UNIVERSITÉ DE STRASBOURG
FACULTÉ DES SCIENCES ÉCONOMIQUES ET DE GESTION

THÈSE

présentée pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE STRASBOURG

Spécialité : Sciences Économiques

par

Frédo JEAN-BAPTISTE

Essays on the Econometric Evaluation of Monetary Business Cycle Models

JURY

Frédéric DUFOURT	Professeur, Université de Strasbourg	Directeur de Thèse
Roger FARMER	Professeur, UCLA	Rapporteur
Patrick FÈVE	Professeur, Université de Toulouse I	Rapporteur
François LAISNEY	Professeur, Université de Strasbourg	Examinateur
Thepthida SOPRASEUTH	Professeur, Université du Maine	Présidente

02 Décembre 2011

REMERCIEMENTS

Voici venu le temps des remerciements. C'est un exercice que le doctorant entreprend avec joie (car la fin est proche), mais aussi avec émotion (car c'est le moment d'apprécier par écrit la contribution de nombreuses personnes au produit final qu'est la thèse).

Mes premiers remerciements vont naturellement à Frédéric Dufourt, mon directeur de thèse. J'ai pu découvrir ses très nombreuses qualités humaines et intellectuelles durant ces quatre années. Je souhaite le remercier en particulier pour sa pédagogie, ses compétences, sa rigueur, sa disponibilité, sa patience. Alors que cette thèse s'approchait de sa fin, il était toujours là, n'hésitant pas à interrompre ses vacances pour m'aider à clarifier certaines idées développées dans la thèse.

Je voudrais également remercier François Laisney. J'ai pu bénéficier de son enseignement en économétrie. J'ai donc une immense dette intellectuelle envers lui à propos du contenu économétrique de cette thèse. De plus, il m'a fait l'honneur de participer à mon jury de thèse. Merci.

Mes remerciements s'adressent à mes deux rapporteurs, Roger Farmer et Patrick Fève. C'est une lourde tâche d'être rapporteur, tâche que vous avez acceptée avec plaisir malgré les différents imprévus qui sont survenus lors de la fin de cette thèse.

Je souhaite remercier Theptidha Sopraseuth, de m'avoir fait l'honneur de participer à mon jury de thèse.

Cette thèse aurait eu du mal à aboutir sans le support informatique de plusieurs personnes de la Faculté des Sciences Économiques et de Gestion. Je voudrais remercier tout spécialement Kene Boun My, Marc Diebold, Bertrand Koebel, Frédéric Olland, Audrey-Rose Ménard. Sans votre aide, il aurait été tout simplement impossible d'aboutir à tous les résultats du chapitre 5 de la thèse.

Mes remerciements vont aux doctorants du BETA avec qui j'ai eu l'occasion d'avoir de véritables discussions, bien au-delà de la thèse : Tayeb, Anta, Jean-Phillipe, Charlotte, Audrey-Rose, Mamoudou, Lionel, Frédéric,

Walliya, Adeline, Xi, Qiao, Tim, Christian, Faustine, Naouel.

Merci à Danielle Génevé, à Géraldine Manderscheidt-Del Fabbro et Claudine Berst pour leur aide dans mes nombreuses démarches administratives. Merci au BETA et à l'École Doctorale Augustin Cournot de m'avoir fourni les appuis logistique et financier nécessaires à la réalisation de cette thèse.

Je remercie mon église EBS, l'association Hibiscus, Pierre-Wildor, Yannick et Nicolas, qui m'ont aidé à comprendre que je devais avoir une vie à côté de la thèse. Cela m'a permis de faire la part des choses à chaque fois.

Merci au proviseur du Centre Alcibiade Pommayrac, Gérard Borne, de m'avoir accordé une bourse pour poursuivre mes études en France. J'espère que cette thèse sera en quelque sorte la rentabilité de son investissement.

Merci à Rose-Irma pour tout. Merci à tous mes amis, trop nombreux pour figurer ici.

Mes profonds remerciements vont à mes parents Prosper et Jeannette Jean-Baptiste, à ma soeur Ennette Jean-Baptiste et à mon frère Ricardo Jean-Baptiste, pour leur soutien, leur encouragement et leur confiance sans faille.

Et enfin, merci à ma chère Farah pour son soutien, sa patience, sa disponibilité et sa compréhension, alors que plusieurs de mes nuits étaient consacrées à l'économétrie des modèles DSGE.

Frédo Jean-Baptiste, Strasbourg, Novembre 2011

À T-Rica

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Contents

1.1. Introduction

In recent years, the academic research on monetary business cycles has been quickly expanding. One reason for this expansion is that monetary Dynamic Stochastic General Equilibrium (DSGE) models have made decisive progress in accounting for observed business cycle, and in explaining the real effects of monetary shocks. Today, many central banks and policy-making institutions give more and more credence to the predictions of monetary DSGE models, as these models are currently used for policy analysis. Thus, *data confrontation* has always been a first-order issue in the success of these models. Data confrontation is the main focus of this thesis.

