tinbergen institute

Şanlı Pınar Ceyhan

Essays on Bayesian Analysis of Time Varying Economic Patterns

Erasmus University Rotterdam

Essays on Bayesian Analysis of Time Varying Economic Patterns

Şanlı Pınar Ceyhan

ISBN: 978 90 3610 213 1

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. **590** of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

Essays on Bayesian Analysis of Time Varying Economic Patterns

Essays over Bayesiaanse analyse van tijdsvariërende patronen in de economie

Thesis

to obtain the degree of Doctor
from Erasmus University Rotterdam
by command of the rector magnificus

Prof.dr. H.A.P Pols

and in accordance with the decision of Doctoral Board

The public defense shall be held
Thursday 4 September 2014 at 15:30 hours
by
Şanlı Pınar Ceyhan
born in Istanbul, Turkey.

ERASMUS UNIVERSITEIT ROTTERDAM

Doctoral Committee

Promoter:	Prof.dr. H.K. Van Dijk
Other members:	Prof.dr. R. Paap
	Prof.dr. D. Fok
	Prof.dr. D.J.C. Van Dijk
Copromoter:	Dr. N. Baştürk

To my granddad

Preface

I am very grateful to the Erasmus University Rotterdam and the Tinbergen Institute for providing me with the opportunity to write this PhD thesis. I would like to thank below all the people that supported me during this journey.

When I was writing my MPhil thesis, working on a macroeconomic model, I was aiming to continue the modelling with the Bayesian estimation of this model. I did not know at that time that this was going to be a long journey with quite a different path than I imagined. I was not only going to make an estimation which is Bayesian in spirit, but I was also going to live in the Bayesian world. This was all made possible thanks to my supervisor, Herman Van Dijk. I feel indebted to Herman not only for supervising this thesis, but also for the endless patience he showed for my work, deep insight he provided me with and nice chats we had during our regular meetings. I would also like to thank my co-supervisor Nalan Basturk for the endless patience in answering my questions and helping me go throughout this process more smoothly. It was not only nice to talk in my native language, but also I learned a lot during our meetings. I should thank my committee member Cem for the insight he provided me with and bringing me back to reality from time to time. I would also like to thank my other committee members for accepting to review this thesis.

I was a teaching assistant for some courses taught by Christiaan Heij and Paul De Boer. These courses were the most organized ones I have ever done TAship for during my master and PhD life. I did not only improve my technical knowledge on econometrics but also learned a lot on how to more efficiently organize a course. It was also enjoyable to have chats with my TAship-mate Chen, I thank her for the super synchronized working and for being a great team-mate. I should also thank my students at EUR and TI who taught me how to be even more patient and how to look at the issues from the other side of the classroom.

I should thank Wouter Den Haan for the great, enjoyable and high standard lectures on Macroeconomics and Jan Brinkhuis for the joyful mathematics lectures. These make the life of a PhD student easier. I thank my MPhil thesis supervisor Sweder Van Wijnbergen

for the insight he provided me with during the preparation of my MPhil thesis. It was great to be part of the MInt group at UvA for a period and to have conversations with Franc, Massimo, Naomi, Markus, Petr, Tim, Matija, Sajjad, Siert and many others.

From the moment I came to the Netherlands, the classmates, friends and the secretaries at the Tinbergen institute did never let me feel lonely and were very helpful. The institute had a warm atmosphere and the yearly TI lectures were a great opportunity to listen to and meet worldwide known academicians. I would like to thank my colleagues and faculty and staff members at the Tinbergen and Econometric Institutes. I had nice chats with Arianne from time to time I was at TI, I should thank Judith for the sense of humor she brings to our conversations. I should also thank Carien, Carolien, Ine and the HR department for their help.

My childhood friends Sehkar and Tulay, thank you for being my friends. Talks with you were very enjoyable and doing nostalgia were more than joyful. Leontine, thank you for making me feel at home in Amsterdam and the valuable friendship. It was a pleasure to share my office and exchange the teaching of new languages (Chinese, Lithuanian and Turkish) with my office-mates Ona and Yun in between our works. It was a nice coincidence to meet here in the Netherlands with my bachelors classmate, Oznur. I loved all the nice chats during our lunch breaks. I thank Asli, Umut, Yildiz, Taylan, Sait, Kleopatra as well as all other class-mates I got to know in Amsterdam. Some, Phillipe, Charles, Pedro, left us early to go back home, but it is nice to still keep in touch. It was such a great pleasure to get to know different people from different cultures, from Russian to Brazilian, from Chinese to Ghanan. I am sure we will have opportunities to meet in the future. I thank Kim, Amelie, Cle who contributed to my positive thinking and helped me increase my endurance during PhD with enjoyable extracurricular activities.

Last but not the least, I must thank my family for the unlimited support that they have provided me with all through my study life. My grandfather has been a friend with whom I exchanged knowledge frequently and whom I follow as a role model. I cannot thank him enough for his invaluable support. My mother and father both knew that I had an aim in my life and have been supportive to me although, I must admit, they were a bit tired of my studies from time to time. I was lucky to meet Alper, who, among other things, helped me realize that I indeed do not have a bad memory:) I also realized that I did not need a mirror to observe myself in all aspects in his presence. I should thank him for all the support.

S.Pinar Ceyhan
Rotterdam, 2014

Contents

Chapter 1

Introduction and Outline

1.1 Introduction

Knowing the history of your topic of interest is important: It teaches what happened in the past, helps to understand the present, and allows one to look ahead in the future. Given my interest in the development of Bayesian econometrics, this thesis starts with a description of its history since the early 1960s.

My aim is to quantify the increasing popularity of Bayesian econometrics by performing a data analysis in the sense of measuring both publication and citation records in major journals. This will give a concrete idea about where Bayesian econometrics came from and in which journals its papers appeared. With this information, one will be able to predict some future patterns. Indeed, the analysis indicates that Bayesian econometrics has a bright future.

I also look at how the topics and authors of the papers in the data set are connected to each other using the bibliometric mapping technique. This analysis gives insight in the most important topics examined in the Bayesian econometrics literature. Among these, I find that a topic like unobserved components models and time varying patterns has shown tremendous progress. Finally, I explore some issues and debates about Bayesian econometrics.

Given that the analysis of time varying patterns has become an important topic, I explore this issue in the following two chapters. The subject of Chapter 3 is twofold. First, I give a basic exposition of the technical issues that a Bayesian econometrician faces in terms of modeling and inference when she is interested in forecasting US real GDP growth by using a time varying parameter model using simulation based Bayesian inference. Having observed particular time varying patterns in the level and volatility of the series, I propose a time varying parameter model that incorporates both level shifts and stochastic volatility components. I further try to explain the GDP growth series using survey data on expectations. Doing posterior and predictive analyses, the forecasting performances of several models are compared. The results of this chapter may become an input for more policy oriented models on growth and stability.

In addition to output growth stability, price stability is also an important policy objective. Both households and businesses are interested in the behavior of prices over time and follow the decisions of policymakers in order to be able to make sound decisions. Moreover, policymakers are interested in making inflation forecasts to be able to make sound policy decisions and guide households and businesses. Therefore, inflation forecasting is important for everybody. I deal with this topic in Chapter 4. In this chapter, I explore forecasting of US inflation via the class of New Keynesian Phillips Curve (NKPC) models using original data. I propose various extended versions of the NKPC models and make a comparative study based on posterior and predictive analyses. I also show results from using models that are misspecified and from using survey inflation expectations data. The latter is done since most macroeconomic series do not contain strong data evidence on typical patterns and using survey data may help strengthening the information in the likelihood. The results indicate that inflation forecasts are better described by the proposed class of extended NKPC models and this information may be useful for policies such as inflation targeting.

Section 1.2 summarizes the contributions of this thesis. Section 1.3 presents an outline of the thesis and summarizes each chapter.

1.2 Contributions

The contributions of this thesis may be classified into two groups. The first group refers to the quantitative analysis of the developments in Bayesian econometrics reported in Chapter 2. New empirical analysis on the publication patterns of Bayesian econometrics give valuable information on its positive development. In terms of more methodology, I make use of recent methods from connectivity analysis in this chapter. Network and connectivity maps are used to find the connectivity among subjects in Bayesian econometrics where distance-based and graph-based maps are used in bibliometric research. I make use of the VOSviewer program, available at `http://www.vosviewer.com` and their software to address proximity, see Waltman et al. (2010); Van Eck and Waltman (2010). In distance-based maps, the strength of relationships between topics is reflected in the distance between their locations in the map: the closer they are, the stronger the topics are related. In addition to this, there are various ways to display a map, one of which is the label view. In the visualization, Bayesian econometrics topics are depicted by a label and a circle. The bigger the label and the circle around it, the more important the topic is. Other ways I use to display a map are the density view and the cluster density view in which the color of a topic depends on its density, i.e., the number and importance of topics surrounding it.

The second group of contributions is made in Chapters 3 and 4, and refers to the use of state of the art techniques and new sources of prior and data information in order to obtain new empirical results on growth and inflation in US macroeconomic series. For posterior analyses, Gibbs sampling with Metropolis-Hastings steps and data augmentation are used in models where new specifications are given on time varying conditional means and variances incorporating information from survey data.

More specifically, to be able to carry out our inferences on unobserved variables of interest (conditional on the model parameters and the information set), I make use of state space modeling. For this purpose, I use the Kalman filtering and smoothing techniques. Kalman filtering is a recursive procedure in which an optimal estimate of the state vector is obtained given all the current information. After reaching the end of the series, future observations can optimally be predicted. Smoothing is a backward recursion which obtains the optimal estimates of the state vector at each time period by using the whole sample. One advantage of the filtering and smoothing techniques is that they have computational advantages, i.e., there exist efficient computing algorithms to get results about NKPC models under the assumption of Calvo pricing and its hybrid version. These models are extended to model both the low and high frequency components of the series explicitly. The time varying behavior (time varying mean, level shifts in the mean, stochastic volatility) of inflation and real marginal cost series are modeled to extend the basic growth models in Chapter 3 and the basic Phillips Curve models in Chapter 4. Furthermore, I use survey data to model the GDP growth in Chapter 3 and I use inflation expectations in the hybrid NKPC model of Chapter 4.

1.3 Outline and Summary

This thesis consists of three chapters. Having gained insight into Bayesian econometrics from its birth until recently in the second chapter, the thesis continues with two chapters proposing models for two series using the techniques which are suggested from the analysis in Chapter 2.

Chapter 2 is based on the paper Baştürk, Çakmaklı, Ceyhan and Van Dijk, (2013a), which is about the history of Bayesian econometrics since 1960s until 1 April 2014. The popularity of empirical probability assessments in economics has its roots in the early steps of Bayesian econometrics in the 1960s after the likelihood based inference reported in the Cowles Foundations Monographs 10 and 14. Many of the papers in these monographs applied R.A. Fisher's likelihood approach to a system of simultaneous equations. However, during the first half of the 1970s, this research area experienced breaks. There appeared the issue of modelling data with high persistence and time varying volatility, vector autoregression models and new simulation methods based on Importance Sampling and Markov Chain Monte Carlo were developed. As a result, the econometrics literature saw the birth of new research areas with important methodological and practical consequences for economic analysis, forecasting and decision strategies.

After a brief description of the first Bayesian steps into econometrics in the 1960s and early 70s, publication and citation patterns are analyzed in ten major econometric journals until 1 April 2014. For the analysis of publication patterns, the number of pages written on Bayesian econometrics are recorded as a percentage of total number of pages for each year for all journals. The results indicate that journals which contain both theoretical and applied papers, such as *Journal of Econometrics*, *Journal of Business and Economic Statistics* and *Journal of Applied Econometrics*, publish the large majority of high quality Bayesian econometric papers in contrast to theoretical journals like *Econometrica* and *the Review of Economic Studies*. These latter journals published, however, some high quality papers that had a substantial impact on Bayesian research. Moreover, I observe that beginning in the 1990s, there is an increasing trend in the publication patterns due to the increase in the computational power.

Next, a visualization technique is used to connect papers and authors around important theoretical and empirical themes such as forecasting, macro models, marketing models, model uncertainty and sampling algorithms. The distance between topics depends on the number of times that the keywords and names appear together or in relation to all keywords and names cited together. The information distilled from this analysis shows also names of authors who contribute substantially to particular themes. This is followed by a discussion of those topics that pose interesting challenges for discussion amongst Bayesian econometricians, namely the computational revolution, unobserved component and flexible model structures, choice models, IV models,the issue of identification and prior information, dynamic models and forecasting. Three issues are summarized where Bayesian and frequentist econometricians differ: Identification, the value of prior information and model evaluation; dynamic inference and nonstationarity; and vector autoregressive versus structural modeling. A major topic of debate amongst Bayesian econometricians is listed as objective versus subjective econometrics, and communication problems and bridges between statistics and econometrics are summarized.

The chapter ends with a list of four important themes that will be a challenge for the twenty-first century Bayesian econometrics: Sampling methods which are suitable for parallelization and GPU calculations, complex economic models which can account for nonlinearities, analysis of implied model features such as risk and instability and incorporating model incompleteness in econometric analysis. I further predict that the popularity of Bayesian econometrics will continue to increase over time.

Chapter 3 is based on Basturk, Ceyhan and Van Dijk (2014). This chapter proposes models for the US real GDP growth series that can be used for forecasting. Time varying patterns in US growth are analyzed using various univariate model structures, starting from a naive model structure where all features change every period to a model where the slow variation in the conditional mean and changes in the conditional variance are specified together with their interaction, including survey data on expected growth in order to strengthen the information in the model. Results indicate that incorporating time variation in mean growth rates as well as volatility of mean growth rates are important for improving the predictive performances of standard growth models. Furthermore, using data information on growth expectations is important for forecasting growth in specific periods, such as the 2000s and around 2008.

Chapter 4 is based on the paper Baştürk, Çakmaklı, Ceyhan and Van Dijk, (2013b). This chapter proposes extensions to the NKPC model, which shows the relationship between inflation and economic activity, and can be used for inflation forecasting. In this chapter, I show that mechanical removal or modeling of simple low frequency movements in the data may yield poor predictive results which depend on the model specification used. Therefore, I model the changing time series properties of US inflation and economic activity, measured as marginal costs, within a set of extended NKPC models, allowing for level shifts and stochastic volatility in inflation and log marginal costs, and applied to quarterly U.S. data over the period 1960-I until 2014-I. I incorporate forward and backward looking expectation components for inflation and evaluate their relative importance. Finally, I model the unobserved inflation expectations using the University of Michigan survey data.

For all models considered, simulation based Bayesian techniques are used for the empirical analysis. Choice among the proposed models depends on the criteria predictive likelihood and Mean Squared Forecast Error (MSFE). The results suggest that the NKPC models which incorporate structural time series features and an inflation expectation term and use survey data for inflation expectations capture the time variation in the inflation and real marginal cost series well. The proposed models perform better than the basic NKPC models using demeaned and/or detrended data as well as the standard stochastic volatility model of Stock and Watson (2007) and extended BVAR models. No credible evidence is found on endogeneity and long run stability between inflation and marginal costs. The backward looking term in the hybrid NKPC model dominates the forward looking term. Levels and volatilities of inflation are estimated more precisely using rich NKPC models. Tails of the complete predictive distributions indicate an increase in the probability of deflation in recent years.

Chapter 2

Historical Developments in Bayesian Econometrics after Cowles Foundation Monographs 10, 14

Chapter 2 is based on Baştürk, Çakmaklı, Ceyhan and Van Dijk, (2013a).

2.1 Introduction

Bayesian econometrics is now widely used for inference, forecasting and decision analysis in economics, in particular, in the fields of macroeconomics, finance and marketing. Three practical examples are: International corporations that sell their goods abroad want to know the risk of foreign exchange rate exposure that they incur at the time they repatriate the proceeds of their sales, see Bos, Mahieu and Van Dijk, (2000); in modern macroeconomics the risk of a liquidity trap, defined as low inflation, low growth and an interest rate close to the zero lower bound, is relevant for an adequate economic policy for several countries; evaluating the effect of a new pricing policy is highly relevant in decision strategies of supermarket chains. Particular references and more examples are given in textbooks like Geweke (2005); Rossi et al. (2005) and Koop et al. (2007). This widespread interest in and the use of empirical probability assessments of important economic issues has come a long way from the early steps of Bayesian econometrics in the 1960s following the likelihood based inference reported in the brilliant Cowles Foundations monographs 10 and 14, see Koopmans (1950) and Hood and Koopmans (1953). Papers in these monographs applied the likelihood approach introduced by R.A. Fisher (Fisher, 1912, 1922) to, predominantly, a system of simultaneous equations where immediate feedback mechanisms or, otherwise stated, a set of jointly endogenous variables posed substantial methodological challenges to estimate systems of equations. In the early and middle part of the 1970s there were several shocks to this line of research. Data series exhibited novel features like strong persistence and time varying volatility; new modeling like the vector autoregressive approach was developed and novel simulation based inferential techniques based on Importance Sampling and Markov Chain Monte Carlo were introduced. This opened a wide set of new research lines that had substantial methodological and practical consequences for economic analysis, forecasting and decision strategies. Structural economic models based on dynamic stochastic general equilibrium concepts, unobserved component models allowing for time varying parameters and using data augmentation, simulation methods and increased focus on the complete forecast distribution and in particular on the tails of the distribution like in Value-at-Risk are only a few examples in this respect.

In the present chapter we sketch historical developments of Bayesian econometrics from the early 1960s until 2014 by collecting and analyzing the publication and citation patterns on Bayesian econometric papers in ten major econometric journals until the end of March 2014. The number of pages written on Bayesian econometrics were recorded as a percentage of total number of pages for each year for all journals. The results indicate that journals which contain both theoretical and applied papers, such as *Journal of Econometrics*, *Journal of Business and Economic Statistics* and *Journal of Applied Econometrics*, publish the large majority of high quality Bayesian econometric papers in contrast to theoretical journals like *Econometrica* and *the Review of Economic Studies*. These latter journals published, however, some high quality papers that had a substantial impact on Bayesian research. The journals *Econometric Reviews* and *Econometric Theory* published key invited papers and/or special issues that received wide attention, while *Marketing Science* shows an ever increasing number of papers since the middle 90s. The *International Economic Review* and the *Review of Economics and Statistics* show a moderate time varying increase. It is noteworthy that since the early 90s there exists an upward movement in publication patterns in most journals probably due to the effect of the 'Computational Revolution'.

Next, a visualization technique is used to connect papers and authors around important themes in theoretical and empirical econometrics. The proximity of topics that we consider is defined by the number of times that the keywords or names appear together or in relation to all keywords and names cited together. The results show the interconnectedness of several topics of interest. Macroeconomics and finance literature is related to simulation and filtering methods as well as methods dealing with model uncertainty.

Macro models used for policy purposes are related to identification issues. Marketing models are linked to flexible model and prior structures such as hierarchical Bayes and Dirichlet processes. The information distilled from this analysis shows also names of authors who contribute substantially to particular themes. This is followed by a discussion of those topics that pose interesting challenges for discussion amongst Bayesian econometricians. The effects of the computational revolution on Bayesian econometrics, advances in several fields, such as flexible and unobserved component model structures, dynamic models and forecasting, the issue of identification and prior information, as well as three main debates between Bayesian and frequentist econometricians are presented. There is probably not a subject area in the literature where everything goes smoothly. Bayesian econometrics is no exception. We only summarize major debates and issues encountered in the Bayesian econometrics literature in the latter part of the twentieth century. The chapter ends with a list of important themes that are predicted to be a challenge for the twenty-first century Bayesian econometrics. These refer to big data, model complexity, parallel computing and model incompleteness.

Influential papers in Bayesian econometrics have been analyzed in Poirier (1989, 1992), which provide quantitative evidence of the impact of the Bayesian viewpoint as measured by the percentage of pages devoted to Bayesian topics in leading journals. We contribute to this literature by extending the bibliographical data with more recent papers and additional leading journals. Our contribution differs from the literature in several ways. First, regarding the influential papers in the field, we consider an alternative measure, the number of citations of each paper in addition to the percentage of pages devoted to Bayesian topics. The impact of papers are found to be different according to the criteria chosen for this purpose. Second, we define the set of influential papers in the field relying on the references in Geweke, Koop and Van Dijk, (2011), who provide an up-to-date set of references in the field with their respective subfields, such as macroeconomics and computational advances. Third, we consider the clustering of papers in the field without defining a measure for their influence. This analysis is based on online bibliographic databases, and the results are not affected by the subjective definition of influential papers in the field of Bayesian econometrics.

The chapter is organized as follows: In Section 2.2 some first Bayesian steps into the field of econometrics are listed. Section 2.3 analyzes the publication and citation patterns of the papers in our database. Section 2.4 visualizes the connectivity of the topics in those papers, linking these to the key topics of interest during the period 1962–2014. Section 2.5 is devoted to several issues and debates that are pivotal in the history of Bayesian econometrics. Section 2.6 contains the authors' expectations about the future of Bayesian econometrics and concludes. We note that many more interesting Bayesian empirical papers exist and are published in major economics journals. These are, however, outside the scope of the present chapter.

2.2 Cowles Foundation Research and Early Bayesian Econometric Steps

The Cowles Foundation Monographs 10 and 14, published in 1950 and 1953, respectively, laid the foundations for modern inference in econometrics after the famous Haavelmo paper (Haavelmo, 1944) on the probability approach to econometrics. Although Haavelmo listed two interpretations of probability: the 'frequentist' concept and the 'a priori confidence' one, see Haavelmo (1944) pp. 48, the first one was used henceforth in econometrics. The focus in Monographs 10 and 14 was on developing the method of maximum likelihood for systems of simultaneous equations. Identification issues, full information maximum likelihood, limited information maximum likelihood, the corresponding numerical optimization methods to find the maximum and also some basic time series problems were analyzed. Among the contributors to this area are Koopmans (1945), Anderson (1947), Anderson and Rubin (1949), Hurwicz (1950), Chernoff and Divinsky (1953) and Chernoff (1954).

Much of this research followed the likelihood approach from R.A. Fisher. Although Fisher rejected Bayesianism, he had an alternative: so-called fiducial inference (see Fisher (1973) and Aldrich (1995) for a review), which was Fisher's attempt to use inverse probability and to analyze the shape of the likelihood without stochastic prior information. This has been characterized by Savage as: 'A bold attempt to make the Bayesian omelette without breaking the Bayesian eggs' (Savage, 1961).

The frequentist interpretation of estimators obtained by using the likelihood approach became known as the 'classical approach' in econometrics. As argued, in Cowles Foundation Monograph 23, by Rothenberg, this classical approach has some very restrictive assumptions, see Rothenberg (1973). First, for efficiency, accuracy, and credibility one usually makes use of the Crámer-Rao lower bound of the variance of the estimator. This holds only for unbiased estimators and is in most cases only asymptotically valid. Second, one makes use of prior conditioning information as exact restrictions, which is often unrealistic and overly restrictive. Third, this ignores completely the decision aspect of inference.

Given all the work on the implementation of the likelihood approach to econometrics and the recognition of its limitations, it is very natural that the Bayesian approach would follow.[1] Among early contributions to this literature are three important books. The first one is Raiffa and Schlaifer (1961) who introduced the concept of conjugate analysis as a way to construct informative prior information on model parameters. The idea is that the model was already in existence in the period before the data were observed or alternatively that the model was in existence for related data sets in other countries or for a different set of agents with similar features. A second book was Schlaifer (1959), later summarized in Pratt et al. (1995), where practical decision problems were explained and analyzed. Here a connection with the field of finance and business was made. Thirdly, there came the very influential 'Bible' of analytical results in Bayesian econometrics by Zellner (1971). All econometric models that were in use at that time were analyzed in this classic book from a Bayesian perspective. Analogies and differences between the classical and Bayesian approach were discussed, often using a weak or non-informative prior approach.

Following these early Bayesian steps, there were several issues that attracted attention in the literature. We document four of these issues in this section.

Rothenberg's problem and natural conjugate priors

The natural conjugate family gives an analytically tractable and convenient family of prior densities for the case of the standard linear regression model. This density is known as the normal-inverted gamma density and it allows to update prior information using Bayes theorem in a simple way: The posterior mean of the regression parameters is a weighted average of prior mean and data mean with weights that are given as the relative accuracy of prior and data mean, respectively.

However, for a system of regression equations the corresponding conjugate prior has an important restriction on the prior variances of the regression equations parameters. This can be shown as follows. Consider the system

$$Y = X\Pi + V, \tag{2.1}$$

where Y and X are appropriate data matrices from T observations and Π is a parameter matrix and the columns of V are $\text{NID}(0, \Sigma)$ where NID denotes the independent

[1] Note that in post World War-II econometrics one has, apart from the likelihood approach, the (dynamic) regression methods and GMM as major schools of econometric inference. We only refer to the likelihood approach in the present chapter.

normal distribution. As shown in Rothenberg (1963), the likelihood of $(\Pi, \Sigma|Y, X)$ can be written as

$$f(\Pi, \Sigma|Y, X) \propto |\Sigma|^{-\frac{T}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}\Sigma^{-1}\left((\Pi - P)'(X'X)(\Pi - P) + TS\right)\right) \tag{2.2}$$

where P and S are the maximum likelihood estimators: $P = (X'X)^{-1}X'Y, \quad S = \frac{1}{T}(Y - XP)'(Y - XP)$.

A prior density that belongs to the conjugate family and that is proportional and of the same functional form as the likelihood in (2.2) implies that the variance of π_{ij} is given as

$$\mathrm{Var}(\pi_{ij}) = \sigma_{ii}a_{jj}, \tag{2.3}$$

where σ_{ii} is the ith diagonal element of Σ and a_{jj} is the jth diagonal element of $(X'X)^{-1}$.

It follows that one has the restriction

$$\frac{\mathrm{Var}(\pi_{ri})}{\mathrm{Var}(\pi_{si})} = \frac{\mathrm{Var}(\pi_{rj})}{\mathrm{Var}(\pi_{sj})}, \tag{2.4}$$

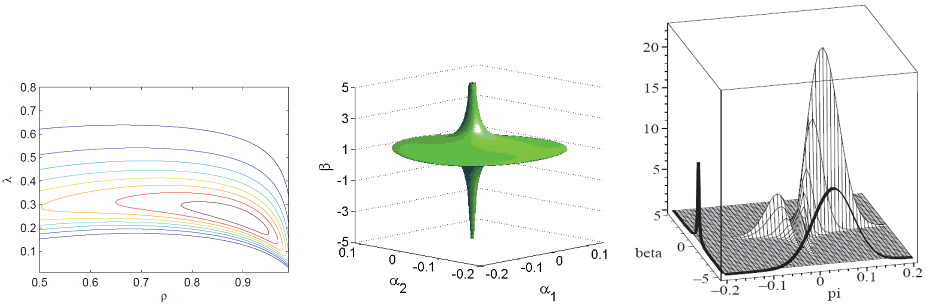
that is, the variances of the parameters in the r^{th} equation have to be proportional to the variances of the corresponding parameters in the s^{th} equation. There is no a priori economic reason why this mathematical restriction should be the case. This creates a problem for inference in systems of equations, like Seemingly Unrelated Regression Equations (SURE) and Vector AutoRegressive (VAR) models, and it is known as ‘Rothenberg’s problem’, see Rothenberg (1963, 1973).

In further research Drèze and Richard followed a path to limit this restriction by either concentrating the analysis on a single equation within a system of equations which is known as the Bayesian limited information approach, see Drèze (1976) and Bauwens and Van Dijk, (1990) for details, or by extending the natural conjugate family for a system of equations, see Drèze and Richard (1983). However cavalier these approaches were, the restriction on the natural conjugate prior family is a severe one and different attempts in research were started to free the analysis from analytical restrictions.

Monte Carlo gives freedom from analytical restrictions and allows for evaluation of uncertainty of policy effectiveness

Monte Carlo simulation has freed the Bayesian approach from very restrictive model structures: It allows Bayesians to apply their inference to a wide range of complex mod-

Figure 2.1: Examples of complex (non-elliptical) posterior distributions



els in many scientific disciplines. Figure 2.1 shows examples of complex (non-elliptical) posterior distributions in models for realistic problems in economics. The posterior distributions occur in finance (modeling daily stock returns), macroeconomics (modeling the joint behavior of variables with a long run equilibrium relationship), and microeconomics (modeling the effect of education on income), see Hoogerheide, Opschoor and Van Dijk, (2012) and Baştürk, Hoogerheide and Van Dijk, (2013c).

Using novel simulation methods one can obtain reliable and accurate estimates of the properties of interest of such posterior distributions. Importance Sampling, introduced into statistics and econometrics by Kloek and Van Dijk, (1975) and later published as Kloek and Van Dijk, (1978), or the independence chain Metropolis-Hastings algorithm, Metropolis et al. (1953) and Hastings (1970), can be used.

Apart from this need to free the Bayesian approach from restrictive priors, there also existed interest in evaluating uncertainty of policy effectiveness. An important example was to obtain the posterior distribution of the multiplier in a system of simultaneous equations, see Brainard (1967). Here one faces the issue that this multiplier is usually a ratio or more general a rational function of structural parameters. The Monte Carlo method approach gives here an easy operational approach to obtain the finite sample distribution, see e.g. Van Dijk, and Kloek (1980).

Testing, signifying nothing, sequential testing and credibility of the final chosen model

The focus on statistical testing in econometrics in the past fifty years has not yielded substantial confidence in obtained results. The focus on statistical testing regularly means that the researcher does not raise the issue whether the results matter from an economic point of view. Statistically significant but economically almost meaningless is something that decision makers will not accept as a sound basis for policy analysis, see McCloskey and Ziliak (1996) and The Economist (2004). A more fundamental statistical weakness of the 'classical approach' is the testing of many different hypotheses in econometric models by a sequential testing procedure. In the analysis it is usually not taken into account that the distribution of the second test depends on the outcome of the first one and so on for further tests. In the end a model is 'accepted' without stating a measure of 'credibility' of the final result. A natural Bayesian solution is to give weights to particular model features by using Bayesian model averaging and pursue forecasting with a weighted average of model structures. This line of research is shown in, e.g. Wright (2008) and Strachan and Van Dijk, (2013).

Conditional probability statements on features of interest

Another fundamental problem with the classical approach is the difficulty of dealing with the issue of conditional probability statements which is a concept that is widely used in practice. Given a set of data decision makers are usually interested in the probability of an unknown feature of a problem at hand. The earlier listed examples are clear indications: how to hedge currency risk for international corporations given data on exchange rate behavior; what monetary policy to be used in the face of a liquidity trap given data for countries like Japan and the European Union; and which advertising policy given scanner data about customer behavior are relevant. The simulation based Bayesian approach is very suitable for such conditional probability statements.

These issues will be dealt with in more detail in section 2.5.

2.3 Exploratory Data Analysis

In this section we analyze the advance of Bayesian econometrics since the middle of the 1970s from a descriptive point of view. Specifically, we analyze how Bayesian econometrics got in the mainstream and high quality econometric journals using the publication and citation patterns of 999 papers in leading journals during the period between 1978 and 2014 (March). We select these papers on the basis of their contributions to theoretical or applied topics in Bayesian econometrics and denote them by 'Bayesian papers'. The list of leading journals consists of 10 journals: *Econometrica* (Ectra), *Econometric Reviews* (ER), *Econometric Theory* (ET), *International Economic Review* (IER), *Journal of Applied Econometrics* (JAE), *Journal of Business and Economic Statistics* (JBES), *Journal of Econometrics* (JE), *Marketing Science* (MS), *Review of Economic Studies* (RES) and *Review of Economics and Statistics* (ReStat). Our analysis extends the one from

Poirier (1992) by including more journals and a longer period. We also make use of citation patterns. Detailed statistics of the papers considered in this section are provided in Appendix 2.A, Tables A.1, A.2, A.3 and A.4.

2.3.1 Publication Patterns

The first criterion we use to analyze the advances in Bayesian econometrics is the percentages of Bayesian pages in the leading econometrics and quantitative economics journals that were listed above.

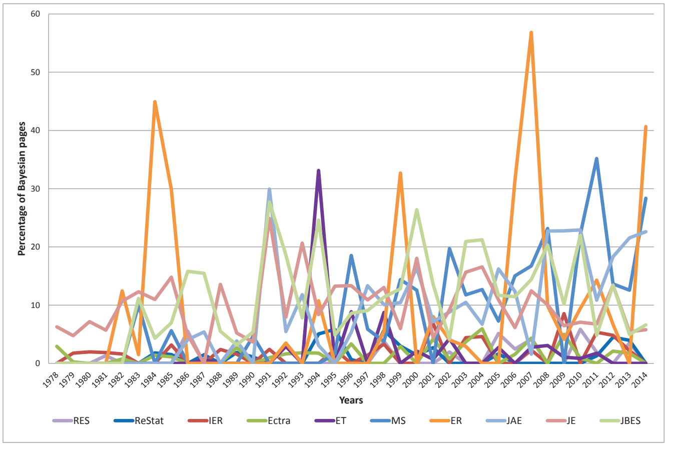
The top panel in Figure 2.2 presents the annual percentages of the pages allocated to Bayesian papers for each journal. These percentages are usually below 30%, with exceptions in ER, ET and MS. There are only two journal issues which have more than 40% Bayesian content. The ER issue in 1984 has four Bayesian papers constituting 44.93% of the total number of pages with as most influential paper the one by Doan et al. (1984). The 2007 issue of ER also has a high percentage of Bayesian pages, where 56.83% of the issue is devoted to 18 Bayesian papers, including An and Schorfheide (2007) as one of the longest papers.

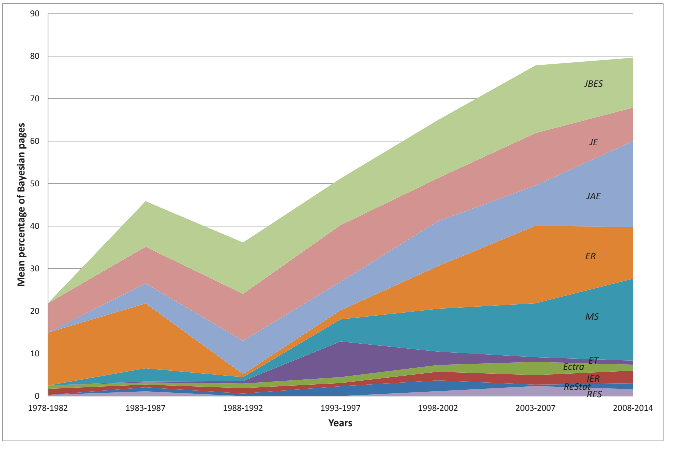
Special issues yield the highest values reported in the top panel of Figure 2.2. These are the ER issues in 1984, 1999, 2007; the ET issue of 1994 (on Bayes methods and unit roots); the JE issue in 1991, and to a lesser extent the ones in 1985, 2004, and 2012.

The bottom panel in Figure 2.2 presents the average percentage of pages for Bayesian papers in each journals over 5-year intervals. These average percentages provide general publication patterns compared to the top panel of Figure 2.2 since the influence of special journal issues related to Bayesian econometrics is now more limited due to the 5-year averaging.

The bottom panel of Figure 2.2 shows that the influence of Bayesian econometrics in terms of the percentage of allocated pages is time varying and journal dependent. Journals such as Ectra and IER typically have low percentages of Bayesian pages. On the other hand, JBES, JAE and MS typically have high percentages of Bayesian pages, with a substantial increase in these percentages after 1990s. The set of journals with a large share of Bayesian papers show that Bayesian inference is mainly present in applied papers rather than theoretical papers. Figure 2.2 indicates two main clusters of journals in terms of their focus on Bayesian econometrics. The first cluster consists of journals with relatively low average number of percentages of Bayesian pages: Ectra, ET, IER, RES and ReStat. The average percentages of Bayesian pages in these journals are less than 5%. The second cluster consists of journals with relatively high average percentages of Bayesian pages:

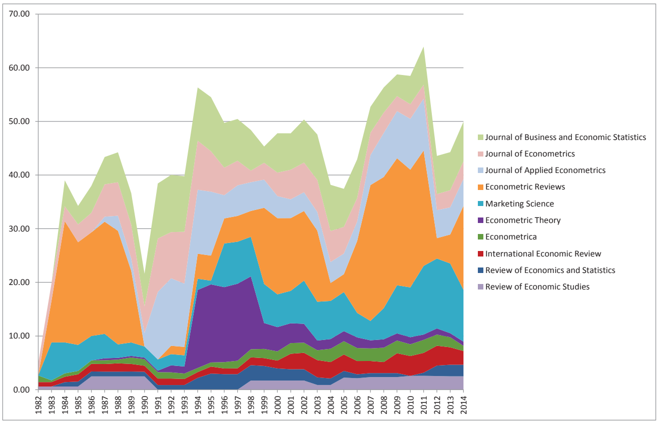
Figure 2.2: Percentages of pages allocated to Bayesian papers for all journals





Note: The figures present the annual percentage of pages (top panel) and 5-year averages of pages (bottom panel) of Bayesian papers for the period between 1978 and 2014 (March). In the bottom panel, the final period consists of 7 years. Abbreviations of journals are as follows: *Econometrica* (Ectra), *Econometric Reviews* (ER), *Econometric Theory* (ET), *International Economic Review* (IER), *Journal of Applied Econometrics* (JAE), *Journal of Business and Economic Statistics* (JBES), *Journal of Econometrics* (JE), *Marketing Science* (MS), *Review of Economic Studies* (RES) and *Review of Economics and Statistics* (ReStat).

Figure 2.3: Rolling st. deviations for the number of Bayesian pages for all journals



ER, JAE, JBES, JE and MS. The increased share of Bayesian pages is most visible for MS, then for ER and JAE, particularly for the period after 1992. To a lesser extent, this increasing pattern holds for Ectra, IER and RES. This general increasing influence of Bayesian econometrics after 1990s can be attributed to computational advances making Bayesian inference easier and the increased number of applied papers using Bayesian inference.

In summary, a high cluster of published Bayesian papers and a low cluster appear, where the low cluster refers to the more theoretical journals. Special issues contain more Bayesian papers and a structural break, indicating an increasing number of Bayesian papers, occurs since the early 1990s due the use of novel computational techniques.

One final point to note here is the variability in the number of Bayesian pages, which is partly caused by the special issues: in Figure 2.3, we can indeed see an increase in the standard deviation when a special issue in Bayesian Econometrics is published such as in 1984, 1999 and 2007 for Econometric Reviews and in 1994 for Econometric Theory.

2.3.2 Citation Patterns and Impact

We next focus on citation patterns of papers in the ten journals, as an additional criterion to define the advances in Bayesian econometrics after the 1970s.[2]

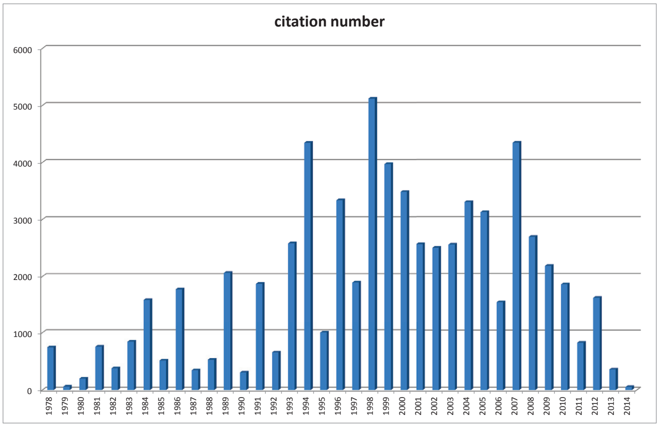
Not surprisingly, the data reveals an increasing pattern in the number of citations for Bayesian papers thanks to its increasing popularity. Additionally, we observe a clearly increasing pattern in the number of Bayesian papers cited over the years. These can be observed in the top and bottom panel of Figure 2.4, respectively. Figure 2.5 is a comparison of these series. We note three observations which stand out in Figure 2.4, namely 1994, 1998 and 2007. Looking at the bottom panel in Figure 2.4, we infer that the reason of the high citation number in 2007 is the high number of papers with modest citation numbers rather than few influential papers with high citation numbers. In 1998, there are fewer papers cited. However, the influential paper Kim et al. (1998) with 1459 citations contribute a lot to the high number of citations. Similarly, in 1994 the high number of citations accompanied by not so many papers reveals that there is at least one influential paper written in this year, which is indeed Jacquier et al. (1994) with 1347 citations.

The top and bottom left panels in Figure 2.7 present the number of citations for papers in leading journals during the period between 1978 and (March) 2014. The top and bottom right panels in the figure show the number of citations for a subset of these papers, with at least 400 citations over the sample period. We note that three papers are highly influential in the field with more than 1000 citations: Geweke (1989) (Ectra) with 1303 citations, Kim et al. (1998) (RES) with 1459 citations and Jacquier et al. (1994) (JBES) with 1347 citations. These papers refer to computational advances and financial econometrics with a focus on time varying volatility.

The top panel of Figure 2.6 shows that the impact analysis is substantially different when we compare the citation pattern with the number of Bayesian pages. Although the journals with a high share of Bayesian pages, ER, JAE, JBES, JE and MS, also have a high share in the total number of citations, there are two high quality theoretical journals, Ectra and RES, with a high influence in the field in terms of the citation numbers despite their relatively low numbers of Bayesian pages. This high impact is more visible when we focus on papers with more than 400 citations, shown in the figures on the right panel of Figure 2.6. There are 28 papers satisfying this criteria. Note that ER has a large share in total citations although its share in terms of the percentage of pages is more time varying.

[2]The citation numbers are collected using Google Scholar on 13–14 April, 2014, which are available at `http://scholar.google.com/`.

Figure 2.4: Histogram of citation numbers and number of papers cited for all journals



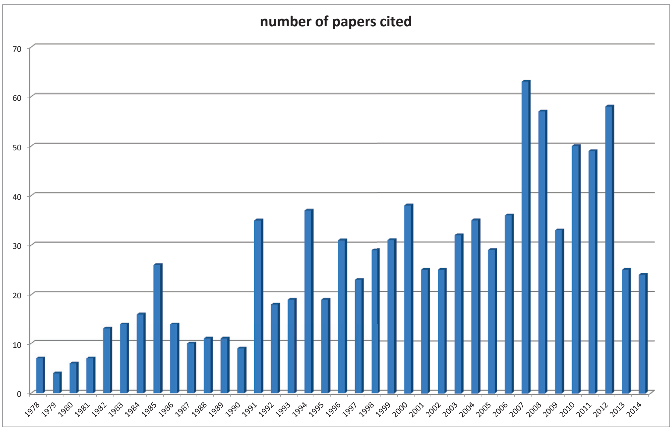
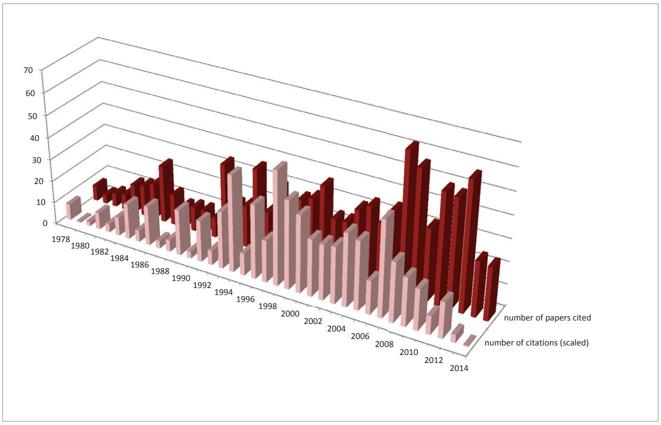


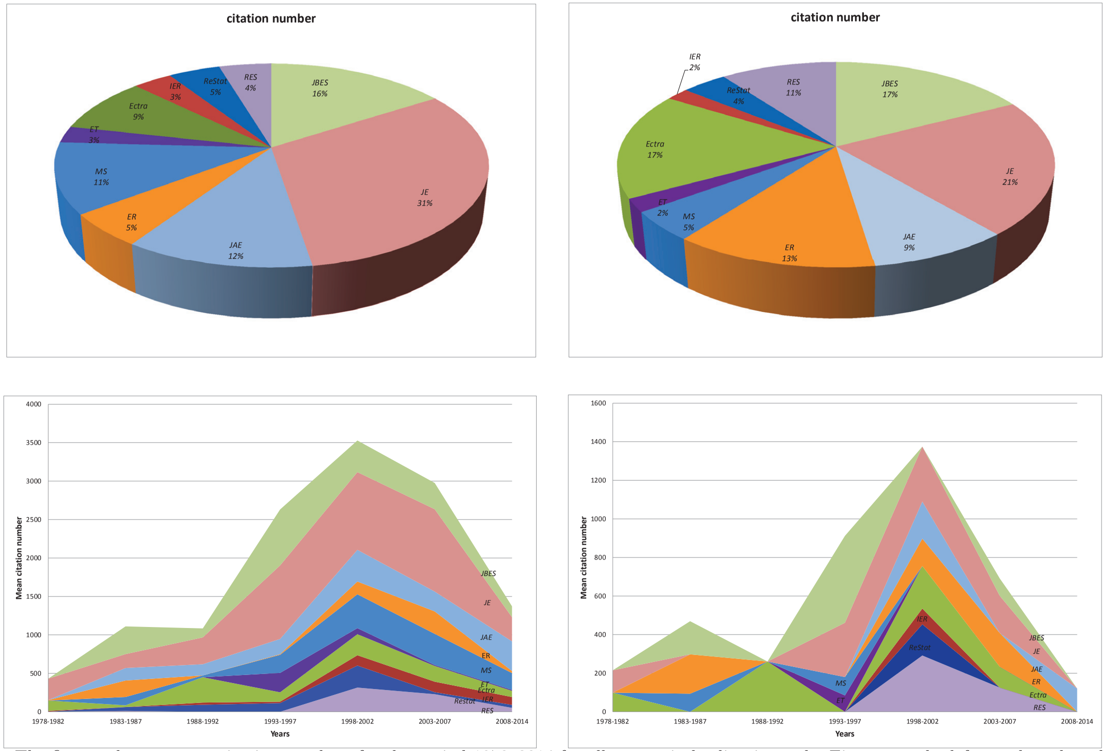
Figure 2.5: Histogram of citation numbers versus number of papers cited for all journals



The bottom left and bottom right panels in Figure 2.6 show average citations of papers over 5 year intervals for all cited papers in the journal list and for papers with at least 400 citations, respectively. We first note that the reported average number of citations are naturally low at the end of the sample period, especially between 2008–2014, since these papers are relatively new. When these recent papers are not taken into account, an increasing pattern in the overall number of citations for Bayesian papers is visible: The total numbers presented in the bottom panel of Figure 2.6 clearly increase between 1978 and 2002. These figures also show that *Marketing Science* papers started to be cited much more heavily after 1996 and in general the 1990s bring more citations to each journal in our data set.

