went above its threshold too early, for about five years. This could trigger unnecessary counter-cyclical buffer measurements from

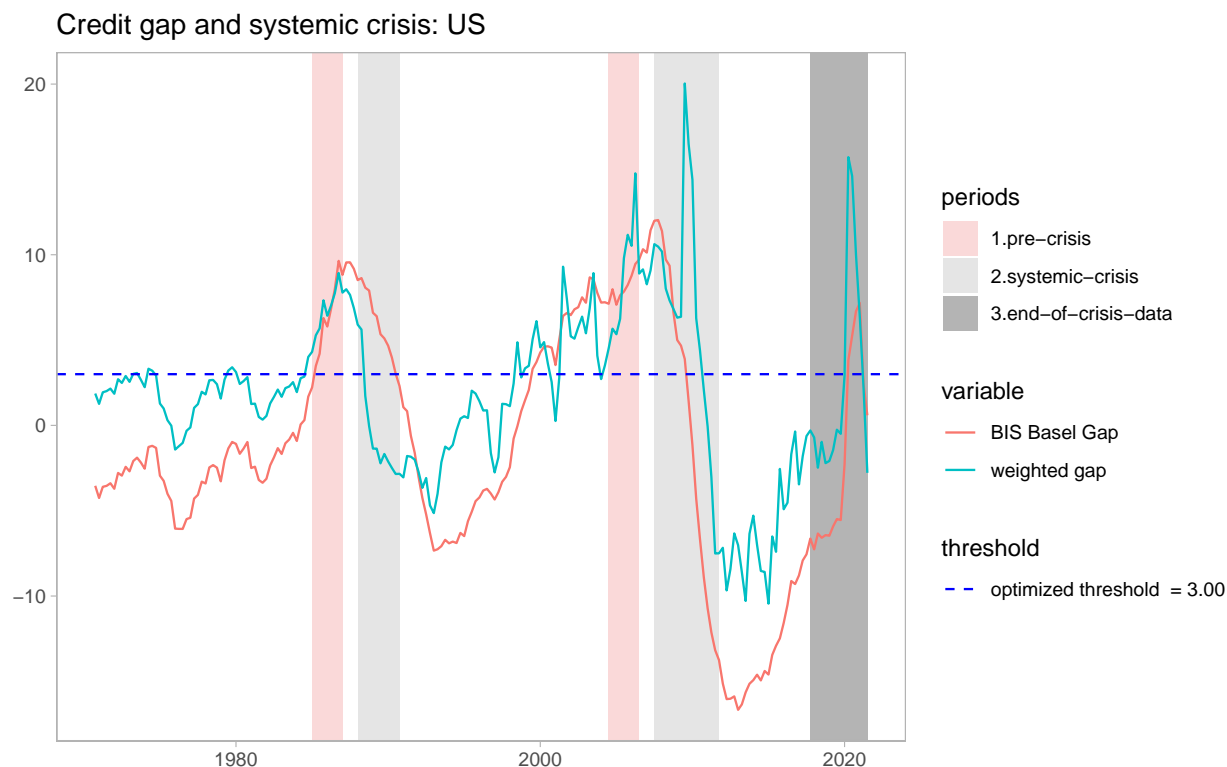


Figure 4.3: US time series graph

the central bank and cause the economy to slow its growth rate because of credit constraints. The weighted gap performed much better than the BIS Basel gap as it only sent out the signal during the pre-crisis period of the early 1990s for the UK.

From the graphical features of the weighted gap, the magnitude and length of the weighted gaps are comparable to that of the BIS Basel gap, which has a long smooth trend generating a large credit gap. Additionally, the weighted gap also retains the more volatile features and short spikes of sudden credit growth that the Beveridge-Nelson and other shorter-term cycle decomposition methods have. These averaged features help explain the averaged method's improved performance as an EWI over other traditional decomposition filters.

From here, to continue my analysis beyond 2017:Q4, I extrapolated the weights on one-sided credit gaps as constants to create a weighted credit gap outside of the crisis data range. This allows me to extrapolate my prediction until 2021:Q3. In response to COVID-19, most countries increased their total credit through multiple stimulus policies, or their GDP growth

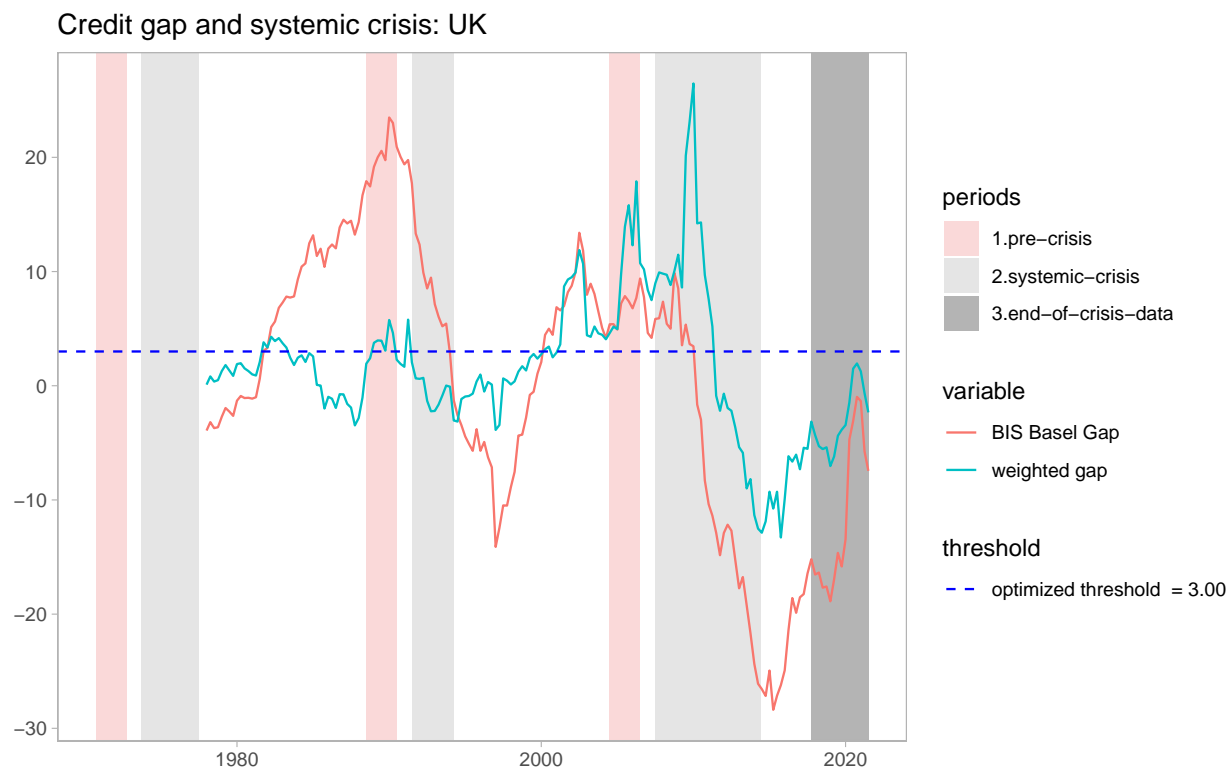


Figure 4.4: UK time series graph

slowed down, both scenarios created a significant upswing in the magnitude of the credit gap measurement above its optimized pre-crisis detection threshold. Per Drehmann & Yetman (2021), a good EWI signal would have the ideal characteristic of a stable signal meaning the credit gap need to be above the threshold and remain stable above the threshold during the pre-crisis period. I assume from my model finding that an economy with a credit gap above its optimized threshold for 8 consecutive quarters has a 2/3rd chance of having a financial crisis after the next 4 quarters.

As for the US, even though its credit gap was above the threshold for a few quarters, the gap eventually dropped below the threshold. This does not give us enough evidence to suggest that the country is at risk of a systemic financial crisis in the short future. However, there was an increase in the risk of crisis as the credit gap went beyond the threshold substantially during COVID-19 period. For the UK, the credit gaps did not exceed its threshold. Therefore, there is no evidence to show that the UK is at risk of a future financial

crisis.

Moreover, I performed a graphical analysis of all the countries in the sample and identified countries with significant credit gaps above the optimized 3.00 credit gap threshold beyond the 2017:Q4 crisis data limit window. I selected nine countries with signs of being at pre-crisis risk and should have performed pre-emptive macroprudential measures. They are Canada, France, Hong Kong (SAR), Japan, South Korea, Saudi Arabia, Switzerland, Sweden, and Thailand.⁹

There are three outliers in those nine countries: Canada, Hong Kong (SAR), and Switzerland. Since the end of the GFC, the credit gaps in these four countries have been substantially above their optimized threshold. Additionally, those three countries and Saudi Arabia never experienced systemic financial crises. Therefore, the EWI model performance for those three countries is not robust, and the evidence only comes from abroad. The other countries: France, Sweden, Japan, South Korea, and Thailand have evidence to show pre-emptive macroprudential measures need to be implemented to reduce the risk of future systemic financial crises.

4.7 Conclusion

Our study is based on the idea that excessive credit growth measured by the total credit to GDP deviation could be used as an early warning indicator for future systemic financial crises. I build a novel framework to select relevant credit gap decomposition filters, average their combinations, and create a crisis-weighted credit gap that inherits the combined features of the selected filters. I also proposed a novel performance metric for early warning indicator psAUC with a specific constraint that has appealing policy implication properties.

Our findings in this paper provide robust evidence to show the superior predictive performance of our proposed weighted credit gap over other univariate early warning indicator models that use individual credit gap filters. The novel proposed metric psAUC also showed

⁹graphs of those countries are included in Appendix C.2.1

that it has appealing properties for policy maker. It simplifies the steps to search for a decomposition filter with minimum loss function in the range of credit gap thresholds that have Type II error rates smaller than $1/3$. Therefore, the weighted gap and novel metric can be regarded as useful additional reference points for policy implications.

References

- Agnello, L., Castro, V., & Sousa, R. M. (2018). Economic Activity, Credit Market Conditions, and the Housing Market. *Macroeconomic Dynamics*, 22, 1769–1789.
- Agnello, L., & Schuknecht, L. (2011). Booms and busts in housing markets: Determinants and implications. *Journal of Housing Economics*, 20, 171–190.
- Aikman, D., Haldane, A. G., & Nelson, B. D. (2015). Curbing the Credit Cycle. *The Economic Journal*, 125, 1072–1109.
- Aldasoro, I., Borio, C. E. V., & Drehmann, M. (2018). *Early Warning Indicators of Banking Crises: Expanding the Family* (SSRN Scholarly Paper No. ID 3139160). Rochester, NY: Social Science Research Network.
- Alessi, L., & Detken, C. (2018). Identifying excessive credit growth and leverage. *Journal of Financial Stability*, 35, 215–225.
- Babecký, J., Havránek, T., Matějů, J., Rusnák, M., Šmídková, K., & Vašíček, B. (2014). Banking, debt, and currency crises in developed countries: Stylized facts and early warning indicators. *Journal of Financial Stability*, 15, 1–17.
- Banking Supervision (BCBS), B. C. for. (2010). Guidance for national authorities operating the countercyclical capital buffer. *Bank for International Settlements*, December.
- Barberis, N., Shleifer, A., & Vishny, R. W. (2005). Chapter 12. A Model of Investor Sentiment. In R. H. Thaler (Ed.), *Advances in Behavioral Finance, Volume II* (pp. 423–459). Princeton University Press.
- Bates, J. M., & Granger, C. W. J. (1969). The Combination of Forecasts. *Journal of the Operational Research Society*, 20, 451–468.
- Beltran, D. O., Jahan-Parvar, M. R., & Paine, F. A. (2021). Optimizing Credit Gaps for Predicting Financial Crises: Modelling Choices and Tradeoffs. *International Finance Discussion Paper*, 2021, 1–40.
- Bernanke, B., & Gertler, M. (1989). Agency Costs, Net Worth, and Business Fluctuations. *American Economic Review*, 19.

