Three Applications of Time-Varying Parameter Models to Macroeconomics

Bowen Fu

Economics Discipline Group

University of Technology Sydney

This dissertation is submitted for the degree of

Doctor of Philosophy

UTS Business School December 2019

Declaration

I would like to dedicate this thesis to my loving parents.

Bowen Fu

December 2019

Certificate of Original Authorship

I, Bowen Fu, declare that this thesis, submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Economics Discipline Group, Business School, at the University of Technology Sydney. This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Signature: Production Note: Signature removed prior to publication.

Date: July 2019

Acknowledgements

I would like to thank my Chair supervisor Joshua Chan for all the help he has given me during my PhD candidacy. Through his direction and guidance I was able to learn and understand what it takes to become a professional researcher. I am also grateful to Joshua in regards to teaching me the fundamentals in regards to Bayesian econometrics and Matlab programming. I would also like to thank my Panel supervisor and coauthor, Mengheng Li who gave me good constructive feedback and contribute to estimating the models of my Chapter 4 paper.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AUROC	Area under the receiver operating characteristic
CC-SV	Constant coefficient model with stochastic volatility
CPI	Consumer price index
EU	European Union
GDP	Gross domestic product
HPD	Highest Posterior Density
LASSO	Least absolute shrinkage and selection operator
MCMC	Markov Chain Monte Carlo
NFE	Non-financial enterprise
SPF	Survey of Professional Forecasters
TVC	Time-varying coefficient model with constant volatility
TVP	Time-varying parameter
TVC-SV	Time-varying coefficient model with stochastic volatility
UIP	Uncovered interest parity
UK	United Kingdom
US	United States
VAR	Vector autoregression

Abstract

This thesis includes chapters that examine the application of time-varying parameter models to three macroeconomic topics: the Phillips curve, early warning system models, and uncovered interest rate parity.

Chapter 2 formally tests for time variation in the slope of the Phillips curve using a variety of measures of inflation expectations and real economic slack. We find that time variation in the slope of the Phillips curve depends on the measure of inflation expectations rather than the measure of real economic slack. We find strong evidence supporting the time-varying slopes of the Phillips curve with different measures of inflation expectations. Thus, we conclude that the slope of the Phillips curve is time varying.

In Chapter 3, we both narrowly and widely replicates the results of Anundsen et al. (2016). Further, we find that allowing for time-varying parameters of early warning system models can considerably improve the in-sample model fit and out-of-sample forecasting performance based on an expanding window forecasting exercise.

In Chapter 4, we consider a time-varying coefficient model with stochastic volatility for the uncovered interest parity regression. We show that jointly estimating time-varying coefficients and stochastic volatility can provide relatively reliable time-varying parameters. Using posterior samples from Bayesian estimation, we determine which United States macroeconomic variables explain the variation in time-varying coefficients and volatility based on least squares with shrinkage. Our empirical study shows that the null hypothesis of uncovered interest parity cannot be unconditionally rejected in the cases of several de-

veloped economies. Further, we show that local breaches of uncovered interest parity are mainly associated with variables from the labour market variables and the output variables in the United States, among other variables.

Chapter 1

Introduction

Many papers highlight the empirical importance of time-varying parameters (TVPs) for modelling financial and macroeconomic variables, including Canova (1993), Cai et al. (2000), Cogley and Sargent (2005), Koop and Potter (2007), Koop and Korobilis (2013), and Chan and Eisenstat (2015). This thesis examines three applications of TVP models in empirical macroeconomics. Chapter 2 studies whether the slope of the Phillips curve is time varying. Chapter 3 investigates whether allowing the time-varying coefficients in the early warning system models can improve these models' forecasting performance. Chapter 4 uses a time-varying coefficients model with stochastic volatility (TVC-SV) to study the uncovered interest rate puzzle.

The original Phillips curve describes the empirical relationship between inflation and unemployment rate (Phillips, 1958). Since many central banks need to maintain both price stability and maximum employment, estimating this relationship is important. In the aftermath of the financial crisis of the 2008-2009, inflation remained stable while there was a surge in the unemployment rate. This is often referred to as the 'missing disinflation' puzzle. One explanation for the puzzle is that the slope of the Phillips curve becomes flatter (e.g., Bean, 2006; Gaiotti, 2008; Ihrig et al., 2010; Kuttner and Robinson, 2010), calling into question of the stability of the Phillips curve. Given these considerations, in Chapter

2, we consider a range of models with an embedded Phillips curve using a variety of measures of inflation expectations and real economic slack. We then test for time variation in the slope of these Phillips curves using the method proposed by Chan (2018). We provide strong evidence in favour of the time-varying slope of the Phillips curve from unobserved components models.

In the early warning system literature, Anundsen et al. (2016) used a dataset covering 20 countries and spanning the period 1976 to 2014 to assess the probability of a financial crisis—specifically, the likelihood that an economy is in a pre-crisis state. They contribute to the early warning system literature by finding that both household credit to the gross domestic product (GDP) gap and the non-financial enterprise (NFE) credit-to-GDP gap affect the likelihood of a financial crisis. In addition, global housing market development can affect domestic financial stability. Finally, their measures of exuberance in housing and credit markets have predictive power of financial stability. Chapter 3 both narrowly and widely replicates the main results of Anundsen et al. (2016). In the narrow replication, using the dataset of Anundsen et al. (2016), we replicate their main results using Matlab. In the wide replication, we replicate their results by using TVP probit models. The wide replication has three main findings. First, the coefficients of financial crisis indicators are quite stable over time, but the time variation in the country dummies is considerable. Second, most of the main findings of Anundsen et al. (2016) are robust to using a TVP approach. The important role of the household credit-to-GDP gap and global housing market developments is very robust to using a TVP approach. The effect of exuberance measures on the probability of a financial crisis also remains robust to using a TVP approach. However, the evidence thatthe NFE credit-to-GDP gap significantly affects the probability of a financial crisis is mixed across different specifications. Third, based on the expanding window forecasting exercise, allowing for TVPs of probit models can considerably improve the in-sample model fit and out-of-sample forecasting performance based on expanding window forecasting exercise.

In Chapter 4, we study the uncovered interest parity relation via a new UIP regression model that extends the model of Fama (1984) by allowing a time-varying coefficient model and stochastic volatility. First, we show that, if we need to estimate the time-varying coefficients or stochastic volatility in a UIP regression model, we need to jointly estimate these two parameters. Compared to alternative models, we show that the TVC-SV can give us sensible weighting schemes on the observations, leading to a reliable inference on the time-varying coefficients and stochastic volatility. Second, different from many other papers that report that the UIP is usually violated, we show the UIP hypothesis unconditionally cannot be rejected for several economies . Our finding does not support the time-varying risk premia explanation for the UIP puzzle. We follow an atheoretical or agnostic approach to find which variables can explain the variation in the slope of the UIP regression model. Third, using least squares with shrinkage, we show that variables associated with the United States (US) labour market play the most important role in explaining the variation in the slope of the UIP regression model.

Chapter 2

Is the Slope of the Phillips Curve Time Varying? Evidence from Unobserved Components Models

2.1 Introduction

The original Phillips curve describes the empirical relationship between inflation and the unemployment rate (Phillips, 1958). Other versions that use related measures of real economic activity are later considered. Estimating this relationship is important for a number of reasons. For example, many central banks need to maintain both price stability and maximum employment. However, these two goals might be not consistent. Therefore, understanding the trade-off between these two goals is important. In addition, in the aftermath of the financial crisis of 2008–2009, inflation remained stable while there was a surge in the unemployment rate. This is often referred to as the 'missing disinflation' puzzle. One explanation for the puzzle is that the slope of the Phillips curve becomes flatter (e.g., Bean, 2006; Gaiotti, 2008; Ihrig et al., 2010; Kuttner and Robinson, 2010), calling into question of the stability of the Phillips curve.

Many papers have documented changes in the slope of the Phillips curve. Examples include Roberts (2006), Atkeson and Ohanian (2001), and Mishkin (2007). To test for time variation in the slope of the Phillips curve, these papers estimate a constant coefficient Phillips curve using split samples to check whether the slope changes considerably across different samples. Rather than model the slope of the Phillips curve as a constant and compare the estimated slope of the Phillips curve in different samples, some studies model the slope as time varying. Examples include Stella and Stock (2012), Chan et al. (2016), and Kim et al. (2014).

However, there are two issues for assuming the slope of the Phillips curve as time varying. First, the conclusion that the slope of the Phillips curve changes is challenged by some recent studies. For example, Gordon (2013) finds that the slope of the Phillips curve is stable by estimating a model with a hybrid Phillips curve. Coibion and Gorodnichenko (2015) estimated many models with a standard expectation-augmented Phillips curve using a variety of measures of inflation expectations and find mixed evidence regarding changes in the slope of the Phillips curve. Second, the TVP specification might lead to over-parameterisation compared to the constant coefficient specification as pointed out by Chan et al. (2012), Nakajima and West (2013), and Belmonte et al. (2014). Therefore, one should be cautious about modelling the slope of the Phillips curve as time varying without testing whether this specification is relevant.

Given these considerations, we consider a range of models with an embedded Phillips curve using a variety of measures of inflation expectations and real economic slack. We then test for time variation in the slope of these Phillips curves using the method proposed by Chan (2018). We find strong evidence in favour of the time-varying slope of the Phillips curve from unobserved components models.

Formal tests of time variation in the slope of the Phillips curve have been recently implemented by Berger et al. (2016) and Karlsson et al. (2018). Karlsson et al. (2018) test for time variation within the framework of a TVP Bayesian vector autoregression using new tools for model selection proposed by Chan and Eisenstat (2018). By comparing a bivariate vector autoregression with constant coefficients with a time-varying VAR, Karlsson et al. (2018) find strong evidence in favour of the latter and conclude that the slope of the Phillips curve is unstable. Instead of jointly testing time variation in all the parameters, our approach is more specific and tests only if the slope coefficient of the Phillips curve is time-varying.

This chapter is most related to Berger et al. (2016). They estimate a model with a New Keynesian Phillips curve in which the trend inflation is interpreted as long-run inflation expectations. They then test for time variation in the slope of the Phillips curves using a stochastic model specification search, an approach proposed by in Frühwirth-Schnatter and Wagner (2010). Berger et al. (2016) find that the time-varying slope specification is rejected by the stochastic model specification search and conclude that the slope of the Phillips curve is not time varying.

This chapter is different from Berger et al. (2016) in three aspects. First, we consider a wider range of measures of inflation expectations and economic slack. In particular, we consider trend inflation as a measure of inflation expectations and also consider survey-based inflation expectations and a variety of measures of real economic slack. Second, we directly compute the Bayes factor in favour of the model with a time-varying Phillips curve via the method proposed by Chan (2018) rather than the stochastic model specification search reported by Berger et al. (2016). Finally, unlike Berger et al. (2016), we find strong evidence supporting the time varying slope of the Phillips curve.

The remainder of this chapter is organised as the follows. In Section 2.2, we describe how we test the time variation in the slope of the Phillips curve. In Section 2.3, we describe the models with different specifications for the Phillips curve. In Section 2.4, we discuss the estimation procedures for the models in Section 2.3. Section 2.5 describes the results of the test for the time variation of the slope of the Phillips curve. In Section 2.6, we conclude that the slope of the Phillips curve is time varying.

2.2 Testing for Time Variation

In this section, we outline the methodology to test for time-variation. We first provide an overview of the Bayes factor and Savage-Dickey density ratio and then introduce a new method, as proposed by Chan (2018), of calculating the Bayes factor.

2.2.1 Bayes Factor and Savage-Dickey Density Ratio

To demonstrate the method of testing for time variation in the slope of the Phillips curve, we first consider the following unobserved components model with stochastic volatility:

$$\pi_t - \mathrm{E}_t\pi_{t+1} = \lambda_t x_t + \varepsilon_t^{\pi}, \qquad \varepsilon_t^{\pi} \sim \mathcal{N}(0, e^{h_t}), \tag{2.1}$$

$$\lambda_t = \lambda_{t-1} + \varepsilon_t^{\lambda}, \qquad \varepsilon_t^{\lambda} \sim \mathcal{N}(0, \omega_{\lambda}^2), \tag{2.2}$$

where π_t is the inflation rate at time t, $\mathrm{E}_t\pi_{t+1}$ is a measure of expected inflation at time $t+1$ given the information at time t, x_t is a measure of real economic slack and λ_t is the slope of the Phillips curve. We model the slope, λ_t, as a random walk process instead of a stationary AR(1).[1] To test whether the slope, λ_t, is time-varying, we can compare the model (2.1)–(2.2) to a restricted version in which the slope is constant, that is, $\omega_{\lambda}^2 = 0$. Denote the former model as Model 1 and the restricted version as Model 2. One popular model comparison criterion for comparing these two models is the Bayes factor that favours

[1]Eisenstat and Strachan (2016) argue that the random walk assumption has two main advantages for macroeconomic applications. First, the random walk specification can be a parsimonious approximation to a stationary specification with high persistence. Second, the random walk specification implies greater smoothness than the stationary model with low persistence.

Model 1 against Model 2, defined as

$$\mathrm{BF}_{12} = \frac{p(\mathbf{y} \mid \text{Model 1})}{p(\mathbf{y} \mid \text{Model 2})},$$

where $p(\mathbf{y} \mid \text{Model i})$ is the marginal likelihood for Model_i. The corresponding posterior odds ratio is defined as

$$\frac{p(\text{Model 1} \mid \mathbf{y})}{p(\text{Model 2} \mid \mathbf{y})} = \frac{p(\text{Model 1})}{p(\text{Model 2})} \times \mathrm{BF}_{12}.$$

Assume that the prior model probabilities are equal, i.e., $p(\text{Model 1}) = p(\text{Model 2})$, and the posterior odds ratio favouring Model 1 reduces to the Bayes factor BF_{12}. For example, $\mathrm{BF}_{12} = 10$ means that model Model 1 is 10 times more likely than model Model 2 given the data.

The Bayes factor is commonly used to compare models. However, the main challenge here is that it is often difficult to compute the marginal likelihood of models with time-varying parameters.

Fortunately, one simpler method is available when we need to compute the Bayes factor for nested models. Specifically, the Bayes factor can be calculated by using the Savage-Dickey density ratio (Verdinelli and Wasserman, 1995). This approach requires only the estimation of the unrestricted model. For example, the Bayes factor favouring Model 1 against Model 2 can be obtained using the Savage-Dickey density ratio as

$$\mathrm{BF}_{12} = \frac{p(\omega_\lambda^2 = 0)}{p(\omega_\lambda^2 = 0 \mid \mathbf{y})},$$

where the numerator is the marginal prior density of ω_λ^2 evaluated at 0, and the denominator is the marginal posterior of ω_λ^2 evaluated at 0. Intuitively, if ω_λ^2 is more likely to be 0 under the prior density relative to the posterior density, this can be viewed as evidence supporting the time-varying slope of the Phillips curve. However, this method cannot be directly applied in our setting due to two related issues. First, the value 0 is at the boundary of the parameter space of ω_λ^2. Therefore, the Savage-Dickey density ratio approach is not applicable. Second, ω_λ^2 is often assumed to have an inverse-gamma prior, which has zero density at zero. To deal with these two difficulties, we follow the method proposed by Chan (2018). Specifically, we use the so-called non-centred parameterisation discussed in Frühwirth-Schnatter and Wagner (2010), working with the unsigned standard deviation, ω_λ, that supports the whole real line. Then we directly calculate the relevant Bayes factor using the Savage-Dickey density ratio.

2.2.2 Non-centred Parameterisation

Next, we briefly discuss the non-centred parameterisation. First, we define $\lambda_t = \lambda_0 + \omega_\lambda \widetilde{\lambda}_t$, then, the state space model in (1)-(2) can be written as follows:

$$\pi_t - \mathrm{E}_t \pi_{t+1} = (\lambda_0 + \omega_\lambda \widetilde{\lambda}_t) x_t + \varepsilon_t^\pi, \qquad \varepsilon_t^\pi \sim \mathcal{N}(0, e^{h_t}), \tag{2.3}$$

$$\widetilde{\lambda}_t = \widetilde{\lambda}_{t-1} + \varepsilon_t^{\widetilde{\lambda}}, \qquad \varepsilon_t^{\widetilde{\lambda}} \sim \mathcal{N}(0,1), \tag{2.4}$$

where $\widetilde{\lambda}_0 = 0$.

In this model, we assume $\omega_\lambda \sim \mathcal{N}(0, V_{\omega_\lambda})$, which has two main advantages. First, by a change of variable (Kroese and Chan, 2014), the implied prior for ω_λ^2 is $\mathscr{G}(\frac{1}{2}, \frac{1}{2V_{\omega_\lambda}})$. This gamma prior has more mass concentrated around small values of ω_λ^2. Therefore, it provides shrinkage—*a priori* it favours the more parsimonious constant—coefficient model. Second, it is a conjugate prior for ω_λ, under the non-centred parameterisation, facilitating computation. The sign of ω_λ is not identified, but alteration of the sign dose not change the likelihood value. After the non-centered parameterisation of model (2.1)–(2.2), the Bayes factor $\mathrm{BF}_{12} = p(\omega_\lambda = 0)/p(\omega_\lambda = 0 \mid \mathbf{y})$, obtained by using Savage-Dickey density ratio, can be directly calculated using the method proposed by Chan (2018).

2.3 Specifications for the Phillips Curve

We consider two classes of models for modelling the Phillips curve: the univariate unobserved components models with stochastic volatility and the bivariate unobserved components models with stochastic volatility.[2] For each model, we need a measure of both inflation expectations and economic slack. For the univariate model, the unobserved component of real economic activities is the trend of real economic activities, z_t, denoted as z_t^*. Then, we use the deviation from the trend, $x_t = z_t - z_t^*$, as a measure of economic slack. We will use observable measures for the inflation expectations, $\mathrm{E}_t\pi_{t+1}$, such as the average of the past four quarters inflation or Survey of Professional Forecasters (SPF) inflation expectations.

A rapidly growing literature highlights that trend inflation has important implications for the specification of the New Keynesian Phillips curve (e.g., Ascari, 2004; Cogley and Sbordone, 2008; Kozicki and Tinsley). Thus, we also consider bivariate unobserved components models to jointly model real economic activities and inflation. In the bivariate case, the additional unobserved component is the trend inflation, denoted as τ_t. In the spirit of Beveridge and Nelson (1981), τ_t can be interpreted as the long-run inflation expectations. The estimated trend inflation usually has substantial variance. To reduce the variance in the estimated trend inflation, Chan et al. (2018) estimate trend inflation by linking Blue Chip 10 years inflation forecasts to trend inflation. With additional information from the Blue Chip 10 years inflation forecasts, the variance of τ_t decreases substantially. Thus, in addition, we also consider the models with Phillips curves linking Blue Chip 10 years inflation forecasts, q_t, to trend inflation, τ_t.

Altogether, we will estimate eight models from these two classes of models, using Blue Chip 10 years inflation forecasts, q_t, different measures of real economic slack, x_t, and

[2] In this paper,‘ univariate’ and ‘bivariate’ refer to the numbers of unobserved components.

different measures of inflation expectations, $\mathrm{E}_t\pi_{t+1}$. We will provide the details of the univariate and bivariate unobserved components models in Section 2.3.1 and Section 2.3.2.

2.3.1 Univariate Unobserved Components Model with Stochastic Volatility

Let π_t and z_t denote the inflation rate and level of economic activities, respectively. Further, let z_t^* denote the trend of real activities. Then $x_t = z_t - z_t^*$ is a measure of the economic slack, such as the unemployment gap or output gap. Considering the following class of univariate unobserved components models with stochastic volatility:

$$\pi_t = \mathrm{E}_t\pi_{t+1} + \lambda_t(z_t - z_t^*) + \varepsilon_t^{\pi}, \qquad \varepsilon_t^{\pi} \sim \mathcal{N}(0, e^{h_t}), \tag{2.5}$$

$$\lambda_t = \lambda_{t-1} + \varepsilon_t^{\lambda}, \qquad \varepsilon_t^{\lambda} \sim \mathcal{N}(0, \omega_\lambda^2), \tag{2.6}$$

$$z_t = z_t^* + e_t, \tag{2.7}$$

$$e_t = \phi_1 e_{t-1} + \phi_2 e_{t-2} + \varepsilon_t^{e}, \qquad \varepsilon_t^{e} \sim \mathcal{N}(0, \omega_e^2), \tag{2.8}$$

where λ_t is the slope of the Phillips curve, $\mathrm{E}_t\pi_{t+1}$ represents different measures for expectations of inflation. λ_t and τ_t are modelled as random walk, and e_t follows an AR(2) process. We consider two different specifications for z_t^*. First, when z_t represents unemployment rate, z_t^* is modelled as a random walk:

$$z_t^* = z_{t-1}^* + \varepsilon_t^{z_t^*}, \qquad \varepsilon_t^{z_t^*} \sim \mathcal{N}(0, \omega_{z^*}^2). \tag{2.9}$$

Second, when z_t represents output level, the growth of z_t, Δz_t^*, is modelled as a random walk:

$$\Delta z_t^* = \Delta z_{t-1}^* + \varepsilon_t^{z_t^*}, \qquad \varepsilon_t^{z_t^*} \sim \mathcal{N}(0, \omega_{z^*}^2). \tag{2.10}$$

2.3.2 Bivariate Unobserved Components Model with Stochastic Volatility

Next, we augment the univariate unobserved components models to model trend inflation. Specifically, the class of bivariate unobserved components models with stochastic volatility can be denoted as

$$\pi_t - \tau_t = \lambda_t(z_t - z_t^*) + \varepsilon_t^\pi, \qquad \varepsilon_t^\pi \sim \mathcal{N}(0, e^{h_t}), \tag{2.11}$$

$$\lambda_t = \lambda_{t-1} + \varepsilon_t^\lambda, \qquad \varepsilon_t^\lambda \sim \mathcal{N}(0, \omega_\lambda^2), \tag{2.12}$$

$$\tau_t = \tau_{t-1} + \varepsilon_t^\tau, \qquad \varepsilon_t^\tau \sim \mathcal{N}(0, e^{g_t}), \tag{2.13}$$

$$z_t = z_t^* + e_t, \tag{2.14}$$

$$e_t = \phi_1 e_{t-1} + \phi_2 e_{t-2} + \varepsilon_t^e, \qquad \varepsilon_t^e \sim \mathcal{N}(0, \omega_e^2), \tag{2.15}$$

As before, either z_t^* or Δz_t^* is modelled as a random walk:

$$z_t^* = z_{t-1}^* + \varepsilon_t^{z_t^*}, \qquad \varepsilon_t^{z_t^*} \sim \mathcal{N}(0, \omega_{z^*}^2), \tag{2.16}$$

$$\Delta z_t^* = \Delta z_{t-1}^* + \varepsilon_t^{z_t^*}, \qquad \varepsilon_t^{z_t^*} \sim \mathcal{N}(0, \omega_{z^*}^2). \tag{2.17}$$

In addition, we also link the trend inflation, τ_t, to the Blue Chip inflation forecasts by adding the following equation:

$$q_t = d_0 + d_1\tau_t + \varepsilon_t^q, \qquad \varepsilon_t^q \sim \mathcal{N}(0, \omega_q^2), \tag{2.18}$$

where q_t is Blue Chip 10 years forecasts. Following Chan et al. (2018), we allow the possibility that the forecasts are unrelated to the trend inflation by introducing the intercept, d_0, and slope coefficient, d_1. When $d_0 = 0$ and $d_1 = 1$, the Blue Chip forecasts are an unbiased measure of the trend inflation.

2.3.3 Specific Models

The following is a brief summary of all the models we consider. Details of these eight models are provided in Section 2.7.

- M1—univariate unobserved components model of the Phillips curve with the unemployment gap, $u_t - \nu_t$, and the backward-looking inflation expectations, $\pi^e_{t|t-1}$, measured as the average of past four quarter inflation, where $z_t = u_t$, $z^*_t = \nu_t$, and then $x_t = u_t - \nu_t$, and $\mathrm{E}_t\pi_{t+1} = \pi^e_{t|t-1}$.

- M2—univariate unobserved components model of the Phillips curve with the unemployment gap, $u_t - \nu_t$, and the forward-looking inflation expectations, $\pi^e_{t+1|t}$, measured as SPF one year inflation forecasts, where $z_t = u_t$, $z^*_t = \nu_t$, then $x_t = u_t - \nu_t$, and $\mathrm{E}_t\pi_{t+1} = \pi^e_{t+1|t}$.

- M3—univariate unobserved components model of the Phillips curve with the output gap, $y_t - y^*_t$, and the backward-looking inflation expectations, $\pi^e_{t|t-1}$, measured as the average of past four quarter inflation, where $z_t = y_t$, $z^*_t = y^*_t$, then $x_t = y_t - y^*_t$, and $\mathrm{E}_t\pi_{t+1} = \pi^e_{t|t-1}$.

- M4—univariate unobserved components model of the Phillips curve with the output gap, $y_t - y^*_t$ and the forward-looking inflation expectations, $\pi^e_{t+1|t}$, measured as SPF one year inflation forecasts,where $z_t = y_t$, $z^*_t = y^*_t$, then $x_t = y_t$-y^*_t, and $\mathrm{E}_t\pi_{t+1} = \pi^e_{t+1|t}$.

- M5—bivariate unobserved components model of the Phillips curve with the unemployment gap, $u_t - \nu_t$, and the trend inflation, τ_t, where $z_t = u_t$, $z^*_t = \nu_t$, and then $x_t = u_t - \nu_t$.

- M6—bivariate unobserved components model of the Phillips curve with the output gap, $y_t - y_t^*$, and the trend inflation, τ_t, where $z_t = y_t$, $z_t^* = y_t^*$, and then $x_t = y_t - y_t^*$.
- M7—bivariate unobserved components model of the Phillips curve with the unemployment gap, $u_t - \nu_t$, and the trend inflation, τ_t, where $z_t = u_t$, $z_t^* = \nu_t$, and then $x_t = u_t - \nu_t$. τ_t is estimated with the additional information from Blue Chip 10 years inflation forecasts, q_t.
- M8—bivariate unobserved components model of the Phillips curve with the output gap, $y_t - y_t^*$, and the trend inflation, τ_t, ,where $z_t = y_t$, $z_t^* = y_t^*$, and then $x_t = y_t - y_t^*$. τ_t is estimated with the additional information of Blue Chip 10 years inflation forecasts, q_t.