In order to compare the influential papers in terms of their shares of pages and in terms of the number of citations, we next consider two clusters of journals according to publication patterns in section 2.3.1 and report the number of citations separately for the journals in these clusters. Figure 2.7 presents the total number of citations for all papers in leading journals and for a subset of influential papers with at least 400 citations for journals classified in cluster 1 and cluster 2. Cluster 1 consists of journals with a relatively low average number of Bayesian pages: Ectra, ET, IER, ReStat and RES. Cluster 2, on

Figure 2.6: Average citation patterns for papers in leading journals



The figures show average citation numbers for the period 1978–2014 for all papers in leading journals. Figures on the left panel are based on all papers published in leading journals. Figures on the right panel are based on a subset of influential papers with at least 400 citations. Reported years correspond to the years that the cited papers are published. Journal abbreviations are as in Figure 2.2.

the other hand, consists of journals with a relatively high average number of Bayesian pages: ER, JE, JBES, JAE and MS.

The left panel in Figure 2.7 shows that the journal classification in terms of the percentage of pages is informative about their influence in terms of citations. Papers in cluster 2, journals with a high number of pages devoted to Bayesian econometrics, are on average cited more than those in cluster 1. Despite this similarity, the right panel in Figure 2.7 shows that highly influential papers with at least 400 citations are more evenly distributed across cluster 1 and cluster 2 journals. Particularly, Ectra and RES have papers that are highly cited.

We conclude that journals which contain both theoretical and applied papers, such as JAE, JE and JBES, publish the large majority of high quality Bayesian econometric papers. Theoretical journals, such as Ectra and RES, on the other hand, publish papers which are highly influential and have a substantial impact on Bayesian research, although the number of these papers are relatively small. Special issues of journals like *Econometric Reviews* and *Econometric Theory* receive more citations than usual issues.

2.4 Subject Connectivity

This section considers connectivity of subjects in Bayesian econometric papers. The list of scientific papers in Bayesian econometrics is extensive. We first consider a large set of papers relying on digital archives in order to analyze subject connectivity. We use a random sample of 1000 papers and *key terms* extracted from each paper provided by the JSTOR digital archive in the field of Bayesian econometrics.[3] We next refine this extensive list of papers by focusing on the influential ones and summarize the connectivity of key subjects based on this refined set. For each set, the proximities are defined by the number of times that keywords or key terms appear together and in relation to their pairwise concurrence.[4]

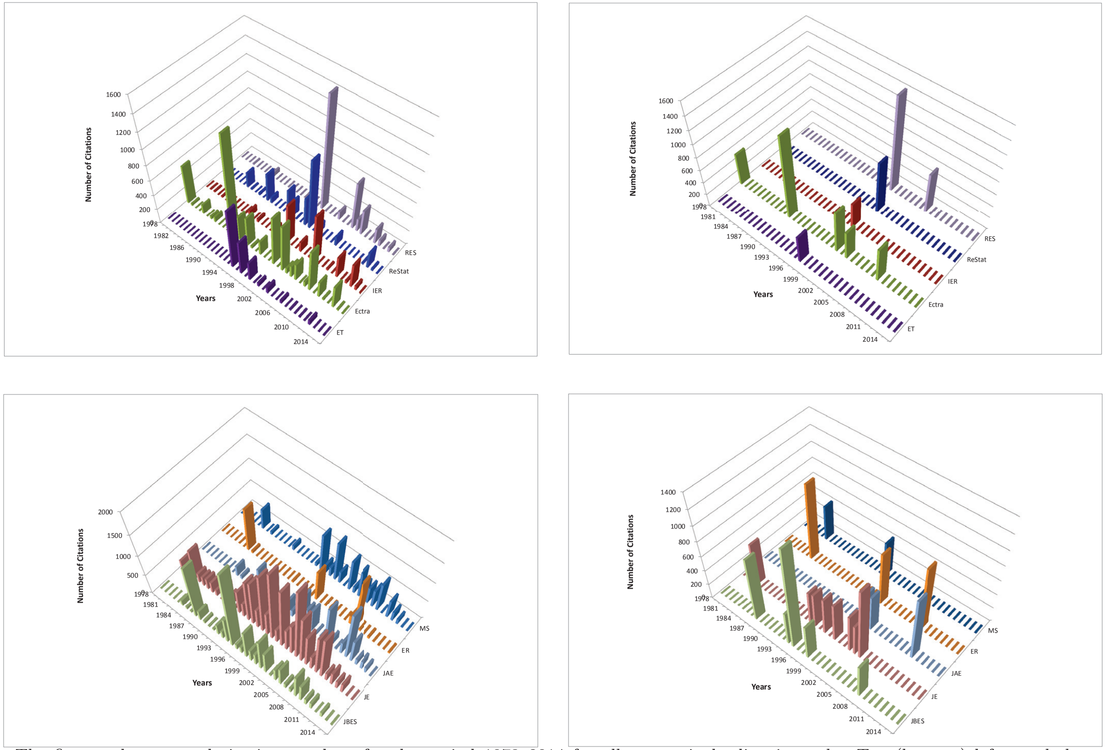
We consider two sets of influential papers in the field of Bayesian econometrics. The first set is taken from Geweke, Koop and Van Dijk, (2011) and has more than 100 citations according to Google Scholar or Web of Knowledge.[5] This selection of papers uses *expert information*, since the set of papers is based on the careful selection of the authors by

[3] `http://www.jstor.org/`

[4] The network maps we present are obtained from the VOSviewer program, available at `http://www.vosviewer.com` and their software to address proximity, see Waltman et al. (2010); Van Eck and Waltman (2010).

[5] `http://scholar.google.com/`, `http://thomsonreuters.com/web-of-science/`.

Figure 2.7: Citation numbers for journals with high and low numbers of pages devoted to Bayesian econometrics



The figures show annual citation numbers for the period 1978–2014 for all papers in leading journals. Top (bottom) left panel shows citations of papers in journals with a low (high) percentage of Bayesian pages. Top (bottom) right panel shows the citations of papers with at least 400 citations in journals with a low (high) percentage of Bayesian pages. Reported years correspond to the years that the cited papers are published. Journal abbreviations are as in Figure 2.2.

Geweke, Koop and Van Dijk, (2011). The second set of influential papers are based on the extension of Poirier (1992) presented in section 2.3.

Figure 2.8 presents the network and connectivity map for the key terms based on 1000 random papers in Bayesian econometrics, published since 1970. This connectivity analysis is solely based on a random sample of papers that JSTOR provides. Three major areas emerge from this connectivity analysis and these are presented in different colors in Figure 2.8. The first cluster of keywords, plotted in dark and light green in the figure, corresponds to theoretical topics, with related keywords of likelihood, moment, statistics, assumption and probability. This cluster is naturally linked to all remaining clusters. The second area, consisting of clusters colored in blue and purple, is centered around the key terms 'forecasting' and 'price'. This cluster shows that particularly forecasting is central to the analysis of macroeconomic and financial data, such as (economic) growth, exchange (rates), (financial) returns and interest (rates). Most common models for these data include autoregressive models. (Forecast) horizon, regime (changes) and testing are related and important issues for this area. The third area, shown in light and dark red in Figure 2.8 has as most prominent key terms 'market', 'choice', 'information' and 'equilibrium'. Other keywords in this area, such as decision, brand, profit, behavior, equilibrium and utility signal market equilibrium models as well as choice models.

The connectivity analysis presented so far does not take into account the amount of influence of each paper, the papers' original keywords or any extra information on the subject area. We next consider a refined set of influential papers in Bayesian econometrics, based on the citations in Geweke, Koop and Van Dijk, (2011). Note that the topics and the references covered in Geweke, Koop and Van Dijk, (2011) are divided to 9 chapters according to the subfields: endogeneity & treatment (Chapter 1), heterogeneity (Chapter 2), unobserved components & time series (Chapter 3), flexible structures (Chapter 4), computational advances (Chapter 5), micro & panel data (Chapter 6), macro & international economics (Chapter 7), marketing (Chapter 8) and finance (Chapter 9). We consider keywords of Bayesian papers cited in each chapter, and include the corresponding subfield as an additional keyword for each paper.

Figure 2.9 shows that the subfields defined in Geweke, Koop and Van Dijk, (2011) are connected to several keywords. This is an expected outcome since we use the chapter information in Geweke, Koop and Van Dijk, (2011) as 'expert knowledge' to relate each paper to a subfield. Besides these subfields, sampling techniques such as the Gibbs sampler, Markov Chain Monte Carlo (MCMC), Metropolis Hastings (MH) algorithm and importance sampling have very large weights indicated by the sizes of the circles in Figure 2.9

Figure 2.8: Connectivity of topics in JSTOR database

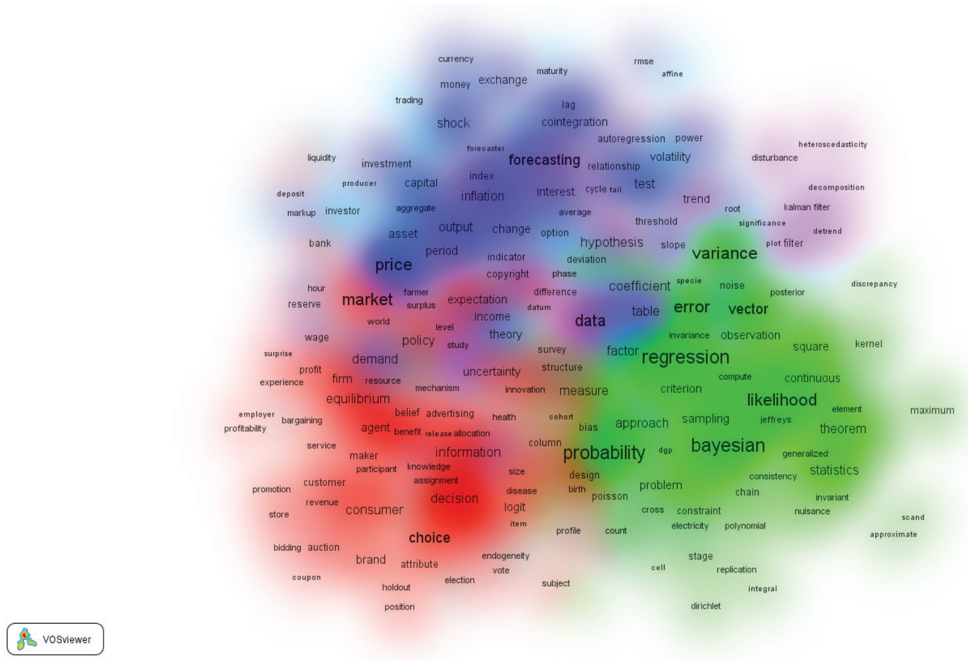
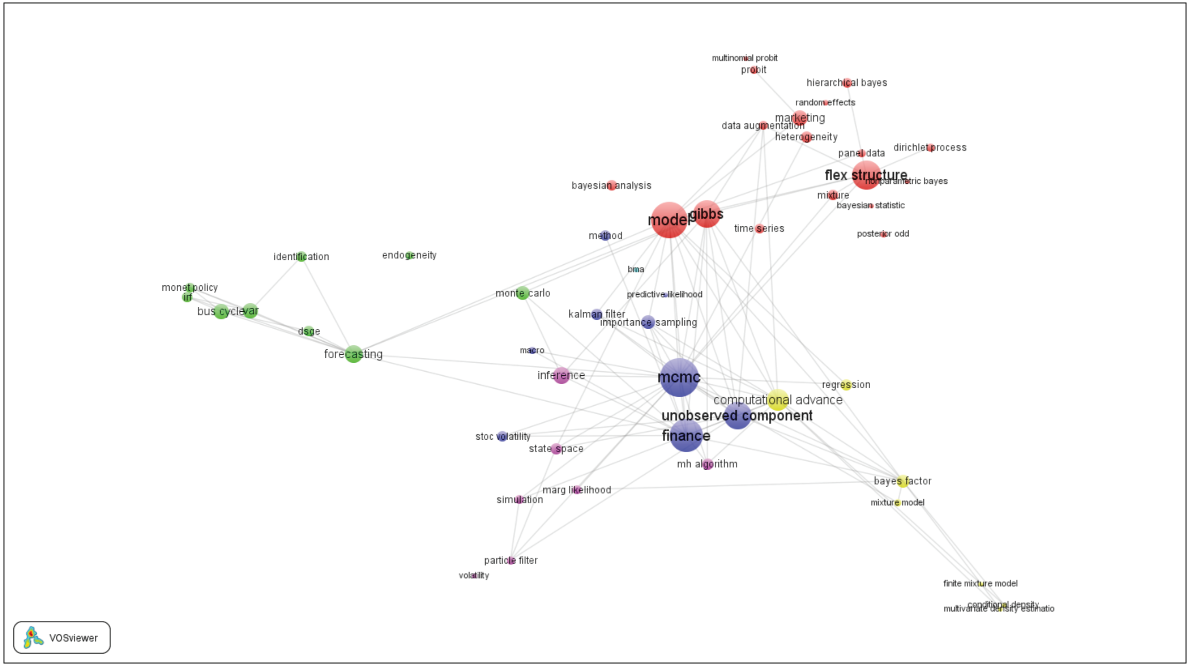


Figure 2.9: Connectivity of topics in papers cited in Geweke, Koop and Van Dijk, (2011)



and they lie in the middle of the keyword connection map. This indicates that sampling algorithms are central to research in all subfields of Bayesian econometrics covered here.

An interesting result from Figure 2.9 is the connectivity of Bayesian methods and economic subfields. Papers in the area of marketing are closely related to flexible model structures (flexible functional forms), and particularly hierarchical Bayes, Dirichlet processes, panel data methods and heterogeneity. Given the increased amount of consumer data in the marketing field, more complex model structures which can handle heterogeneity across consumers are becoming important for this field.

Figure 2.9 also indicates a strong relation between the macroeconomics and finance literature and Bayesian methods. First, the topic of forecasting is central for macroeconomics and finance as this keyword occurs very frequently and is linked to both areas. Second, state space models, particle filters, Monte Carlo methods, Kalman filter, predictive likelihood analysis and Bayesian Model Averaging (BMA) are closely related to the macroeconomics and finance literature. These close relations indicate the need for sophisticated simulation techniques, such as particle filters, for the estimation and forecasting of complex models used for financial and macroeconomic data. Furthermore, the issue of (parameter) identification is central for macro models used for policy analysis, such as the VAR, Impulse Response Functions (IRF), and business cycle models. This relation is shown in the lower right corner of Figure 2.9.

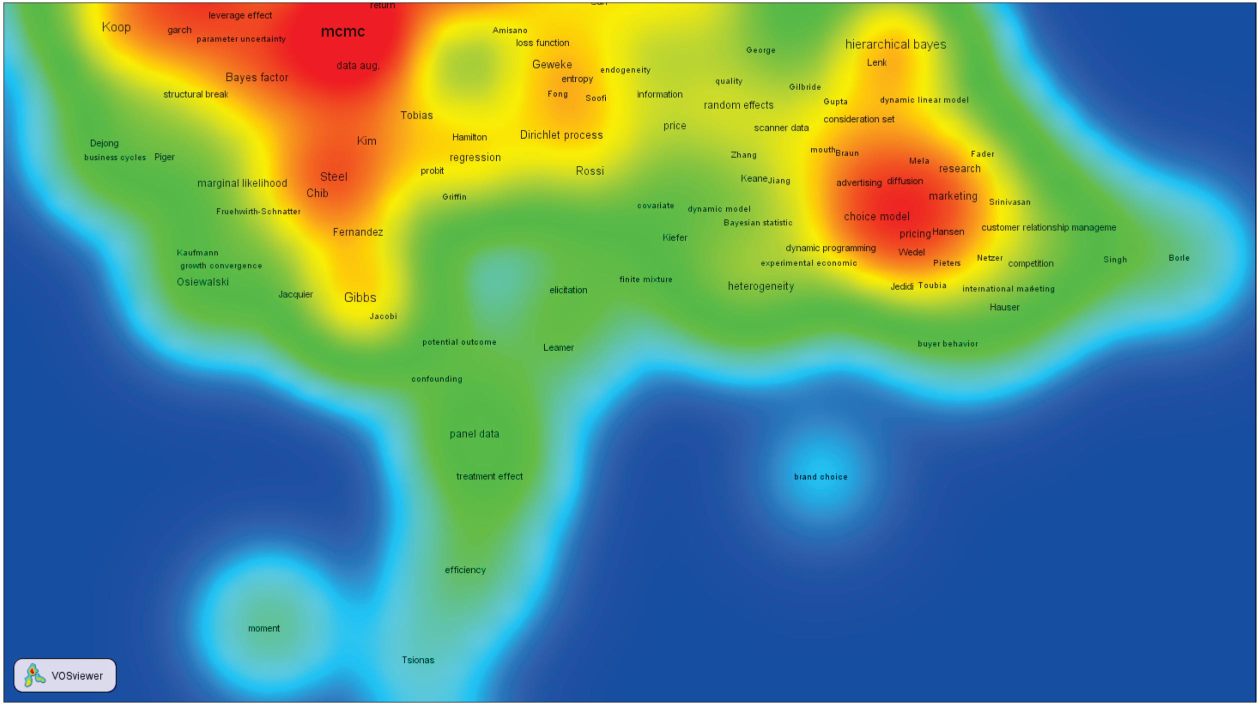
We finally note that computational advances have a large weight according to Figure 2.9. This topic is naturally linked to simulation methods, as speeding up computations is a central topic for the wide applicability of simulation methods. Computational advances are central especially for finite mixture models, and are in close relation to the areas of marketing and macro models.

We next select the influential papers in Bayesian econometrics based on the highly (more than 100 times) cited papers published in leading journals in section 2.3, and analyze the connectivity of the authors and the keywords of each paper. The connectivity of keywords and authors of these papers are shown in Figure 2.10 using a heatmap of the terms' density estimated by the concurrence of each keyword and author. We note that we leave some authors, such as Atkinson, Dorfman, Gelfand, Griffith and Trivedi, outside Figure 2.10 for visualization purposes. Despite a high number of papers by these authors, our clustering method separates these authors from the central part of the heatmap, most probably due to the diversity of the keywords in these authors' papers.

According to Figure 2.10, MCMC is central for Bayesian inference and Bayesian analysis. Macroeconomic and finance topics, such as stochastic volatility, time series, DSGE and option pricing, occur frequently in Bayesian econometrics. Marketing and choice mod-

Figure 2.10: Connectivity of topics and authors in papers in leading journals

Figure 2.10: Connectivity of topics and authors in papers in leading journals (continued)



els also occur frequently since the second dense area in the heatmap is centered around keywords such as pricing, choice model and advertising.

2.5 Topics, Issues and Debates

In this section we distill topics from section 2.4 where Bayesian econometrics has shown tremendous progress by itself and also compare it to the frequentist approach. We continue with a set of issues where Bayesian and Non-Bayesian econometricians disagree on and finally we list two debates that Bayesian econometricians hold amongst themselves.

2.5.1 Topics

In this subsection we summarize the advances in five topics which are shown to be central to Bayesian econometrics in section 2.4.

The Computational Revolution

Applicability of Bayesian methods in econometric analysis relies heavily on the feasibility of the estimation of models. For most econometric models of interest the posterior distribution is not of a known form like the normal one, and analyzing this distribution and its corresponding model probability using analytical methods or numerical integration methods is infeasible. Monte Carlo (MC) methods have been very useful for tackling these problems. One may characterize this as a 'Computational Revolution' for Bayesian inference leading to statements like 'Monte Carlo saved Bayes'. The popular Markov Chain Monte Carlo method, known as Gibbs sampling, contributed in particular to this. Therefore 'Gibbs saved Bayes' is a more appropriate statement.

There are at least three features of MC simulation techniques that make it attractive for Bayesian inference. First, there exists the traditional one of being able to directly simulate a nonlinear function of parameters of a model. Obvious examples are: Given a set of generated draws from parameters of a structural model one can directly evaluate the distribution of the multiplier to study the uncertainty of policy effectiveness; and given a set of generated draws from the parameters of a dynamic model one can obtain the distribution of the eigenvalues to study the stability of that system and the random walk nature of the process. These are examples of direct simulation of a model feature given appropriate parameter draws from a model. A second, more methodological feature is how to obtain these parameter draws from models where the posterior is not of a known form and it is not known how to generate draws directly from the posterior. Research in this topic focuses on indirect sampling methods using approximations to the posterior density labeled as importance densities or candidate densities. At first, Importance Sampling (IS), see Hammersley and Handscomb (1964), was introduced in Bayesian inference by Kloek and Van Dijk, (1975) and published in Kloek and Van Dijk, (1978), further developed by Van Dijk, and Kloek (1980, 1985) and given a complete detailed treatment in Geweke (1989). The construction of an approximately correct importance function in high dimensions is not trivial and given that the theory of Markov chain Monte Carlo (MCMC) was developed by Metropolis et al. (1953) and Hastings (1970), and extended in several influential papers such as Tierney (1994), this simulation method became the popular one. A major pioneering advance in this first computational revolution is Gibbs sampling developed in Geman and Geman (1984) and extended in Tanner and Wong (1987) and Gelfand and Smith (1990). See Robert and Casella (2004) for a recent and detailed discussion on the Gibbs sampling method and its extensions. The use of sampling methods turned out to be crucial for a third feature of Monte Carlo. Limited dependent variable models including Probit and Tobit models and unobserved component models, in particular State Space models using the Kalman filter, became popular due to their added flexibility in describing nonlinear data patterns. However, these models have an integral in the likelihood that refers to the underlying unobserved continuous data for the limited dependent variable models and unobserved state for the state space models. Bayesian simulation methods that were already used for integration in the parameter space can easily be extended and are the natural technical tools to also integrate these unobserved data and states.

These three features of Monte Carlo contributed greatly to the development of Bayesian econometrics, however, Monte Carlo became operational only with the improvements in the hardware of computing power, i.e. how fast a computer can perform an operation. The issue of computing power is central in econometric analysis in general, but it is even more central to Bayesian econometrics when the MC methods are applied. The improvements in computing power since the 1970s are clearly not negligible, but a more recent improvement has been observed with the introduction of clusters of computers, super computers and the possibility of performing operations in Graphical Processing Units (GPUs). These computing power improvements have been immediately adopted in the Bayesian econometrics literature. Using such computing power efficiently, however, mostly requires a careful engineering or modifications in the posterior sampler. Certain sampling methods, such as the importance sampler, are naturally suited for efficient use of computational power, see Cappé et al. (2008) for a discussion. A recent study specifically focusing on enabling Bayesian estimation using the GPU is Durham and Geweke (2013).

Flexible structures, unobserved components models and data augmentation in macroeconomics and finance

Unobserved components models constitute a field in econometrics where Bayesian inference is heavily used. As an example we focus on the state space models in time series analysis. The reason for the extensive use of Bayesian methods in this context is that simulation based Bayesian inference allows for much flexibility in the model structure as well as in the distributional assumptions. Flexible nonlinear structures can be modeled by introducing an extra latent space in such a way that the conditional structure of the model is linear given this unobserved state, see the local level model from Harvey (1990). Then from an estimation point of view, since the unobserved patterns underlying the behavior of observables need to be integrated out of the model, Bayesian integration methods can be used for inference and are very suitable for this class of models. That is, from the inference point of view, Bayesian inference takes the uncertainty of the unobserved patterns into account while estimating the model parameters. This is an important issue where the frequentist approach is more restrictive since the unobserved patterns are estimated conditional on the estimates of the model parameters (one takes the mode of the distribution rather than the whole distribution). Carlin et al. (1992) provide an exposition of the estimation methodology based on simulation to estimate the unobserved components and the model parameters jointly. Shortly after, Jacquier et al. (1994) show how an exact inference, unlike the quasi maximum likelihood approximation, can be obtained for the stochastic volatility models, a popular class of models in finance for modeling time varying volatility, using a similar approach. While the basic Bayesian inference principle remains unchanged, more efficient simulation algorithms are proposed in Carter and Kohn (1994), Frühwirth-Schnatter (1994), Carter and Kohn (1996), De Jong and Shephard (1995) and Koopman and Durbin (2000).

Although standard models using unobserved components allow for only continuous changes, models with discrete changes in parameters allowing for structural changes or discrete Markov processes are also feasible using Bayesian techniques. Gerlach et al. (2000) and Giordani and Kohn (2008), among others, provide efficient algorithms for obtaining Bayesian inference in case of such discrete changes in parameters. Interesting applications on regime analysis in economics are provided by Paap and Van Dijk, (1998) and Sims and Zha (2006).

When the observed variables to be modeled using unobserved components do not follow the standard normal distribution or dependence structures in the model are not linear, other estimation strategies, denoted as Particle Filter or Sequential Monte Carlo techniques that approximate the target distribution to be estimated well, can be conducted.

Bayesian inference backed up with advanced simulation algorithms have proved to be very useful in these circumstances, see for example Gordon et al. (1993), Pitt and Shephard (1999), Andrieu and Doucet (2002) and Andrieu et al. (2010). This type of inference is also the key ingredient of the volatility modeling in finance and micro founded macroeconomic models, among others, if the researcher does not resort to linear approximations to estimate the model. This makes it feasible to obtain exact online inference in these settings providing more accurate outcomes. Omori et al. (2007), Fernandez-Villaverde and Rubio-Ramirez (2007) and Fernandez-Villaverde and Rubio-Ramirez (2008) are some examples of this approach.

Choice models, robustness and policy effectiveness

Hierarchical models that refer to choice processes are a prominent research topic in recent decades due to the, often, unobserved heterogeneity of individual agents in commodity and labor markets. Flexibility in structure and distribution like the Dirichlet process are important features of the modeling process. Latent processes such as Probit models are used to describe unobserved components in models. Panel data are more and more used with scanner data giving rise to massive computing. Basic papers that deal with these issues are McCulloch and Tsay (1994); Rossi et al. (1996) and Hansen et al. (2006). More references are given in Geweke et al. (2011) chapter 8.

Econometric issues in this area are the presence of endogenous regressors, treatment effect problems, latent variables and many parameters. Shrinkage priors are regularly used in this class of models. Gibbs-based MCMC sampling are standard but new simulation based Bayesian techniques are developed using Dirichlet processes in order to gain robustness of results. It is expected that parallel processing will become important in this area of Bayesian research.

Instrumental variables

One of the prominent issues in econometric analysis is endogeneity arising from the correlation between a right hand side variable in an equation and the disturbance of that equation, the so-called direct feedback mechanism. A conventional solution to this problem is the use of instrumental variables (IV), equivalently IV regression models, see e.g. Sargan (1958), Goldberger (1972), and Bowden and Turkington (1990) for a detailed review.

Instrumental Variable (IV) regression model is a single equation Simultaneous Equations Model (SEM). SEMs, incorporating possibly complex feedback mechanisms between series, have been analyzed in the Cowles Commission monographs (Koopmans, 1950; Hood and Koopmans, 1953) and have been widely employed to analyze the behavior of markets, macroeconomic and other multivariate systems.

Bayesian analysis has been popular in these class of models due to the parameter identification issues plaguing inference in these models, in line with identification issues in SEMs in general. Bayesian analysis of the IV regression models are introduced by Lindley and El-Sayyad (1968), Drèze (1976) and Drèze (1977). Earlier literature suggests that the posterior densities in these class of models under flat priors may be improper, see e.g. Zellner, Bauwens and Van Dijk, (1988) and Bauwens and Van Dijk, (1990). An exception of this result is the case of over-identification where the posterior densities are shown to be proper under flat priors as shown in Zellner, Ando, Baştürk, Hoogerheide and Van Dijk, (2014).

Due to the identification issues in IV regression models, the use of alternative prior structures, such as the Jeffrey's prior, are proposed in Kleibergen and Van Dijk, (1998) and Hoogerheide, Kaashoek and Van Dijk, (2007). More recent advances in the Bayesian estimation of these models are the introduction of semiparametric models by Conley et al. (2008) and Florens and Simoni (2012) among others, and efficient posterior sampling algorithms as in Zellner, Ando, Baştürk, Hoogerheide and Van Dijk, (2014). For a detailed discussion of Bayesian approaches to IV and many examples, we refer to Lancaster (2004) and Rossi et al. (2005).

Dynamic models and forecasting

Bayesian analysis has become a dominant forecasting and counterfactual analysis tool in recent decades. There are four main reasons for this phenomenon. First, many of the complex, otherwise non-estimable, models can be estimated using simulation based Bayesian methodology. Perhaps, the most important example of these models includes the class of structural micro founded macroeconomic models, such as Dynamic Stochastic General Equilibrium models, that are used both for policy analysis and for forecasting, see for example Smets and Wouters (2003), Smets and Wouters (2007), An and Schorfheide (2007). Currently, many of the central banks employ such models to obtain short and long term projections of the economy. An advantage of the Bayesian methodology is that it provides a solid statistical ground for efficient analysis using these structural models. As Bayesian inference provides the distribution of many key parameters that play a crucial role in economic analysis it is often used as a tool for counterfactual analysis. For instance, the questions such as 'if quantitative easing were not conducted in US, would the course of the recession differ?' could be answered by estimating relevant structural models. Bayesian analysis provides a statistically coherent tool for employing counterfactual analysis by forecasting under counterfactuals.

Second, prior distributions can play an integral part of the forecasting especially for the overparametrized models. Vector Auto Regression models (VAR) are major examples where Bayesian inference facilitates forecasting using the prior distributions for shrinking the parameters towards zero and thereby decreasing the dimensionality of the models. Decreasing the dimension of the overparametrized models using clever prior distributions has proved to be very useful in many applications. Prominent examples of this approach constitute Doan et al. (1984), Kadiyala and Karlsson (1997), Banbura et al. (2010).

Third, Bayesian methodology takes the parameter uncertainty into account which may be of crucial importance in many applications. This enables researchers to obtain the entire predictive distribution rather than point forecasts based on the mode of the parameter distribution as in the frequentist approach. An important advantage of this feature is that different parts of the predictive distribution can be analyzed easily. This yields an obvious advantage for the analysis of various types of risk in finance and macroeconomics. A recent example is given in Baştürk, Çakmaklı, Ceyhan and Van Dijk, (2013b) where the probability of deflation is evaluated for the US.

Fourth, the Bayesian methodology provides a natural and statistically solid way to take model uncertainty into account and to combine models to increase the predictive ability of many competing models. Bayesian model averaging technique provides one elegant way to do so, see for example Min and Zellner (1993), Fernandez et al. (2001), Avramov (2002), Cremers (2002) and Wright (2008). Recent advances in Bayesian model combination also allows to combine models where the model space is not complete implying that none of the competing models might be the true model. In that case, optimal combinations are proposed in Geweke and Amisano (2010) and Geweke and Amisano (2011). For a recent example where Sequential Monte Carlo is used to obtain density combinations from different model structures we refer to Billio, Casarin, Ravazzolo and Van Dijk, (2013).

2.5.2 Issues

We next summarize three issues in Bayesian econometrics that non-Bayesian repeatedly referred to in the literature, and we discuss the advances in Bayesian econometrics regarding these issues.

Identification, the value of prior information and model evaluation

The choice of a reasonable informative prior distribution is a crucial part of the Bayesian analysis and often subject to criticism by frequentist econometricians. However, sensible prior distributions provides valuable improvements in inference for many of the econometric issues that are hard to tackle without it. Hence, it is a bless rather than a curse. Prior distributions play a key role in many aspects such as identification and incorporation of prior knowledge or evidence from other studies, dimension reduction and forecasting, and nonparametric Bayesian inference.

In many cases, data are not sufficiently informative about the appropriate parameter values and may yield similar likelihood values with different parameter combinations, which is referred to as 'the weak identification problem'. Weak identification occurs in models with nearly reduced rank, which occurs in simultaneous equations models, instrumental variable regression models, dynamic models with cointegration and in factor models. Weak identification gives usually irregular behavior of the likelihood; see papers by Bauwens and Van Dijk, (1990), Kleibergen and Van Dijk, (1994), Kleibergen and Van Dijk, (1998) and Hoogerheide, Kaashoek and Van Dijk, (2007). In such cases, assigning reasonable priors from other studies or other evidence alleviate the identification problem since Bayesian inference incorporates prior knowledge with the data information using the Bayes rule. This is often used in micro founded macroeconomic studies where priors are constructed using economic theory or from other studies or from micro data such as households surveys, see for example Del Negro and Schorfheide (2008). Another well known example of prior information is the use of reasonable regions in the parameter space, restricted by inequality conditions. Frequentist inference is extremely difficult using such restrictions. Examples of Bayesian inference where the implied prior on the range of an economics multiplier or a prior on the length of the period of oscillation of the business cycle yield plausible restrictions are given in Van Dijk, and Kloek (1980) and Harvey, Trimbur and Van Dijk, (2007).

One important issue is the existence of the posterior distribution and its moments in case of weak or noninformative priors and an almost flat likelihood. Many Bayesians argue that the posterior distribution does not exist in such a case. However, Zellner, Ando, Baştürk, Hoogerheide and Van Dijk, (2014) shows that this may not be correct and give very weak conditions for existence of moments of zero and higher order for the famous weak instrument problem. Priors play an instrumental role in forecasting especially when the models are overparametrized. Often overparametrized models are condemned with poor forecasting performance as any additional parameters exploits the data information further and yielding prediction uncertainty. To tackle this problem sensible prior distributions are assigned on parameters shrinking those parameters that are not informative towards zero and thereby decreasing the dimension of the problem. In macroeconomic forecasting the priors proposed by Doan et al. (1984) have become a standard tool among econometricians in academia and in other institutions such as central banks. In more general cases, many tailored priors are used for shrinkage of the model parameters towards zero and therefore they are efficiently used in variable selection when there are many candidate variables to select from. Prominent examples include, George and McCulloch (1993), Ishwaran and Rao (2005) and Park and Casella (2008) among others.

Prior distributions are also the key ingredients for flexible modeling strategies in Bayesian analysis. This is especially of key importance for density estimation. Bayesian nonparametric analysis is one evolving area where such prior distributions or processes are heavily used. While some of the theoretical achievements were already accomplished during 1970s, see Ferguson (1973), Antoniak (1974) for example and Sethuraman (1994) for a more recent paper, extensive use of such priors were only possible with the advance of computing power. Indeed with increasing computing power simulation schemes used for Bayesian nonparametric inference proved to be very useful for such complex analysis, see for example Escobar and West (1995), Neal (2000) and Walker (2007). Currently, many applications have emerged in different fields using such flexible prior distributions, see for example Chib and Hamilton (2002); Hirano (2002); Griffin and Steel (2004); Jensen (2004); Jensen and Maheu (2010).

Lindley's paradox - or Bartlett's or Jeffreys' paradox; see Lindley (1957) and Bartlett (1957) - implies that one has to choose very carefully the amount of prior information compared to the amount of sample information, when comparing alternative hypotheses on model structures with the intention to let the information from the data in the likelihood dominate that of the prior. Typically a naive or malevolent researcher could 'force' the posterior probability of a certain model M, the 'restricted model' in case of two nested models, to tend to go to unity by letting the priors in all alternative models tend to diffuse priors, thereby decreasing the marginal likelihoods of all alternative models, even if the particular model M does not make sense and poorly describes the data. In an attempt to make the posterior model probabilities 'fair', one could use predictive likelihoods instead of marginal likelihoods; see Gelfand and Dey (1994), O'Hagan (1995), and Berger and Pericchi (1996). However, the use of predictive likelihoods brings several questions and issues. First, one must choose the training sample and the hold-out sample. Examples of important questions are: How many observations are included in these samples? Is one training sample used or does one average over multiple (or all possible) training samples? In the latter case, what does one average: e.g., marginal likelihoods, logarithms of marginal likelihoods, Savage-Dickey Density Ratios or posterior model probabilities? Second, if one chooses to average results over multiple (or all possible) training samples, then the computing time that is required for obtaining all Monte Carlo simulation results for all training samples may be huge. In other words, the Lindley paradox and the computation of predictive likelihoods enlarge the relevance of simulation methods that efficiently provide reliable and accurate results in case of non-elliptical credible sets. A suitable method must deal with large numbers of different non-elliptical shapes in a feasible computing time. For time series models computing the marginal likelihood for a random subsample implies that the estimation must be performed for an irregularly observed time series (with many 'missing values'), which is typically only feasible using an appropriate formulation and estimation of a state space model. In future research computationally efficient and accurate simulation methods need to be developed here.[6]

Dynamic inference and nonstationarity

Dynamic inference and methods to handle nonstationary data constitute a second Bayesian topic subject to criticism by frequentist econometricians. As a start we summarize the perfect duality between Bayesian inference in the parameter space and frequentist inference in the sample space for the well-known class of the linear regression model $y = X\beta + \epsilon$. In both frequentist and Bayesian econometrics, the parameter β has a student-t density. However, the interpretations are different. A graphical illustration of the difference for this model is provided in the left panel in Figure 2.11. Table 2.1 presents a summary of the duality and differences between Bayesian and frequentist inference. In practice one finds that many empirical researchers are 'closet' Bayesians in interpreting the obtained value of a t-test as indicating possible positive strength of the empirical result. In the strict frequentist sense one can only reject a null hypothesis if there is sufficient evidence for it.

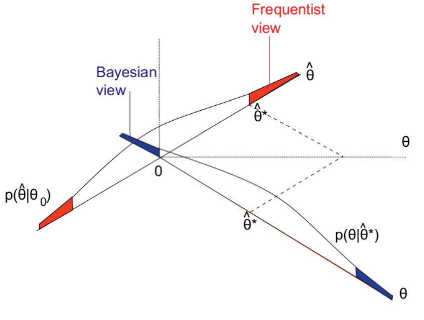
This equivalence breaks for dynamic regression models. In inference about stationary dynamic models, while Bayesian econometrics suggests student-t density for the parameters, the frequentist econometrics has finite sample bias problems. This divergence between the Bayesian and the frequentist inference is also observed for nonstationary dynamic models, see Sims and Uhlig (1991). In the frequentist case, the inferential statement: 'No falsification of the Null Hypothesis of a Unit Root leads to the acceptance

[6]Part of this paragraph is taken from Van Dijk, (2013).

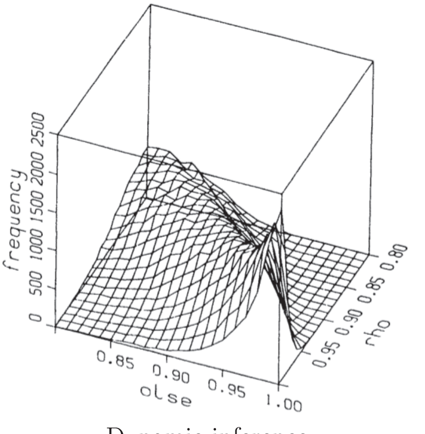
Table 2.1: Duality and differences between Bayesian and frequentist inference

Classical inference	**Bayesian inference**
Parameters β are fixed unknown constants	Parameters β are stochastic variables. One defines a prior distribution on parameter space.
Data y are used to estimate β and check validity of postulated model, by comparing data with (infinitely large, hypothetical) data set from model.	Data y are used as evidence to update state of the mind: data transform prior into posterior distribution using the likelihood.
Objective concept of probability: Probability is the fraction of occurrences when the process is repeated infinitely often.	Subjective concept of probability: Probability is a degree of belief that an event occurs.
One can use the maximum likelihood estimator as an estimator of β.	One uses Bayes' theorem to obtain the posterior distribution of β. One can use $E(\beta\|y)$ or minimize a loss function to estimate θ.
R^2 is used to compare models.	Model comparison is carried out by using posterior odds ratio.

Figure 2.11: Frequentist versus Bayesian econometrics



Static inference



Dynamic inference
(Sims and Uhlig, 1991)

of the Unit Root' yields fragile and often incorrect conclusions, for example, if a break occurs in the series. That is, several alternatives may be more plausible.

Sims and Uhlig (1991) and Schotman and Van Dijk, (1991b), Schotman and Van Dijk, (1991a) suggest that Bayesian inference for models with unit root is more sensible, as well as much easier to handle analytically, than the classical confidence statements. Even under the assumptions of linearity and Gaussian disturbances, and even if conditioning on initial conditions is maintained, classical small-sample distribution theory for autoregressions is complicated. Classical asymptotic theory breaks discontinuously at the boundary of the stationary region. Therefore, the usual simple normal asymptotic approximations are not available. The likelihood function, however, is well known to be the same in autoregressions and non-dynamic regressions, assuming independence of disturbances from lagged dependent variables. Thus inference satisfying the likelihood principle has the same character in autoregressions whether or not the data may be non-stationary. The illustration of the likelihood for this autoregressive model in Sims and Uhlig (1991) is given in the right panel of Figure 2.11. Phillips (1991) stresses the fragility of Bayesian inference to the specification of the prior and warns against the mechanical use of a flat prior. Schotman and Van Dijk, (1991b) present a solution to this problem in a different parameterizations of this model. Schotman and Van Dijk, (1991a) suggest to use posterior odds test (for the choice between a unit root model and an AR(1) stationary model). In the Bayesian approach, a unit root is not a testing problem but a choice problem on the relative weights of two states of nature: the stationary and the nonstationary case, see Sims and Uhlig (1991). One can use these weights in evaluating forecasts and impulse response functions. De Pooter, Ravazzolo, Segers and Van Dijk, (2009) suggest to use the Schotman-van-Dijk (SVD) prior in this context.

The difference between Bayesian and frequentist econometrics is particularly apparent for the case of multivariate dynamic models with possible nonstationarity. Kleibergen and Van Dijk, (1993) propose a Bayesian procedure to inference and show that by using flat priors, the marginal posteriors of the cointegration vectors are ill-behaved when certain parameters become non-identified. This problem also plagues standard frequentist inference. Kleibergen and Van Dijk, (1993) solve this problem by proposing the Jeffrey's prior. The key problem with the frequentist approach is that a sequential testing procedure is used to determine the number of stationary and the number of nonstationary relations, in other words how many stable and unstable relations exist. In the Bayesian approach weights can be evaluated for each member of the set of stable and unstable relations. Forecasts can be made and impulse responses evaluated with a weighted average of such relations using marginal and predictive likelihoods, see Wright (2008) and Strachan and Van Dijk, (2013) for details.

Vector autoregressive versus structural modeling

Given that in the early 1970s oil price shocks and high inflation affected macroeconomic time series both in levels and volatility, the existing classes of econometric models often based on Keynesian structures did not fit and forecast well. In the words of Christopher Sims these models were not ‘realistic’. In his 1980 paper in Econometrica, Sims advocated the use of Vector Autoregressive Models (VAR) to describe better the time series patterns in the data. One may characterize his work as: Sims against the ‘Econometric Establishment’. However pragmatic the VAR approach was, there quickly were discussions on the fact that the unrestricted VAR had the curse of parameter dimensionality or otherwise stated an over-parametrization danger. Several approaches to overcome this criticism were developed of which we discuss the following ones. One approach is to make use of shrinkage priors that were of great help in forecasting. This class of priors became known as the Minnesota prior from Doan et al. (1984). A useful alternative is the dummy based observation prior due to amongst others Kadiyala and Karlsson (1997)). In the late 1990s structural economic priors Del Negro and Schorfheide (2004) came into existence parallel to the use of more structural VAR models like the DSGE model from (Smets and Wouters, 2007) and many other Structural VAR’s. This latter topic has been discussed before. Nowadays structural VARs with informative priors are used everywhere in macroeconomics in academia and professional organizations both for forecasting and policy analysis. Given the recent economic crisis it is clear that this class of models needs to be developed further to include financial sectors.

2.5.3 Debates

This section summarizes a major debate between Bayesian econometricians, namely the choice of objective versus subjective econometrics. We then summarize the attempts to improve communication between statistics and econometrics.

Objective versus subjective econometrics

In general, probabilities are not physical quantities like weight, distance, or heat that one can measure using physical instruments. As De Finetti (1989) suggested, ‘Probabilities are a state of the mind’. Therefore, in general Bayesians are subjectivists and probabilities are personal statements. However, some are more subjectivist/personal than others.

Table 2.2: Objective versus subjective econometrics

Objective	**Subjective**
Let the model speak: analyze the shape of the likelihood	Everything is personal
Scientific evidence dominates	Personal probabilities should be solicited
Experts opinion may fail	Experts opinions matter
Reach a large public	

The alternative, objectivist, viewpoint is based on the idea that experts opinions may fail, and aims to 'Let the model speak' in reporting scientific evidence. This viewpoint, due to limited inclusion of personal or expert statements, reaches a large public. The subjective viewpoint, on the other hand, argues that even when experts opinions fail, the likelihood will show that this prior is not relevant. A brief summary of the differences between the subjectivist and objectivist viewpoints are provided in table 2.2.

In conclusion, there exist 'true' Bayesian econometricians who belong in the right hand column of table 2.2; Instrumental and pragmatic Bayesian econometricians who belong more in the left column of table 2.2; Pragmatic Bayesian econometricians and 'Closet' Bayesian econometricians those that use regression outcomes and talk about 'strongly significant' t-values and 'accept' the null hypothesis. They all apply Bayesian techniques nowadays!

Communication between Statistics and Econometrics

Statistics and Econometrics has had a difficult relationship, with several switches in the past 50 years, due to the fact that econometric models are high-dimensional while statisticians prefer a maximum of 3 dimensions. Early statistics was applied to economic time series while recent statistics is applied more to biology and is becoming very computational. Econometrics is more model-oriented with a large number of parameters.

There have been attempts to construct bridges between statistics and econometrics. Among these are the Seminar on Bayesian Inference in Econometrics and Statistics (SBIES) that was pioneered by Arnold Zellner from 1970 onwards and now actively steered by Siddharta Chib[7], the European Seminar in Bayesian Econometrics (ESOBE) that started in 2010 by Herman K. van Dijk[8] and the Economics, Finance and Business

[7] `apps.olin.wustl.edu/conf/sbies/Home/`
[8] `www.esobe.org/`

Section (EFAB) of the International Society of Bayesian Analysis (ISBA) that was started in 2013 by Mike West.[9]

2.6 Conclusion

Inspired by the path that Bayesian econometrics has followed for the last half century, this section presents the authors' personal expectations of the important topics/subjects for the future of Bayesian econometrics in the 21^{st} Century. One such prediction is that 'the second computational simulation revolution' where efficient information distillation from 'big data' with sophisticated Bayesian methodology using parallelization techniques is going to play an important role. Another topic that is predicted to gain popularity is complex economic structures with nonlinearities and complete predictive densities. A third topic that is expected to have importance in the future is the analysis of implied model features, such as risk or instability due to diverging ratios, and decision analysis. Finally, model incompleteness, which refrains from the assumption that the true data generating process is in the defined model set, is predicted to be important topics in Bayesian econometrics, see Geweke (2010).

Besides focusing on important topics in Bayesian econometrics, we further predict that the influence of Bayesian econometrics in the econometrics field will continue to increase over time. This final prediction is in line with the statement 'Econometrics should always and everywhere be Bayesian' in Sims (2007). We refer to Sims (2007) for a detailed discussion on this topic and on how Bayesian approaches might become more prevalent in several areas of economic applications.