This dissertation consists of two parts, each part being subdivided into two chapters. The first part is related to the traditional "frequentist" economet-

ric approach. It presents two essays that investigate the empirical properties of the New Keynesian Phillips curve (chapters 2 and 3). The second part is related to the "Bayesian" econometric approach that has recently become popular in the macroeconomic literature. Chapter 4 surveys Bayesian inference tools used to estimate and evaluate dynamic stochastic general equilibrium models in the presence of regime switches in parameters or in shock variances. Chapter 5 applies these tools to a *state-of-the-art* medium scale DSGE model with regime switches, and uses the estimated model to shed new light on the issue of identifying the sources of the "Great Moderation". This expression refers to the decline in macroeconomic volatility observed in most industrialized economies during the 80's until the recent financial crisis.

A common denominator of the four chapters is the systematic use of empirical methods that take monetary business cycle models to the data. Accordingly, this introduction provides a short historical background where we contrast real versus monetary models of the business cycle. Then, we briefly present the econometric tools that have been developed over the years to take these models to the data. Finally, we outline the content of the thesis by summarizing our main findings.

1.1.1. From Real Business Cycle to Monetary Business Cycle theories

A central issue of modern macroeconomics has been to identify the sources of the volatility in macroeconomic time series. Over the years, macroeconomic theory has been divided between two conflicting views on this issue: the first view attributes to monetary factors a dominant role, and emphasizes the importance of monetary policy as a powerful stabilization tool. The second view attributes to *real* factors the leading cause of business cycles, and underlines the potentially destabilizing role of active monetary policies through their

effects on long-run inflation.

The publication in 1971 of Friedman and Schwartz's essay, *A Monetary history of the United States*, constitutes a landmark day in our understanding of business cycles. By observing that periods of deep recessions in the US economy usually coincide with periods where the stock of money declines, the authors provide narrative evidence that money is an important factor of observed fluctuations. This narrative evidence oriented the macroeconomic research toward models where money plays an important role. Accordingly, during the 70's and the early 80's, two main theories of money non-neutrality were developed. In the first one, formulated in two contributions by Lucas (1972) and Lucas (1975), money is important because monetary shocks are treated "as a source of confusion that makes it difficult for agents to separate relative price changes from aggregate price changes" (Cooley and Hansen, 1995). In the second strand, money matters because nominal prices and wages are sticky in the short-run (Fischer, 1977; Taylor, 1979).

Shortly after this, the publication in 1982 of Kydland and Prescott's seminal article, "*Time to build and aggregate fluctuations*", came as another landmark day for the development of modern macroeconomic theory. This paper, which led to the emergence of the Real Business Cycle (RBC) theory, strongly revived the proposition that economic fluctuations are mostly the result of *non-monetary* forces.[1] In the canonical RBC model (see King et al. (1988), for a presentation), business cycles are viewed as the outcome of an economy submitted to exogenous technological shocks. In this economy, rational agents continuously adjust their decisions to exogenous variations in the environment. Thus, the theory conveys two new important messages: (i) business cycles do not necessarily reflect inefficiencies in the allocation of resources ; (ii) monetary factors such as, for example, variations in the stock of money,

[1] The term 'Real Business Cycle' was first used by Long and Plosser (1983)

do not necessarily play a significant role in observed business cycles, since a model submitted to technological shocks alone can account for a significant fraction of the volatility of output and of other macroeconomic variables.

As emphasized by Galí (2008), a second reason why the RBC theory was considered a "revolution" for macroeconomic analysis is in its methodological contribution. The RBC theory is the first approach to combine, within the same model, three decisive ingredients: First, the use of dynamic stochastic general equilibrium theory, which gives strong microeconomic foundations to the behavioral equations used to describe the aggregate variables in the large-scale macroeconometric models of the 70's. Second, agents in these models form *rational expectations*, i.e. expectations consistent with the information they have. Third, the methodology aims to provide *quantitative* (as opposed to qualitative) predictions, by simulating the model and generating time series for the main variables than can be compared to their empirical counterpart. Based on this comparison, models may be validated or rejected.

While decisive in terms of methodology, the Real Business Cycle theory (at least in its initial form) did not really survive the data confrontation step it advocated. Indeed, the predictions of the canonical RBC model on the role of money and on the covariations between real and monetary variables were quickly shown to be at odds with the large body of *empirical* evidence analyzing this question. For instance, Cooley and Hansen (1995) show that in the data, there is a positive correlation between monetary aggregates and output, while the canonical RBC model augmented with a cash-in-advance constraint predicts a negative correlation. More recently, Christiano et al. (1999, 2005) provide empirical evidence that an expansionary monetary shock increases output in the short run, with a peak occurring after a few periods, while inflation adjusts very little at the time of the shock. Such evidence is clearly inconsistent with the predictions of the augmented RBC model.