2.4 Estimation

In this section, we discuss the estimation procedures for the class of bivariate unobserved components models with stochastic volatility, specified in (11)–(17). The class of univariate unobserved components model with stochastic volatility, specified in (5)–(10) can be estimated similarly. Estimation details for these eight models with different measures for inflation expectations and real economic slack are provided in Section 2.7.

2.4.1 Prior

We assume $\omega_\lambda \sim \mathcal{N}(0, V_{\omega_\lambda})$. This assumption, discussed in Section 2.2.2, can provide shrinkage—*a priori* it favours the more parsimonious constant-coefficient model. Under the non-centred parameterisation, this assumption can also facilitate computation. We set V_{ω_λ}=0.25^2 so that the implied prior means of ω_λ^2 are $\mathrm{E}\omega_\lambda^2 = 0.25^2$. The priors for other parameters will be discussed in detail in Section 2.7.

2.4.2 Likelihood

In this Section, we derive the densities of $\pi = (\pi_1, \ldots, \pi_T)'$ and $z = (z_1, \ldots, z_T)'$ that will be used to construct the posterior sampler. Let

$$\Lambda_\lambda = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_T)$$

and

$$\lambda_t = \lambda_0 + \omega_\lambda \widetilde{\lambda}_t.$$

Then,

$$\Lambda_\lambda = \mathrm{diag}(\lambda_0 + \omega_\lambda \widetilde{\lambda}_1, \lambda_0 + \omega_\lambda \widetilde{\lambda}_2, \lambda_0 + \omega_\lambda \widetilde{\lambda}_3, \ldots, \lambda_0 + \omega_\lambda \widetilde{\lambda}_T).$$

Stack (11) over t, and, we have

$$\pi - \tau - \Lambda_\lambda (z - z^*) = \varepsilon^\pi.$$

Then, the log conditional density of π is

$$\log p(\pi \mid \tau, z, z^*, e, \widetilde{\lambda}, \lambda_0, \omega_h, \omega_\lambda, \omega_{z^*}, \omega_e, h, g) \propto -\frac{1}{2}(\pi - \tau - \Lambda_\lambda (z - z^*))' S_\pi^{-1} (\pi - \tau - \Lambda_\lambda (z - z^*)),$$

where

$$S_\pi = \mathrm{diag}(e^{h_1}, e^{h_2}, e^{h_3}, \ldots, e^{h_T}).$$

Let

$$H_\phi = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ -\phi_1 & 1 & 0 & 0 & \dots & 0 \\ -\phi_1 & -\phi_2 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & \dots & \dots & -\phi_1 & -\phi_2 & 1 \end{bmatrix}.$$

Stack (15) over t. Then, we have

$$H_\phi e = \varepsilon^e.$$

Finally, the log conditional density of z is

$$\log p(z \mid z^*, e, \omega_{z^*}, \omega_e, \phi) \propto -\frac{T}{2}\log\omega_e^2 - \frac{1}{2\omega_e^2}(z - z^*)' H_\phi' H_\phi (z - z^*).$$

2.4.3 Posterior Sampler

The bivariate unobserved components model is estimated using Markov Chain Monte Carlo (MCMC) methods. Specifically, given the priors and the likelihood derived in Section 2.4.2, posterior draws can be obtained by sequentially sampling from the following seven conditional densities:

1. $p(\tau \mid \pi, z, z^*, \widetilde{\lambda}, \lambda_0, \omega_\lambda, g, h, e, \phi, \omega_g^2, \omega_{z^*}^2, \omega_e^2)$;

2. $p(z^* \mid \pi, \tau, \widetilde{\lambda}, \lambda_0, \omega_\lambda, z, e, \phi, \omega_{z^*}, \omega_e^2)$;

3. $p(h, g \mid \pi, z, z^*, \tau, \widetilde{\lambda}, \lambda_0, \omega_\lambda)$;

4. $p(\phi \mid z, z^*, \omega_{z^*}, \omega_e^2)$;

5. $p(\widetilde{\lambda} \mid \pi, z, z^*, \tau, h, \lambda_0, \omega_\lambda)$;

6. $p(\lambda_0, \omega_\lambda \mid \pi, z, z^*, \widetilde{\lambda}, h)$;

7. $p(\omega_e^2 \mid z, z^*, \phi)$.

The estimation is standard and details are provided in Section 2.7. We also provide in Section 2.7 the estimation details of all other models.

2.5 Results

In this section, we first estimate six different models embedded with the Phillips curve, M1–M6, and test the time variation in the slopes. In addition, we estimate two additional models, M7 and M8, that use additional information from the Blue Chip 10 years inflation forecasts. We then test the time variation in the slopes of the Phillips curve under these two models.

Our data consist of quarterly consumer price index (CPI) inflation rates, (civilian seasonally adjusted) unemployment rates, and real gross domestic product (GDP) from 1948Q1 to 2013Q1, SPF one- year inflation forecasts from 1982Q1 to 2013Q1, and Blue Chip 10 years inflation forecasts from 1982Q1 to 2013Q1.

To formally test if there is substantial time variation in the slope of the Phillips curve λ_t, we compute the Bayes factor favouring the six different unrestricted models, M1–M6, against their corresponding restricted versions, where λ_t is constant ($\omega_\lambda = 0$). The test results for the time variation of the slopes of different models with the Phillips curve are shown in Table 2.1.

Table 2.1: Estimated Log Bayes Factors and the Numerical Standard Errors.

	M1	M2	M3	M4	M5	M6
Log BF	4.8 (0.08)	4.1 (0.12)	8.4 (2.05)	57.2 (3.57)	0.2 (0.03)	0.7 (0.06)

Overall, for most models, the data prefer the version with time variation. Specifically, the log Bayes factors associated with M1, M2, M3 and M4 are all larger than 4, indicating substantial time variation in the slope, λ_t. Conversely, the log Bayes factors associated with M5 and M6 are small but positive, suggesting slight evidence supporting time variation in λ_t.

To corroborate these model comparison results, we plot the posterior estimates of λ_t and ω_λ in Figures 2.1 and Figure 2.2. First, Figure 2.1 shows the results for the Phillips curve specified as the univariate unobserved components model with stochastic volatility. Figure 2.2 shows the corresponding results for the bivariate unobserved components model with stochastic volatility.

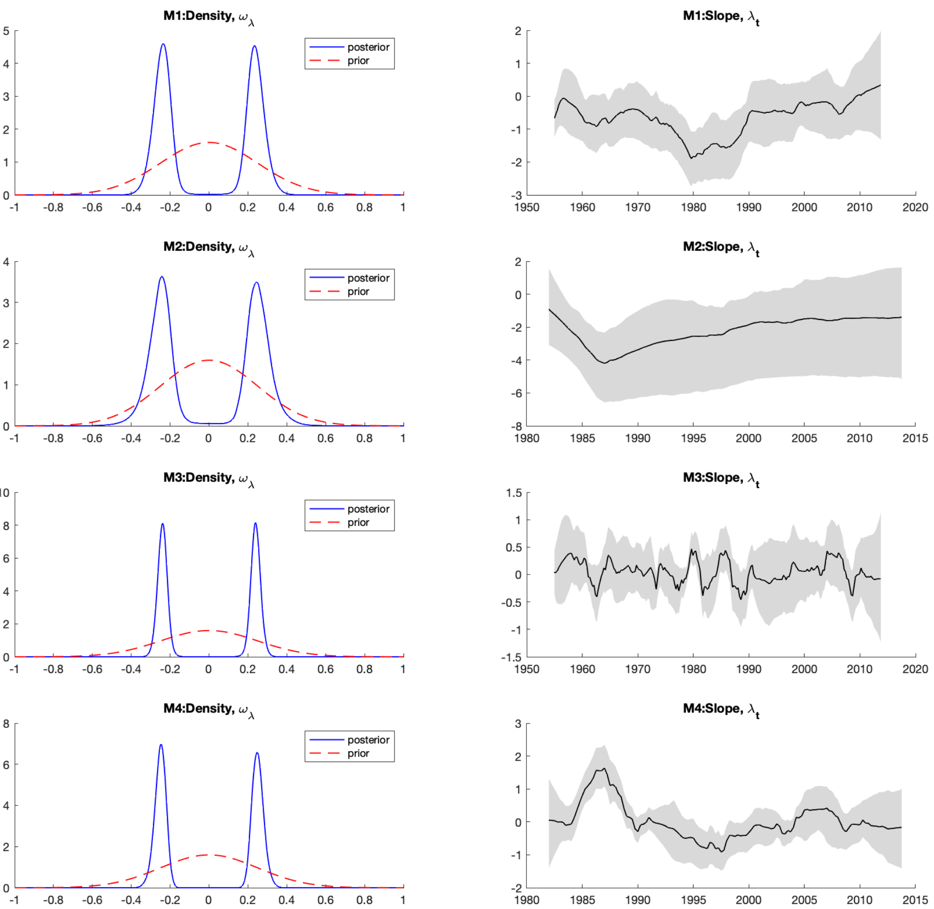


Figure 2.1: Estimated Slope λ_t and Density of ω_λ for the Phillips Curves Specified as Univariate Unobserved Components Model with Stochastic Volatility.

Consistent with model comparison results, the right panel of Figure 2.1 shows that estimates of the slopes of the Phillips curve, λ_t, under M1 and M2 are volatile and time-varying. They are always negative, consistent with the idea of the Phillips curve that there is a trade-off between inflation and the unemployment gap. Also, starting from the 1980s, λ_t becomes flatter and is closer to zero. Estimates of the slopes of the Phillips curve, λ_t, under M3 and M4 are also volatile but they mostly move around 0. This suggests that the real GDP gap has little effects on inflation. These results are similar to those reported in Berger et al. (2016) and Chan and Grant (2017). They both find the magnitude of λ_t is small when the economic slack is measured as the output gap.

The left panel of Figure 2.1 shows that the posterior densities of ω_λ under M1, M2, M3, and M4 are all bimodal and have almost no mass around 0. This can be viewed as strong evidence supporting the time-varying λ_t. Figure 2.1 indicates that when the Phillips curve is specified as the univariate unobserved components model with stochastic volatility, where the unobserved component for inflation is the trend of real economic activities, the slope, λ_t, of the Phillips curve has substantial time variation.

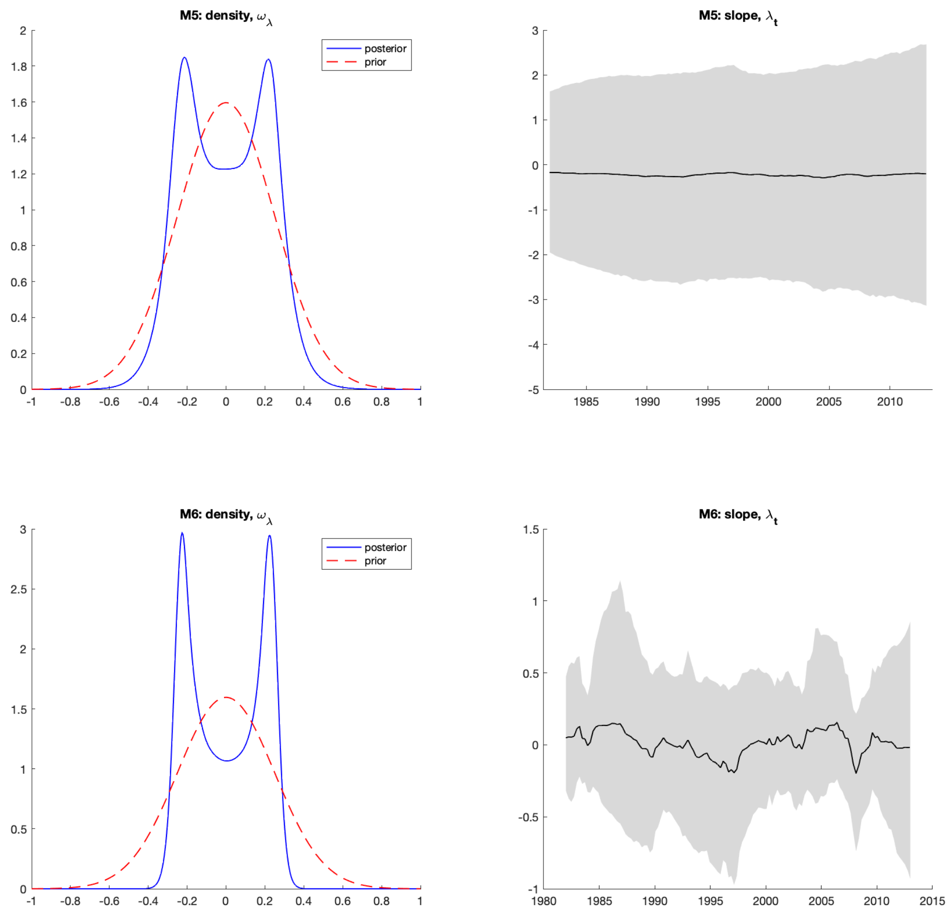


Figure 2.2: Estimated Slope λ_t and Density of ω_λ for the Phillips Curves with Bivariate Unobserved Components Model with Stochastic Volatility.

The right panel of Figure 2.2 shows that estimates of the time-varying slopes of the Phillips curve, λ_t, of M5 and M6 are insignificant and stable around 0. The left panel of Figure 2.2 shows that the posterior densities of ω_λ under M5 and M6 are bimodal but have a considerable mass around 0. However, compared to the prior density, the posterior density at 0 is lower, suggesting ω_λ is less likely to be 0 given the data. This is consistent with the model comparison result that shows moderate evidence on the time variation of λ_t.

To summarise our results so far, in the univariate case, the slope of the Phillips curve is conclusively time varying. In the bivariate case, evidence of the time variation on the slope of the Phillips curve is suggestive but not conclusive.

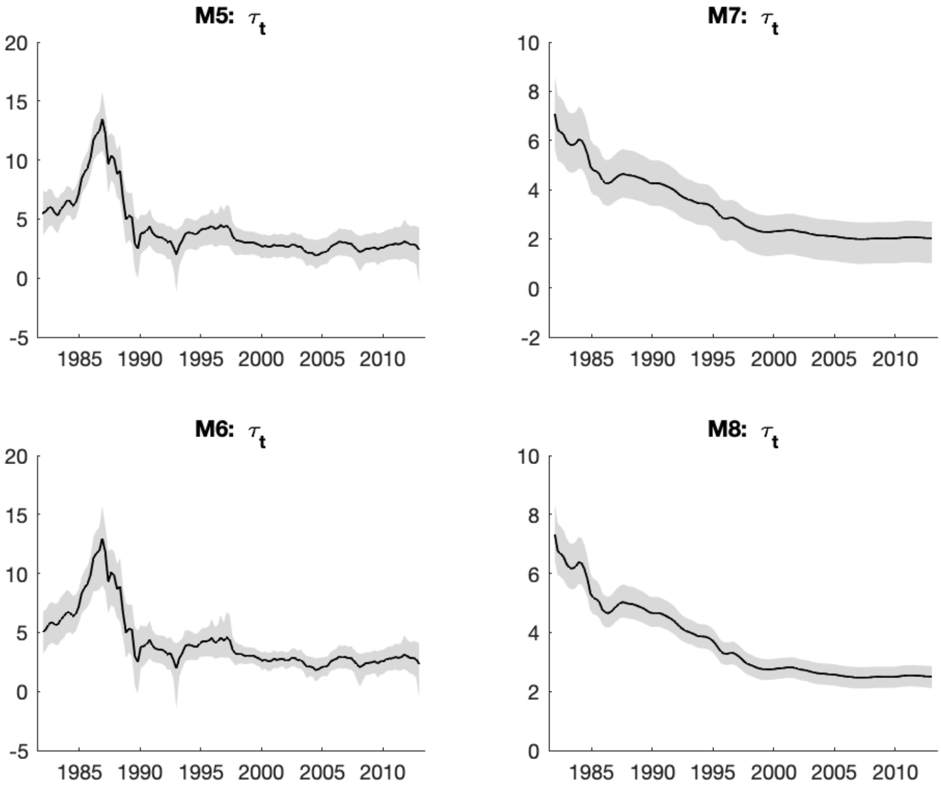


Figure 2.3: Trend Inflation: τ_t.

The inconclusive evidence in the bivariate case could be due to the substantial variance in the estimated trend inflation. To investigate this possibility, we follow Chan et al. (2018), who link trend inflation, τ_t, to the blue chip inflation forecasts that substantially reduce the variance in the estimated trend inflation. Following Chan et al. (2018), we add an additional measurement equation linking trend inflation to the blue chip 10 years inflation forecasts to M5 and M6 respectively, resulting in M7 and M8. Figure 2.3 shows that with the additional information from the blue chip 10 years inflation forecasts, M7 and M8 have a substantially smaller variance in the estimated trend inflation than do M5 and M6.

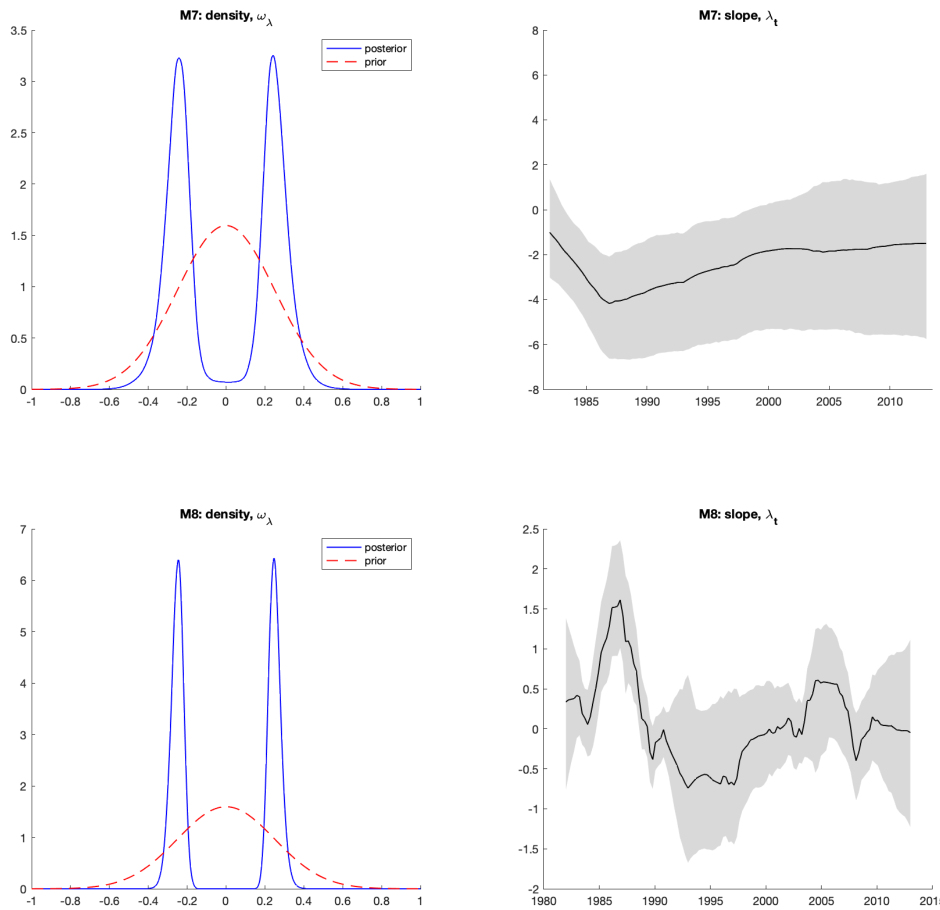


Figure 2.4: Estimated slope λ_t and density of ω_λ of M7 and M8

Next, we test the time variation in the slopes under M7 and M8. The log Bayes factors associated with M7 and M8 are 3.5 (0.24) and 51.0 (3.46) respectively. These values are large, indicating substantial time variation in the slope, λ_t. Consistent with model comparison results of M7 and M8, Figure 2.4 shows that estimates of the slopes of the Phillips curve, λ_t, of M7 and M8 are volatile and the posterior densities of ω_λ under M7 and M8 are bimodal and have almost no mass around 0. This shows strong evidence favouring time variation in the slope.

In summary, in the univariate case, the slope of the Phillips curve is conclusively time varying. Moreover, in the bivariate case, the slope of the Phillips curve is also conclusively time varying with more precise estimates of the trend inflation, τ_t.

2.6 Conclusion

In this chapter, we estimate eight Phillips curve models and test for time variation in the slopes of the Phillips curve under these models. First, we find that CPI inflation is much more sensitive to changes in the unemployment gap than to changes in the output gap. Models with the unemployment gap have a much larger λ_t in magnitude than models with the output gap. Second, we find that time variation of the slope of the Phillips curve mainly depends on specifications of inflation expectations. When the measures of inflation expectations are observable, the slope of the Phillips curve is time varying. However, when the measure of inflation expectations is trend inflation with large variance, the slope of the Phillips curve is constant. By reducing the variance in the estimated trend inflation through linking blue chip 10 years inflation forecast to trend inflation, the Phillips curve model with trend inflation has the time-varying slope. We consider different measures of inflation expectations: backward-looking inflation expectations, forward-looking inflation expectations, and trend inflation. The slopes of these Phillips curves with different measures of inflation expectations are volatile and time varying. Thus, we conclude that the slope of the Phillips curve is time varying.

2.7 Appendix

2.7.1 Details of the Specific Models

In this section, we outline the eight model, M1-M8, in detail. In general, we have two classes of models: univariate and bivariate unobserved components models. The former consists of

M1–M4 that use different measures of economic slack and inflation expectations. The latter comprises M5–M8. The specifications for each model are discussed below.

M1

M1 is a univariate unobserved components model where π_t is inflation. λ_t is the slope of the Phillips curve. $\mathrm{E}_t\pi_{t+1}$ is measured as the average of past four quarter inflation, $\pi^e_{t|t-1}$=$(\pi_{t|t-1}+\pi_{t|t-2}+\pi_{t|t-3}+\pi_{t|t-4})/4$. λ_t is modeled as a random walk. u_t represents the unemployment rate. e_t follows an AR(2) process. NAIRU, ν_t, is modeled as a random walk. The log of stochastic volatility, h_t, is modelled as a random walk.

$$\pi_t-\pi^e_{t|t-1}=\lambda_t(u_t-\nu_t)+\varepsilon^\pi_t, \qquad \varepsilon^\pi_t\sim\mathcal{N}(0,e^{h_t}), \tag{2.19}$$

$$u_t=\nu_t+e_t, \tag{2.20}$$

$$e_t=\phi_1 e_{t-1}+\phi_2 e_{t-2}+\varepsilon^e_t, \qquad \varepsilon^e_t\sim\mathcal{N}(0,\omega^2_e), \tag{2.21}$$

$$\nu_t=\nu_{t-1}+\varepsilon^\nu_t, \qquad \varepsilon^\nu_t\sim\mathcal{N}(0,\omega^2_\nu), \tag{2.22}$$

$$\lambda_t=\lambda_{t-1}+\varepsilon^\lambda_t, \qquad \varepsilon^\lambda_t\sim\mathcal{N}(0,\omega^2_\lambda), \tag{2.23}$$

$$h_t=h_{t-1}+\varepsilon^h_t, \qquad \varepsilon^h_t\sim\mathcal{N}(0,\omega^2_h). \tag{2.24}$$

M2

M2 is a univariate unobserved components model where π_t is inflation. λ_t is the slope of the Phillips curve. $\mathrm{E}_t\pi_{t+1}$ is measured as SPF one-year inflation forecasts, $\pi^e_{t+1|t}$. λ_t is modelled as a random walk. u_t represents the unemployment rate. e_t follows an AR(2) process. NAIRU, ν_t, is modelled as a random walk. The log of stochastic volatility, h_t, is modelled as a random walk.

$$\pi_t - \pi^e_{t+1|t} = \lambda_t(u_t - \nu_t) + \varepsilon^\pi_t, \qquad \varepsilon^\pi_t \sim \mathcal{N}(0, e^{h_t}), \tag{2.25}$$

$$u_t = \nu_t + e_t, \tag{2.26}$$

$$e_t = \phi_1 e_{t-1} + \phi_2 e_{t-2} + \varepsilon^e_t, \qquad \varepsilon^e_t \sim \mathcal{N}(0, \omega^2_e), \tag{2.27}$$

$$\nu_t = \nu_{t-1} + \varepsilon^\nu_t, \qquad \varepsilon^\nu_t \sim \mathcal{N}(0, \omega^2_\nu), \tag{2.28}$$

$$\lambda_t = \lambda_{t-1} + \varepsilon^\lambda_t, \qquad \varepsilon^\lambda_t \sim \mathcal{N}(0, \omega^2_\lambda), \tag{2.29}$$

$$h_t = h_{t-1} + \varepsilon^h_t, \qquad \varepsilon^h_t \sim \mathcal{N}(0, \omega^2_h). \tag{2.30}$$

M3

M3 is a univariate unobserved components model where π_t is inflation. λ_t is the slope of the Phillips curve. $\mathrm{E}_t\pi_{t+1}$ is measured as the average of past four quarter inflation, $\pi^e_{t|t-1}$=$(\pi_{t|t-1} + \pi_{t|t-2} + \pi_{t|t-3} + \pi_{t|t-4})/4$. λ_t is modelled as a random walk. The cyclical component, c_t, follows an AR(2) process. y_t represents real output level. Underlying output trend growth, Δy^*_t, is modelled as a random walk. The log of stochastic volatility, h_t, is modelled as a random walk.

$$\pi_t - \pi^e_{t|t-1} = \lambda_t(y_t - y^*_t) + \varepsilon^\pi_t, \qquad \varepsilon^\pi_t \sim \mathcal{N}(0, e^{h_t}), \tag{2.31}$$

$$y_t = y^*_t + c_t, \tag{2.32}$$

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \varepsilon^c_t, \qquad \varepsilon^c_t \sim \mathcal{N}(0, \omega^2_c), \tag{2.33}$$

$$\Delta y^*_t = \Delta y^*_{t-1} + \varepsilon^{y*}_t, \qquad \varepsilon^{y*}_t \sim \mathcal{N}(0, \omega^2_{y*}), \tag{2.34}$$

$$\lambda_t = \lambda_{t-1} + \varepsilon^\lambda_t, \qquad \varepsilon^\lambda_t \sim \mathcal{N}(0, \omega^2_\lambda), \tag{2.35}$$

$$h_t = h_{t-1} + \varepsilon^h_t, \qquad \varepsilon^h_t \sim \mathcal{N}(0, \omega^2_h). \tag{2.36}$$

M4

M4 is a univariate unobserved components model where π_t is inflation. λ_t is the slope of the Phillips curve. $\mathrm{E}_t\pi_{t+1}$ is measured as SPF one-year inflation forecasts, $\pi^e_{t+1|t}$. λ_t is modeled as random walk. y_t represents real output level. The cyclical component, c_t, follows an AR(2) process. Underlying output trend growth, Δy^*_t, is modelled as a random walk. The log of stochastic volatility, h_t, is modelled as a random walk.