We end this chapter with a bit of a game. The citation numbers we analyze for Bayesian papers can be related to the h-index of authors, a conventional measure for the impact of published work by scholars, see Hirsch (2005). Nowadays, the h-index is sometimes used in the career path and promotion stages of young researchers. We employ a simple simulation study to assess the expected h-index of a 'random Bayesian' publishing a predefined number of papers in the leading journals we consider. In order to assess this expected h-index, we consider a random sample of size J from the 999 papers in our database and calculate the h-index. The average h-index for 1000 such random samples is used to approximate the expected h-index for an author with J publications in leading journals. For a young Bayesian econometrician who is the author of 5 such publications, we find that the h-index is approximately 4, i.e. very high compared to the total number of publications of this author, and the expected number of citations for

[9] `bayesian.org/sections/EFaB/bylaws`

this author's papers is 334. For an author, coming up for tenure, with 12 publications in leading journals the expected h-index is approximately 9 with an expected number of citations of 765. For a very established author with 50 publications in these journals, the expected h-index is approximately 25, with an expected number of citations of 1644. The papers and the journals considered therefore have a considerable impact in the field, according to the calculated h-indexes. Conditional upon our data set we conclude that young Bayesian econometricians have a very good chance to follow an academic career successfully.

Appendices

2.A Bayesian Papers in Leading Journals

Table A.1: Percentages of pages devoted to Bayesian papers in leading journals

Years	RES	ReStat	Ectra	IER	ET	MS	ER	JE	JAE	JBES
1978	0.00	0.00	2.92	0.00	–	–	–	6.27	–	–
1979	0.00	0.00	0.25	1.77	–	–	–	4.76	–	–
1980	0.00	0.00	0.00	5.12	–	–	–	7.17	–	–
1981	1.32	0.00	0.00	1.87	–	–	–	5.70	–	–
1982	0.00	0.97	0.75	1.60	–	0.00	12.46	10.69	–	–
1983	0.00	0.00	0.00	0.00	–	10.00	1.50	12.31	–	11.14
1984	0.00	1.81	1.40	0.00	–	0.00	44.93	10.97	–	4.32
1985	0.00	1.52	1.11	3.17	0.00	5.61	29.86	15.64	–	6.95
1986	5.54	0.00	0.00	0.00	0.00	0.00	0.00	4.89	4.08	15.80
1987	0.00	1.47	0.00	0.00	0.65	0.00	0.00	0.00	5.40	15.47
1988	0.00	0.00	0.00	2.38	0.00	0.00	0.00	13.57	0.00	5.52
1989	0.00	2.83	2.73	1.58	0.00	0.00	0.00	5.21	3.84	3.17
1990	0.00	1.10	0.00	0.00	0.00	4.66	0.00	3.66	0.00	5.39
1991	1.24	0.00	0.98	2.41	0.00	0.00	0.00	24.91	29.92	27.73
1992	0.00	0.00	1.63	0.00	2.92	0.00	3.49	7.95	5.47	18.62
1993	0.00	0.00	1.83	0.00	0.00	0.00	0.00	20.65	11.79	7.82
1994	0.00	5.07	1.76	0.00	33.13	0.00	10.75	9.47	3.10	24.60
1995	0.00	5.83	0.00	2.36	0.00	1.57	0.00	14.88	0.00	4.87
1996	0.00	0.65	3.35	0.00	8.87	18.53	0.00	13.34	5.34	8.60
1997	0.00	0.00	0.00	1.41	0.00	5.88	0.00	10.91	15.75	9.07
1998	3.79	8.01	0.00	3.32	8.74	3.76	5.48	13.04	10.09	11.37
1999	0.00	2.98	2.90	0.00	0.00	14.38	32.67	5.98	10.45	12.82
2000	0.00	1.41	0.00	0.00	2.11	12.56	0.00	18.05	16.40	26.41
2001	0.00	2.69	4.19	6.88	0.69	0.00	8.04	4.40	7.17	13.59
2002	1.96	0.00	0.75	0.00	4.25	19.71	4.03	9.19	8.81	4.20
2003	0.00	0.00	3.70	4.41	0.00	11.76	2.92	15.65	10.49	20.93
2004	0.00	0.00	5.94	4.61	0.00	12.68	0.00	16.57	6.73	21.22
2005	5.15	1.65	0.00	0.00	2.72	7.25	0.00	11.92	16.23	11.67
2006	2.47	0.00	1.56	0.00	0.00	15.03	31.43	6.14	12.37	11.44
2007	4.14	0.00	4.35	2.27	2.78	16.74	56.83	12.45	1.57	14.40
2008	0.00	0.00	0.00	0.00	3.07	23.15	9.87	10.62	22.73	20.32
2009	0.00	0.00	4.98	8.54	1.06	0.00	4.13	7.28	24.65	10.18
2010	5.85	0.00	0.90	0.00	0.86	22.16	9.44	7.06	24.22	21.97
2011	1.91	1.30	0.00	5.35	1.75	35.17	14.29	6.75	13.57	4.82
2012	0.00	4.49	2.10	4.82	0.00	13.65	6.70	13.44	18.32	13.42
2013	3.51	3.96	1.74	2.18	0.00	12.56	0.00	5.43	21.56	4.97
2014	0.00	0.00	0.00	0.00	0.00	28.34	40.65	5.78	22.61	6.62
1978-2014	0.96	1.18	1.40	1.70	2.45	8.94	9.98	9.91	11.18	12.48

The table presents the percentages of pages devoted to Bayesian papers in the journals for each year and average percentages for the period 1978-2014. The table is an extension of Table 2 in Poirier (1992). The numbers in red correspond to years with special issues. *Econometric Reviews*, *Econometric Theory*, *Journal of Applied Econometrics* and *Marketing Science* did not exist before 1982, 1985, 1986 and 1982, respectively. Average numbers of Bayesian pages only include years for which the journal existed. Journal abbreviations are as in Figure 2.2.

Table A.2: Citation information for papers in leading journals

	Total	RES	ReStat	IER	Ectra	ET	MS	ER	JAE	JE	JBES
All papers											
Number of citations	67891	3143	3104	2351	6000	1799	7590	3592	8103	21432	10777
Number of cited papers	969	16	28	29	34	30	121	61	125	329	178
Number of Bayesian papers	1020	16	29	29	34	31	129	82	130	336	206
Papers with at least 100 citations											
Number of citations	46895	2748	2392	1581	5147	1164	4922	2711	5340	13361	7293
Number of Bayesian papers	170	6	9	6	13	7	23	4	20	55	26
Papers with at least 400 citations											
Number of citations	20460	2086	808	408	3442	425	944	2603	1806	4376	3562
Number of Bayesian papers	31	2	1	1	5	1	2	3	3	8	5

The table presents citation information for each leading journal and for the leading journals jointly. Total number of cited papers is based on papers which are cited at least once. Journal abbreviations are as in Figure 2.2.

Table A.3: Average number of citations for papers in leading journals for 5-year intervals

Years	RES	ReStat	Ectra	IER	ET	MS	ER	JAE	JBES	JE
All papers										
1978-1982	2	0	134	10	–	0	6	–	279	–
1983-1987	12	50	22	1	0	108	214	163	181	360
1988-1992	0	91	330	28	0	23	1	145	349	117
1993-1997	0	113	122	18	252	235	7	199	956	727
1998-2002	315	286	274	134	78	440	166	412	1011	411
2003-2007	229	26	205	137	12	406	292	258	1071	342
2008-2014	51	39	80	102	12	219	26	386	315	141
1978-2014	85	84	162	64	60	230	109	279	579	337
Papers with at least 100 citations										
1978-1982	0	0	124	0	–	0	0	–	116	–
1983-1987	0	35	0	0	0	94	204	121	79	291
1988-1992	0	75	324	0	0	22	0	88	197	54
1993-1997	0	87	116	0	182	207	0	163	660	618
1998-2002	315	255	259	108	51	381	142	322	746	236
2003-2007	212	26	132	129	0	235	197	166	698	217
2008-2014	16	0	53	57	0	33	0	200	126	30
1978-2014	74	65	139	43	39	149	82	184	361	228

The table presents the average citation numbers of the Bayesian papers in the journals for the 5 year periods for all papers in leading journals (top panel) and a subset of papers with at least 100 citations (bottom panel). The table is an extension of Table 2 in Poirier (1992). Note that the period of observation is different for the following journals: Econometric Theory did not exist before 1985. Therefore the mean for the period 1983-1987 is taken over the periods 1985-1987. Journal of Applied Econometrics did not exist before 1986. Therefore, the mean over the years 1983-1987 is equal to the mean for 1986-1987. Econometric Reviews did not exist before 1982. Therefore, the mean for the period 1978-1982 is equal to the value in 1982. Marketing Science did not exist before 1982. So, the mean for the period 1978-1982 is equal to the value in 1982. Journal abbreviations are as in Figure 2.2.

Table A.4: Papers in leading journals with at least 100 citations

Paper	Cite'd	Paper	Cite'd	Paper	Cite'd	Paper	Cite'd	Paper	Cite'd
Kloek and Van Dijk (1978)	495	Goodwin (1993)	163	Filardo and Gordon (1998)	196	Eraker (2001)	394	Rust and Verhoef (2005)	125
Akaike (1981)	580	Albert and Chib (1993)	446	Phillips (1998)	241	Tsionas (2002)	134	Ainslie et al. (2005)	146
Leamer (1982)	126	Uhlig (1994)	105	Chib (1998)	370	Engle (2002)	382	Primiceri (2005)	627
Monahan (1983)	107	Elliott and Stock (1994)	118	Allenby and Rossi (1998)	508	Chib et al. (2002)	416	Yu and Meyer (2006)	108
Gersch and Kitagawa (1983)	128	Kleibergen and Van Dijk (1994)	120	So et al. (1998)	174	Moe and Fader (2002)	109	Chib et al. (2006)	198
Hauser and Shugan (1983)	469	Phillips and Ploberger (1994)	141	Kim and Nelson (1998)	344	Hofstede et al. (2002)	121	Hansen et al. (2006)	100
Doan et al. (1984)	1020	Van den Broeck et al. (1994)	235	Ainslie and Rossi (1998)	227	Kim et al. (2002)	204	Pesaran et al. (2006)	170
Zellner and Rossi (1984)	179	Chib and Greenberg (1994)	333	Otrok and Whiteman (1998)	130	Talukdar et al. (2002)	231	Das et al. (2007)	540
Connolly and Hirschey (1984)	173	McCulloch and Rossi (1994)	466	Sims and Zha (1998)	408	Lancaster (2002)	392	An and Schorfheide (2007)	875
Geweke (1986)	242	Andrews and Chen (1994)	380	Kim et al. (1998)	1459	Geweke et al. (2003)	120	Omori et al. (2007)	168
Hamilton (1986)	110	Jacquier et al. (1994)	1347	Sims and Zha (1999)	661	Chernozhukov and Hong (2003)	284	Amisano and Giacomini (2007)	193
Ohta and Griliches (1986)	105	Geweke et al. (1994)	332	Geweke (1999)	708	Jones (2003)	355	Del Negro et al. (2007)	450
Clemen and Winkler (1986)	230	Chib and Greenberg (1995)	170	Zha (1999)	112	Hamilton (2003)	964	Heerde et al. (2007)	110
Litterman (1986)	859	Chern et al. (1995)	104	Koop and Potter (1999a)	114	Jedidi et al. (2003)	108	Fernández-Villaverde and Rubio-Ramirez (2007)	263
Garcia-Ferrer et al. (1987)	135	Phillips (1996)	187	Koop and Potter (1999b)	137	Canova and Ciccarelli (2004)	100	Masanjala and Papageorgiou (2008)	113
Geweke (1988)	185	Chib and Greenberg (1996)	425	Waggoner and Zha (1999)	125	Fernández-Villaverde and Rubio-Ramirez (2004)	213	Diebold et al. (2008)	115
Geweke (1989a)	1303	Terrell (1996)	139	Kim and Nelson (1999)	808	Hansen et al. (2004)	251	Wright (2008)	123
Zellner and Hong (1989)	120	Pfann et al. (1996)	113	Neelamegham and Chintagunta (1999)	237	Jacquier et al. (2004)	339	Mol et al. (2008)	179
Geweke (1989b)	144	Geweke (1996)	126	Manchanda et al. (1999)	281	Berg et al. (2004)	147	Lee and Card (2008)	355
Karp and Perloff (1989)	111	Chib (1996)	368	Rust et al. (1999)	309	Kim et al. (2004)	181	Chauvet and Piger (2008)	110
Koray and Lastrapes (1989)	265	Smith and Kohn (1996)	494	Martin (2000)	133	Gilbride and Allenby (2004)	238	Manchanda et al. (2008)	115
Lenk and Rao (1990)	109	DeJong et al. (1996)	106	Schorfheide (2000)	481	Montgomery et al. (2004)	240	Netzer et al. (2008)	116
Sims and Uhlig (1991)	317	Lenk et al. (1996)	352	Nakatsuma (2000)	115	Canova (2004)	274	Imai et al. (2009)	110
Schotman and Van Dijk (1991b)	158	Rossi et al. (1996)	475	Geweke and Keane (2000)	126	Negro and Schorfheide (2004)	373	Ley and Steel (2009)	154
Phillips (1991)	280	Kadiyala and Karlsson (1997)	300	DeJong et al. (2000)	189	Fernández-Villaverde and Rubio-Ramirez (2005)	137	Mountford and Uhlig (2009)	845
Zellner et al. (1991)	131	Fernández et al. (1997)	121	McCulloch et al. (2000)	212	Peersman (2005)	189	Faust and Wright (2009)	101
Schotman and Van Dijk (1991a)	158	Koop et al. (1997)	196	Lancaster (2000)	392	Canova (2005)	213	Canova and Ciccarelli (2009)	119
Hamilton (1991)	159	Maddala et al. (1997)	189	Aguilar and West (2000)	333	Smets and Wouters (2005)	292	Justiniano and Preston (2010)	100
Chib (1992)	246	Faust and Leeper (1997)	460	Wedel and Pieters (2000)	185	Dehejia (2005)	119	Nbura et al. (2010)	190
Shea (1992)	110	Montgomery (1997)	208	Hamilton (2001)	190	O'Donnell and Coelli (2005)	122	Leeper et al. (2010)	109
Andrews (1993)	395	Kleibergen and Van Dijk (1998)	118	Elerian et al. (2001)	443	Imai and van Dyk (2005)	133	Rubio-Ramirez et al. (2010)	114
Geweke (1993)	378	Poirier (1998)	137	Fernandez et al. (2001b)	480	Yu (2005)	245	Schmitt-Grohe and Uribe (2012)	260
Min and Zellner (1993)	232	Chib et al. (1998)	102	Fernandez et al. (2001a)	501	Owyang et al. (2005)	130	Kose et al. (2012)	277
Chib (1993)	447	Chiang et al. (1998)	134	Chib and Winkelmann (2001)	143	Borle et al. (2005)	107		

The table gives the number of citations for papers which are published in the ten journals we consider and are cited at least 100 times. The citation numbers are collected using Google Scholar on 13–14 April, 2014.

2.B Subject Connectivity

Network, connectivity and heat maps in this chapter have been produced using the computer program VOSviewer, which uses the VOS (Visualisation Of Similarities) mapping technique that combines the approaches to mapping and clustering of bibliometric networks (Waltman et al. (2010),Van Eck and Waltman (2010)). These techniques aim to shed light on the interpretation of a network. Suppose there are n nodes in a network. The VOS mapping technique is based on the minimization of the below function with respect to $x_1, ..., x_n$;

$$V(x_1, ..., x_n) = \sum_{i<j} s_{ij} d_{ij}^2 - \sum_{i<j} d_{ij} \tag{2.5}$$

where s_{ij} is the association strength of nodes i and j $s_{ij} = \frac{2mc_{ij}}{c_i c_j}$, c_{ij} is the number of links between nodes i and j with $c_{ij} = c_{ji} \geq 0$, c_i is the total number of links of node i $c_i = \sum_{i \neq j} c_{ij}$, m is the total number of links in the network $m = \frac{1}{2} \sum_i c_i$. Moreover, d_{ij} is the distance between nodes i and j $d_{ij} = \|x_i - x_j\| = \sqrt{\sum_{k=1}^{p} (x_{ik} - x_{jk})^2}$ in the case of mapping, and in the case of clustering

$$d_{ij} = \begin{cases} 0 & if \quad x_i = x_j \\ \frac{1}{\gamma} & if \quad x_i \neq x_j \end{cases}$$

where $\gamma > 0$ is the resolution parameter. As the resolution parameter is higher, the number of clusters obtained becomes higher.

The network and connectivity map in Figure 2.8 has been created using the cluster density view of the map, the network and connectivity map in Figure 2.9 has been created using the label view of the map and the heat map in Figure 2.10 has been created using the density view of the map.

In the label view, the font size of the labels of the items and of the circles around them depend on the weight of the items. The greater the weight is and therefore the more important the item is, the bigger the font size and the circle are. Moreover, in case an item belongs to a cluster, its circle is colored in the color of its cluster. If, however, the items have scores, the colors of their circles depend on their scores: blue if the score is 0, green if the score is 1 and red if the score is 2. Another way to set the color of the circle is according to the specification in a map file (red, green and blue).

In the density view, each item has a color, between red and blue, depending on its density. The item is colored close to red (the color of the highest item density) if it has a lot of items in its neighbourhood that are close to each other and those neighbouring items have high weights (importance) while in the opposite case, the color of the circle is closer to blue (the color of the lowest item density).

Similar to the density view, in the cluster density view, which is available in case the items are clustered, the color of the circle of an item depends on its density, but the density is displayed separately depending on the cluster that the item belongs to. Therefore, the color of the item is close to the color of its cluster. Moreover, items weighted highly count more heavily than items with low weights.

Chapter 3

Bayesian Forecasting of US Growth using Basic Time Varying Parameter Models and Expectations Data

Chapter 3 is based on Basturk, Ceyhan and Van Dijk (2014)

3.1 Introduction

Quarterly economic growth in the USA, measured as the quarterly change in log real Gross Domestic Product, has shown typical data features in the time period 1947–I until 2013–IV with, as most important ones, a time varying mean and variance. It is important to model these stylized data features which in turn may serve as inputs for forecasting growth gaps and/or as indications of the development of economic welfare. These data features have been analyzed in many papers. A complete literature analysis is beyond the scope of this chapter, but a summary is given in the next section.

In this chapter these time varying patterns in US growth are analyzed using various univariate model structures for time variation, starting from naive model structures where all features change every period to a model where the slow variation in the conditional mean and changes in the conditional variance are specified together with their interaction. Use is made of a simulation based Bayesian inferential method to determine the forecasting performance of the various model specifications. The extension of a basic growth model with a constant mean to models including time variation in the mean and variance requires careful investigation of possible identification issues of the parameters and existence conditions of the posterior under a diffuse prior. The use of diffuse priors leads to a focus on the likelihood function and it enables a researcher and policy adviser to evaluate the scientific information contained in model and data. As Hildreth (1963) argued: 'Reporting the shape of the likelihood and its properties is an important task for a Bayesian econometrician.' For this reason, in section 3.3, these topics are analyzed and a connection is made with the well known Hierarchical Linear Mixed Model (HLMM), see Hobert and Casella (1996). Results are illustrated using simulated data.

Macroeconomic data are usually not so informative on detailed data characteristics and an additional source of information is the use of expectations data on growth, for instance, from the Survey of Professional Forecasters; see Milani (2011). We make use of these data in section 3.5.5.

The use and analysis of flexible model structures about the mean and variance of US economic growth and the use of expectations data allow one to compare empirical results obtained from alternative models. This gives information on their relative strengths and weaknesses with respect to posterior accuracy and predictive performance. As stated above such models may also serve as input for other purposes like estimating a growth gap.

The contents of this chapter are organized as follows. Some stylized facts on the data are presented in section 3.2 together with a brief survey of the literature. In section 3.3 the properties and shapes of the likelihood functions for some naive models for growth with time varying parameters are analyzed together with the existence conditions of posteriors under diffuse priors. The proposed model extensions, and the properties of the posterior parameter distributions in each model are presented in section 3.4. Empirical results on prediction are presented in section 3.5 while section 3.6 presents some final remarks including suggestions for further research.

3.2 Stylized Facts about US Real GDP Growth

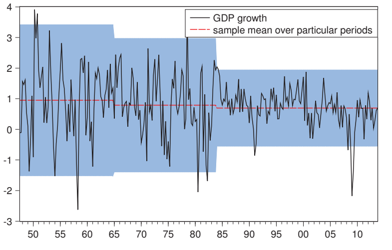
As stated in the introduction, modeling quarterly change in the log of the US real GDP series, which has undergone periodic changes in the level during the post World War II period, is important for measuring and forecasting actual and potential US growth. Changes in the pattern of US real GDP growth have been modelled by various authors in the literature. We present a selective survey of papers that relate to the topic of this chapter and refer further to the references cited in these papers. McConnell and Perez-Quiros (2000) suggest that the volatility of US real GDP decreased in early 1980s and that there is a change in the mean in US real GDP growth that occurred in the first quarter of 1984. These authors model US GDP growth as a Markov switching model with two features: (i) the mean and variance follow independent switching processes, (ii) a two-state process for the mean changes according to the state of the variance. Blanchard and Simon (2001) report on the decline of US output volatility, suggesting a decline in the standard deviation of output fluctuations over time, from about 1.5 % a quarter in the early 1950s to less than 0.5 % in the late 1990s. Blanchard and Simon (2001) shows that this decline is not continuous since volatility increases from the late 1960s to the mid-1980s, and this is followed by a sharp decline in late 1980s. Their paper mentions two possible ways of modeling the output volatility: (i) a declining trend with interruptions in 1970s and early 1980s, and (ii) a downward shift in mid-1980s. Clarida et al. (2000) consider two sub-periods for different output volatility: (i) 1960Q1-1979Q2 (pre-Volcker) and (ii) 1979Q3-1996Q4 (Volcker and Greenspan). Kim et al. (2004) suggest that the decline in the volatility of the US real GDP growth beginning in the early 1980s is concentrated in the cyclical component rather than in the trend component. Penelope and Summers

(2009) find a decline in the US GDP volatility beginning in late 1984, using a model in which the mean and variance of GDP growth are influenced by latent state variables following independent Markov chain processes. Kim and Nelson (1999a) use a Bayesian approach to identify a change in the mean at an unknown change point in a Markov-switching model of the business cycle. Their paper suggests a decline in the variance of shocks and a change in the mean in 1984–1. It also finds a narrowing gap between growth rates during recessions and booms. Several other authors studied US growth data but as indicated earlier a detailed analysis is beyond the scope of this chapter. The lesson from this literature analysis is to seriously consider and model time varying parameter models in order to model the observed low and high frequencies in the data.

Before the modeling step, it is useful to perform an exploratory data analysis on the data series that are extended over a longer period than the earlier studies listed above. An important stylized fact observed from Figure 3.1 about US real GDP growth is that the mean of the series is changing during the post WWII period, albeit not continuously like a random walk where at each period of time a shift in the series is observed. Instead, we observe occasional and discrete shifts in the mean of the series around data points like 1965 and 1984.

The mean of quarterly (annual) growth stays around 0.95% (3.8%) until the first break 1965–I at which time an downward move to a mean quarterly (annual) growth rate to around 0.83% (3.2%) is shown. This level is preserved until the second break date around 1984–I at which time another negative shock decreases the mean quarterly (annual) growth rate to approximately 0.68% (2.7%). We note that these estimates of growth are sensitive to the selection of the change point. Mean growth level at the beginning of the sample is much lower if observations prior to 1953 are disregarded. Furthermore, if the second change point is set later than 1984–I, the decline in mean growth in the last period is

Figure 3.1: Quarterly US GDP growth for the period between 1947–I and 2013–IV



Note: The figure shows GDP growth data (in percentages) and sample means over three sub-periods: 1947Q1–1964Q4, 1965Q1–1984Q4 and 1985Q1–2013Q4. The area bands are constructed as the area of two standard deviations around the sample mean, based on the sample standard deviation for each sub-periods.

more pronounced, hence it is important to assess break dates from data instead of this a priori definition[1].

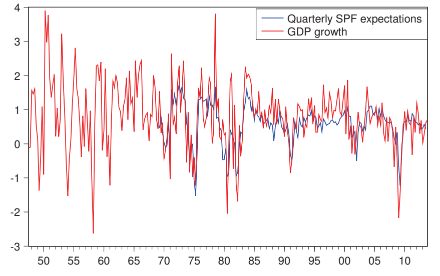
A second stylized feature that the data suggests is that the volatility of the real GDP growth series is changing in particular in the mid 1960s and mid 1980s. Basic findings regarding the volatility of the series are shown in Figure 3.1, which suggests that the volatilities in different sub-periods of the real GDP growth series is not the same, but show a declining behaviour in later years, especially after mid 1980s (the Great Moderation)[2]. Therefore, a second contribution would be adding the stochastic volatility assumption taking into account the fact that the volatility of the series is not constant, but changing over time.

Given that the information in macroeconomic series is usually not very strong about the typical data features of low and high frequencies, use is made of expectations data. These data are obtained from the Survey of Professional Forecasters (SPF), conducted

[1]Given the range of their data, which spans the time period 1953Q2-1999Q2, McConnell and Perez-Quiros (2000) conclude that the US real GDP series is stable. In our analysis,however, with an expanded data set, we reach a different conclusion.

[2]Indeed, for different sub-periods, we find that the sample standard deviations are 1.23,1.00 and 0.62, which are close to the volatility values reported in Blanchard and Simon (2001) using rolling estimation windows.

Figure 3.2: One quarter ahead US GDP growth expectations from SPF for the period between 1969–I and 2013–IV



by Federal Reserve Bank of Philadelphia and are shown in Figure 3.2[3]. For the analysis, expected GDP growth for the next quarter is used for the period between 1969–I and 2013–IV. Prior to 1992, SPF data provides expected GNP growth. This difference in the survey results are corrected by adding the difference between GNP and GDP growth to data for each quarter before 1992. Given that these data follow the actual growth data reasonably well, in particular in recession periods, it may be expected that this information will strengthen the information in the likelihood function and as a consequence the empirical results.

3.3 Basic Model Structures with Time Varying Mean

This section presents basic univariate models which allow for changes in the mean of the series. Possible identification issues are discussed in relation to relevant model parameters. The existence conditions for posterior distributions are analyzed under diffuse and informative priors.

[3]Expectations data is available from `http://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/`.

In section 3.3.1, it is shown that time variation is not straightforward to incorporate even when a simple model for the data mean is considered. Identification and existence conditions need to be examined carefully. Specifically, one cannot completely free this standard model from assumptions on the long run expected mean: an intuitive and tractable as well as flexible structure must be assumed on the long run expected mean parameter. In section 3.3.2, we show that two alternative models, a simple local level model and the hierarchical linear mixed model, provide intuitive and flexible structures for the long run mean. We show that these alternative models have common properties in terms of identification, and a careful definition of priors enables Bayesian inference of these alternative models. We conclude, however, that these alternative models should be extended to incorporate macroeconomic data features of low and high frequencies by defining a realistic time series structure for the time varying means.

3.3.1 Basic models for time varying means

Consider a simple model allowing for time varying means for a univariate time series y_t, $t = 1, \ldots, T$:

$$y_t = \mu_t + \epsilon_t, \;\; \epsilon_t \sim NID(0, \sigma_\epsilon^2), \tag{3.1}$$

where μ_t is a scalar, time-varying parameter and $\sigma_\epsilon^2 > 0$.

Equation (3.1) is a general time-varying form of a standard model for the data mean, where the time variation in this mean, μ_t, is not linked to any specific model or assumptions. As the basic case one has $\mu_t = \mu, \forall t$ and this is the well known constant mean model. Given a flat prior the marginal posterior density of the constant mean, μ, is proportional to the likelihood function and has the usual bell shaped curve with the posterior mean equal to the sample mean. This latter estimate is also used for forecasting. In this section, we consider different specifications for the time variation in the mean, and illustrate identification issues and existence conditions of the posterior for each specification.

The likelihood of the model in (3.1) is:

$$\ell(y|\mu,\sigma_\epsilon^2) = \prod_{t=1}^{T} \ell(y_t|\mu_t,\sigma_\epsilon^2) = \prod_{t=1}^{T} \phi(y_t;\mu_t,\sigma_\epsilon^2) \propto \left(\sigma_\epsilon^2\right)^{-T/2} \prod_{t=1}^{T} \exp\left(-\frac{(y_t-\mu_t)^2}{2\sigma_\epsilon^2}\right), \quad (3.2)$$

where the $\phi(x;\mu,\sigma^2)$ denotes the normal density function with mean μ and variance σ^2. In order to construct this likelihood, we make use of (3.1) which indicates that observations are independent of each other, given the model parameters, and that the observation at time t depends on the time-invariant parameter σ_ϵ^2 and the time varying parameter μ_t at time t.

Non existence of the posterior in the general time-varying mean model under a flat prior

We show first that the model in (3.1) leads to an unbounded likelihood function without further restrictions on the time varying parameters μ_t. For this purpose, consider the parameter points where $\mu_t = y_t, \forall t$ and let $\sigma_\epsilon^2 \to 0$. The likelihood with these parameter settings is:

$$\lim_{\sigma_\epsilon^2\to 0} \ell(y|\mu = y,\sigma_\epsilon^2) \propto \lim_{\sigma_\epsilon^2\to 0} \left(\left(\sigma_\epsilon^2\right)^{-T/2} \prod_{t=1}^{T} \exp\left(-\frac{(y_t-\mu_t)^2}{2\sigma_\epsilon^2}\right)\right) \quad (3.3)$$

$$= \lim_{\sigma_\epsilon^2\to 0} \left(\left(\sigma_\epsilon^2\right)^{-T/2} \left(\exp\left(-\frac{0}{2\sigma_\epsilon^2}\right)\right)^T\right) = \infty. \quad (3.4)$$

For a similar result for a model with time varying variances, we refer to De Pooter et al. (2009). Hence, under flat priors on the parameters, the posterior in (3.1) does not exist. The reason for this result is that we are in a situation of overfitting and lack of degrees of freedom: There are T observations and one has $T+1$ parameters to estimate. A mechanical way to deal with this degree of freedom problem is to assume that one has several sub periods where the mean growth is constant, albeit different in each sub period, and that one knows a priori the switching points in the mean growth. Given at least three observations where the mean is constant, the posterior is proper under a flat prior, see also De Pooter et al. (2009). However, this is a too mechanical solution of the existence problem since in most cases one does not have such precise information on switch points. A different way of dealing with the existence issue is to impose a probability distribution for the time varying parameters, as in a local level model or a random coefficients model. In the remainder of this section, we show that such restrictions can be employed in a Bayesian setting with adequate prior distributions for the mean parameter.

3.3.2 Shapes of likelihood functions for alternative model structures

A Simple Local Level Model (SLLM)

In order to impose identifying restrictions on the model in (3.1), we first consider a simple local level model for the time-varying means, μ_t such that:

$$y_t = \mu_t + \epsilon_t, \;\; \epsilon_t \sim NID(0, \sigma^2_\epsilon), \tag{3.5}$$

$$\mu_t = \mu^\star + \eta_t, \;\; \eta_t \sim NID(0, \sigma^2_\eta), \tag{3.6}$$

where $\mu^\star$ is a scalar, $\sigma^2_\epsilon > 0$, $\sigma^2_\eta > 0$ and $E(\epsilon_t, \eta_k) = 0$, $\forall t, k$.

The model in (3.5) and (3.6) imposes a shrinkage structure on the time varying parameters μ_t towards a constant value $\mu^\star$ and the posterior is proper but a new problem arises. The two variances of the disturbances defined in this model are not identified in this model under flat priors. This result is also known as a label switching problem, see Frühwirth-Schnatter (2006). To show this, we obtain the reduced form model by inserting

(3.6) in (3.5):

$$y_t = \mu^\star + \epsilon_t + \eta_t = \mu^\star + \upsilon_t, \;\; \upsilon_t \sim NID(0, \sigma_\epsilon^2 + \sigma_\eta^2), \tag{3.7}$$

where the corresponding likelihood function is:

$$\ell(y|\mu^\star, \sigma_\epsilon^2, \sigma_\eta^2) = \left(2\pi(\sigma_\epsilon^2 + \sigma_\eta^2)\right)^{-T/2} \prod_{t=1}^{T} \exp\left(-\frac{(y_t - \mu^\star)^2}{2(\sigma_\epsilon^2 + \sigma_\eta^2)}\right). \tag{3.8}$$

Assuming flat priors for $\mu^\star, \sigma_\epsilon^2$ and σ_η^2, the marginal posterior of the variance parameters can be calculated as follows:

$$p(\sigma_\epsilon^2, \sigma_\eta^2|y) = \int p(\mu^\star, \sigma_\epsilon^2, \sigma_\eta^2|y) d\mu^\star \tag{3.9}$$

$$\propto \int \left(\sigma_\epsilon^2 + \sigma_\eta^2\right)^{-T/2} \prod_{t=1}^{T} \exp\left(-\frac{(y_t - \mu^\star)^2}{2(\sigma_\epsilon^2 + \sigma_\eta^2)}\right) d\mu^\star \tag{3.10}$$

$$= \int \left(\sigma_\epsilon^2 + \sigma_\eta^2\right)^{-T/2} \prod_{t=1}^{T} \exp\left(-\frac{y_t^2 - 2y_t\mu^\star + (\mu^\star)^2}{2(\sigma_\epsilon^2 + \sigma_\eta^2)}\right) d\mu^\star \tag{3.11}$$

$$= \int \left(\sigma_\epsilon^2 + \sigma_\eta^2\right)^{-T/2} \exp\left(-\frac{\sum y_t^2 - 2T\bar{y}\mu^\star + T(\mu^\star)^2}{2(\sigma_\epsilon^2 + \sigma_\eta^2)}\right) d\mu^\star \tag{3.12}$$

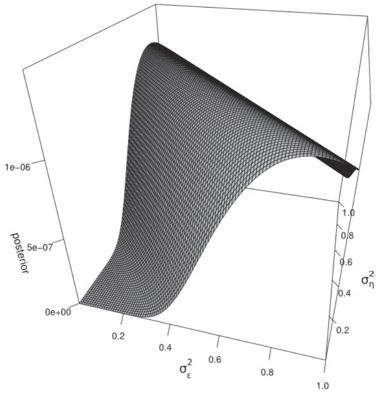
$$= \left(\sigma_\epsilon^2 + \sigma_\eta^2\right)^{-T/2} \exp\left(-\frac{\sum y_t^2 - T\bar{y}^2}{2(\sigma_\epsilon^2 + \sigma_\eta^2)}\right) \int \exp\left(-\frac{T(\mu^\star - \bar{y})^2}{2(\sigma_\epsilon^2 + \sigma_\eta^2)}\right) d\mu^\star \tag{3.13}$$

where the last equation shows that $\mu^\star$ has a conditional normal distribution with sample mean and variance $(\sigma_\epsilon^2 + \sigma_\eta^2)/T$. Using this property, (3.13) becomes:

$$p(\sigma_\epsilon^2, \sigma_\eta^2|y) \propto \left(\sigma_\epsilon^2 + \sigma_\eta^2\right)^{-(T-1)/2} \exp\left(-\frac{\sum_{t=1}^{T} (y_t - \bar{y})^2}{2(\sigma_\epsilon^2 + \sigma_\eta^2)}\right), \tag{3.14}$$

i.e. the joint posterior is of an unknown form, posterior for the sum of variance terms $(\sigma_\epsilon^2 + \sigma_\eta^2)$ is an inverted Gamma density, and the conditional posteriors $p(\sigma_\epsilon^2|\sigma_\eta^2, y)$ and $p(\sigma_\eta^2|\sigma_\epsilon^2, y)$ are identical.

Figure 3.3: Marginal posterior of the variance parameters for simulated data from the local level model in (3.5) and (3.6) for different values of σ_ϵ^2 and σ_η^2 satisfying the restriction $(\sigma_\epsilon^2 + \sigma_\eta^2) = 1$.



From (3.5) and (3.6), it can be seen that the conditional posteriors of the variance parameters are inverted Gamma densities, given flat priors for all model parameters. Similarly, the conditional posterior of $\mu^\star$ is a normal density from (3.6). Standard Gibbs sampling using these conditionals is possible despite the identification issue for the two variance parameters.

An illustration of the marginal posterior of the variance parameters in (3.14) is given in Figure 3.3 for simulated data with $T = 10$ and $(\sigma_\epsilon^2 + \sigma_\eta^2) = 1$ for different values of σ_ϵ^2 and σ_η^2 satisfying this restriction. Note that the posterior distribution exists under a flat prior on the parameters, but the two variance terms are not identified separately, hence the marginal posterior density has a ridge in Figure 3.3.

The identification or label switching problem in the SLLM can be avoided in several ways. A simple one is to impose the inequality condition that σ_η^2 is greater then σ_ϵ^2 and to use proper prior distributions for σ_ϵ^2 and σ_η^2. An alternative solution is to define a proper prior distribution on the so-called 'signal-to-noise ratio', defined by $q = \sigma_\eta^2/\sigma_\epsilon^2$. Rewriting the likelihood in (3.8) with this change of parameters we obtain:

$$\ell(y|\mu^\star, \sigma_\epsilon^2, q) = \left(2\pi(1+q)\sigma_\epsilon^2\right)^{-T/2} \prod_{t=1}^{T} \exp\left(-\frac{(y_t - \mu^\star)^2}{2(1+q)\sigma_\epsilon^2}\right). \tag{3.15}$$

A 'regularizing prior' for the model can be obtained, e.g. with a normal prior for the signal-to-noise ratio, $q \sim N_{[0,\infty)}(0, \sigma_q^2)$ and flat priors for $\mu^\star$ and σ_ϵ^2. Combining the likelihood in (3.15) with these priors, the marginal posterior of the variance parameters are:

$$p(\sigma_\epsilon^2, q|y) \propto \int \left((1+q)\sigma_\epsilon^2\right)^{-T/2} (\sigma_q^2)^{-1/2} \exp\left(-\frac{q^2}{2\sigma_q^2}\right) \prod_{t=1}^{T} \exp\left(-\frac{(y_t - \mu^\star)^2}{2(1+q)\sigma_\epsilon^2}\right) d\mu^\star, \tag{3.16}$$

where the integration steps for $\mu^\star$ can be followed as in (3.9)–(3.13), and the marginal posterior of the variance parameters is:

$$p(\sigma_\epsilon^2, q|y) \propto \left((1+q)\sigma_\epsilon^2\right)^{-(T-1)/2} (\sigma_q^2)^{-1/2} \exp\left(-q^2/(2\sigma_q^2)\right) \prod_{t=1}^{T} \exp\left(-\frac{(y_t - \bar{y})^2}{2(1+q)\sigma_\epsilon^2}\right), \tag{3.17}$$

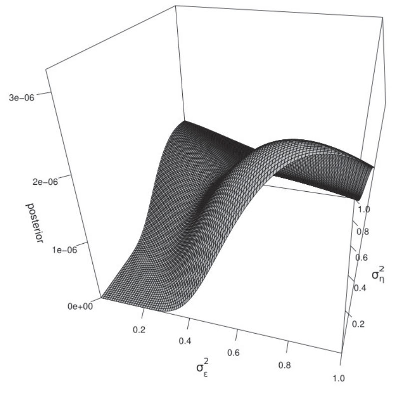
i.e. the regularizing prior only adds the factor $\exp\left(-q^2/(2\sigma_q^2)\right)$ to the posterior density in (3.14). Figure 3.4 illustrates the marginal posterior density of σ_ϵ^2 and σ_η^2 in (3.17) with $q \sim N_{[0,\infty)}(0, 0.5)$ for simulated data with $T = 10$ observations, where the marginal posterior has a clear peak and no ridge is visible. Notice that the definition of a prior for the signal-to-noise ratio relates to the properties of the SLLM. In the extreme case of $q = 0$, there is no time variation in this model. On the other hand, when $q \to \infty$, all variation in the data are captured by the state equation disturbances. [4]

A Hierarchical Linear Mixed Model (HLMM)

An equivalent simplification of the general time varying parameter model in (3.1) is

[4] In the empirical analysis, we implicitly use this ratio of variances to specify proper and regularizing priors for the variance parameters.

Figure 3.4: Marginal posterior of the variance parameters for simulated data from the local level model in (3.5) and (3.6) for different values of σ^2_ϵ and σ^2_η, under the prior $q \sim N_{[0,\infty)}(0, 0.5)$ and flat priors for the rest of the parameters.



achieved by interpreting the SLLM as a random coefficients model that is known as the Hierarchical Linear Mixed Model, see for instance Hobert and Casella (1996). The HLMM defines the model restrictions by specifying priors for the time-varying parameters.

We illustrate the HLMM model by defining normal prior hyperparameters $\mu^\star$ and σ^2_η and using a normal prior for the time-varying mean:

$$\mu_t | \mu^\star, \sigma^2_\eta \sim NID(\mu^\star, \sigma^2_\eta). \tag{3.18}$$

It is easily seen that the HLMM model with the prior distribution in (3.18) and flat priors for the rest of the parameters is equivalent to the simple local level model in (3.5) and (3.6). Therefore, the identification issue, and the possibility of avoiding the identification issue using a proper prior for the signal-to-noise ratio are valid for this model as well.

Given the results on the basic models for time-varying means, the SLLM and HLMM, the next step is to incorporate data features which are more complex than the simple shrinkage mean considered so far. Specifically, incorporating macroeconomic data features of low and high frequencies should be incorporated in a meaningful and tractable way in order to model the time series features of time varying means. In section 3.4 we present a set of alternative extended models and the corresponding prior distributions which aim to avoid the identification issue in the basic time-varying models.

3.4 Prior and Posteriors of Extended Model Structures

In this section, we present alternative models for US GDP growth, where different forms of time variation in long run growth are defined. In addition, the standard growth model is extended using data on growth expectations. The extended models are based on the conventional AR(1) model used for US GDP growth.

The standard AR(1) model is defined as follows

$$y_t = \mu + \upsilon_t \qquad t = 1, ..., T \tag{3.19}$$

$$\upsilon_t = \rho\upsilon_{t-1} + \epsilon_t \qquad \epsilon_t \overset{iid}{\sim} N(0, \sigma_{\epsilon_t}) \tag{3.20}$$

where y_t for $t = 1, \ldots, T$ is the GDP growth at time t, ρ is the autocorrelation coefficient, ϵ_t and υ_t are the disturbances and $\sigma_{\epsilon_t} > 0$. The parameter μ in this model provides the expected long run growth rate.

Inserting (3.20) in (3.19), we obtain the following AR(1) model for GDP growth:

$$y_t - \mu = \rho(y_{t-1} - \mu) + \epsilon_t \qquad \epsilon_t \overset{iid}{\sim} N(0, \sigma_{\epsilon_t}) \tag{3.21}$$

which provides a model for growth in deviation from its expected mean. Note that when $\rho = 1$ one has an identification problem since μ is not identified in this case. From the exploratory data analysis, it can be seen that this event is unlikely for this dataset. Otherwise, the approach of Schotman and Van Dijk, (1991b) and Schotman and Van Dijk, (1991a) can be used.

A straightforward extension of this model, which allows for time-variation in the long-run means is the well known local level time series model with $\mu_t = \mu_{t-1} + \eta_t$, see Harvey (1989, ch. 2). Despite its general specification in terms of time-varying means, this local level model implies an I(1) process for growth data, which is unlikely to hold. Therefore, we propose alternative model structures with different time variation in parameters that do not assume an I(1) process for growth.

3.4.1 Level shifts

The first extension of the model in (3.21), denoted by 'LS', introduces changes in long-run growth through 'occasional' shifts over time:

$$y_t - \mu_t = \rho(y_{t-1} - \mu_{t-1}) + \epsilon_t \tag{3.22}$$

$$\mu_t = \mu_{t-1} + \kappa_t \eta_{1,t} \tag{3.23}$$

for $t = 1, \ldots, T$ observations where μ_t is the time-varying long run growth, with initial value $\mu_0 \sim N(\hat{\mu}_0, \sigma_{\mu_0})$, κ_t has a Binomial distribution with parameter p_κ, $\epsilon_t \sim NID(0, \sigma_\epsilon^2)$, $\eta_{1,t} \sim NID(0, \sigma_{\eta_1}^2)$ and $E(\epsilon_t, \eta_{1,k}) = 0, \forall t, k$. In this model, changes in levels depend on p_κ, hence only a single additional parameter is introduced in the model. Occasional and large level shifts correspond to low values of p_κ together with high values of σ_{η_1}. When p_κ is 1, the model becomes standard local level model.

3.4.2 Stochastic volatility

The second extension of the model in (3.21), denoted by 'SV', introduces time variation in the volatility of growth through a stochastic volatility component in the observation equation in (3.22). For this extension we define time-varying variances such that $\epsilon_t \sim N(0, \sigma_{\epsilon_t}^2)$ with $\sigma_{\epsilon_t} = \exp(h_t/2)$, and the time varying stochastic volatility component is incorporated in the model through an additional state equation:

$$h_t = h_{t-1} + \eta_{2,t} \tag{3.24}$$

where $\eta_{2,t} \sim NID(0, \sigma^2_{\eta_2})$. Note that parameter $\sigma^2_{\eta_2}$ can be estimated from data, but the empirical identification of this variance parameter is not trivial.

3.4.3 Survey data on growth expectations

The final extension to the standard growth model is the inclusion of growth expectations data, denoted by 'EXP', as an explanatory variable in the model. Since the growth model is defined in deviations from (possibly time-varying) means, we consider a model specification where the growth expectations are also modeled in deviation from expected means. For this purpose, the observation equation in the generalized model in (3.22) is rewritten as follows:

$$y_t - \mu_t = \rho(y_{t-1} - \mu_{t-1}) + \beta(S_t - \mu_t) + \epsilon_t,\ \epsilon_t \sim N(0, \sigma^2_{\epsilon_t}), \tag{3.25}$$

where S_t denotes the expected growth from survey data for time t, obtained in the previous period $t-1$. In this model, mean growth can be defined as a constant mean or as the level shift specification in (3.23). Similarly, a constant volatility or SV model in (3.24) can be adopted depending on the definition of $\sigma^2_{\epsilon_t}$.

3.4.4 Bayesian inference for proposed models

Given the above alternative model extensions, conventional priors can be used for each model parameter and posterior samplers are based on Gibbs sampling for the parameters.

The conventional priors refer to normal or flat priors for the mean μ or the auto-correlation coefficient ρ, and inverse gamma priors for the variance parameters. The conditional posterior densities of these parameters then follow normal and inverse gamma distributions.

For the models with level shifts, conditional on other model parameters and level shifts κ_t, draws from unobserved states are obtained using the Kalman Filtering and Smoothing algorithm, see Carter and Kohn (1994) and Frühwirth-Schnatter (1994). Conditional on the unobserved states, draws from level shifts are obtained using the algorithm proposed by Gerlach et al. (2000), see also Giordani and Kohn (2008) for details.

Finally, for models with a stochastic volatility component, draws from the unobserved variable h_t can be obtained using the Kalman filter and smoother, similar to the unobserved states μ_t. However, since the SV model is based on the logarithmic transformation of the variance, the resulting conditional density follows a log-χ^2 distribution. Noticing the properties of log-χ^2 distribution, Kim et al. (1998) and Omori et al. (2007) approximate this distribution using a mixture of Gaussian distributions. Hence, conditional on these mixture components the system remains Gaussian allowing for standard inference outlined above. For details, see Omori et al. (2007).

For a detailed description of the all the sampling steps, we refer to Chapter 4 of this thesis.