- Beveridge, S., & Nelson, C. R. (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the “business cycle”. *Journal of Monetary Economics*, 7, 151–174.
- Blake, A. P., & Mumtaz, H. (2017). Applied Bayesian econometrics for central bankers. In *Technical Books*. Centre for Central Banking Studies, Bank of England.
- Boissay, F., Collard, F., & Smets, F. (2016). Booms and Banking Crises. *Journal of Political Economy*, 124, 489–538.
- Borio, C. (2014). The financial cycle and macroeconomics: What have we learnt? *Journal of Banking & Finance*, 45, 182–198.
- Borio, C. E. V., & Drehmann, M. (2009). *Assessing the Risk of Banking Crises – Revisited* (SSRN Scholarly Paper No. ID 1513316). Rochester, NY: Social Science Research Network.
- Borio, C. E. V., & Lowe, P. W. (2002). *Asset Prices, Financial and Monetary Stability: Exploring the Nexus* (SSRN Scholarly Paper No. ID 846305). Rochester, NY: Social Science Research Network.
- Borio, C., & Lowe, P. (2002). Assessing the risk of banking crises. *BIS Quarterly Review*, 7, 43–54.
- Campagnoli, P., Petrone, S., & Petris, G. (2009). *Dynamic Linear Models with R*. New York, NY: Springer New York.
- Capozza, D. R., Hendershott, P. H., & Mack, C. (2004). An Anatomy of Price Dynamics in Illiquid Markets: Analysis and Evidence from Local Housing Markets. *Real Estate Economics*, 32, 1–32.
- Carpenter, J., & Bithell, J. (2000). Bootstrap confidence intervals: When, which, what? A practical guide for medical statisticians. *Statistics in Medicine*, 19, 1141–1164.
- Clark, P. K. (1987). The Cyclical Component of U. S. Economic Activity. *The Quarterly Journal of Economics*, 102, 797–814.
- Cogley, T., & Nason, J. M. (1995). Effects of the Hodrick-Prescott filter on trend and

- difference stationary time series Implications for business cycle research. *Journal of Economic Dynamics and Control*, 19, 253–278.
- Committee, B. (2010). An assessment of the long-term economic impact of stronger capital and liquidity requirements. *Basel Committee on Banking Supervision*.
- Davis, M. A., & Van Nieuwerburgh, S. (2015). Housing, Finance, and the Macroeconomy. In *Handbook of Regional and Urban Economics* (Vol. 5, pp. 753–811). Elsevier.
- Deghi, A., Katagiri, M., Shahid, M. S., & Valckx, N. (2020). *Predicting Downside Risks to House Prices and Macro-Financial Stability*. International Monetary Fund.
- Detken, C., Weeken, O., Alessi, L., Bonfim, D., Boucinha, M., Castro, C., ... Welz, P. (2014). Operationalising the Countercyclical Capital Buffer: Indicator Selection, Threshold Identification and Calibration Options. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3723336>
- Di Maggio, M., & Kermani, A. (2017). Credit-Induced Boom and Bust. *The Review of Financial Studies*, 30, 3711–3758.
- Drehmann, M., Borio, C. E. V., Gambacorta, L., Jiménez, G., & Trucharte, C. (2010). Countercyclical Capital Buffers: Exploring Options. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1648946>
- Drehmann, M., Borio, C. E. V., & Tsatsaronis, K. (2011). *Anchoring Countercyclical Capital Buffers: The Role of Credit Aggregates* (SSRN Scholarly Paper No. ID 1959299). Rochester, NY: Social Science Research Network.
- Drehmann, M., & Juselius, M. (2014). Evaluating early warning indicators of banking crises: Satisfying policy requirements. *International Journal of Forecasting*, 30, 759–780.
- Drehmann, M., & Tsatsaronis, K. (2014). *The Credit-to-GDP Gap and Countercyclical Capital Buffers: Questions and Answers* (SSRN Scholarly Paper No. ID 2457108). Rochester, NY: Social Science Research Network.
- Drehmann, M., & Yetman, J. (2018). *Why you should use the Hodrick-Prescott filter - at least to generate credit gaps*. 24.

- Drehmann, M., & Yetman, J. (2021). Which Credit Gap Is Better at Predicting Financial Crises? A Comparison of Univariate Filters. *International Journal of Central Banking*, 31.
- Durbin, J., & Koopman, S. J. (2012). *Time series analysis by state space methods*. Oxford university press.
- Favara, G., & Imbs, J. (2015). Credit Supply and the Price of Housing. *American Economic Review*, 105, 958–992.
- Favilukis, J., Kohn, D., Ludvigson, S., & Van Nieuwerburgh, S. (2012). *International Capital Flows and House Prices: Theory and Evidence* (No. w17751; p. w17751). Cambridge, MA: National Bureau of Economic Research.
- Favilukis, J., Ludvigson, S. C., & Van Nieuwerburgh, S. (2017). The Macroeconomic Effects of Housing Wealth, Housing Finance, and Limited Risk Sharing in General Equilibrium. *Journal of Political Economy*, 125, 140–223.
- Furnival, G. M., & Wilson, R. W. (2000). Regressions by leaps and bounds. *Technometrics*, 42, 69–79.
- Galán, J. (2019). Measuring Credit-To-GDP Gaps. The Hodrick-Prescott Filter Revisited. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3384613>
- Galati, G., Hindrayanto, I., Koopman, S. J., & Vlekke, M. (2016). Measuring financial cycles in a model-based analysis: Empirical evidence for the United States and the euro area. *Economics Letters*, 145, 83–87.
- Gerdrup, K., Kvinlog, A. B., & Schaanning, E. (2013). *Key indicators for a countercyclical capital buffer in Norway - Trends and Uncertainty*. 44.
- Glaeser, E. L., Gyourko, J., & Saiz, A. (2008). Housing supply and housing bubbles. *Journal of Urban Economics*, 64, 198–217.
- Glaeser, E. L., & Nathanson, C. G. (2017). An extrapolative model of house price dynamics. *Journal of Financial Economics*, 126, 147–170.
- Guerrieri, V., & Uhlig, H. (2016). Housing and Credit Markets: Bubbles and Crashes.

- Handbook of Macroeconomics*, 110.
- Hamilton, J. D. (2018). Why You Should Never Use the Hodrick-Prescott Filter. *The Review of Economics and Statistics*, 100, 831–843.
- He, C., Wright, R., & Zhu, Y. (2015). Housing and liquidity. *Review of Economic Dynamics*, 18, 435–455.
- Head, A., Lloyd-Ellis, H., & Sun, H. (2014). Search, Liquidity, and the Dynamics of House Prices and Construction. *American Economic Review*, 104, 1172–1210.
- Hodrick, R. J., & Prescott, E. C. (1997). Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking*, 29, 1–16.
- Hodrick, R. J., & Zhang, X. (2003). Forecasting real GDP growth with Hodrick–Prescott filtered components: An out-of-sample analysis. *Discussion Paper*.
- Holopainen, M., & Sarlin, P. (2017). Toward robust early-warning models: A horse race, ensembles and model uncertainty. *Quantitative Finance*, 17, 1933–1963.
- Hong, H., & Stein, J. C. (1999). A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets. *The Journal of Finance*, 54, 2143–2184.
- Huang, A., & Kishor, N. K. (2019). The rise of dollar credit in emerging market economies and US monetary policy. *The World Economy*, 42, 530–551.
- Huo, Z., & Ríos-Rull, J.-V. (2016). *Financial Frictions, Asset Prices, and the Great Recession* [Preprint]. Staff Report.
- Jordà, Ò., Schularick, M., & Taylor, A. M. (2015). Betting the house. *Journal of International Economics*, 96, S2–S18.
- Jordà, Ò., Schularick, M., & Taylor, A. M. (2016). The great mortgaging: Housing finance, crises and business cycles. *Economic Policy*, 31, 107–152.
- Jordà, Ò., Schularick, M., & Taylor, A. M. (2017). Macrofinancial History and the New Business Cycle Facts. *NBER Macroeconomics Annual*, 31, 213–263.
- Justiniano, A., Primiceri, G. E., & Tambalotti, A. (2019). Credit supply and the housing boom. *Journal of Political Economy*, 127, 1317–1350.

- Kaminsky, G. L., & Reinhart, C. M. (1999). The Twin Crises: The Causes of Banking and Balance-of-Payments Problems. *THE AMERICAN ECONOMIC REVIEW*, 89, 132.
- Kermani, A. (2012). Cheap credit, collateral and the boom-bust cycle. *University of California-Berkeley, Working Paper*.
- Kim, C.-J., & Nelson, C. R. (1999). State-space models with regime switching: Classical and Gibbs-sampling approaches with applications. *MIT Press Books*, 1.
- Kishor, N. K. (2020). Forecasting real-time economic activity using house prices and credit conditions. *Journal of Forecasting*. Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1002/for.2710>
- Kishor, N. K., Kumari, S., & Song, S. (2015). Time variation in the relative importance of permanent and transitory components in the U.S. Housing market. *Finance Research Letters*, 12, 92–99.
- Kiyotaki, N., & Moore, J. (1997). Credit Cycles. *Journal of Political Economy*, 105, 211–248.
- Knoll, K. (2016). *Return predictability in international housing markets, 1870-2014*. Dissertation draft, University of Bonn.
- Knoll, Katharina, Schularick, M., & Steger, T. (2017). No Price Like Home: Global House Prices, 1870-2012. *American Economic Review*, 107, 331–353.
- Laeven, L., & Valencia, F. (2013). Systemic Banking Crises Database. *IMF Economic Review*, 61, 225–270.
- Lo Duca, M., Koban, A., Basten, M., Bengtsson, E., Klaus, B., Kusmierczyk, P., ... Peltonen, T. (2017). A new database for financial crises in European countries. *ECB Occasional Paper*.
- Madigan, D., York, J., & Allard, D. (1995). Bayesian Graphical Models for Discrete Data. *International Statistical Review / Revue Internationale de Statistique*, 63, 215.
- Manias, K. C. (1978). Panics, and crashes: A history of financial crises. *Kindleberger—Basic Books*.—1978.—272.
- McClish, D. K. (1989). Analyzing a Portion of the ROC Curve. *Medical Decision Making*,

9, 190–195.

- Mian, A., & Sufi, A. (2011). House Prices, Home Equity-Based Borrowing, and the US Household Leverage Crisis. *American Economic Review*, 101, 2132–2156.
- Mian, A., & Sufi, A. (2018). *Credit Supply and Housing Speculation* (No. w24823). National Bureau of Economic Research.
- Minsky, H. P. (1977). The financial instability hypothesis: An interpretation of Keynes and an alternative to “standard” theory. *Challenge*, 20, 20–27.
- Morley, J. C. (2007). The Slow Adjustment of Aggregate Consumption to Permanent Income. *Journal of Money, Credit and Banking*, 39, 615–638.
- Myerson, R. B. (2012). A Model of Moral-Hazard Credit Cycles. *Journal of Political Economy*, 120, 847–878.
- Nelson, C. R. (2008). The Beveridge–Nelson decomposition in retrospect and prospect. *Journal of Econometrics*, 146, 202–206.
- Orphanides, A., & Norden, S. van. (2002). The Unreliability of Output-Gap Estimates in Real Time. *The Review of Economics and Statistics*, 84, 569–583.
- Ortalo-Magne, F., & Rady, S. (2006). Housing market dynamics: On the contribution of income shocks and credit constraints. *The Review of Economic Studies*, 73, 459–485.
- Raftery, A. E. (1995). Bayesian Model Selection in Social Research. *Sociological Methodology*, 25, 111.
- Ravn, M. O., & Uhlig, H. (2002). On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations. *Review of Economics and Statistics*, 84, 371–376.
- Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, 12, 77.
- Rotemberg, J. J., & Woodford, M. (1996). Real-Business-Cycle Models and the Forecastable Movements in Output, Hours, and Consumption. *The American Economic Review*, 86, 71–89.

- Schularick, M., & Taylor, A. M. (2012). Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870-2008. *American Economic Review*, 102, 1029–1061.
- Schüler, Y. S., Hiebert, P., & Peltonen, T. A. (2015). Characterising the Financial Cycle: A Multivariate and Time-Varying Approach. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2664126>
- Stein, J. C. (1995). Prices and Trading Volume in the Housing Market: A Model with Down-Payment Effects*. *The Quarterly Journal of Economics*, 110, 379–406.
- Wakerly, E. C., Scott, B. G., & Nason, J. M. (2006). Common trends and common cycles in Canada: Who knew so much has been going on? *Canadian Journal of Economics/Revue Canadienne d'économique*, 39, 320–347.

Appendix A

Chapter 2

A.1 Model Estimation - Initial Values Selection

The starting values/priors for vector autoregressive parameters in matrix F from equation 2.10 are taken from VAR regression of the one-sided HP filter cycle decomposition of the series with $\lambda = 125,000$.

For $\beta_{0|0}$, I set $\tau_{0|0}$ as the value HP filtered trend component and omit the first observation from the regression. $c_{0|0}$ cycle components are also set to be equal to their HP filter counterpart. Variance $var(\tau_{0|0}) = 100 + 50 * random$, with the random value drawn from a uniform distribution (0,1). While other measures of the starting covariance are set to be at their unconditional mean values.

Means and standard deviations for the prior distributions shown in subsection A.2 are also derived from the information extracted using the method above. All autoregressive parameters are jointly correlated in a multivariate normal distribution, while the variance parameters have inverse gamma distributions. The shape and scale parameters of a gamma distribution can be deducted using the methods of moments from the estimated means and variances. Lastly, the prior distribution for correlation coefficients are set to be normal distribution with arbitrary unassuming means of near zero and standard deviations of 0.25.

A.2 Random-Walk Metropolis Hasting Sampler

The Bayesian method is a full information-based approach that uses all moments of the observations. Together with reasonable constraints on parameters to ensure the stability of the model, we can estimate complex time series state-space models. I use the last 1 million

draws from the 1.5 million (Markov chain Monte Carlo) MCMC chains for the analysis of the posterior distribution. The posterior and prior distribution graphs and the posterior chains can show us evidence of convergence or estimation stability. The posterior chain in Figure A.7 shows that there is a lack of convergence of the estimates compared to other models. This could be a signal that the model estimation is misspecified.

The steps to implement a Random-Walk Metropolis Hasting sampler in a state-space setting with a Kalman filter can be referred to in chapter 5 from (Blake & Mumtaz, 2017).