$$\pi_t - \pi^e_{t|t+1} = \lambda_t(y_t - y^*_t) + \varepsilon^\pi_t, \qquad \varepsilon^\pi_t \sim \mathcal{N}(0, e^{h_t}), \tag{2.37}$$

$$y_t = y^*_t + c_t, \tag{2.38}$$

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \varepsilon^c_t, \qquad \varepsilon^c_t \sim \mathcal{N}(0, \omega^2_c), \tag{2.39}$$

$$\Delta y^*_t = \Delta y^*_{t-1} + \varepsilon^{y*}_t, \qquad \varepsilon^{y*}_t \sim \mathcal{N}(0, \omega^2_{y*}), \tag{2.40}$$

$$\lambda_t = \lambda_{t-1} + \varepsilon^\lambda_t, \qquad \varepsilon^\lambda_t \sim \mathcal{N}(0, \omega^2_\lambda), \tag{2.41}$$

$$h_t = h_{t-1} + \varepsilon^h_t, \qquad \varepsilon^h_t \sim \mathcal{N}(0, \omega^2_h). \tag{2.42}$$

M5

M5 is a bivariate unobserved components model where π_t is inflation. λ_t is the slope of the Phillips curve. Long-run inflation expectations are measured as trend inflation, τ_t, and follow a random walk. λ_t is modelled as a random walk. u_t represents the unemployment rate. e_t follows an AR(2) process. NAIRU, ν_t, is modelled as a random walk. Two variables of log of stochastic volatility, h_t and g_t, are modelled as a random walk.

$$\pi_t - \tau_t = \lambda_t (u_t - \nu_t) + \varepsilon_t^{\pi}, \qquad \varepsilon_t^{\pi} \sim \mathcal{N}(0, e^{h_t}), \tag{2.43}$$

$$\tau_t = \tau_{t-1} + \varepsilon_t^{\tau}, \qquad \varepsilon_t^{\tau} \sim \mathcal{N}(0, e^{g_t}), \tag{2.44}$$

$$u_t = \nu_t + e_t,$$

$$e_t = \phi_1 e_{t-1} + \phi_2 e_{t-2} + \varepsilon_t^{e}, \qquad \varepsilon_t^{e} \sim \mathcal{N}(0, \omega_e^2), \tag{2.45}$$

$$\nu_t = \nu_{t-1} + \varepsilon_t^{\nu}, \qquad \varepsilon_t^{\nu} \sim \mathcal{N}(0, \omega_\nu^2), \tag{2.46}$$

$$\lambda_t = \lambda_{t-1} + \varepsilon_t^{\lambda}, \qquad \varepsilon_t^{\lambda} \sim \mathcal{N}(0, \omega_\lambda^2), \tag{2.47}$$

$$h_t = h_{t-1} + \varepsilon_t^{h}, \qquad \varepsilon_t^{h} \sim \mathcal{N}(0, \omega_h^2), \tag{2.48}$$

$$g_t = g_{t-1} + \varepsilon_t^{g}, \qquad \varepsilon_t^{g} \sim \mathcal{N}(0, \omega_g^2). \tag{2.49}$$

M6

M6 is a bivariate unobserved components model where π_t is inflation. λ_t is the slope of the Phillips curve. Long-run inflation expectations are measured as trend inflation, τ_t, and follow a random walk. λ_t is modelled as a random walk. y_t represents the real output level. The cyclical component, c_t, follows an AR(2) process. Underlying output trend growth, Δy_t^*, is modelled as the random walk. Two variables of log of stochastic volatility, h_t and g_t, are modelled as a random walk.

$$\pi_t - \tau_t = \lambda_t(y_t - y_t^*) + \varepsilon_t^\pi, \qquad \varepsilon_t^\pi \sim \mathcal{N}(0, e^{h_t}), \tag{2.50}$$

$$\tau_t = \tau_{t-1} + \varepsilon_t^\tau, \qquad \varepsilon_t^\tau \sim \mathcal{N}(0, e^{g_t}), \tag{2.51}$$

$$y_t = y_t^* + c_t, \tag{2.52}$$

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \varepsilon_t^c, \qquad \varepsilon_t^c \sim \mathcal{N}(0, \omega_c^2), \tag{2.53}$$

$$\Delta y_t^* = \Delta y_{t-1}^* + \varepsilon_t^{y*}, \qquad \varepsilon_t^{y*} \sim \mathcal{N}(0, \omega_{y^*}^2), \tag{2.54}$$

$$\lambda_t = \lambda_{t-1} + \varepsilon_t^\lambda, \qquad \varepsilon_t^\lambda \sim \mathcal{N}(0, \omega_\lambda^2), \tag{2.55}$$

$$h_t = h_{t-1} + \varepsilon_t^h, \qquad \varepsilon_t^h \sim \mathcal{N}(0, \omega_h^2), \tag{2.56}$$

$$g_t = g_{t-1} + \varepsilon_t^g, \qquad \varepsilon_t^g \sim \mathcal{N}(0, \omega_g^2). \tag{2.57}$$

M7

M7 is a bivariate unobserved components model where π_t is inflation. λ_t is the slope of the Phillips curve. Long-run inflation expectations are measured as trend inflation, τ_t, and follow a random walk. λ_t is modelled as a random walk. u_t represents the unemployment rate. e_t follows an AR(2) process. NAIRU, ν_t, is modelled as a random walk. Two variables of log of stochastic volatility, h_t and g_t, are modeled as a random walk. q_t is the blue chip 10 years inflation forecasts.

$$\pi_t - \tau_t = \lambda_t(u_t - \nu_t) + \varepsilon_t^{\pi}, \qquad \varepsilon_t^{\pi} \sim \mathcal{N}(0, e^{h_t}), \tag{2.58}$$

$$\tau_t = \tau_{t-1} + \varepsilon_t^{\tau}, \qquad \varepsilon_t^{\tau} \sim \mathcal{N}(0, e^{g_t}), \tag{2.59}$$

$$u_t = \nu_t + e_t,$$

$$e_t = \phi_1 e_{t-1} + \phi_2 e_{t-2} + \varepsilon_t^{e}, \qquad \varepsilon_t^{e} \sim \mathcal{N}(0, \omega_e^2), \tag{2.60}$$

$$\nu_t = \nu_{t-1} + \varepsilon_t^{\nu}, \qquad \varepsilon_t^{\nu} \sim \mathcal{N}(0, \omega_\nu^2), \tag{2.61}$$

$$\lambda_t = \lambda_{t-1} + \varepsilon_t^{\lambda}, \qquad \varepsilon_t^{\lambda} \sim \mathcal{N}(0, \omega_\lambda^2), \tag{2.62}$$

$$h_t = h_{t-1} + \varepsilon_t^{h}, \qquad \varepsilon_t^{h} \sim \mathcal{N}(0, \omega_h^2), \tag{2.63}$$

$$g_t = g_{t-1} + \varepsilon_t^{g}, \qquad \varepsilon_t^{g} \sim \mathcal{N}(0, \omega_g^2), \tag{2.64}$$

$$q_t = d_0 + d_1 \tau_t + \varepsilon_t^{q}, \qquad \varepsilon_t^{q} \sim \mathcal{N}(0, \omega_q^2). \tag{2.65}$$

M8

M8 is a bivariate unobserved components model where π_t is inflation. λ_t is the slope of the Phillips curve. Long-run inflation expectations are measured as trend inflation, τ_t, and follow a random walk. λ_t is modelled as a random walk. y_t represents the real output level. The cyclical component, c_t, follows an AR(2) process. Underlying output trend growth, Δy_t^*, is modeled as a random walk. Two variables of log of stochastic volatility, h_t and g_t, are modelled as a random walk. q_t is the blue chip 10 years inflation forecasts.

$$\pi_t - \tau_t = \lambda_t(y_t - y_t^*) + \varepsilon_t^\pi, \qquad \varepsilon_t^\pi \sim \mathcal{N}(0, e^{h_t}), \tag{2.66}$$

$$\tau_t = \tau_{t-1} + \varepsilon_t^\tau, \qquad \varepsilon_t^\tau \sim \mathcal{N}(0, e^{g_t}), \tag{2.67}$$

$$y_t = y_t^* + c_t, \tag{2.68}$$

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \varepsilon_t^c, \qquad \varepsilon_t^c \sim \mathcal{N}(0, \omega_c^2), \tag{2.69}$$

$$\Delta y_t^* = \Delta y_{t-1}^* + \varepsilon_t^{y*}, \qquad \varepsilon_t^{y*} \sim \mathcal{N}(0, \omega_{y^*}^2), \tag{2.70}$$

$$\lambda_t = \lambda_{t-1} + \varepsilon_t^\lambda, \qquad \varepsilon_t^\lambda \sim \mathcal{N}(0, \omega_\lambda^2), \tag{2.71}$$

$$h_t = h_{t-1} + \varepsilon_t^h, \qquad \varepsilon_t^h \sim \mathcal{N}(0, \omega_h^2), \tag{2.72}$$

$$g_t = g_{t-1} + \varepsilon_t^g, \qquad \varepsilon_t^g \sim \mathcal{N}(0, \omega_g^2), \tag{2.73}$$

$$q_t = d_0 + d_1\tau_t + \varepsilon_t^q, \qquad \varepsilon_t^q \sim \mathcal{N}(0, \omega_q^2). \tag{2.74}$$

2.7.2 Estimation Details

In this section, we provide the details of the priors and estimation for M5, M6 and M7 are outlined in this section. Estimation for M1 and M2 is similar to M5, estimation for M3 and M4 is similar to M6, and estimation for M8 is similar to M7. Thus, for brevity, we omit estimation details for these five models.

M5

Prior

The parameters under M5 are τ, ν, ϕ,$\widetilde{\lambda}$, λ_0, ω_λ, ω_e^2, ω_ν^2, h, and g.

We assume the following priors:

$$\begin{array}{cccc}
\tau_0 = 0, & \tau_1 \sim \mathcal{N}(\tau_0, V_\tau e^{g_t}), & \lambda_0 \sim \mathcal{N}(a_0, V_{\lambda_0}), & e_0 = 0, \\
\omega_\lambda \sim \mathcal{N}(0, V_{\omega_\lambda}), & \omega_g = V_{\omega_g}^{1/2}, & \omega_h = V_{\omega_h}^{1/2}, & \omega_e^2 \sim \mathcal{IG}(\nu_e, S_{\omega_e}), \\
V_{\omega_h} = 0.2, & V_{\omega_g} = 0.2, & V_{\omega_\lambda} = 0.25^2, & \nu_e = 3, \\
V_{\lambda_0} = 0.25^2, & V_\tau = 10, & V_g = 10, & \nu_\nu = 3, \\
a_0 = -0.25, & V_\beta = (V_{\lambda_0}, V_{\omega_\lambda}), & \widehat{\beta} = (a_0, 0), & \phi \sim \mathcal{N}(\phi_0, V_\phi), \\
\phi_0 = (0.5; 0.2), & e_{-1} = 0, & \omega_\nu^2 \sim, \mathcal{IG}(\nu_\nu, S_{\omega_\nu}), & S_{\omega_\nu} = 1 * (\nu_\nu - 1), \\
S_{\omega_e} = 1 * (\nu_e - 1), & V_\phi = \mathrm{I}_2. & &
\end{array}$$

Likelihood

In this section, we derive the densities of $\pi = (\pi_1, \ldots, \pi_T)'$ and $u = (u_1, \ldots, u_T)'$, that will be used to construct the posterior sampler.

Let

$$\Lambda_\lambda = \mathrm{diag}(\lambda_0 + \omega_\lambda \widetilde{\lambda}_1, \lambda_0 + \omega_\lambda \widetilde{\lambda}_2, \lambda_0 + \omega_\lambda \widetilde{\lambda}_3, \ldots, \lambda_0 + \omega_\lambda \widetilde{\lambda}_T).$$

Then, we have

$$\pi - \tau - \Lambda_\lambda (u - \nu) = \varepsilon^\pi.$$

Then, the log conditional density of π is

$$\log p(\pi \mid \tau, u, \nu, c, \widetilde{\lambda}, \lambda_0, \omega_h, \omega_\lambda, \omega_\nu, \omega_c, h) \propto -\frac{1}{2}(\pi - \tau - \Lambda_\lambda (u - \nu))' S_\pi^{-1} (\pi - \tau - \Lambda_\lambda (u - \nu)),$$

where

$$S_\pi = \mathrm{diag}(e^{h_1}, e^{h_2}, e^{h_3}, \ldots, e^{h_T}).$$

Let

$$H_\phi = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ -\phi_1 & 1 & 0 & 0 & \dots & 0 \\ -\phi_1 & -\phi_2 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & \dots & \dots & -\phi_1 & -\phi_2 & 1 \end{bmatrix}.$$

Then, we have

$$H_\phi e = \varepsilon^e.$$

Then, the log conditional density of u is

$$\log p(u \mid \nu, e, \omega_\nu, \omega_e, \phi) \propto -\frac{T}{2}\log \omega_e^2 - \frac{1}{2\omega_e^2}(u-\nu)' H_\phi' H_\phi (u-\nu).$$

Sampling τ

In this section, we derive the joint prior density of $\tau = (\tau_1, \dots, \tau_T)'$, that will be used to construct the posterior sampler of τ.

Let

$$S_g = (V_\tau e^{g_1}, e^{g_2}, e^{g_3}, \dots, e^{g_T}).$$

and

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ -1 & 1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 \\ 0 & 0 & -1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & \dots & \dots & 0 & -1 & 1 \end{bmatrix}.$$

Then, we have

$$H\tau = \varepsilon^{\tau}.$$

Then,

$$\tau \sim \mathcal{N}(0, H^{-1}S_g H'^{-1}).$$

Then, the log prior density for τ is

$$\log p(\tau) = -\frac{1}{2}\tau' H' S_g^{-1} H\tau.$$

Then, we have

$$\begin{aligned}
&\log p(\tau \mid \pi, u, \nu, e, \phi, \gamma, e, \omega_\nu, \widetilde{\lambda}, \lambda_0, \omega_h, \omega_\lambda, \omega_g, h, g) \\
&\propto -\frac{1}{2}(\tau' S_\pi^{-1}\tau - 2\tau' S_\pi^{-1}(\pi - \Lambda_\lambda(u - \nu))) - \frac{1}{2}\tau' H' S_g^{-1} H\tau.
\end{aligned}$$

Then, the conditional distribution of τ is

$$\tau \sim \mathcal{N}(\widehat{\tau}, K_\tau^{-1}),$$

where

$$\widehat{\tau} = K_\tau^{-1}((S_\pi^{-1})(\pi - \Lambda_\lambda(u - \nu))), \qquad K_\tau = S_\pi^{-1} + H' S_g^{-1} H.$$

Since K_τ is a band matrix, τ can be sampled using the precision sampler proposed by Chan and Jeliazkov (2009).

Sample h and g

we sample h and g, following Kim et al. (1998).

Sample ν

In this section, we construct the posterior sampler of ν.

We have

$$\nu \sim \mathscr{N}(0, H'^{-1}\omega_\nu^2 H^{-1}).$$

Then, the log prior density of ν is

$$\log p(\nu) = -\frac{T}{2}\log\omega_\nu^2 - \frac{1}{2\omega_\nu^2}\nu' H' H \nu.$$

Then, the posterior distribution of ν is

$$\mathscr{N}(\widehat{\nu}, K_\nu^{-1}),$$

where

$$K_\nu = \frac{H_\phi' H_\phi}{\omega_e^2} + \frac{H'H}{\omega_\nu^2} + \Lambda_\lambda' S_\pi^{-1}\Lambda_\lambda$$

and

$$\widehat{\nu} = K_\nu^{-1}(\frac{H_\phi' H_\phi u}{\omega_e^2} - S_\pi^{-1}\Lambda_\lambda(\pi - \tau - \Lambda_\lambda u)).$$

Sample ϕ

In this section, we construct the posterior sampler of ϕ.

Let

$$\phi = \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix}$$

and

$$X_\phi = \begin{bmatrix} c_0 & c_{-1} \\ c_1 & c_0 \\ c_3 & c_2 \\ \vdots & \vdots \\ c_{T-1} & c_{T-2} \end{bmatrix},$$

then, we have

$$e = X_\phi \phi + \varepsilon^e.$$

Then, the conditional distribution of ϕ is

$$\phi \sim \mathcal{N}(\widehat{\phi}, K_\phi^{-1}) 1(\phi \in R)$$

where

$$\widehat{\phi} = K_\phi^{-1}(V_\phi^{-1}\phi_0 + \frac{X_\phi' e}{\omega_e^2}), \qquad K_\phi^{-1} = V_\phi^{-1} + \frac{X_\phi' X_\phi}{\omega_e^2}.$$

Sample $\widetilde{\lambda}$

In this section, we construct the posterior sampler of $\widetilde{\lambda}$.

We have

$$H\widetilde{\lambda} = \varepsilon^{\widetilde{\lambda}}.$$

Then, $\widetilde{\lambda}$ is distributed as

$$\mathcal{N}(0, H^{-1}H'^{-1}).$$

Then, the log prior density of $\widetilde{\lambda}$ is

$$\log p(\widetilde{\lambda}) = -\frac{1}{2}(\widetilde{\lambda}' H' H \widetilde{\lambda}).$$

Let

$$\Lambda_u = \mathrm{diag}(u_1 - v_1, u_2 - v_2, u_3 - v_3, \ldots, u_T - v_T).$$

Then, we have

$$-\frac{1}{2}(\omega_\lambda^2 \widetilde{\lambda}' \Lambda_u S_\pi^{-1} \Lambda_u \widetilde{\lambda}) - 2(\widetilde{\lambda}' \Lambda_u \omega_\lambda S_\pi^{-1}(\pi - \tau - \lambda_0 \Lambda_u)).$$

$\widetilde{\lambda}$ is distributed as

$$\mathcal{N}(\widehat{\widetilde{\lambda}}, K_{\widetilde{\lambda}}^{-1}),$$

where

$$\widehat{\widetilde{\lambda}} = K_{\widetilde{\lambda}}^{-1}(\Lambda_u \omega_\lambda S_\pi^{-1}(\pi - \tau - \lambda_0 \Lambda_u)), \qquad K_{\widetilde{\lambda}} = H'H + \omega_\lambda^2 \Lambda_u S_\pi^{-1} \Lambda_u.$$

Sample λ_0 and ω_λ

In this section, we construct the posterior sampler of λ_0 and ω_λ.

Let $X_\beta = (u - v, \Lambda_u \widetilde{\lambda})$ and $\beta = (\lambda_0, \omega_\lambda)'$,

Then, we have

$$\pi - \tau = X_\beta \beta + \varepsilon^\pi.$$

Then, β is distributed as

$$\mathcal{N}(\widehat{\beta}, K_\beta^{-1}),$$

where

$$K_\beta = V_\beta^{-1} + X_\beta' S_\pi^{-1} X_\beta, \qquad \widehat{\beta} = K_\beta^{-1}(V_\beta^{-1}\beta_0 + X_\beta' S_\pi^{-1}(\pi - \tau)).$$

Sample ω_e^2

In this section, we show the posterior sampler of ω_c^2.

The conditional distribution of ω_e^2 is

$$\mathscr{IG}(\nu_e + \frac{T}{2}, S_{\omega_e} + \frac{1}{2}(e - X_\phi \phi)'(e - X_\phi \phi)).$$

Sample ω_ν^2

In this section, we show the posterior sampler of ω_ν^2.

The conditional distribution of ω_ν^2 is

$$\mathscr{IG}(\nu_\nu + \frac{T}{2}, S_{\omega_\nu} + \frac{1}{2}\nu' H' H \nu).$$

M6

Prior

The parameters under M5 are τ, y^*, γ, ϕ, $\widetilde{\lambda}$, λ_0, ω_λ, ω_c^2, and $\omega_{y^*}^2$.

We assume the following priors:

$$\begin{array}{cccc}
\tau_0 = 0, & \tau_1 \sim \mathscr{N}(\tau_0, V_\tau e^{g_t}), & \lambda_0 \sim \mathscr{N}(a_0, V_{\lambda_0}), & c_0 = 0, \\
\omega_\lambda \sim \mathscr{N}(0, V_{\omega_\lambda}), & \omega_g = V_{\omega_g}^{1/2}, & \omega_h = V_{\omega_h}^{1/2}, & \omega_c^2 \sim \mathscr{IG}(\nu_c, S_{\omega_c}), \\
V_{\omega_h} = 0.2, & V_{\omega_g} = 0.2, & V_{\omega_\lambda} = 0.25^2, & \nu_c = 3, \\
V_{\lambda_0} = 0.25^2, & V_\tau = 10, & V_g = 10, & \gamma \sim \mathscr{N}(\gamma_0, V_\gamma), \\
a_0 = -0.25, & V_\beta = (V_{\lambda_0}, V_{\omega_\lambda}), & \widehat{\beta} = (a_0, 0), & V_\gamma = 100 * \mathrm{I}_2, \\
\gamma_0 = (750; 750), & \phi \sim \mathscr{N}(\phi_0, V_\phi), & \phi_0 = (1.34; -0.7), & c_{-1} = 0, \\
\omega_{y^*}^2 \sim \mathscr{U}(0, V_{\omega_{y^*}}), & S_{\omega_c} = 1 * (\nu_c - 1), & V_{\omega_{y^*}} = 0.001, & V_\phi = \mathrm{I}_2.
\end{array}$$

Likelihood

In this section, we derive the densities of $\pi = (\pi_1, \ldots, \pi_T)'$ and $y = (y_1, \ldots, y_T)'$, that will be used to construct the posterior sampler.

Let

$$\Lambda_\lambda = \mathrm{diag}(\lambda_0 + \omega_\lambda \widetilde{\lambda}_1, \lambda_0 + \omega_\lambda \widetilde{\lambda}_2, \lambda_0 + \omega_\lambda \widetilde{\lambda}_3, \ldots, \lambda_0 + \omega_\lambda \widetilde{\lambda}_T).$$

Then, we have

$$\pi - \tau - \Lambda_\lambda (y - y^*) = \varepsilon^\pi.$$

Then, the log conditional density of π is

$$\log p(\pi \mid \tau, y, y^*, c, \widetilde{\lambda}, \lambda_0, \omega_h, \omega_\lambda, \omega_{y^*}, \omega_c, h) \propto -\frac{1}{2}(\pi - \tau - \Lambda_\lambda (y - y^*))' S_\pi^{-1} (\pi - \tau - \Lambda_\lambda (y - y^*)),$$

where

$$S_\pi = \mathrm{diag}(e^{h_1}, e^{h_2}, e^{h_3}, \ldots, e^{h_T}).$$

Let

$$H_\phi = \begin{bmatrix} 1 & 0 & 0 & 0 & \ldots & 0 \\ -\phi_1 & 1 & 0 & 0 & \ldots & 0 \\ -\phi_1 & -\phi_2 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & \ldots & \ldots & -\phi_1 & -\phi_2 & 1 \end{bmatrix}.$$

Then, we have

$$y = y^* + c,$$

$$H_\phi c = \varepsilon^c.$$

Then, the log conditional density of y is

$$\log p(y \mid y^*, c, \omega_{y^*}, \omega_c, \phi, y_0^*, y_{-1}^*) \propto -\frac{T}{2} \log \omega_c^2 - \frac{1}{2\omega_c^2}(y - y^*)' H_\phi' H_\phi (y - y^*)$$

Sampling τ

In this section, we derive the densities of $\tau = (\tau_1, \dots, \tau_T)'$, that will be used to construct the posterior sampler of τ.

Let

$$S_g = (V_\tau e^{g_1}, e^{g_2}, e^{g_3}, \dots, e^{g_T})$$

and

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ -1 & 1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 \\ 0 & 0 & -1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & \dots & \dots & 0 & -1 & 1 \end{bmatrix}.$$

Then, we have

$$H\tau = \varepsilon^\tau.$$

Then,

$$\tau \sim N(0, H^{-1} S_g H'^{-1}).$$

Then, the log prior density for τ is

$$\log p(\tau) = -\frac{1}{2}\tau' H' S_g^{-1} H\tau.$$

and we have

$$\log p(\tau \mid \pi, y, y^*, c, \phi, \gamma, \omega_c, \omega_{y^*}, \widetilde{\lambda}, \lambda_0, \omega_h, \omega_\lambda, \omega_g, h, g)$$
$$\propto -\frac{1}{2}(\tau' S_\pi^{-1}\tau - 2\tau' S_\pi^{-1}(\pi - \Lambda_\lambda(y - y^*))) - \frac{1}{2}\tau' H' S_g^{-1} H\tau.$$

Then, the conditional distribution of τ is

$$\tau \sim \mathcal{N}(\widehat{\tau}, K_\tau^{-1}),$$

where

$$\widehat{\tau} = K_\tau^{-1}((S_\pi^{-1})(\pi - \Lambda_\lambda(y - y^*))), \qquad K_\tau = S_\pi^{-1} + H' S_g^{-1} H.$$

Sample h and g

we sample h and g, following Kim et al. (1998).

Sample y^*

In this section, we construct the posterior sampler of y^*. Let

$$H_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & -2 & 0 & 0 & \dots & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & \dots & \dots & 1 & -2 & 1 \end{bmatrix}.$$

then we have

$$H_2 y^* = \widetilde{\alpha}_{y^*} + \varepsilon^{y^*}$$

where

$$\widetilde{\alpha}_{y^*} = (y_0^* + \Delta y_0^*, -y_0^*, 0, \dots, 0)'.$$

Let

$$\alpha_{y^*} = H_2^{-1} \widetilde{\alpha}_{y^*}.$$

Then,

$$y^* \sim N(\alpha_{y^*}, \omega_{y^*}^2 (H_2' H_2)^{-1}).$$

Then, the log prior density of y^* is

$$\log p(y^*) = -\frac{T}{2}\log \omega_{y^*}^2 - \frac{1}{2\omega_{y^*}^2}(y^* - \alpha_{y^*})' H_2' H_2 (y^* - \alpha_{y^*}).$$

The posterior distribution of y^* is

$$\mathcal{N}(\widehat{y^*}, K_{y^*}^{-1}),$$

where

$$K_{y^*} = \frac{H_\phi' H_\phi}{\omega_c^2} + \frac{H_2' H_2}{\omega_{y^*}^2} + \Lambda_\lambda' S_\pi^{-1} \Lambda_\lambda$$

and

$$\widehat{y^*} = K_{y^*}^{-1}\left(\frac{H_\phi' H_\phi y}{\omega_c^2} + \frac{H_2' H_2 \alpha_{y^*}}{\omega_{y^*}^2} - S_\pi^{-1} \Lambda_\lambda (\pi - \tau - \Lambda_\lambda y)\right).$$

Sample ϕ

In this section, we construct the posterior sampler of ϕ.