3.5 Empirical Results

In this section we apply the standard and extended growth models to quarterly US GDP growth for the period between 1947–I and 2013–IV. As described in section 3.4, five different model structures are considered for US GDP data:

AR autocorreltion/serial correlation in errors

LL local level model for GDP growth levels, i.e. GDP growth level is subject to continuous changes over time

LS level shifts in GDP growth, i.e. GDP growth level is subject to occasional changes over time

SV stochastic volatility in GDP growth, i.e. GDP growth volatility changes over time

EXP growth model incorporating growth expectations from professional forecasters as an additional explanatory variable

where one or more of these structures are combined in alternative models.

3.5.1 Prior definitions

The priors for the parameters of alternative model structures are defined independently, as outlined in section 3.4. For models with serial correlation (AR), a flat prior is defined for the correlation coefficient: $\rho \sim U(0,1)$. For standard growth models with no changes in GDP growth level, a flat prior is defined for mean GDP growth $\mu \sim U(-\infty,\infty)$. For models incorporating GDP growth expectations, a flat prior is defined for the effect of expectations, i.e. $\beta \sim U(-\infty,\infty)$.

For models with constant error variance for GDP growth variance, we define proper but uninformative priors: $\sigma_\epsilon^2 \sim IG(T/2, \sigma^2/2 \times T/2)$, where $IG(a,b)$ denotes the inverse Gamma distribution with shape a and scale b, $T = 267$ is the number of observations and σ^2 is the data variance. This data based prior implies that the prior is equivalent to $T/2$ observations, hence we expect the data with T observations to dominate the prior.

In addition, for models with level changes in GDP growth (LL, LS), we define inverse Gamma priors: $\sigma_v^2 \sim IG(T/2, \sigma^2/20 \times T/2)$. Note that the a priori mean of this variance is 10 times smaller than the a priori mean of the observation variance defined above. This specification is based on the intuition that GDP growth levels, if modeled as time varying components, should not have drastic changes over time but capture longer run GDP growth levels. Despite this definition, the prior we define is still relatively uninformative compared to the information in the likelihood: it is equivalent to $T/2$ observations only and the data information is likely to dominate this prior.

For models with occasional level changes in GDP growth (LS) we set a prior level change probability of 0.04, indicating around 5 a priori expected changes in gdp levels during the sample period. For models with the stochastic volatility component and occasional level changes in GDP growth, we set a prior level change probability of 0.02. This lower prior probability for a level change is intended to empirically identify more accurately the mean and variance changes at the same time.

3.5.2 Estimation results

In this subsection, we report estimation results from eight alternative models combining one or more of the AR, SV, LL and LS model structures.[5] Table 3.1 presents parameter estimates from all compared models based on 10000 draws where we consider the first 5000 draws as burn-in draws.

Despite the differences in model structures, Table 3.1 shows that parameter estimates from alternative models are similar across all models we consider. For models with constant mean, mean quarterly growth is around 0.8, which is in line with the visual inspection of the data. For models with serial correlation, the posterior mean for the correlation coefficient is around 0.3, i.e. correlation levels seem to be far from the nonidentification region $\rho = 1$. Two unique models in these alternative models are LL and AR-LL models, which imply a unit root in GDP growth. Despite this counter-intuitive assumption, results of the LL and AR-LL models are roughly in line with the constant mean models in terms of parameter estimates.

Table 3.1 shows that models which allow for time variation in means, with LL and LS components, in general lead to smaller observation variances compared to the constant mean models (AR and AR-SV). This result is intuitive since part of the data variation is

[5] Extended models incorporating growth expectations are analyzed separately since the sample size is much smaller due to the range of the data on growth expectations.

Table 3.1: Estimation results from alternative models

model	μ	ρ	σ^2_ϵ	σ^2_υ
models without level changes				
constant mean	0.793	-	0.885	-
	(0.059)	-	(0.072)	-
AR	0.792	0.380	0.775	-
	(0.089)	(0.056)	(0.064)	-
AR-SV*	0.804	0.377	0.356	-
	(0.044)	(0.056)	(0.241)	-
models with frequent level changes				
LL	-	-	0.569	0.054
	-	-	(0.038)	(0.006)
AR-LL	-	0.251	0.562	0.087
	-	(0.077)	(0.042)	(0.034)
models with occasional level changes				
LS	-	-	0.743	0.024
	-	-	(0.055)	(0.003)
AR-LS	-	0.368	0.685	0.023
	-	(0.055)	(0.049)	(0.003)
AR-LS-SV*	-	0.321	0.305	0.064
	-	(0.078)	(0.238)	(0.012)

Note: The table reports posterior means and standard deviations (in parentheses) of model parameters. σ^2_ϵ and σ^2_υ are the observation and state variances, respectively. (*) indicates that the observation variance is time varying in these models. For these models we report mean observation variance and standard deviation for the sample period. Results are based on 10000 draws of which 5000 are burn-in draws.

explained by the changing means in models with the time variation in mean growth. Furthermore, estimated state variances are very small compared to the observation variance, indicating rather smooth changes in growth levels over time.

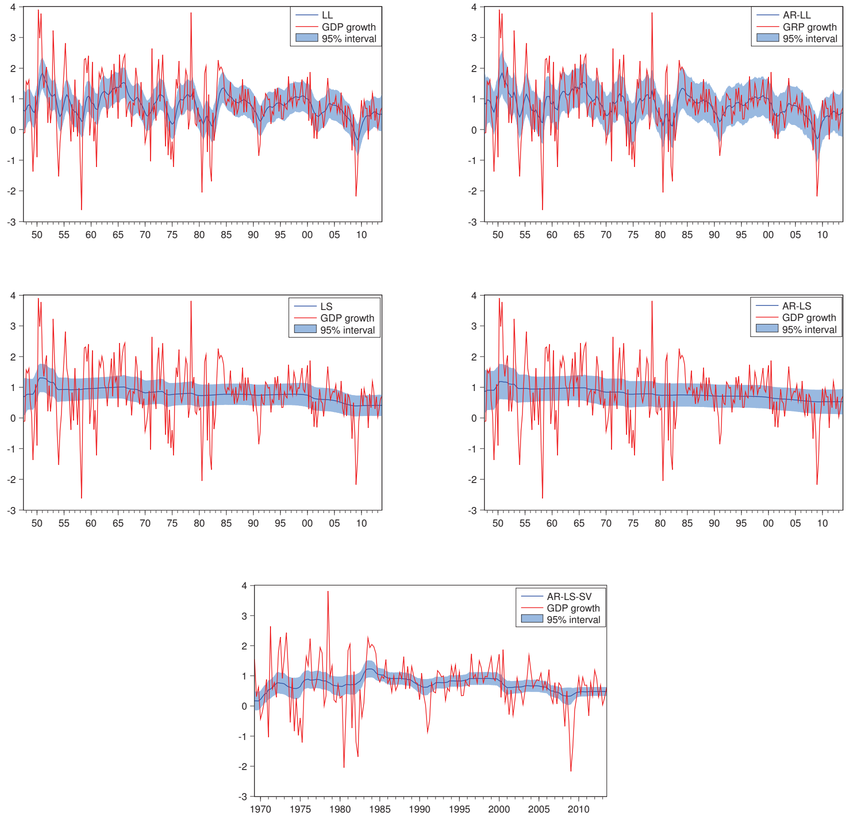
We next report estimated growth means for the five models with time variation in means, namely LL, AR-LL, LS, AR-LS and AR-LS-SV models. Figure 3.5 presents these estimates together with the 95% HPDI for estimated mean growth. All models imply changing growth means over time. In line with the model definitions, AR-LL and LL models lead to more volatile growth levels over time while AR-LS and AR-LS-SV models lead to less volatile growth levels. All models imply relatively high growth levels at the beginning of the sample and relatively low growth levels towards the end of the sample. These results confirm stylized facts reported in section 3.2 although we find more than two periods of level changes in AR-LS and AR-LS-SV models, as shown in Figure 3.6.

Finally, we note that incorporating the stochastic volatility component in the model seems to be important. Estimated volatility levels from the two models with SV component, AR-SV and AR-LS-SV, are given in Figure 3.7. Both models indicate substantial changes in volatility over time. Estimated volatility also confirms the stylized facts covered in section 3.2: the period of the 50s and the beginning of the 60s are characterized by relatively high volatility in growth and the volatility level decreases towards the end of the sample period, apart from the recession of 2000 and the recent crisis. Comparing the two panels in Figure 3.7, it can be seen that allowing for time variation in means (AR-LS-SV) leads to much lower volatility levels compared to the counterpart model with constant means (AR-SV) at the beginning of the sample. In other words, the volatile period of 50s is explained both by the changing means and changing volatility in AR-LS-SV model.

3.5.3 Prediction results

In order to assess the predictive performance of the compared models in section 3.5.2, we consider 1 quarter ahead forecasts of GDP growth for the period between 1997–I and

Figure 3.5: GDP growth level estimates from alternative models with time varying growth means



Note: The figure presents mean posterior levels for GDP growth ($\times 100$) for compared models for the whole sample period. Results are based on 10000 draws of which 5000 are burn-in draws.

Figure 3.6: Estimated level shifts for models with occasional level shifts (AR-LS and AR-LS-SV)

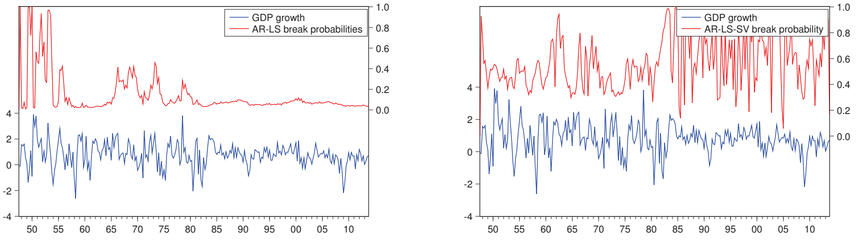


Figure 3.7: Estimated mean volatility for models with stochastic volatility (AR-SV and AR-LS-SV)

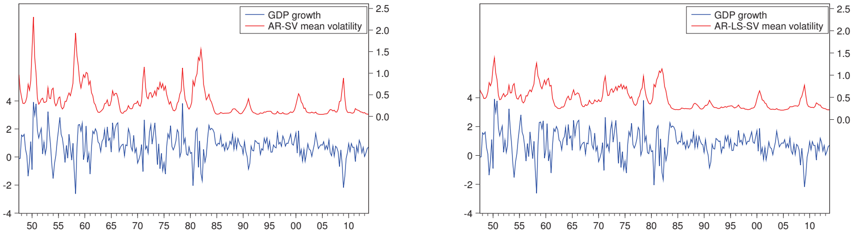


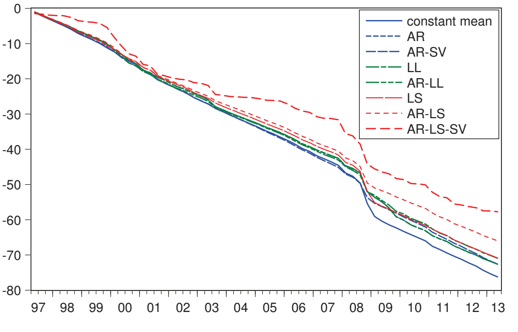
Table 3.2: 1 period ahead MSFE and cumulative predictive likelihoods

model	MSFE	Cumulative Pred. Lik.
models without level changes		
constant mean	0.5170	-76.2513
AR	0.3916	-69.4748
AR-SV	0.3990	-72.6564
models with frequent level changes		
LL	0.5347	-72.6307
AR-LL	0.5035	-70.9595
models with occasional level changes		
LS	0.4748	-70.9052
AR-LS	0.3905	-66.0502
AR-LS-SV	0.3888	-57.7819

2013–IV. We compare the predictive performance of models using the average MSFE and cumulative predictive likelihood for the prediction period. These results are given in Table 3.2 and Figure 3.8.

Table 3.2 and Figure 3.8 show that predictive performances of alternative models are quite different, despite the similar parameter estimates reported in section 3.5.2.

Figure 3.8: Cumulative predictive likelihoods for the prediction sample from alternative models



First, the serial correlation component is found to be important for the predictive performance of growth models since the standard AR model, the AR-LL and AR-LS models perform better in terms of MSFE and predictive likelihoods compared to their counterparts without the AR component, namely the constant mean model, LL model and LS model.

Second, models with frequent level shifts, the LL and AR-LL models, do not perform well compared to models with occasional level shifts, LS and AR-LS, in terms of the predictive likelihoods. According to these predictive performances, models allowing for occasional level shifts, i.e. slowly varying long-run growth levels, are more appropriate to forecast growth.

Third, models with occasional level shifts (AR-LS and AR-LS-SV) perform better than the standard growth model with serial correlation, hence changes in growth levels should definitely be accounted for.

Finally, a comparison of the prediction results for AR-LS and AR-LS-SV models in Table 3.8 indicate that adding a stochastic volatility component improves the predictive performance of the growth model substantially, particularly according to the predictive likelihoods.

From these results, we conclude that the most appropriate model for predicting growth is the AR-LS-SV model, which takes into account the slowly changing levels of growth, serial correlation in disturbances and changing disturbance variances over time.

3.5.4 Prior predictive likelihood analysis

The predictive performance comparisons in section 3.5.3 are naturally affected by the prior choices for parameters, especially since the existence and identification issues illustrated in section 3.3 require proper priors for part of the model parameters in the alternative models we consider. In this subsection, we report the effect of the selected priors on prediction

Table 3.3: Prior predictive analysis: 1 period ahead MSFE and cumulative predictive likelihoods

model	MSFE	Cumulative Pred. Lik.
constant mean	34.131	-2297.532
AR	11.696	-813.025
AR-SV	0.482	-58.386
models with frequent level changes		
LL	0.495	-70.142
AR-LL	0.606	-77.464
models with occasional level changes		
LS	0.499	-66.682
AR-LS	0.503	-62.260
AR-LS-SV	0.486	-58.141

results. An important point is whether the superior performance of the AR-LS-SV can be attributed solely to prior definitions.

The effect of the priors are analyzed using a prior-predictive likelihood analysis. For this purpose, we do not update model parameters using data, but instead use parameter draws from prior densities, i.e. Gibbs sampling steps are performed for each model without updating the model parameters. For models with level changes and stochastic volatility, unobserved levels are updated using the data since these variables are not model parameters but unobserved states. For all considered models, MSFE and cumulative predictive likelihoods resulting from the prior-predictive likelihood analysis are presented in Table 3.3.

According to the prior-predictive analysis in Table 3.3, the AR-SV model is the ‘best’ alternative model in comparison to the preferred model, AR-LS-SV. MSFE and predictive likelihoods of these models are very similar without the likelihood information. Hence the superior predictive performance of the AR-LS-SV model in section 3.5.3 is not solely attributed to the prior definitions.

The 'least preferred' models according to the prior-predictive likelihoods are the constant mean model and the standard AR model. This result is expected for two reasons. First, uninformative prior densities for these models lead to very wide prediction intervals.[6] Second, in models with level changes and stochastic volatility, unobserved states are updated using the data even though model parameters are not updated. Hence these models make use of data information to update unobserved states regardless of the prior definition.

Table 3.3 also shows that the models with occasional level shifts (LS, AR-LS, AR-LS-SV) are not necessarily favored by the prior as the AR-SV model, without mean changes, leads to better prior-predictive results than most models with occasional level shifts. The conclusion on the necessity to include occasional level shifts in section 3.5.3 does not seem to be the result of the employed prior densities. Similarly, models with AR components are also not necessarily favored by the prior definitions.

Comparing the AR-LS and AR-LS-SV models in Table 3.3, we conclude that the priors favor the model stochastic volatility component. Despite this result, the prior domination seems to be less than the data domination particularly for predictive likelihoods: The improvement in predictive likelihoods of the AR-LS-SV model is much more pronounced in Table 3.2 compared to the prior-predictive likelihood comparison in Table 3.3. Furthermore, the priors favor the simple AR-SV model substantially, according to the MSFE and cumulative predictive likelihoods. Despite this finding, the preferred model, AR-LS-SV, is found to perform better than the simple AR-SV model in the predictive analysis in section 3.5.3.

[6] One important point is the flat prior we use for the standard growth models with no change in levels. For the prior predictive analysis, we truncate this prior density between $[-10, 10]$ since none of the posterior draws lie outside this region.

Table 3.4: Estimation results from alternative models for US growth between 1969–I and 2013–III

model	μ	ρ	β	σ^2_ϵ	σ^2_υ
models without level changes					
AR	0.804	0.377	-	0.356	-
	(0.044)	(0.056)	-	(0.241)	-
AR-EXP	0.679	0.131	0.456	0.570	-
	(0.488)	(0.079)	(0.133)	(0.059)	-
models with occasional level changes					
AR-LS-SV*	-	0.280	-	0.294	0.036
	-	(0.097)	-	(0.144)	(0.006)
AR-LS-SV-EXP*	-	0.118	0.106	0.288	0.035
	-	(0.083)	(0.039)	(0.138)	(0.005)

Note: For all models, sample is determined as the period for which expectation data is available. Results are based on 10000 draws of which 5000 are burn-in draws.

3.5.5 Results using growth expectations

In this subsection, we consider growth models incorporating growth expectations as a final addition to the proposed growth models. We note that the expectation data is only available for the period between 1969–I and 2013–IV. Therefore, for a fair comparison of models, alternative models are estimated for this smaller sample period. Similarly, prediction comparisons, in terms of MSFE and predictive likelihoods, are also based on the time period for which expectation data is available.

Table 3.4 presents estimation results for 4 models, the constant mean, AR, AR-EXP, AR-LS-SV and AR-LS-SV-EXP models. For both models with expectations, AR-EXP and AR-LS-SV-EXP, 95% HPDI for parameter β is found to be on the positive region, i.e. expectations are found to drive GDP growth (in deviation from the long-run mean). Furthermore, posterior means for the serial correlation parameter, ρ, is lower when expectations are incorporated in the models hence part of the persistence in growth rates are captured by the persistence in expectations.

Table 3.5: Prediction results from alternative models for US growth between 1969–I and 2013–III

model	MSFE	Cumulative Pred. Lik.
models without level changes		
AR	0.360	-76.78
AR-EXP	0.367	-77.52
models with occasional level changes		
AR-LS-SV	0.371	-65.16
AR-LS-SV-EXP	0.375	-64.54

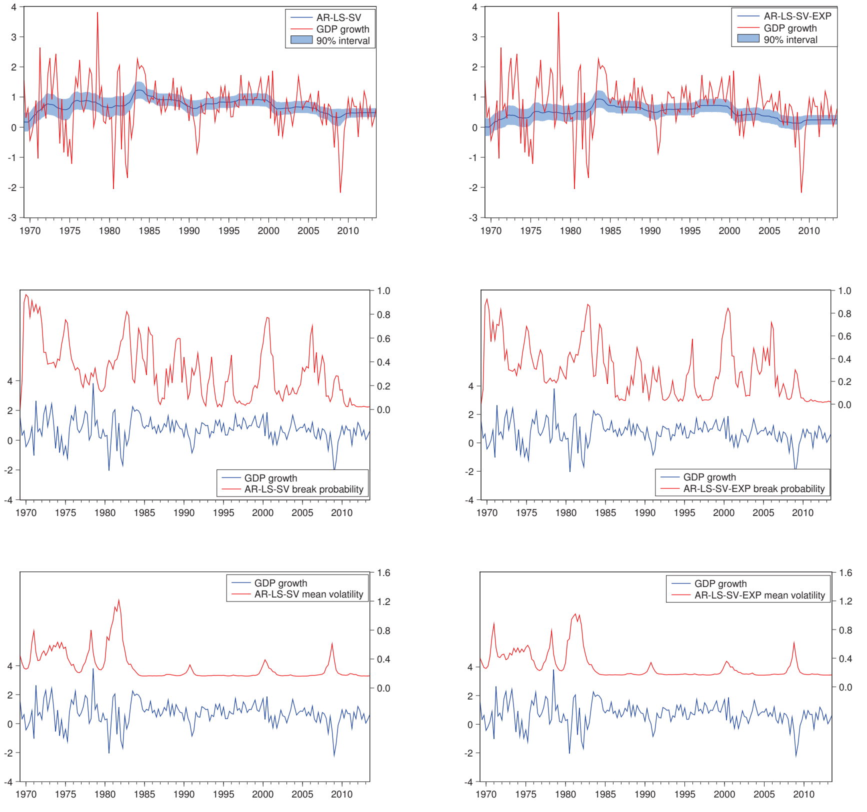
Note: For all models, sample is determined as the period for which expectation data is available. Forecast sample includes data between 1994–IV and 2013–IV. Results are based on 10000 draws of which 5000 are burn-in draws.

Figure 3.9 presents estimated levels, break probabilities and mean volatilities for AR-LS-SV and AR-LS-SV-EXP models, i.e. models with time varying levels. Comparing the graphs top panel in Figure 3.9, it can be seen that including growth expectations in the model leads to smoother estimated long-run levels since part of the variation in growth is captured by the variation in growth expectations. Despite this finding, the middle and bottom panels in Figure 3.9 shows that estimated break probabilities and volatility of growth follow similar patterns regardless of the inclusion of expectations in the model.

For a detailed comparison of the models with and without growth expectations, we consider the predictive performances of the four models in Table 3.4, using 1 quarter ahead MSFE and cumulative predictive likelihoods. The predictive performances are reported in Table 3.5 for the forecast period 1994–IV and 2013–IV.

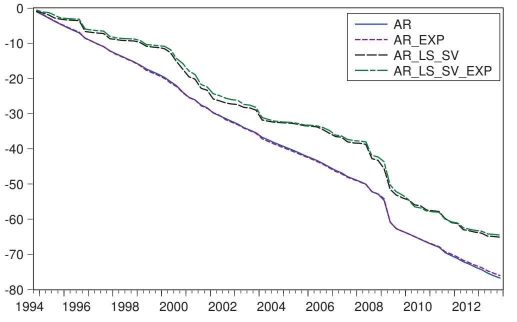
The results in Table 3.5, which are based on a smaller sample period compared to section 3.5.3, confirm that modeling occasional shifts (AR-LS-SV or AR-LS-SV-EXP) in mean growth improve the predictive power of standard growth models. Secondly, according to the MSFE values reported in Table 3.5, adding expectations to AR or AR-LS models causes a slight deterioration in point forecasts. However, according to the cumulative predictive likelihoods, expectations clearly improve predictive power of these

Figure 3.9: Estimation results for the AR-LS-SV model and AR-LS-SV-EXP models



Note: The figure presents estimated levels, break probabilities and mean volatility for GDP growth (in percentages) for the sample period 1969Q2–2013Q4, the period for which growth expectations are available. Results are based on 10000 draws of which 5000 are burn-in draws.

Figure 3.10: Cumulative predictive likelihoods from alternative models for US growth between 1969–I and 2013–III



Note: For all models, sample is determined as the period for which expectation data is available. Forecast sample includes data between 1994–IV and 2013–IV. Results are based on 10000 draws of which 5000 are burn-in draws.

models. Hence, if the purpose is to obtain density forecasts of growth, incorporating expectations in the model structure is important.

The evolution of predictive likelihoods for all compared models through the forecast sample period are shown in Figure 3.10. The model incorporating expectations, AR-LS-SV-EXP, leads to improvements in predictive likelihoods in two important periods in the sample: the period between 2000 and 2004, and between 2008 and 2009, i.e. the recession in early 2000s and the recent Great recession. We therefore conclude that incorporating expectations in growth models can be useful especially in turbulent times, even if the performance of the models point forecasts (measured by MSFE) does not improve with the addition of expectation data.

3.6 Final Remarks

We analyzed time varying patterns in US growth using various univariate model structures, starting from a naive model structure where all features change every period to a model where the slow variation in the conditional mean and changes in the conditional variance are specified together with their interaction, including survey data on expected growth in order to strengthen the information in the model. Our results indicate that incorporating time variation in mean growth rates as well as volatility of mean growth rates are important for improving the predictive performances of standard growth models. Furthermore, using data information on growth expectations is important for forecasting growth in specific periods, such as the 2000s and around 2008.

The empirical analysis leads to several topics for further research in relation to incorporating time varying patterns and growth expectations in growth models. First, the approach of this chapter can be extended to analyze the difference between actual and potential growth, the so-called growth gap and also to analyze features of the business cycle like the dating procedure of Pagan and Harding (2005). Second, instead of 'selecting' a specific model, the proposed models with time variation or with GDP expectations can be 'combined', as in Billio et al. (2013) to combine the weights of the expectations data and the model information over time, with possibly an improvement of the overall predictive performance of the growth model. Finally, the proposed models can be applied to GDP growth of countries other than the US in order to analyze international evidence on the predictive gains from changing growth levels, variances and incorporating growth expectations in standard growth models. This requires extending our work and including a panel data model.

Chapter 4

Posterior-Predictive Evidence on US Inflation using Extended New Keynesian Phillips Curve Models with non-filtered Data

Chapter 4 is based on Baştürk, Çakmaklı, Ceyhan and Van Dijk, (2013b).

4.1 Introduction

Modeling the relation between inflation and fluctuations in economic activity has been one of the building blocks of macroeconomic policy analysis. Often, the analysis of this relation, denoted as New Keynesian Phillips Curve (NKPC) models, is conducted using the short-run variations in inflation and economic activity. The conventional method for extracting this short run variation in the observed series is to demean and detrend the data prior to analysis, see Galí and Gertler (1999). However, mechanical removal of the low frequency movements in the data may lead to misspecification in the models, as suggested in Ferroni (2011) and Canova (2012) for DSGE models. The existence of complex low frequency movements, such as structural breaks and level shifts in the observed series, in particular, in the inflation series, is well documented in the literature, see McConnell and Perez-Quiros (2000) and Stock and Watson (2008). Distinct periods with different patterns can be observed for the non-filtered inflation series. The period between the early 1970s and the early 1980s is often labeled as a high inflationary period compared to earlier and later periods. A similar type of statement holds for economic activity. The real marginal cost series, often used as a proxy for economic activity, see Galí and Gertler (1999), follows a negative trend which is amplified in the recent decades. Note that in the literature, there are papers which use the output gap variable instead of real marginal cost as a proxy for economic activity. However, as Rudd and Whelan (2007) suggests, the coefficient of output gap turns out to have the wrong sign. Moreover, the construction of the output gap variable using the potential output is not trivial. Therefore, in this study we use the real marginal cost. The importance of joint analysis of such high and low frequency movements in macroeconomic data has recently been documented, see Ferroni (2011), Delle Monache and Harvey (2011), Canova (2012), and Faust and Wright (2013).

In this chapter we aim to contribute to this literature in four ways. We illustrate and discuss possible effects that simple prior filtering of the low frequencies in the data may have on posterior and predictive inference using a basic NKPC model. The issue is that the observed inflation and marginal cost data have more complex low frequency structures than just a simple constant mean and/or a basic linear or HP trend. We show that this misspecification affects posterior inference of the structural NKPC parameters and gives poor forecasting results depending on the model specification. In appendix 4.A, we present extensive evidence on this feature using a set of simulated and real data and a range of NKPC model structures. Obviously, in well specified models and in series with relatively constant means and linear trends the misspecification effects are not severe. However, from the outset, the use of mechanical filters without properly examining the frequency features of the data is not advisable.

We extend the basic NKPC model by specifying structural time series models which allow for stochastic trends, structural breaks and stochastic volatility in inflation and log marginal costs and integrate these with the basic model. The more complex model structure enables the identification of the relation between macroeconomic variables inherent in the NKPC model, together with possible long and short run dynamics in each series.

Next, we enrich the extended NKPC models to include both forward and backward looking expectation components. There is a debate in the literature on the relative weights of these two components in explaining and forecasting inflation patterns in the U.S.. Our combined model structure can provide valuable information on that point.

As a final contribution we make use of survey data on inflation expectations from the University of Michigan Research Center, which provides quarterly one year ahead inflation expectations. It is well known that the class of NKPC models including complex time series features and basic expectation mechanisms is not easy to estimate given the usually weak data information and the few available weak instrumental variables. The proposed richer expectation mechanism and making use of survey data strengthen the likelihood information and are expected to make inference more efficient and forecasting more accurate.

Several alternatives to structural time series models for efficiently combining the NKPC model with explicit low frequency movements in the data are available. One alternative is to focus only on the high frequencies by rewriting the likelihood in the frequency domain and maximizing the (log)likelihood only over a portion of fluctuations, see e.g. Christiano and Vigfusson (2003). Another alternative is to utilize multiple prior filters, to capture possibly incorrectly specified low frequency components, see Canova and Ferroni (2011). Here we focus on explicitly modeling the low frequency movements to improve the predictive performances of the structural form models while we keep the theoretical model at a simple tractable level.

We apply the proposed set of models to quarterly U.S. data over the period 1960-I until 2012-I. For all models considered, posterior and predictive results are obtained using a simulation based Bayesian approach. Our results indicate that NKPC structures with three additional components (structural time series features, expectation mechanisms and inflation survey data) capture time variation in the low and high frequency movements of both inflation and marginal cost data. For the inflation series, the extended model identifies distinct periods with different inflation levels and volatilities. In terms of marginal costs, the local linear trend specification accommodates the smoothly changing trend observed in the series, specifically after 2000. We also find improved forecasting performance of the extended NKPC models when these are compared with basic NKPC models with demeaned and/or detrended data and with the standard stochastic volatility model proposed by Stock and Watson (2007) and, further, with an extended Bayesian Vector Autoregressive (BVAR) model which accounts for changing levels, trends and volatility in the data. The model comparison is based on predictive likelihood and Mean Squared Forecast Error (MSFE) comparisons. The Bayesian approach we adopt has additional appealing features for the models considered. In terms of inflation predictions, several measures of interest, such as deflation probabilities obtained from the lower tail of the complete predictive densities, are obtained as a by-product of Bayesian inference. Furthermore, for the most general model with good fit and forecasting features, weak endogeneity and almost non-existence of a stable long-run relationship between inflation and marginal costs can easily be assessed using the posterior draws of the trends and levels.

The structure of this chapter is as follows: Section 4.2 presents the three extensions to the standard NKPC model structure. Section 4.3 provides the application of the proposed models and the standard NKPC model on U.S. inflation and marginal cost data. Section 4.4 concludes. Additional illustrations, results, details of the posterior sampling algorithms and references are provided in the appendix 4.E and appendix 4.F.

4.2 Extended New Keynesian Phillips Curve Models

We start with a standard NKPC model based on a priori filtered data. Next, we extend this model with a structural time series model in order to deal with low and high frequencies that are present in U.S. inflation and the low frequency property in the U.S. log marginal cost series. Thirdly, we extend the latter NKPC model by introducing a Hybrid NKPC model (HNKPC) with both backward and forward looking inflation expectations making use of observed inflation expectations from survey data.

The standard NKPC can be derived by the approximation of the equilibrium conditions of the firms under staggered price setting using the Calvo formulation, see Calvo (1983). The Calvo model implies that a fraction of firms optimize their prices while the remaining fraction, i.e. non-optimizing firms, keep their prices unchanged. Assuming zero inflation at the steady state the basic NKPC model derived from the firm's price setting is given as

$$\begin{aligned} \tilde{\pi}_t &= \lambda \tilde{z}_t + \gamma_f E_t(\tilde{\pi}_{t+1}) + \epsilon_{1,t}, \\ \tilde{z}_t &= \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \epsilon_{2,t}, \end{aligned} \tag{4.1}$$

where $\tilde{\pi}_t$ is the filtered inflation and $\tilde{z}_t$ is the filtered (log) real marginal cost, $(\epsilon_{1,t}, \epsilon_{2,t})' \sim NID(0, \Sigma)$, λ is the slope of the Phillips curve, γ_f is the weight given to the forward looking inflation, and standard stationary restrictions hold for (ϕ_1, ϕ_2).

One way to estimate this model is to replace the expectation term with actual inflation values relying on rational expectations. Another option is to use survey data on expected inflation as 'observed' expectations. Still, direct substitution of survey data for expected inflation does not exploit the full model structure. In our modeling strategy, we use the full data generating process for real marginal costs together with the Phillips curve relation to form inflation expectations.[1] Iterating the model forward and computing future expected inflation, model (4.1) implies that inflation can be expressed as the sum of the current and future discounted stream of the real marginal costs. Given the AR(2) dynamics for

[1]We also estimate the model by inserting the survey expectations directly in model (4.1). The results are provided in the appendix 4.K.

the long run deviation of the marginal costs, one can compute this sum and obtain a closed-form solution of model (4.1). The NKPC model takes the form of an instrumental variable model with nonlinear parameters in the inflation equation[2]

$$\begin{aligned} \tilde{\pi}_t &= \frac{\lambda}{1-(\phi_1+\phi_2\gamma_f)\gamma_f}\tilde{z}_t + \frac{\phi_2\gamma_f\lambda}{1-(\phi_1+\phi_2\gamma_f)\gamma_f}\tilde{z}_{t-1} + \epsilon_{1,t} \\ \tilde{z}_t &= \phi_1\tilde{z}_{t-1} + \phi_2\tilde{z}_{t-2} + \epsilon_{2,t}. \end{aligned} \tag{4.2}$$

One way to estimate the structural parameters is by estimating the parameters of the unrestricted reduced form model using a uniform prior and solve for the structural form parameters, see the appendix 4.B and Kleibergen and Mavroeidis (2011) for details. However, this parameter transformation involves a complex Jacobian determinant that may seriously obscure posterior inference on the structural parameters. Hence we opt for estimating structural parameters directly.

Extended NKPC models: low frequency components, non-filtered data

We depart from the standard NKPC model by avoiding a priori data filtering and emphasize that data filtering is an integral part of modeling from an econometric point of view. Specifically, we make use of models with time varying levels as well as volatility for capturing both the low and high frequency changes in the U.S. inflation and marginal costs. Furthermore, estimating data filters together with other model parameters concerns the uncertainty related to long run specifications. Modeling the data filters explicitly takes this uncertainty into account while the use of filtered data does not. Finally, prior data filtering also has important effects on the predictive performance of the models as shown in section 4.3.

There exists a substantial literature on the connection between actual inflation and target inflation and the firms' pricing behavior. We summarize the major issues here. In full equilibrium DSGE models with explicit monetary policy modeling, the mean level of the inflation is related to the target inflation rate. In these specifications, the target inflation rate is either assumed to be constant or is allowed to change to accommodate

[2] The model in (4.2) can be written as a triangular simultaneous equations model (SEM).

variation in inflation level. Prominent examples include Woodford (2003) and Sargent et al. (2006), who fix the target inflation and Erceg and Levin (2003), Schorfheide (2005), Ireland (2007) and Liu et al. (2011), who allow for discrete or continuous changes in the target inflation level.

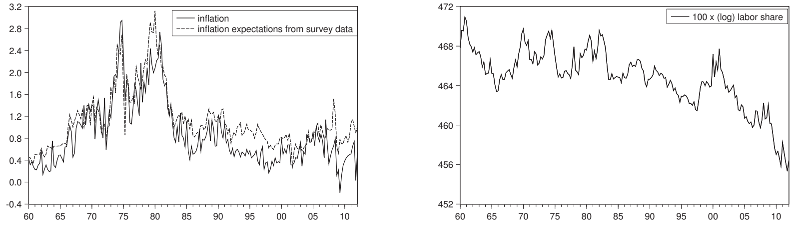
In our partial equilibrium NKPC models, the specification of the steady state inflation level and the firms' decision process are of key importance for the final model structure. In the standard NKPC models as in Galí and Gertler (1999), assuming zero steady state inflation and keeping the prices fixed for the non-optimizing firms results in the standard form of NKPC as in model (4.1). Ascari (2004) and Ascari and Sbordone (2013), extend these models to allow for constant positive trend inflation and they analyze the implications of the trend inflation on the NKPC structure. Cogley and Sbordone (2008) take one-step further and derive the NKPC model with time-varying trend inflation modeled as a driftless random walk. Adding a trend inflation to standard NKPC assumptions, while preserving the assumption that non-optimizing firms keep their prices fixed causes the resulting NKPC coefficients to depend on the trend inflation. Thus, the interpretation of the coefficients differs from the standard model in these extended models.

Other assumptions on non-optimizing firm's pricing behaviour include indexation on past inflation (i.e. non-optimizing firms change their prices based on past inflation), see Smets and Wouters (2003) and Christiano et al. (2005). Alternatively, Smets and Wouters (2007) and Liu et al. (2011) make use of steady state inflation. As discussed in Ascari (2004) and Levin and Yun (2007) the structure of the NKPC remains as in model (4.1) with constant parameters if the non-optimizing firms adjust their price by the steady state inflation. Indeed, this is the route taken in Yun (1996), Jeanne (1998) and Schorfheide (2005). Moreover, Nason and Smith (2008) provide empirical evidence in favor of stable structural parameters. In our extended NKPC models with non-filtered data we follow this assumption and keep the structural parameters constant focusing on short and long run inflation levels.

The proposed joint modeling of data filters and other model parameters is also motivated by the stylized facts for the non-filtered U.S. inflation and log marginal cost data, shown in Figure 4.1 over the period between 1960-I and 2012-I.[3] The left panel displays distinct periods with differing inflation patterns. The period between the early 1970s and the early 1980s can be labeled as a high inflationary period with high volatility compared with the remaining periods. Existing evidence shows that the decline in inflation level and volatility is due to credible monetary policy that stabilized inflationary expectations since the early eighties, see McConnell and Perez-Quiros (2000) and Stock and Watson (2007). One way to model this changing inflation behavior is to allow for regime changes in parameters, see Sims and Zha (2006) and Cogley and Sbordone (2008). We consider two cases for the inflation process. In the first case, we assume continuous level shifts in inflation using a random walk process

$$c_{\pi,t+1} = c_{\pi,t} + \eta_{1,t+1}, \quad \eta_{1,t} \sim NID(0, \sigma^2_{\eta_1}). \tag{4.3}$$

Figure 4.1: Inflation, inflation expectations and log real marginal cost (×100) series over the period 1960-I until 2012-I



Alternatively, we consider an inflation level subject to occasional and discrete shifts. Such level shifts are modeled as follows

$$c_{\pi,t+1} = c_{\pi,t} + \kappa_t \eta_{1,t+1}, \quad \eta_{1,t} \sim NID(0, \sigma^2_{\eta_1}), \tag{4.4}$$

[3]Inflation is computed as the continuously compounded growth rate of the implicit GDP deflator and for the real marginal cost series we use labor share in non-farm business sector obtained from http://research.stlouisfed.org/fred2/, see Galí and Gertler (1999) for details. The right panel in Figure 4.1 displays real marginal costs, in natural logarithms and multiplied by 100.

where κ_t is a binary variable taking the value of 1 with probability p_κ if there is level shift and the value 0 with probability $1-p_\kappa$ if the level does not change. This model structure allows for level shifts to depend on p_κ while preserving a parsimonious model structure with only a single additional parameter. Occasional and large level shifts correspond to low values of p_κ together with high values of σ_{η_1}. When p_κ is 1, the model becomes the local level model of (4.3). We use both specifications (4.3) and (4.4) in the empirical analysis.

The real marginal costs, shown in the right panel of Figure 4.1, does not exhibit discrete changes as observed in the inflation series. These data instead have a continuously changing pattern around a negative trend, which can be attributed to technology shocks. Since this trend is more prominent in the second half of the sample period, we allow for a changing trend using a local linear trend specification

$$\begin{aligned} c_{z,t+1} &= \mu_{z,t} + c_{z,t} + \eta_{2,t+1}, \quad && \eta_{2,t} \sim NID(0, \sigma^2_{\eta_2}) \\ \mu_{z,t+1} &= \mu_{z,t} + \eta_{3,t+1}, && \eta_{3,t} \sim NID(0, \sigma^2_{\eta_3}). \end{aligned} \tag{4.5}$$

This specification is flexible enough to encompass many types of filters used for detrending, see Delle Monache and Harvey (2011) and Canova (2012) for a similar specification in the more general context of DSGE models. When $\sigma^2_{\eta_3} = 0$, the level of the real marginal costs follow a random walk with a drift, μ_z. Additionally, when $\sigma^2_{\eta_2} = 0$, a deterministic trend is obtained. Note that, setting only $\sigma^2_{\eta_2} = 0$ but allowing $\sigma^2_{\eta_3}$ to be positive results in an integrated random walk process which can approximate nonlinear trends including the Hodrick-Prescott (HP) trend.

Together with the level specifications of the inflation and real marginal costs, the NKPC model in (4.2) using (4.4) and (4.5) takes the following form

$$\begin{aligned}
\pi_t - c_{\pi,t} &= \tfrac{\lambda}{1-(\phi_1+\phi_2\gamma_f)\gamma_f}\left(z_t - c_{z,t}\right) + \tfrac{\phi_2\gamma_f\lambda}{1-(\phi_1+\phi_2\gamma_f)\gamma_f}\left(z_{t-1} - c_{z,t-1}\right) + \epsilon_{1,t}, \\
z_t - c_{z,t} &= \phi_1\left(z_{t-1} - c_{z,t-1}\right) + \phi_2\left(z_{t-2} - c_{z,t-2}\right) + \epsilon_{2,t}, \\
c_{\pi,t+1} &= c_{\pi,t} + \kappa_t\eta_{1,t+1}, \\
c_{z,t+1} &= \mu_{z,t} + c_{z,t} + \eta_{2,t+1}, \\
\mu_{z,t+1} &= \mu_{z,t} + \eta_{3,t+1},
\end{aligned} \tag{4.6}$$

where $(\epsilon_{1,t},\ \epsilon_{2,t})' \sim NID\left(0, \begin{pmatrix} \sigma^2_{\epsilon_1} & \rho\sigma_{\epsilon_1}\sigma_{\epsilon_2} \\ \rho\sigma_{\epsilon_1}\sigma_{\epsilon_2} & \sigma^2_{\epsilon_2} \end{pmatrix}\right)$, $(\eta_{1,t},\ \eta_{2,t},\ \eta_{3,t})' \sim NID\left(0, \begin{pmatrix} \sigma^2_{\eta_1} & 0 & 0 \\ 0 & \sigma^2_{\eta_2} & 0 \\ 0 & 0 & \sigma^2_{\eta_3} \end{pmatrix}\right)$ and the disturbances $(\epsilon_{1,t},\ \epsilon_{2,t})'$ and $(\eta_{1,t},\ \eta_{2,t},\ \eta_{3,t})'$ are independent for all t.

Adding stochastic volatility as high frequency component

A further refinement in the NKPC model can be achieved allowing for time variation in the variances of the disturbances. This extension is particularly appealing for the inflation series, as the inflation variance changes over time substantially, see e.g. Stock and Watson (2007) for a reduced form model with a stochastic volatility component. The following state equation extends the NKPC model with a stochastic volatility process for inflation

$$h_{t+1} = h_t + \eta_{4,t+1},\ \eta_{4,t+1} \sim NID(0, \sigma^2_{\eta_4}), \tag{4.7}$$

where we specify a time-varying volatility, $\sigma_{\epsilon_{1,t}} = \exp(h_t/2)$, in the first equation in (4.6). We follow Stock and Watson (2007) by fixing the value of $\sigma^2_{\eta_4}$ prior to analysis to facilitate inference. We set $\sigma_{\eta_4} = 0.5$, which seems to work well for U.S. inflation.

An important estimation challenge in this extended model is the close relation between the changing inflation levels and volatilities. These changing data patterns can be captured by either of these model components which makes it hard to identify these components unless one makes strong prior restrictions. We fix the value of $\sigma^2_{\eta_4}$ prior to analysis to facilitate inference and in order to impose smoothness in the volatility process. It is straightforward to extend the analysis with a more flexible, strong, stochastic prior so that the parameter $\sigma^2_{\eta_4}$ is estimated together with the rest of the parameters. We report on this in section 4.

Hybrid NKPC: forward and backward expectations using survey data

The specification of the HNKPC can be derived using an assumption on the firm's behavior, where a fraction ω of the firms, that are unable to reset their prices, adjust their price by the lagged inflation rate. The HNKPC model takes then the form of

$$\begin{aligned} \tilde{\pi}_t &= \lambda^H \tilde{z}_t + \gamma_f^H E_t(\tilde{\pi}_{t+1}) + \gamma_b^H \tilde{\pi}_{t-1} + \epsilon_{1,t}, \\ \tilde{z}_t &= \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \epsilon_{2,t}, \end{aligned} \tag{4.8}$$

where parameters of the HNKPC model, indicated by a superscript H are functions of the price stickiness parameter, a discount factor and the fraction of firms with backward looking pricing behavior. We note that the HNKPC has the same forward looking inflation expectation term in the model as the NKPC but the HNKPC has both a backward and forward looking component due to the specification of the lagged inflation deviation.

As in the NKPC case, we opt for using the full information approach by exploiting the information in the data generating process for real marginal costs.[4] Iterating the first equation forward and solving for the expected inflation, the HNKPC implies the triangular simultaneous equations model which is nonlinear in parameters

$$\begin{aligned} \tilde{\pi}_t &= \frac{\lambda^H}{(1-\gamma_b^H\gamma_f^H)(1-(\phi_1+\phi_2\gamma_f^H)\gamma_f^H)}\tilde{z}_t + \frac{\phi_2\gamma_f^H\lambda^H}{(1-\gamma_b^H\gamma_f^H)(1-(\phi_1+\phi_2\gamma_f^H)\gamma_f^H)}\tilde{z}_{t-1} \\ &+ \frac{\gamma_b^H\gamma_f^H}{(1-\gamma_b^H\gamma_f^H)}\sum_{k=1}^{\infty}(\gamma_f^H)^k E_t(\tilde{\pi}_{t+k}) + \frac{\gamma_b^H}{(1-\gamma_b^H\gamma_f^H)}\tilde{\pi}_{t-1} + \frac{1}{(1-\gamma_b^H\gamma_f^H)}\epsilon_{1,t} \\ \tilde{z}_t &= \phi_1\tilde{z}_{t-1} + \phi_2\tilde{z}_{t-2} + \epsilon_{2,t}. \end{aligned} \tag{4.9}$$

Unlike the NKPC solution, this system has a lagged inflation term and an infinite sum of inflation expectations. A closed form solution for the latter expression only exists under certain assumptions such as rational expectations.