Facing these difficulties, researchers have gradually started to include some "Keynesian" features into the model, leading to the progressive development of what is known today as the *New Keynesian model* of the business cycle. There are three main ingredients at the heart of the New Keynesian model: The first one is the introduction of monopolistically competitive firms, who endogenously set their price in order to maximize profits. The second one is the assumption of *nominal rigidities*. This means that, for some reason, firms cannot reset their price optimally in any period. For example, in the traditional Calvo (1983) price-setting framework, firms are only given (in each period) a constant probability of resetting their price optimally. The third ingredient is the specification of a *monetary policy rule* such as, for example, a *Taylor rule*, which describes the reaction of the monetary authority to changes in the economic environment.

In addition to these ingredients, the New Keynesian model also features an important concept which will be our main research interest in the first two chapters of this thesis. This concept is the so-called *New Keynesian Phillips curve*. The New Keynesian Phillips curve (NKPC) is derived from the log-linearization of the optimal pricing decision of firms in the Calvo framework. In its simplest form, the curve relates current and expected inflation to a measure of real activity (average marginal cost). Yet, as we will see, there also exist *hybrid* versions of the NKPC, which usually include past inflation in the equation. A significant part of our research program has been to develop or to test hybrid versions of the NKPC.

The New Keynesian model is also important because it gives a key role to monetary factors in the business cycle. Recently, Christiano et al. (2005) and Smets and Wouters (2007) have illustrated the ability of medium scale New-keynesian DSGE models to account for the real effects of monetary shocks. For this reason, the New Keynesian model of Christiano et al. (2005) and Smets

and Wouters (2007) has become the reference model used to investigate the sources of empirical business cycles. The Smets and Wouters (2007) model is the benchmark model we use in the second part of the thesis, when estimating regime-switching medium scale DSGE models.

1.1.2. Taking the models to the data

Simultaneously with the development of DSGE models, the empirical methods used to assess the fit of these models have undergone a rapid evolution. The first empirical method, advocated by Kydland and Prescott (1982), is termed "calibration". DeJong and Dave (2007) define the calibration step as "an exercise under which a set of empirical targets is used to pin down the parameters of the model under investigation, and a second set of targets is used to judge the model's empirical performance". Thus, according to this definition, the DSGE model is not considered as a data generating process. For proponents of the calibration methodology, the main reason is that any DSGE model, being highly stylized, is *de facto* false. Thus, any formal statistical test should reject it. Hence, it is preferable to judge the empirical performance of this model relatively to the set of quantitative facts it is supposed to explain.

Despite its numerous advantages, the main shortcoming of the calibration approach is that it does not attach any measure of uncertainty to the predictions of the model. To address this problem, the next empirical method that has been considered in the literature is the Generalized methods of moments (GMM), due to Hansen (1982). Like calibration, the GMM methodology focuses on matching only a limited set of empirical targets, called *moments.* But unlike calibration, GMM takes uncertainty seriously since it implies that the model could be interpreted as a data generating process from which the moments were obtained.[2]

[2]For an early example of application of GMM to macroeconomic time series, see **?**. For

Other empirical methods in the same family of GMM has also been used to estimate DSGE models. Among them, we can mention the Simulated Methods of Moments (SMM), the method of Indirect Inference (II), and the method of Minimum Distance Estimation (MDE). In the late 80's and during the 90's, Maximum Likelihood has been at the centre of the estimation procedure. An early reference is Altug (1989). In contrast to the calibration and moment matching procedures, Maximum Likelihood is a full-information method under which the DSGE model is assumed to provide a complete characterization of the data. Hence, in theory, ML should deliver more reliable estimates. However, misspecification remains a concern since maximum likelihood estimation requires some assumption about the distribution of the stochastic components of the model.

Over recent years, several researchers have preferred to favor a Bayesian approach to estimating DSGE models (an early example is Schorfheide (2000)). This choice partly reflects the willingness to avoid some traditional difficulties encountered with the frequentist approach. But the main advantage of Bayesian inference is that it allows researchers to incorporate prior information into the model. This is important because forming prior opinions is a natural device among economists. Because of its increasing importance in the macroeconomic literature, the Bayesian approach is the central theme of the second part of our thesis.

1.1.3. Outline of the thesis

This thesis follows closely the chronology of empirical methods that have been used in recent decades to fit DSGE models to the data. Chapter 2, *Forecasting with the New Keynesian Phillips curve: Evidence from survey data* estimates

an application of the GMM approach to a DSGE model, see e.g. Christiano and Eichenbaum (1992).