Let

$$\phi = \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix}$$

and

$$X_\phi = \begin{bmatrix} c_0 & c_{-1} \\ c_1 & c_0 \\ c_3 & c_2 \\ \vdots & \vdots \\ c_{T-1} & c_{T-2} \end{bmatrix}$$

then, we have

$$c = X_\phi \phi + \varepsilon^c.$$

Then, the conditional distribution of ϕ is

$$\phi \sim \mathcal{N}(\widehat{\phi}, K_\phi^{-1}) 1(\phi \in R)$$

where

$$\widehat{\phi} = K_\phi^{-1}(V_\phi^{-1}\phi_0 + \frac{X'_\phi c}{\omega_c^2}), \qquad K_\phi^{-1} = V_\phi^{-1} + \frac{X'_\phi X_\phi}{\omega_c^2}$$

Sample $\boldsymbol{\gamma}$

In this section, we construct the posterior sampler of γ.

Let

$$\boldsymbol{\gamma} = (y_0^*, y_{-1}^*)$$

and

$$\alpha_{y^*} = \begin{bmatrix} 2y_0^* - y_{-1}^* \\ 3y_0^* - 2y_{-1}^* \\ \dots \\ (T+1)y_0^* - Ty_{-1}^* \end{bmatrix} = X_\gamma \gamma$$

where

$$X_{\boldsymbol{\gamma}} = \begin{bmatrix} 2 & -1 \\ 3 & -2 \\ \vdots & \vdots \\ T+1 & -T \end{bmatrix}.$$

Then,

$$y^* = X_\gamma \boldsymbol{\gamma} + H_2^{-1} \varepsilon^{y^*}.$$

Then, the conditional distribution of γ is

$$\mathcal{N}(\widehat{\gamma}, K_{\boldsymbol{\gamma}}^{-1})$$

where

$$\widehat{\gamma} = K_{\boldsymbol{\gamma}}^{-1}(V_\gamma^{-1}\gamma_0 + \frac{X_\gamma' H_2' H_2 y^*}{\omega_{y*}^2}), \qquad K_\gamma = V_\gamma^{-1} + \frac{X_\gamma' H_2' H_2 X_\gamma}{\omega_{y^*}^2}.$$

Sample $\widetilde{\lambda}$

In this section, we construct the posterior sampler of $\widetilde{\lambda}$.

Let

$$H\widetilde{\lambda} = \varepsilon^{\widetilde{\lambda}}.$$

Then, $\widetilde{\lambda}$ is distributed as

$$N(0, H^{-1}H'^{-1}).$$

Then, the log prior density of $\widetilde{\lambda}$ is

$$\log p(\widetilde{\lambda}) = -\frac{1}{2}(\widetilde{\lambda}' H' H \widetilde{\lambda}).$$

Let

$$\Lambda_y = \mathrm{diag}(y_1 - y_1^*, y_2 - y_2^*, y_3 - y_3^*, \ldots, y_T - y_T^*).$$

Then, we have

$$-\frac{1}{2}(\omega_\lambda^2 \widetilde{\lambda}' \Lambda_y S_\pi^{-1} \Lambda_y \widetilde{\lambda}) - 2(\widetilde{\lambda}' \Lambda_y \omega_\lambda S_\pi^{-1}(\pi - \tau - \lambda_0 \Lambda_y)).$$

$\widetilde{\lambda}$ is distributed as

$$\mathcal{N}(\widehat{\widetilde{\lambda}}, K_{\widetilde{\lambda}}^{-1}),$$

where

$$\widehat{\widetilde{\lambda}} = K_{\widetilde{\lambda}}^{-1}(\Lambda_y \omega_\lambda S_\pi^{-1}(\pi - \tau - \lambda_0 \Lambda_y)), \qquad K_{\widetilde{\lambda}} = H'H + \omega_\lambda^2 \Lambda_y S_\pi^{-1} \Lambda_y.$$

Sample λ_0 and ω_λ

In this section, we construct the posterior sampler of λ_0 and ω_λ.

Let $X_\beta = (y - y^*, \Lambda_y \widetilde{\lambda})$ and $\beta = (\lambda_0, \omega_\lambda)'$, then, (1) can be written as

$$\pi - \tau = X_\beta \beta + \varepsilon^\pi.$$

Then, β is distributed as

$$\mathcal{N}(\widehat{\beta}, K_\beta^{-1}),$$

where

$$K_\beta = V_\beta^{-1} + X_\beta' S_\pi^{-1} X_\beta, \qquad \widehat{\beta} = K_\beta^{-1}(V_\beta^{-1}\beta_0 + X_\beta' S_\pi^{-1}(\pi - \tau)).$$

Sample ω_c^2

In this section, we show the posterior sampler of ω_c^2.

The conditional distribution of ω_c^2 is

$$\mathscr{IG}(\nu_c + \frac{T}{2}, S_{\omega_c} + \frac{1}{2}(c - X_\phi \phi)'(c - X_\phi \phi)).$$

Sample ω_{y^*}

In this section, we show the posterior sampler of ω_{y^*}.

The conditional density of ω_{y^*} is not a standard density, but it can be sampled by using Griddy-Gibbs .

M7

In equation (65), we link the blue chip 10 years inflation forecasts to trend inflation. Thus, the differences of estimation details between M7 and M5 are that M7 has a different sampler for τ_t and has two more samplers for $d = (d_0 \quad d_1)$ and ω_q^2. For brevity, we only display the estimation details for τ_t, $d = (d_0 \quad d_1)$, and ω_q^2.

Prior

$$\begin{array}{cccc}
\tau_0 = 0, & \tau_1 \sim \mathscr{N}(\tau_0, V_\tau e^{g_t}), & \lambda_0 \sim \mathscr{N}(a_0, V_{\lambda_0}), & e_0 = 0, \\
\omega_\lambda \sim \mathscr{N}(0, V_{\omega_\lambda}), & \omega_g \sim \mathscr{N}(0, V_{\omega_g}), & \omega_h \sim \mathscr{N}(0, V_{\omega_h}), & \omega_e^2 \sim \mathscr{IG}(\nu_e, S_{\omega_e}), \\
V_{\omega_h} = 0.2, & V_{\omega_g} = 0.2, & V_{\omega_\lambda} = 0.25^2, & \nu_e = 3, \\
V_{\lambda_0} = 0.25^2, & V_\tau = 10, & V_g = 10, & \nu_\nu = 3, \\
a_0 = -0.25, & V_\beta = (V_{\lambda_0}, V_{\omega_\lambda}), & \widehat{\beta} = (a_0, 0), & \phi \sim \mathscr{N}(\phi_0, V_\phi), \\
\phi_0 = (0.5; 0.2), & \omega_q^2 \sim \mathscr{IG}(\nu_q, S_{\omega_q}), & \nu_q = 3, & S_{\omega_q} = 1 * (\nu_q - 1), \\
V_d = \mathrm{I}_2, & e_{-1} = 0, & \omega_\nu^2 \sim, \mathscr{IG}(\nu_\nu, S_{\omega_\nu}), & S_{\omega_e} = 1 * (\nu_e - 1), \\
S_{\omega_\nu} = 1 * (\nu_\nu - 1), & \mu_d = (0 \quad 1), & V_\phi = \mathrm{I}_2. &
\end{array}$$

Likelihood

In this section, we derive the densities of $\pi = (\pi_1, \ldots, \pi_T)'$, that will be used to construct the posterior sampler.

Let

$$\Lambda_\lambda = \operatorname{diag}(\lambda_0 + \omega_\lambda \widetilde{\lambda}_1, \lambda_0 + \omega_\lambda \widetilde{\lambda}_2, \lambda_0 + \omega_\lambda \widetilde{\lambda}_3, \ldots, \lambda_0 + \omega_\lambda \widetilde{\lambda}_T).$$

Then, we have

$$\pi - \tau - \Lambda_\lambda (u - \nu) = \varepsilon^\pi.$$

Then, the log conditional density of π is

$$\log p(\pi \mid \tau, u, \nu, c, \widetilde{\lambda}, \lambda_0, \omega_h, \omega_\lambda, \omega_\nu, \omega_c, \omega_q, d, h) \propto -\frac{1}{2}(\pi - \tau - \Lambda_\lambda (u - \nu))' S_\pi^{-1} (\pi - \tau - \Lambda_\lambda (u - \nu)),$$

where

$$S_\pi = \operatorname{diag}(e^{h_1}, e^{h_2}, e^{h_3}, \ldots, e^{h_T}).$$

Sampling τ

In this section, we construct the posterior sampler of τ.

Let

$$S_g = (V_\tau e^{g_1}, e^{g_2}, e^{g_3}, \ldots, e^{g_T}).$$

and

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & \ldots & 0 \\ -1 & 1 & 0 & 0 & \ldots & 0 \\ 0 & -1 & 1 & 0 & \ldots & 0 \\ 0 & 0 & -1 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & -1 & 1 \end{bmatrix}.$$

Then, we have

$$H\tau = \varepsilon^{\tau}.$$

Then,

$$\tau \sim \mathcal{N}(0, H^{-1}S_g H'^{-1}).$$

Then, the log prior density for τ is

$$\log p(\tau) = -\frac{1}{2}\tau' H' S_g^{-1} H\tau.$$

(72) can be written as

$$q = d_0 1_T + d_1\tau + \varepsilon^z, \qquad \varepsilon^z \sim \mathcal{N}(0, \omega_q^2 \mathrm{I}_T).$$

Therefore, we have

$$\log(p(q|d_0, d_1, \tau, \omega_q^2)) \propto -\frac{1}{2\omega_q^2 \mathrm{I}_T}(q - d_0 1_T - d_1\tau)'(q - d_0 1_T - d_1\tau).$$

Then, we have

$$\log p(\tau \mid \pi, u, \nu, e, \phi, \gamma, \omega_e, \omega_\nu, \widetilde{\lambda}, \lambda_0, \omega_h, \omega_\lambda, \omega_g, \omega_q, d, h, g)$$
$$\propto -\frac{1}{2}(\tau' S_\pi^{-1}\tau - 2\tau' S_\pi^{-1}(\pi - \Lambda_\lambda(u - \nu))) - \frac{1}{2}\tau' H' S_g^{-1} H\tau - \frac{1}{2}(\tau' \frac{d_1^2}{\omega_q^2 I_T}\tau - 2\tau' \frac{d_1(q - d_0 1_T)}{\omega_q^2}).$$

Then, the conditional distribution of τ is

$$\tau \sim \mathcal{N}(\widehat{\tau}, K_\tau^{-1}),$$

where

$$\widehat{\tau} = K_\tau^{-1}((S_\pi^{-1})(\pi - \Lambda_\lambda(u-\nu)) + \frac{d_1(q - d_0 1_T)}{\omega_q^2}), \qquad K_\tau = S_\pi^{-1} + H' S_g^{-1} H + \frac{d_1^2}{\omega_q^2 \mathrm{I}_T}.$$

Sampling d

In this section, we construct the posterior sampler of d. Let

$$X_\tau = (1_T \quad \tau),$$
$$d = (d_0 \quad d_1).$$

Then (72) can be written as

$$q = X_\tau d + \varepsilon^q.$$

Then we have

$$\log(p(q|d_0, d_1, \tau, \omega_q^2)) \propto -\frac{1}{2\omega_q^2}(q - X_\tau d)'(q - X_\tau d).$$

we also have

$$\log(p(d)) \propto -\frac{1}{2}(d - \mu_d)' V_d^{-1}(d - \mu_d).$$

Then, the posterior d is distributed as $\mathcal{N}(\widehat{d}, K_d^{-1})$, where

$$K_d = (\frac{X_\tau' X_\tau}{\omega_q^2} + V_d^{-1}),$$
$$\widehat{d} = K_d^{-1}(\frac{X_\tau' q}{\omega_q^2} + V_d^{-1}\mu_d).$$

Sampling ω_q^2

In this section, we show the posterior sampler of ω_q^2.

$$\omega_q^2 \sim \mathscr{I}\mathscr{G}(\nu_{\omega_q^2} + \frac{T}{2}, S_{\omega_q} + \frac{1}{2}\sum_{t=1}^{T}\varepsilon_{q,t}^2).$$

Chapter 3

Bubbles and Crises: Replicating the Results of Anundsen et al. (2016)

3.1 Introduction

There is a growing literature of early warning systems for financial crises (e.g., Alessi and Detken, 2017; Bussiere and Fratzscher, 2006; Büyükkarabacak and Valev, 2010; Jordà et al., 2015a,b; Schularick and Taylor, 2012). In particular, Anundsen et al. (2016) use a dataset covering 20 countries and a period spanning 1976 to 2014 to assess the probability of a financial crisis—specifically, the likelihood that an economy is in a pre-crisis state. They contribute to the early warning system literature by finding that both the household credit-to-GDP gap and the NFE credit-to-GDP gap affect the likelihood of a financial crisis. In addition, global housing market development can affect domestic financial stability. Finally, their measures of exuberance in housing and credit markets have predictive power of financial stability.

This chapter first narrowly replicates the main results of Anundsen et al. (2016) using their dataset and Matlab. Next, we widely replicate their results by using TVP probit models. Many papers highlight the empirical importance of TVPs (e.g., Cai et al., 2000; Canova, 1993; Chan and Eisenstat, 2015; Cogley and Sargent, 2005; Koop and Korobilis, 2013; Koop and Potter, 2007). However, few papers consider TVP models in the early warning system literature. Thus, to fill this gap, this chapter estimates a set of TVP probit models and assesses their performance relative to their constant parameter counterparts.

The narrow replication produces the same results as those of Anundsen et al. (2016). In particular, we obtain identical out-of-sample forecasting performance of their models. The wide replication has three main findings. First, the coefficients of financial crisis indicators are quite stable over time, but the time variation in the country dummies is considerable. Second, most of the main finding of Anundsen et al. (2016) are robust to using a TVP approach. The important role of the household credit-to-GDP gap and global housing market developments is very robust to using a TVP approach. The effect of exuberance measures on the probability of a financial crisis also remains robust to using a TVP approach. However, the evidence that the NFE credit-to-GDP gap significantly affects the probability of a financial crisis is mixed across different specifications. Third, allowing for time-varying parameters of probit models can considerably improve the in-sample model fit and out-of-sample forecasting performance based on an expanding window forecasting exercise.

3.2 Results

In this section, we first narrowly replicate the main results of Anundsen et al. (2016) using maximum likelihood estimation and reproduce an out-of-sample forecasting performance for their models.[1] Second, we widely replicate their results by using a set of TVP probit models and show the out-of-sample forecasting performance of these models.

[1]Their main results are summarised in Table II of Anundsen et al. (2016).

3.2.1 Narrow Replication

To assess the likelihood that the economy is in a pre-crisis state, Anundsen et al. (2016) estimate a standard logit model using the maximum likelihood method. Their dataset covers 20 countries and the period spanning 1976 to 2014 and includes a dummy variable for the pre-crisis state and financial crisis indicators such as the credit-to-GDP gap (see Section 3.4).

We follow Anundsen et al. (2016) to use the area under the receiver operating characteristic (AUROC) curve as the main model evaluation criterion. A model issues a crisis signal if the estimated probability of crisis exceeds some threshold level, c, otherwise not. The true positive rate $TPR(c)$ is the proportion of correct crisis warning signals. The false positive rate $FPR(c)$ is the proportion of false crisis warning signals. The receiver operating characteristic curve (ROC) is the plot of $TPR(c)$ against $FPR(c)$ for all threshold parameters $c \in [0,1]$. Among the two models, given the same $FPR(c)$, if a model has a higher AUROC, it has a higher $TPR(c)$. Therefore, the model with a higher AUROC is preferred. AUROC is widely used to compare the predictive performance of alternative early warning system models (e.g., Berge and Jordà, 2011; Drehmann and Juselius, 2014; Jordà and Taylor, 2011, 2012).

We use the model and dataset of Anundsen et al. (2016) to reproduce their results. The computations are implemented using Matlab. The main results are identical to those in Anundsen et al. (2016)— specifically, the coefficients of explanatory variables and the forecasting performance measured as in-sample and out-of-sample AUROC are identical to theirs (see Section 3.4).[2] From the results, this narrow replication exercise shows that Anundsen et al. (2016)'s results are robust to using Matlab instead of Stata.

[2] The standard errors are similar, though not exactly the same as those in Anundsen et al. (2016). The differences may be attributed to using different software. The different standard errors are reported in Section 3.4.

3.2.2 Wide Replication

Considering the potential time-varying nature of the underlying structure in the economy, we widely replicate the results of Anundsen et al. (2016) using a set of TVP probit models. The specifications of these models are the same as their corresponding constant parameter logit models, but the parameters are allowed to evolve over time as a random walk. The details of the TVP-models are provided in Section 3.4.

Section 3.4 also shows the estimated time-varying coefficients of the TVP probit models. First, there is considerable time variation in the coefficients of country dummy variables. However, the coefficients of determinants of financial crises are stable.[3] This time variation may capture the dynamics of time-varying unobserved country heterogeneity. The changing unobserved country heterogeneity may approximate an important omitted variable in predicting financial crises, such as the changing country-specific institution or debt structure. Second, the effect of the household credit-to-GDP gap is significant at the 5% significance level across different specifications for all the sample periods. Anundsen et al. (2016) also find the significant effect of the household credit-to-GDP gap. Therefore, we conclude that the important role of household credit- to-GDP gap in affecting the likelihood of a crisis is robust to using TVP probit model. Third, the evidence that the NFE credit-to- GDP gap affects the probability of a financial crisis is mixed across different specifications. The effect of the NFE credit-to-GDP gap is significant at the 10% significance level for TVP Model 1 and TVP Model 2, but not significant for TVP Model 3 and TVP Model 4.[4] Fourth, the effect of the global house price-to-income gap on the probability of a financial crisis is

[3]Though the coefficients of economic variables are quite stable over time, the TVP model approach has an advantage over the constant parameter model in providing more precise estimated parameters. For example, the standard error of the coefficient of the GDP gap is around 5 in TVP Model 4, but in its corresponding constant parameter model, the standard error of the coefficient of the GDP gap is 13.27. Given the smaller standard errors, a stronger conclusion can be made using TVP-models.

[4]In Section 3.4 , we estimate the constant parameter logit model of Anundsen et al. (2016) using balanced panel data. In this case, the effect of the NFE credit-to-GDP gap is also significant for their Model 1 and Model 2, but not significant for their Model 3 and Model 4. This shows that when using balanced panel data, the significance of the coefficient of the NFE credit-to-GDP gap is more varying.

significant at the 5% significance level for all the sample periods. This is consistent with Anundsen et al. (2016). Fifth, the effects of the exuberance house price-to-income gap and the exuberance credit-to-GDP are significant at the 5% significance level for most of the sample periods. This is also consistent with Anundsen et al. (2016).

The in-sample fit of these models is improved considerably by allowing for TVPs. The AUROCs of TVP models 1–4, are 0.988, 0.988, 0.989, and 0.994, respectively. They are much higher than the AUROCs of the corresponding constant parameter logit model. The better in-sample fit in TVP models may be due to overfitting. Therefore, we do the expanding window forecasting exercise to access the out-of-sample forecasting performance. The evaluation period is 2000-2014 and the model is re-estimated for every new forecast.

Anundsen et al. (2016) do not use expanding window forecasting exercise to evaluate the out-of-sample forecasting performance of their models due to their unbalanced panel dataset. To compare the out-of-sample forecasting performance of their models with those of TVP probit models under the expanding window exercise, we estimate the models of Anundsen et al. (2016) using the balanced panel dataset. Post-crisis bias may lead to a poor in-sample fit of constant parameter logit models. However, we include observations that during or immediately after crisis only slightly change the in-sample fit of their constant logit models (see Section 3.4), measured as AUROC and pseudo-R^2.

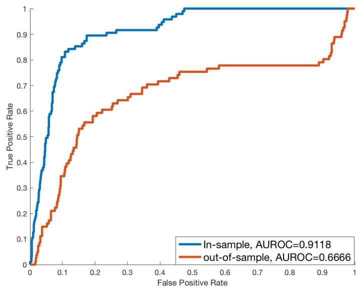
Table 3.1 shows that TVP probit models generally outperform constant parameter logit models used by Anundsen et al. (2016) for out-of-sample forecasts. TVP Model 2 and TVP Model 3 outperform their corresponding constant parameter models for all the forecasting horizons. TVP Model 1 outperforms its corresponding constant parameter logit model at one-ahead and two-ahead steps. The performance of TVP Model 4 with the TVP is not significantly different from its corresponding constant parameter logit model. Since most of the other parameters are relatively stable, allowing time variation in country dummies would contribute to the better in-sample fit and performance of the expanding window forecasting exercise.

Table 3.1: Area under the Receiver Operating Characteristic of Time-Varying Parameter Probit Models and Constant Logit Models with Alternative Specifications.

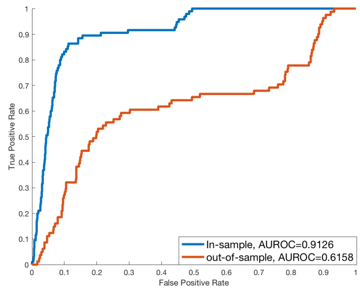
Time-Varying Parameter Probit Model				
	One-step	Two-step	Three-step	Four-step
TVP Model 1	0.904**	0.840**	0.785	0.735
TVP Model 2	0.919**	0.848**	0.797**	0.735**
TVP Model 3	0.923**	0.863**	0.798**	0.741**
TVP Model 4	0.908	0.830	0.765	0.702
Constant Parameter Logit Model				
Model 1	0.781	0.751	0.705	0.683
Model 2	0.805	0.743	0.680	0.611
Model 3	0.829	0.774	0.722	0.662
Model 4	0.867	0.800	0.723	0.657

** Means that the AUROC of the model under consideration is statistically different from its corresponding constant parameter model at the 5% significance level.

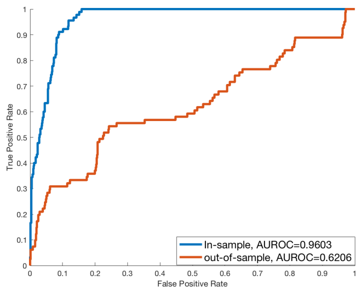
Then, we do same forecasting exercise as in Anundsen et al. (2016) using TVP probit models. The evaluation period covers the period from 2000 to 2014. I use the estimated coefficients at the end of the training sample to compute out-of-sample forecasts. From Figure 3.1, we find that constant parameter logit models have higher out-of-sample forecasting AUROC than do the TVP probit models. Therefore, in this case, the out-of-sample forecasting performance of the constant parameter logit models is better than that of the TVP probit models. The better performance of the constant parameter logit models may be attributed to using more information.



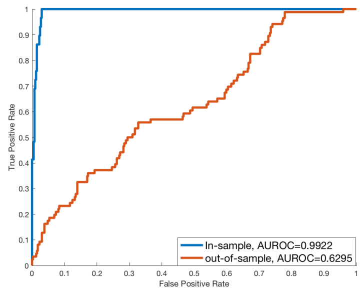
(a) Model 1



(b) Model 2



(c) Model 3



(d) Model 4

Figure 3.1: Out-of-Sample Forecasting Receiver Operating Characteristic Curve of the Time-Varying Parameter Probit Models.

3.3 Conclusions

This chapter replicates the results of Anundsen et al. (2016) in both a narrow and wide sense. We establish the narrow replication by reproducing the same results as theirs by using Matlab. First, the wide replication shows that first, the coefficients of financial crisis indicators are quite stable over time, but the country dummies have considerable time variation. Second, through the lens of the TVP probit models, most of the main findings of Anundsen et al. (2016) are robust to using a TVP approach, but the evidence that the NFE credit-to-GDP gap significantly affects the probability of a financial crisis is mixed across different specifications. Third, based on expanding window forecasting exercise, the warning system models' in-sample model fit and out-of-sample forecasting performance based on expanding window forecasting exercise can be considerably improved by allowing for

TVPs.

3.4 Appendix

3.4.1 Models

Constant Parameter Logit Model

The constant parameter logit model, which is used by Anundsen et al. (2016), is specified as follows:

$$\text{logit}(p_{it}) = a_i + \beta x_{i,t} + e_{it}, \tag{3.1}$$

where p_{it} is the probability of a financial crisis at country i and time t, a_i represents the unobserved heterogeneity of country i, and $\mathbf{x_{it}}$ is a vector of explanatory variables. β is the coefficients vector that is specified to be common for all the countries.

Time-Varying Parameter Probit Model

Next, we consider a TVP probit model using an equivalent latent variable representation. Specifically, let

$$Z_{it} \sim \mathcal{N}(x_{it}'\beta_t, 1), \tag{3.2}$$

where β_t is a $K \times 1$ vector. To avoid over-parameterisation and to be consistent with the model of Anundsen et al. (2016) displayed in equation (1), β_t is specified to be common for all the countries.

These latent variables are then linked to the observed binary variables Y_{it} as follows:

$$Y_{it} = \begin{cases} 1, Z_{it} > 0 \\ 0, Z_{it} \leqslant 0 \end{cases} . \tag{3.3}$$

With latent variable Z_{it}, we then have

$$Z_{it} = x_{it}{}'\beta_t + \varepsilon_{it}, \qquad \varepsilon_{it} \sim \mathcal{N}(0,1), \tag{3.4}$$

where $t = 1 \dots T$, $i = 1 \dots N$, β_t is a $K \times 1$ vector, x'_{it} is a $1 \times K$ vector, $x'_{it} = (\text{Dummy1}, \text{Dummy2}, \dots, \text{Dummy}$

Stack over i, we have

$$Z_t = X_t\beta_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, I_N), \tag{3.5}$$

$$\beta_t = \beta_{t-1} + u_t \qquad u_t \sim \mathcal{N}(0, Q), \tag{3.6}$$

$$\beta_0 \sim \mathcal{N}(a_0, B_0), \tag{3.7}$$

where Z_t is $N \times 1$, β_t is $K \times 1$ vector, and X_t is a

$$\begin{pmatrix} x'_{1t} \\ x'_{2t} \\ \vdots \\ x'_{Nt} \end{pmatrix} \tag{3.8}$$

$N \times K$ matrix, $Q = \text{diag}(q_1, \dots, q_P)$.