We do not follow this route but proceed differently. Consider $E_t(\tilde{\pi}_{t+k}) = E_t(\pi_{t+k}) - E_t(c_{\pi,t+k})$ which is the difference between expected future inflation and the expected future

[4] We also estimate the model by inserting the survey expectations directly in model (4.8). The results are provided in appendix 4.K.

value of the low frequency component of inflation that we modeled in (4.4) as a process that is similar to a random walk but subject to occasional and discrete level shifts and it has a bounded variance. One can interpret this difference as the difference between short and long run inflation expectations. As a next step we substitute the observed survey data on next period's expected inflation, denoted by μ_t, for the expected inflation in period $t+1$, i.e. $\mu_t = E_t(\pi_{t+1})$ and we assume the following partial adjustment mechanism

$$\mu_t - c_{\pi,t+1} = \beta(\mu_{t-1} - c_{\pi,t}) + \eta_{5,t+1}, \tag{4.10}$$

where $|\beta| < 1$ and $\eta_{5,t+1}$ is iid and $E_t(\eta_{5,t+1}) = 0$. Iterating this equation forward and taking expectations one obtains $E_t(\mu_{t+k-1} - c_{\pi,t+k}) = \beta^{k-1}(\mu_t - c_{\pi,t+1})$. That is, the partial adjustment mechanism described in (4.10) implies that the further one gets into the future the smaller will be the difference between short and long run inflation expectations. Estimates of β will indicate the empirical speed of adjustment. For instance, for a value of the posterior mean of β equal to 0.5 it follows that within a few periods one has almost complete adjustment. Given the restriction on β one can solve (4.10) for μ_t and obtain $\mu_{t-1} = c_{\pi,t} + \sum_{j=0}^{\infty} \beta^j \eta_{5,t-j}$. That is, the observed survey inflation expectations are equal to the long run unobserved inflation pattern and an infinite moving average of errors with declining weights that are determined by the adjustment mechanism given in (4.10). This adaptive mechanism has a Bayesian learning and updating interpretation on the difference between short and long run expected inflation. Using this mechanism, the term $\sum_{k=1}^{\infty}(\gamma_f^H)^k E_t(\tilde{\pi}_{t+k})$ in (4.9) can be rewritten and the HNKPC model becomes

$$\begin{aligned}
\pi_t - c_{\pi,t} = \; & \frac{\lambda^H}{(1-\gamma_b^H\gamma_f^H)(1-(\phi_1+\phi_2\gamma_f^H)\gamma_f^H)}\left(z_t - c_{z,t}\right) + \frac{\phi_2\gamma_f^H\lambda^H}{(1-\gamma_b^H\gamma_f^H)(1-(\phi_1+\phi_2\gamma_f^H)\gamma_f^H)}\left(z_{t-1} - c_{z,t-1}\right), \\
& + \frac{\gamma_b^H\gamma_f^H}{(1-\gamma_b^H\gamma_f^H)}\frac{\gamma_f^H}{1-\gamma_f^H\beta}(\mu_t - c_{\pi,t}) + \frac{\gamma_b^H}{(1-\gamma_b^H\gamma_f^H)}\left(\pi_{t-1} - c_{\pi,t-1}\right) + \frac{1}{(1-\gamma_b^H\gamma_f^H)}\epsilon_{1,t}, \\
z_t - c_{z,t} = \; & \phi_1\left(z_{t-1} - c_{z,t-1}\right) + \phi_2\left(z_{t-2} - c_{z,t-2}\right) + \epsilon_{2,t}.
\end{aligned} \tag{4.11}$$

We emphasize that alternative models on inflation expectations exist, see Mankiw et al. (2003). For instance, a skew density for $\eta_{5,t}$ allows systematic under- or overoptimism. This is an interesting topic, but outside the scope of this chapter.[5]

Note that the model-implied expectation is for GDP inflation while the overlaid data is CPI inflation expectations. For this reason we subtract the average difference between CPI and GDP inflation from the survey data.[6] Furthermore, since the survey data provide four-steps-ahead (one-year) expectations, we divide the survey data by 4, assuming constant expectations over the year.

The NKPC model in (4.6) is a special case of (4.11) when $\gamma_b^H = 0$. Then the model becomes purely forward looking. Similar to the NKPC model, we consider three case of the HNKPC model with different specifications for inflation: (i) continuous level changes; (ii) discrete occasional level changes; and (iii) discrete occasional level changes and stochastic volatility.

4.3 Posterior and Predictive Evidence

In this section we present posterior and predictive evidence on several features of the extended NKPC models using U.S. data on inflation and marginal costs. We note that the MCMC sampler for the full conditional posterior distribution is based on Gibbs sampling with a Metropolis-Hastings step and data augmentation, combining the methodologies in Geman and Geman (1984); Tanner and Wong (1987); Gerlach et al. (2000) and Çakmaklı et al. (2011). The posterior sampler, the prior distributions and a prior sensitivity analysis using prior-predictive likelihoods are given in appendix 4.J. We compare the results with those obtained from alternative reduced form models like BVAR models and the stochastic volatility model from Stock and Watson (2007). Specifically, we estimate two NKPC

[5] Alternatively, survey expectations may be measured with an error. In this case one can specify unobserved inflation expectations anchored around observed survey expectations. We consider this possibility and report these estimation results in appendix 4.L. Such extensions do not seem to alter the results.

[6] We thank an anonymous referee for pointing this out. Our approach of recalculating the inflation expectations is similar to Del Negro and Schorfheide (2013).

models with demeaned inflation series and with detrended real marginal costs using a linear trend or the HP filter, which are labeled NKPC-LT and NKPC-HP, respectively. In six extended NKPC models we make use of structural time series models to specify low and high frequencies. The first three of these models allow for continuous changes in the level of inflation (NKPC-TV), in addition discrete occasional level shifts (NKPC-TV-LS), and in further addition stochastic volatility for inflation (NKPC-TV-LS-SV). The final three models use the HNKPC framework with forward and backward looking expectations and using survey data. The corresponding extensions are denoted as HNKPC-TV, HNKPC-TV-LS and HNKPC-TV-LS-SV. All six models use the local linear trend specification in (4.6) for the real marginal costs. A summary of the eight models used in this chapter is given in Table 4.1.

Table 4.1: Standard and extended NKPC models

model structure / low/high frequencies	NKPC	HNKPC
Inf: constant level RMC: linear trend	NKPC-LT	HNKPC-LT*
Inf: constant level RMC: Hodrick-Prescott filter	NKPC-HP	HNKPC-HP*
Inf: time varying levels RMC: local linear trend	NKPC-TV (4.2)-(4.3)-(4.5)	HNKPC-TV (4.3)-(4.5)-(4.10)-(4.11)
Inf: time varying levels and switching RMC: local linear trend	NKPC-TV-LS (4.2)-(4.4)-(4.5)	HNKPC-TV-LS (4.4)-(4.5)-(4.10)-(4.11)
Inf: ... and stochastic volatility RMC: local linear trend	NKPC-TV-LS-SV (4.2)-(4.4)-(4.5)-(4.7)	HNKPC-TV-LS-SV (4.4)-(4.5)-(4.7)-(4.10)-(4.11)

Results for the models indicated by (*) are provided in the appendix. 'Inf' ('RMC') stands for Inflation (Real Marginal Cost).

Posterior evidence

We display the estimation results in Table 4.2 and focus on four features: slope of the Phillips Curve; weight of forward and backward inflation expectations; degree of endogeneity and persistence in survey expectations. First, the slope of the NKPC ($\lambda^{(H)}$)

is estimated around 0.07 and 0.09 which is slightly higher than the conventional estimates of the Phillips curve slope, that indicate an almost flat curve, see e.g. Galí and Gertler (1999); Galí et al. (2005); Nason and Smith (2008). When we model the levels of the series explicitly, $\lambda^{(H)}$ drops to values around 0.05 for both NKPC and HNKPC models. A possible explanation for this difference is the departure from the zero steady state inflation assumed in the traditional NKPC models. As shown in Ascari (2004) and Ascari and Ropele (2007) among others, when firms that cannot re-optimize their prices keep their prices fixed, trend inflation can affect the slope of the NKPC. In this case, this slope is a decreasing function of the trend inflation. Still, in both NKPC and HNKPC models, the estimated slopes are substantially different from zero as point 0 is outside the 95% Highest Posterior Density Interval (HPDI) for most cases.

Table 4.2: Posterior results of alternative NKPC models

Model	$\lambda^{(H)}$	$\gamma_f^{(H)}$	γ_b^H	β	ρ	ϕ_1	ϕ_2
NKPC-LT	0.07 (0.03)	0.36 (0.24)	–	–	-0.01 (0.02)	0.84 (0.05)	0.08 (0.05)
NKPC-HP	0.10 (0.05)	0.43 (0.27)	–	–	-0.05 (0.04)	0.66 (0.05)	-0.01 (0.05)
NKPC-TV	0.06 (0.03)	0.39 (0.25)	–	–	-0.09 (0.06)	0.82 (0.05)	0.06 (0.05)
NKPC-TV-LS	0.05 (0.02)	0.36 (0.24)	–	–	-0.06 (0.05)	0.82 (0.05)	0.07 (0.05)
NKPC-TV-LS-SV	0.06 (0.02)	0.32 (0.23)	–	–	-0.02 (0.07)	0.87 (0.05)	0.10 (0.05)
HNKPC-TV	0.04 (0.02)	0.01 (0.01)	0.42 (0.12)	0.52 (0.29)	0.01 (0.06)	0.81 (0.05)	0.07 (0.05)
HNKPC-TV-LS	0.04 (0.02)	0.01 (0.01)	0.47 (0.10)	0.50 (0.19)	0.02 (0.01)	0.81 (0.06)	0.16 (0.07)
HNKPC-TV-LS-SV	0.06 (0.02)	0.03 (0.05)	0.21 (0.11)	0.56 (0.21)	-0.01 (0.01)	0.87 (0.05)	0.10 (0.05)

The table presents posterior means and standard deviations (in parentheses) of parameters for the competing NKPC type models estimated for quarterly inflation and real marginal costs over the period 1960-I until 2012-I. λ (λ^H) and γ_f (γ_f^H) are the slope of the Phillips curve and the coefficient of inflation expectations in NKPC (HNKPC) model in (4.2) ((4.11)). γ_b^H is the coefficient of the backward looking component in the HNKPC model in (4.11). H superscript denotes the parameters of the hybrid models while these parameters without H superscript correspond to the NKPC model counterparts. β is the autoregressive parameter for the deviation of the short run expectations from the long run, as defined in (4.10). ρ is the correlation coefficient of the residuals ϵ_1 and ϵ_2. ϕ_1 and ϕ_2 are the autoregressive parameters for the real marginal cost specification in model (4.2). Posterior results are based on 40000 simulations of which the first 20000 are discarded for burn-in. Model abbreviations are as in Table 4.1.

Second, with respect to inflation expectations, the coefficient of the short-run inflation expectations in Table 4.2, $\gamma_f^{(H)}$, is much lower than the conventional estimates, which are above 0.9 in most cases. A potential reason for this finding is the methodology used. Conventional Bayesian analyses often impose dogmatic priors on this parameter unlike our uninformative prior specification. When we consider the NKPC model with the subjective discount factor γ_f, the (implied) prior for the discount factor (either directly or through other parameter's priors in the steady state relations) is either fixed to the values around 0.99, see Smets and Wouters (2003) for example, or it is tightly centered around 0.99, see for example Schorfheide (2005); An and Schorfheide (2007). We also notice a relatively higher posterior standard deviation for this parameter, hence another potential cause of this finding is the relatively low information content in the data about this parameter. This is in accordance with the discussion in the appendix 4.B on the shape of the likelihood in these macro-models. Note that the more conventional values of this parameter are still inside the 95% HPDI.

Another reason might be the fact that, even if the models are estimated without a restriction, in most cases inflation expectations are replaced by the real leading value of the inflation relying on the rational expectations hypothesis, see e.g. Galí and Gertler (1999) and Sims (2002). However, we opt for explicitly solving for expectations resulting in a highly nonlinear system of simultaneous equations.

A striking result from Table 4.2 is the relative importance of the forward and backward looking components of the HNKPC, measured by parameters γ_f^H and γ_b^H. On the one hand, the evidence in Galí et al. (2005) suggests a dominant forward looking effect. Cogley and Sbordone (2008) document that the forward looking component of the HNKPC model dominates once the trend variation in inflation is taken into account. Similarly, Benati (2008) shows that under stable monetary regimes with clearly defined nominal anchors, inflation appears to be (nearly) forward looking. On the other hand, many studies including Fuhrer and Moore (1995); Rudd and Whelan (2005) document a dominant backward looking effect. Our results favor the latter view since the effect of the backward looking component of inflation estimated by the HNKPC models in the bottom panel of Table 4.2 are substantially higher than those of the forward looking components. More specifically, Table 4.2 shows that the HNKPC and NKPC model results differ in terms of the forward looking components' coefficient $\gamma_f^{(H)}$. From an economic point of view, these results maybe driven by the model assumptions on firm behavior that differs from those of Cogley and Sbordone (2008) and Benati (2008). From an econometric point of view, as in the NKPC case, the specification of the prior distribution is crucial. In many analyses, the implied prior on these parameters suggests a support of the distribution in the interval [0.5,1] ([0,0.5]) for γ_f^H (γ_b^H), see Smets and Wouters (2003, 2007); Benati (2008); Del Negro and Schorfheide (2008) and Del Negro and Schorfheide (2013). Hence, the difference may be partly due to the presence of only one weak instrument (second order lagged marginal costs), see Nason and Smith (2008) for further empirical results and a discussion on this topic.

Third, the contemporaneous correlation between the observation disturbances determines the degree of endogeneity of the log real marginal costs in the NKPC. The estimates of this correlation parameter, ρ, are displayed in the fifth column of Table 4.2. Posterior means of ρ from all NKPC models are negative and close to 0, with high standard deviations and point 0 is inside the 95% HPDI. For the HNKPC models, posterior means of ρ are mostly positive with an even smaller magnitude. Therefore, the endogeneity problem does not seem to be severe and single equation inference may yield credible results for inflation and marginal costs. Still, we refrain from doing so since one neglects several cross-equation restrictions in that case.

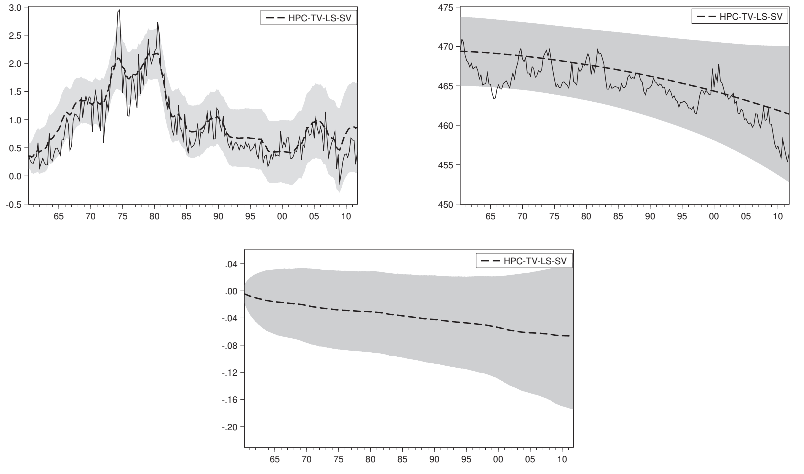
Fourth, the β parameter, which indicates the adaptation of the short run survey expectations to the long run inflation, has posterior means given in the fifth column of Table 4.2. All HNKPC models indicate relatively quick adjustment, as the posterior means are around 0.5.

Estimated Levels, Volatilities and Breaks

We present estimated levels, trends, inflation volatilities and break probabilities for the proposed HNKPC models in Figures H.1, 4.3 and 4.4, respectively. Estimates for the NKPC counterparts are similar, and are provided in appendix 4.E.

The top-left panel of Figure H.1 shows estimated levels for the HNKPC-TV-LS-SV model. We first stress that models that only allow for discrete and occasional level shifts

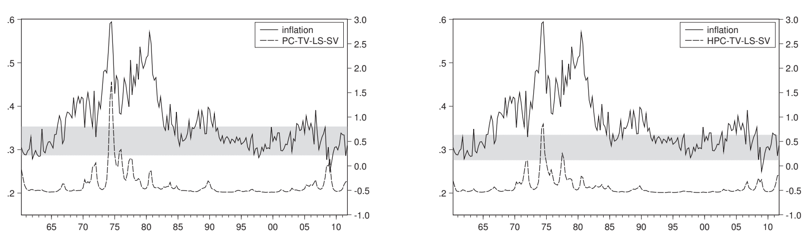
Figure 4.2: Level, trend and slope estimates from the HNKPC-TV-LS-SV model



Note: The top-left panel exhibits estimated inflation levels, $c_{\pi,t}$ in model (4.6). The top-right and bottom panels show estimated (log) real marginal cost levels and the slopes, $c_{z,t}$ and $\mu_{z,t}$ in model (4.6), respectively. Grey shaded areas correspond to the 95% HPDI. Model abbreviations are as in Table 4.1. However, for notational convenience we use the abbreviation '(H)PC' instead of the abbreviation '(H)NKPC' in the figures. Results are based on 40000 simulations of which the first 20000 are discarded for burn-in.

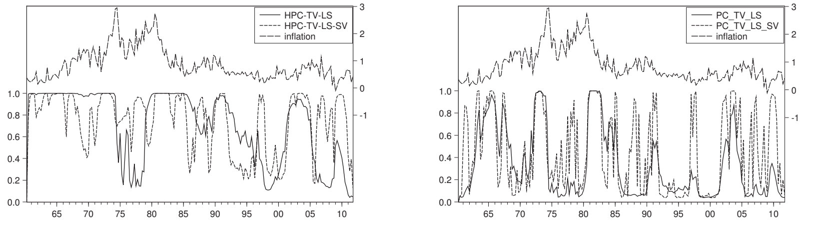
lead to smoother inflation levels compared to the model that allows for continuous level changes. Detailed results on this issue are provided in appendix 4.F. In DSGE models, mean inflation is generally connected to the inflation target in the central bank's policy rule. Hence movements in trend inflation reflect to a large extent changes in the monetary policy target (see also Schorfheide (2005); Cogley and Sbordone (2008)). Adding stochastic volatility to the model with level shifts creates more frequent discrete changes in the

Figure 4.3: Estimated inflation volatility from the (H)NKPC-TV-LS-SV models



Note: The dashed and solid lines show the posterior mean of the time varying inflation volatility and the observed inflation level. The shaded areas are the 90% HPDI of inflation volatility estimated by the equivalent models without the stochastic volatility components. Results are based on 40000 simulations of which the first 20000 are discarded for burn-in.

Figure 4.4: Estimated level shift probabilities for the NKPC and HNKPC models



Note: The solid and dotted lines are the posterior means of the estimated level shift probabilities from the (H)NKPC-TV-LS model and the (H)NKPC-TV-LS models, respectively. The dashed line is the observed inflation level. Results are based on 40000 simulations with the first 20000 discarded for burn-in.

inflation level, possibly reflecting the uncertainty in monetary policy target captured by volatility changes. Estimated marginal cost levels for the HNKPC-TV-LS-SV are given in the top-right panel of Figure H.1 and indicate a slightly nonlinear trend during the sample period.

Figure 4.3 presents estimated volatility levels for the (H)NKPC model with level shifts and the stochastic volatility component. The stochastic volatility pattern coincides nicely with data features of the Great Moderation. The decline in inflation level and volatility after the 1980s is linked to credible monetary policy that stabilized inflationary expectations at a low level via commitment to a nominal anchor since the early eighties, see Ahmed et al. (2004); Stock and Watson (2007). The effect of this is also seen in the inflation levels presented in Figure H.1. This period of low volatility is replaced by a volatile period after 2005 and during the recent financial crisis. A slight difference between NKPC and HNKPC models occurs during the volatility peaks around 1975. High volatility is distributed more evenly in the HNKPC model with stochastic volatility, whereas for the NKPC counterpart, high volatility is concentrated around 1975. Peak points of estimated volatilities coincide with rapid and substantial changes in inflation.

Estimated break probabilities for the NKPC and HNKPC models with and without the stochastic volatility component are presented in Figure 4.4. The estimated level shift probabilities for the NKPC-TV-LS model identify four major shifts in the inflation level around 1966, 1973, 1982 and 2005. Note that the estimated shift probabilities in the NKPC-TV-LS-SV model demonstrate the complementarity of level shifts and changing volatility. The probabilities follow a similar pattern with the NKPC-TV-LS model but the periods subject to level shifts are much longer. During the highly volatile periods of the 1970s, the model produces clear signals of changing inflation levels, as high volatility causes rapid changes in inflation. Accordingly, low volatility periods are characterized by mild but significant changes in inflation. This shows the complementarity of the stochastic volatility component and level shifts.

Predictive Performance

Predictive performances of the models are reported using predictive likelihoods, MSFEs and predictive densities which enable us to report the deflation probabilities.

The first metric we consider is the predictive likelihoods of all models in order to compare the density forecasts of the models. The one-step ahead predictive likelihood of the observation at $t_0 + 1$, y_{t_0+1}, conditional on the previous observations $y_{1:t_0}$, is

$$f(y_{t_0+1}|y_{1:t_0}) = \int p(y_{t_0+1}|X_{t_0+1}, \theta)p(X_{t_0+1}, \theta|y_{1:t_0})dX_{t_0+1}d\theta, \tag{4.12}$$

which can be computed by first generating $\{X_{t_0+1}\}_{m=1}^{M}$ for M posterior draws, using the corresponding state equations. Next, the predictive likelihood of the observation at $t_0 + 1$ can be approximated by $\frac{1}{M}\sum_{m=1}^{M} p(y_{t_0+1}|X_{t_0+1}^m, \theta^m)$, where $p(y_{t_0+1}|X_{t_0+1}^m, \theta_{1:t_0}^m)$ is a multivariate normal density and M is a sufficiently large number.

We base the MSFE and predictive likelihood comparisons on the inflation predictions. For the general case of $h \geq 1$ period ahead forecasts, the predictive density of inflation at time t is calculated conditional on the inflation and marginal cost data up to time t, the estimated mean marginal cost values for the periods $t+1, \ldots, t+h$ and, if $h > 1$, on the estimated mean inflation levels for the periods $t+1, \ldots, t+h-1$. For all models using survey expectations, predictive likelihoods are also conditioned on the observed survey expectations up to time t.

A feature of the predictive likelihoods is that these can be evaluated by $p(y_{t_0+1:T}) = \prod_{t=t_0}^{T} f(y_{t+1}|y_{1:t})$, which provides a tool to analyze the contribution of each observation at time period t, see Geweke and Amisano (2010). For the models with a priori demeaned and detrended data predictive likelihoods do not take into account the parameter uncertainty arising from this a priori step. We choose to calculate the predictive likelihoods this way, which is a fair replication of the literature.

Accurate point predictions of inflation are of key importance to economic agents such as investors and central banks. Therefore, we consider MSFE, computed as the mean of the sum of squares of the prediction errors. Point forecasts for inflation are defined as the mean of the predictive distribution, which is consistent with a quadratic loss function. We report MSFE for one and four period ahead forecasts in order to examine the forecasting ability of the models for longer horizons.

As a third performance criteria, we report the deflation risk indicated by each model, computed as the lower tail probability of the one step ahead predictive distributions.

Apart from the models considered so far, we include alternative reduced form models that are proven to have good predictive abilities. The first model is the unobserved components model proposed by Stock and Watson (2007), henceforth denoted as SW2007. This model captures the unobserved trend in inflation where both inflation and trend volatility follow a stochastic process, see SW2007 for details. The second model is an unrestricted Bayesian VAR (BVAR-SV) model with two lags and with stochastic volatility for inflation. BVAR models are one of the workhorse models used for forecasting macroeconomic series. For the sake of brevity, we do not provide details of this class of models and refer to Del Negro and Schorfheide (2013). We use the proposed ‘TV’ model extension in the BVAR-SV model, which allows for continuous changes in the level of inflation and a smoothly changing trend for the marginal costs. Both SW2007 and BVAR-TV-SV models are strong competitors for the models we propose. In all considered models, the data based prior distributions given in appendix 4.E, calculated using the full sample data, are used.

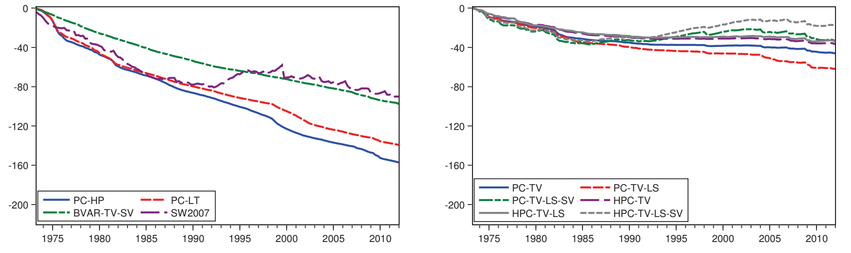
Predictive likelihoods and MSFE of the alternative models are presented in Table K.2. The likelihood contribution of each observation and the corresponding cumulative predictive likelihoods are displayed in Figure 4.5. We present the log predictive likelihoods of the competing models in the first column of Table K.2. These values together with Figure 4.5 indicate three groups of models in terms of their predictive performances. The first group of models include the conventional NKPC models with demeaned and detrended data (NKPC-LT and NKPC-HP). The second group consists of the NKPC models with time variation in inflation levels (NKPC-TV, NKPC-TV-LS) together with BVAR-TV-SV and the SW2007 model. The models in the second group have much superior performance in terms of the predictive likelihood values. A second increase in the predictive likelihood values can be observed when we consider the models in the third group, namely the HNKPC models (HNKPC-TV, HNKPC-TV-LS, HNKPC-TV-LS-SV) and the NKPC model together with discrete level shifts and stochastic volatility for inflation (NKPC-TV-LS-SV).

Table 4.3: Predictive performance of NKPC models and reduced form alternatives

Model	Cumulative (Log) Pred. Likelihood	MSFE 1 period ahead	MSFE 4 period ahead
SW2007	-78.03	0.17	0.25
BVAR-TV-SV	-97.98	0.10	0.25
NKPC-LT	-139.33	0.35	0.36
NKPC-HP	-157.19	0.46	0.37
NKPC-TV	-46.16	0.14	0.26
NKPC-TV-LS	-61.97	0.14	0.28
NKPC-TV-LS-SV	-33.48	0.13	0.21
HNKPC-TV	-36.38	0.12	0.28
HNKPC-TV-LS	-35.05	0.11	0.24
HNKPC-TV-LS-SV	-18.15	0.09	0.18

Note: The table reports the predictive performances of all competing models for the prediction sample over the period 1973-II until 2012-I. 'Cumulative (Log) Pred. Likelihood' stands for the sum of the natural logarithms of predictive likelihoods. 'MSFE' stands for the Mean Squared Forecast Error. Results are based on 10000 simulations of which the first 5000 are discarded for burn-in. 'SW2007' stands for the model proposed by Stock and Watson (2007), and 'BVAR-TV-SV' stands for the Bayesian VAR model with time varying levels and trends and a stochastic volatility component for the inflation equation. Remaining abbreviations are as in Table 4.1.

Figure 4.5: Predictive likelihoods from competing models



Note: The figure displays the evolution of the (log) predictive likelihoods for the computing models over the period 1973-II until 2012-I. Model abbreviations are as in Table 4.1. Results are based on 5000 simulations of which the first 10000 are discarded for burn-in.

A similar clustering of models is observed when we compare model performances using the one period ahead MSFE with the exception of the BVAR-TV-SV model. BVAR-TV-SV model performs considerably better in terms of point prediction compatible with the HNKPC models.

Three main conclusions can be drawn from these findings. First, the conventional NKPC models with demeaned and detrended data (NKPC-LT and NKPC-HP) perform worse than the competing models both in terms of MSFE and in terms of the cumulative predictive likelihood metric. The difference between HNKPC and NKPC models in terms of point forecasts is less pronounced compared to the increase in precision when switching from models using demeaned and detrended data to the models that use the raw data. Hence it is important to estimate levels and trends together with the structural model parameters.

Second, the difference between the NKPC model with level shifts and stochastic volatility and the basic NKPC models is substantial. The former model delivers more accurate point predictions considering MSFE and predictive likelihood values. Thus, it is important to incorporate both high and low frequency movements in structural models. This performance increases further in the HNKPC models, which incorporate the survey data and the backward looking component.

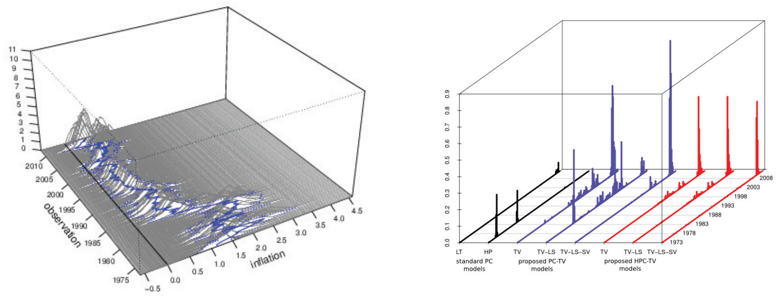
Third, structural models perform at least as well as the strong reduced form candidates, the SW2007 and BVAR-TV-SV models. These findings are crucial since structural models deliver both structural macroeconomic information and predictive performance, whereas the reduced form models are solely designed for improving the predictive performance.

The evolution of the model performance over the forecast sample is shown in Figure 4.5. An important finding from the figure is the increasing predictive performance of the HNKPC models and the models with stochastic volatility components after mid 1980s. This period is characterized by a decrease in inflation volatility during the Great Moderation, which the stochastic volatility component captures accurately. Moreover, the effect of the level shifts can be observed when we compare the NKPC-TV-LS-SV model with the SW2007 model. Much of the difference in the performance of these models can be attributed to the changes in inflation levels. This shows that the inflation process exhibits several regime changes.

The last metric we use for model comparison considers the implied deflationary risk. The left panel in Figure I.1 shows the entire density of the inflation predictions for the HNKPC-LS-SV model where the levels and trends are estimated together with the structural parameters. The mean predicted inflation is represented by the solid line, and the width of the predictive distribution is indicated by the white area under the inflation density. As expected, inflation predictions are concentrated around high (low) values during the high (low) inflationary periods. The uncertainty around the inflation predictions are also high for these periods, together with the periods when inflation is subject to a transition to low values around 1980s. When the observed inflation values are close to the zero bound, the predictive densities indicate deflationary risk.

The right panel in Figure I.1 displays this deflationary risk, which is of key importance especially for policy making. The figure shows that NKPC models with a priori demeaned and detrended data do not signal any pronounced deflation risk except for the low deflation

Figure 4.6: Predicted inflation densities from HNKPC-LS-SV model and deflation probabilities implied by different Phillips curve models



Note: The left figure presents one period ahead predictive distribution of inflation from the HNKPC-LS-SV model, over the period 1973-II until 2012-I. The right figure presents deflation probabilities computed using these predictive distributions of inflation over the same period. Model abbreviations are as in Table 4.1. Results are based on 5000 simulations of which the first 10000 are discarded for burn-in.

probabilities during mid 1970s and mid 1980s. However, extended NKPC and HNKPC models exploiting the high and low frequency movements produce clear signals of deflation risk and deflationary pressure during the recent recession.

Note that actual deflation only occurs around 2009 in this sample period and the models signal deflationary risk slightly later than this period. This result can be explained by the agents' learning process. As indicated in Schorfheide (2005), if agents learn about the monetary policy changes later than the inflation level changes, the perceived target inflation in general equilibrium happens only gradually. In Schorfheide (2005), this is incorporated as Bayesian learning of the agents which is in line with the econometric assumption underlying our models. As the modelled state-space updating incorporates Bayesian learning, changes in the inflation level occur gradually and deflationary risk signals are delayed. Our models are still able to capture this deflationary pressure successfully.

We conclude this section with two remarks. First, the models we considered so far rely on the implicit assumption of the absence of a long-run cointegrating relationship. We assess whether this assumption is plausible for the U.S. data considering the HPC-

TV-LS-SV model, and find credible evidence that the existence of such a cointegrating relationship is very unlikely. Second, the proposed models extend the standard NKPC model in several ways. However, the superiority of the most extensive model, HNKPC-TV-LS-SV, is based on all proposed model extensions jointly. Details on these results are provided in appendix 4.M and appendix 4.K.

4.4 Conclusion

NKPC models constitute an integral part of macroeconomic models used for forecasting and policy analysis. These models are often estimated after demeaning and/or detrending the data. In this chapter it is shown that mechanical removal of the low frequency movements in the data may lead to poor forecasts. Potential structural breaks and level shifts as well as changing volatility in the observed series require more complex models, which can handle these time variation together with the standard NKPC parameters. We have proposed a set of models where levels and trends of the series together with the volatility process are integrated with a structural NKPC model. Furthermore, we consider richer expectational mechanisms for the inflation series in enlarged Hybrid-NKPC models using survey data for inflation expectations.

The proposed models capture time variation in the low frequency movements of both inflation and marginal cost data. For the inflation series we identify three distinct periods with high and low inflation. The high inflationary period corresponds to 1970s, following a low inflationary period of 1960s. The last period starting with 1980s is characterized by low inflation levels corresponding to an annual inflation level around 2%. When this model is blended with the stochastic volatility component, the level shifts can be identified even more precisely.

The use of macroeconomic information in the structural models together with the remaining high and low frequency movements in the data improves the predictive ability also compared to celebrated reduced form models, including the Bayesian VAR and the stochastic volatility model, see Stock and Watson (2007). Furthermore, modeling inflation expectations using survey data and adding stochastic volatility to the NKPC model structure improves in sample fit and out of sample predictive performance substantially. We also analyze deflation probabilities indicated by each competing model. The complete predictive densities, most notably from the enlarged models, indicate an increase in the probability of deflation in the U.S. in recent years.

Modeling forward and backward looking components of inflation has important effects on empirical results. Endogeneity and persistence do not appear to be very important empirical issues in NKPC model structures. Finally, we analyze the existence of a long-run relation between the low frequency movements of both series. No credible evidence is found on such a long run stable cointegrating relation for the U.S. series.

Given that incorporating low and high frequency movements explicitly in macroeconomic models provides additional insights for both policy analysis and more accurate predictions, we plan to enlarge the proposed model to a more general DSGE framework in future work. Another interesting possibility of future research is to combine different NKPC models using their predictive performances, which seem to be time varying.

Appendices

In section 4.A we present the effects of misspecified levels on posterior results in a standard New Keynesian Phillips Curve (NKPC). This analysis provides a straightforward motivation for the extended NKPC and HNKPC models in the chapter. Specifically, we show that a priori demeaning and detrending of the data, without considering the short and long-run data properties obscure inference in these standard models.

In section 4.B we elaborate and compare the inference of the NKPC model using structural form and unrestricted reduced form. This section illustrates the difficulty of inferring the unrestricted reduced form parameters and to obtain the main parameters of interest, the structural parameters, using these. This difficulty is based on the non-linear parameter transformations required to link the structural and the reduced form models. Through simulation examples, we show that flat prior distributions used in one of the model representations can be very informative in the other model representation. This section motivates the structural parameter estimation approach we follow throughout the chapter. In section 4.C, we solve for the expectation term in the NKPC model and derive the restricted reduced form of the model.

Sections 4.E and 4.F provide the details of the posterior sampling algorithms for the extended NKPC and HNKPC models proposed in the main text of the chapter. In these sections, the state space representations of the extended models and the appropriate sampling scheme are explained in detail. We further report the exact prior parameters used for the results in the chapter and present a sketch of a prior sensitivity analysis based on prior-predictive likelihood comparisons.

Sections 4.G, 4.H and 4.I provide posterior and predictive results for the extended NKPC and HNKPC models which are not included in the chapter due to space constraints. In section 4.G we present additional posterior and predictive results for the extended NKPC models. Main conclusions from these models are similar to the extended HNKPC model results discussed in the chapter. Nonetheless, we provide these results for clarity and the ease of comparison. In section 4.H we present additional results for the HNKPC models which are in line with the main conclusions of the chapter. Section 4.I presents the entire distribution of the inflation predictions for extended NKPC and HNKPC models we propose.

Section 4.J presents the results of the prior-predictive likelihood analysis for the proposed models. The main conclusion of this section is that the adopted priors in the chapter do not dominate the results. The data information is the main factor favoring the extended models we propose.

Section 4.K presents the posterior and predictive results of the alternative NKPC and HNKPC models, considered for robustness checks, in detail. Several alternative models are compared with the extended models in the chapter. We show that our main conclusions on the improved model performance through modeling the trends and levels in the data, and the use of survey data hold. We further disentangle the predictive gains from these two sources of extensions.

In section 4.L we present a further alternative HNKPC model to the proposed HNKPC models in this chapter. This model aims at accounting for the possibility of measurement errors in survey expectations. The results obtained from this alternative model are very similar to the corresponding results in the chapter, thus, we conclude that the effect of the measurement errors in survey expectations is negligible.

Section 4.M presents a straightforward cointegration analysis for inflation and marginal cost series, based on the time-varying NKPC model structure. This analysis is performed to justify an implicit assumption in the proposed models namely the assumption that there is no stable long-run relationship between the inflation and marginal cost series. The results of this cointegration analysis are in line with the implicit assumption we make in the proposed NKPC model structures.

4.A Effect of Misspecified Level Shifts on Posterior Estimates of Inflation Persistence

The linear NKPC captures the relation between real marginal cost $\tilde{z}_t$ and inflation $\tilde{\pi}_t$. We illustrate in this section that model misspecification resulting from ignoring level shifts in inflation data leads to overestimation of persistence in the inflation equation within a linear NKPC.

The linear NKPC model can be written as

$$\begin{aligned} \tilde{\pi}_t &= \lambda \tilde{z}_t + \gamma_b \tilde{\pi}_{t-1} + \epsilon_{1,t}, \\ \tilde{z}_t &= \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \epsilon_{2,t}, \end{aligned} \tag{4.A.1}$$

with $(\epsilon_{1,t}, \epsilon_{2,t})' \sim NID(0, \Sigma)$. This model is a triangular simultaneous equations model and can also be interpreted as an instrumental variable model with two instruments. We specify an AR(2) model for the marginal costs in order to mimic for the cyclical behavior of the observed series, see Basistha and Nelson (2007); Kleibergen and Mavroeidis (2011) for a similar specification. The AR(2) parameters are restricted to the stationary region $|\phi_1| + \phi_2 < 1$, $|\phi_2| < 1$, and the lagged adjustment parameter in the inflation equation is restricted as $0 \leq \gamma_b < 1$. The structural parameter λ, the slope of the Phillips curve, is restricted as $0 \leq \lambda < 1$ which is in line with previous evidence on the slope of the NKPC.

Since NKPC in (4.A.1) specifies the relation between the *short-run* stationary fluctuations in real marginal costs and inflation, $\tilde{\pi}_t$ and $\tilde{z}_t$ can be interpreted as the *transitory* components of inflation and marginal costs, in deviation from their long-run components. In fact, the observed non-filtered data can be decomposed into permanent and transitory components in a straightforward way as

$$\begin{aligned} \pi_t &= \tilde{\pi}_t + c_{\pi,t}, \\ z_t &= \tilde{z}_t + c_{z,t}, \end{aligned} \tag{4.A.2}$$

where π_t and z_t are the inflation and marginal cost data, respectively, and $c_{\pi,t}$ and $c_{z,t}$ are the permanent components of the series.

In our simulation experiment, we model the steady state inflation as a constant level subject to regime shifts in order to mimic the high inflationary period during the 1970s. For modelling the permanent component of the real marginal cost series, we use a trend specification mimicking the declining real marginal cost levels in the U.S. over the sample starting from the 1960s. This specification can be formulated as follows

$$\begin{aligned} c_{\pi,t} &= c_{\pi,t-1} + \kappa_t \eta_{t-1}, \;\; c_{z,t} = c_{z,t-1} + \mu_{z,t-1}, \\ \mu_{z,t} &= \mu_{z,t-1}, \;\; \eta_t \sim NID(0, \omega^2), \end{aligned} \tag{4.A.3}$$

where κ_t is a binary variable indicating a level shift in the level series, $c_{\pi,t}$ and $c_{z,t}$ indicate the level value of inflation and real marginal costs, respectively, in period t and $\mu_{z,t}$ is the slope of the trend in the real marginal cost series. By excluding the stochastic component for the slope and the trend of the real marginal costs in (4.A.3), we specify a deterministic trend for this series.

We simulate three sets of data from the model in (4.A.1)–(4.A.3). For the first set, the inflation series show no level shifts, i.e. $\kappa_t = 0, \;\; \forall t$. For the other two sets of data, we impose different level shifts with moderate ($\omega^2 = 2.5$) and large ($\omega^2 = 5$) changes in the level values, respectively. For each specification we simulate 100 datasets with $T = 200$ observations, where two level shifts occur in periods $t = 50$ and $t = 150$. The observation error variance is set to $\left(\begin{smallmatrix} 1 & 0.01 \\ 0.01 & 0.01 \end{smallmatrix}\right)$, which leads to a correlation of 0.1 between the disturbances, and parameter λ is set to 0.1. Note that parameters $\phi_1 = 0.1$ and $\phi_2 = 0.5$ are chosen such that the transitory component of the series is stationary.

In order to capture the effect of model misspecification on posterior inference, when computing the transitory component, we ignore level shifts in the simulated inflation series and simply demean the series. For the marginal cost series, we remove the linear trend prior to the analysis and only focus on the effect of misspecification in the inflation series. This implies that for the simulated data with no level shifts, the model is correctly specified and the posterior results should be close to the true values. For each simulated data set we estimate the model in (4.A.1) using flat priors on restricted parameter regions:

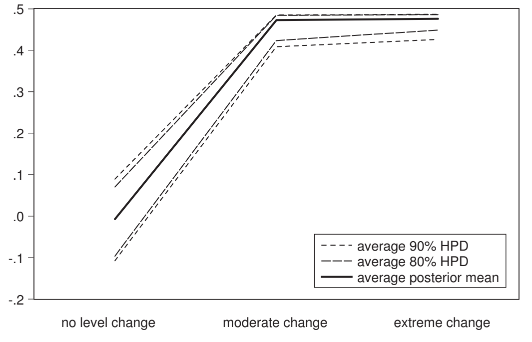
$$p(\phi_1, \phi_2, \gamma_b, \lambda) \propto \begin{cases} 1, & \text{if } |\phi_1| + \phi_2 < 1,\ |\phi_2| < 1,\ 0 \leq \gamma_b < 1,\ 0 \leq \lambda < 1 \\ 0, & \text{otherwise} \end{cases}. \quad (4.A.4)$$

Given that model (4.A.1) is equivalent to an instrumental variables model with 2 instruments, it can be shown that the likelihood function for such a model combined with the flat prior on a large space yields a posterior distribution that exists but it has no first or higher moments. Due to the bounded region condition on the parameters, where the structural parameter λ is restricted to the unit interval, all moments exist. For details, we refer to Zellner, Ando, Baştürk, Hoogerheide and Van Dijk, (2014). We mention this existence result since it provides an econometric explanation why it is often difficult to estimate a structural model for macro-economic data such as (4.A.1). Indeed, the rather flat posterior surface plagues the inference, in particular, when ϕ_2 is close to zero. Posterior moments are in our case computed by means of standard Metropolis-Hastings method on ϕ_1 and ϕ_2 and λ and γ_b. Other Monte Carlo methods like Gibbs sampling are also feasible in this case.

Figure A.1 presents the overestimation results from 100 different simulations for each setting we consider. We report the average overestimation in posterior γ_b estimates and 95% highest posterior density intervals (HPDI) for this overestimation.

The persistence parameter γ_b is overestimated in all cases except for the correctly specified model. The degree of overestimation becomes larger with a larger shift in the level of inflation. Note that the average 95% HPDI of overestimation becomes tighter for data with extreme changes in levels. Hence the effect of model misspecification on the persistence estimates is more pronounced if the regime shifts are extreme.

Figure A.1: Overestimation illustration for the backward looking NKPC model



Note: The figure presents overestimation probability of parameter γ_b for simulated data from the NKPC model with different structural breaks structures. We report average quantiles of overestimation based on 100 simulation replications for each parameter setting.

In summary, our simulation experiments using NKPC show that when the shifts in the inflation level are not modelled, inference on model persistence parameters may be severely biased due to the model misspecification. This will also hold for predictive estimates.

We note that we focused on misspecification effects on persistence measures when level shifts in the series are ignored. Similar experiments can be set up for the NKPC with weak identification (or weak instruments) by setting $\phi_2 \approx 0$. The effect of misspecification on posterior and predictive estimates in the case of weak identification is a topic outside the scope of the present chapter. We refer to Kleibergen and Mavroeidis (2011) for details on Bayesian estimation in case of weak identification.

4.B Structural and Reduced Form Inference of the NKPC Model

This section presents the unrestricted reduced form inference (URF) of the NKPC model, and the inference of the corresponding structural form (SF) model parameters. The structural form (SF) representation for the basic NKPC model derived from the firm's price setting for filtered data is given as

$$\begin{aligned} \tilde{\pi}_t &= \lambda \tilde{z}_t + \gamma_f E_t(\tilde{\pi}_{t+1}) + \epsilon_{1,t}, \\ \tilde{z}_t &= \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \epsilon_{2,t}, \end{aligned} \tag{4.B.1}$$

where $(\epsilon_{1,t}, \epsilon_{2,t})' \sim NID\,(0, \Sigma)$ and standard stationary restrictions hold for ϕ_1, ϕ_2.

We show that the posterior draws from the structural form parameters can be obtained using the reduced form representation of (4.B.1):

$$\begin{aligned} \tilde{\pi}_t &= \alpha_1 \tilde{z}_{t-1} + \alpha_2 \tilde{z}_{t-2} + \epsilon_{1,t}, \\ \tilde{z}_t &= \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \epsilon_{2,t}, \end{aligned} \tag{4.B.2}$$

where $(\epsilon_{1,t}, \epsilon_{2,t})' \sim NID\,(0, \Sigma)$, and the restricted reduced form (RRF) representation is obtained by introducing the following restrictions on parameters in (4.B.1):

$$\alpha_1 = \frac{\lambda(\phi_1 + \gamma_f \phi_2)}{1 - \gamma_f(\phi_1 + \gamma_f \phi_2)}, \;\; \alpha_2 = \frac{\lambda \phi_2}{1 - \gamma_f(\phi_1 + \gamma_f \phi_2)}. \tag{4.B.3}$$

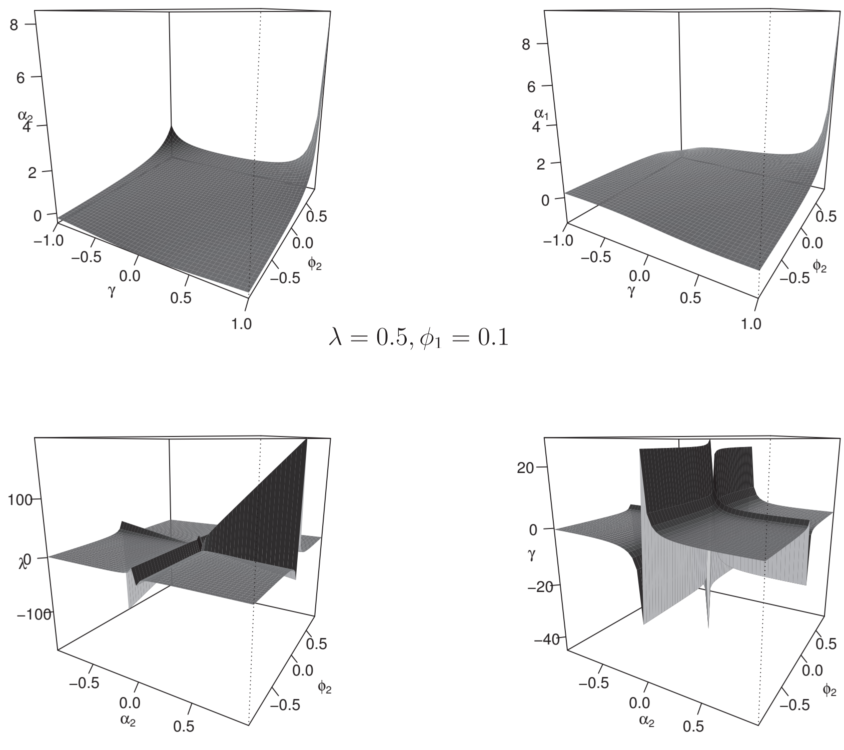
Finally, the model in (4.B.1) is related to an Instrumental Variables (IV) model with exact identification. Bayesian estimation of the unrestricted reduced form model in (4.B.2) is straightforward under flat or conjugate priors. Given the posterior draws of reduced form parameters, posterior draws of structural form parameters in (4.B.1) can be obtained using the transformation in (4.B.3). This nonlinear transformation, however, causes difficulties in setting the priors in an adequate way. The determinant of the Jacobian of this nonlinear transformation is $\mid J \mid = \frac{\lambda \phi_2^2}{\left(1 - \gamma_f(\phi_1 + \gamma_f \phi_2)\right)^2}$, where the Jacobian is non-zero and finite if $\gamma_f(\phi_1 + \gamma_f \phi_2) \neq 1$, $\phi_2 \neq 0$ and $\lambda \neq 0$.[7]

Figure B.1 illustrates the nonlinear transformation for the SF and RRF representations, for a grid of parameter values from SF representations, and plot the corresponding RRF parameter values, and vice versa. The top panel in Figure B.1 shows the trans-

[7] We only consider the transformation from $\{\lambda, \gamma_f, \phi_1, \phi_2\}$ to $\{\alpha_1, \alpha_2, \phi_1, \phi_2\}$, i.e. variance parameters in the transformed model are left as free parameters.

formations from SF to RRF. Reduced form parameters α_1 and α_2 tend to infinity when persistence in inflation and marginal cost series are high, i.e. when the structural form parameters λ and $\phi_1 + \phi_2$ tend to 1. The bottom panel in Figure B.1 shows the RRF to SF transformations. The corresponding SF parameters lead to an irregular shape, for example, when the instrument z_{t-2} has no explanatory power with $\phi_2 = 0$ or when $\alpha_2 = 0$.