A.2.1 Posterior and Prior Distribution

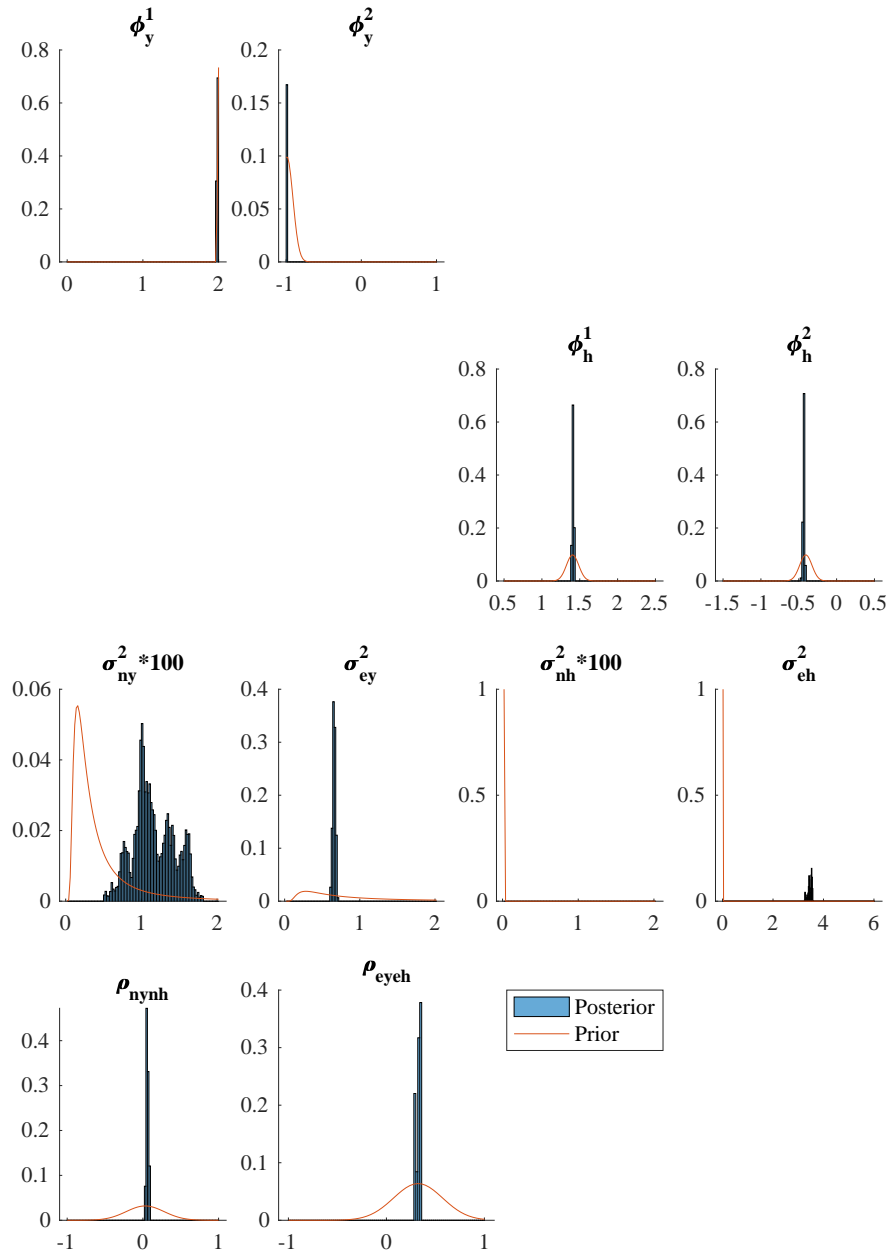


Figure A.1: UK VAR(2) - Posterior and Prior Distribution

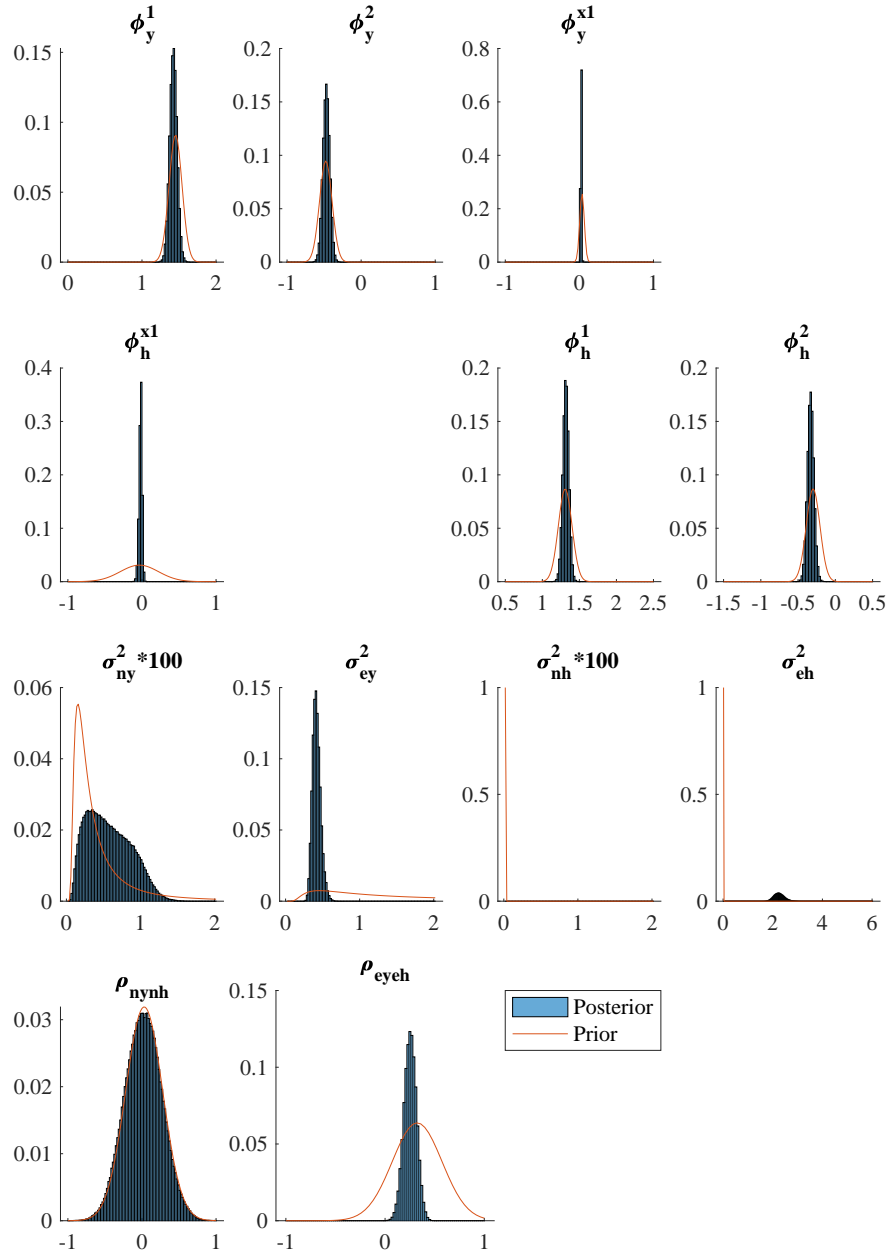


Figure A.2: UK VAR(2) 1 cross-lag - Posterior and Prior Distribution

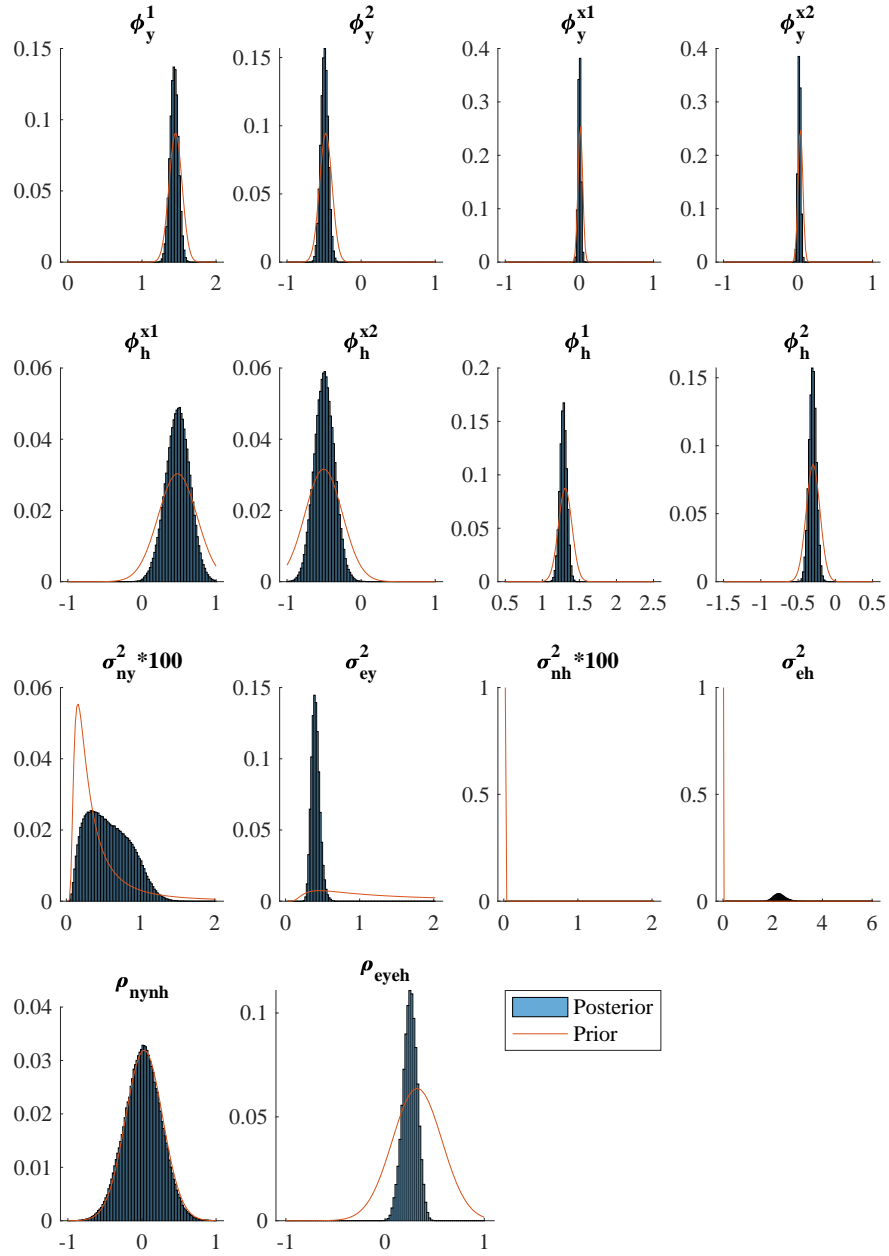


Figure A.3: UK VAR(2) 2 cross-lags - Posterior and Prior Distribution

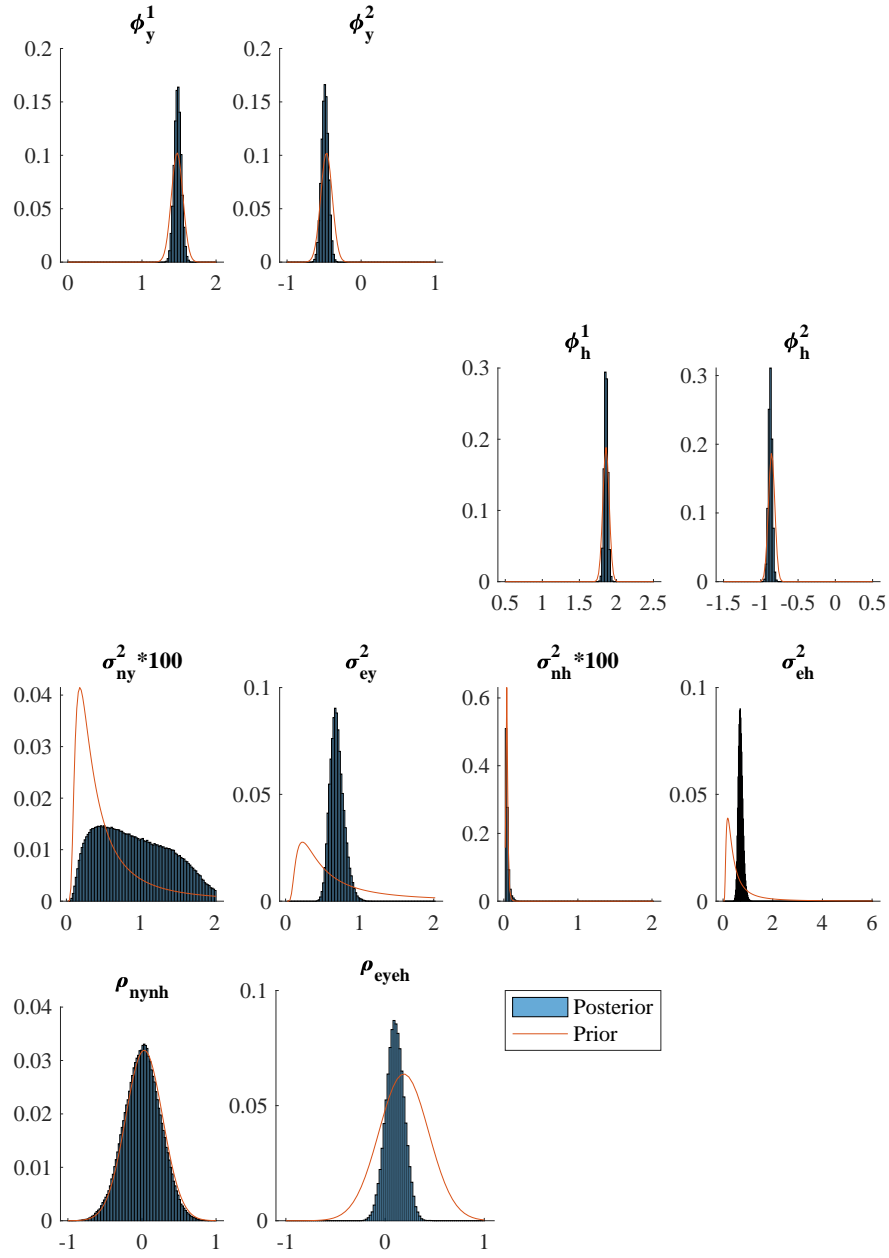