Then, stack over t, we have

$$Z = X\beta + \varepsilon, \qquad \varepsilon \sim \mathcal{N}(0, I_{NT}) \tag{3.9}$$

where Z is $NT \times 1$, β is $TK \times 1$ vector, and X is a $NT \times TK$ matrix,

$$\begin{pmatrix} X_1 & 0 & \dots & 0 \\ 0 & X_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & \dots & \dots & X_T \end{pmatrix}. \tag{3.10}$$

The model is estimated using Bayesian methods. Specifically, we use the precision sampler technique of Chan and Jeliazkov (2009) to estimate this set of models. The results are obtained by sampling 200,000 draws and discarding the first 100,000. The estimation details are shown as the following.

Sampler Z_{it}

$$(Z_{it}|y_{it} = 1; \beta_t) \sim TN_{(0,\infty)}(x_{it}{}'\beta_t, 1) \tag{3.11}$$

and

$$(Z_{it}|y_{it} = 0; \beta_t) \sim TN_{(-\infty,0)}(x_{it}{}'\beta_t, 1), \tag{3.12}$$

Sampler β

Let

$$H = \left\{ \begin{matrix} I_K & 0 & 0 & \dots & 0 \\ -I_K & I_K & 0 & \dots & 0 \\ 0 & -I_K & I_K & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & -I_K & I_K \end{matrix} \right)$$

Then (7) can be written as

$$H\beta = \widetilde{\alpha}_\beta + u, \tag{3.13}$$

where $u \sim \mathscr{N}(0, I_{NT} \otimes Q)$, $\widetilde{\alpha}_\beta = (\beta_0', 0, \ldots, 0)'$ The prior of β is given by

$$(\beta|\beta_0, Q) \sim \mathscr{N}(1_{NT} \otimes \beta_0, (H'(I_{NT} \otimes Q^{-1})H)^{-1})$$

Then,

$$(\beta|\beta_0, Q, Z, Y) \sim \mathscr{N}(\widehat{\beta}, K_\beta^{-1}),$$

where

$$\begin{aligned} K_\beta &= H'(I_{NT} \otimes Q^{-1})H + X'X, \\ \widehat{\beta} &= K_\beta^{-1}(H'(I_{NT} \otimes Q^{-1})H(1_{NT} \otimes \beta_0) + X'z) \end{aligned}$$

Sampler q_p

$$q_p \sim \mathscr{IG}(\nu + \frac{T}{2}, S + \frac{1}{2}\sum_{t=1}^{T}(\beta_{it} - \beta_{i(t-1)}))$$

where $\mathrm{S} = 0.09^2 \times (12-1)$, $\nu = 12$.

Sampler β_0

$$(\beta_0|\beta, Q, Z, Y) \sim \mathscr{N}(\widehat{\beta}_0, K_{\beta_0}^{-1}),$$

where

$$K_{\beta_0} = B_0^{-1} + Q^{-1},$$
$$\widehat{\beta}_0 = K_{\beta_0}^{-1}(B_0^{-1}a_0 + Q^{-1}\beta_1)$$

where $a_0 = 0$ and $B_0 = I_K$

3.4.2 Data

we use the same dataset of Anundsen et al. (2016) that is available on the Journal of Applied Econometrics Data Archive. This dataset covers 20 countries and a period spanning1976 to 2014. To avoid post-crisis bias, we follow Anundsen et al. (2016) to omit all observations during or immediately after a crisis. [5]

The dependent variable is a dummy variable for the pre-crisis state, defined as

$$Y_{it} = \begin{cases} 1, & \text{Financial crisis}_{i,t+k} = 1, \quad k \in [5,12] \\ 0, & \text{Otherwise} \end{cases}.$$

The independent variables include private credit growth, household credit-to-GDP gap, NFE credit-to-GDP gap, house price-to-income gap, output gap, global credit/GDP gap, global house price-to-income gap, exuberance house price-to-income gap, exuberance credit/GDP, non-core funding gap, and equity ratio.

We also use the dataset of Anundsen et al. (2016) to estimate the TVP probit models. For the expanding window forecasting exercise in the next section, we need to ensure the panel dataset balanced. Therefore, we do not remove the observations during or immediately after a financial crisis and drop observations before 1981. Since the TVPs are estimated using

[5]Bussiere and Fratzscher (2006) note that macroeconomic variables and relationships may be different after a crisis, thus, using information from the period staying or immediately following a crisis to predict a crisis may lead to poor predictions.

current information, it is not necessary to remove the observations during or six quarters succeeding a crisis to avoid post-crisis bias.

3.4.3 Details of Results

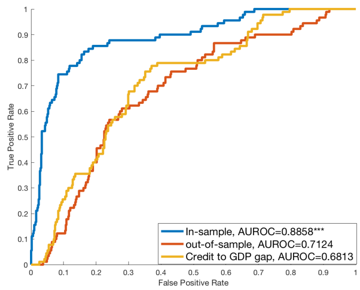
The Results of Replication of the Table II of Anundsen et al. (2016)

Table 3.2: Estimated Results from the Models in Anundsen et al. (2016).

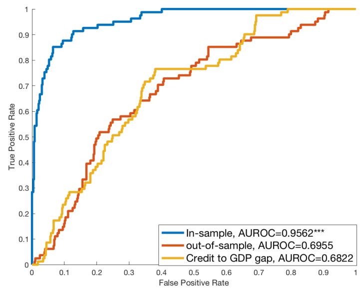
	Model 1	Model 2	Model 3	Model 4
Real credit growth	6.64**	8.66**	9.08**	10.19
	(2.881)	(3.34)	(4.19)	(5.65)
Household credit/GDP gap	17.46**	14.36**	15.28**	17.44**
	(3.578)	(4.29)	(5.43)	(7.32)
NFE credit/GDP gap	16.70**	23.83**	13.08**	14.57**
	(2.883)	(3.35)	(3.81)	(4.69)
Global credit/GDP gap		2.88	15.93**	-9.29
		(4.33)	(6.05)	(8.68)
Global house price-to-income gap		18.44**	19.61**	23.11**
		(2.94)	(3.71)	(4.85)
Exuberance house price to income			1.07**	2.12**
			(0.31)	(0.41)
Exuberance credit/GDP			1.85**	1.51**
			(0.30)	(0.35)
Non-core funding gap				56.37**
				(11.36)
Equity ratio				-59.67**
				(16.54)
House prices to income gap	12.29**	11.33**	9.53**	4.20**
	(1.608)	(1.76)	(1.94)	(1.85)
GDP gap	47.68**	43.16**	54.94**	45.35**
	(7.356)	(7.77)	(9.22)	(13.27)
Pseudo-R^2	0.290	0.334	0.404	0.446
AUROC	0.865	0.879	0.904	0.919

** Means that the coefficient is statistically different from 0 at the 5% significance level.

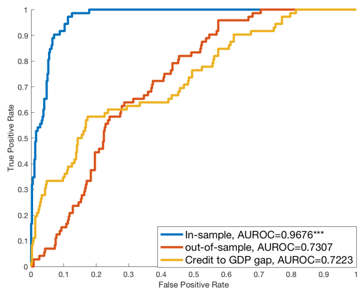
Results of replication of the Figure 6 of Anundsen et al. (2016)



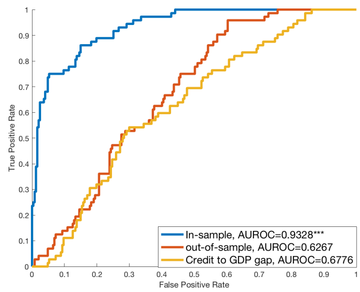
(a) Model 1



(b) Model 2



(c) Model 3

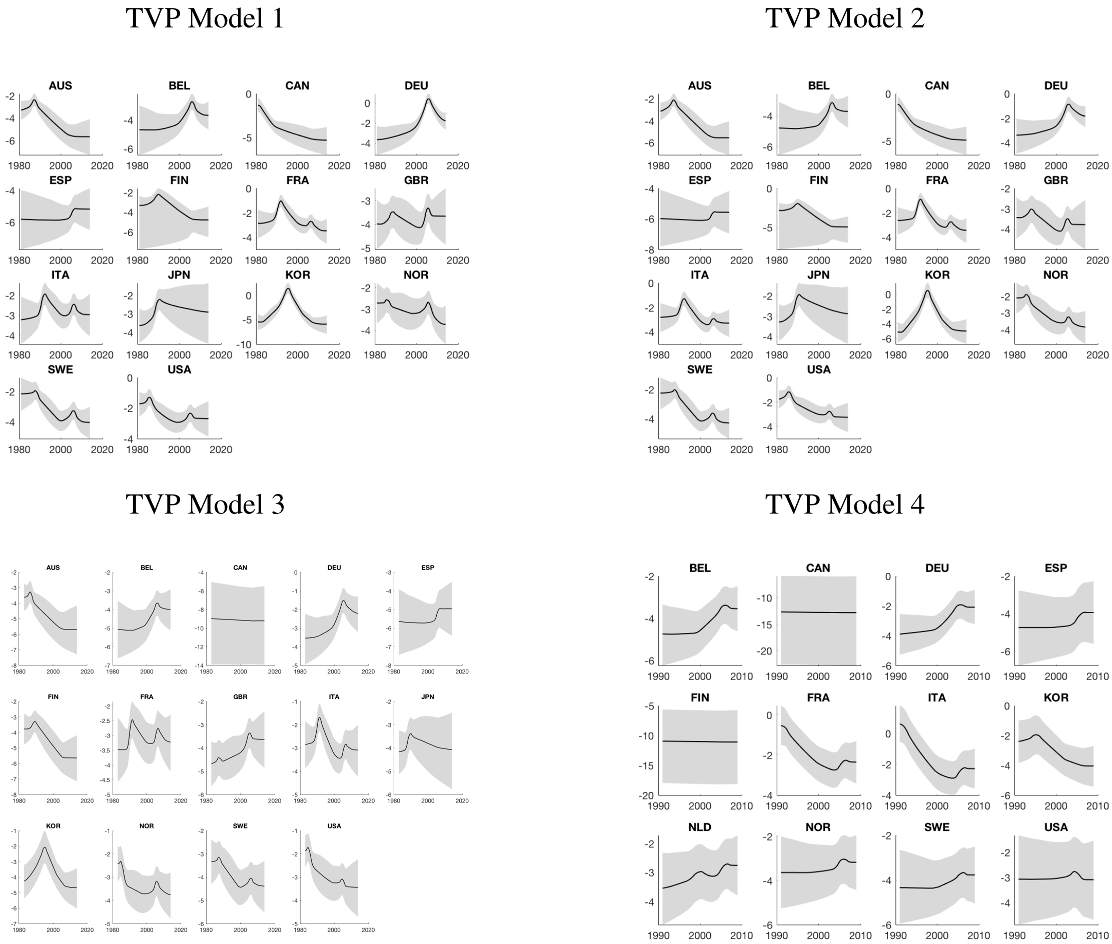


(d) Model 4

** Means that the AUROC of the model under consideration is statistically different to that of the credit-to-GDP gap at the 5% significance level.

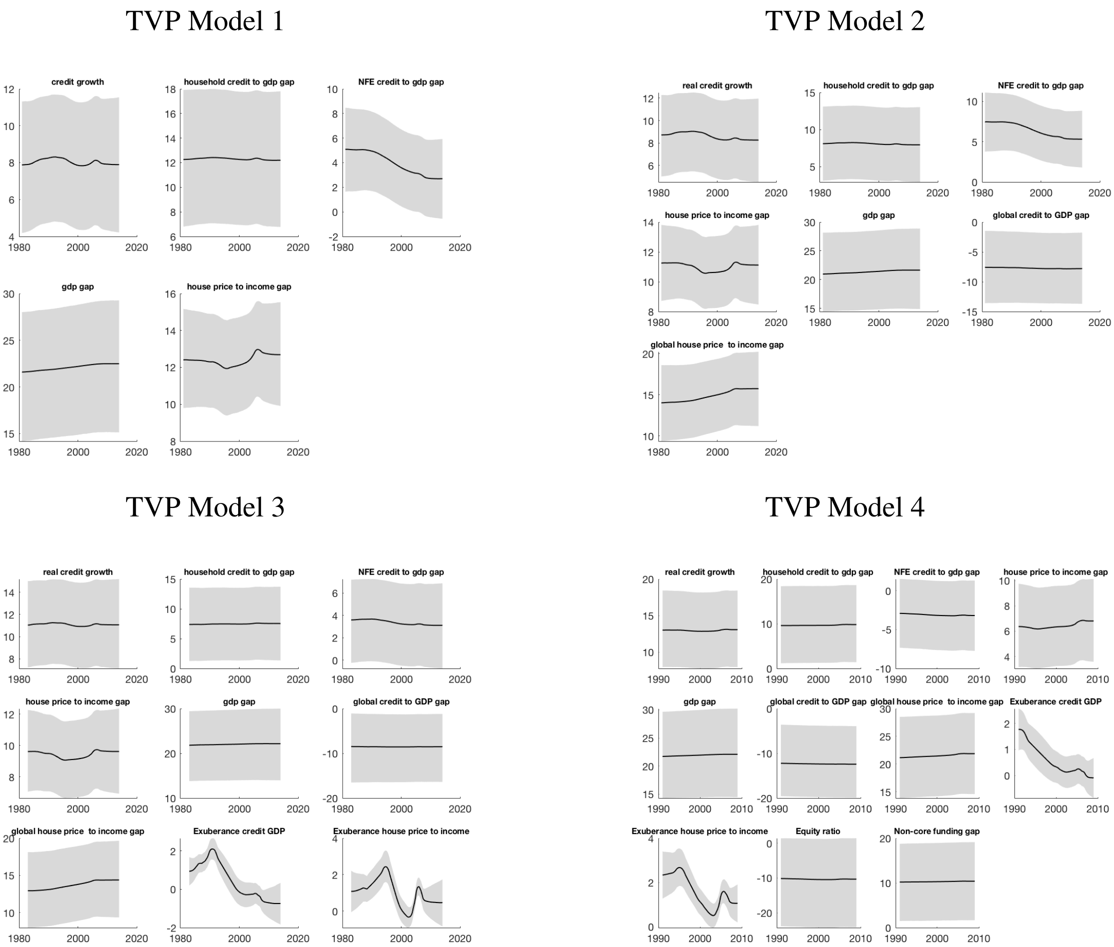
Figure 3.2: Out-of-Sample Forecasting Performance of the Models in Anundsen et al. (2016).

The Estimated Time-Varying Coefficients



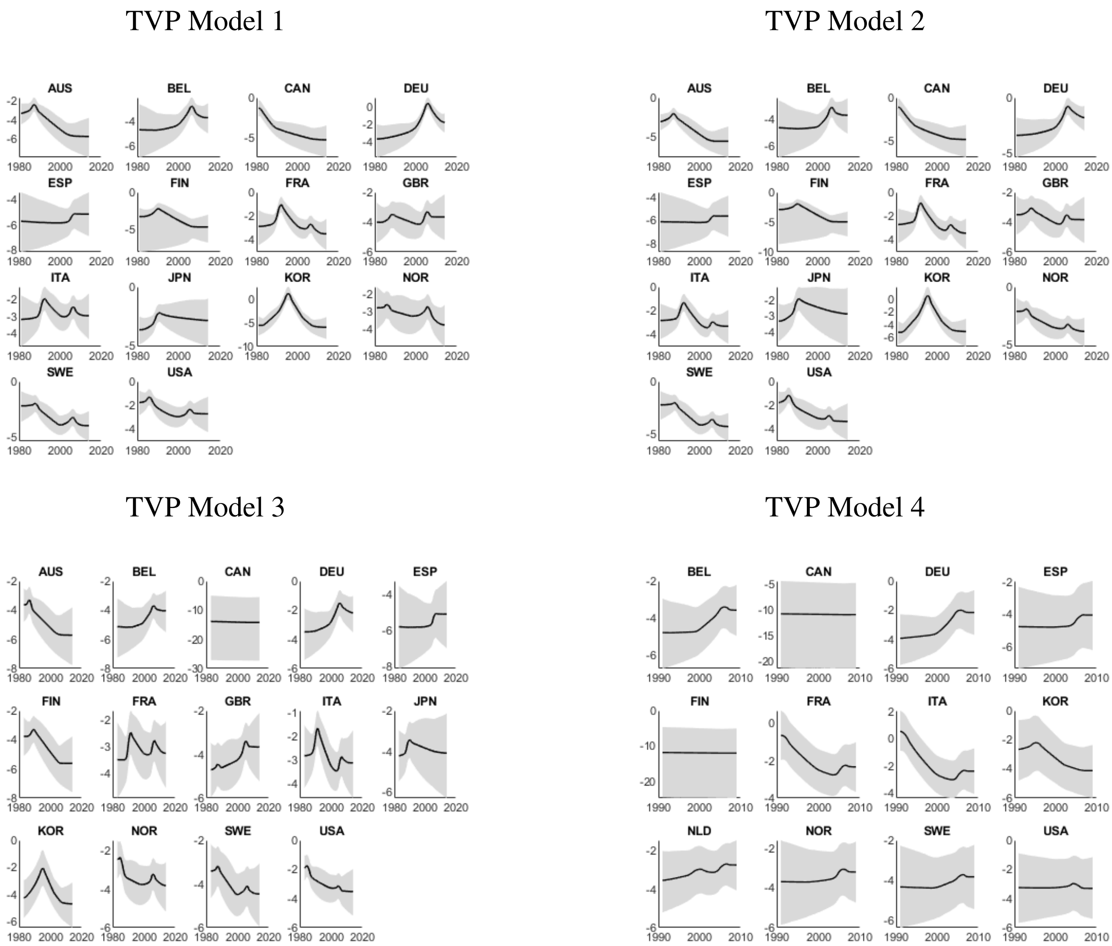
Note: Shaded areas denote 90% credible intervals.

Figure 3.3: The Plots of Coefficients of Country Dummy Variables.



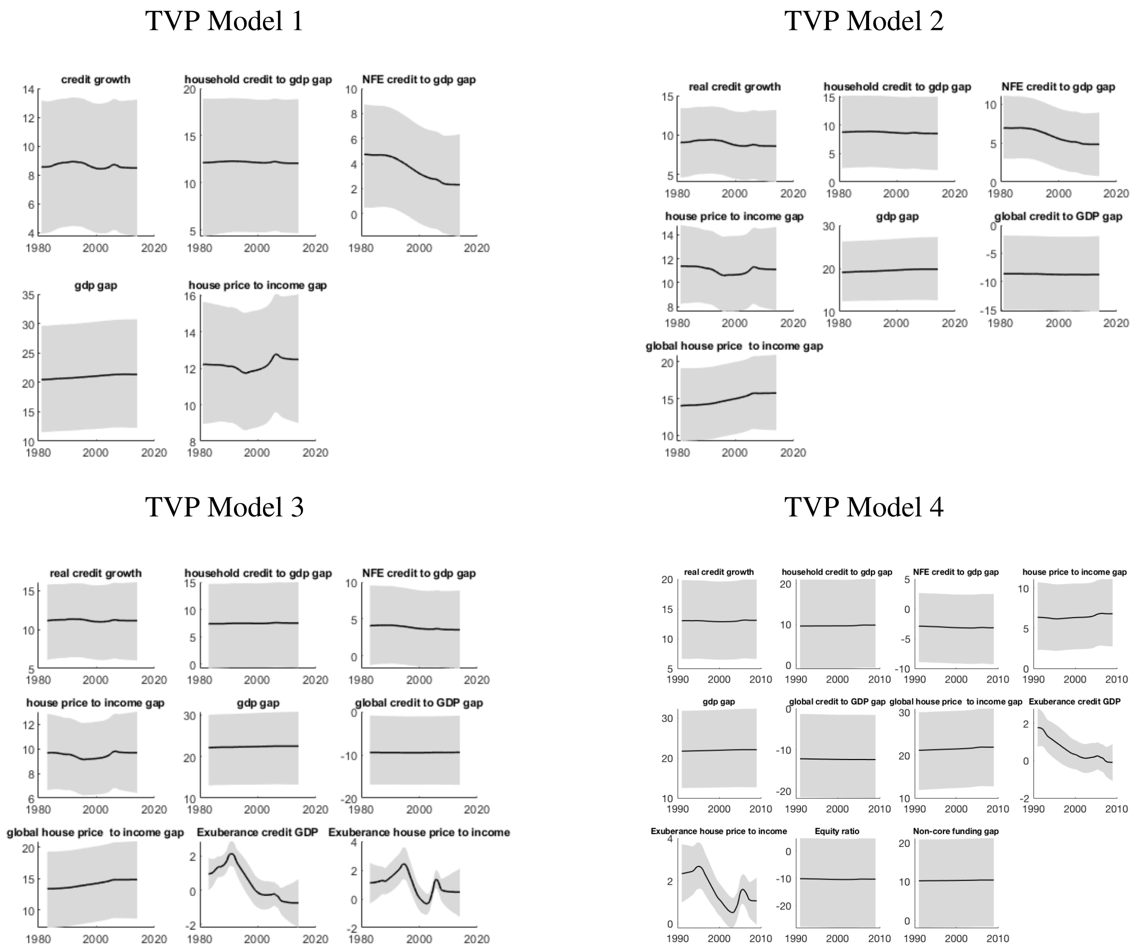
Note: Shaded areas denote 90% credible intervals.

Figure 3.4: The Plots of Coefficients of Determinants of Financial Crises.



Note: Shaded areas denote 95% credible intervals.

Figure 3.5: The Plots of Coefficients of Country Dummy Variables.



Note: Shaded areas denote 95% credible intervals.

Figure 3.6: The Plots of Coefficients of Determinants of Financial Crises.

Results of constant parameter logit models with dataset not omit observations during or immediately after a crisis

Table 3.3: Alternative Specification for the Determinants of Financial Crisis

	Model 1	Model 2	Model 3	Model 4
Real credit growth	12.665**	12.822**	16.391**	21.669*
	[6.338 18.991]	[6.289 19.356]	[8.923 23.858]	[7.968 35.369]
Household credit/GDP gap	8.280 **	6.442	5.149	8.093
	[0.668 15.893]	[-1.951 14.834]	[-4.385 14.683]	[-8.873 25.058]
NFE credit/GDP gap	10.936**	11.973**	4.350	-8.822
	[5.334 16.538]	[6.124 17.821]	[-1.898 10.598]	[-20.989 3.346]
Global credit/GDP gap		-1.56	1.628	-8.822
		[-10.40 7.29]	[-8.448 11.705]	[-20.989 3.346]
Global house price-to-income gap		14.65**	12.787**	30.960**
		[8.51 20.79]	[5.771 19.802]	[8.051 60.338]
Exuberance house price to income			1.352**	2.505**
			[0.785 1.918]	[1.660 3.350]
Exuberance credit/GDP			1.405**	1.493**
			[0.868 1.918]	[0.559 2.426]
Non-core funding gap				34.195**
				[8.051 60.338]
Equity ratio				-53.682 **
				[-103.381 -3.984]
House prices-to-income gap	13.980**	12.216 **	9.608**	2.093
	[10.499 17.462]	[8.652 15.781]	[6.201 13.015]	[-2.676 6.862]
GDP gap	42.427**	44.411**	46.961**	33.164 **
	[26.608 58.246]	[27.891 60.931]	[29.336 64.586]	[2.262 64.065]
Pseudo-R^2	0.276	0.296	0.362	0.451
AUROC	0.860	0.869	0.906	0.9307

Note: 95% confidence interval is in square brackets.

Chapter 4

United States Shocks and the Uncovered Interest Rate Parity

4.1 Introduction

The UIP is an important building block of standard open economy models. It states that, under no arbitrage, the returns from the investment on two economies should be equal, if the returns are converted into the same currency. UIP implies that, first, risk should be zero in the exchange rate market. Second, positive (negative) interest rate differentials should predict bilateral nominal exchange rate depreciation (appreciation). Specifically, UIP gives $(1+i^*_{t+h})E_t(S_{t+h})/S_t = 1+i_{t+h}$, where $E_t(.)$ is the conditional expectation using information up to time t, S_t is the nominal bilateral exchange rate, which is the price of one US dollar in terms of units of home currency. i^*_{t+h} is the US interest rate between time t and $t+h$, and i_{t+h} is the home bond interest rate between time t and $t+h$. Log-linearising this equation gives

$$E_t(\Delta s_{t+h}) = \alpha + \beta(i_{t+h} - i^*_{t+h})$$

where $\Delta s_{t+h} = s_{t+h} - s_t$, $s_t = \log S_t$, $\alpha = 0$ implies that zero risk premium, and then $\beta = 1$ implies that interest rate differential can perfectly predict changes in the bilateral nominal exchange rate.

The vast majority of papers empirically study this UIP relation by estimating the following benchmark model proposed by Fama (1984),

$$\Delta s_{t+h} = \alpha + \beta(i_{t+h} - i^*_{t+h}) + \varepsilon_{t+h}, \tag{4.1}$$

where $E_t(\varepsilon_{t+h}) = 0$. Based on equation (4.1), UIP can be tested in a standard linear regression framework with null hypothesis $H_0 : \alpha = 0$ and $\beta = 1$. However, this hypothesis is widely rejected by empirical evidence from numerous papers, (e.g., for a recent survey, see Rossi, 2013). This is known as 'UIP puzzle'. Ismailov and Rossi (2018) document another puzzling empirical fact: α and β are not only inconsistent with their predicted values but also time-varying or unstable over time.

Some explanations for the UIP puzzle have been proposed in the literature. These explanations include the presence of a time-varying risk premium (Fama, 1984; Li et al.), imprecise standard errors(Baillie and Bollerslev, 2000; Rossi, 2007), small samples (Chen and Tsang, 2013; Chinn and Meredith, 2004; Chinn and Quayyum, 2012), and rare disasters (Brunnermeier et al., 2008; Farhi and Gabaix, 2015).

Recently, two papers find that the UIP puzzle and unstable coefficients in equation 4.1 can be due to exchange rate uncertainty. Ismailov and Rossi (2018) find that the coefficients tend to be close to the values predicted by UIP at times of low uncertainty. They also find that the time variation is partly attributed to that UIP holds when uncertainty is low but does not when uncertainty is high. Ichiue and Koyama (2011) use a regime-switching model to examine how exchange rate volatility is related to the failure of UIP. First, they find that UIP tends to hold in a low volatility environment, and vice versa. Second, they find that the slope, β, and volatility regimes are partially dependent.