Figure B.1: Nonlinear parameter transformations



Note: The top panel presents the implied unrestricted reduced form parameters in (4.B.2) given structural form parameters in (4.B.1). The bottom panel presents implied structural form parameters in (4.B.1) given unrestricted reduced form parameters in (4.B.2). Parameter transformations are obtained using the RRF restrictions in (4.B.3).

4.C Structural Form and Restricted Reduced Form Derivations for the NKPC and HNKPC Models

4.C.1 Structural Form Derivations for the NKPC Model

In this section we derive the structural form equations for the NKPC model.Consider the NKPC model in deviation from (possibly time-varying) levels:

$$\tilde{\pi}_t = \lambda \tilde{z}_t + \gamma_f E_t(\tilde{\pi}_{t+1}) + \epsilon_{1,t}, \tag{4.C.1}$$

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \epsilon_{2,t}, \tag{4.C.2}$$

where $(\epsilon_{1,t}, \epsilon_{2,t})' \sim NID(0, \Sigma)$ and the unobserved variables $\tilde{\pi}_t$ and $\tilde{z}_t$ are the inflation and (log) real marginal cost in deviation from levels at time t, respectively. For stationarity in real marginal cost (in deviation from levels) standard stationarity restrictions should hold for ϕ_1, ϕ_2. Note that in (4.C.1), to clarify the expectation assumption in this model, the term $E(\tilde{\pi}_{t+1})$ is explicitly written as the conditional expectation $E_t(\tilde{\pi}_{t+1}) = E(\tilde{\pi}_{t+1}|I_t)$ where I_t denotes the information available at time t.

In order to obtain the structural form representation of this model used in the estimation, the expectation term $E_t(\tilde{\pi}_{t+1})$ in (4.C.1) has to be solved[8]. The solution of the expectation mechanism in (4.C.1) is obtained by induction.

First, we replace the expectation term in (4.C.1) using the same equation, iterated by one period:

$$\tilde{\pi}_t = \lambda \tilde{z}_t + \gamma_f E_t(\lambda \tilde{z}_{t+1} + \gamma_f E_{t+1}(\tilde{\pi}_{t+2}) + \epsilon_{1,t+1}) + \epsilon_{1,t}, \tag{4.C.3}$$

$$= \lambda \tilde{z}_t + \lambda \gamma_f E_t(\tilde{z}_{t+1}) + \gamma_f^2 E_t(\tilde{\pi}_{t+2}) + \epsilon_{1,t}, \tag{4.C.4}$$

[8]Alternatively, one can replace the inflation expectations with survey data directly. We do not opt for this solution but consider alternative models with this possibility.

where the second equality follows from the law of iterated expectations and the assumption that $E_t(\epsilon_{1,t+1}) = 0$.

Iterating (4.C.1) for k periods we obtain:

$$\tilde{\pi}_t = \lambda \tilde{z}_t + \lambda \sum_{j=1}^{k} \gamma_f^j E_t(\tilde{z}_{t+j}) + \gamma_f^{k+1} E_t(\tilde{\pi}_{t+k+1}) + \epsilon_{1,t}. \tag{4.C.5}$$

The limiting equation from this iteration is:

$$\begin{aligned}\tilde{\pi}_t &= \lambda \tilde{z}_t + \lambda \sum_{k=1}^{\infty} \gamma_f^k E_t(\tilde{z}_{t+k}) + \lim_{k\to\infty} \gamma_f^{k+1} E_t(\tilde{\pi}_{t+k+1}) + \epsilon_{1,t} \\ &= \lambda \tilde{z}_t + \lambda \gamma_f \sum_{k=0}^{\infty} \gamma_f^k E_t(\tilde{z}_{t+k+1}) + \epsilon_{1,t},\end{aligned} \tag{4.C.6}$$

where the last equality holds under the assumption that $|\gamma_f| < 1$.

The iterative solution to the inflation expectations in (4.C.6) depends on the infinite sum of marginal costs in (4.C.6). In order to solve for this infinite sum, we use the marginal cost equation (4.C.2) of the NKPC model. Rearranging (4.C.6) and inserting (4.C.2) in (4.C.6) we obtain:

$$
\begin{aligned}
\tilde{\pi}_t - \epsilon_{1,t} &= \lambda \tilde{z}_t + \lambda\gamma_f \sum_{k=0}^{\infty} \gamma_f^k E_t\left(\phi_1 \tilde{z}_{t+k} + \phi_2 \tilde{z}_{t+k-1} + \epsilon_{2,t+k+1}\right), \\
&= \lambda \tilde{z}_t + \phi_1\lambda\gamma_f \sum_{k=0}^{\infty} \gamma_f^k E_t\left(\tilde{z}_{t+k}\right) + \phi_2\lambda\gamma_f \sum_{k=0}^{\infty} \gamma_f^k E_t\left(\tilde{z}_{t+k-1}\right), \\
&= \lambda \tilde{z}_t + \phi_2\lambda\gamma_f \tilde{z}_{t-1} + \phi_1\lambda\gamma_f \sum_{k=0}^{\infty} \gamma_f^k E_t\left(\tilde{z}_{t+k}\right) + \phi_2\lambda\gamma_f \sum_{k=1}^{\infty} \gamma_f^k E_t\left(\tilde{z}_{t+k-1}\right), \\
&= \lambda \tilde{z}_t + \phi_2\lambda\gamma_f \tilde{z}_{t-1} + \phi_1\lambda\gamma_f \sum_{k=0}^{\infty} \gamma_f^k E_t\left(\tilde{z}_{t+k}\right) + \phi_2\lambda\gamma_f^2 \sum_{k=0}^{\infty} \gamma_f^k E_t\left(\tilde{z}_{t+k}\right), \\
&= \lambda \tilde{z}_t + \phi_2\lambda\gamma_f \tilde{z}_{t-1} + \lambda\left(\phi_1\gamma_f + \phi_2\gamma_f^2\right) \sum_{k=0}^{\infty} \gamma_f^k E_t\left(\tilde{z}_{t+k}\right), \\
&= \lambda \tilde{z}_t + \phi_2\lambda\gamma_f \tilde{z}_{t-1} + \lambda\left(\phi_1\gamma_f + \phi_2\gamma_f^2\right) \left(E_t\left(\tilde{z}_t\right) + \sum_{k=1}^{\infty} \gamma_f^k E_t\left(\tilde{z}_{t+k}\right)\right), \\
&= \lambda \tilde{z}_t + \phi_2\lambda\gamma_f \tilde{z}_{t-1} + \lambda\left(\phi_1\gamma_f + \phi_2\gamma_f^2\right) \left(E\left(\tilde{z}_t | \tilde{z}_1, ..., \tilde{z}_t\right) + \sum_{k=1}^{\infty} \gamma_f^k E_t\left(\tilde{z}_{t+k}\right)\right), \\
&= \lambda \tilde{z}_t + \phi_2\lambda\gamma_f \tilde{z}_{t-1} + \lambda\left(\phi_1\gamma_f + \phi_2\gamma_f^2\right) \left(\tilde{z}_t + \gamma_f \sum_{k=0}^{\infty} \gamma_f^k E_t\left(\tilde{z}_{t+k+1}\right)\right), \\
&= \lambda \tilde{z}_t + \phi_2\lambda\gamma_f \tilde{z}_{t-1} + \left(\phi_1\gamma_f + \phi_2\gamma_f^2\right)\left(\tilde{\pi}_t - \epsilon_{1,t}\right),
\end{aligned}
$$

where the last equality follows from (4.C.6).

Hence the solution for the infinite sum of inflation expectations in (4.C.1) leads to the following structural inflation equation:

$$
\tilde{\pi}_t = \frac{\lambda}{1 - \left(\phi_1 + \phi_2\gamma_f\right)\gamma_f} \tilde{z}_t + \frac{\phi_2\lambda\gamma_f}{1 - \left(\phi_1 + \phi_2\gamma_f\right)\gamma_f} \tilde{z}_{t-1} + \epsilon_{1,t}, \tag{4.C.7}
$$

where, together with the explicit level specifications of the inflation and labor share series, the extended NKPC model takes the form of equation (4.E.3).

The restricted reduced form of the model can be derived by replacing $\tilde{z}_t$ in (4.C.7) using (4.C.2):

$$\tilde{\pi}_t = \frac{\lambda\phi_1 + \phi_2\lambda\gamma_f}{1-(\phi_1+\phi_2\gamma_f)\,\gamma_f}\tilde{z}_{t-1} + \frac{\lambda\phi_2}{1-(\phi_1+\phi_2\gamma_f)\,\gamma_f}\tilde{z}_{t-2} + \eta_t, \tag{4.C.8}$$

where $\eta_t = \lambda/(1-(\phi_1+\phi_2\gamma_f)\,\gamma_f)\epsilon_{2,t} + \epsilon_{1,t}$.

4.C.2 Structural Form Derivations for the HNKPC Model

In this section we derive the structural form equations of the HNKPC model. Consider the HNKPC model in deviation from (possibly time varying) levels:

$$\tilde{\pi}_t = \lambda^H \tilde{z}_t + \gamma_f^H E_t(\tilde{\pi}_{t+1}) + \gamma_b^H \tilde{\pi}_{t-1} + \epsilon_{1,t}, \tag{4.C.9}$$

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \epsilon_{2,t}. \tag{4.C.10}$$

where, similar to the NKPC model, the term $E\left(\tilde{\pi}_{t+1}\right)$ is explicitly written as the conditional expectation $E_t\left(\tilde{\pi}_{t+1}\right) = E\left(\tilde{\pi}_{t+1}|I_t\right)$ where I_t denotes the information available at time t.

Similar to the NKPC model, we replace the expectation term in (4.C.9) using the same equation, iterated 1 period ahead:

$$\tilde{\pi}_t = \lambda^H \tilde{z}_t + \gamma_f^H E_t\left(\lambda^H \tilde{z}_{t+1} + \gamma_f^H E_{t+1}\left(\tilde{\pi}_{t+2}\right) + \gamma_b^H \tilde{\pi}_t + \epsilon_{1,t+1}\right) + \gamma_b^H \tilde{\pi}_{t-1} + \epsilon_{1,t}, \tag{4.C.11}$$

$$= \lambda^H \tilde{z}_t + \gamma_f^H \lambda^H E_t\left(\tilde{z}_{t+1}\right) + (\gamma_f^H)^2 E_t\left(\tilde{\pi}_{t+2}\right) + \gamma_f^H \gamma_b^H E_t\left(\tilde{\pi}_t\right) + \gamma_b^H \tilde{\pi}_{t-1} + \epsilon_{1,t}, \tag{4.C.12}$$

where we again use the law of iterated expectations and the assumption on residuals: $E_t(\epsilon_{1,t+1})$=0.

Iterating (4.C.9) k periods ahead we obtain:

$$\tilde{\pi}_t = \lambda^H \tilde{z}_t + \lambda^H \sum_{j=1}^{k} \left(\gamma_f^H\right)^j E_t\left(\tilde{z}_{t+j}\right) + \left(\gamma_f^H\right)^{k+1} E_t\left(\tilde{\pi}_{t+k+1}\right) + \gamma_b^H \sum_{j=2}^{\infty} \left(\gamma_f^H\right)^j \tilde{\pi}_{t+j-1}$$
$$+ \gamma_b^H \tilde{\pi}_{t-1} + \gamma_b^H \gamma_f^H \tilde{\pi}_t + \epsilon_{1,t}. \quad (4.C.13)$$

Assuming $|\gamma_f^H| < 1$ and using the property $E_t\left(\tilde{z}_t\right) = \tilde{z}_t$ the limiting equation from this iterative solution is:

$$\begin{aligned} \tilde{\pi}_t &= \lambda^H \sum_{k=0}^{\infty} \left(\gamma_f^H\right)^k E_t\left(\tilde{z}_{t+k}\right) + \gamma_b^H \tilde{\pi}_{t-1} + \gamma_b^H \gamma_f^H \tilde{\pi}_t + \gamma_b^H \sum_{k=2}^{\infty} \left(\gamma_f^H\right)^k \tilde{\pi}_{t+k-1} + \epsilon_{1,t} \\ &= \frac{\lambda^H}{1-\gamma_b^H\gamma_f^H} \sum_{k=0}^{\infty} \left(\gamma_f^H\right)^k E_t\left(\tilde{z}_{t+k}\right) + \frac{\gamma_b^H}{1-\gamma_b^H\gamma_f^H}\tilde{\pi}_{t-1} + \frac{\gamma_b^H\gamma_f^H}{1-\gamma_b^H\gamma_f^H} \sum_{k=1}^{\infty} \left(\gamma_f^H\right)^k E_t\left(\tilde{\pi}_{t+k}\right) \\ &\quad + \frac{1}{1-\gamma_b^H\gamma_f^H}\epsilon_{1,t}. \end{aligned} \quad (4.C.14)$$

Similar to the NKPC model, the infinite sum of marginal cost expectations in (4.C.14) is solved using the marginal cost equation of the model in (4.C.10):

$$\begin{aligned} \sum_{k=0}^{\infty} \left(\gamma_f^H\right)^k E_t\left(\tilde{z}_{t+k}\right) &= \tilde{z}_t + \phi_1 \sum_{k=1}^{\infty} \left(\gamma_f^H\right)^k E_t\left(\tilde{z}_{t+k-1}\right) + \phi_2 \sum_{k=1}^{\infty} \left(\gamma_f^H\right)^k E_t\left(\tilde{z}_{t+k-2}\right) \\ &\quad + \sum_{k=1}^{\infty} \left(\gamma_f^H\right)^k E_t\left(\epsilon_{t+k}\right) \\ &= \tilde{z}_t + \gamma_f^H \phi_1 \tilde{z}_t + \gamma_f^H \phi_2 \tilde{z}_{t-1} + \left(\gamma_f^H\right)^2 \phi_2 \tilde{z}_t + \phi_1 \sum_{k=2}^{\infty} \left(\gamma_f^H\right)^k E_t\left(\tilde{z}_{t+k-1}\right) \\ &\quad + \phi_2 \sum_{k=3}^{\infty} \left(\gamma_f^H\right)^k E_t\left(\tilde{z}_{t+k-2}\right) \\ &= \tilde{z}_t + \gamma_f^H \phi_2 \tilde{z}_{t-1} + \left(\phi_1 \gamma_f^H + \phi_2 \left(\gamma_f^H\right)^2\right) \sum_{k=0}^{\infty} \left(\gamma_f^H\right)^k E_t\left(\tilde{z}_{t+k}\right) \\ &= \frac{1}{1-\left(\theta_1+\theta_2\gamma_f^H\right)\gamma_f^H}\tilde{z}_t + \frac{\gamma_f^H\phi_2}{1-\left(\theta_1+\theta_2\gamma_f^H\right)\gamma_f^H}\tilde{z}_{t-1}. \end{aligned} \quad (4.C.15)$$

Inserting (4.C.15) in (4.C.14), we obtain the structural form of the inflation equation in the HNKPC model:

$$\begin{aligned}\tilde{\pi}_t = & \frac{\lambda^H}{\left(1-\gamma_b^H\gamma_f^H\right)\left(1-\left(\phi_1\gamma_f^H+\phi_2\left(\gamma_f^H\right)^2\right)\right)}\tilde{z}_t + \\ & \frac{\lambda^H\gamma_f^H\phi_2}{\left(1-\gamma_b^H\gamma_f^H\right)\left(1-\left(\phi_1\gamma_f^H+\phi_2\left(\gamma_f^H\right)^2\right)\right)}\tilde{z}_{t-1} + \frac{\gamma_b^H}{1-\gamma_b^H\gamma_f^H}\tilde{\pi}_{t-1} + \\ & \frac{\gamma_b^H\gamma_f^H}{1-\gamma_b^H\gamma_f^H}\sum_{k=1}^{\infty}\left(\gamma_f^H\right)^k\tilde{E}_t\left(\tilde{\pi}_{t+k}\right) + \frac{1}{1-\gamma_b^H\gamma_f^H}\epsilon_{1,t},\end{aligned} \tag{4.C.16}$$

i.e. we obtain (4.9) as the structural form of the HNKPC model. Unlike the NKPC model, this solution still includes the infinite sum of inflation expectations. A closed form solution for the latter expression only exists under certain assumptions such as rational expectations. We refrain from these assumptions but instead model inflation expectations in this model using unobserved components, as outlined in section 4.L.

4.D Derivations for the NKPC Equations Using the Calvo Formulation

In this section, we present the derivations of the standard NKPC model, i.e. the relation between inflation and log real marginal cost in deviation from their respective steady state levels, using the staggered price setting in the Calvo formulation (Calvo, 1983), as outlined in Gali (2008). In this staggered price setting, at each time period t, $(1-\theta)$ fraction of firms can reoptimize their prices while θ fraction of firms cannot change the prices, with $\theta \in [0,1]$. Firms which can optimize their prices have identical cost functions and they set their optimal prices to maximize the expected sum of discounted real profits subject to the demand function:

$$\max_{P_t^\star} \quad E_t \left\{ \sum_{j=0}^{\infty} \theta^j \gamma_f^j \left(\frac{P_t^\star}{P_{t+j}} Y_{i,t+j} - TC_{t+j} \left(Y_{i,t+j} \right) \right) \right\} \tag{4.D.1}$$

$$\text{s.t.} \quad Y_{i,t+j} = \left(\frac{P_t^\star}{P_{t+j}} \right)^{-\epsilon} Y_{t+j} \tag{4.D.2}$$

where the constraint in the optimization stems from the demand curve and $\epsilon \in (0,1]$ is the elasticity of substitution between goods for the households. Furthermore, $\gamma_f \in [0,1)$ is the discount factor, P_{t+j} and Y_{t+j} denote the aggregate price and aggregate output at time $t+j$, respectively, and $Y_{i,t+j}$ and $TC_{t+j}(Y_{i,t+j})$ denote the output and the total cost (as a function of the output) of firm i at time $t+j$, respectively.

Inserting (4.D.2) in (4.D.1), the constrained optimization is equivalent to the following maximization:

$$\max_{P_t^\star} \quad E_t \left\{ \sum_{j=0}^{\infty} \theta^j \gamma_f^j \left(\left(\frac{P_t^\star}{P_{t+j}} \right)^{1-\epsilon} Y_{t+j} - TC_{t+j} \left(\left(\frac{P_t^\star}{P_{t+j}} \right)^{-\epsilon} Y_{t+j} \right) \right) \right\}, \tag{4.D.3}$$

where the firms optimizing prices take the aggregate output and price levels as given.

The first order condition for the maximization in (4.D.3) is:

$$P_t^\star = \frac{\epsilon}{1-\epsilon} \frac{E_t \left\{ \sum_{j=0}^{\infty} \theta^j \gamma_f^j \left(\frac{1}{P_{t+j}} \right)^{-\epsilon} Y_{t+j} z_{t+j} \left(\left(\frac{P_t^\star}{P_{t+j}} \right)^{-\epsilon} Y_{t+j} \right) \right\}}{E_t \left\{ \sum_{j=0}^{\infty} \theta^j \gamma_f^j \left(\frac{1}{P_{t+j}} \right)^{1-\epsilon} Y_{t+j} \right\}}, \tag{4.D.4}$$

where $z_{t+j}(.)$ is the marginal cost level of the firm i at time $t+j$.

Since the NKPC model explains the relation between inflation and real marginal cost levels, the next step is to rewrite (4.D.4) in terms of the ratios of prices rather than price levels. Dividing both sides of (4.D.4) by P_t we obtain:

$$\frac{P_t^\star}{P_t} = \frac{\epsilon}{1-\epsilon} \frac{E_t \left\{ \sum_{j=0}^{\infty} \theta^j \gamma_f^j \left(\pi_{t,t+j}\right)^{\epsilon} Y_{t+j} z_{t+j} \right\}}{E_t \left\{ \sum_{j=0}^{\infty} \theta^j \gamma_f^j \left(\pi_{t,t+j}\right)^{\epsilon-1} Y_{t+j} \right\}}, \tag{4.D.5}$$

where $\pi_{t,t+j} = P_{t+j}/P_t = \prod_{k=1}^{j} \pi_{t+k-1,t+k}$ denotes the ratio of prices between periods $t+j$ and t, and we simplify the notation for marginal costs, z_{t+j}, by removing its dependence on variables.

The next step in deriving the NKPC model is to write (4.D.5) in deviations from the steady state levels. At the steady state output, marginal cost and (one period ahead) inflation are assumed to be constant and these steady state values are denoted by $\bar{Y}$, $\bar{z}$ and $\bar{\pi}$, respectively. We further assume zero steady state inflation, i.e. $\bar{\pi} = 1$. From (4.D.5), the relation between the steady state variables are derived as follows:

$$\frac{\bar{P}_\star}{\bar{P}} = \frac{\epsilon}{1-\epsilon} \frac{E_t \left\{ \sum_{j=0}^{\infty} \theta^j \gamma_f^j \bar{\pi}^{j\epsilon} \bar{Y} \bar{z} \right\}}{E_t \left\{ \sum_{j=0}^{\infty} \theta^j \gamma_f^j \bar{\pi}^{j(\epsilon-1)} \bar{Y} \right\}} = \frac{\epsilon}{1-\epsilon} \bar{z}, \tag{4.D.6}$$

where we use the steady state relation for inflation: $\bar{\pi}_{t,t+j} = \prod_{k=1}^{j} \pi_{t+k-1,t+k} = \prod_{k=1}^{j} \bar{\pi} = \bar{\pi}^j$.

Deviations of variables from the respective steady states are denoted by $\tilde{Y}$, $\tilde{z}$ and $\tilde{\pi}$.[9] For this purpose, we write all individual components in (4.D.5) using the transformation $x_t = \bar{x} e^{\tilde{x}_t}$:

[9] Note that this log-linearization implies that $\tilde{z}$ is the deviation of log real marginal costs from their steady state value. Therefore in most empirical analyses in the literature log real marginal costs are used.

$$\frac{\bar{P}^{\star}e^{\tilde{P}_t^{\star}}}{\bar{P}e^{\tilde{P}_t}} = \frac{\epsilon}{1-\epsilon}\frac{E_t\left\{\sum_{j=0}^{\infty}\theta^j\gamma_f^j\left(\bar{\pi}^je^{\tilde{\pi}_{t,t+j}}\right)^{\epsilon}\bar{Y}e^{\tilde{Y}_{t+j}}\overline{z}e^{\widetilde{z}_{t+j}}\right\}}{E_t\left\{\sum_{j=0}^{\infty}\theta^j\gamma_f^j\left(\bar{\pi}^je^{\tilde{\pi}_{t,t+j}}\right)^{\epsilon-1}\bar{Y}e^{\tilde{Y}_{t+j}}\right\}} \tag{4.D.7}$$

$$= \frac{\epsilon\overline{z}}{1-\epsilon}\frac{E_t\left\{\sum_{j=0}^{\infty}\theta^j\gamma_f^j\bar{\pi}^{j\epsilon}e^{\epsilon\tilde{\pi}_{t,t+j}+\tilde{Y}_{t+j}+\widetilde{z}_{t+j}}\right\}}{E_t\left\{\sum_{j=0}^{\infty}\theta^j\gamma_f^j\bar{\pi}^{j(\epsilon-1)}e^{(\epsilon-1)\tilde{\pi}_{t,t+j}+\tilde{Y}_{t+j}}\right\}}. \tag{4.D.8}$$

Using the steady state relation in (4.D.6), and the assumption of zero steady state inflation, $\bar{\pi} = 1$, equation (4.D.8) simplifies to the following:

$$E_t\left\{\sum_{j=0}^{\infty}\theta^j\gamma_f^je^{\epsilon\tilde{\pi}_{t,t+j}+\tilde{Y}_{t+j}+\widetilde{z}_{t+j}-\tilde{P}_t^{\star}}\right\} = E_t\left\{\sum_{j=0}^{\infty}\theta^j\gamma_f^je^{(\epsilon-1)\tilde{\pi}_{t,t+j}+\tilde{Y}_{t+j}-\tilde{P}_t}\right\}. \tag{4.D.9}$$

Next, we simplify (4.D.9) using the approximation property $e^x \approx 1 + x$:

$$\sum_{j=0}^{\infty}\theta^j\gamma_f^jE_t\left\{1+\epsilon\tilde{\pi}_{t,t+j}+\tilde{Y}_{t+j}+\widetilde{z}_{t+j}-\tilde{P}_t^{\star}\right\} = \sum_{j=0}^{\infty}\theta^j\gamma_f^jE_t\left\{1+(\epsilon-1)\tilde{\pi}_{t,t+j}+\tilde{Y}_{t+j}-\tilde{P}_t\right\}. \tag{4.D.10}$$

Assuming $|\theta\gamma_f| < 1$ and rearranging (4.D.10) we obtain

$$\tilde{P}_t^{\star} - \tilde{P}_t = (1-\theta\gamma_f)\sum_{j=0}^{\infty}\theta^j\gamma_f^jE_t\left\{\widetilde{z}_{t+j}+\tilde{\pi}_{t,t+j}\right\}, \tag{4.D.11}$$

i.e. the difference between short run deviations of the optimizing firm's price and the aggregate price depends on current marginal costs, the discounted sums future marginal cost changes and future inflation.

In order to obtain the NKPC model, the infinite sum of marginal cost and inflation expectations in (4.D.11) have to be solved. For this purpose, we use the fact that $E_t(\widetilde{z}_t) = \widetilde{z}_t$ and $\tilde{\pi}_{t,t} = 0$, i.e. contemporaneous marginal cost change is known and price change within the same period is 0. Using these relations, the infinite sum of expectations can be simplified:

$$\tilde{P}_t^{\star} - \tilde{P}_t = (1-\theta\gamma_f)\, E_t \left\{\widetilde{z}_t + \tilde{\pi}_{t,t}\right\} + (1-\theta\gamma_f) \sum_{j=1}^{\infty} \theta^j \gamma_f^j E_t \left\{\widetilde{z}_{t+j} + \tilde{\pi}_{t,t+j}\right\} \tag{4.D.12}$$

$$= (1-\theta\gamma_f)\, \widetilde{z}_t + (1-\theta\gamma_f)\, \theta\gamma_f \sum_{j=0}^{\infty} \theta^j \gamma_f^j E_t \left\{\widetilde{z}_{t+j+1} + \tilde{\pi}_{t,t+j+1}\right\} \tag{4.D.13}$$

$$= (1-\theta\gamma_f)\, \widetilde{z}_t + \theta\gamma_f E_t \left\{ E_{t+1} \left\{ (1-\theta\gamma_f) \sum_{j=0}^{\infty} \theta^j \gamma_f^j \left(\widetilde{z}_{t+j+1} + \tilde{\pi}_{t,t+j+1}\right) \right\} \right\} \tag{4.D.14}$$

where the last equality holds from the law of iterated expectations.

The infinite sum of expectations in (4.D.14) can be solved using (4.D.11). For this purpose we rearrange this term:

$$E_{t+1} \left\{ (1-\theta\gamma_f) \sum_{j=0}^{\infty} \theta^j \gamma_f^j \left(\widetilde{z}_{t+j+1} + \tilde{\pi}_{t,t+j+1}\right) \right\} =$$

$$E_{t+1} \left\{ (1-\theta\gamma_f) \sum_{j=0}^{\infty} \theta^j \gamma_f^j \left(\widetilde{z}_{t+j+1} + \tilde{\pi}_{t+1,t+j+1} + \tilde{\pi}_{t,t+1}\right) \right\} \tag{4.D.15}$$

$$= E_{t+1} \left\{ (1-\theta\gamma_f) \sum_{j=0}^{\infty} \theta^j \gamma_f^j \left(\widetilde{z}_{t+j+1} + \tilde{\pi}_{t+1,t+j+1}\right) \right\} + E_{t+1} \left\{\pi_{t,t+1}\right\} \tag{4.D.16}$$

where we use the equality $\tilde{\pi}_{t,t+j} = \tilde{P}_{t+j} - \tilde{P}_t = \tilde{P}_{t+j} - \tilde{P}_{t+1} + \tilde{P}_{t+1} - \tilde{P}_t = \tilde{\pi}_{t+1,t+j} + \tilde{\pi}_{t,t+1}$. From (4.D.11), the last sum of expectations is the price difference in the next period: $\tilde{P}_{t+1}^{\star} - \tilde{P}_{t+1}$. Therefore (4.D.14) simplifies to:

$$\tilde{P}_t^{\star} - \tilde{P}_t = (1-\theta\gamma_f)\, \widetilde{z}_t + \theta\gamma_f E_t \left(\tilde{P}_{t+1}^{\star} - \tilde{P}_{t+1}\right) + \theta\gamma_f E_t \left\{\pi_{t,t+1}\right\} \tag{4.D.17}$$

One relation which has not been set until now is the relation between the optimizing firms' prices, past prices which are adopted by the non-optimizing firms and the aggregate price level, where the latter is directly related to the inflation variable, one of the two variables in the NKPC model. The Calvo price setting assumes that all the firms that re-optimize will choose the same price P_t^*. From the definition of the aggregate price level, we can write[10]:

$$\begin{aligned} P_t = \left[\int_0^1 P_t(i)^{1-\epsilon} di\right]^{\frac{1}{1-\epsilon}} &= \left[\int_{S(t)} P_{t-1}(i)^{1-\epsilon} di + (1-\theta)(P_t^*)^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}} \\ &= \left[\theta(P_{t-1})^{1-\epsilon} + (1-\theta)(P_t^*)^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}} \end{aligned} \tag{4.D.18}$$

where the first integral is over the continuum of all firms, $S(t) \subset [0,1]$ represents the set of firms which do not optimize their prices, and it is assumed that the distribution of prices among firms not adjusting in period t corresponds to the distribution of prices in period $t-1$. Log-linearizing (4.D.18) around zero steady state inflation, we obtain[11]:

$$\tilde{P}_t = \theta \tilde{P}_{t-1} + (1-\theta)\tilde{P}_t^*. \tag{4.D.19}$$

Hence the the difference between the optimizing firms' price deviation and aggregate price deviation is:

$$\tilde{P}_t^* - \tilde{P}_t = \frac{\tilde{P}_t - \theta\tilde{P}_{t-1}}{1-\theta} - \tilde{P}_t = \frac{\theta}{1-\theta}(\tilde{P}_t - \tilde{P}_{t-1}) \tag{4.D.20}$$

and this difference is 0 in the limiting case, $\theta = 0$, if all firms can optimize their prices.

[10] This relation follows from the price setting assumptions of firms and the underlying elasticity of substitution between goods. We refer to Gali (2008) for further details.

[11] We define the log-linearized series, $\tilde{Y}_t$, around its steady state Y as $\tilde{Y} \equiv \ln Y_t - \ln \bar{Y}$. Rearranging, we can solve this for Y_t:

$$\ln Y_t = \ln \bar{Y} + \tilde{Y}_t \Rightarrow Y_t = e^{(\ln \bar{Y} + \tilde{Y}_t)} = e^{\ln \bar{Y}} e^{\tilde{Y}} = \bar{Y} e^{\tilde{Y}_t} \approx \bar{Y}(1 + \tilde{Y}_t)$$

where the last equality comes from the first order Taylor approximation of $e^{\tilde{Y}_t}$ around $\tilde{Y}_t = 0$, and $\bar{Y}$ and $\tilde{Y}_t$ denote the steady state level and deviations from the steady state level, respectively.

Inserting (4.D.20) in (4.D.17) for time t and time $t+1$, we obtain the NKPC relation between inflation and marginal costs:

$$\frac{\theta}{1-\theta}\tilde{\pi}_t = (1-\theta\gamma_f)\,\widetilde{z}_t + \theta\gamma_f E_t\left(\frac{\theta}{1-\theta}\tilde{\pi}_{t+1}\right) + \theta\gamma_f E_t\left\{\pi_{t+1}\right\} \tag{4.D.21}$$

where the notation is simplified such that $\tilde{\pi}_t = \tilde{\pi}_{t-1,t}$ denotes the one period ahead inflation in deviation from its steady state at time t.

Rearranging (4.D.21), we obtain the structural form of the inflation equation in the NKPC model:

$$\tilde{\pi}_t = \frac{(1-\theta)(1-\theta\gamma_f)}{\theta}\widetilde{z}_t + \gamma_f E_t\left(\tilde{\pi}_{t+1}\right) = \lambda\widetilde{z}_t + \gamma_f E_t\left(\tilde{\pi}_{t+1}\right). \tag{4.D.22}$$

4.E Bayesian Inference of the Extended NKPC Model

In this section we summarize the prior specifications, our use of prior predictive likelihoods, and the posterior sampling algorithms for the extended NKPC and HNKPC models. We further present a prior sensitivity analysis for the proposed models using a prior-predictive analysis.

4.E.1 Prior Specification for Parameters

The extended NKPC and HNKPC models contain several additional parameters compared to the standard NKPC model. We classify the model parameters in five groups, and assign independent priors for each group. The first group includes the common parameters in the NKPC and HNKPC models, $\theta_N = \{\lambda, \gamma_f, \phi_1, \phi_2, \Sigma\}$, in (4.B.1). For the structural parameters $\{\lambda, \gamma_f, \phi_1, \phi_2\}$ we define flat priors on restricted regions, which also ensure that the autoregressive parameters, ϕ_1 and ϕ_2, are in the stationary region and the (observation) variance priors are of inverse-Wishart type[12]

$$\begin{aligned} p(\lambda, \gamma_f, \phi_1, \phi_2|\Sigma) &\propto \text{ constant for } |\lambda| < 1,\ |\gamma_f| < 1,\ |\phi_1| + \phi_2 < 1,\ |\phi_2| < 1, \\ \Sigma &\sim IW(1, 20 \times \tilde{\Sigma}), \end{aligned} \tag{4.E.1}$$

where $IW(\nu, \Psi)$ is the inverse Wishart density with scale Ψ and degrees of freedom ν. It is possible to use economic theory or steady state relationships to construct priors for these parameters, see Del Negro and Schorfheide (2008). We do not follow this approach but let the data information dominate our relatively weak prior information. For the same reason, we perform a prior-predictive analysis and investigate the sensitivity of our posterior results with respect to the prior.

Note that the prior specifications of the observation and state covariances are important in this class of models and for macroeconomic data. Since the sample size is typically small, differentiating the short-run variation in series (the observation variances) from the variation in the long-run (the state variation) can be cumbersome, see Canova (2012). We therefore impose a data based prior on the observation covariances. We first estimate an unrestricted reduced form VAR model using demeaned inflation series and (linear) detrended (log) real marginal cost series, and base the observation variance prior on this covariance estimate, $\tilde{\Sigma}$. This specification imposes smoothness for the estimated levels and trends, and ensures that the state errors do not capture all variation in the observed variables. Second, prior distributions for the extra model parameters stemming from the hybrid models, $\theta_H = \{\gamma_b^H, \beta\}$ are defined as uniform priors on restricted regions $|\gamma_b^H| < 1$, $|\beta| < 1$. Third, we define independent inverse-Gamma priors for the state variances

$$\sigma_{\eta_1} \sim IG(20, 20 \times 10^{-2}), \quad \sigma_{\eta_2} \sim IG(20, 20 \times 10^{-3}), \quad \sigma_{\eta_3} \sim IG(1, 1 \times 10^{-5}), \tag{4.E.2}$$

[12] We experimented with wider truncated uniform densities for the λ and γ_f parameters. The prior truncation does not seem to have a substantial affect on the posterior results.

where $IG(\alpha, \alpha\xi)$ is the inverse-Gamma distribution with shape α and scale $\alpha\xi$. Parameters α and ξ are the a priori number and variance of dummy observations.

Similar to the standard counterparts, the extended NKPC and HNKPC models may also suffer from flat likelihood functions. We therefore set weakly informative priors for the state parameters, such that not all variation in inflation and marginal cost series are captured by the time-varying trends and levels. For example, the number of prior dummy observations for σ_{η_1} and σ_{η_2} is much less than the number of observations to limit the prior information.

The fourth prior distribution we consider is applicable to the NKPC and HNKPC models with level shifts. For these models, we consider a fixed level shift probability of 0.04. This choice leads to an a priori expected number of shifts of 8 for 200 observations in the sample. Alternatively, this parameter can be estimated together with other model parameters. However, often the limited number of level shifts plague the inference of this parameter. Hence, we set this value, obtained trough an extensive search over intuitive values of this parameter, prior to analysis.

Finally, for the stochastic volatility models, we specify an inverse-gamma prior for the marginal cost variances. For the correlation coefficient, ρ, we take an uninformative prior $p(\rho) \propto (1-\rho^2)^{-3/2}$, see Çakmaklı et al. (2011).

4.E.2 Posterior Existence and the Sampling Algorithm

We summarize the Bayesian inference for the proposed models. An important point regarding the posterior of the structural parameters is the existence of a posterior distribution and its moments, which depends on the number of instruments and the prior. Given one relatively weak instrument (the second lag of the marginal cost series) the posterior will have very fat tails and the existence of the posterior distribution is ensured through priors defined on a bounded region, see Zellner et al. (2014) for a detailed analysis of a linear IV model with small numbers of weak instruments.

The MCMC sampler for the full conditional posterior distribution is based on Gibbs sampling with a Metropolis-Hastings step and data augmentation, combining the methodologies in Geman and Geman (1984); Tanner and Wong (1987); Gerlach et al. (2000) and Çakmaklı et al. (2011).

Together with the level specifications of the inflation and real marginal cost series the proposed extended NKPC model takes the following form

$$\begin{aligned}
\pi_t - c_{\pi,t} &= \tfrac{\lambda}{1-(\phi_1+\phi_2\gamma_f)\gamma_f}\left(z_t - c_{z,t}\right) + \tfrac{\phi_2\gamma_f\lambda}{1-(\phi_1+\phi_2\gamma_f)\gamma_f}\left(z_{t-1} - c_{z,t-1}\right) + \epsilon_{1,t}, \\
z_t - c_{z,t} &= \phi_1\left(z_{t-1} - c_{z,t-1}\right) + \phi_2\left(z_{t-2} - c_{z,t-2}\right) + \epsilon_{2,t}, \\
c_{\pi,t+1} &= c_{\pi,t} + \kappa_t\eta_{1,t+1}, \\
c_{z,t+1} &= \mu_{z,t} + c_{z,t} + \eta_{2,t+1}, \\
\mu_{z,t+1} &= \mu_{z,t} + \eta_{3,t+1},
\end{aligned} \tag{4.E.3}$$

where $(\epsilon_{1,t},\ \epsilon_{2,t})' \sim NID\left(0, \left(\begin{smallmatrix} \sigma^2_{\epsilon_1} & \rho\sigma_{\epsilon_1}\sigma_{\epsilon_2} \\ \rho\sigma_{\epsilon_1}\sigma_{\epsilon_2} & \sigma^2_{\epsilon_2} \end{smallmatrix}\right)\right)$, $(\eta_{1,t},\ \eta_{2,t},\ \eta_{3,t})' \sim NID\left(0, \left(\begin{smallmatrix} \sigma^2_{\eta_1} & 0 & 0 \\ 0 & \sigma^2_{\eta_2} & 0 \\ 0 & 0 & \sigma^2_{\eta_3} \end{smallmatrix}\right)\right)$ and the disturbances $(\epsilon_{1,t},\ \epsilon_{2,t})'$ and $(\eta_{1,t},\ \eta_{2,t},\ \eta_{3,t})'$ are independent for all t.

The NKPC model in (4.E.3) can be cast into the state-space form as follows

$$\begin{aligned}
Y_t &= HX_t + BU_t + \epsilon_{\mathbf{t}}, & \epsilon_{\mathbf{t}} &\sim N(0, Q_t) \\
X_t &= FX_{t-1} + R_t\eta_{\mathbf{t}}, & \eta_{\mathbf{t}} &\sim N(0, I)
\end{aligned} \tag{4.E.4}$$

where

$$Y_t = \begin{pmatrix} \pi_t \\ z_t \end{pmatrix}, \quad X_t = \begin{pmatrix} c_{\pi,t}, & c_{z,t}, & \mu_{z,t}, & c_{z,t-1}, & c_{z,t-2} \end{pmatrix}', \quad U_t = \begin{pmatrix} z_t \\ z_{t-1} \\ z_{t-2} \end{pmatrix}, \quad \epsilon_{\mathbf{t}} = \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix},$$

$$H = \begin{pmatrix} 1 & -\alpha_1 & 0 & -\alpha_2 & 0 \\ 0 & 1 & 0 & -\phi_1 & -\phi_2 \end{pmatrix}, \quad B = \begin{pmatrix} \alpha_1 & \alpha_2 & 0 \\ 0 & \phi_1 & \phi_2 \end{pmatrix}, \quad Q_t = \begin{pmatrix} \sigma^2_{\epsilon_1,t} & \rho\sigma_{\epsilon_1,t}\sigma_{\epsilon_2} \\ \rho\sigma_{\epsilon_1,t}\sigma_{\epsilon_2} & \sigma^2_{\epsilon_2} \end{pmatrix},$$

$$F = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \quad R_t = \begin{pmatrix} \kappa_t \sigma_{\eta_1} & 0 & 0 \\ 0 & \sigma_{\eta_2} & 0 \\ 0 & 0 & \sigma_{\eta_3} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \eta_{\mathbf{t}} = \begin{pmatrix} \eta_{1,t} \\ \eta_{2,t} \\ \eta_{3,t} \end{pmatrix},$$

where $\alpha_1 = \frac{\lambda}{1-(\phi_1+\phi_2\gamma_f)\gamma_f}$ and $\alpha_2 = \frac{\lambda\gamma_f\phi_2}{1-(\phi_1+\phi_2\gamma_f)\gamma_f}$.

Once the state-space form of the model is set as in (4.E.4) standard inference techniques in state-space models can be carried out. Let $Y_{1:T} = (Y_1, Y_2, \ldots, Y_T)'$, $X_{1:T} = (X_1, X_2, \ldots, X_T)'$, $U_{1:T} = (U_1, U_2, \ldots, U_T)'$, $\sigma^2_{\epsilon_1,1:T} = (\sigma^2_{\epsilon_1,1}, \sigma^2_{\epsilon_1,2}, \ldots, \sigma^2_{\epsilon_1,T})'$ and $\theta = (\phi_1, \phi_2, \gamma_f, \lambda)'$. For the most general NKPC model with level shifts and stochastic volatility, the simulation scheme is as follows

1. Initialize the parameters by drawing κ_t using the prior for level shift probability, p_κ, and by drawing unobserved states X_t, h_t for $t = 1, 2, \ldots, T$ from standard normal distribution and conditional on κ_t for $t = 0, 1, \ldots, T$. Initialize $m = 1$.

2. Sample $\theta^{(m)}$ from $p(\theta | Y_{1:T}, X_{1:T}, U_{1:T}, R_{1:T}, Q_{1:T})$.

3. Sample $X_t^{(m)}$ from $p(X_t | \theta^{(m)}, Y_{1:T}, U_{1:T}, R_{1:T}, Q_{1:T})$ for $t = 1, 2, \ldots, T$.

4. Sample $h_t^{(m)}$ from $p(h_t | X_{1:T}^{(m)}, \theta^{(m)}, Y_{1:T}, U_{1:T}, R_{1:T}, \rho, \sigma^2_{\epsilon_2}, \sigma^2_{\eta_4})$ for $t = 1, 2, \ldots, T$.

5. Sample $\kappa_t^{(m)}$ from $p(\kappa_t^{(m)} | \theta^{(m)}, Y_{1:T}, h_{1:T}^{(m)}, U_{1:T}, R_{1:T}, \rho, \sigma^2_{\epsilon_2})$ for $t = 1, 2, \ldots, T$.

6. Sample $\sigma^{2,(m)}_{\eta_i}$ from $p(\sigma^{2,(m)}_{\eta_i} | X_{1:T}^{(m)}, h_{1:T}^{(m)}, \kappa_{1:T}^{(m)})$ for $i = 1, 2, 3, 4$.

7. Sample $\rho^{(m)}$ from from $p(\rho^{(m)} | X_{1:T}^{(m)}, h_{1:T}^{(m)}, Y_{1:T}, U_{1:T}, \theta^{(m)}, \sigma^{2,(m-1)}_{\epsilon_2})$.

8. Sample $\sigma^{2,(m)}_{\epsilon_2}$ from from $p(\sigma^{2,(m)}_{\epsilon_2} | \rho^{(m)}, X_{1:T}^{(m)}, h_{1:T}^{(m)}, Y_{1:T}, U_{1:T}, \theta^{(m)})$.

9. Set $m = m + 1$, repeat (2)-(9) until $m = M$.

Steps (3)-(5) are common to many models in the Bayesian state-space framework, see for example Kim and Nelson (1999b); Gerlach et al. (2000); Çakmaklı (2012).