Figure A.4: US VAR(2) - Posterior and Prior Distribution

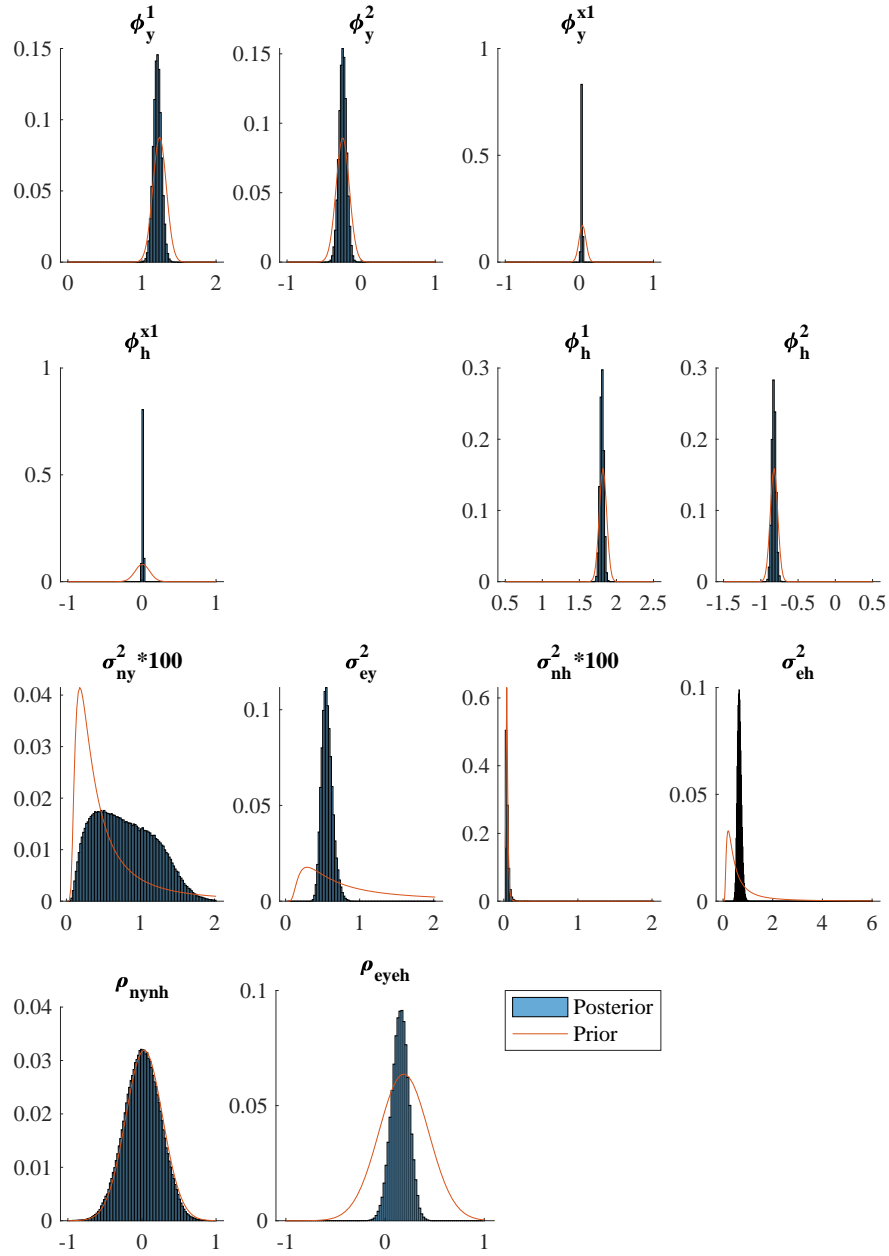


Figure A.5: US VAR(2) 1 cross-lag - Posterior and Prior Distribution

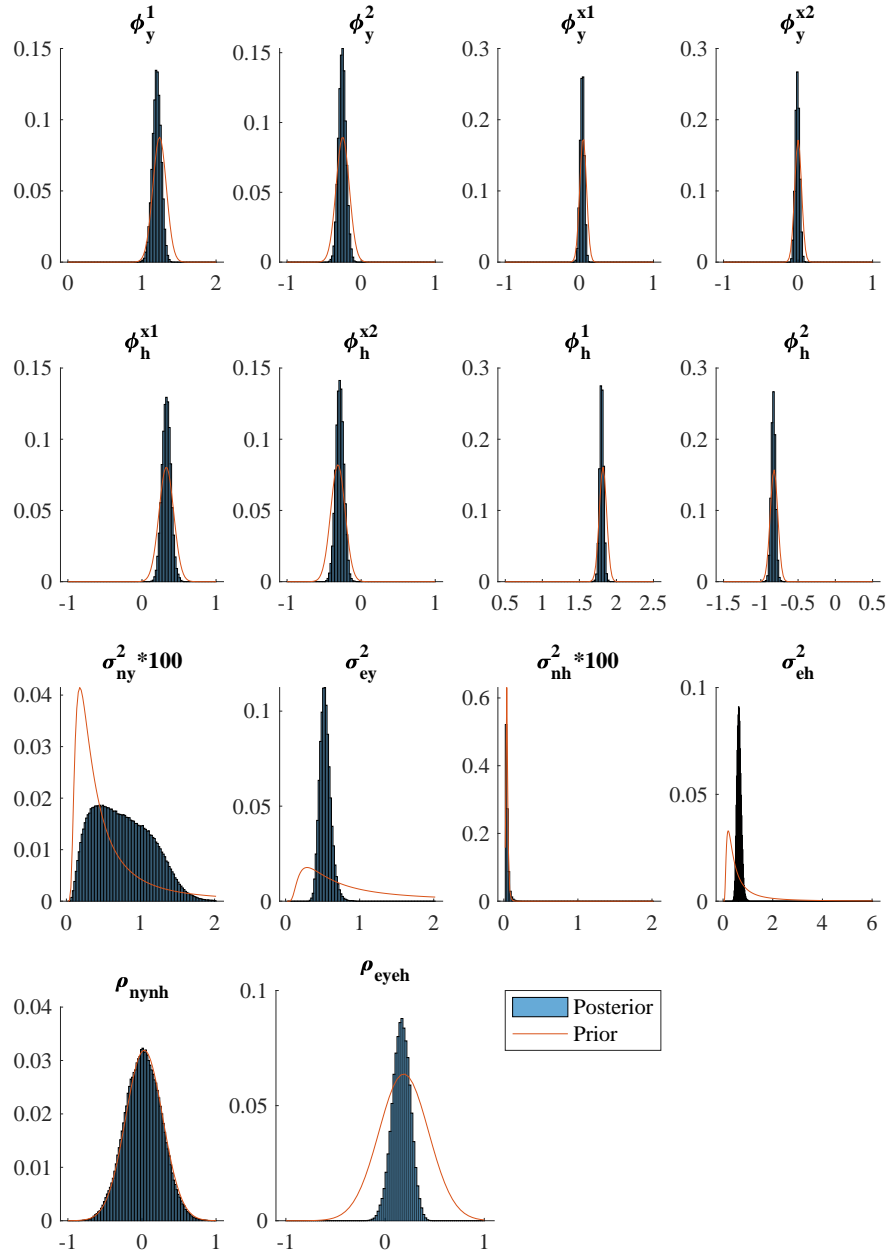


Figure A.6: US VAR(2) 2 cross-lags - Posterior and Prior Distribution

A.2.2 Posterior Chain

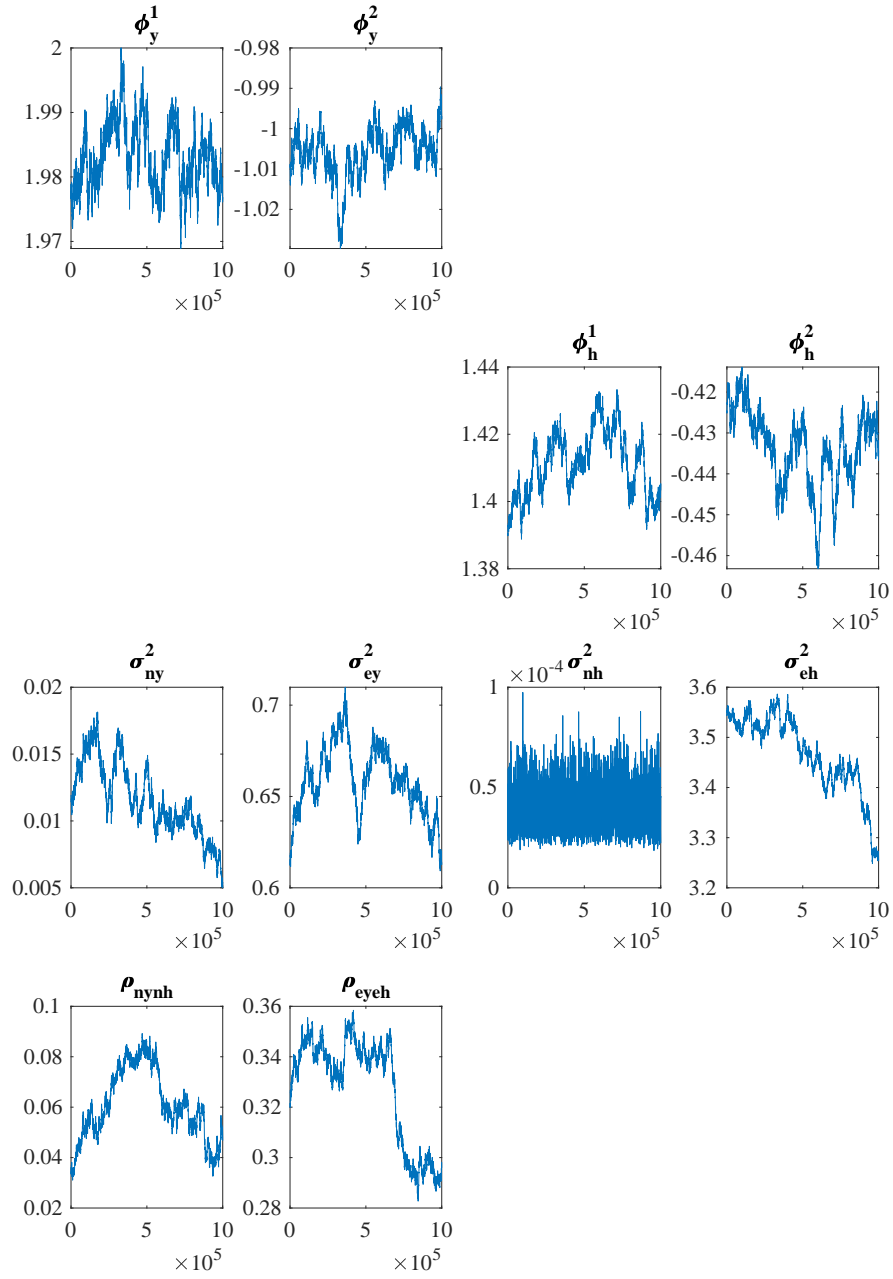


Figure A.7: UK VAR(2) - Posterior chain

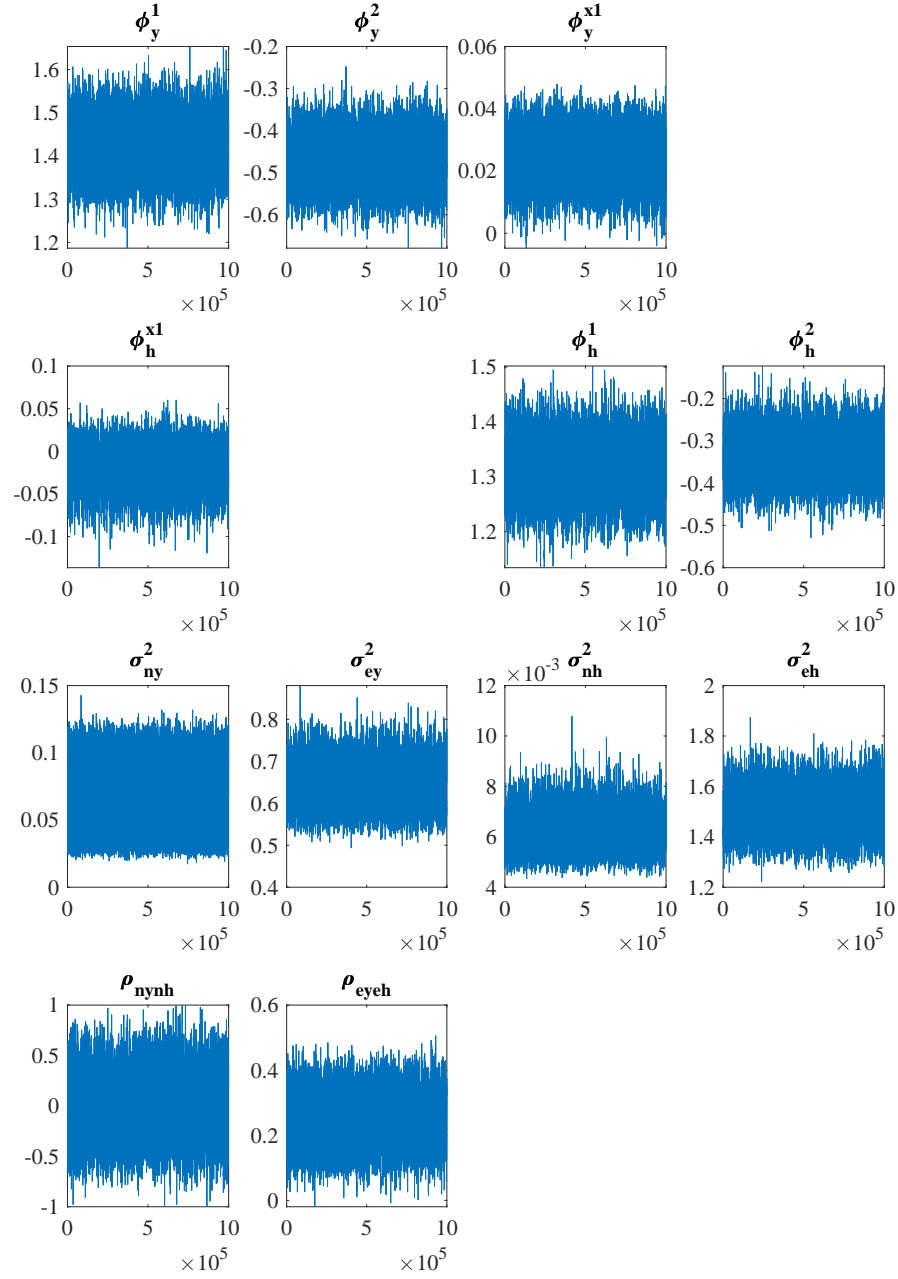


Figure A.8: UK VAR(2) 1 cross-lag - Posterior chain