Given the unstable coefficients in the UIP regression and the relation between uncertainty, the two studies by Ismailov and Rossi (2018) and Ichiue and Koyama (2011) suggest that that time-varying coefficients and stochastic volatility should be considered for estimating equation 4.1. Conversely, there is a growing literature that highlights the empirical importance of TVP models in analysing the relations of time series (e.g., Canova, 1993; Chan and Eisenstat, 2015; Cogley and Sargent, 2005; Koop and Korobilis, 2013; Koop and Potter, 2007).

Surprisingly, few papers in the UIP literature jointly consider time-varying coefficients and the stochastic volatility in the UIP literature. To fill this gap, we study the UIP relation via a new UIP regression model that extends the model of Fama (1984) by allowing a time-varying coefficient model and stochastic volatility. We also explore which variables can explain the variation in the time-varying slope of the UIP regression, β_t. We follow an atheoretical or agnostic approach to find which variables can explain the variation in β_t. More specifically, we link β_t to a large data set of U.S. macroeconomic variables (McCracken and Ng, 2016) and find variables that are related to β_t. Joint estimation is infeasible due to using many variables and the excess Monte Carlo noise brought by sampling β_t.[1] Thus, we consider a two-stage estimation approach. Specifically, we first estimate β_t and then use least absolute shrinkage and selection operator (LASSO) to find which macroeconomic variables are relevant to β_t. A similar approach is commonly used in the literature. For example, Cecchetti et al. (2017) regress the estimates of trend inflation on other variables to find which variable can determine the trend inflation.

This chapter has three contributions to the UIP literature. First, we show that if we need to estimate the time-varying coefficients or stochastic volatility in a UIP regression model, we need to jointly estimate these two parameters. Compared to alternative models, we find that the TVC-SV can provide sensible weighting schemes on the observations, leading to a reliable inference on the time-varying coefficients and stochastic volatility. Second,

[1]The details of our argument can be seen in Section 4.2.3.

different from many other papers that show the UIP is usually violated, this chapter finds that the null hypothesis of UIP unconditionally cannot be rejected for several developed economies. Also, our finding does not support the time-varying risk premium explanation for the UIP puzzle, since the estimated α_t are tightly around zero. Conversely, the violation of UIP is mainly attributed to β_t. Third, we find that variables associated with the US labour market play the most important role in explaining the variation in β_t.

The remainder of this chapter is organized as follows. In Section 4.2, we introduce our TVC-SV for modelling the UIP relation, describe how we test UIP under our UIP regression model and discuss the variable selection method for find variables that can explain the variation in β_t. In Section 4.3, we discuss our results of the estimated TVPs, the test of UIP, and variable selection. Section 4.4 is the conclusion.

4.2 Empirical Methodology

In this section, we first introduce our model that features time-varying coefficients and stochastic volatility. Second, we show how we test the UIP hypothesis. Third, we discuss the method we use to find which macroeconomic variables explain the time-varying UIP coefficients.

4.2.1 The Model

Our TVC-SV is specified by

$$\begin{aligned} \Delta s_{t+h} &= \alpha_t + \beta_t (i_{t+h} - i^*_{t+h}) + \sigma_t \varepsilon_t, \quad t = 1, ..., T, \\ f_{t+1} &= \mu_f (1 - \phi_f) + \phi_f f_t + \sigma_f \eta_{f,t}, \\ f_1 &\sim N(\mu_f, \frac{\sigma_f^2}{1 - \phi_f}), \quad f \in \{\alpha, \beta, \log \sigma\}, \end{aligned} \tag{4.2}$$

where ε_t, $\eta_{\alpha,t}$, $\eta_{\beta,t}$ and, $\eta_{\log\sigma,t}$ are uncorrelated standard normal variates.[2]. The model is a (nonlinear) state space model and we estimate it using the standard Markov chain Monte Carlo (MCMC) method. Let θ collect all hyperparameters, the joint prior distribution is given by

$$p(\theta) = p(\mu_\alpha)p(\mu_\beta)p(\log\sigma)p(\phi_\alpha)p(\phi_\beta)p(\phi_{\log\sigma})p(\sigma_\alpha)p(\sigma_\beta)p(\sigma_{\log\sigma}).$$

For $f \in \{\alpha, \beta, \log\sigma\}$, we choose non-informative Gaussian prior for the unconditional mean, that is. $p(\mu_f) \overset{d.}{=} N(u_f, v)$ with $(u_\alpha, u_\beta, u_{\log\sigma}) = (0, 1, \mathrm{Var}[\Delta s_{t+h} - (i_{t+h} - i^*_{t+h})])$ and $v = 10$,; and a beta prior $p(\frac{1+\phi_f}{2}) \overset{d.}{=} Beta(a,b)$ with $a = 20$ and $b = 1.5$ commonly used in the Bayesian time series literature for the autoregression coefficient and; a non-informative inverse gamma prior $\sigma_f \overset{d.}{=} IG(\gamma, \delta)$ with $\gamma = 0.5$ and $\delta = 0.5$. The beta prior imposes persistence on the time evolution of latent processes, whereas the non-informative prior leaves ample room for allowing the data speak. The Bayesian sampling procedure for generating the posterior samples of the hyperparamters and latent processes are provided in Section 4.5.

4.2.2 Test for Unconditional Uncovered Interest Parity

Due to variation in the UIP coefficients, equation (4.1) can be thought of as an equilibrium relationship between the exchange rate and interest rate differential. Theoretically, no arbitrage has to hold in equilibrium. Thus, any local movements in UIP parameters have to show mean-reverting dynamics to the equilibrium. Further, we can determine which US macroeconomic variables explain the local movements (see Section 4.2.3). We can also study if the equilibrium relationship holds by testing if the unconditional mean of α_t and β_t equal to their theoretical values. For economies in our empirical study, α_t is tightly esti-

[2]We also consider the model with heavy-tailed disturbances, namely $\Delta s_{t+h} = \alpha_t + \beta_t(i_{t+h} - i^*_{t+h}) + \sigma_t\sqrt{w_t}\varepsilon_t$ with w_t an inverse gamma mixing variable. Our empirical results carry over without any changes, so we stick to the simpler framework in which the measurement equation has normal disturbance.

mated to show no time variation and that is close to zero, so our focus is simply to test the null hypothesis $H_0 : \mu_\beta = 1$ against $H_1 : \mu_\beta \neq 1$.

To this end, we propose to use the Bayes factor (BF) based on the Savage-Dickey density ratio (*SDDR*) (see e.g. Kass and Raftery, 1995). The BF calculates a data density ratio conditional on the alternative and null hypothesis respectively, respectively, or $BF_{10} = p(\widetilde{Y}_T, \widetilde{X}_T | H_1)/p(\widetilde{Y}_T, \widetilde{X}_T | H_0)$. It can be shown that

$$BF_{10} = SDDR = \frac{p(\mu_\beta = 1|\widetilde{Y}_T, \widetilde{X}_T)}{N(\mu_\beta = 1; u_\beta, v)};$$

that is, the ratio of the posterior and prior ordinate. The denominator can be readily computed, but the numerator cannot be computed analytically. Since we have a closed-form conditional posterior distribution of μ_β, we can use the Rao-Blackwellisation procedure of Gelfand et al. (1992) to construct a consistent and unbiased estimator of $p(\mu_\beta = 1|\widetilde{Y}_T, \widetilde{X}_T)$. It follows that

$$\widehat{p}(\mu_\beta|\widetilde{Y}_T, \widetilde{X}_T) = \frac{1}{S}\sum_{s=1}^{S} p(\mu_\beta|\theta^{(s)}_{-\mu_\beta}, \widetilde{Y}_T, \widetilde{X}_T) = \frac{1}{S}\sum_{s=1}^{S} p(\mu_\beta|\beta_1^{(s)}, ..., \beta_T^{(s)}, \phi_\beta^{(s)}, \sigma_\beta^{(s)}),$$

where S is the number of random draws from the posterior sample, superscript (s) indicates the s-th draw and; $\theta^{(s)}_{-\mu_\beta}$ is $\theta^{(s)}$ without $\mu_\beta^{(s)}$. In our case, the conditional posterior $p(\mu_\beta|\beta_1^{(s)}, ..., \beta_T^{(s)}, \phi_\beta^{(s)}, \sigma_\beta^{(s)})$ is Gaussian with mean $\underline{u_\beta}^{(s)}$ and variance $\underline{v_\beta}^{(s)}$ (see Section 4.5). Thus, a consistent and unbiased estimator of *SDDR* is

$$\widehat{SDDR} = \frac{1}{S}\sum_{s=1}^{S} \frac{N(\mu_\beta = 1; \underline{u_\beta}^{(s)}, \underline{v_\beta}^{(s)})}{N(\mu_\beta = 1; u_\beta, v)} = \frac{1}{S}\sum_{s=1}^{S} \sqrt{\frac{v}{\underline{v_\beta}^{(s)}}} \exp\left(\frac{-(\underline{u_\beta}^{(s)} - 1)^2}{2\underline{v_\beta}^{(s)}}\right).$$

$\widehat{SDDR}$ can then be used to determine statistical significance based on the scale reported by Kass and Raftery (1995).

4.2.3 Variable Selection

To answer the question of which U.S. macroeconomic variables can explain the variation in UIP parameters, one can modify the TVC-SV model in several ways. For example, f_t can be directly linked to a vector of zero mean stationary explanatory variables z_t that may include the lags of f_t and modelled by $f_t = z_t' \gamma_f + \sigma_f \eta_{f,t}$, and a distributed lag autoregressive dynamics with parameter vector γ_f (Dufour and Kiviet, 1998). This design, though appealing, becomes infeasible when the dimension of z_t is large as in our case, in which we use a high-dimensional U.S. macroeconomic dataset, and one has to opt for variable selection techniques. There is a large literature in Bayesian variable selection that uses some form of sparsity prior on the regression coefficients to generate sparse posterior (see *e.g.,* O'Hara et al., 2009 and Ghosh and Clyde, 2011). Diverging from this literature in which the dependent variable is the data, in our case, f_t is latent and thus different across each draw in the MCMC algorithm. The excess Monte Carlo noise brought by this stochasticity leads the Gibbs sampler for Bayesian variable selection to converge extremely slowly, if at all.

Alternatively, one can respect the autoregression specification of f_t but link it to another equation that treats f_t as the signal extracted from a linear combination of elements in z_t. This is to augment the TVC-SV model by $\Lambda_f z_t = f_t + \eta_t^*$ with the $1 \times K$ loading matrix, Λ_f, where K is the dimension of z_t. This design is related to the literature on inflation dynamics with surveyed inflation expectation. In this literature, $\Lambda_f = 1$ and z_t is a survey-based inflation expectation that aims to provide additional information to pin down the unobserved expectation process. In our case, z_t is high-dimensional, so immediately Λ_f cannot be identified. Further, as far as we know, it is unclear how Bayesian variable selection can be implemented in this setting.

Since these modifications are infeasible, we consider running a second stage estimation based on the posterior sample generated by the MCMC algorithm for the TVC-SV model. The posterior sample of UIP parameters α_t, β_t and $\log \sigma_t$ are drawn from $p(f_t | \widetilde{Y}_T, \widetilde{X}_T)$

where $f \in \{\alpha, \beta, \log\sigma\}$, $\widetilde{Y}_T = (\widetilde{y}_1, ..., \widetilde{y}_T)'$ with $\widetilde{y}_t = \Delta s_{t+h}$ and $\widetilde{X}_T = (\widetilde{x}_1, ..., \widetilde{x}_T)'$ with $\widetilde{x}_t = i_{t+h} - i^*_{t+h}$. Though the two-step approach is indirect as compared with the aforementioned approaches, it can be viewed as an approximation to exact inference. Let $Z = (z_1, ..., z_T)'$ and $F_T = (f_1, ..., f_T)'$ and suppose we are interested in the posterior distribution of γ_f, that is,

$$p(\gamma_f|\widetilde{Y}_T, \widetilde{X}_T, Z) = \int p(\gamma_f|F_T, \widetilde{Y}_T, \widetilde{X}_T, Z)p(F_T|\widetilde{Y}_T, \widetilde{X}_T, Z)dF_T. \tag{4.3}$$

If we assume the following two conditional independence assumptions: 1). γ depends on $\widetilde{Y}_T$ and $\widetilde{X}_T$ only through F_T and Z; and 2). F_T depends on Z only though $\widetilde{Y}_T$ and $\widetilde{X}_T$, the left-hand side of equation (4.3) can be unbiasedly estimated by

$$\frac{1}{M}\sum_{i=1}^{M} p(\gamma_f|F_T^{(i)}, Z), \quad F_T^{(i)} \sim p(F_T|\widetilde{Y}_T, \widetilde{X}_T).$$

$F_T^{(i)}$ can be taken from the MCMC outputs of the TVC-SV model. γ_f can be obtained directly from the linear regression

$$F_T^{(i)} = Z\gamma_f + \xi^{(i)}, \tag{4.4}$$

where $\xi^{(i)}$ is a normal error vector of size T. It is known that under certain conditions, both the least absolute shrinkage and selection operator (LASSO) and Bayesian variable selection achieve model selection consistency (see *e.g.* Zhao and Yu, 2006 and Casella et al., 2009). Thus, we make use of the easy computational procedure of LASSO for choosing relevant variables in (4.4). Another advantage of following the two-step procedure is that instead of running (4.4) with a randomly chosen $F_T^{(i)}$ from its posterior, we can observe the importance of some variables in explaining the variation in F_T changes along its quantile by directly examining the quantile functions of $(F_T|\widetilde{Y}_T, \widetilde{X}_T)$. To this end, regression coefficients are determined by the LASSO criterion function

$$\widehat{\gamma}_f^{(i)} = \arg\min_{\gamma_f} \frac{1}{2T}(F_T^{(i)} - Z\gamma_f)'(F_T^{(i)} - Z\gamma_f) + \lambda \sum_{i=1}^{K} |\gamma_i|,$$

where $i \in \{10, 25, 50, 75, 90\}$ indicates a certain percentile of interest. The shrinkage parameter λ is chosen such that no more than 30 variables in Z receive non-zero coefficients.

In practice, it is more sensible for variable selection and the determination of the statistical significance of US macroeconomic variables to be based on the filtering distribution $p(f_t|\widetilde{Y}_t, \widetilde{X}_t)$ instead of smoothed or posterior distribution. This is because samples from $p(f_t|\widetilde{Y}_T, \widetilde{X}_T)$ are functions of all data, including future information. This artificially creates endogeneity that biases the estimate of regression coefficients. Suppose we aim to sample from $p(\beta_t|\widetilde{Y}_t, \widetilde{X}_t)$, this can be easily combined with the MCMC algorithm detailed in Section 4.5. It follows that

$$p(\beta_t|\widetilde{Y}_t, \widetilde{X}_t) = \int_\theta \int_{\sigma_1,\ldots,\sigma_t} p(\beta_t|\theta, \sigma_1, \ldots, \sigma_t, \widetilde{Y}_t, \widetilde{X}_t) p(\sigma_1, \ldots, \sigma_t|\theta, \widetilde{Y}_t, \widetilde{X}_t) p(\theta|\widetilde{Y}_t, \widetilde{X}_t) d\theta d\sigma_1, \ldots, d\sigma_t.$$

To reduce Monte Carlo noise, we approximate $p(\theta|\widetilde{Y}_t, \widetilde{X}_t)$ via the posterior distribution $p(\theta|\widetilde{Y}_T, \widetilde{X}_T)$. Thus, an estimate of $p(\beta_t|\widetilde{Y}_T, \widetilde{X}_T)$ is given by

$$\frac{1}{NM} \sum_{n=1}^{N} \sum_{m=1}^{M} N(E_t^{n,m}(\beta_t), Var_t^{n,m}(\beta_t)), \;\; n = 1, \ldots, N, \;\; m = 1, \ldots, M,$$

where $E_t^{n,m}(\beta_t)$ and $Var_t^{n,m}(\beta_t)$ is the filtering mean and variance of β_t given by Kalman filter conditional on the n-th draw of hyperparameter $\theta^{(n)}$ and m-th draw of the sequence of volatility $\{\sigma_s^{(m)}\}_{s=1}^t$ from its filtering distribution, respectively. The filtering distribution of stochastic volatility comes from the Kalman filter conditional on some auxiliary mixture components (e.g., Kim et al., 1998; see Section 4.5 for details).[3]

[3]Though one can use a particle filter (Doucet and Johansen, 2009) to find the filtering distribution, our approach can be directly implemented within the MCMC algorithm, generating the desired distribution as soon as the MCMC terminates.

To ensure interpretability and equal scale, we normalise F_T and columns in Z before applying LASSO. Once the variables are selected, we can re-estimate the regression model (4.4) using least squares to determine the statistical significance of each selected variable with respect to the i-th percentile of $(F_t|\widetilde{Y}_t, \widetilde{X}_t)$.

4.3 Results

In this section, we will first argue the importance of jointly modelling time-varying coefficients and stochastic volatility in the inference on the TVPs in our model. In particular, we will show that modelling time-varying coefficients and stochastic volatility jointly can provide sensible weighting schemes on the observations and therefore relatively reliable inference on the TVPs α_t, β_t, and σ_t. Second, we will show the estimated time-varying coefficients for diverse economies and then show the test results of the UIP hypothesis from using the TVC-SV model. Finally, we will show which US macroeconomic variables explain β_t.

4.3.1 The Importance of Time-varying Coefficients and Stochastic Volatility

To show the importance of jointly modeling time-varying coefficients and stochastic volatility in the inference on the TVPs in the UIP model, we take the Canadian dollars as an example and estimate the UIP model under different specifications: a time-varying coefficient model with constant volatility (TVC), a constant coefficient model with stochastic volatility (CC-SV), and a TVC-SV.

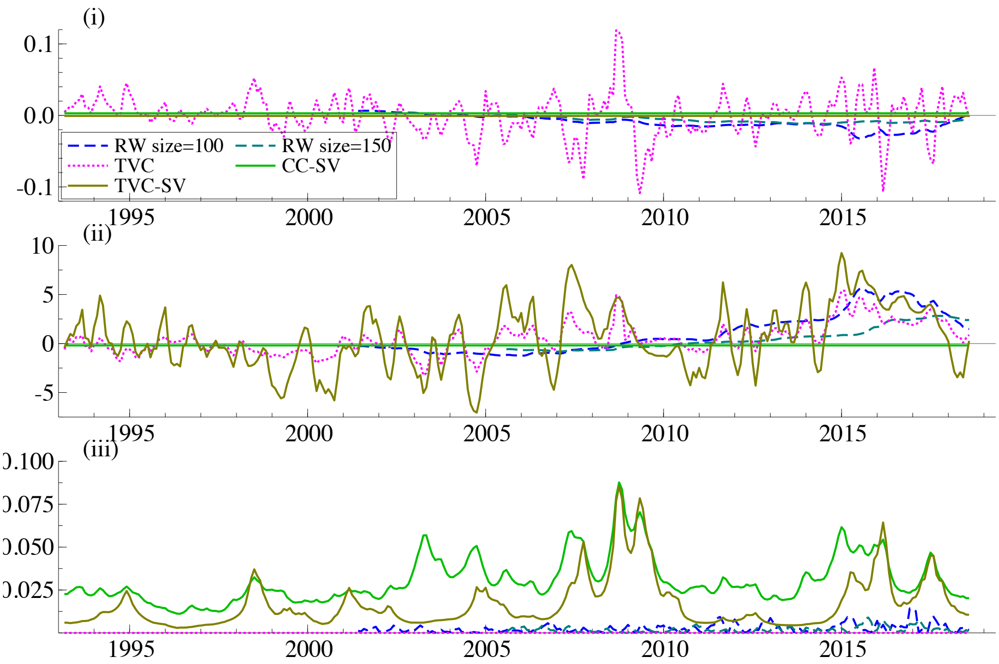


Figure 4.1: The Time Variation of Uncovered Interest Parameters for Canada. Note: (i): Estimate of α_t, (ii): Estimate β_t, (iii): Estimate of σ_t.

Figure 4.1 illustrates the estimates of α_t, β_t, and σ_t against time. First, if we consider stochastic volatility and assume the coefficients are constant and the estimates from CC-SV of both α_t and β_t are indistinguishable from zero. CC-SV has the highest volatility, σ_t, among all specifications as observed in the bottom panel of Figure 4.1. We have also tried relaxing the autoregressive dynamics assumption in the volatility process by running a generalised least squares with heteroskedasticity, and the results are similar. Second, if we consider time-varying coefficients and assume the volatility is constant, the estimated α_t and β_t from TVC are volatile and TVC gives essentially zero volatility, σ_t. Third, when account for time-varying coefficients and stochastic volatility jointly, TVC-SV gives essentially zero α_t, volatile β_t and σ_t.

To summarise our results so far, the estimated TVPs—α_t, β_t, and σ_t– are different under different models. If we were to make conclusions on the UIP TVPs, we might have reached very different results using different models. Thus, it is natural to ask why we have different estimated parameters under different models and which model provides reliable estimated TVPs.

The main reason can be that the weighting scheme on data points that is used to conduct the UIP regression varies with different model specifications. Specifically, in TVP models or models with latent processes, say β_t, if one ignores the hyperparameters which are functions of all data, the estimate of β_t at time t is effectively a function of data points around t, each receiving an observation weight. Since the models we considering are (conditionally) linear and Gaussian state space models, we can write

$$E(\beta_t|\widetilde{Y}_T, \widetilde{X}_T) = \sum_{j=1}^{T} \omega_{jt}(\widetilde{x}_j)\widetilde{y}_j,$$

where $\omega_{jt}(\widetilde{x}_j)$, a function of $\widetilde{x}_j$, is the weight associated with the posterior estimator $E(\beta_t|\widetilde{Y}_T, \widetilde{X}_T)$ and the j-th dependent variable (Durbin and Koopman, 2012). $\omega_{jt}(\widetilde{x}_j)$ can be different with different model specifications.

To investigate this possibility, we observe the effect that accounting for the time-varying coefficients and stochastic volatility has on the observation weights used to estimate β_t in June of 2003 and November of 2008.

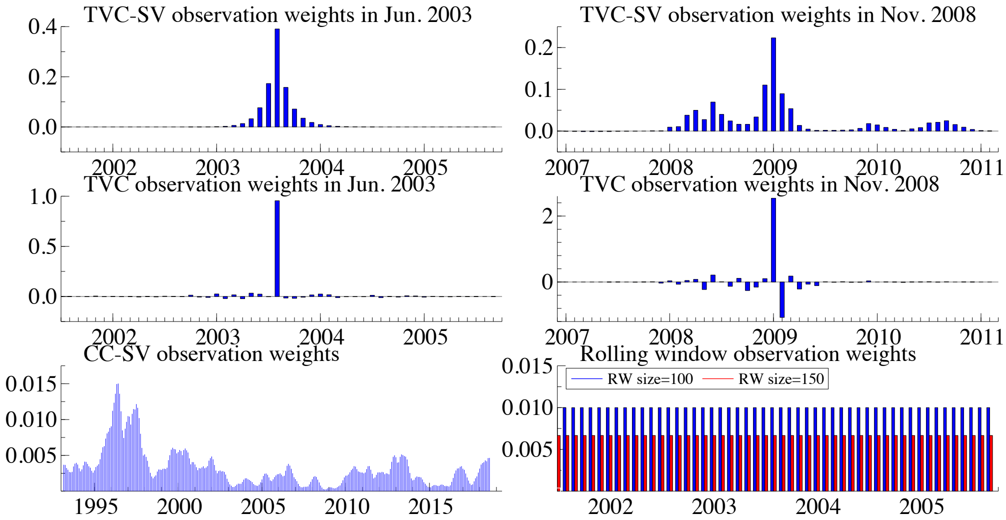


Figure 4.2: Observation Weights for Estimating β_t under Different Specifications. Note: For TVC-SV and CC-SV, models are cast into conditionally linear and Gaussian state space form with stochastic volatility evaluated at its posterior mean. Observation weights for estimating β_t under TVC-SV, CC-SV and TVC models are computed using output from the Kalman filter and smoother. The observation weights for rolling window estimations are constant due to equal weighting.

First, Figure 4.2 shows that during the high-volatility period identified by CC-SV model such as 2003, 2008 and 2016, observations receive near zeros weights, and effectively most of the weights are assigned to the low-volatility period prior to 2003. This may lead to bias if, during the low-volatility period, the variation in interest rate differentials is not informative about the variation in the exchange rate despite the bias introduced by ignoring parameter uncertainty.

Second, if we switch off stochastic volatility and switch on time-varying coefficients, we observe in Figure 4.1 that TVC estimates zero variance for the idiosyncratic errors, overfitting the data with local movements of UIP coefficients. This explains the near full weight on the observation at t shown in Figure 4.2 for both periods. Thus, changes in the UIP relationship are fully absorbed by local movements in coefficients or volatility, if one considers time-varying coefficients or stochastic volatility, respectively.

Third, compared with the previous two cases, nesting together the time-varying coefficient and stochastic volatility provides more sensible weighting schemes on observations. TVC-SV puts exponentially declining weight on observations away from t during normal times such as June 2003 and puts more weight on observations further away from t during volatile times such as November 2008. The fact that TVC-SV identifies both non-overfitting TVPs and non-zero stochastic volatility highlights the importance of both time-varying coefficient and stochastic volatility when we estimate the TVP in the UIP regression.

Our results indicate that how much information to discount when estimating the model should be subject to the economic environment. In particular, during volatile times, the local estimate of β_t should rely on more observations prior to and after t or down-weight the relative weights around t. However, if the economic condition is stable, the local movement of β_t can safely reply on observations tightly around t. Regarding Ismailov and Rossi, 2018, the observation weights for the rolling window estimations are constant due to equation equal weighting as shown in Figure 4.2. This may lead to biased estimated time-varying UIP parameters shown in Figure 4.1, in which the rolling window estimates of α_t, β_t and σ_t are smooth and generally close to zero.

4.3.2 Estimated Time-Varying Coefficients and Stochastic Volatility

In this section, we show the time-varying coefficient, β_t, and stochastic volatility, σ_t, for 11 economies—Canada, Denmark, the European Union (EU), Japan, Norway, New Zealand, South Africa, Sweden, Switzerland, and the United Kingdom (UK), –in Figure 4.3 and Figure 4.4, respectively. [4]

Figure 4.3 shows that for all the economies we consider, β_t is unstable and has substantial time variation in β_t. In addition, β_t seems to be a mean-reverting dynamic. From Figure 4.4, for all the economies shown here, stochastic volatility is also volatile and has substantial time variation. Regarding the relationship between β_t and σ_t, for all the economies here, in

[4] For all the economies we consider here, the estimated α_t is stable and close to 0. For brevity, we omit the Figure of α_t.