the New Keynesian Phillips Curve developed in Galí and Gertler (1999) and assesses the forecasting performance of that curve. Despite the success of the New Keynesian Phillips Curve in explaining the dynamics of inflation, many empirical studies document its weakness by showing that purely statistical models, like ARIMA models, do a better job in forecasting inflation. This chapter tries to revisit this empirical finding. Specifically, I find that a plausible explanation for the poor forecasting performance of the NKPC is due to the way inflation expectations are measured. Following Galí and Gertler (1999), a large body of empirical studies estimate the NKPC, assuming rational expectations. Under the rational expectations assumption, the error in the forecast of expected inflation is uncorrelated with information in the current and past periods. Hence, provided the existence of a vector of variables (called *instruments*) in the current or earlier periods, the NKPC can be estimated via GMM. However, if agents do not make (fully) rational forecasts, estimates could be seriously biased. Thus, I consider an alternative methodology, which consists in constructing an expected inflation series using qualitative survey data. The survey data collect qualitative answers of consumers on their expectation about the evolution of prices for the coming year in Great Britain. Two important results are in order. First, the estimates obtained with the alternative measure of inflation expectations are better than those obtained with the traditional rational expectations assumptions. Second, survey forecasts on inflation expectations greatly improve the forecasting performance of the NKPC.

Chapter 3, *Time-varying inflation target and the New Keynesian Phillips Curve* also focuses on the NKPC but with a different perspective. Our starting point is the two empirical findings documented in Cogley and Sbordone (2008). The authors first argue that once trend inflation is taken into account, the New Keynesian Phillips Curve can replicate the amount of inflation per-

sistence found in the data without requiring the inclusion of *ad hoc* backward-looking terms. Second, they show that the resulting reduced-form NKPC has time-varying coefficients.

We consider the consequences of introducing a time-varying inflation target in the specification of the New Keynesian Phillips Curve. Specifically, we assume that when firms cannot adjust their price, they follow an indexation rule which consists in indexing their price on an inflation rate which is different from the long-run inflation rate. Our main idea is that, in order to limit relative price distortions, firms should index their price on a target which is close to the expected inflation rate prevailing during the average duration of the price contract. In the presence of trend inflation, this target is likely to be (i) significantly different from the sample mean of the inflation rate and (ii) time-varying. We derive a new specification for the NKPC that follows from this assumption, and we take it to the data. Compared to Cogley and Sbordone (2008), the chapter provides three main conclusions. First, in contrast with their paper, our specification of the NKPC features constant coefficients. This enables us to relate more easily the reduced form NKPC to the deep (structural) parameters. Second, as in Cogley and Sbordone (2008), our specification of the NKPC leads to non-significant backward looking coefficients. Third, using identification-robust methods, our alternative NKPC slightly improves the estimated value of the degree of price rigidity in the Calvo price setting mechanism.

While our research in the first two chapters was using frequentist estimation methods, the next two chapters deal with the Bayesian approach. Chapter 4, *A review of Bayesian analysis of DSGE models*, surveys Bayesian econometric methods that have recently been used to estimate DSGE models. We show how such methods can be modified to account for the presence of regime switches in DSGE models. Farmer et al. (2009b) offer an excellent treatment

of forward-looking Markov-Switching DSGE models.

Chapter 5, *Great Moderation and endogenous monetary policy switches* applies these tools to address an important macroeconomic question: what are the sources of the Great moderation that the US and the Euro economies have experienced in the period spanning the mid 80's until the recent financial crisis? The literature suggests two main explanations that have not reached a consensus. For some economists, e.g. (Stock and Watson, 2003b; **?**), the Great Moderation is mainly the outcome of *good luck*. By luck, we mean smaller shocks faced by these economies during this period. For others, the Great Moderation is due to the virtues of monetary authorities in their conduct of monetary policy. Economists have observed that the reduction in output volatility was accompanied by a similar reduction in the volatility of inflation, as documented in Blanchard and Simon (2001). Given the broad consensus that monetary policy is a crucial determinant of inflation, a reduction in the volatility of output may have been the result of better monetary policy. Proponents of this viewed are, inter alia, Bernanke (2004), Lubik and Schorfheide (2004) and Clarida et al. (2000).

To take account of these explanations, we modify the *state-of-the-art* Smets and Wouters (2007) medium-scale DSGE model to incorporate the possibility of regime switches in the variance of shocks and in the coefficients of the monetary policy rule. We use this model to consider three alternative specifications: The first one only introduces changes in the variance of shocks. The second one allows for regime switches in both the variance of shocks and the policy rule coefficients, but assumes that these changes are independent (i.e., changes in the monetary policy regime are independent of the current state of the economy). Finally, the third specification introduces synchronized changes in shocks variance and in monetary policy.

We estimate the three versions of the model with Bayesian methods, and use

the estimation results to shed new lights on the following questions: (i) What are the sources of the Great Moderation ; (ii) Are regime switches in monetary policy exogenous, or does the conduct of monetary policy change according to the economic situation ? The possibility of endogenous monetary policy regime changes has been recently emphasized by Davig and Leeper (2008).