Sampling of θ

Conditional on the states $c_{\pi,t}, c_{z,t}$ and h_t for $t = 1, 2, \ldots, T$, redefining the variables such that $\tilde{\pi}_t = \pi_t - c_{\pi,t}$, $\tilde{z}_t = z_t - c_{z,t}$ and $\varepsilon_t = \epsilon_t / \exp(h_t/2)$, the measurement equation in (4.E.4) can be rewritten as

$$\begin{aligned} \tilde{\pi}_t &= \tfrac{\lambda}{1-(\phi_1+\phi_2\gamma_f)\gamma_f}\tilde{z}_t + \tfrac{\phi_2\gamma_f\lambda}{1-(\phi_1+\phi_2\gamma_f)\gamma_f}\tilde{z}_{t-1} + \varepsilon_t \\ \tilde{z}_t &= \phi_1\tilde{z}_{t-1} + \phi_2\tilde{z}_{t-2} + \epsilon_{2,t}. \end{aligned} \tag{4.E.5}$$

Posterior distributions of the structural parameters under flat priors are non-standard since z_t term also is on the right hand side of (4.E.5) and the model is highly non-linear in parameters. We therefore use two Metropolis Hastings steps to sample these structural parameters, see Metropolis et al. (1953) and Hastings (1970). For sampling ϕ_1, ϕ_2 conditional on λ, γ_f and other model parameters, the candidate density is a multivariate student-t density on the stationary region with a mode and scale with the posterior mode and scale using only the second equation in (4.E.5) and 1 degrees of freedom. For sampling λ, γ_f conditional on ϕ_1, ϕ_2 and other model parameters, the candidate is a uniform density.

Sampling of states, X_t

Conditional on the remaining model parameters, drawing $X_{0:T}$ can be implemented using standard Bayesian inference. This constitutes running the Kalman filter first and running a simulation smoother using the filtered values for drawing smoothed states as in Carter and Kohn (1994) and Frühwirth-Schnatter (1994). We start the recursion for

$t = 1, \ldots, T$

$$\begin{aligned}
X_{t|t-1} &= FX_{t-1|t-1} \\
P_{t|t-1} &= FP_{t-1|t-1}F' + R_t'R_t \\
\eta_{t|t-1} &= y_t - HX_{t|t-1} - BU_t \\
\zeta_{t|t-1} &= HP_{t|t-1}H' + Q_t \\
K_t &= P_{t|t-1}H'\zeta_{t|t-1}' \\
X_{t|t} &= X_{t|t-1} + K_t\eta_{t|t-1} \\
P_{t|t} &= P_{t|t-1} - K_tH'\zeta_{t|t-1}',
\end{aligned} \tag{4.E.6}$$

and store $X_{t|t}$ and $P_{t|t}$. The last filtered state $X_{T|T}$ and its covariance matrix $P_{T|T}$ correspond to the smoothed estimates of the mean and the covariance matrix of the states for period T. Having stored all the filtered values, simulation smoother involves the following backward recursions for $t = T-1, \ldots, 1$

$$\begin{aligned}
\eta_{t+1|t}^* &= X_{t+1} - FX_{t|t} \\
\zeta_{t+1|t}^* &= FP_{t|t}F' + R_{t+1}'R_{t+1} \\
X_{t|t,X_{t+1}} &= X_{t|t} + P_{t|t}F'\zeta_{t+1|t}^{*-1}\eta_{t+1|t}^* \\
P_{t|t,P_{t+1}} &= P_{t|t} - P_{t|t}F'\zeta_{t+1|t}^{*-1}FP_{t|t}.
\end{aligned} \tag{4.E.7}$$

Intuitively, the simulation smoother updates the states using the same principle as in the Kalman filter, where at each step filtered values are updated using the smoothed values obtained from backward recursion. For updating the initial states, using the state equation $X_{0|t,X_1} = F^{-1}X_1$ and $P_{0|t,P_1} = F^{-1}(P_1 + R_1'R_1)F'^{-1}$ can be written for the first observation. Given the mean $X_{t|t,X_{t+1}}$ and the covariance matrix $P_{t|t,P_{t+1}}$, the states can be sampled from $X_t \sim N(X_{t|t,X_{t+1}}, P_{t|t,P_{t+1}})$ for $t = 0, ..., T$.

Sampling of inflation volatilities, h_t

Conditional on the remaining model parameters, we can draw $h_{0:T}$ using standard Bayesian inference as in the case of X_t. One important difference, however, stems from the logarithmic transformation of the variance in the stochastic volatility model. As the transformation concerns the error structure, the square of which follows a χ^2 distribution, the system is not Gaussian but follows a log-χ^2 distribution. Noticing the properties of log-χ^2 distribution, Kim et al. (1998) and Omori et al. (2007) approximate this distribution using a mixture of Gaussian distributions. Hence, conditional on these mixture components the system remains Gaussian allowing for standard inference outlined above. For details, see Omori et al. (2007). For the estimation of the volatilities in the BVAR-TV-SV model we use the extension of the algorithm following Kastner and Frühwirth-Schnatter (2013) for improving the efficiency of the MCMC algorithm.

Sampling of structural break parameters, κ_t

Sampling of structural break parameters, κ_t relies on the conditional posterior of the binary outcomes, i.e. the posterior value in case of a structural break in period t and the posterior value of the case of no structural breaks. However, evaluating this posterior requires one sweep of filtering, which is of order $O(T)$. As this evaluation should be implemented for each period t the resulting procedure would be of order $O(T^2)$. When the number of sample size is large this would result in an infeasible scheme. Gerlach et al. (2000) propose an efficient algorithm for sampling structural break parameters, κ_t, conditional on the observed data, which is still of order $O(T)$. We implement this algorithm for estimation of the structural breaks and refer to Gerlach et al. (2000); Giordani and Kohn (2008) for details.

Sampling of state error variances, σ^2_η

Using standard results from a linear regression model with a conjugate prior for the variances in (4.E.4), it follows that the conditional posterior distribution of $\sigma^2_{\eta_i}$, with $i = 1, 2, 3, 4$ is an inverted Gamma distribution with scale parameter $\Phi_{\eta_i} + \sum_{t=1}^{T} \eta^2_{i,t}$ and with $T + \nu_{\eta_i}$ degrees of freedom for $i = 2, 3, 4$ where Φ_{η_i} and ν_{η_i} are the scale and degrees of freedom parameters of the prior density. For $i = 1$ the parameters of the inverted Gamma distribution becomes $\Phi_{\eta_1} + \sum_{t=1}^{T} \kappa_t \eta^2_{1,t}$ and $\sum_{t=1}^{T} \kappa_t + \nu_{\eta_1}$.

Sampling of marginal costs variance and correlation coefficient

To sample the variance of marginal costs and correlation coefficient, we decompose the multivariate normal distribution of $\epsilon_{\mathbf{t}}$ into the conditional distribution of $\epsilon_{2,t}$ given $\epsilon_{1,t}$ and the marginal distribution of $\epsilon_{1,t}$, as in Çakmaklı et al. (2011). This results in

$$\prod_{t=1}^{T} f(\epsilon_{\mathbf{t}}) = \prod_{t=1}^{T} \frac{1}{\sigma_{\epsilon_{1,t}}} \, \phi\left(\frac{\epsilon_{1,t}}{\sigma_{\epsilon_{1,t}}}\right) \frac{1}{\sigma_{\epsilon_{2,t}} \sqrt{(1-\rho^2)}} \, \phi\left(\frac{\epsilon_{2,t} - \rho\epsilon_{1,t}}{\sigma_{\epsilon_{2,t}}(1-\rho^2)}\right), \tag{4.E.8}$$

Hence, together with prior for the variance in (4.E.4), variance of the marginal cost series can be sampled using (4.E.8) by setting up a Metropolis-Hasting step using an inverted Gamma candidate density with scale parameter $\sum_{t=1}^{T} \epsilon_{2,t}^2$ and with T degrees of freedom. To sample ρ from its conditional posterior distribution we can again use (4.E.8). Conditional on the remaining parameters the posterior becomes

$$(1-\rho^2)^{-\frac{3}{2}} \prod_{t=1}^{T} \left(\frac{1}{\sqrt{(1-\rho^2)}} \, \phi\left(\frac{\epsilon_{2,t} - \rho\epsilon_{1,t}}{\sigma_{\epsilon_{2,t}}(1-\rho^2)}\right)\right). \tag{4.E.9}$$

We can easily implement the griddy Gibbs sampler approach of Ritter and Tanner (1992). Given that $\rho \in (-1, 1)$ we can setup a grid in this interval based on the precision we desire about the value of ρ.

4.E.3 Prior-predictive Likelihood Analysis

In the proposed models, it is important to assess the effects of the specified prior distributions on the predictive likelihoods. Due to the nonlinear structure of the models, assessing the amount of prior information on the predictive results is not trivial. We present a prior-predictive analysis as in Geweke (2010). For each of the extended NKPC and HNKPC model, we consider 1000 parameter draws from the joint prior distribution and compute the prior predictive likelihoods for the period between 1973-II and 2012-I. Hence a comparison of the resulting prior predictions will indicate which model is preferred by the priors.

4.F Bayesian Inference of the Extended HNKPC Model

Posterior inference of the HNKPC models with time varying parameters follow similar to section 4.E, using the Gibbs sampler with data augmentation. Together with the level specifications of the inflation and real marginal cost series the proposed extended HNKPC model takes the following form

$$\begin{aligned}
\pi_t - c_{\pi,t} &= \frac{\lambda^H}{(1-\gamma_b^H\gamma_f^H)(1-(\phi_1+\phi_2\gamma_f^H)\gamma_f^H)}\left(z_t - c_{z,t}\right) + \frac{\phi_2\gamma_f^H\lambda^H}{(1-\gamma_b^H\gamma_f^H)(1-(\phi_1+\phi_2\gamma_f^H)\gamma_f^H)}\left(z_{t-1} - c_{z,t-1}\right), \\
&+ \frac{\gamma_b^H\gamma_f^H}{(1-\gamma_b^H\gamma_f^H)}\frac{\gamma_f^H}{1-\gamma_f^H\beta}(\mu_t - c_{\pi,t}) + \frac{\gamma_b^H}{(1-\gamma_b^H\gamma_f^H)}\left(\pi_{t-1} - c_{\pi,t-1}\right) + \frac{1}{(1-\gamma_b^H\gamma_f^H)}\epsilon_{1,t}, \\
z_t - c_{z,t} &= \phi_1\left(z_{t-1} - c_{z,t-1}\right) + \phi_2\left(z_{t-2} - c_{z,t-2}\right) + \epsilon_{2,t}, \\
c_{\pi,t+1} &= c_{\pi,t} + \kappa_t\eta_{1,t+1}, \\
c_{z,t+1} &= \mu_{z,t} + c_{z,t} + \eta_{2,t+1}, \\
\mu_{z,t+1} &= \mu_{z,t} + \eta_{3,t+1}.
\end{aligned} \tag{4.F.1}$$

This can be cast into the state-space form as in (4.E.4)

$$\begin{aligned}
Y_t &= HX_t + BU_t + \epsilon_{\mathbf{t}}, \quad \epsilon_{\mathbf{t}} \sim N(0, Q_t) \\
X_t &= FX_{t-1} + R_t\eta_{\mathbf{t}}, \quad \eta_{\mathbf{t}} \sim N(0, I)
\end{aligned} \tag{4.F.2}$$

using the following definitions

$$Y_t = \begin{pmatrix} \pi_t \\ z_t \end{pmatrix}, \; X_t = \begin{pmatrix} c_{\pi,t} & c_{z,t} & \mu_{z,t}, & c_{z,t-1} & c_{z,t-2} & c_{\pi,t-1} \end{pmatrix}', \; \epsilon_{\mathbf{t}} = \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix},$$

$$U_t = \begin{pmatrix} z_t & z_{t-1} & z_{t-2} & \pi_{t-1} & \mu_t \end{pmatrix}', \; B = \begin{pmatrix} \alpha_1 & \alpha_2 & 0 & \alpha_4 & \alpha_3 \\ 0 & \phi_1 & \phi_2 & 0 & 0 \end{pmatrix},$$

$$H = \begin{pmatrix} 1-\alpha_3 & -\alpha_1 & 0 & -\alpha_2 & 0 & -\alpha_4 \\ 0 & 1 & 0 & -\phi_1 & -\phi_2 & 0 \end{pmatrix}, \; Q_t = \begin{pmatrix} \sigma^2_{\epsilon_1,t} & \rho\sigma_{\epsilon_1,t}\sigma_{\epsilon_2} \\ \rho\sigma_{\epsilon_1,t}\sigma_{\epsilon_2} & \sigma^2_{\epsilon_2} \end{pmatrix},$$

$$F_t = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad R_t = \begin{pmatrix} \kappa_t \sigma_{\eta_1} & 0 & 0 \\ 0 & \sigma_{\eta_2} & 0 \\ 0 & 0 & \sigma_{\eta_3} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \eta_{\mathbf{t}} = \begin{pmatrix} \eta_{1,t} \\ \eta_{2,t} \\ \eta_{3,t} \end{pmatrix},$$

where parameters $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ are defined as functions of the structural form parameters

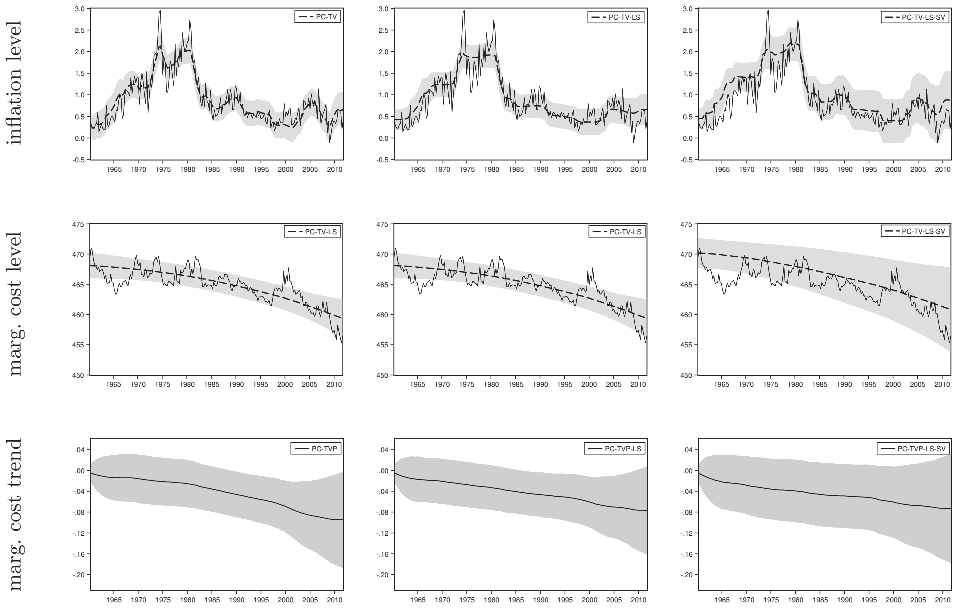
$$\alpha_1 = \frac{\lambda^H}{\left(1 - (\phi_1 + \phi_2 \gamma_f^H)\gamma_f^H\right)\left(1 - \gamma_b^H \gamma_f^H\right)}, \quad \alpha_2 = \frac{\lambda^H \gamma_f^H \phi_2}{\left(1 - (\phi_1 + \phi_2 \gamma_f^H)\gamma_f^H\right)\left(1 - \gamma_b^H \gamma_f^H\right)},$$
$$\alpha_3 = \frac{\gamma_b^H \gamma_f^H}{\left(1 - \gamma_b^H \gamma_f^H\right)} \frac{\gamma_f^H}{\left(1 - \gamma_f^H \beta\right)}, \quad \alpha_4 = \frac{\gamma_b^H}{\left(1 - \gamma_b^H \gamma_f^H\right)}.$$

Given this setup, posterior inference can be carried out using the steps outlined in section 4.E.

4.G Posterior Results for the NKPC Models with non-filtered Time Series

This section presents additional estimation results for the NKPC models with non-filtered time series. We summarize the estimated levels, volatilities, breaks and inflation expectations obtained from the NKPC-TV, NKPC-TV-LS and NKPC-TV-LS-SV models. Figure G.1 shows the estimated levels from the three NKPC models. Estimated inflation levels, computed as the posterior mean of the smoothed states, are given in the first row of Figure G.1. Shaded areas around the posterior means represent the 95% HPDI for the estimated levels. For all three models, estimated inflation levels nicely track the observed inflation. Effects of the level specification are reflected in the estimates in various ways. First, when we model inflation level changes as discrete level shifts rather than continuous changes, we observe a relatively smoother pattern in estimated inflation levels. This effect can be seen by comparing the second and first graphs in the first row of Figure G.1. While estimated inflation level in the first graph follows the observed inflation patterns closely, estimated inflation level in the second (and third to a less extent) graph mostly indicates three distinct periods. These periods are the high inflation periods capturing 1970s with a constant inflation level around 1.7% (quarterly inflation) following a low inflation period in 1960s, and the period after the beginning of 1980s with a stable inflation level around 0.5%, see Cecchetti et al. (2007) for similar findings. Second, adding the stochastic volatility together with level shifts results in discrete level shifts in inflation which are more frequent than the model with only level shifts.

Figure G.1: Level, trend and slope estimates from the NKPC models



Note: The top panel exhibits estimated inflation levels. The middle and the bottom panels show estimated real marginal cost levels and slopes, respectively. Grey shaded areas correspond to the 95% HPDI. NKPC-TV refers to the NKPC model with time varying levels and trends. NKPC-TV-LS refers to the NKPC model with time varying levels and trends. NKPC-TV-LS-SV refers to the NKPC model with time varying levels, trends and volatility. HNKPC-TV refers to the Hybrid NKPC model with time varying levels, trends and inflation expectations. HNKPC-TV-LS refers to the HNKPC model with time varying levels, trends and inflation expectations. HNKPC-TV-LS-SV refers to the HNKPC model with time varying levels, trends, inflation expectations and volatility. Results are based on 40000 simulations of which the first 20000 are discarded for burn-in.

The second panel in Figure G.1 presents the estimated levels for the real marginal cost series for all models. A common feature of all these estimates is the smoothness of the estimated levels. In all models, marginal cost series follows a slightly nonlinear trend during the sample period. The estimated slopes of these trends for all models are given in the bottom panel of Figure G.1, together with the 95% HPDIs. Nonlinearity of the negative trend is reflected in the negative values for the slope of the trend, with an increasing magnitude at the end of the sample. This change in the slope of the trend is accompanied by the increasing uncertainty about the slope. The difference between the models in terms of the estimated marginal cost structures is negligible.

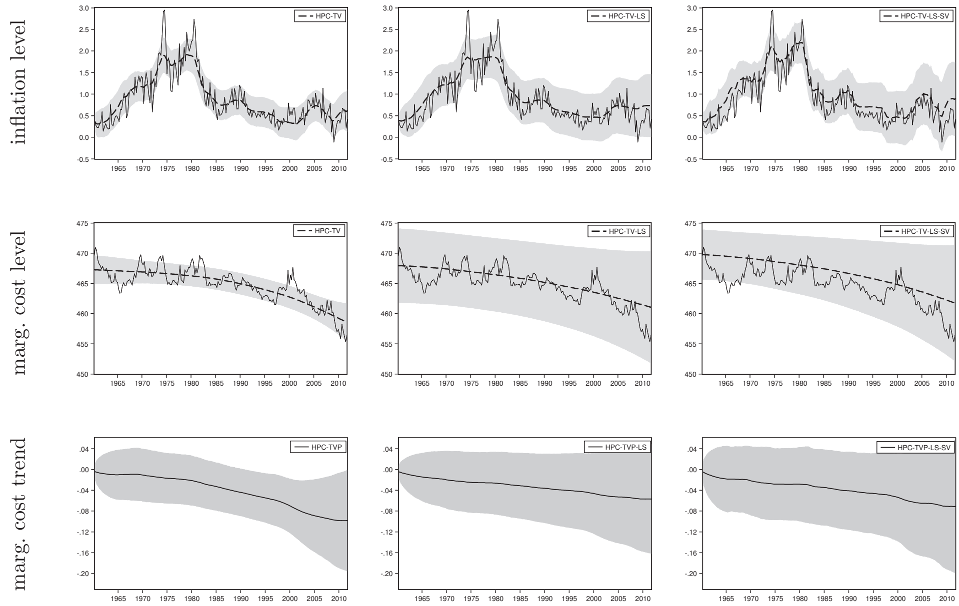
4.H Posterior Results for the HNKPC Models with non-filtered Time Series

This section presents additional estimation results for the HNKPC models with non-filtered time series. We summarize the estimated levels, volatilities, breaks and inflation expectations obtained from the HNKPC-TV, HNKPC-TV-LS and HNKPC-TV-LS-SV models.

Figure H.1 presents the estimated inflation levels, together with estimated levels and trends of the marginal cost series.

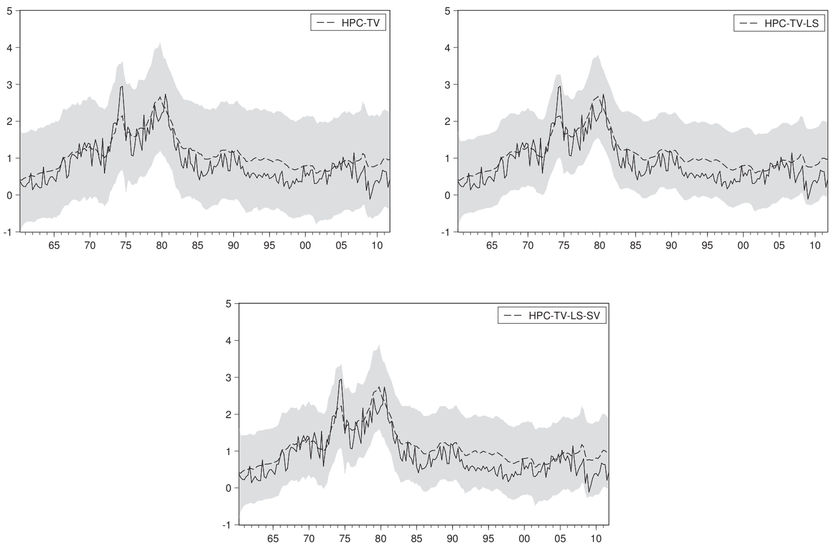
Figure H.2 presents the estimated inflation expectations together with observed survey based inflation expectations.

Figure H.1: Level, trend and slope estimates from the HNKPC models



Note: The top panel exhibits estimated inflation levels. The middle and the bottom panels show estimated real marginal cost levels and slopes, respectively. Grey shaded areas correspond to the 95% HPDI. Results are based on 40000 simulations of which the first 20000 are discarded for burn-in.

Figure H.2: Implied inflation expectations by HNKPC models



Note: The thick solid lines are the posterior means of inflation expectations from the HNKPC models. The thin solid lines are the observations of inflation expectations from survey data. Grey shaded areas are the 95% HPDI for estimated inflation expectations. Results are based on 40000 simulations of which the first 20000 are discarded for burn-in.

4.I Predicted Inflation Densities from All Proposed Models

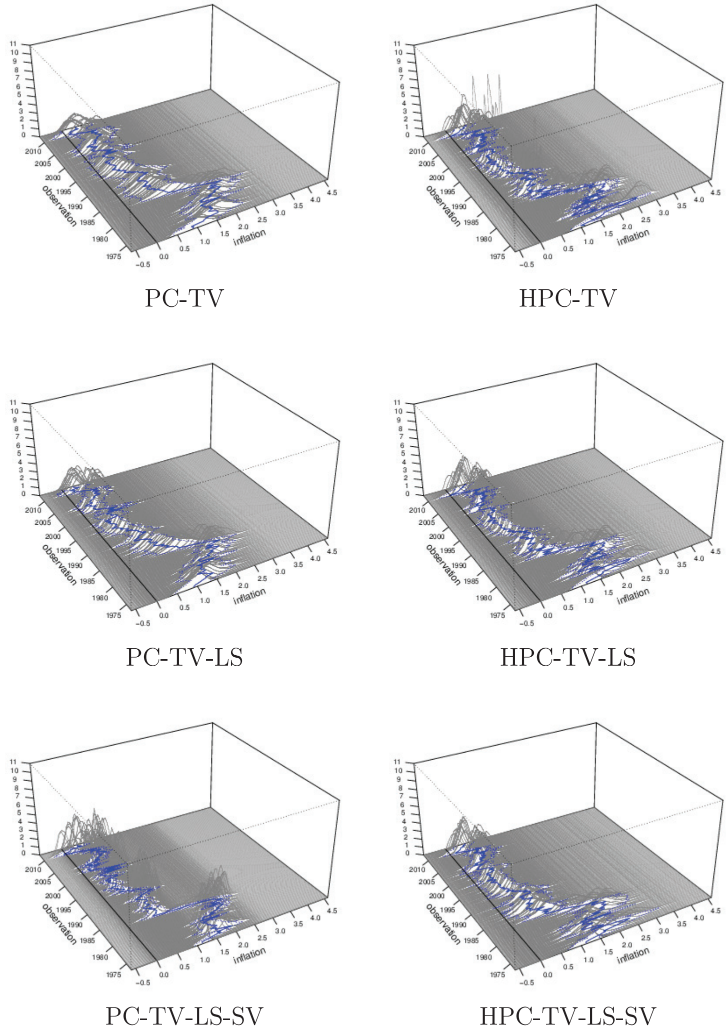
This section presents the entire distribution of the inflation predictions for all NKPC and HNKPC models. Predicted inflation densities from all proposed models are presented in Figure I.1. In these figures, the solid lines represent the posterior mean of predicted inflation, and the white areas under the inflation densities show the inflation levels with non-zero posterior probability. For all models we propose, inflation predictions are concentrated around high (low) values during the high (low) inflationary periods. The uncertainty around the inflation predictions are also high for these periods, together with the periods when inflation is subject to a transition to low values around 1980s.

When the observed inflation values are close to the zero bound, the predictive densities indicate disinflationary risk, computed as the fraction of the predictive distribution below zero.

4.J Prior-predictive Likelihoods of Proposed Models

Due to the complex model structures in the proposed models, it is important to address the effects of the specified prior distributions on the predictive performances. We therefore perform the prior-predictive analysis outlined in section 4.E for the extended NKPC models, for the forecast sample analyzed earlier, covering the period between 1973-II and 2012-I. Table J.1 presents the average and cumulative prior predictive likelihoods for the forecast sample. Prior predictive likelihoods, not using the data information and also using weak prior information, naturally perform worse than the predictive results reported in Table K.2. Table J.1 also shows that the adopted prior distributions clearly favor the less parameterized model, NKPC-TV. Moreover, the priors clearly do not favor models with stochastic volatility components. Most importantly, the 'best performing model' according to the predictive results in Table K.2, HNKPC-TV-LS-SV, is the least

Figure I.1: Predicted inflation densities from NKPC and HNKPC models



Note: The figure presents one period ahead predictive distributions of inflation from the NKPC and HNKPC models, for the period between the third quarter of 1973 and the first quarter of 2012. Model abbreviations are as in Figure G.1 . Results are based on 40000 simulations of which the first 20000 are discarded for burn-in.

favorable one according to the adopted prior distributions using the same forecast sample. We therefore conclude that data information is dominant, and the superior predictive performance of the HNKPC-TV-LS-SV model is not driven by the prior distribution.

Table J.1: Prior-predictive results for the NKPC models

Model	Average (Log) Pred. Likelihood	Cumulative (Log) Pred. Likelihood
NKPC-TV	-1.16	-180.88
NKPC-TV-LS	-1.36	-210.91
NKPC-TV-LS-SV	-1.45	-224.66
HNKPC-TV	-1.28	-199.22
HNKPC-TV-LS	-1.27	-197.68
HNKPC-TV-LS-SV	-2.04	-318.77

Note: The table reports the prior-predictive performances of all competing models for the prediction sample over the period 1973-II until 2012-I. 'Average (Cumulative) Log Pred. Likelihood' stands for the average (sum) of the natural logarithms of predictive likelihoods. Results are based on 1000 simulations from the joint priors of model parameters. Model abbreviations are as in Table 1 in the chapter.

4.K Posterior and Predictive Results from Alternative Models for Robustness Checks

The proposed NKPC and HNKPC models extend the standard models in several ways. First, both model structures introduce time variation in the long and short run dynamics of inflation and marginal cost series. Second, the introduction and the iterative solution of the expectational mechanisms and the survey data in the extended HNKPC models enables the use of more data information. Furthermore, extended and standard HNKPC models use the additional information from a backward looking component for the inflation series compared to the HNKPC counterparts. According to the predictive results, the most comprehensive model, HNKPC-TV-LS-SV is also the best performing model. However, a deeper analysis is needed in order to see the added predictive gains from each of these extensions. In this section we consider several alternative models and their predictive performances to separately address the predictive gains from each of these extensions in the model structure. Table K.1 presents all NKPC and HNKPC model structures we compare to differentiate these effects.

Table K.1: Standard and extended NKPC models: robustness check

low/high frequencies \ model structure	iterated expectations solution		direct expectations data	
	NKPC	HNKPC	NKPC	HNKPC
linear trend	NKPC-LT	n/a ⋆	NKPCS-LT	HNKPCS-LT
Hodrick-Prescott filter	NKPC-HP	n/a ⋆	NKPCS-HP	HNKPCS-HP
time varying levels	NKPC-TV	HNKPC-TV	NKPCS-TV	HNKPCS-TV
time varying levels and switching	NKPC-TV-LS	HNKPC-TV-LS	NKPCS-TV-LS	HNKPCS-TV-LS
time varying levels and stochastic volatility	NKPC-TV-SV	HNKPC-TV-SV	NKPCS-TV-SV	HNKPCS-TV-SV
time varying levels, switching and stochastic volatility	NKPC-TV-LS-SV	HNKPC-TV-LS-SV	NKPCS-TV-LS-SV	HNKPCS-TV-LS-SV

The first two columns present the standard and extended (H)NKPC models presented in the main text of the chapter, for which expectational mechanisms are solved explicitly. The last two columns present alternative model structures for (H)NKPC models. For these models, we do not iterate inflation expectations in the models, but instead replace them with survey data directly. NKPC(S)-LT (NKPC-HP(S)) refers to the NKPC model where the real marginal cost series is detrended using linear trend (Hodrick-Prescott) filter. For the remaining models real marginal cost series' trend is modeled using local linear trend model. NKPC(S)-TV refers to the NKPC model with time varying inflation levels. NKPC(S)-TV-LS refers to the NKPC model with time varying inflation levels together with level shifts. NKPC(S)-TV-SV refers to the NKPC model with time varying inflation levels and stochastic volatility. NKPC(S)-TV-LS-SV refers to the NKPC model with time varying inflation levels together with level shifts and stochastic volatility. HNKPC(S)-TV refers to the Hybrid NKPC model with time varying levels and inflation expectations. HNKPC(S)-TV-LS refers to the HNKPC model with time varying levels together with level shifts and inflation expectations. HNKPC(S)-TV-SV refers to the HNKPC model with time varying levels, inflation expectations and stochastic volatility. HNKPC(S)-TV-LS-SV refers to the HNKPC model with time varying levels together with level shifts, inflation expectations and stochastic volatility.

⋆ Iterative solution of these models without using the survey data does not exist.

The first set of alternative models we consider are the standard NKPC and HNKPC models combined with data from survey expectations, without introducing explicit time variation in the low frequency structure of data but instead demeaning the inflation series, and detrending the marginal cost series prior to analysis. These models are given in the first two rows of the right panel of Table K.1 and are abbreviated by NKPCS-LT, NKPCS-HP, HNKPCS-LT and HNKPCS-HP, according to linear detrending or HP detrending prior to analysis. The improved predictive performances of NKPCS-LT and NKPCS-HP models compared to the standard NKPC counterparts show predictive gains from incorporating survey expectations in the models. Furthermore, comparing the predictive performances of the HNKPCS-LT and HNKPCS-HP models with the time-varying hybrid models, such as the HNKPC-TV or HNKPC-TV-LS models show the gains from incorporating time variation alone, since all these models use survey data and the backward looking component for inflation.

The second set of alternative models we consider, on the right panel of Table K.1, are NKPC models with time-varying levels, where we incorporate the survey expectations in the model directly rather than solving the model iteratively. These models correspond to (4.B.1) where the expectation term is replaced by survey expectations. We denote these models by NKPCS-TV, NKPCS-TV-LS and NKPCS-TV-LS-SV, for the time-varying levels, time-varying levels with regimes shifts in inflation and time-varying levels with regime shifts and stochastic volatility component in inflation, respectively. Comparing the predictive results of these models to the HNKPC counterparts provide the predictive gains solely from the HNKPC extension, i.e. they separate the gains from incorporating the backward looking inflation component in the model from the other model extensions.

The third set of alternative models we consider are the HNKPC models using the survey expectations directly, without solving for the expectational mechanisms. We denote these models by HNKPCS-TV, HNKPCS-TV-LS and HNKPCS-TV-LS-SV, for the time-varying levels, time-varying levels with regimes shifts in inflation and time-varying levels with regime shifts and stochastic volatility component, respectively. Comparing the predictive performance of these models with the proposed HNKPC models clarifies the predictive gains from solving for the inflation expectations iteratively in the hybrid models.

The final set of alternative models aim to pinpoint predictive gains from introducing level shifts in inflation in the models with a stochastic volatility component. The comparison of the predictive results of models with time-varying levels and stochastic volatility, (H)NKPC-TV-SV, and with level shifts and stochastic volatility, (H)NKPC-TV-LS-SV, highlights predictive gains solely from introducing level shifts when changes in inflation volatility are taken into account.

One period ahead MSFE and log marginal likelihoods of these models, together with the standard (H)NKPC models and the models proposed in the chapter, are given in Table K.2. The prediction results are based on the forecast sample, which covers the period between the second quarter of 1973 and the first quarter of 2012. Comparing the first block and the first two rows of the second block in Table K.2, we see that the gains from using survey data inflation are substantial even in the standard NKPC models. In terms of predictive gains, the biggest improvement in predictive likelihoods and the MSFE are achieved with this contribution in the models. However, the predictive performances of these improved models are still far from the more involved models. Hence the gains from the proposed models do not only stem from the inclusion of the survey data information alone.

We also report the predictive gains resulting specifically from introducing time-variation in the inflation and marginal cost series, by comparing the results of the HNKPCS-LT and HNKPCS-HP models with the HNKPC-TV or HNKPC-TV-LS models in the table. The more involved models with time variation clearly perform better according to the predictive results. Especially the difference in marginal likelihoods of these models enables us to conclude that incorporating time variation in the data is also important.

As a third possible reason for predictive gains, we focus on the models with backward looking components. One way to separate the added value from this component is to consider the second block of Table K.2. The prediction results from the NKPC and HNKPC models in this block are very similar, with slight improvements in the hybrid models, where the backward looking component is incorporated. Another way to see the effect of the backward looking component is to compare the NKPCS-TV, NKPCS-TV-LS and NKPCS-TV-LS-SV models with HNKPCS-TV, HNKPCS-TV-LS and HNKPCS-TV-LS-SV models, respectively. In all these comparisons, the models without the backward looking component performs slightly better (worse) in terms of MSFE (marginal likelihood), hence the backward looking component does not seem to improve predictive results in general and the improvements in the hybrid models mainly stem from incorporating the survey expectations.

From the considered alternative models, time-varying level models with a stochastic volatility component using survey data directly (NKPCS-TV-LS-SV and HNKPCS-TV-LS-SV) clearly perform best. In terms of the predictive likelihoods, these models are also comparable to the 'best performing' model we propose.

A final source of possible predictive gains in the proposed models is the iterative solution of inflation expectations. This comparison is based on the comparison of the models in the third (fourth) block and the fifth (sixth) block of Table K.2, where only the third (fourth) block uses the iterative solution. According to the MSFE, predictive results deteriorate slightly when we solve the system. We find this result rather counterintuitive since the iterative solution is based on the complete model structure. As we show briefly, despite this slight increase in the predictive performances, models without the iterative solutions suffer from identification issues.

We next focus on changes in parameter estimates for the alternative models proposed in this section. Table K.3 presents the parameter estimates for all alternative models. Despite the predictive gains from these alternative models, parameter estimates are rather different from those obtained from the proposed models. Specifically for the hybrid models considered, uncertainty in posterior distributions increase substantially if the iterative model solution is not used. Furthermore, posterior densities of some parameters are quite irregular in most of these models which use expectations data directly. Figure K.1 shows this irregularity for the HNKPCS-TV model, parameters $\lambda^{(H)}$, $\gamma_b^{(H)}$ and $\gamma_f^{(H)}$. The bimodality problem in posterior densities is most apperant in the NKPC slope, $\lambda_b^{(H)}$. Furthermore, the backward looking component $\gamma_b^{(H)}$ is spread over a wide region with multiple modes. Similar results hold for the remaining alternative models which make

Table K.2: Predictive performance of additional NKPC models

Model	(Log) Marg. Likelihood	MSFE 1 period ahead
NKPC-LT	-139.327	0.353
NKPC-HP	-157.195	0.458
NKPCS-LT	-79.141	0.105
NKPCS-HP	-85.397	0.130
HNKPCS-LT	-81.047	0.105
HNKPCS-HP	-85.200	0.119
NKPC-TV	-46.162	0.142
NKPC-TV-LS	-61.972	0.138
NKPC-TV-SV	-22.761	0.134
NKPC-TV-LS-SV	-33.476	0.126
HNKPC-TV	-36.385	0.123
HNKPC-TV-LS	-35.052	0.105
HNKPC-TV-SV	-19.695	0.106
HNKPC-TV-LS-SV	-18.150	0.091
NKPCS-TV	-34.407	0.129
NKPCS-TV-LS	-32.004	0.099
NKPCS-TV-LS-SV	-15.390	0.092
HNKPCS-TV	-40.465	0.176
HNKPCS-TV-LS	-38.082	0.297
HNKPCS-TV-LS-SV	-12.977	0.139
BVAR (constant)	-166.226	0.085
BVAR-TV-SV	-97.980	0.100
SW2007	-78.033	0.168

Note: The table reports the predictive performances of alternative models for the period between the second quarter of 1973 and the first quarter of 2012. '(Log) Marg. Likelihood' stands for the natural logarithm of the marginal likelihoods. 'MSFE' stands for the Mean Squared Forecast Error. Marginal likelihood values in the first column are calculated as the sum of the predictive likelihood values in the prediction sample. Results are based on 10000 simulations of which the first 5000 are discarded for burn-in. Model abbreviations are based on Table K.1. BVAR (constant) denotes the BVAR model with 2 lags and with constant parameters. 'BVAR-TV-SV' denotes the 'BVAR' model with 2 lags, time varying levels for both series and stochastic volatility for inflation. SW2007 stands for the model proposed by Stock and Watson (2007).

use of the survey expectations data directly. We therefore conclude that replacing the expectational term in the (H)NKPC models with survey expectations deteriorate posterior inference compared to the iterative solution of these expectational terms.

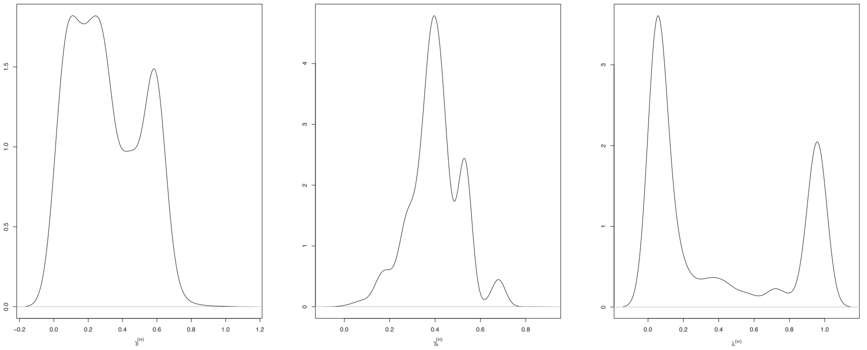
To conclude, predictive gains obtained from including the survey expectations in the models are substantial and incorporating the low and high frequency data movements in the model is crucial. These two conclusions are in line with Faust and Wright (2013), who consider a large set of alternative models for inflation forecasting, including unrestricted reduced form models, and compare their forecast performances based on MSFE. Our model incorporates both these features in the NKPC model structure. Third, once survey

Table K.3: Posterior results of alternative NKPC models: robustness check

Model	$\lambda^{(H)}$	$\gamma_f^{(H)}$	γ_b^H	ρ	ϕ_1	ϕ_2
NKPCS-LT	0.011 (0.051)	0.611 (0.055)	–	-0.016 (0.021)	0.824 (0.047)	0.075 (0.044)
NKPCS-HP	0.064 (0.051)	0.627 (0.081)	–	-0.045 (0.064)	0.681 (0.096)	0.014 (0.081)
HNKPCS-LT	0.154 (0.205)	0.350 (0.236)	0.408 (0.202)	-0.114 (0.155)	0.823 (0.058)	0.069 (0.057)
HNKPCS-HP	0.234 (0.235)	0.333 (0.180)	0.472 (0.154)	-0.216 (0.197)	0.614 (0.079)	-0.018 (0.057)
NKPCS-TV	0.057 (0.028)	0.142 (0.086)	–	-0.034 (0.061)	0.815 (0.052)	0.067 (0.052)
NKPCS-TV-LS	0.049 (0.023)	0.430 (0.125)	–	-0.027 (0.050)	0.821 (0.054)	0.072 (0.052)
NKPCS-TV-LS-SV	0.058 (0.025)	0.307 (0.165)	–	-0.015 (0.068)	0.826 (0.052)	0.078 (0.053)
HNKPCS-TV	0.383 (0.395)	0.308 (0.197)	0.401 (0.111)	-0.322 (0.349)	0.593 (0.314)	0.007 (0.100)
HNKPCS-TV-LS	0.557 (0.432)	0.375 (0.196)	0.393 (0.094)	-0.468 (0.367)	0.432 (0.328)	-0.031 (0.101)
HNKPCS-TV-LS-SV	0.151 (0.178)	0.216 (0.161)	0.368 (0.149)	-0.024 (0.095)	0.871 (0.027)	0.112 (0.032)

Note: Posterior results are based on 40000 simulations of which the first 20000 are discarded for burn-in. Model abbreviations are based on Table K.1.

Figure K.1: Posterior density of $\lambda^{(H)}$, $\gamma_b^{(H)}$ and $\gamma_f^{(H)}$ from the HNKPCS-TV model



Note: The figure presents posterior densities of parameters from the HNKPCS-TV model. Model abbreviations are based on Table K.1. Results are based on 40000 simulations of which the first 20000 are discarded for burn-in.

data and time variation are included in the model, there are still additional predictive gains from the backward looking component in the hybrid models.

4.L Modelling Inflation Expectations using Unobserved Components

The HNKPC models implicitly assume that survey based inflation expectations capture 'real' inflation expectations for the next period accurately. However, survey expectations are likely to reflect real inflation expectations with a measurement error. In this section we extend the HNKPC model by including a latent variable for unobserved inflation expectations, aiming to account for the possibility of measurement errors in survey expectations. Specifically, we propose an adaptive rule under which inflation expectations partially adjust to survey expectations at each period:

$$S_{t+1} = \mu_{t+1} + \beta_S(S_t - \mu_t) + \eta_{S,t+1}, \tag{4.L.1}$$

where $|\beta_S| < 1$ and μ_t is the survey observation for inflation expectation at time t. This adaptive rule implies that unobserved inflation expectations converge to the survey based expectations in the long run. Given the restriction on parameter β_S, one can solve (4.L.1) for S_t and obtain $S_t = \mu_t + \sum_{j=0}^{\infty} \beta_S^j \eta_{S,t-j}$. This specification allows for the interpretation that expected inflation is equal to the survey values with a measurement error that is specified as an infinite moving average with declining weights.

We next consider the HNKPC model given the specified adaptive rule for the unobserved inflation expectations. Notice that we can factorize the expectation term in equation (9) in the main text of the chapter, $E_t\left(\tilde{\pi}_{t+k}\right)$, into two parts related to the measurement error and the relation between survey based expectations and long run expectations, as $E_t\left(\tilde{\pi}_{t+k}\right) = E_t\left(S_{t+k-1} - \mu_{t+k-1}\right) + E_t\left(\mu_{t+k-1} - c_{\pi,t+k}\right)$. Then the weighted sum of expectations in equation (9) in the chapter becomes

$$\sum_{k=1}^{\infty} \gamma_f^k E_t\left(\tilde{\pi}_{t+k}\right) = \sum_{k=1}^{\infty} \gamma_f^k E_t\left(S_{t+k-1} - \mu_{t+k-1}\right) + \sum_{k=1}^{\infty} \gamma_f^k E_t\left(\mu_{t+k-1} - c_{\pi,t+k}\right). \qquad (4.L.2)$$

The first part of the summation, $\sum_{k=1}^{\infty} \gamma_f^k E_t\left(S_{t+k-1} - \mu_{t+k-1}\right)$, is related to the measurement error and can be computed from (4.L.1). For the second part of the summation, $\sum_{k=1}^{\infty} \gamma_f^k E_t\left(\mu_{t+k-1} - c_{\pi,t+k}\right)$, we specify a similar partial adjustment process as the process specified in the chapter $\mu_t - c_{\pi,t+1} = \beta_\mu(\mu_{t-1} - c_{\pi,t}) + \eta_{\mu,t+1}$. The partial adjustment mechanism implies that the further one gets into the future the smaller will be the difference between short and long run inflation expectations. Estimates of β_μ will indicate the empirical speed of adjustment. For instance, for a value of the posterior mean of β_μ equal to 0.5 it follows that within a few periods one has almost complete adjustment.

Replacing the infinite sum of expectations of inflation deviations using the two specifications for the measurement error and for the deviation of the survey expectations from the long run inflation expectations in (4.L.2), the HNKPC model becomes

$$\begin{aligned}
\pi_t - c_{\pi,t} &= \frac{\lambda^H}{(1-\gamma_b^H\gamma_f^H)(1-(\phi_1+\phi_2\gamma_f^H)\gamma_f^H)}\left(z_t - c_{z,t}\right) + \frac{\phi_2\gamma_f^H\lambda^H}{(1-\gamma_b^H\gamma_f^H)(1-(\phi_1+\phi_2\gamma_f^H)\gamma_f^H)}\left(z_{t-1} - c_{z,t-1}\right) \\
&+ \frac{\gamma_b^H\gamma_f^H}{(1-\gamma_b^H\gamma_f^H)}\left(\frac{\gamma_f^H}{1-\gamma_f^H\beta_S}(S_t-\mu_t) + \frac{\gamma_f^H}{1-\gamma_f^H\beta_\mu}(\mu_t - c_{\pi,t})\right) \\
&+ \frac{\gamma_b^H}{(1-\gamma_b^H\gamma_f^H)}\left(\pi_{t-1} - c_{\pi,t-1}\right) + \frac{1}{(1-\gamma_b^H\gamma_f^H)}\epsilon_{1,t}, \\
z_t - c_{z,t} &= \phi_1\left(z_{t-1} - c_{z,t-1}\right) + \phi_2\left(z_{t-2} - c_{z,t-2}\right) + \epsilon_{2,t}.
\end{aligned} \tag{4.L.3}$$

Notice that if the speed of adjustment for both specifications are equal, i.e. $\beta_S = \beta_\mu$, then the HNKPC reduces to

$$\begin{aligned}
\pi_t - c_{\pi,t} &= \frac{\lambda^H}{(1-\gamma_b^H\gamma_f^H)(1-(\phi_1+\phi_2\gamma_f^H)\gamma_f^H)}\left(z_t - c_{z,t}\right) + \frac{\phi_2\gamma_f^H\lambda^H}{(1-\gamma_b^H\gamma_f^H)(1-(\phi_1+\phi_2\gamma_f^H)\gamma_f^H)}\left(z_{t-1} - c_{z,t-1}\right) \\
&+ \frac{\gamma_b^H\gamma_f^H}{(1-\gamma_b^H\gamma_f^H)}\frac{\gamma_f^H}{1-\gamma_f^H\beta_S}(S_t - c_{\pi,t}) + \frac{\gamma_b^H}{(1-\gamma_b^H\gamma_f^H)}\left(\pi_{t-1} - c_{\pi,t-1}\right) + \frac{1}{(1-\gamma_b^H\gamma_f^H)}\epsilon_{1,t}, \\
z_t - c_{z,t} &= \phi_1\left(z_{t-1} - c_{z,t-1}\right) + \phi_2\left(z_{t-2} - c_{z,t-2}\right) + \epsilon_{2,t}.
\end{aligned} \tag{4.L.4}$$

We next compare the models specified in (4.L.3) and in (4.L.4) with a HNKPC-TV parametrization in terms of their forecast performances. For the forecast sample considered in the chapter, the cumulative predictive likelihood for the HNKPC-TV model in (4.L.3) is −36.19 while for the model in (4.L.4) this value is −36.44. The cumulative predictive likelihood values for the HNKPC-TV model with and without the restriction $\beta_S = \beta_\mu$ indicate that this restriction is statistically valid as the difference between the likelihood values are very small. Following this evidence we display the parameter estimates of all extended HNKPC models using the expectation specification in (4.L.4) in Table L.1. We further report the cumulative predictive likelihood values and 1 step ahead MSFE for these models in Table L.2.