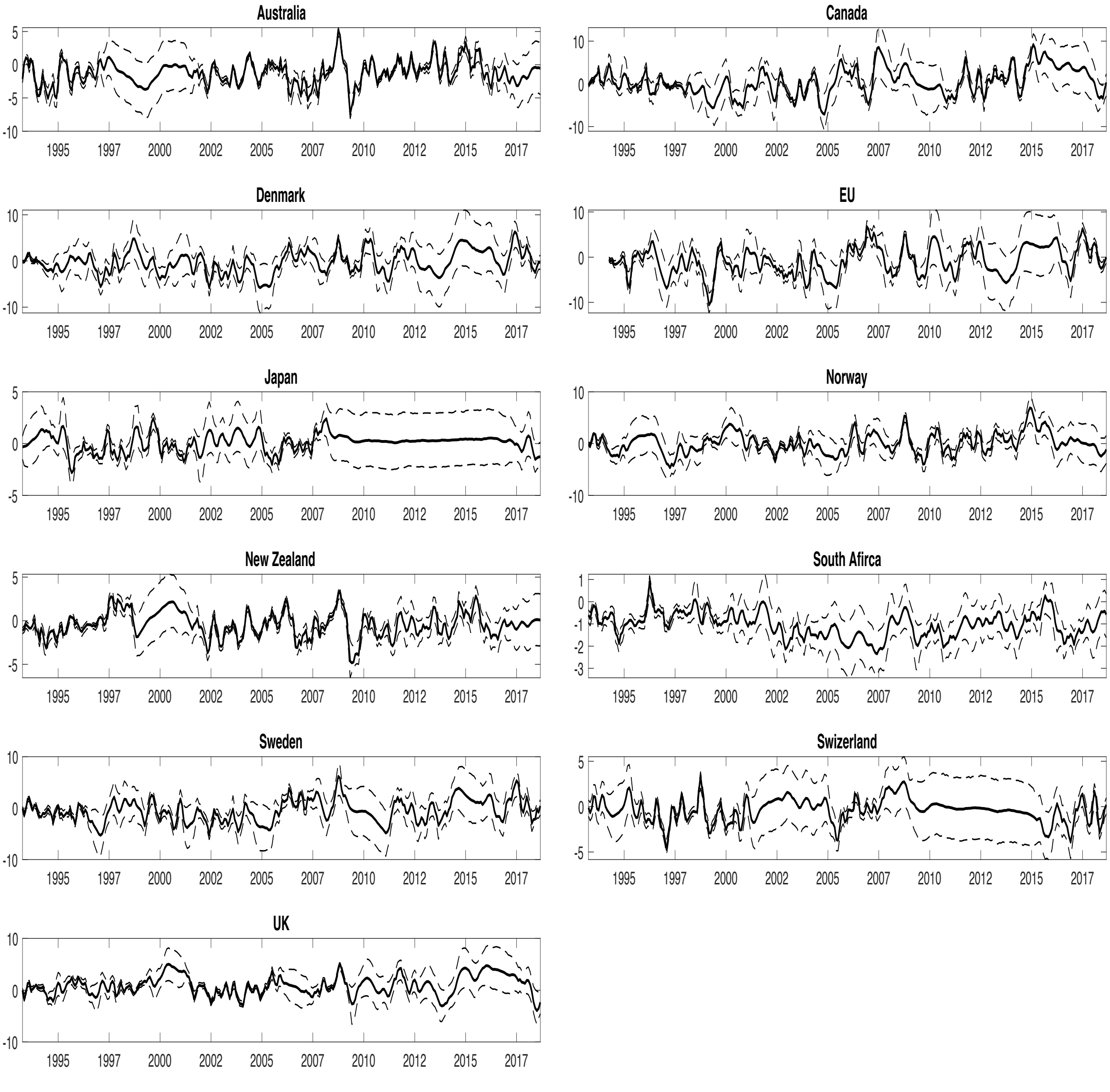


Figure 4.3: The estimated time-varying coefficents, β_t. The dashed lines indicates 95% credible intervals.

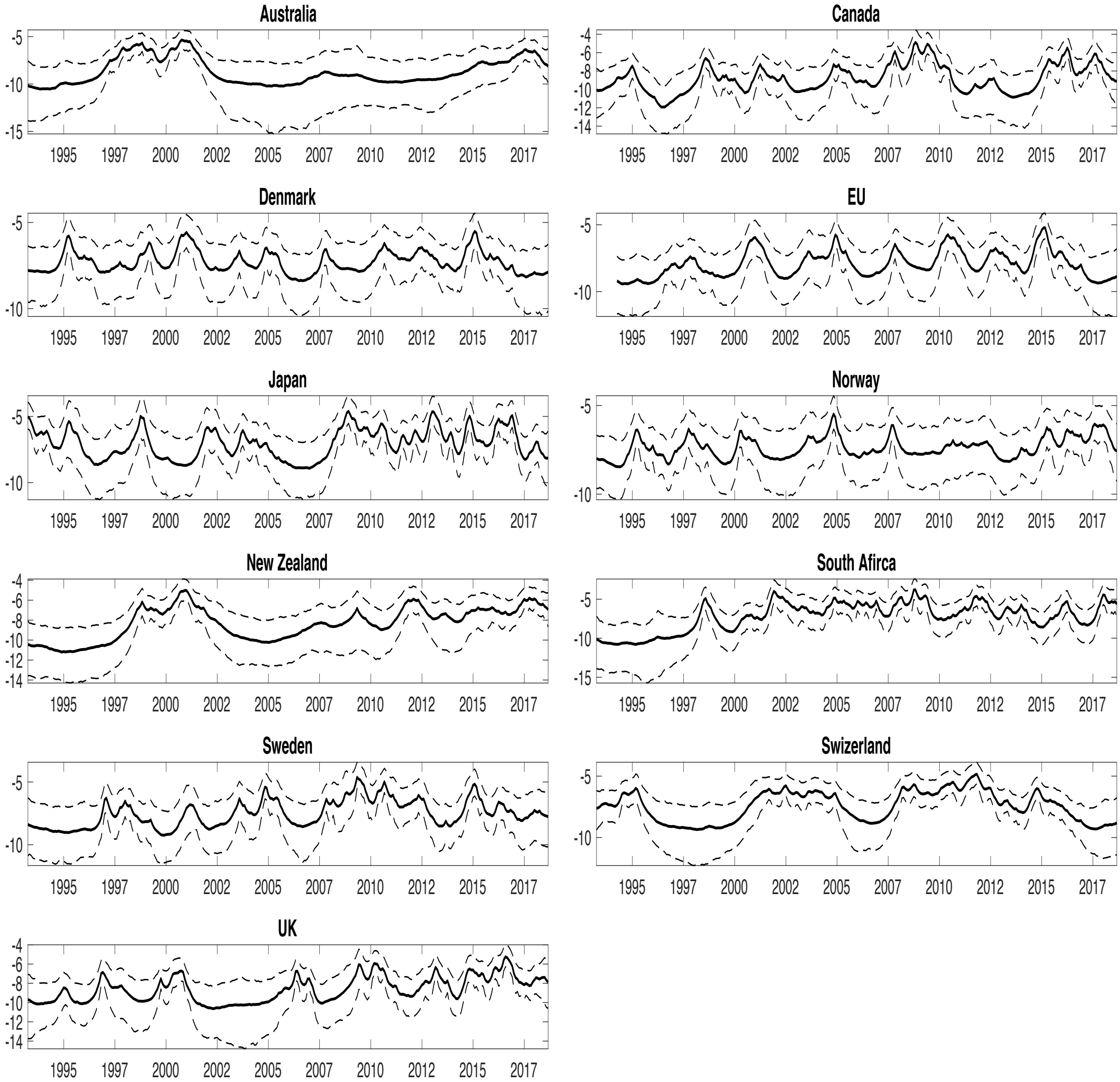


Figure 4.4: The Estimated Stochastic Volatility, σ_t.

the periods of high stochastic volatility, β_t tends to be more stable and has a large variance. In particular, for Japan during the period 2007-2017, the stochastic volatility is at a high level while the β_t is quite stable and the variance is quite large.

4.3.3 Test of the Uncovered Interest Parity Hypothesis

Due to the time-variation in the UIP coefficients, equation (4.1) can be thought of as an equilibrium relationship between the exchange rate and interest rate differential. Theoretically, no arbitrage has to hold in equilibrium. And local movements in the UIP parameters have to display mean-reverting dynamics to the equilibrium. Thus, we study if the equilibrium relationship holds by testing if the unconditional mean of α_t and β_t are equal to their theoretical values.

For all the economies, the second column of Table 4.1 shows that the values of α are close to 0, consist with the UIP hypothesis that states $\alpha = 0$. Conversely, the magnitudes of β_μ, shown in the third column of Table 4.1, vary with different economies and have a large variance. It is not clear whether β_μ is consistent with its theoretical value.

Thus, following on from Section 4.2.2, we test the null $H_0 : \mu_\beta = 1$ against $H_1 : \mu_\beta \neq 1$. The fourth column of Table 4.1 shows that the log of SDDRs for Canada, Denmark, the EU, Norway, and the UK are larger than -2. This suggests that for those economies, we cannot reject the null hypothesis, $\beta_\mu = 1$. The log of SDDRs for Japan, New Zealand, Sweden, and Switzerland are greater than -6 and less than -2. These log of SDDRs show that for these four economies, we find positive evidence that against the UIP null hypothesis. The log of SDDRs for South Africa are much less than -6, suggesting that we can strongly reject the UIP hypothesis for South Africa. The log of SDDRs could vary as the prior of β_μ changes from loose to tight. Therefore,as a sensitivity test, in column 5 and 6, we show the log of SDDRs in cases of loose prior and tight prior, respectively. This sensitivity test shows that the log of SDDRs for Canada, Denmark, the EU, Norway, and the UK are still larger than

Table 4.1: Estimated Uncovered Interest Parity Coefficients and Hypothesis Test Results

	α	β_μ	log of SDDR		
			$V_\beta = 1$	$V_\beta = 10$	$V_\beta = 0.1$
Canada	-0.0003	0.681	0.616	2.304	0.065
	[-0.004 0.002]	[-0.519 2.046]			
Denmark	0.125	-1.263	0.003	-0.462	-0.391
	[-0.009 0.012]	[-0.985 1.373]			
EU	0.002	-0.075	-1.676	-1.661	-0.323
	[-0.006 0.010]	[-1.482 1.415]			
Japan	-0.004	0.126	-2.477	-1.077	-1.210
	[-0.011 0.005]	[-0.598 0.979]			
Norway	-0.002	0.256	-0.993	0.325	-0.532
	[-0.010 0.005]	[-0.635 1.249]			
New Zealand	0.008628	-0.251	-5.7155	-8.320	-1.880
	[-0.004 0.026]	[-1.087 0.668]			
South Africa	0.079	-0.957	-15.575	-18.026	-7.626
	[0.060 0.104]	[-1.464 -0.472]			
Sweden	0.010	-0.136	-3.000	-2.372	-0.779
	[0.0001 0.0205]	[-1.116 0.977]			
Switzerland	-0.014	-0.197	-4.339	-3.526	-1.456
	[-0.023 -0.007]	[-1.089 0.754]			
UK	-0.010	0.856	1.1681	3.1995	0.221
	[-0.019 -0.003]	[-0.172 1.974]			

Note: A 95% HPD is in square brackets.

-2. For these economies, we still cannot reject the UIP hypothesis. The log of SDDRs for Japan, New Zealand, Sweden, and Switzerland show that as the prior changes from loose to tight, the positive evidence rejecting the UIP hypothesis becomes weak. The log of SDDR for South Africa is still less than -6, suggesting for this economy that the UIP hypothesis can still be very strongly rejected.

4.3.4 Influential Variable

Given the UIP coefficient β_t is unstable, it is natural to ask what leads to changes in β_t. To answer this question, we first employ the LASSO method to find the important U.S macroeconomic variables from McCracken and Ng, 2016 to explain β_t.[5] Second, we regress the U.S variables on β_t using ordinary least squares and find the variables that have statistically significant effects on β_t. We do this exercise for each economy that we consider in this chapter. Our agnostic approach is similar to Cecchetti et al. (2017). They estimate an unobserved component model with stochastic volatility to obtain the estimates of trend inflation and regress the estimates of trend inflation on other variables to find which variable can determine the trend inflation.

Table 4.2 shows 29 statistically significant variables and their effects on β_t for each economy. To give an overview of our results, we classify these variables into 6 categories: output, money and credit, stock market, labour market, and price.

From Table 4.2, the U.S labour market has significant effects on β_t for all the economies which we consider. US output has significant effects on β_t for most of the economies we consider in this chapter, except Sweden. The US stock market has significant effects on β_t for most of the economies considering in this chapter, except Canada and Switzerland. US consumption is relatively less important in explaining β_t since U.S consumption has significant effects on β_t for the EU, Norway, South Africa, Sweden, and New Zealand but

[5]We use the median estimated β_t as the proxy of β_t. The results of other quartiles can be viewed in Section 4.5. Our main results are robust to different quartiles.

not for other economies. US money and credit and U.S price are not important in explaining β_t. US money and credit only have significant effects on β_t for two economies: Denmark and Norway.

Now, we explain how the US macroeconomic variables affect β_t for each economy. For Canada, Denmark, the EU and Norway, a one standard deviation increase in the number of all employees from the non-durable goods industry leads to a standard deviation increase in β_t of 0.362, 0.320 and 0.206, respectively. For Japan, a one standard deviation increase in capacity utilisation rate for manufacturing leads to a one standard deviation increase in β_t of 1.015. For New Zealand and South Africa, a one standard deviation increase in the capacity utilisation rate for manufacturing leads to a standard deviation decrease in β_t of 0.752 and 1.050, respectively. For Sweden, a one standard deviation increase in the number of all employees from the manufacturing industry leads to a standard deviation increase in β_t of 0.283. For Switzerland, a one standard deviation increase in total business inventories leads to a standard deviation increase in β_t of 0.267. For the UK, a one standard deviation increase in the number of all employees from the wholesale trade industry leads to a standard deviation decrease in β_t of 0.532.

In summary, first, the output and labour markets have a statistically significant effect on β_t for all the economies we consider in this chapter.[6] Second, the variables that have the largest effect on β_t for each economy are from either the output or labour market. Therefore, we can conclude that the output and labour markets plays the most important role in β_t. We can only find strong evidence rejecting the UIP null hypothesis for South Africa since, as for these two economies the log of SDDRs are less than -10. Unlike Ismailov and Rossi (2018) who find that the UIP hypothesis does not hold, we find evidence to support the UIP hypothesis.

[6] The variation is associated with output or labor market can be due to that the US output and labour markets are closely related to the US short-term interest rate in the UIP regression.

4.4 Conclusion

In this chapter, we use a UIP regression model with time-varying coefficients and stochastic volatility to explore the UIP relation. First, we highlight the importance of jointly modelling time-varying coefficients and stochastic volatility in estimating the TVPs in our UIP regression model. Second, we provide more evidence in support of UIP. The time-varying risk premium explanation is not favoured by our estimated α_t that is tightly around zero. Our test shows that the UIP hypothesis unconditionally cannot be rejected for Canada, Denmark, the EU, Norway and the UK. Among 94 US macroeconomic series, using LASSO, we find 29 variables that can explain the variation in β_t for the 11 economies considered in this chapter. Most of the influential variables come from the US labour market and the output category. Therefore, the US labour market and output factors play an important role in the β_t.

Table 4.2: Determinants for Uncovered Interest Parity Coefficient

	Canada	Denmark	EU	Japan	Norway	New Zealand	South Africa	Sweden	Switzerland	UK
Output										
Capacity utilization: Manufacturing		-0.169		1.015***		-0.752***	-1.050***			
IP: Manufacturing (SIC)				-0.979***			0.816***			
IP: Nondurable consumer goods							0.102***			
IP: Materials		-0.114***								-0.221***
IP: Durable materials	-0.23***									
IP: Nondurable materials					-0.101**					
IP: Residential utilities						-0.281***				
IP: Final products (market group)						0.466***				
Money and credit										
Real M2 money stock		0.157***			0.156***					0.271***
M2 money stock		-0.124***			-0.130***					-0.173***
Stock market										
S&P's common stock price index: composite				0.090**			0.172***	0.233***		
S&P's composite common stock: Price-earnings ratio		-0.212***	-0.274***		-0.199***	-0.169***		0.233***		-0.142***
Labour market										
All employees: Durable goods						0.301***	0.325***			
All Employees: Financial activities		0.125***		-0.217***		0.187***	0.149***			
All Employees: Service-providing industries				0.222***			0.162***			
All Employees: Wholesale trade					-0.151***	-0.285***	-0.221***			-0.532***
All Employees: Manufacturing								0.283***		0.249**
All Employees: Mining and logging: Mining	-0.156***			-0.206***					0.095**	-0.134***
All Employees: Nondurable goods	0.362***	0.320***	0.320***		0.206***					
All Employees: Retail trade	-0.166***		-0.124***							
Civilian unemployment rate					-0.109**					
All employees: Total nonfarm										0.261***

Table 4.2: Determinants for Uncovered Interest Parity Coefficient

	Canada	Denmark	EU	Japan	Norway	New Zealand	South Africa	Sweden	Switzerland	UK
All employees: Construction									-0.176***	
Civilians unemployed for 27 weeks and over									0.148***	
Consumption and inventories										
Unfilled orders for durable goods			0.137***				-0.157***			
New orders for consumer goods					-0.131***			-0.107**		
Real personal consumption expenditures						0.129***				
Total business inventories			-0.234***	0.223***			-0.320***		0.267***	
Price										
PPI: Metals and metal products:										

Note: The dependent variable is the median estimates of UIP coefficient, β_t. The asterisks denote significance levels: *10%; **5%; ***1%.

4.5 Appendix

4.5.1 Bayesian Estimation Procedure

The MCMC algorithm used for the Bayesian inference iterates over the following three blocks:

1. Sample stochastic volatility process σ_t from $p(\sigma_t|\widetilde{y}_t, \widetilde{x}_t, \alpha_t, \beta_t, \mu_{\log\sigma}, \phi_{\log\sigma}, \sigma_{\log\sigma})$ for $t = 1, ..., T$;

2. Sample UIP coefficients from $p(\alpha_t, \beta_t|\widetilde{Y}_T, \widetilde{X}_T, \mu_\alpha, \mu_\beta, \phi_\alpha, \phi_\beta, \sigma_\alpha, \sigma_\beta)$;

3. Sample hyperparameters from $p(\theta|\{\alpha_t\}_{t=1}^T, \{\beta_t\}_{t=1}^T, \{\log\sigma_t\}_{t=1}^T)$.

In the first block, the conditional posterior of $\log\sigma_t$ comes from the following standard stochastic volatility model with measurement equation

$$\zeta_t = \sigma_t \varepsilon_t, \quad \zeta_t = \Delta s_{t+h} - \alpha_t - \beta_t(i_{t+h} - i^*_{t+h}).$$

The model is equivalent to the linear but non-Gaussian state space model

$$\begin{aligned} \log(\zeta_t^2) &= 2\log\sigma_t + \xi_t, \quad \xi_t \sim \log\chi_1^2, \\ \log\sigma_{t+1} &= \mu_{\log\sigma}(1 - \phi_{\log\sigma}) + \phi_{\log\sigma}\log\sigma_t + \sigma_{\log\sigma}\eta_{\log\sigma,t}. \end{aligned} \tag{4.5}$$

According to Kim et al. (1998), the $\log \chi_1^2$ distribution can be closely approximated using a Gaussian mixture with seven components tabulated by the triple (q_j, m_j, ψ_j), $j = 1, ..., 7$ where q_j is the probability that ξ_t is described by component $N(m_j, \psi_j)$. Thus given a sequence of auxiliary variables $s_t \in \{1, ..., 7\}$ indicating which component is chosen, the measurement equation can be written as

$$\log(\zeta_t^2) = m_{j,s_t} + 2\log\sigma_t + \sqrt{\psi_{j,s_t}}\xi_t^*, \quad \xi_t^* \sim N(0,1). \tag{4.6}$$

Thus, conditional on $S_T = (s_1, ..., s_T)$, system (4.5) and (4.6) form a linear Gaussian state space model with time-variant but pre-determined transition matrices. The simulation smoother of Frühwirth-Schnatter (1994) or De Jong and Shephard (1995) can be used to efficiently draw $\log\sigma_t$ for $t = 1, ..., T$ as one block. The sampling of S_T can be easily done due to the fact $p(s_t = j | \log(\zeta_t^2), \log\sigma_t) \propto q_j N(\log(\zeta_t^2); m_j + 2\log\sigma_t, \psi_j)$. Importantly, S_T is sampled at the end of each MCMC iteration to ensure it is generated from the correct conditional posterior (Del Negro and Primiceri, 2015).

Given the volatility process σ_t, the TVS-SV model (4.2) becomes a linear Gaussian state space model with pre-determined time-variant system matrices. A simulation smoother is used to generate draws of α_t and β_t for $t = 1, ..., T$ as one block.

For $f = (\alpha, \beta, \log\sigma)$, the unconditional mean, μ_f, is drawn from a Gaussian distribution $N(\underline{u_f}, \underline{v_f})$ with

$$\underline{v_f} = \left(\frac{1-\phi_f^2}{\sigma_f^2} + \frac{(T-1)(1-\phi_f)^2}{\sigma_f^2} + \frac{1}{v} \right)^{-1},$$
$$\underline{u_f} = \underline{v_f} \left(\frac{u_f}{v} + \frac{(1-\phi_f^2) f_1}{\sigma_f^2} + (1-\phi_f) \frac{\sum_{t=2}^T (f_t - \phi_f f_{t-1})}{\sigma_f^2} \right).$$

The conditional posterior distribution of innovation variance σ_t is $IG(\underline{\gamma_f}, \underline{\delta_f})$ with $\underline{\gamma_f} = \gamma + T/2$ and $\underline{\delta_f} = \delta + \frac{1}{2}(\sum_{t=2}^T (f_t - \phi_f f_{t-1})^2 + (1-\phi_f^2)(f_1 - \mu_f)^2)$. Let $f_t^* = f_t - \mu_f$. Given the $Beta(a, b)$ prior, the conditional posterior distribution of autoregression coefficient is given by

$$p(\phi_f|f_1,...,f_T,\mu_f,\sigma_f) \propto Beta(\phi_f;a,b)\sqrt{1-\phi_f^2}\exp\left(-\frac{(1-\phi_f^2)f_1^2}{2\sigma_f^2}-\sum_{t=2}^{T}\frac{(f_t^*-\phi_f f_{t-1}^*)^2}{2\sigma_f^2}\right)$$
$$\propto (1+\phi_f)^{a-1}(1-\phi_f)^{b-1}\sqrt{1-\phi_f^2}\exp\left(\frac{(\phi_f-\underline{a})^2}{2\underline{b}}\right),$$

where $\underline{a}=(\sum_{t=2}^{T} f_t^* f_{t-1}^*)/\sum_{t=1}^{T-1} f_t^{*2}$ and $\underline{b}=\sigma_f^2/\sum_{t=1}^{T-1} f_t^{*2}$. To sample from this distribution, we apply the Metropolis-Hastings accept-reject algorithm (Chib and Greenberg, 1995) by drawing a candidate ϕ^{new} from $N(\underline{a},\underline{b})$ truncated between $(-1,1)$ to ensure stationarity, and the draw is accepted with probability

$$\min\left(\frac{(1+\phi_f^{new})^{a-1}(1-\phi_f^{new})^{b-1}\sqrt{(1-\phi_f^{new2})}\sqrt{1-\phi_f^{new2}}}{(1+\phi_f)^{a-1}(1-\phi_f)^{b-1}\sqrt{(1-\phi_f^2)}\sqrt{1-\phi_f^2}},\ 1\right).$$

4.5.2 Observation Weights

Conditional on θ and $\sigma_1,...,\sigma_T$, such as an MCMC draw or a posterior estimate, a Kalman filter and smoother can be used to compute the observation weights of the TVC-SV model (4.2) with respect to the UIP coefficient β_t. For simplicity, we assume $\alpha_t=0$ for all t, which is in line with our empirical results. Let $\beta_t^*=\beta_t-\mu_\beta$. The state space model becomes

$$\bar{y}_t=\beta_t^*+\frac{\sigma_t}{\widetilde{x}_t}\varepsilon_t,$$
$$\beta_{t+1}^*=\phi_\beta\beta_t^*+\sigma_\beta\eta_{\beta,t},$$
$$\beta_1^*\sim N\left(0,\frac{\sigma_\beta^2}{1-\phi_\beta^2}\right),$$

where $\bar{y}_t$ is equal to $\widetilde{y}_t/\widetilde{x}_t-\mu$. Suppose the system were time-invariant, say $\sigma_1/\widetilde{x}_1=...=\sigma_T/\widetilde{x}_T=\sigma_\varepsilon$.[7] Kalman filter outputs reach their steady state quickly (Durbin and Koopman, 2012, see Chapter 4). Let a_t and P_t denote the filtering expectation $E(\beta_t|\widetilde{Y}_{t-1},\widetilde{X}_{t-1})$ and

[7]One can think of σ_ε as the unconditional mean of $\sigma_1/\widetilde{x}_1$, because the interest rate differential in the denominator is stationary between developed economies, and the stochastic volatility is assumed to be mean-reverting.

variance $Var(\beta_t|\widetilde{Y}_{t-1},\widetilde{X}_{t-1})$, respectively. The Kalman filter iterates forward over

$$a_{t+1} = \phi_\beta a_t + \frac{P_t}{P_t+\sigma_\varepsilon^2}(\bar{y}_t - A_t), \;\; P_{t+1} = \frac{P_t\sigma_\varepsilon^2}{P_t+\sigma_\varepsilon^2} + \sigma_\beta^2.$$

The steady state is given by the fixed-point solution to the second equation, and it is $\bar{P} = (q+\sqrt{q^2+4q})/2$ where $q = \sigma_\beta^2/\sigma_\varepsilon^2$ is the signal-to-noise ratio. Let b_t denote the smoothed expectation $E(\beta_t|\widetilde{Y}_T,\widetilde{X}_T)$. The Kalman smoother iterates backward from $r_T = 0$ and $Var(r_t) = N_t$ over

$$r_{t-1} = \frac{\bar{y}_t - a_t}{P_t+\sigma_\varepsilon^2} + \frac{\sigma_\varepsilon^2}{P_t+\sigma_\varepsilon^2} r_t, \;\; b_t = a_t + P_t r_{t-1}, \;\; N_{t-1} = \frac{1}{P_t+\sigma_\varepsilon^2} + \left(\frac{\sigma_\varepsilon^2}{P_t+\sigma_\varepsilon^2}\right)^2 N_t.$$

Since $\frac{\sigma_\varepsilon^2}{P_t+\sigma_\varepsilon^2} < 1$, the steady state of N_t exists and is $\bar{N} = (\bar{P}+\sigma_\varepsilon^2)/(\bar{P}^2+2\bar{P}\sigma_\varepsilon^2)$.