We perform this exercise for the US and the Euro Area economies and obtain the following findings: first, we find strong evidence in favor of regime switches in both policy parameters and shock variances, whether these switches are assumed to be synchronized or independent. This finding holds true for both the US and the Euro Area. Second, for both economies, the specification with synchronized regime shifts fits the data equally well as the specification with independent changes in regime. Third, our findings do not support the view that the US monetary policy has been endogenous. According to our results, the conduct of monetary policy in the US was more strongly determined by the chairmen in office than by the ongoing economic situation. However, this result does not hold true for the Euro Area, and we do find strong evidence of endogeneity of monetary policy in Europe.

Summarizing, the contents of thesis are organized as follows: Chapter 2 estimates the New Keynesian Phillips Curve using survey data and compares its forecasting performance to an AR(1) model. Chapter 3 considers the consequences of introducing a time-varying inflation target in the New Keynesian Phillips Curve. Chapter 4 provides a survey of Bayesian analysis of DSGE models, while chapter 5 estimates a Markov-switching DSGE models and uses the estimated model to shed new lights on the sources of the Great moderation and on the endogeneity of monetary policy in the US and the Euro Area. Finally, a technical appendix for chapters 3 and 5 is provided.

Chapter 2

Forecasting with the New Keynesian Phillips curve: Evidence from survey data

Contents

2.1. Introduction

This chapter is based on my forthcoming paper, Jean-Baptiste (2011). It estimates the New Keynesian Phillips Curve version derived in Galí and Gertler (1999) and assesses the forecasting performance of that curve.

Our motivation is the following. Empirical studies often find that the hybrid New Keynesian Phillips curve developed by Galí and Gertler (1999), while

theoretically justified, has a poor predictive content of inflation, compared to the variety of ARIMA models.[1] In this chapter, I estimate the hybrid NKPC and use the estimation results to forecast inflation. I follow two estimation methods. First, I impose the rational expectations assumption of agents as Galí and Gertler (1999) and estimate the Phillips curve by GMM. Second, I use inflation forecasts obtained from survey data and estimate the Phillips curve by OLS.

The results are the following. First, estimation with survey data performs well in quantifying the backward and forward coefficients of the hybrid NKPC. Second, output gap is not significant and enters the hybrid NKPC with a negative sign On the contrary, output gap is found to be significant and enters the NKPC positively when survey-based inflation forecasts are used. Third, the forecasting performance of the hybrid NKPC is superior to the benchmark AR(1) model when output gap and survey-based inflation forecasts are used.

The rest of the paper is organized as follow. The second section briefly presents the hybrid NKPC. The third section describes the data and presents the estimation results. The fourth section presents the results of the forecasting experiments and the last section concludes.

2.2. The hybrid New-Keynesian Phillips curve

To derive the hybrid NKPC, Galí and Gertler (1999) use a staggered price-setting scheme *à la* Calvo (1983), where a fraction of firms, $1 - \theta$, change prices in a given period. In contrast to the original Calvo (1983) model, Galí and Gertler (1999) assume that among firms being able to change prices in a given period, only a fraction $1 - \omega$ sets price optimally in a forward-looking

[1] The literature uses the term *hybrid* in opposition to the pure forward-looking New Keynesian Phillips curve.

manner. The remaining part sets prices by simply augmenting last period's average price by the inflation rate of that period. This assumption leads to the following form of the hybrid NKPC:

$$\pi_t = \lambda x_t + \gamma_b \pi_{t-1} + \gamma_f E_t \pi_{t+1} + u_t, \tag{2.2.1}$$

where π_t is the inflation rate, $E_t\pi_{t+1}$ the expected inflation rate of the next period, x_t the real marginal cost and u_t a cost-push shock. The reduced form in (2.2.1) is related to the structural form of the NKPC by the following combination of parameters:

$$\gamma_f = \frac{\beta\theta}{\phi},$$
$$\gamma_b = \frac{\omega}{\phi},$$
$$\lambda = \frac{(1-\omega)(1-\theta)(1-\beta\theta)}{\phi}$$

where β is a discount rate and $\phi = \theta + \omega[1-\theta(1-\beta)]$. Since the reduced form of the hybrid NKPC in (2.2.1) is consistent with various pricing schemes, including the Calvo (1983) scheme, estimation and forecasting results are reported only for the reduced form.

2.3. Estimation

In (2.2.1), the term $E_t\pi_{t+1}$ is not directly observed. This is a fundamental challenge with estimating the hybrid NKPC parameters. There are two solutions to circumvent this challenge. The first one, which is the most standard, is to use the law of iterated expectations to obtain a forecast of $E_t\pi_{t+1}$[2]. Replacing $E_t\pi_{t+1}$ by $\pi_{t+1} - \eta_{t+1}$, with η_{t+1} being the one-step-ahead forecast error in π_{t+1}, I obtain the new equation $\pi_t = \lambda x_t + \gamma_b \pi_{t-1} + \gamma_f \pi_{t+1} + e_t$, where $e_t = u_t - \gamma_f \eta_{t+1}$. From $E_t(u_t) = 0$, it follows that $E_t(e_t) = 0$ so that the