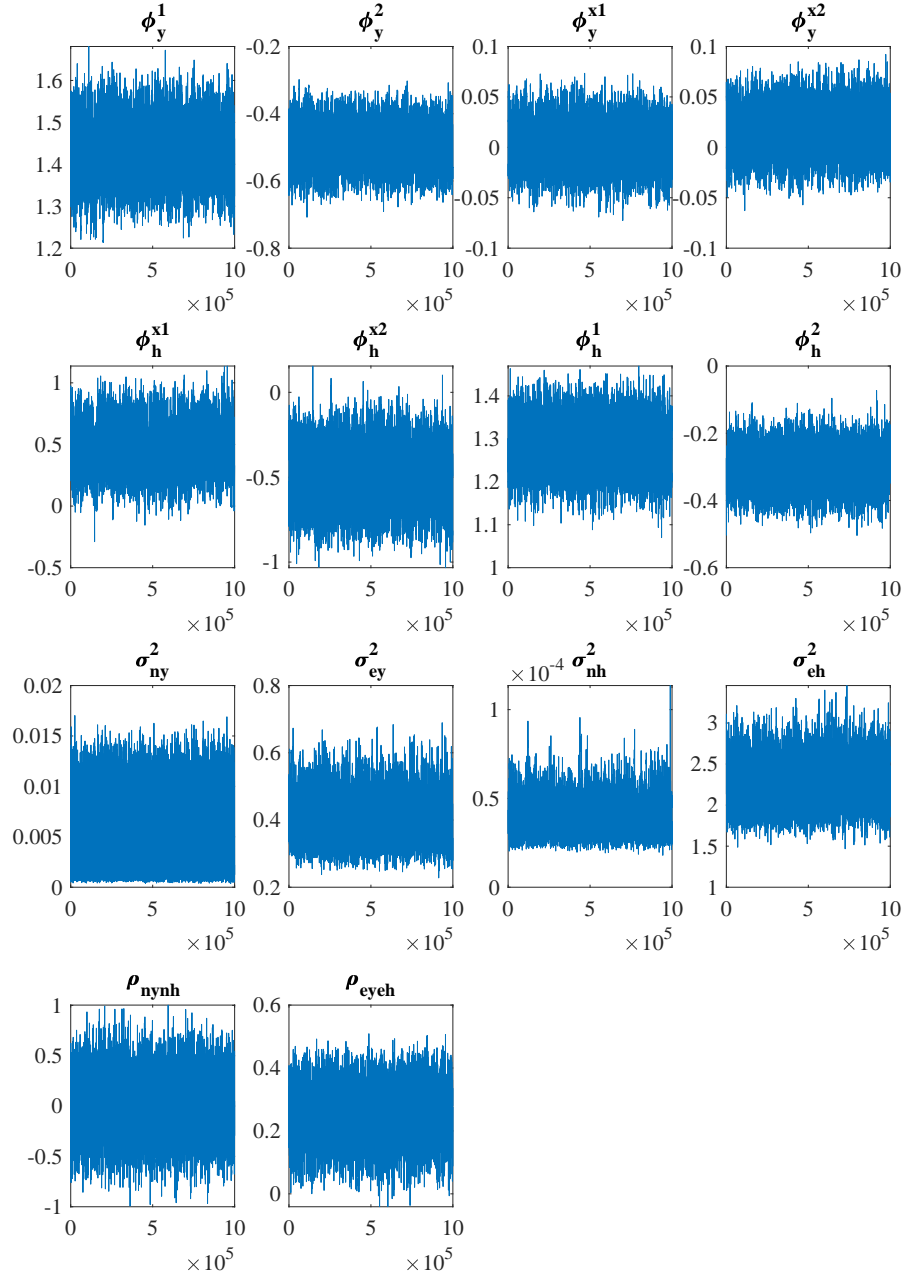


Figure A.9: UK VAR(2) 2 cross-lags - Posterior chain

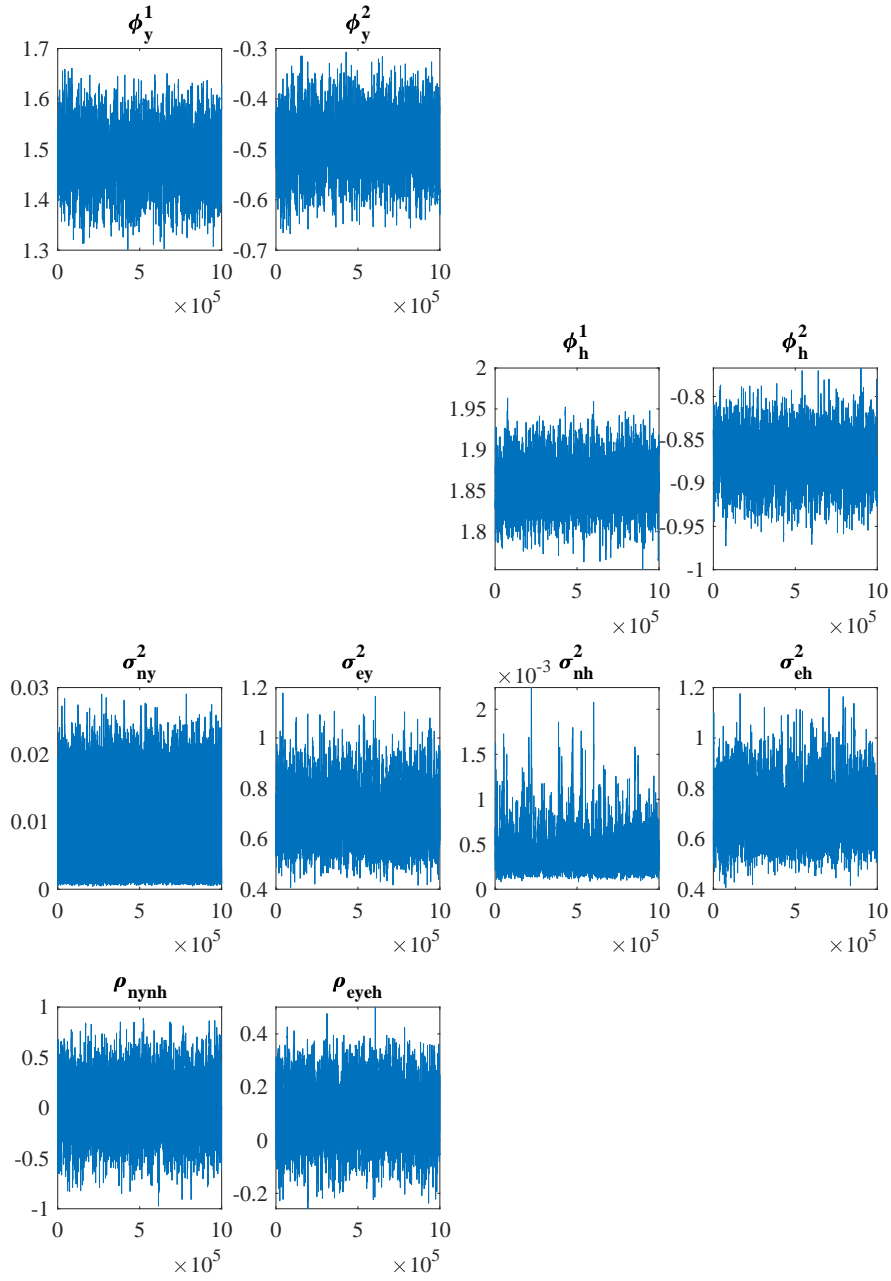


Figure A.10: US VAR(2) - Posterior chain

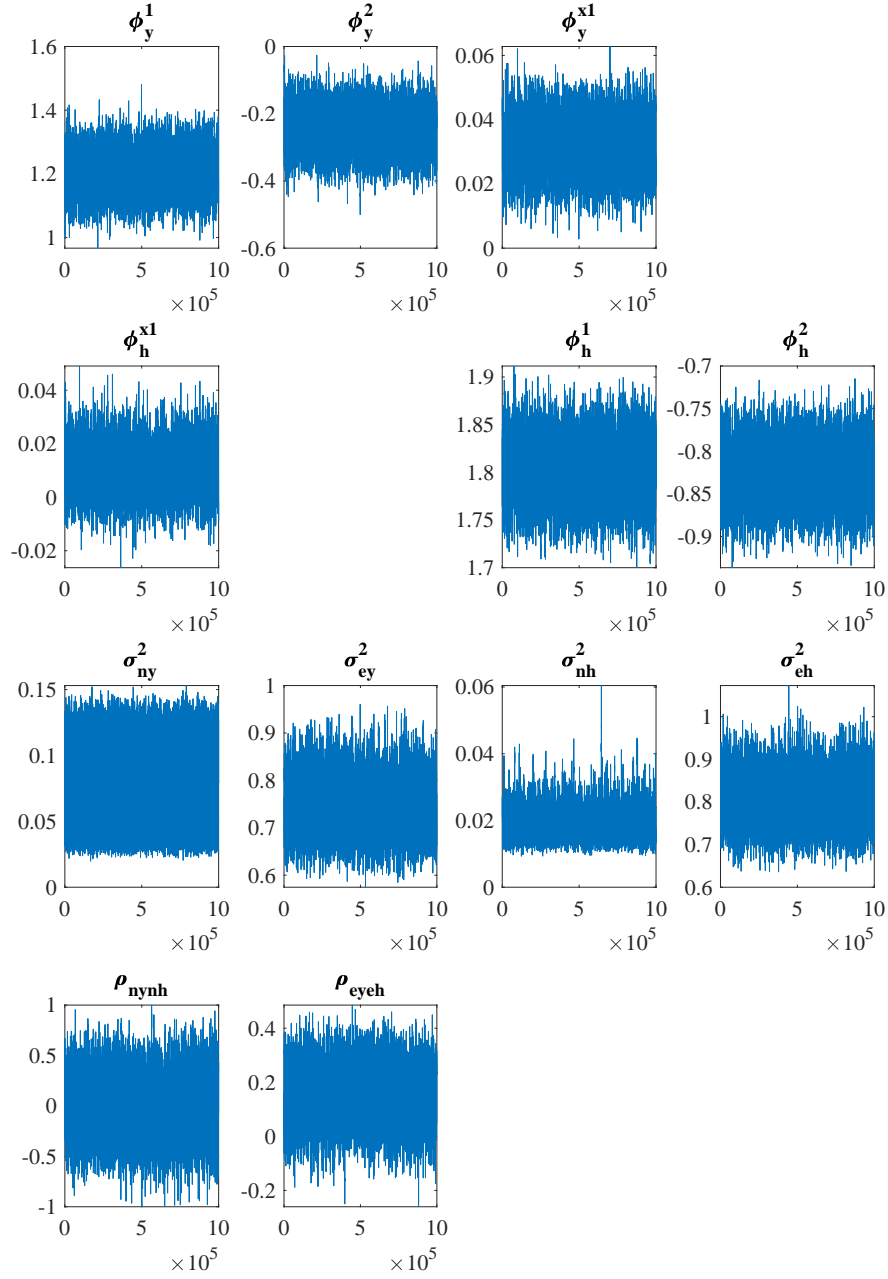


Figure A.11: US VAR(2) 1 cross-lag - Posterior chain

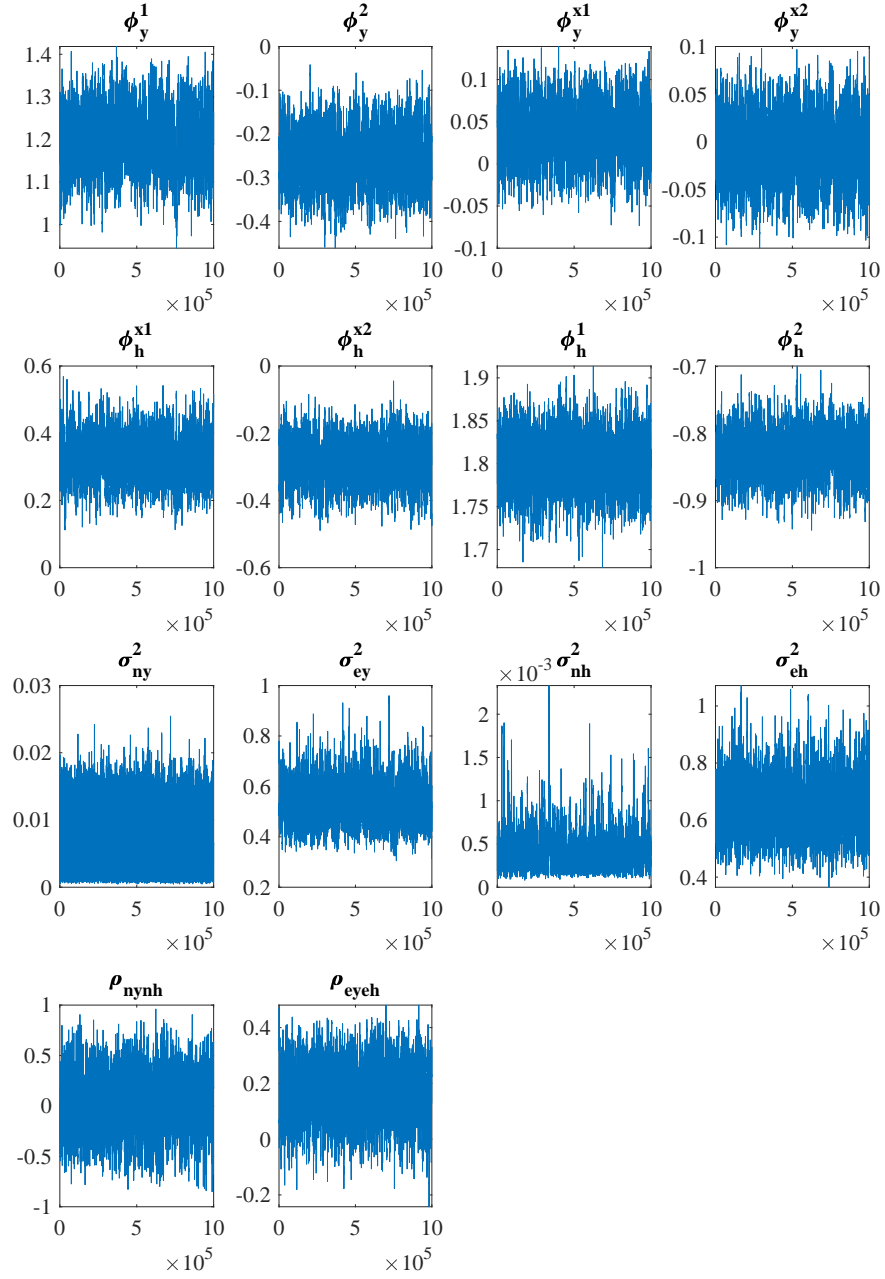


Figure A.12: US VAR(2) 2 cross-lags - Posterior chain

A.2.3 Graphs of other Countries