Suppose we can write $b_t = \sum_{j=1}^T \omega_{jt}\bar{y}_j$ with weight ω_{jt} associated with the j-th observation corresponding to t-th smoothed estimate. Then we have $E(b_t\varepsilon_j) = \omega_{jt}E(\bar{y}_j\varepsilon_j) = \omega_{jt}$, but

$$E(b_t\varepsilon_j) = \begin{cases} -Cov(\varepsilon_j - E(\varepsilon_j|\widetilde{Y}_T,\widetilde{X}_T), \beta_t - b_t), & \text{for } j < t; \\ -Cov(\beta_t - b_t, \varepsilon_j - E(\varepsilon_j|\widetilde{Y}_T,\widetilde{X}_T)), & \text{for } j \geqslant t. \end{cases}$$

After minor algebraic manipulation, the steady state gives for $j < t$

$$Cov(\varepsilon_j - E(\varepsilon_j|\widetilde{Y}_T,\widetilde{X}_T), \beta_t - b_t) = E(\varepsilon_j(\beta_t - b_t)) = -\frac{\bar{P}}{\bar{P}+\sigma_\varepsilon^2}\left(\frac{\sigma_\varepsilon^2}{\bar{P}+\sigma_\varepsilon^2}\right)^{t-j}\frac{\bar{P}\sigma_\varepsilon^2}{\bar{P}^2+2\bar{P}\sigma_\varepsilon^2}.$$

Similarly, for $j \geqslant t$ we have

$$Cov(\beta_t - b_t, \varepsilon_j - E(\varepsilon_j|\widetilde{Y}_T,\widetilde{X}_T)) = -\frac{\bar{P}}{\bar{P}+\sigma_\varepsilon^2}\left(\frac{\sigma_\varepsilon^2}{\bar{P}+\sigma_\varepsilon^2}\right)^{j-t}\frac{\bar{P}\sigma_\varepsilon^2}{\bar{P}^2+2\bar{P}\sigma_\varepsilon^2}.$$

Thus, as j moves away from t, observation $\bar{y}_j$ receives exponentially declining weight proportional to

$$\omega_{jt} \propto \left(\frac{\sigma_\varepsilon^2}{\bar{P}+\sigma_\varepsilon^2}\right)^{|t-j|} = \left(\frac{2\sigma_\varepsilon^4}{\sigma_\beta^2 + \sqrt{\sigma_\beta^4 + 4\sigma_\beta^2\sigma_\varepsilon^2 + 2\sigma_\varepsilon^4}}\right)^{|t-j|}.$$

This result is heuristic in our case because when stochastic volatility is present, the steady state of the Kalman filter ceases to exist. However, as the first-order approximation of the weighting function can be computed by replacing σ_ε by $\sigma_t/\widetilde{x}_t$. So if σ_t increases or the interest rate differential $\widetilde{x}_t$ decreases, ω_{jt} becomes larger. This means that during volatile times, the accounting of the UIP coefficient β_t relies on more backward and forward information, and vice versa.

Table 4.3: Determinants for the 10 Percentile of Uncovered Interest Parity Coefficient, β_t.

	Canada	Denmark	EU	Japan	Norway	New Zealand	South Africa	Sweden	Swizerland	UK
Output										
Capacity utilization: Manufacturing						-0.269***	-0.990***			
IP: Manufacturing (SIC)					-0.337**	0.801***		-0.306**		
IP: Nondurable consumer goods							0.109***			0.117***
IP index		-0.116**								
IP: Durable consumer goods					0.177***					
IP: Durable materials	-0.309***									
IP: Business Equipment				-0.106**						
IP: Residential utilities						-0.096**				
IP: Materials										
Money and credit										
Real M2 money stock					0.136**					0.124***
M2 money stock					-0.130***					
Stock market										
S&Ps Composite common stock: Dividend yield								0.265***		0.132***
S&Ps composite common stock: Price-earnings ratio		-0.232***	-0.269***		0.210***	-0.170***		0.230***		
Labour market										
All employees: Nondurable goods	0.364***	0.323***	0.319***					0.215***	-0.176***	
All employees: Wholesale trade										-0.292***
All employees: Retail trade	-0.174***	-0.139***	-0.138**							
All employees: Financial activities		0.177***				0.237***	0.134***			0.143***
All employees: construction									-0.252***	0.227***
All employees: Mining and logging: mining	-0.159***								0.120***	
All employees: Durable goods							0.305***			
Civilian unemployment rate					-0.147***				0.121***	
Consumption and inventories										
Real personal consumption expenditures						0.245***				
Unfilled orders for durable goods			0.139***	-0.107**		0.141***	-0.147***			
New orders for consumer goods		-0.109**			-0.176***			-0.125***		
Retail and food services sales						-0.195***				
New orders for durable goods						-0.102**				
Total business inventories		-0.228***	-0.256***	0.205***			-0.335***		0.239***	
Price										
PPI: Metals and metal products:										

Note: The dependent variable is median estimates of the uncovered interest parity coefficient, β_t. The asterisks denote significance levels: *10%; **5%; ***1%.

Table 4.4: Determinants of the 25 Percentile of the Uncovered Interest Parity Coefficient, β_t.

	Canada	Denmark	EU	Japan	Norway	New Zealand	South Africa	Sweden	Swizerland	UK
Output										
IP: Materials										-0.249***
Capacity utilization: manufacturing						-0.255***	-0.990***			
IP: Durable materials	-0.230***							-0.206***		
IP: Business equipment					-0.171***					
IP: Manufacturing (SIC)							0.801***			
IP index		-0.127***								
IP: Nondurable materials					-0.099**					
IP: Nondurable consumer goods							0.109***			
Money and credit										
M2 money stock		-0.119**			-0.141***					-0.167***
Real M2 money stock		0.153***			0.168***					0.249***
Stock market										
S&Ps composite common stock: Price-earnings ratio		-0.223***	-0.274***		-0.203***	-0.202***		0.224***		
S&Ps composite common stock: dividend yield								0.245***	0.113***	0.132***
Labour market										
All employees: Mining and logging: mining	-0.156***									-0.108***
All employees: Nondurable goods	0.362***	0.323***	0.343***		0.192***				-0.150***	
All employees: Wholesale trade					-0.145***					-0.535***
All employees: Financial activities		0.159***				0.285***	0.134***			
All employees: total nonfarm										0.358***
All employees: Retail trade	-0.166***		-0.114**							
All employees: Construction									-0.243***	
All employees: Durable goods							0.305***			0.298***
Civilians unemployed for 27 weeks and over									0.167***	
All employees: Trade, transportation & utilities						-0.134***				
Civilian labor force										
Civilian unemployment rate					-0.105**					
Consumption and inventories										
Total business inventories		-0.207***	-0.239***	0.180***			-0.335***		0.255***	
New orders for consumer goods					-0.1553***	-0.103**		-0.108**		
Unfilled orders for durable goods			0.146***			0.163***	-0.147***			
New orders for durable goods						-0.101**				
Real personal consumption expenditures						0.121***				
Price										
PPI: Metals and metal products										

Note: The dependent variable is 25 percentile estimates of the uncovered interest parity coefficient, β_t. The asterisks denote significance levels: *10%; **5%; ***1%.

Table 4.5: Determinants of the 75 Percentile of the Uncovered Interest Parity Coefficient, β_t.

	Canada	Denmark	EU	Japan	Norway	New Zealand	South Africa	Sweden	Swizerland	UK
Output										
Capacity utilization: Manufacturing				0.841***			-0.341***			
IP: Materials		-0.118***								-0.147***
IP: Manufacturing (SIC)				-0.797***						
IP: Durable materials	-0.288***				-0.130**			-0.229***		-0.163***
IP: Nondurable materials					-0.101**					
IP: Residential utilities						-0.281***				
IP: Final products (market group)						0.466***				
IP: Nondurable consumer goods							0.144***			
Money and credit										
Real M2 money stock		0.175***								0.2503***
M2 money stock		-0.134***	-0.111**							
Commercial and industrial loans			0.096**							
Stock market										
S&Ps composite common stock: dividend yield				0.087**			0.156***	0.233***	0.113***	
S&Ps composite common stock: Price-earnings Ratio		-0.217***	-0.252***		-0.210***	-0.169***		0.233***		
Labour market										
All employees: Mining and logging: Mining	-0.185***			-0.211***		-0.752***	-0.154***			-0.108**
All employees: Nondurable goods	0.358***	0.293***	0.290***		0.230***				-0.150***	
All employees: wholesale trade		-0.146**			-0.212***	-0.285***	-0.266***			-0.572***
All employees: Financial activities		0.111**		-0.211***		0.187***	0.184***			-0.151***
All employees: Durable goods						0.301***	0.323***			
All employees: Total nonfarm										0.325***
All employees: Retail trade	-0.176***									
Civilians unemployed for 27 weeks and over									0.167***	
Avg weekly overtime hours : Manufacturing										
All employees: construction					0.128**		0.158***		-0.243***	
All employees: Service-providing industries							0.207***			
All employees: Manufacturing								0.283***		0.308**
Consumption and inventories										
New orders for consumer goods								-0.107**		
Total business inventories		-0.1349**	-0.247***	0.261***			-0.315***		0.255***	
Unflled orders for durable goods			0.139***							
Real personal consumption expenditures						0.129***				
Price										
PPI: Metals and metal products										

Note: The dependent variable is the75 percentile estimates of the uncovered interest parity coefficient, β_t. The asterisks denote significance levels: *10%; **5%; ***1%.

Table 4.6: Determinants of the 90 Percentile of the Uncovered Interest Parity Coefficient, β_t.

	Canada	Denmark	EU	Japan	Norway	New Zealand	South Africa	Sweden	Swizerland	UK
Output										
Capacity utilization: manufacturing				-0.865***		-0.752***	-0.346***	-0.136**	0.319**	
IP: Manufacturing (SIC)				-0.865***						
IP: Durable materials	-0.230***				-0.130**					
IP: Final products (market group)						0.466***				
IP: Nondurable materials					-0.101**				-0.110**	
IP: Materials		-0.115**								-0.147***
IP: Residential utilities						-0.281***				
IP: Nondurable consumer goods							0.162***			
IP: Business equipment										
Money and credit										
Real M2 money stock		0.232***								0.250***
M2 money stock		-0.150***								-0.151**
Stock market										
S&Ps composite common stock: Price-earnings ratio		-0.185***	-0.272***		-0.210***	-0.169***	0.125**			
S&Ps composite common stock: Dividend yield							0.204***			
Labour market										
All employees: Mining and logging: Mining	-0.156***			-0.211***						
All employees: Nondurable goods	0.362***	0.227***	0.303***		0.230***			0.319***		
All employees: Wholesale trade		-0.219***			-0.212***	-0.285***	-0.210***			-0.572***
All employees: manufacturing										0.308**
All employees: Total nonfarm									-0.231***	0.325***
All employees: Financial activities		0.114**		-0.217***		0.187***	0.214***			
All employees: Durable goods						0.301***	0.364***			
All employees: Retail trade	-0.166***									
All employees: Goods-producing industries										
All employees: Construction					0.128**		0.222***			
Average duration of unemployment (weeks)									0.111***	
Avg weekly overtime hours : Manufacturing										
All employees: government							0.107***			
Consumption and inventories										
Total business inventories			-0.206***	0.244***			-0.323***		0.250***	
New orders for consumer goods								-0.116**		
Unfilled orders for durable goods			0.150***							
Real personal consumption expenditures						0.129***				
Price										
PPI: Metals and metal products:										

Note: The dependent variable is the 90 percentile estimates of the Uncovered Interest Parity Coefficient, β_t. The asterisks denote significance levels: *10%; **5%; ***1%.

Chapter 5

Conclusion

This thesis examines three applications of TVP models in macroeconomics. First, in Chapter 2, we estimate eight Phillips curve models and test for time variation in the slopes of the Phillips curve under these models. We find that CPI inflation is much more sensitive to changes in the unemployment gap than in the output gap. Models with the unemployment gap have a much larger λ_t in magnitude than do models with the output gap. Second, we find that time variation of the slope of the Phillips curve mainly depends on specifications of inflation expectations. When the measures of inflation expectations are observable, the slope of the Phillips curve is time varying. However, when the measure of inflation expectations is trend inflation with large variance, the slope of the Phillips curve is constant. By reducing the variance in the estimated trend inflation through linking blue chip 10 years inflation forecasts to trend inflation, the Phillips curve model with trend inflation has the time-varying slope. We consider different measures of inflation expectations: backward-looking inflation expectations, forward-looking inflation expectations and trend inflation. The slopes of these Phillips curves with different measures of inflation expectations are volatile and time varying. Thus, we conclude that the slope of the Phillips curve is time varying.

Second, chapter 3 replicates the results of Anundsen et al. (2016) in both a narrow and wide sense. We establish the narrow replication by reproducing the same results as theirs by using Matlab. First, the wide replication shows that the coefficients of financial crisis indicators are quite stable over time, but the country dummies have considerable time variation. Second, through the lens of time-varying probit models, most of the main findings of Anundsen et al. (2016) are robust to using a TVP approach. However, the evidence that the NFE credit-to-GDP gap significantly affects the probability of a financial crisis is mixed across different specifications. Third, based on expanding window forecasting exercise, the warning system models' in-sample model fit and out-of-sample forecasting performance based on expanding window forecasting exercise can be considerably improved by allowing for TVPs.

Finally, in Chapter 4, we use a UIP regression model with time-varying coefficients and stochastic volatility to explore the UIP relation. First, we highlight the importance of jointly modelling time-varying coefficients and stochastic volatility in estimating the TVPs in our UIP regression model. Second, we find that the UIP hypothesis cannot be rejected for several advanced economies. The time-varying risk premium explanation is not favoured by our estimated α_t, which is tightly around zero. Third, using LASSO, among 94 US macroeconomic series we find 29 variables that can explain the variation in β_t for 11 economies considered in this paper. Most of the influential variables come from the US labour market and the output category. Therefore, the US labour market and output factors can play an important role in the β_t.

Bibliography

Alessi, L. and Detken, C. (2017). Identifying excessive credit growth and leverage. *Journal of Financial Stability*, 35:215–225.

Anundsen, A. K., Gerdrup, K., Hansen, F., and Kragh-Sørensen, K. (2016). Bubbles and crises: The role of house prices and credit. *Journal of Applied Econometrics*, 31(7):1291–1311.

Ascari, G. (2004). Staggered prices and trend inflation: some nuisances. *Review of Economic dynamics*, 7(3):642–667.

Atkeson, A. and Ohanian, L. E. (2001). Are Phillips curves useful for forecasting inflation? *Quarterly Review*, (Winter):2–11.

Baillie, R. T. and Bollerslev, T. (2000). The forward premium anomaly is not as bad as you think. *Journal of International Money and Finance*, 19(4):471–488.

Bean, C. R. (2006). Globalisation and inflation. Retrieved from https://www.bankofengland.co.uk/-/media/boe/files/speech/2006/globalisation-and-inflation.

Belmonte, M. A., Koop, G., and Korobilis, D. (2014). Hierarchical shrinkage in time-varying parameter models. *Journal of Forecasting*, 33(1):80–94.

Berge, T. J. and Jordà, Ò. (2011). Evaluating the classification of economic activity into recessions and expansions. *American Economic Journal: Macroeconomics*, 3(2):246–277.

Berger, T., Everaert, G., and Vierke, H. (2016). Testing for time variation in an unobserved components model for the U.S. economy. *Journal of Economic Dynamics and Control*, 69:179–208.

Beveridge, S. and 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(2):151–174.

Brunnermeier, M. K., Nagel, S., and Pedersen, L. H. (2008). Carry trades and currency crashes. *NBER Macroeconomics Annual*, 23(1):313–348.

Bussiere, M. and Fratzscher, M. (2006). Towards a new early warning system of financial crises. *Journal of International Money and Finance*, 25(6):953–973.

Büyükkarabacak, B. and Valev, N. T. (2010). The role of household and business credit in banking crises. *Journal of Banking & Finance*, 34(6):1247–1256.

Cai, Z., Fan, J., and Li, R. (2000). Efficient estimation and inferences for varying-coefficient models. *Journal of the American Statistical Association*, 95(451):888–902.

Canova, F. (1993). Modelling and forecasting exchange rates with a Bayesian time-varying coefficient model. *Journal of Economic Dynamics and Control*, 17(1–2):233–261.

Casella, G., Girón, F. J., Martínez, M. L., Moreno, E., et al. (2009). Consistency of Bayesian procedures for variable selection. *The Annals of Statistics*, 37(3):1207–1228.

Cecchetti, S. G., Feroli, M., Hooper, P., Kashyap, A. K., and Schoenholtz, K. (2017). Deflating inflation expectations: The implications of inflation's simple dynamics. (SEPR Discussion Paper 11925). Retrieved from https://ideas.repec.org/p/cpr/ceprdp/11925.html.

Chan, J. C., Koop, G., Leon-Gonzalez, R., and Strachan, R. W. (2012). Time varying dimension models. *Journal of Business & Economic Statistics*, 30(3):358–367.

Chan, J. C. C. (2018). Specification tests for time-varying parameter models with stochastic volatility. *Econometric Reviews*, 37(8):807–823.

Chan, J. C. C., Clark, T. E., and Koop, G. (2018). A new model of inflation, trend inflation, and long-run inflation expectations. *Journal of Money, Credit and Banking*, 50(1):5–53.

Chan, J. C. C. and Eisenstat, E. (2015). Marginal likelihood estimation with the Cross-Entropy method. *Econometric Reviews*, 34(3):256–285.

Chan, J. C. C. and Eisenstat, E. (2018). Bayesian model comparison for time-varying parameter VARs with stochastic volatility. *Journal of Applied Econometrics*, 33(4):509–532.

Chan, J. C. C. and Grant, A. (2017). Measuring the output gap using stochastic model specification search. (Working Paper 2/2017). Retrieved from https://cama.crawford.anu.edu.au/sites/default/files/publication/cama_crawford_anu_edu_au/2017-01/2_2017_chan_grant.pdf.

Chan, J. C. C. and Jeliazkov, I. (2009). Efficient simulation and integrated likelihood estimation in state space models. *International Journal of Mathematical Modelling and Numerical Optimisation*, 1(1–2):101–120.

Chan, J. C. C., Koop, G., and Potter, S. M. (2016). A bounded model of time variation in trend inflation, NAIRU and the Phillips curve. *Journal of Applied Econometrics*, 31(3):551–565.

Chen, Y.-C. and Tsang, K. P. (2013). What does the yield curve tell us about exchange rate predictability? *Review of Economics and Statistics*, 95(1):185–205.

Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician*, 49(4):327–335.

Chinn, M. D. and Meredith, G. (2004). Monetary policy and long-horizon uncovered interest parity. *IMF Staff Papers*, 51(3):409–430.

Chinn, M. D. and Quayyum, S. (2012). Long horizon uncovered interest parity re-assessed. (Technical Report) Washington, DC: National Bureau of Economic Research.

Cogley, T. and Sargent, T. J. (2005). Drifts and volatilities: monetary policies and outcomes in the post WWII US. *Review of Economic Dynamics*, 8(2):262–302.

Cogley, T. and Sbordone, A. M. (2008). Trend inflation, indexation, and inflation persistence in the New Keynesian Phillips curve. *American Economic Review*, 98(5):2101–2126.

Coibion, O. and Gorodnichenko, Y. (2015). Is the Phillips curve alive and well after all? Inflation expectations and the missing disinflation. *American Economic Journal: Macroeconomics*, 7(1):197–232.

De Jong, P. and Shephard, N. (1995). The simulation smoother for time series models. *Biometrika*, 82(2):339–350.

Del Negro, M. and Primiceri, G. E. (2015). Time varying structural vector autoregressions and monetary policy: A corrigendum. *The Review of Economic Studies*, 82(4):1342–1345.

Doucet, A. and Johansen, A. M. (2009). A tutorial on particle filtering and smoothing: Fifteen years later. *Handbook of Nonlinear Filtering*, 12(3):656–704.

Drehmann, M. and Juselius, M. (2014). Evaluating early warning indicators of banking crises: Satisfying policy requirements. *International Journal of Forecasting*, 30(3):759–780.

Dufour, J.-M. and Kiviet, J. F. (1998). Exact inference methods for first-order autoregressive distributed lag models. *Econometrica*, 66(1):79–104.

Durbin, J. and Koopman, S. J. (2012). *Time series analysis by state space methods*. Oxford University Press.

Eisenstat, E. and Strachan, R. W. (2016). Modelling inflation volatility. *Journal of Applied Econometrics*, 31(5):805–820.

Fama, E. F. (1984). Forward and spot exchange rates. *Journal of Monetary Economics*, 14(3):319–338.

Farhi, E. and Gabaix, X. (2015). Rare disasters and exchange rates. *The Quarterly Journal of Economics*, 131(1):1–52.

Frühwirth-Schnatter, S. (1994). Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, 15(2):183–202.

Frühwirth-Schnatter, S. and Wagner, H. (2010). Stochastic model specification search for Gaussian and partial non-Gaussian state space models. *Journal of Econometrics*, 154(1):85–100.

Gaiotti, E. (2008). Has globalisation changed the Phillips curve? Firm-level evidence on the effect of activity on prices. 676. (Working Paper No. 676). Rome, IT: Bank of Italy.

Gelfand, A. E., Smith, A. F., and Lee, T.-M. (1992). Bayesian analysis of constrained parameter and truncated data problems using Gibbs sampling. *Journal of the American Statistical Association*, 87(418):523–532.

Ghosh, J. and Clyde, M. A. (2011). Rao–blackwellization for Bayesian variable selection and model averaging in linear and binary regression: A novel data augmentation approach. *Journal of the American Statistical Association*, 106(495):1041–1052.

Gordon, R. J. (2013). The Phillips curve is alive and well: Inflation and the NAIRU during the slow recovery. Technical report, National Bureau of Economic Research.

Ichiue, H. and Koyama, K. (2011). Regime switches in exchange rate volatility and uncovered interest parity. *Journal of International Money and Finance*, 30(7):1436–1450.

Ihrig, J., Kamin, S. B., Lindner, D., and Marquez, J. (2010). Some simple tests of the globalization and inflation hypothesis. *International Finance*, 13(3):343–375.

Ismailov, A. and Rossi, B. (2018). Uncertainty and deviations from uncovered interest rate parity. *Journal of International Money and Finance*, 88:242–259.

Jordà, Ò., Schularick, M., and Taylor, A. M. (2015a). Betting the house. *Journal of International Economics*, 96:S2–S18.

Jordà, Ò., Schularick, M., and Taylor, A. M. (2015b). Leveraged bubbles. *Journal of Monetary Economics*, 76:S1–S20.

Jordà, Ò. and Taylor, A. M. (2011). Performance evaluation of zero net-investment strategies. (Technical Report). Washington, DC: National Bureau of Economic Research.

Jordà, Ò. and Taylor, A. M. (2012). The carry trade and fundamentals: Nothing to fear but FEER itself. *Journal of International Economics*, 88(1):74–90.

Karlsson, S., Österholm, P., et al. (2018). Is the US Phillips curve stable? Evidence from bayesian VARs. (Working Paper 2018:5). Retrieved from https://ideas.repec.org/p/hhs/oruesi/2018_005.html.

Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795.

Kim, C.-j., Manopimoke, P., and Nelson, C. R. (2014). Trend inflation and the nature of structural breaks in the New Keynesian Phillips curve. *Journal of Money, Credit and Banking*, 46(2-3):253–266.

Kim, S., Shephard, N., and Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with arch models. *The Review of Economic studies*, 65(3):361–393.

Koop, G. and Korobilis, D. (2013). Large time-varying parameter VARs. *Journal of Econometrics*, 177(2):185–198.

Koop, G. and Potter, S. M. (2007). Estimation and forecasting in models with multiple breaks. *The Review of Economic Studies*, 74(3):763–789.

Kozicki, S. and Tinsley, P. A. Alternative sources of the lag dynamics of inflation. (Research Working Paper RWP 02-12). Kansas, TX: Federal Reserve Bank of Kansas City.

Kroese, D. P. and Chan, J. C. C. (2014). *Statistical modeling and computation*. New York, NY: Springer.

Kuttner, K. and Robinson, T. (2010). Understanding the flattening Phillips curve. *The North American Journal of Economics and Finance*, 21(2):110–125.

Li, D., Ghoshray, A., Morley, B., et al. Uncovered interest parity and the risk premium. (Working Paper), Bath, UK: University of Bath.

McCracken, M. W. and Ng, S. (2016). Fred-md: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4):574–589.

Mishkin, F. S. (2007). Inflation dynamics. *International Finance*, 10(3):317–334.

Nakajima, J. and West, M. (2013). Bayesian analysis of latent threshold dynamic models. *Journal of Business & Economic Statistics*, 31(2):151–164.

O'Hara, R. B., Sillanpää, M. J., et al. (2009). A review of Bayesian variable selection methods: What, how and which. *Bayesian Analysis*, 4(1):85–117.

Phillips, A. W. (1958). The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861–1957. *Economica*, 25(100):283–299.

Roberts, J. M. (2006). Monetary policy and inflation dynamics. *International Journal of Central Banking*, 2(3):193–230.

Rossi, B. (2007). Expectations hypotheses tests at long horizons. *The Econometrics Journal*, 10(3):554–579.

Rossi, B. (2013). Exchange rate predictability. *Journal of Economic Literature*, 51(4):1063–1119.

Schularick, M. and Taylor, A. M. (2012). Credit booms gone bust: monetary policy, leverage cycles, and financial crises, 1870–2008. *The American Economic Review*, 102(2):1029–1061.

Stella, A. and Stock, J. H. (2012). A state-dependent model for inflation forecasting. (Discussion Paper, International Finance). Retrieved from https://www.federalreserve.gov/pubs/ifdp/2012/1062/ifdp1062.pdf.

Verdinelli, I. and Wasserman, L. (1995). Computing Bayes factors using a generalization of the Savage-Dickey density ratio. *Journal of the American Statistical Association*, 90(430):614–618.

Zhao, P. and Yu, B. (2006). On model selection consistency of LASSO. *Journal of Machine Learning Research*, 7(Nov):2541–2563.