[2]The law of iterated expectations is a principle of rational expectations.

equation can be estimated by GMM. The second solution is to use inflation forecasts from survey data, obtained by asking economic agents at one period what are their expectations of price for the next period. Since these expectations do not necessarily match inflation expectations, I assume that they are given by $E_t\pi_{t+1} = \pi^s_{t+1} + \eta^s_t$ where π^s_{t+1} is the inflation forecast provided by the survey and η^s_t is an error term uncorrelated with π^s_{t+1}. Finally, I estimate by OLS the resulting equation $\pi_t = \lambda x_t + \gamma_b \pi_{t-1} + \gamma_f \pi^s_{t+1} + \epsilon_t$ where $\epsilon_t = \gamma_f \eta^s_t$.

2.3.1. The data

Survey data come from the European Commission website from 1987:1 to 2007:12. The European Commission conducts a monthly survey about the general economic situation, unemployment and price expectations for the European economy, the Euro area and the single European countries. For the purpose of the paper, I focus on price expectations. The survey participants are asked to give qualitative expectations about the evolution of prices in the next year. As a consequence, qualitative expectations are quantified with the Carlson and Parkin (1975) approach, also called the "probability approach", and the results are converted to quarterly frequency, in order to fit a quarterly hybrid NKPC.[3] To save space, I refer the reader to Carlson and Parkin (1975) and Nielsen (2003) for a detailed explanation of that approach. Data on other variables are provided by the OECD Economic Outlook database. Actual inflation is measured as 100 times log difference of quarterly consumer price index (CPI), from 1987:1 to 2007:4. I consider two proxies for real marginal

[3]The survey asks the following question to the participants:"By comparison with the past 12 months, how do you expect consumer prices will develop in the next 12 months? They will: 1) increase more rapidly 2) increase at the same rate 3) increase at a slower rate 4) stay about the same 5) fall 6) do not know."

Table 2.1: OLS estimation of the hybrid NKPC

Parameters	Unrestricted		Restricted	
	Gap	ULC	Gap	ULC
γ_b	0.236	0.120	–	–
	(0.001)	(0.061)		
γ_f	0.726	0.876	0.779	0.880
	(0.000)	(0.000)	(0.000)	(0.000)
λ	0.131	–	0.096	–
	(0.008)		(0.031)	
λ'	–	-0.022	–	-0.024
		(0.378)		(0.318)

Notes: p-values are given in brackets.

cost: CPI-deflated unit labour cost of the total economy and the output gap.[4] I report empirical results for the United Kingdom. Figure (2.1) plots the actual inflation versus the mean expected inflation and the detrended real unit labour cost versus the output gap.

2.3.2. Estimation results

While the paper focuses on estimates using survey data, I report, for comparison purpose, estimates based on the traditional iterated expectations estimation for comparison.

Table (2.1) reports OLS estimates of the hybrid NKPC. The unrestricted coefficients γ_f and γ_b are positive and significant. Thus, both forward and backward looking components are important in the dynamics of inflation. Irrespective of the proxy used for real marginal cost, the degree of forwardness is

[4] Galí and Gertler (1999) compute output gap as the difference of real gross domestic product and its linear trend. In this paper, I use directly the output gap published by the OECD. The OECD output gap is measured as the percentage difference between GDP (constant prices) and potential GDP.

Table 2.2: GMM estimation of the hybrid NKPC

Parameters	Unrestricted	
	Gap	ULC
γ_b	0.275	0.391
	(0.069)	(0.055)
γ_f	0.735	0.616
	(0.006)	(0.011)
λ	-0.028	–
	(0.325)	
λ'	–	0.0003
		(0.0001)
Hansen's J test	1.36	1.17
	(0.56)	(0.51)

Notes: Instruments include three lags of inflation and two lags of output gap and real unit labour cost. p-values are given in brackets.

more important than the degree of backwardness. This finding is in line with Galí et al. (2005) who find that the forward-looking component of inflation is very important, using IV estimation. The estimates indicate that real unit labour costs are not significant for the dynamics of inflation, at least in the sample considered here. Furthermore, since the slope λ is negative, real unit labour costs enters the hybrid NKPC with the wrong sign, which is possibly a result from an errors-in-variable problem associated with the expected inflation.[5] The Output gap is significant and enters the hybrid NKPC with the correct sign. Thus, the estimates indicate that the output gap is a good proxy for marginal cost. Under the theoretical restriction $\beta \approx 0.99$, which implies

[5]Nason and Smith (2008) using survey data obtained from the Survey of Professional Forecasters (SPF), find a similar result for the US economy and Henzel and Wollmershaeuser (2008) find a similar result for Italy, using survey data from the CESifo World Economic Survey.

$\gamma_f + \gamma_b = 1$, Table (2.1) shows that more weight is given to the forward looking component. The slope remains positive and significant when the output gap is used as proxy, negative and insignificant when real unit labor costs are used. These findings contradict the widespread view that a cost-based formulation of inflation is better than output gap-based formulation of inflation (Galí and Gertler, 1999).