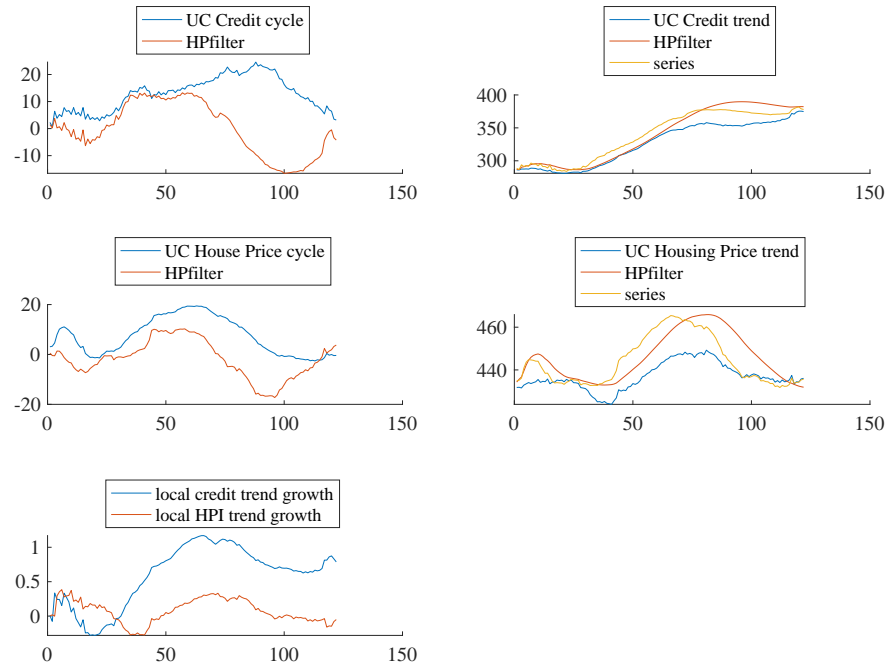


Figure A.13: Italy VAR(2) 1 cross-lag - Unobserved Components

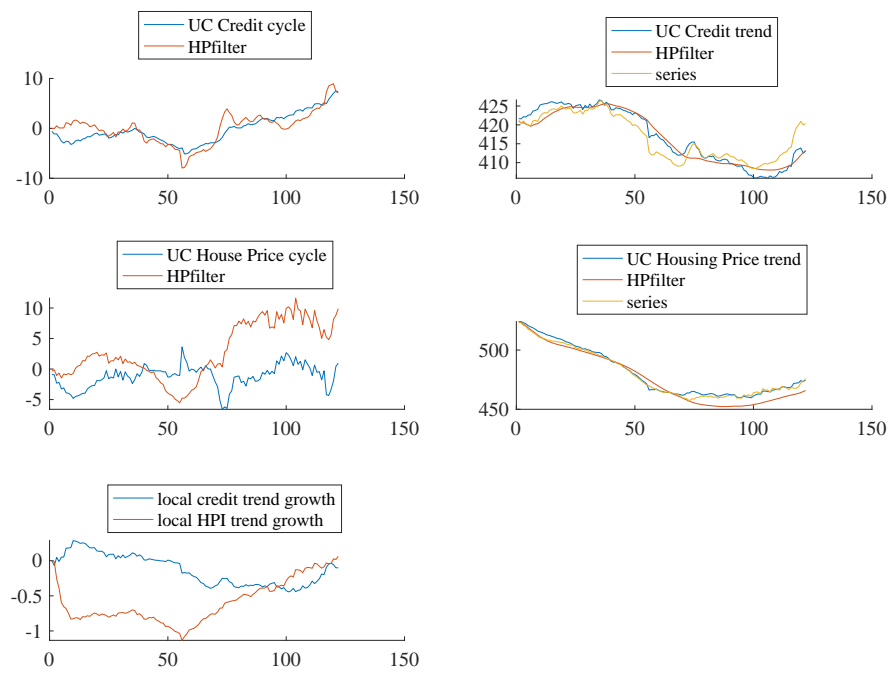


Figure A.14: Japan VAR(2) 1 cross-lag - Unobserved Components

Appendix B

Chapter 3

B.1 Graphs of Credit to GDP Ratios and Credit Gaps

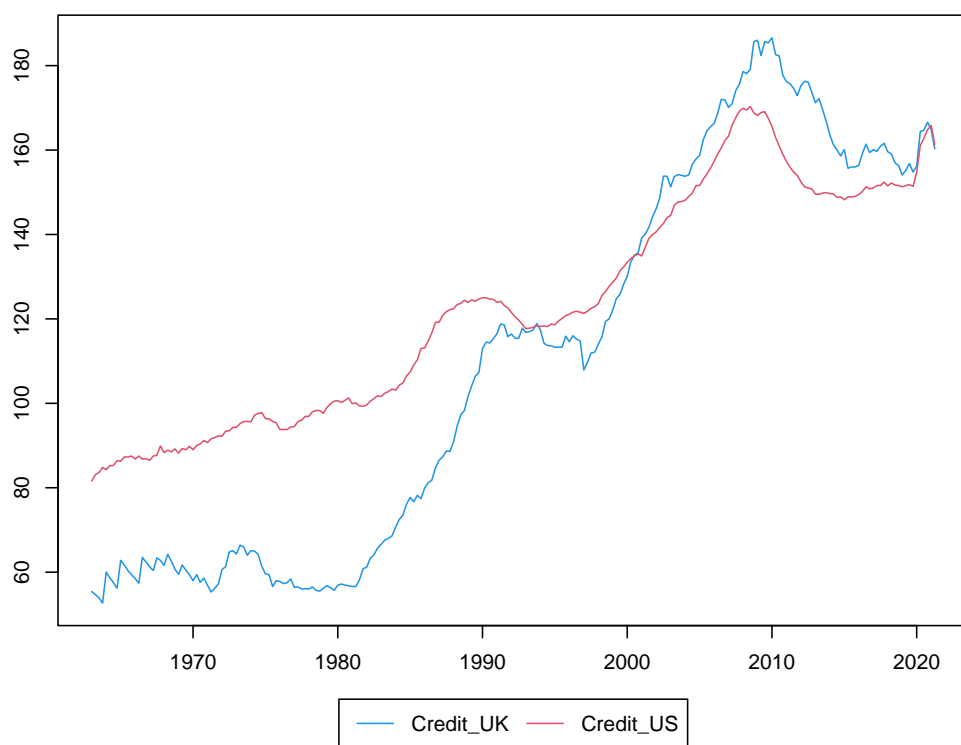


Figure B.1: Credit-to-GDP Ratios

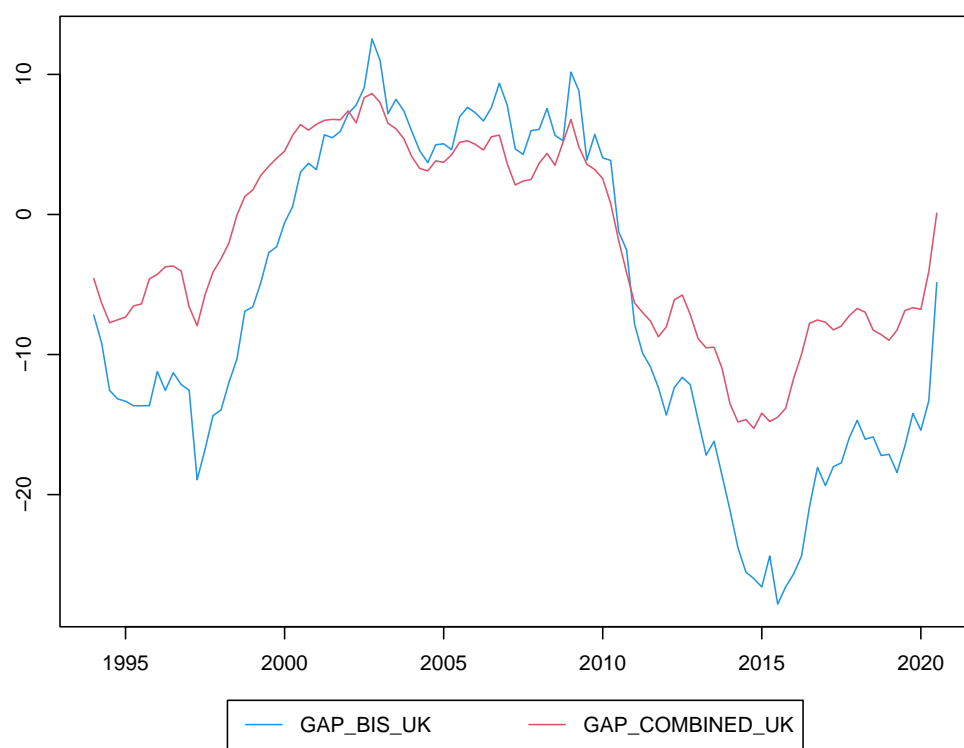


Figure B.2: Credit Gap Comparison (U.K.)