Results are very similar to the corresponding table in the main text of the chapter (Table 2), thus, we conclude that the effect of the measurement errors in survey expectations is negligible.

Table L.1: Posterior results of HNKPC models with unobserved inflation expectations

Model	λ^H	γ_f^H	γ_b^H	β_S	ρ	ϕ_1	ϕ_2
HNKPC-TV	0.05 (0.03)	0.02 (0.03)	0.38 (0.14)	0.49 (0.28)	0.01 (0.06)	0.81 (0.05)	0.07 (0.05)
HNKPC-TV-LS	0.04 (0.02)	0.01 (0.01)	0.49 (0.11)	0.52 (0.18)	0.02 (0.01)	0.79 (0.09)	0.19 (0.08)
HNKPC-TV-LS-SV	0.06 (0.02)	0.04 (0.10)	0.22 (0.12)	0.44 (0.24)	-0.01 (0.01)	0.82 (0.05)	0.15 (0.04)

The table presents posterior means and standard deviations (in parentheses) of parameters for the competing HNKPC type models estimated for quarterly inflation and real marginal costs over the period 1960-I until 2012-I. λ^H and γ_f^H are the slope of the Phillips curve and the coefficient of inflation expectations in HNKPC model in (4.L.4). γ_b^H is the coefficient of the backward looking component in the HNKPC model in (4.L.4). β_S is the autoregressive parameter for the deviation of inflation expectations, as used in (4.L.4). ρ is the correlation coefficient of the residuals ϵ_1 and ϵ_2. ϕ_1 and ϕ_2 are the autoregressive parameters for the real marginal cost specification. Posterior results are based on 40000 simulations of which the first 20000 are discarded for burn-in. Model abbreviations are as in Table 1 in the main text of the chapter.

Table L.2: Predictive performance of HNKPC models with unobserved inflation expectations

Model	Cumulative (Log) Pred. Likelihood	MSFE 1 period ahead
HNKPC-TV	-36.44	0.12
HNKPC-TV-LS	-35.77	0.09
HNKPC-TV-LS-SV	-17.96	0.09

Note: The table reports the predictive performances of competing models for the prediction sample over the period 1973-II until 2012-I. 'Cumulative (Log) Pred. Likelihood' stands for the sum of the natural logarithms of predictive likelihoods. 'MSFE' stands for the Mean Squared Forecast Error. Results are based on 10000 simulations of which the first 5000 are discarded for burn-in. Remaining abbreviations are as in Table 1 in the chapter.

4.M Analysis of Cointegration in Inflation and Marginal Cost Levels

The models in the chapter considered rely on the implicit assumption of the absence of a long-run cointegrating relationship between the inflation and marginal cost series. We assess whether this assumption is plausible for the U.S. data. For this reason, we consider the NKPC-TV model that provides the unobserved levels of both series at each posterior draw. For each of these obtained posterior draws, we perform a simple two-step analysis to check the existence of the cointegrating relationship, which can be seen as a Bayesian extension of the method of Engle and Granger (1987).

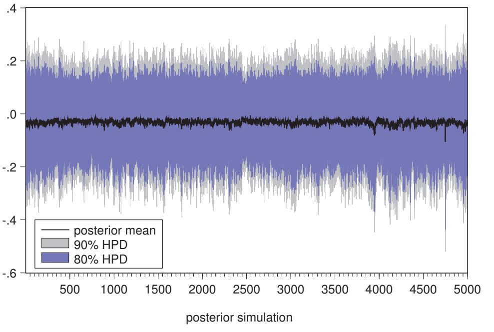
We perform a two step analysis, where in the first step we obtain the residuals from the regression of the estimated level of inflation on a constant and the estimated level of marginal costs, for each posterior draw. This implies that we take the estimation uncertainty in the analysis into account. Next, we obtain the posterior distribution of the autoregressive parameter, ρ, for each set of residuals from the following regression using flat priors on the identified region $\rho \in [-1, 1]$

$$\Delta\hat{\epsilon}_t = \rho\hat{\epsilon}_{t-1} + \eta_t, \;\; \eta_t \sim NID(0, \sigma^2), \tag{4.M.1}$$

where $\hat{\epsilon}_t$ denotes the residuals from the first stage, and $\rho = 0$ implies that there is no cointegrating relationship between the series. An HPDI including the value of 0 indicates that a cointegrating relation between inflation and marginal cost is unlikely.

We compute the mean and the quantiles of these individual densities using 5000 posterior draws, and report the average values of the mean and the quantiles of ρ based on 3000 simulations. These results are presented in Figure M.1. Posterior means of parameter ρ are around 0 for all posterior draws of inflation and marginal cost levels, and the 80% and 90% percent quantiles of the distribution are around 0 as well. Hence this simulation experiment does not indicate a cointegrating relationship between the inflation and marginal cost levels. This pattern is also found for other TV-NKPC models we considered for the U.S. data, but these results are not reported for the sake of brevity. We conclude that the underlying assumption of 'no cointegrating relationship' is found to be feasible for the NKPC models we consider.

Figure M.1: Cointegration analysis for the marginal costs and inflation series



Note: The figure presents the posterior means and quantiles of the ρ parameter from 5×10^3 posterior draws from the NKPC-TV models, where for each draw, the the reported values are calculated using 3000 simulations. $\rho = 0$ implies that there is no cointegrating relationship between the series.

Nederlandse Samenvatting (Summary in Dutch)

Dit proefschrift behandelt toepassing van Bayesiaanse econometrische methoden in de economische wetenschap, in het bijzonder in de macroeconomie, en wil specifiek de volgende drie vragen beantwoorden: 1) Hoe heeft Bayesiaanse econometrie zich de laatste decennia ontwikkeld? 2) Hoe kunnen groeiramingen voor economische variabelen verbeterd worden? 3) Hoe kunnen de analyses en voorspellingen van inflatieprognoses verbeterd worden? Deze vragen worden beantwoord in drie opeenvolgende hoofdstukken van dit proefschrift.

Hoofdstuk 2 beschrijft de geschiedenis van de Bayesiaanse econometrie sinds het begin van de jaren zestig (van de vorige eeuw). We kwantificeren de toenemende populariteit van Bayesiaanse econometrie op basis van een dataset dat bestaat uit ongeveer 1000 artikelen. Hiervoor tellen we het aantal publicatie- en citatie-records van deze artikelen die verschenen zijn in tien econometrische tijdschriften. De aantallen geven inzicht in de wortels van de Bayesiaanse econometrie en maken een voorspelling voor de toekomst mogelijk. De analyse wijst op een positieve toekomst voor Bayesiaanse econometrie wat toepassingsmogelijkheden op verschillende deelgebieden van de economie betreft. Daarnaast analyseert hoofdstuk 2 de verbindingen tussen onderwerpen en auteurs van de artikelen in de dataset met behulp van een bibliometrische techniek. Deze analyse geeft een beeld van de belangrijkste onderwerpen binnen de Bayesiaanse econometrie. De resultaten suggereren dat onderwerpen zoals 'modellen met niet waargenomen variabelen' en 'modellen met tijdsvarierende patronen' steeds belangrijker zijn geworden.

Gezien het belang van tijdsvarierende patronen in de economie richten hoodstuk 3 en 4 zich op het verbeteren van modellen voor het voorspellen van de groei van het Bruto-Binnenlands-Product (BBP) en het voorspellen van inflatie, waarbij gebruike wordt gemaakt van het tijdsvarierend karakter van veel economische reeksen. De klasse van Nieuw-Keynesiaanse Philips Curve (NKPC) modellen voor het voorspellen van inflatie maakt doorgaans gebruik van traditionele methoden voor het opschonen van data vooraf-gaand aan de analyse. Dit kan echter leiden tot onzuivere schattingen en voorspellingen. Om dit probleem aan te pakken en de voorspellingen van macro-economische modellen te verbeteren biedt dit proefschrift een Bayesiaanse methode voor het modelleren van het NKPC-model en van reële Amerikaanse BBP reeksen.

Hoofdstuk 3 begint met een eenvoudige uiteenzetting van de technische kwesties waarmee een Bayesiaanse econometrist geconfronteerd wordt, wanneer zij de reële groei van het Amerikaanse BBP wil ramen met behulp van een tijdsvarierend parameter model en met Bayesiaanse schattingsmethoden. Het gemiddelde van de groeireeks verandert namelijk in de loop van de 20e eeuw met structurele verschuivingen rond 1965 en 1984. Wij willen met deze typische kenmerken in de data rekening houden bij het modelleren van de reële BBP-reeksen. Daarom onderzoeken wij een tijdsvarierend parameter model voor het gemid-delde, met continue verschuivingen in het niveau van het gemiddelde. De Bayesiaaanse analyse wordt uitgevoerd met behulp van Gibbs sampling. Vervolgens breiden we de analyse uit met discrete structurele verschuivingen in het gemiddelde en met stochastis-che volatiliteit in de variantie. Een verdere uitbreiding is het gebruik van data over verwachtingen die gebaseerd zijn op enquête data. De voorspelkracht van alle modellen wordt vergeleken via Bayesiaanse voorspeltechnieken.

De klasse van Philips Curve modellene is populair in de economische wetenschap omdat deze het verband tussen inflatie en economische activiteit weergeeft, wat essentieel is om de werking van de economie te verklaren en te voorspellen. Hoofdstuk 4 biedt varianten van het NKPC model waarin zowel het veranderende gemiddelde van inflatie als ook de trend in de reeksen van marginale kosten expliciet zijn opgenomen in de structuur van het model. Vervolgens wordt een zgn. 'posterior-predictive' analyse uitgevoerd. In één van de varianten van het NKPC model, namelijk het hybride NKPC, wordt een terugwaarts kijkende term opgenomen, en het relatieve belang van terugwaarts en voorwaarts kijkend gedrag wordt geëvalueerd. In nog een andere variant gebruiken we inflatie-verwachtingen uit enquête data, met als doel de informatie in de aannemelijkheidsfunctie en de posterior dichtheid te verbeteren. De bereikte resultaten suggereren dat de precisie van de schattingen verbetert wanneer structurele tijdsvariatie in de reeks(en) expliciet gemodelleerd worden. Bovendien zijn de resultaten van het model, zowel voor het schatten als voor het voorspellen beter of tenminste even goed als bij de gangbare Bayesiaanse vectorautoregressieve en stochastische volatiliteits modellen.

Bibliography

Ahmed S, Levin A, Wilson BA. 2004. Recent U.S. macroeconomic stability: Good policies, Good practices, or Good luck? *The Review of Economics and Statistics* **86**: 824–832.

Aldrich J. 1995. R. A. Fisher and the making of maximum likelihood 1912-22. Discussion Paper Series In Economics And Econometrics 9504, Economics Division, School of Social Sciences, University of Southampton.

An S, Schorfheide F. 2007. Bayesian analysis of DSGE models. *Econometric Reviews* **26**: 113–172.

Anderson T. 1947. A note on a maximum-likelihood estimate. Cowles Foundation paper 21c, Cowles Commission. Reprinted in *Econometrica* (1947) 15(3): 241–244.

Anderson T, Rubin H. 1949. Estimation of the parameters of a single equation in a complete system of stochastic equations. Cowles Foundation paper 36a, Cowles Commission. Reprinted in *Annals of Mathematical Statistics* (1949) 20(1): 46–63.

Andrieu C, Doucet A. 2002. Particle filtering for partially observed Gaussian state space models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **64**: 827–836.

Andrieu C, Doucet A, Holenstein R. 2010. Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **72**: 269–342.

Antoniak CE. 1974. Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *The annals of statistics* : 1152–1174.

Ascari G. 2004. Staggered prices and trend inflation: Some nuisances. *Review of Economic Dynamics* **7**: 642–667.

Ascari G, Ropele T. 2007. Optimal monetary policy under low trend inflation. *Journal of Monetary Economics* **54**: 2568–2583.

Ascari G, Sbordone AM. 2013. The macroeconomics of trend inflation. Working papers, Federal Reserve Bank of New York Staff Reports 628, August 2013.

Avramov D. 2002. Stock return predictability and model uncertainty. *Journal of Financial Economics* **64**: 423–458.

Banbura M, Giannone D, Reichlin L. 2010. Large Bayesian vector auto regressions. *Journal of Applied Econometrics* **25**: 71–92.

Bartlett M. 1957. Comment on 'A Statistical Paradox' by D.V. Lindley. *Biometrika* **44**: 533–534.

Basistha A, Nelson CR. 2007. New measures of the output gap based on the forward-looking New Keynesian Phillips Curve. *Journal of Monetary Economics* **54**: 498–511.

Baştürk N, Çakmaklı C, Ceyhan ŞP, Van Dijk, HK. 2013a. Historical developments in Bayesian econometrics after Cowles Foundation Monographs 10, 14. Technical report, Tinbergen Institute Discussion Paper.

Baştürk N, Çakmaklı C, Ceyhan ŞP, Van Dijk, HK. 2013b. Posterior-predictive evidence on US inflation using extended Phillips Curve models with non-filtered data. Accepted in *Journal of Applied Econometrics*.

Basturk N, Ceyhan P, Van Dijk HK. 2014. Bayesian forecasting of US growth using basic time varying parameter models and expectations data.

Baştürk N, Hoogerheide L, Van Dijk, HK. 2013c. Measuring returns to education: Bayesian analysis using weak or invalid instrumental variables. Unpublished manuscript.

Bauwens L, Van Dijk, HK. 1990. Bayesian limited information analysis revisited. In Gabszewicz J, Richard J, Wolsey L (eds.) *Economic Decision-Making: Games, Econometrics and Optimisation: Contributions in Honour of Jacques H. Drèze*, chapter 18. North Holland, 385–424.

Benati L. 2008. Investigating inflation persistence across monetary regimes. *The Quarterly Journal of Economics* **123**: 1005–1060.

Berger J, Pericchi L. 1996. The intrinsic Bayes factor for linear models. *Bayesian statistics* **5**: 25–44.

Billio M, Casarin R, Ravazzolo F, Van Dijk, HK. 2013. Time-varying combinations of predictive densities using nonlinear filtering. *Journal of Econometrics* **177**: 213–232.

Blanchard O, Simon J. 2001. The long and large decline in U.S. output volatility. *Brookings Papers on Economic Activity* **32**: 135–174.
URL `http://ideas.repec.org/a/bin/bpeajo/v32y2001i2001-1p135-174.html`

Bos C, Mahieu R, Van Dijk, HK. 2000. Daily exchange rate behaviour and hedging of currency risk. *Journal of Applied Econometrics* **15**: 671–696.

Bowden R, Turkington D. 1990. *Instrumental variables*. Cambridge University Press.

Brainard W. 1967. Uncertainty and the effectiveness of policy. *The American Economic Review* **57**: pp. 411–425.

Çakmaklı C. 2012. Bayesian semiparametric dynamic Nelson-Siegel model. Working Paper Series 59-12, The Rimini Centre for Economic Analysis.

Çakmaklı C, Paap R, Van Dijk, DJ. 2011. Modeling and estimation of synchronization in multistate Markov-switching models. Working Papers 11–002/4, Tinbergen Institute.

Calvo G. 1983. Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics* **12**: 383–398.

Canova F. 2012. Bridging DSGE models and the raw data. Working Papers 635, Barcelona Graduate School of Economics.

Canova F, Ferroni F. 2011. Multiple filtering devices for the estimation of cyclical DSGE models. *Quantitative Economics* **2**: 73–98.

Cappé O, Douc R, Guillin A, Marin J, Robert C. 2008. Adaptive importance sampling in general mixture classes. *Statistics and Computing* **18**: 447–459.

Carlin B, Polson N, Stoffer D. 1992. A Monte Carlo approach to nonnormal and nonlinear state-space modeling. *Journal of the American Statistical Association* **87**: 493–500.

Carter CK, Kohn R. 1994. On Gibbs sampling for state space models. *Biometrika* **81**: 541–553.

Carter CK, Kohn R. 1996. Markov Chain Monte Carlo in conditionally Gaussian state space models. *Biometrika* **83**: 589–601. ISSN 00063444.

Cecchetti SG, Hooper P, Kasman BC, Schoenholtz KL, Watson MW. 2007. Understanding the evolving inflation process. In *US Monetary Policy Forum*, volume 8.

Chernoff H. 1954. Rational selection of decision functions. Cowles Foundation paper 91, Cowles Commission. Reprinted in *Econometrica* (1954), 22: 422–443.

Chernoff H, Divinsky N. 1953. The computation of maximum-likelihood estimates of linear structural equations. *Studies in Econometric Method* **14**: 236–302.

Chib S, Hamilton B. 2002. Semiparametric Bayes analysis of longitudinal data treatment models. *Journal of Econometrics* **110**: 67–89. ISSN 0304–4076.

Christiano LJ, Eichenbaum M, Evans CL. 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy* **113**: 1–45.

Christiano LJ, Vigfusson RJ. 2003. Maximum likelihood in the frequency domain: the importance of time-to-plan. *Journal of Monetary Economics* **50**: 789–815.

Clarida R, Galí J, Gertler M. 2000. Monetary policy rules and macroeconomic stability: Evidence and some theory. *The Quarterly Journal of Economics* **115**: 147–180. URL `http://ideas.repec.org/a/tpr/qjecon/v115y2000i1p147-180.html`

Cogley T, Sbordone AM. 2008. Trend inflation, indexation, and inflation persistence in the New Keynesian Phillips Curve. *The American Economic Review* **98**: 2101–2126.

Conley T, Hansen C, McCulloch R, Rossi P. 2008. A semi-parametric Bayesian approach to the instrumental variable problem. *Journal of Econometrics* **144**: 276–305.

Cremers K. 2002. Stock return predictability: A Bayesian model selection perspective. *Review of Financial Studies* **15**: 1223–1249.

De Finetti B. 1989. Probabilism. *Erkenntnis* **31**: 169–223.

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

De Pooter M, Ravazzolo F, Segers R, Van Dijk, HK. 2009. Bayesian near-boundary analysis in basic macroeconomic time-series models. In Chib S, Koop G, Griffiths W, Terrell D (eds.) *Advances in Econometrics, (Bayesian econometrics)*, volume 23. Emerald Group Publishing Limited, 331–402.

Del Negro M, Schorfheide F. 2004. Priors from general equilibrium models for VARs. *International Economic Review* **45**: 643–673. ISSN 00206598.

Del Negro M, Schorfheide F. 2008. Forming priors for DSGE models (and how it affects the assessment of nominal rigidities). *Journal of Monetary Economics* **55**: 1191–1208.

Del Negro M, Schorfheide F. 2013. DSGE model-based forecasting. In Elliott G, Timmermann A (eds.) *Handbook of Economic Forecasting*, volume 2 of *Handbook of Economic Forecasting*. Elsevier, 57–140.

Delle Monache D, Harvey AC. 2011. The effect of misspecification in models for extracting trends and cycles. Working Papers EWP 2011/013, Euroindicators.

Doan T, Litterman R, Sims C. 1984. Forecasting and conditional projection using realistic prior distributions. *Econometric reviews* **3**: 1–100.

Drèze J. 1976. Bayesian limited information analysis of the simultaneous equations model. *Econometrica* **44**: 1045–75.

Drèze J. 1977. Bayesian regression analysis using poly-t densities. *Journal of Econometrics* **6**: 329–354.

Drèze J, Richard J. 1983. Bayesian analysis of simultaneous equation systems. In Griliches Z, Intriligator M (eds.) *Handbook of Econometrics*, volume 1 of *Handbook of Econometrics*, chapter 9. Elsevier, 517–598.

Durham G, Geweke J. 2013. Adaptive sequential posterior simulators for massively parallel computing environments. Working Paper Series 9, Economics Discipline Group, UTS Business School, University of Technology, Sydney.

Engle RF, Granger CWJ. 1987. Co-integration and error correction: Representation, estimation, and testing. *Econometrica* : 251–276.

Erceg CJ, Levin AT. 2003. Imperfect credibility and inflation persistence. *Journal of Monetary Economics* **50**: 915–944.

Escobar M, West M. 1995. Bayesian density estimation and inference using mixtures. *Journal of the american statistical association* **90**: 577–588.

Faust J, Wright JH. 2013. Forecasting inflation. In Elliott G, Timmermann A (eds.) *Handbook of Economic Forecasting*, volume 2 of *Handbook of Economic Forecasting*. Elsevier, 2–56.

Ferguson TS. 1973. A Bayesian analysis of some nonparametric problems. *The annals of statistics* : 209–230.

Fernandez C, Ley E, Steel M. 2001. Model uncertainty in cross-country growth regressions. *Journal of Applied Econometrics* **16**: 563–576. ISSN 08837252.

Fernandez-Villaverde J, Rubio-Ramirez J. 2007. Estimating macroeconomic models: A likelihood approach. *The Review of Economic Studies* **74**: 1059–1087. ISSN 00346527.

Fernandez-Villaverde J, Rubio-Ramirez J. 2008. How structural are structural parameters? *NBER Chapters* : 83–137.

Ferroni F. 2011. Trend agnostic one-step estimation of DSGE models. *The B.E. Journal of Macroeconomics* **11**: 1–36.

Fisher R. 1912. On an absolute criterion for fitting frequency curves. *Messenger of Mathmatics* **41**: 155–160.

Fisher R. 1922. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society* **A**: 309–368.

Fisher R. 1973. Statistical methods and scientific inference hafner. *New York* .

Florens J, Simoni A. 2012. Nonparametric estimation of an instrumental regression: a quasi-Bayesian approach based on regularized posterior. *Journal of Econometrics* **170**: 458–475.

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

Frühwirth-Schnatter S. 2006. *Finite Mixture and Markov Switching Models: Modeling and Applications to Random Processes*. Springer Series in Statistics. Springer.

Fuhrer J, Moore G. 1995. Inflation persistence. *The Quarterly Journal of Economics* **110**: 127–59.

Gali J. 2008. *Monetary Policy, Inflation and the Business Cycle: An Introduction to the New Keynesian Framework*. The USA: Princeton University Press, 1 edition.

Galí J, Gertler M. 1999. Inflation dynamics: A structural econometric analysis. *Journal of Monetary Economics* **44**: 195–222.

Galí J, Gertler M, Lopez-Salido JD. 2005. Robustness of the estimates of the hybrid New Keynesian Phillips Curve. *Journal of Monetary Economics* **52**: 1107–1118.

Gelfand A, Dey D. 1994. Bayesian model choice: asymptotics and exact calculations. *Journal of the Royal Statistical Society. Series B (Methodological)* : 501–514.

Gelfand A, Smith A. 1990. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* **85**: 398–409.

Geman S, Geman D. 1984. Stochastic relaxations, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transaction on Pattern Analysis and Machine Intelligence* **6**: 721–741.

George E, McCulloch R. 1993. Variable selection via Gibbs sampling. *Journal of the American Statistical Association* **88**: 881–889.

Gerlach R, Carter C, Kohn R. 2000. Efficient Bayesian inference for dynamic mixture models. *Journal of the American Statistical Association* : 819–828.

Geweke J. 1989. Bayesian inference in econometric models using Monte Carlo integration. *Econometrica* **57**: 1317–39.

Geweke J. 2005. *Contemporary Bayesian econometrics and statistics*. Wiley.

Geweke J. 2010. *Complete and Incomplete Econometric Models (The Econometric and Tinbergen Institutes Lectures)*. Princeton University Press.

Geweke J, Amisano G. 2010. Comparing and evaluating Bayesian predictive distributions of asset returns. *International Journal of Forecasting* **26**: 216–230.

Geweke J, Amisano G. 2011. Hierarchical Markov normal mixture models with applications to financial asset returns. *Journal of Applied Econometrics* **26**: 1–29.

Geweke J, Koop G, Van Dijk, HK (eds.) . 2011. *The Oxford handbook of Bayesian econometrics*. Oxford University Press.

Giordani P, Kohn R. 2008. Efficient Bayesian inference for multiple change-point and mixture innovation models. *Journal of Business & Economic Statistics* **26**: 66–77.

Goldberger A. 1972. Structural equation methods in the social sciences. *Econometrica* **40**: 979–1001.

Gordon N, Salmond D, Smith A. 1993. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. In *IEE Proceedings F (Radar and Signal Processing)*, volume 140. IET, 107–113.

Griffin J, Steel M. 2004. Semiparametric Bayesian inference for stochastic frontier models. *Journal of Econometrics* **123**: 121–152.

Haavelmo T. 1944. The probability approach in econometrics. Cowles Foundation paper 4, Cowles Commission. Reprinted in *Econometrica* (1944) 12: 1–118.

Hammersley J, Handscomb D. 1964. *Monte Carlo methods*. Chapman & Hall, London.

Hansen K, Singh V, Chintagunta P. 2006. Understanding store-brand purchase behavior across categories. *Marketing Science* **25**: pp. 75–90. ISSN 07322399.

Harvey A, Trimbur T, Van Dijk, HK. 2007. Trends and cycles in economic time series: A Bayesian approach. *Journal of Econometrics* **140**: 618–649.

Harvey AC. 1989. *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press.

Harvey AC. 1990. *Forecasting, structural time series models and the Kalman filter*. Cambridge university press.

Hastings WK. 1970. Monte Carlo sampling using Markov chains and their applications. *Biometrika* **57**: 97–109.

Hildreth C. 1963. Bayesian statisticians and remote clients. *Econometrica* **31**: 422–438.

Hirano K. 2002. Semiparametric Bayesian inference in autoregressive panel data models. *Econometrica* **70**: 781–799. ISSN 00129682.

Hirsch J. 2005. An index to quantify an individual's scientific research output. *Proceedings of the National academy of Sciences of the United States of America* **102**: 165–169.

Hobert JP, Casella G. 1996. The effect of improper priors on Gibbs sampling in hierarchical linear mixed models. *Journal of the American Statistical Association* **91**: 1461–1473.

Hood W, Koopmans T (eds.) . 1953. *Studies in Econometric Method*. John Wiley & Sons. Cowles Commission Monograph No. 14.

Hoogerheide L, Kaashoek J, Van Dijk, HK. 2007. On the shape of posterior densities and credible sets in instrumental variable regression models with reduced rank: An application of flexible sampling methods using neural networks. *Journal of Econometrics* **139**: 154–180. ISSN 0304–4076.

Hoogerheide L, Opschoor A, Van Dijk, HK. 2012. A class of adaptive importance sampling weighted EM algorithms for efficient and robust posterior and predictive simulation. *Journal of Econometrics* **171**: 101–120. ISSN 0304-4076.

Hurwicz L. 1950. Bayes and minimax interpretation of the maximum likelihood estimation criterion. Cowles Commission Discussion Paper Economics 352, Cowles Commission.

Ireland PN. 2007. Changes in the Federal Reserve's inflation target: Causes and consequences. *Journal of Money, Credit and Banking* **39**: 1851–1882.

Ishwaran H, Rao J. 2005. Spike and slab variable selection: frequentist and Bayesian strategies. *Annals of Statistics* : 730–773.

Jacquier E, Polson N, Rossi P. 1994. Bayesian analysis of stochastic volatility models. *Journal of Business & Economic Statistics* **12**: 371–389.

Jeanne O. 1998. Generating real persistent effects of monetary shocks: How much nominal rigidity do we really need? *European Economic Review* **42**: 1009–1032.

Jensen M. 2004. Semiparametric Bayesian inference of long-memory stochastic volatility models. *Journal of Time Series Analysis* **25**: 895–922.

Jensen M, Maheu J. 2010. Bayesian semiparametric stochastic volatility modeling. *Journal of Econometrics* **157**: 306–316. ISSN 0304–4076.

Kadiyala K, Karlsson S. 1997. Numerical methods for estimation and inference in Bayesian VAR-models. *Journal of Applied Econometrics* **12**: 99–132. ISSN 08837252.

Kastner G, Frühwirth-Schnatter S. 2013. Ancillarity-sufficiency interweaving strategy (asis) for boosting MCMC estimation of stochastic volatility models. *Computational Statistics & Data Analysis* .

Kim CJ, Nelson CR. 1999a. Has the u.s. economy become more stable? a Bayesian approach based on a markov-switching model of the business cycle. *The Review of*

Economics and Statistics **81**: pp. 608–616. ISSN 00346535.
URL http://www.jstor.org/stable/2646710

Kim CJ, Nelson CR. 1999b. *State-space models with regime switching: Classical and Gibbs-Sampling approaches with applications*, volume 1. The MIT Press, 1 edition.

Kim CJ, Nelson CR, Piger J. 2004. The less-volatile US economy: a Bayesian investigation of timing, breadth, and potential explanations. *Journal of Business & Economic Statistics* **22**: 80–93.

Kim S, Shephard N, Chib S. 1998. Stochastic volatility: Likelihood inference and comparison with ARCH models. *Review of Economic Studies* **65**: 361–393.

Kleibergen F, Mavroeidis S. 2011. Identification robust priors for Bayesian analysis in DSGE models. Manuscript.

Kleibergen F, Van Dijk, HK. 1993. Non-stationarity in GARCH models: A Bayesian analysis. *Journal of Applied Econometrics* **8**: 41–61. ISSN 08837252.

Kleibergen F, Van Dijk, HK. 1994. On the shape of the likelihood/posterior in cointegration models. *Econometric Theory* **10**: 514–551.

Kleibergen F, Van Dijk, HK. 1998. Bayesian simultaneous equations analysis using reduced rank structures. *Econometric Theory* **14**: 701–743.

Kloek T, Van Dijk, HK. 1975. *Bayesian estimates of equation system parameters: An unorthodox application of Monte Carlo*. Erasmus University.

Kloek T, Van Dijk, HK. 1978. Bayesian estimates of equation system parameters: An application of integration by Monte Carlo. *Econometrica* **46**: 1–19.

Koop G, Poirier D, Tobias J. 2007. *Bayesian Econometric Methods*, volume 7. Cambridge University Press.

Koopman S, Durbin J. 2000. Fast filtering and smoothing for multivariate state space models. *Journal of Time Series Analysis* **21**: 281–296.

Koopmans T. 1945. Statistical estimation of simultaneous economic relations. *Journal of the American Statistical Association* **40**: 448–466.

Koopmans T (ed.) . 1950. *Statistical inference in dynamic economic models*. John Wiley & Sons. Cowles Commission Monograph No. 10.

Lancaster T. 2004. *An Introduction to Modern Bayesian Econometrics*. Blackwell Oxford.

Levin A, Yun T. 2007. Reconsidering the natural rate hypothesis in a New Keynesian framework. *Journal of Monetary Economics* **54**: 1344–1365.

Lindley D. 1957. A statistical paradox. *Biometrika* **44**: 187–192.

Lindley D, El-Sayyad G. 1968. The Bayesian estimation of a linear functional relationships. *Journal of the Royal Statistical Society. Series B (Methodological)* : 190–202.

Liu Z, Waggoner DF, Zha T. 2011. Sources of macroeconomic fluctuations: A regime-switching DSGE approach. *Quantitative Economics* **2**: 251–301.

Mankiw NG, Reis R, Wolfers J. 2003. Disagreement about inflation expectations. In *NBER Macroeconomics Annual 2003*, volume 18 of *NBER Chapters*. National Bureau of Economic Research, Inc, 209–270.

McCloskey D, Ziliak S. 1996. The standard error of regressions. *Journal of Economic Literature* **34**: 97–114.

McConnell MM, Perez-Quiros G. 2000. Output fluctuations in the united states: What has changed since the early 1980's? *The American Economic Review* **90**: pp. 1464–1476. ISSN 00028282.
URL `http://www.jstor.org/stable/2677860`

McCulloch R, Tsay R. 1994. Bayesian inference of trend- and difference-stationarity. *Econometric Theory* **10**: 596–608.

Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. 1953. Equations of state calculations by fast computing machines. *Journal of Chemical Physics* **21**: 1087–1092.

Milani F. 2011. Expectation shocks and learning as drivers of the business cycle. *Economic Journal* **121**: 379–401.
URL `http://EconPapers.repec.org/RePEc:ecj:econjl:v:121:y:2011:i:552:p:379-401`

Min C, Zellner A. 1993. Bayesian and non-Bayesian methods for combining models and forecasts with applications to forecasting international growth rates. *Journal of Econometrics* **56**: 89–118. ISSN 0304-4076.

Nason JM, Smith GW. 2008. The New Keynesian Phillips curve: lessons from single-equation econometric estimation. *Economic Quarterly* **94**: 361–395.

Neal R. 2000. Markov Chain sampling methods for Dirichlet process mixture models. *Journal of computational and graphical statistics* **9**: 249–265.

O'Hagan A. 1995. Fractional Bayes factors for model comparison. *Journal of the Royal Statistical Society. Series B (Methodological)* : 99–138.

Omori Y, Chib S, Shephard N, Nakajima J. 2007. Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics* **140**: 425–449.

Paap R, Van Dijk, HK. 1998. Distribution and mobility of wealth of nations. *European Economic Review* **42**: 1269–1293.

Pagan AR, Harding D. 2005. A suggested framework for classifying the modes of cycle research. *Journal of Applied Econometrics* **20**: 151–159.

URL `http://EconPapers.repec.org/RePEc:jae:japmet:v:20:y:2005:i:2:p:151-159`

Park T, Casella G. 2008. The Bayesian lasso. *Journal of the American Statistical Association* **103**: 681–686.

Penelope S, Summers PM. 2009. Regime switches in GDP growth and volatility: Some international evidence and implications for modeling business cycles. *The B.E. Journal of Macroeconomics* **9**: 1–19.

Phillips P. 1991. To criticize the critics: An objective Bayesian analysis of stochastic trends. *Journal of Applied Econometrics* **6**: 333–364. ISSN 08837252.

Pitt M, Shephard N. 1999. Filtering via simulation: auxiliary particle filters. *Journal of the American Statistical Association* **94**: 590–599.

Poirier D. 1989. A report from the battlefront. *Journal of Business & Economic Statistics* **7**: 137–139.

Poirier D. 1992. A return to the battlefront. *Journal of Business & Economic Statistics* **10**: 473–474.

Pratt J, Raiffa H, Schlaifer R. 1995. *Introduction to statistical decision theory*. MIT press.

Raiffa H, Schlaifer R (eds.) . 1961. *Applied statistical decision theory*. Harvard University Press.

Ritter C, Tanner MA. 1992. Facilitating the Gibbs sampler: The Gibbs stopper and the griddy-Gibbs sampler. *Journal of the American Statistical Association* **87**: 861–868.

Robert C, Casella G. 2004. *Monte Carlo statistical methods*. Springer Verlaag.

Rossi P, Allemby G, McCulloch R. 2005. *Bayesian statistics and marketing*. Wiley Series in Probability and Statistics.

Rossi P, McCulloch R, Allenby G. 1996. The value of purchase history data in target marketing. *Marketing Science* **15**: 321–340.

Rothenberg T. 1963. A Bayesian analysis of simultaneous equation system. Econometric Institute Report 6315, Erasmus University Rotterdam.

Rothenberg T. 1973. *Efficient estimation with a priori information*. Yale University Press. Cowles Commission Monograph No. 23.

Rudd J, Whelan K. 2005. New tests of the New Keynesian Phillips Curve. *Journal of Monetary Economics* **52**: 1167–1181.

Rudd J, Whelan K. 2007. Modeling inflation dynamics: A critical review of recent research. *Journal of Money, Credit and Banking* **39**: 155–170.

Sargan J. 1958. The estimation of economic relationships using instrumental variables. *Econometrica: Journal of the Econometric Society* **26**: 393–415.

Sargent T, Williams N, Zha T. 2006. Shocks and government beliefs: The rise and fall of American inflation. *American Economic Review* **96**: 1193–1224.

Savage L. 1961. *The subjective basis of statistical practice*. University of Michigan.

Schlaifer R. 1959. *Probability and statistics for business decisions: An introduction to magagerial economics under uncertainty*. McGraw-Hill.

Schorfheide F. 2005. Learning and monetary policy shifts. *Review of Economic Dynamics* **8**: 392–419.

Schotman P, Van Dijk, HK. 1991a. A Bayesian analysis of the unit root in real exchange rates. *Journal of Econometrics* **49**: 195–238. ISSN 030-4-4076.

Schotman P, Van Dijk, HK. 1991b. On Bayesian routes to unit roots. *Journal of Applied Econometrics* **6**: 387–401. ISSN 08837252.

Sethuraman J. 1994. A constructive definition of Dirichlet priors. *Statistica Sinica* **4**: 639–650.

Sims C. 2007. Bayesian methods in applied econometrics, or, why econometrics should always and everywhere be Bayesian.

Sims C, Uhlig H. 1991. Understanding unit rooters: A helicopter tour. *Econometrica* **59**: 1591–1599. ISSN 00129682.

Sims CA. 2002. Solving linear rational expectations models. *Computational Economics* **20**: 1–20.

Sims CA, Zha T. 2006. Were there regime switches in U.S. monetary policy? *American Economic Review* **96**: 54–81.

Smets F, Wouters R. 2003. An estimated Dynamic Stochastic General Equilibrium model of the Euro Area. *Journal of the European Economic Association* **1**: 1123–1175.

Smets F, Wouters R. 2007. Shocks and frictions in US business cycles: A Bayesian DSGE approach. *American Economic Review* **97**: 586–606.

Stock JH, Watson MW. 2007. Why has U.S. inflation become harder to forecast? *Journal of Money, Credit and Banking* **39**: 3–33.

Stock JH, Watson MW. 2008. Phillips Curve inflation forecasts. Working Paper 14322, National Bureau of Economic Research.

Strachan R, Van Dijk, HK. 2013. Evidence on features of a DSGE business cycle model from Bayesian model averaging. *International Economic Review* **54**: 385–402.

Tanner MA, Wong WH. 1987. The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association* **82**: 528–550.

The Economist. 2004. Sugnifying nothing? *Economics focus* January 31st, pp. 63.

Tierney L. 1994. Markov Chains for exploring posterior distributions. *the Annals of Statistics* : 1701–1728.

Van Dijk, HK. 2013. Bridging two key issues in Bayesian inference: The relationship between the Lindley paradox and non-elliptical credible sets. In Singpurwalla N, Dawid P, O'Hagan A (eds.) *Festschrift for Dennis Lindley's ninetienth birthday.*

Van Dijk, HK, Kloek T. 1980. Further experience in Bayesian analysis using Monte Carlo integration. *Journal of Econometrics* **14**: 307–328.

Van Dijk, HK, Kloek T. 1985. Experiments with some alternatives for simple importance sampling in Monte Carlo integration. In Bernardo J, Degroot M, Lindley D, Smith A (eds.) *Bayesian Statistics*, volume 2. North-Holland, Amsterdam, 511–530.

Van Eck N, Waltman L. 2010. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* **84**: 523–538.

Walker S. 2007. Sampling the Dirichlet mixture model with slices. *Communications in Statistics, Simulation and Computation* **36**: 45–54.

Waltman L, Van Eck N, Noyons E. 2010. A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics* **4**: 629–635.

Woodford M. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton University Press.

Wright J. 2008. Bayesian model averaging and exchange rate forecasts. *Journal of Econometrics* **146**: 329–341. ISSN 0304-4076. Honoring the research contributions of Charles R. Nelson.

Yun T. 1996. Nominal price rigidity, money supply endogeneity, and business cycles. *Journal of Monetary Economics* **37**: 345–370.

Zellner A. 1971. *An introduction to Bayesian inference in econometrics.* Wiley.

Zellner A, Ando T, Baştürk N, Hoogerheide L, Van Dijk, HK. 2014. Bayesian analysis of instrumental variable models: Acceptance-rejection within Direct Monte Carlo. *Econometric Reviews* **33**: 3–35.

Zellner A, Bauwens L, Van Dijk, HK. 1988. Bayesian specification analysis and estimation of simultaneous equation models using Monte Carlo methods. *Journal of Econometrics* **38**: 39–72.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

540 E. ERDOGAN CIFTCI, *Health Perceptions and Labor Force Participation of Older Workers.*
541 T.WANG, *Essays on Empirical Market Microstructure.*
542 T. BAO, *Experiments on Heterogeneous Expectations and Switching Behavior.*
543 S.D. LANSDORP, *On Risks and Opportunities in Financial Markets.*
544 N. MOES, *Cooperative decision making in river water allocation problems.*
545 P. STAKENAS, *Fractional integration and cointegration in financial time series.*
546 M. SCHARTH, *Essays on Monte Carlo Methods for State Space Models.*
547 J. ZENHORST, *Macroeconomic Perspectives on the Equity Premium Puzzle.*
548 B. PELLOUX, *the Role of Emotions and Social Ties in Public On Good Games: Behavioral and Neuroeconomic Studies.*
549 N. YANG, *Markov-Perfect Industry Dynamics: Theory, Computation, and Applications.*
550 R.R. VAN VELDHUIZEN, *Essays in Experimental Economics.*
551 X. ZHANG, *Modeling Time Variation in Systemic Risk.*
552 H.R.A. KOSTER, *The internal structure of cities: the economics of agglomeration, amenities and accessibility.*
553 S.P.T. GROOT, *Agglomeration, globalization and regional labor markets: micro evidence for the Netherlands.*
554 J.L. MÖHLMANN, *Globalization and Productivity Micro-Evidence on Heterogeneous Firms, Workers and Products.*
555 S.M. HOOGENDOORN, *Diversity and Team Performance: A Series of Field Experiments.*
556 C.L. BEHRENS, *Product differentiation in aviation passenger markets: The impact of demand heterogeneity on competition.*
557 G. SMRKOLJ, *Dynamic Models of Research and Development.*
558 S. PEER, *The economics of trip scheduling, travel time variability and traffic information.*
559 V. SPINU, *Nonadditive Beliefs: From Measurement to Extensions.*
560 S.P. KASTORYANO, *Essays in Applied Dynamic Microeconometrics.*
561 M. VAN DUIJN, *Location, choice, cultural heritage and house prices.*
562 T. SALIMANS, *Essays in Likelihood-Based Computational Econometrics.*
563 P. SUN, *Tail Risk of Equidity Returns.*
564 C.G.J. KARSTEN, *The Law and Finance of M&A Contracts.*

565 C. OZGEN, *Impacts of Immigration and Cultural Diversity on Innovation and Economic Growth.*

566 R.S. SCHOLTE, *The interplay between early-life conditions, major events and health later in life.*

567 B.N.KRAMER, *Why do not they take a card? Essays on the demand for micro health insurance.*

568 M. KILIÇ, *Fundamental Insights in Power Futures Prices.*

569 A.G.B. DE VRIES, *Venture Capital: Relations with the Economy and Intellectual Property.*

570 E.M.F. VAN DEN BROEK, *Keeping up Appearances.*

571 K.T. MOORE, *A Tale of Risk: Essays on Financial Extremes.*

572 F.T. ZOUTMAN, *A Symphony of Redistributive Instruments.*

573 M.J. GERRITSE, *Policy Competition and the Spatial Economy.*

574 A. OPSCHOOR, *Understanding Financial Market Volatility.*

575 R.R. VAN LOON, *Tourism and the Economic Valuation of Cultural Heritage.*

576 I.L. LYUBIMOV, *Essays on Political Economy and Economic Development.*

577 A.A.F. GERRITSEN, *Essays in Optimal Government Policy.*

578 M.L. SCHOLTUS, *The Impact of High-Frequency Trading on Financial Markets.*

579 E. RAVIV, *Forecasting Financial and Macroeconomic Variables: Shrinkage, Dimension reduction, and Aggregation.*

580 J. TICHEM, *Altruism, Conformism, and Incentives in the Workplace.*

581 E.S. HENDRIKS, *Essays in Law and Economics.*

582 X. SHEN, *Essays on Empirical Asset Pricing.*

583 L.T. GATAREK, *Econometric Contributions to Financial Trading, Hedging and Risk Measurement.*

584 X. LI, *Temporary Price Deviation, Limited Attention and Information Acquisition in the Stock Market.*

585 Y. DAI, *Efficiency in Corporate Takeovers.*

586 S.L. VAN DER STER, *Approximate feasibility in real-time scheduling: Speeding up in order to meet deadlines.*

587 A. SELIM, *An Examination of Uncertainty from a Psychological and Economic Viewpoint.*

588 B.Z. YUESHEN, *Frictions in Modern Financial Markets and the Implications for Market Quality.*

589 D. VAN DOLDER, *Game Shows, Gambles, and Economic Behavior.*

This thesis deals with the application of Bayesian econometrics for macroeconomic modelling. In particular, this thesis aims to answer three questions: 1) How did Bayesian econometrics evolve over time? 2) How can output growth forecasts be improved? 3) How can inflation forecasts be improved? These questions are answered in three consecutive chapters of this thesis.

Chapter 2 describes the history of Bayesian econometrics since the early 1960s. It quantifies the increasing popularity of Bayesian econometrics by analyzing the publication and citation records of papers in ten econometrics journals. These numbers give insight into the roots of Bayesian econometrics and a prediction about its future. Additionally, this chapter examines the connections among the topics and authors of the papers in the data set using the bibliometric mapping technique. Given the importance of time varying patterns suggested by these analyses, the following two chapters aim to improve models for forecasting GPD growth and inflation taking into account the time varying behavior of the series.

Chapter 3 starts with a basic exposition of the technical issues that a Bayesian econometrician faces in terms of modeling and inference when she is interested in forecasting US real GDP growth by using a time varying parameter model using simulation based Bayesian inference. It then proposes models for the US real GDP growth series in level and volatility dimensions. New Keynesian Phillips Curve models used for inflation forecasting typically rely on traditional ways of cleaning data before analysis. However, this may lead to poor performance. Therefore, motivated to fill in this gap in the literature and improve model performance, Chapter 4 proposes models for the NKPC model for the US in a Bayesian way.

S. Pinar Ceyhan obtained a BA degree in Business Administration and MA degree in Economics in 2007 from Bogazici University in Istanbul, Turkey. Following this, she started the MPhil program in Economics at the Tinbergen Institute in 2008. Upon completing the MPhil program in 2010, she joined the PhD program at the Econometric Institute of Erasmus University Rotterdam.