Table (2.2) reports estimates based on the continuous updating GMM estimator (CUE-GMM) of Hansen et al. (1996) where the covariance matrix is corrected with a bandwidth of 12 lags. p-values for the Hansen test provide no evidence against the validity of the instruments. The forward looking component continue to play the predominant role. However, there are some differences in magnitude with estimates based on survey forecasts. In particular, more weight is attached to the backward component. The output gap coefficient, while not significant, is negative. When using survey forecasts, I obtain the correct sign and the coefficient is significant.[6] The results suggest that survey-based estimates perform better than estimation methods based on rational expectations of agents.

2.4. Forecasting experiments

This section is motivated by empirical evidence reported by Ang et al. (2007). Using four forecasting models based on macro, asset markets variables and inflation surveys data, Ang et al. (2007) find that in terms of inflation forecasts, the forecasting model with survey data outperforms the other models for the US economy. I use the hybrid NKPC to forecast annual inflation and compare the forecasting results with those of a benchmark autoregressive model of order one, AR(1). The specification of the AR(1) is standard and is given

[6] See Nunes (2010) for similar findings.

by the following equation:

$$\pi_t = \mu + \alpha\pi_{t-1} + \nu_t$$

where ν_t is an error term and μ a drift.

Table 2.3: Out-of-sample percent rmses : AR(1) vs NKPC

		forecast horizon		
		2	4	8
United Kingdom	AR	0.610	0.706	0.896
	RULC	0.727	0.806	0.995
	GAP	0.564	0.636	0.821

Note: Forecast of annual inflation, out-of-sample from 2000:Q1 to 2007:Q4. RULC and GAP refer to the hybrid NKPC estimated with real unit labor cost and output gap. The root mean squared error forecasts are reported in percentage terms.

I consider pseudo out-of-sample forecast of inflation, from 2000:Q1 onwards. Table (2.3) reports the root mean squared error (RMSE) statistics in percentage terms for the AR(1) model and the hybrid NKPC model. Forecasting performance of the hybrid NKPC depends on the proxy used to measure real marginal cost. At all forecast horizons, the autoregressive model beats the hybrid NKPC estimated with real unit labour cost. Using the output gap as proxy for real marginal cost considerably improves the forecasting performance of the hybrid NKPC. Compared to the AR(1) model, all the RMSE are lower at all forecast horizons. This result is encouraging since empirical studies (see for instance Stock and Watson (2003a)) have found that the NKPC, while theoretically justified, has a poor predictive content, compared to the variety of ARIMA models.