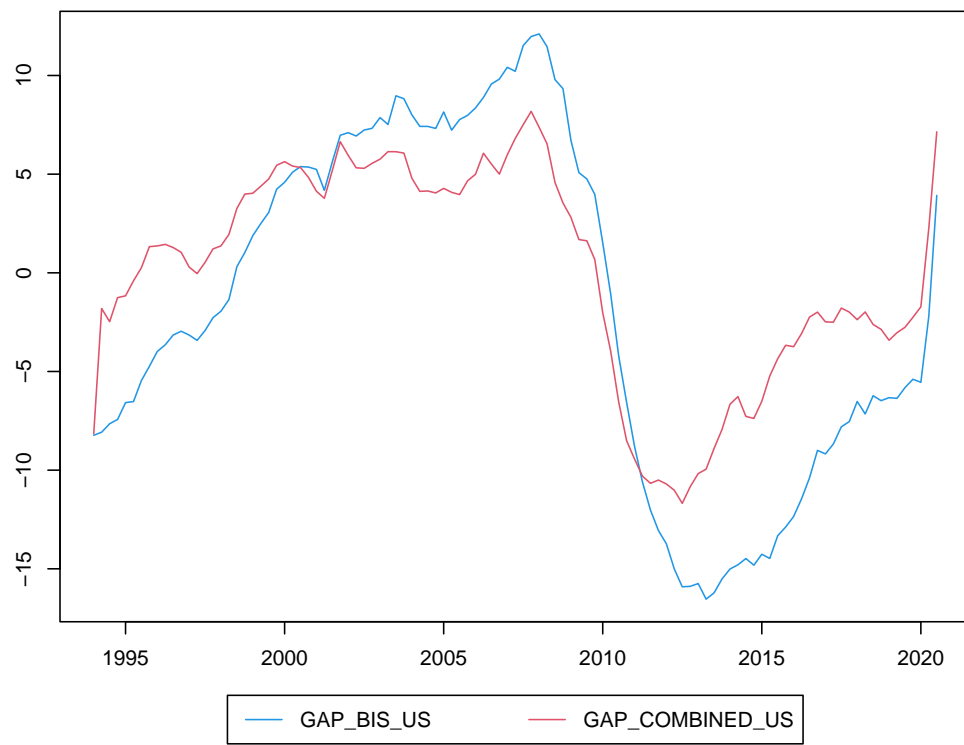


Figure B.3: Credit Gap Comparison (U.S.)

Appendix C

Chapter 4

C.1 List of Decomposition Filters

All filters are in (quasi-real time) one-sided fashion. I store the value of the decomposed cycles for the current period permanently as data becomes available and will not change it when new data comes in the next period.

Hodrick Prescott (HP) filters with different smoothing parameters λ :

- c.hp, c.hp3k, c.hp25k, c.hp125k, c.hp221k, c.hp400k

Hamilton filters with different smoothing parameters θ (distance of past lags):

- c.hamilton13, c.hamilton20, c.hamilton24, c.hamilton28

Linear and polynomial filter models:

- c.linear, c.quad, c.poly3, c.poly4, c.poly5, c.poly6

Beveridge-Nelson decomposition filters with different smoothing parameters (number of lags):

- c.bn2, c.bn3, c.bn4, c.bn5, c.bn6, c.bn7, c.bn8

Structural time series model:

- c.stm

Rolling sample with 15 and 20 years window (of all previous filters):

- c.hp.r15, c.hp3k.r15, c.hp25k.r15, c.hp125k.r15, c.hp221k.r15, c.hp400k.r15,
c.hamilton13.r15, c.hamilton20.r15, c.hamilton24.r15, c.hamilton28.r15, c.linear.r15,
c.quad.r15, c.poly3.r15, c.poly4.r15, c.poly5.r15, c.poly6.r15, c.bn2.r15, c.bn3.r15, c.bn4.r15,
c.bn5.r15, c.bn6.r15, c.bn7.r15, c.bn8.r15, c.stm.r15

- c.hp.r20, c.hp3k.r20, c.hp25k.r20, c.hp125k.r20, c.hp221k.r20, c.hp400k.r20,
c.hamilton13.r20, c.hamilton20.r20, c.hamilton24.r20, c.hamilton28.r20, c.linear.r20,
c.quad.r20, c.poly3.r20, c.poly4.r20, c.poly5.r20, c.poly6.r20, c.bn2.r20, c.bn3.r20, c.bn4.r20,
c.bn5.r20, c.bn6.r20, c.bn7.r20, c.bn8.r20, c.stm.r20

Hamilton filters in panel setting (with rolling samples):

- c.hamilton13.panel, c.hamilton20.panel, c.hamilton24.panel, c.hamilton28.panel,
c.hamilton13.panellr15, c.hamilton20.panellr15, c.hamilton24.panellr15, c.hamilton28.panellr15,
c.hamilton13.panellr20, c.hamilton20.panellr20, c.hamilton24.panellr20, c.hamilton28.panellr20

Moving Average filter:

- c.ma

C.1.1 Statistical Methods for Trend-Cycle Decomposition

Bayesian Structural Time Series Model (STM)

$$y_t = u_t + v_t, v_t \sim N(0, V)$$

$$u_t = u_{t-1} + \beta_{t-1} + w_{1,t}, w_{1,t} \sim N(0, \sigma_{w1}^2)$$

$$\beta_t = \beta_{t-1} + w_{2,t}, w_{2,t} \sim N(0, \sigma_{w2}^2)$$

The implementation of this filter is discussed in Campagnoli et al. (2009). One feature of this filter that allows for smoother trend component is its inclusion of a time-varying local growth rate β_t . Beltran et al. (2021) optimized values of the smoothing parameters at $\sigma_{w1}^2 = 1$, $\sigma_{w2}^2 = 0.01$ and $V = 1100$.

Moving Average

$$MA_t = y_t - \frac{\sum_{i=t-q+1}^t y_i}{q}$$

With q being the smoothing parameter and the length of the moving average window. Beltran et al. (2021) optimized the value for q to be 16.

C.2 Out-of-Sample Forecast Graphical Analysis

C.2.1 Graphs of Selected Countries

In this subsection, I will show graphs of credit gaps for countries that the model deems to be at risk of a future systemic financial crisis.

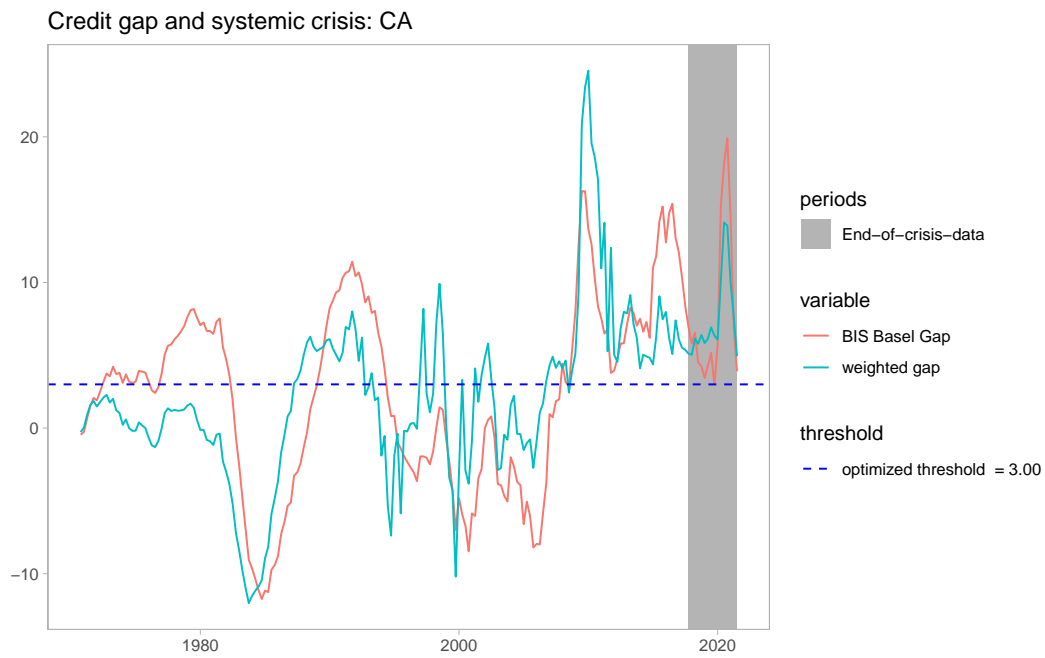


Figure C.1: Canada Graph

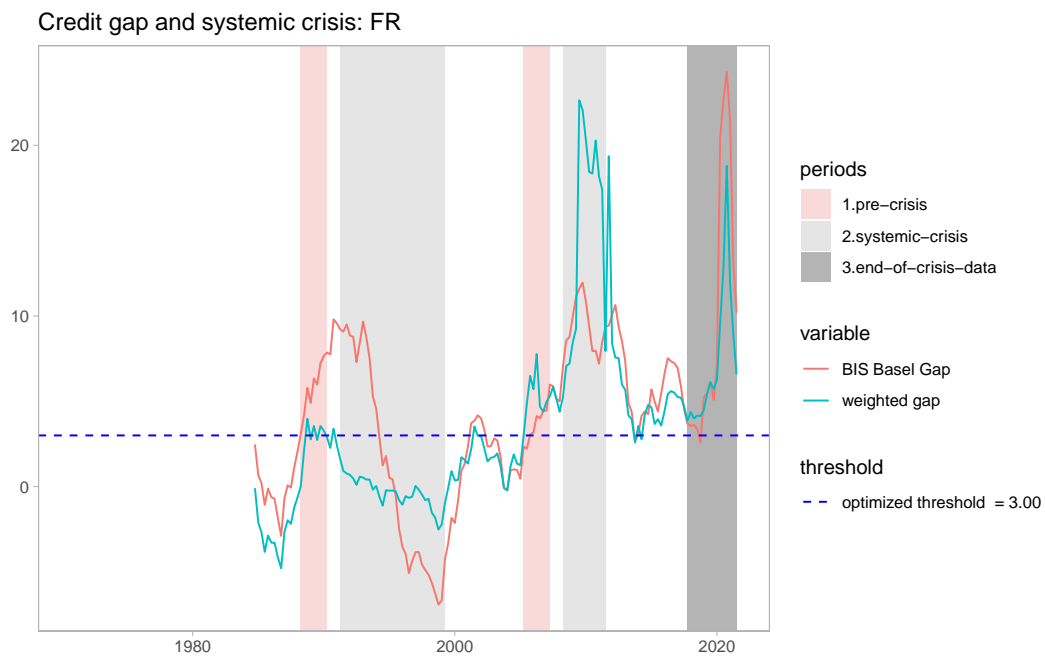


Figure C.2: France Graph

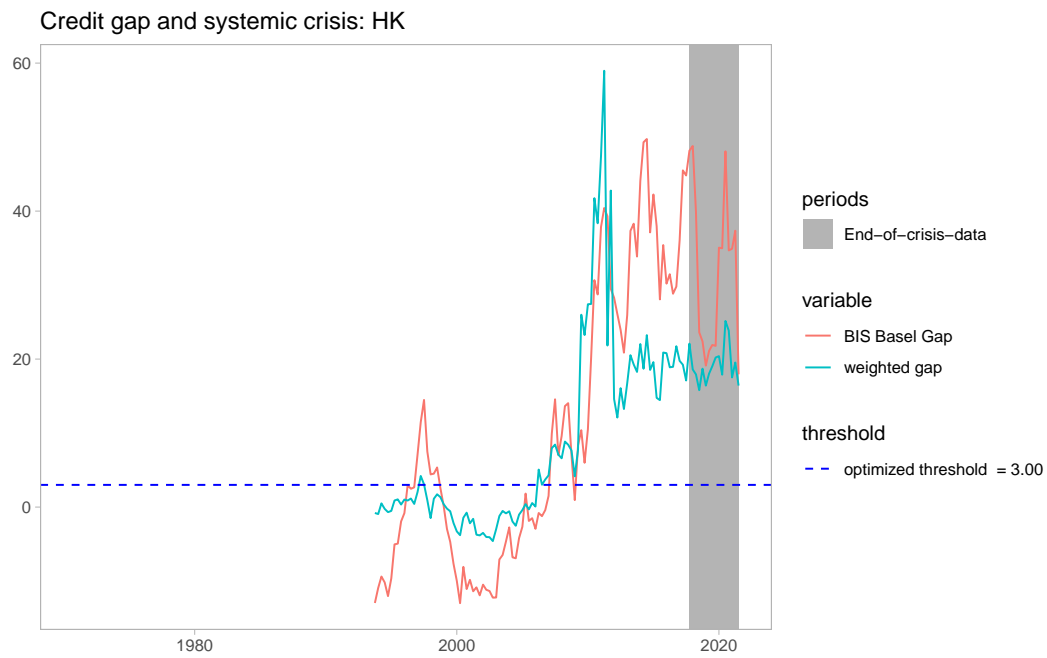


Figure C.3: Hong Kong (SAR) Graph

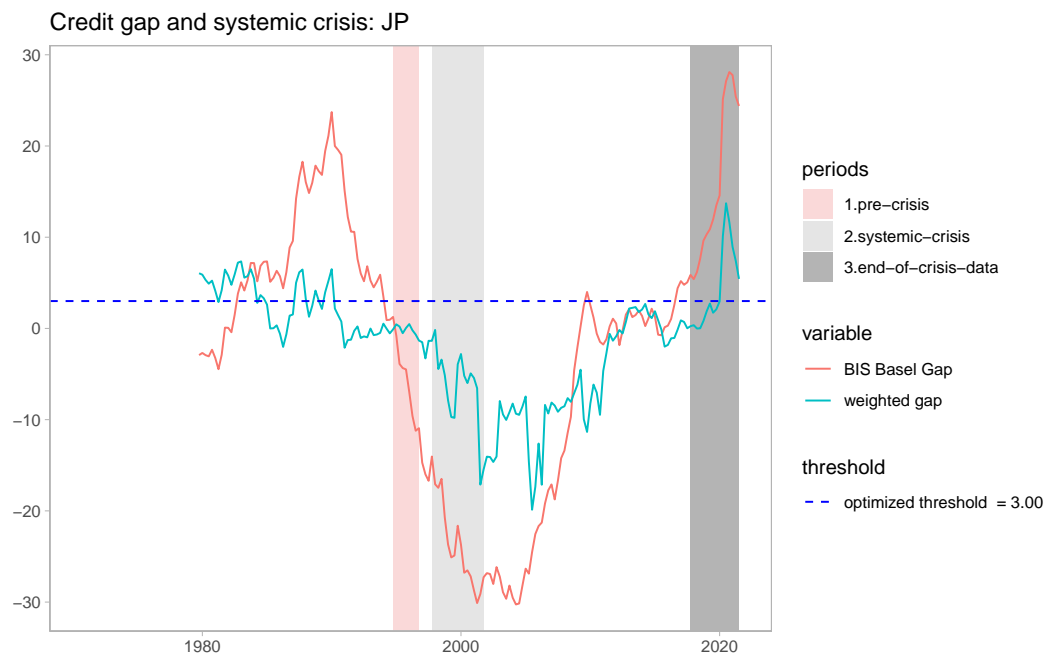


Figure C.4: Japan Graph

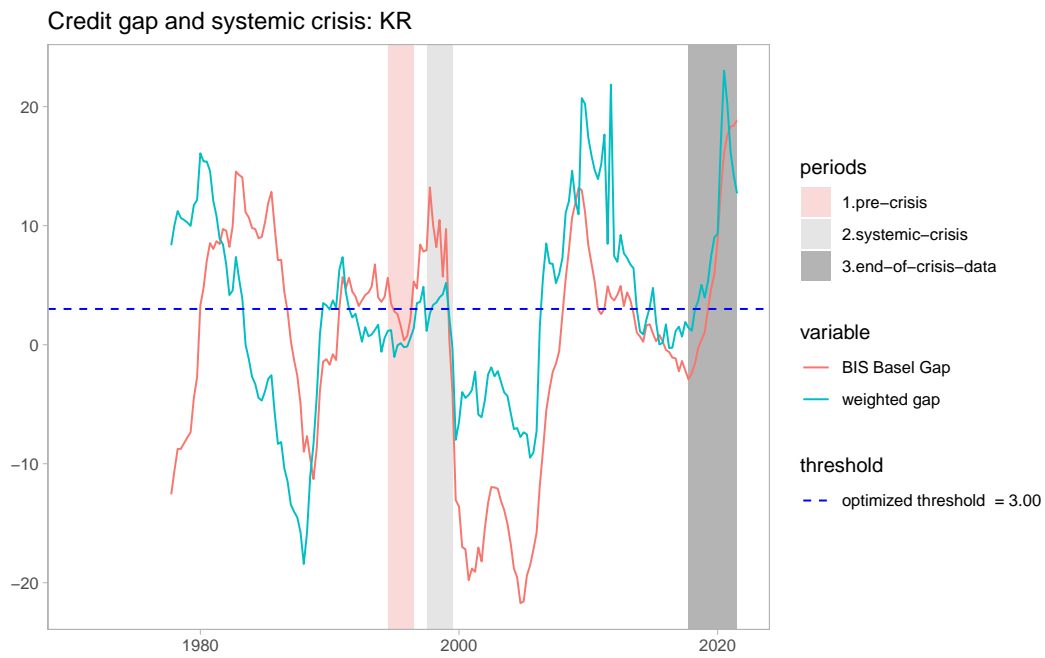


Figure C.5: South Korea Graph

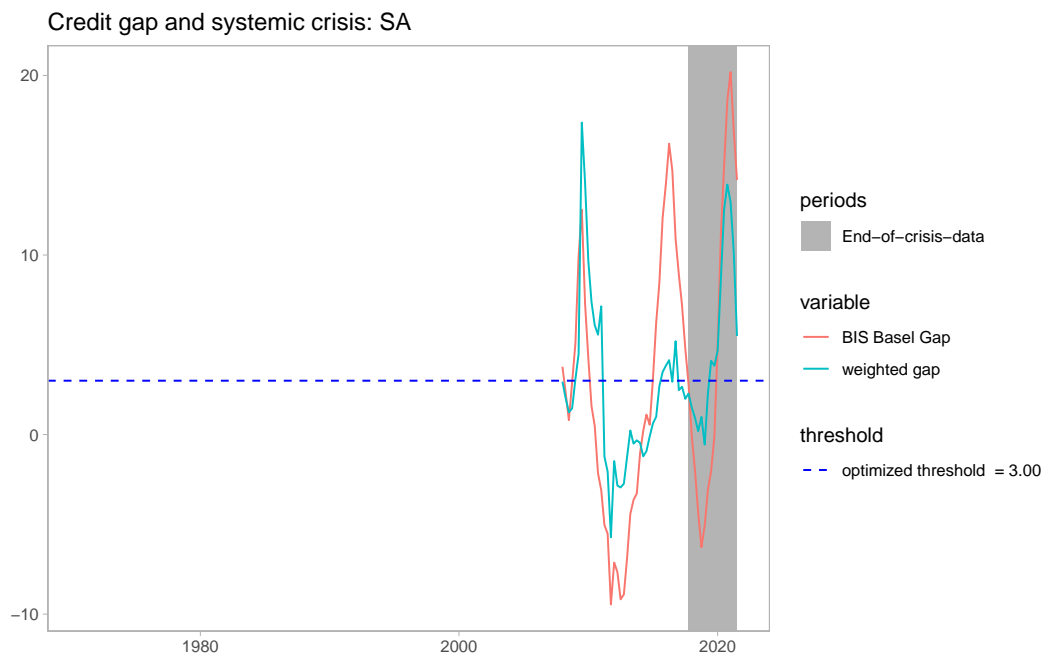


Figure C.6: Saudi Arabia Graph

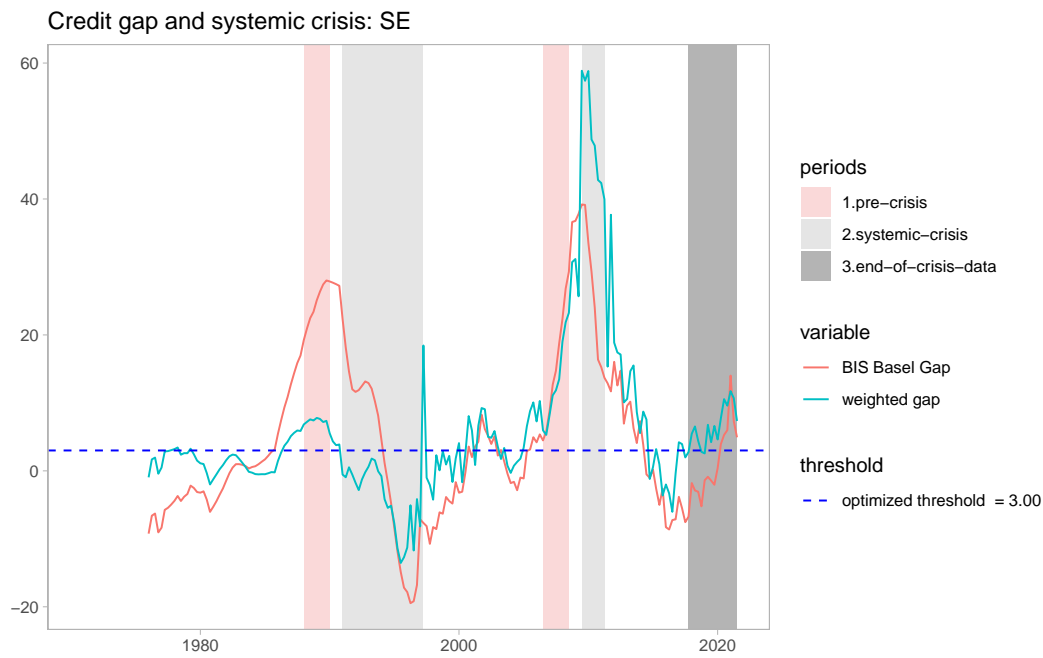


Figure C.7: Sweden Graph

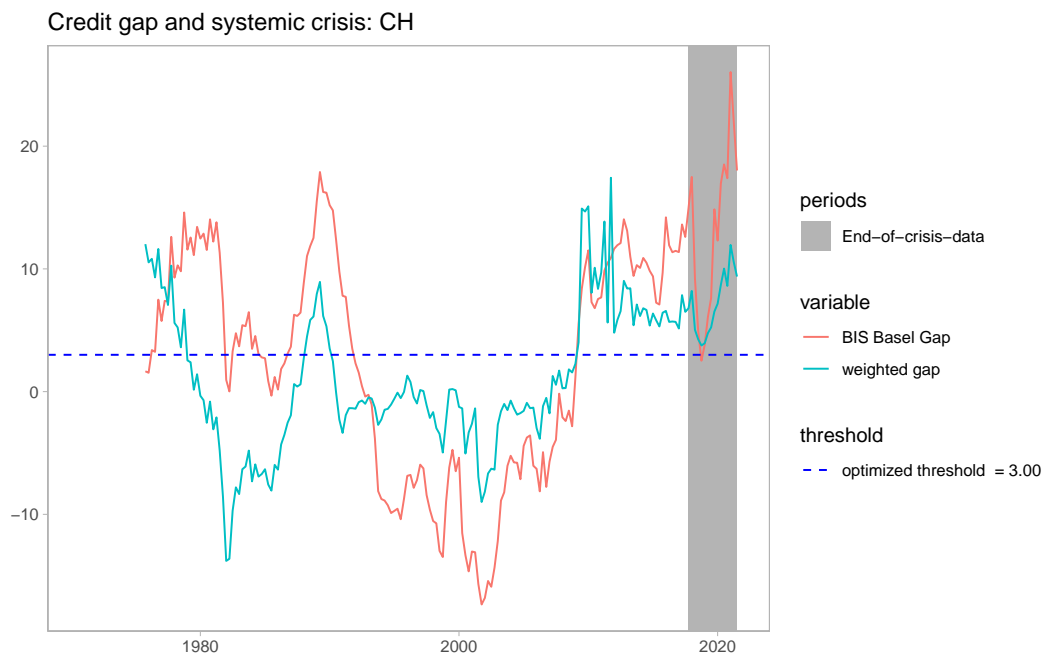


Figure C.8: Switzerland Graph

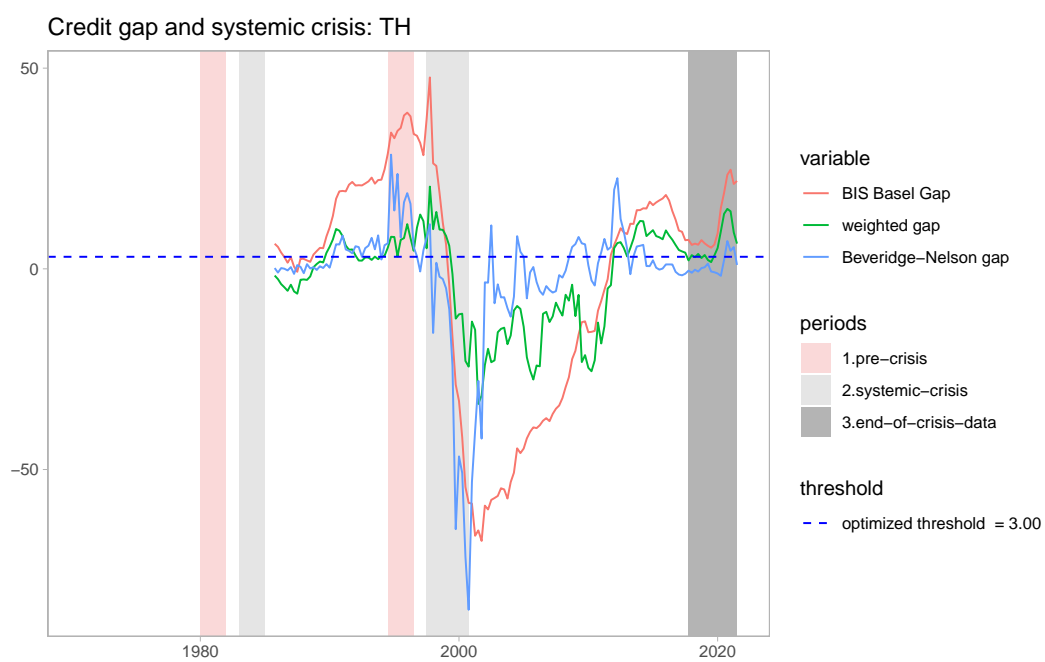


Figure C.9: Thailand Graph