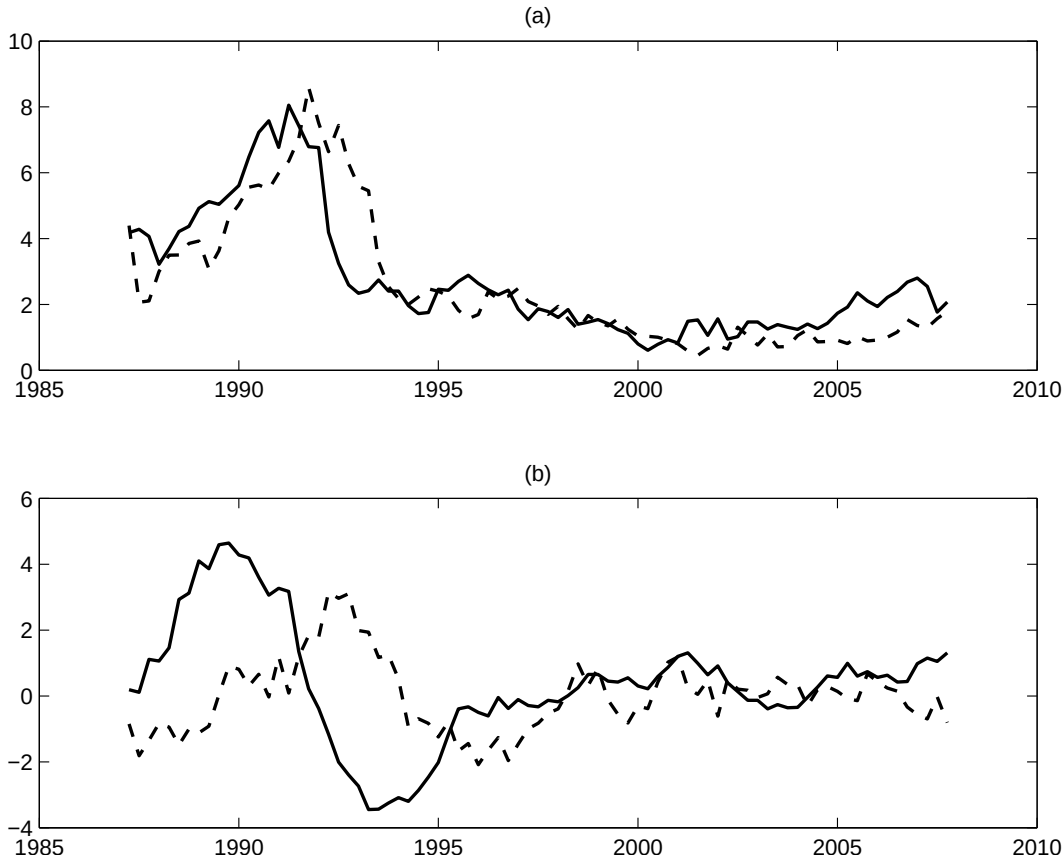

Figure 2.1: (a): Actual inflation (solid) and expected inflation (dashed). (b): Output gap (solid) and real unit labour cost (dashed).

2.5. Conclusions

In this paper, I have found that survey-based inflation forecasts make the Phillips curve predominantly forward looking. The output gap enters positively and significantly, while methods based on traditional rational expectations deliver a negative and insignificant role to the output gap. Furthermore, the root mean squared errors of the Phillips curve are inferior to those of an AR(1) model when inflation forecasts and output gap are used.

Chapter 3

Time-varying inflation target and the New Keynesian Phillips curve

Contents

3.1. Introduction

The aim of this chapter is to propose and to estimate an alternative specification for the New Keynesian Phillips Curve.

The main idea underlying our approach is that, in the traditional specification of the hybrid NKPC (with partial indexation and positive steady-state inflation), the indexing rule of firms that are not allowed to reset their price does not fit well with the observation that inflation in the short run may substantially differ from its long-run (average) value. This is notably the case, for example, if the inflation target of the central bank occasionally switches between a high and a low value (as documented in Schorfheide (2005)), or if there is a trend in the evolution of inflation. Indeed, under such circumstances, using the long-run inflation level to index non-resetted prices may generate substantial price distortions that are ultimately costly to firms.

In the presence of varying trend inflation, a better indexing rule would index prices on an inflation target that does not differ too much from the implicit inflation rate applied by reoptimizing firms to their former price. In the presence of trend inflation, this target is likely to be (i) significantly different from the sample mean of the inflation rate and (ii) time-varying. We derive a new specification for the NKPC that follows from this assumption, and we estimate it with the minimum distance estimation documented in **?**. Following a recent literature on the issue that NKPC might be poorly identified (Ma, 2002; Dufour et al., 2010; Kleibergen and Mavroeidis, 2009), we estimate our NKPC with methods that are robust to identification problem with the approach documented in Magnusson and Mavroeidis (2010). To check the robustness of our conclusions, we also estimate the NKPC with the Generalized method of moments (GMM).

Our work also relates to recent empirical results by Cogley and Sbordone (2008), who emphasized the importance of taking trend inflation into account for the empirical success of the NKPC. The authors criticize the common practice of introducing backward-looking terms in the NKPC to improve its empirical performance, since in its initial formulation the curve is purely

forward-looking. Cogley and Sbordone (2008) find that once drifting trend inflation is taken into account in the Calvo price setting mechanism, backward-looking terms are no longer necessary to account for inflation persistence. They also show that the introduction of drifting trend inflation results in a New Keynesian Phillips Curve which has the characteristic that its reduced form coefficients are time-varying.

Compared to Cogley and Sbordone (2008), our approach leads to three main conclusions. First, in contrast with Cogley and Sbordone (2008), the reduced form coefficients of our modified NKPC are constant. This enables us to relate more easily the reduced form NKPC to the deep (structural) parameters. Second, similarly to Cogley and Sbordone (2008), estimation of our modified NKPC leads to non-significant backward looking coefficients. This conclusion is confirmed, using both non-robust and robust methods to identification issues. Third, time-varying inflation target is a key variable for the evolution of inflation.

The remaining of the chapter is organized as follow. Section (3.2) derives the NKPC, section (3.3) presents the econometric methodology, while section (3.4) presents the results. The last section concludes the chapter.

3.2. The generalized NKPC

The NKPC is derived using the Calvo (1983) pricing mechanism according to which, in each and every period, a firm faces a constant probability $1 - \xi_p$ to reset its price optimally. In the traditional specification of the hybrid NKPC with partial indexation and positive long run inflation, a firm who cannot reset its price optimally is assumed to apply the following indexation rule (see e.g. Smets and Wouters (2007)):

$$p_t^i = (\pi_{t-1})^{\gamma_p} (\overline{\pi})^{1-\gamma_p} p_{t-1}^i$$

where $\overline{\pi}$ is the long-run inflation factor. As mentioned above, if the inflation target of the central bank is changing through time, or if there is a trend in the evolution of inflation, the current inflation level may be significantly different from its sample-mean level $\overline{\pi}$. In this case, the indexing rule is likely to generate substantial price distortions between firms who reset their price optimally and firms who do not.

To avoid this criticism, we consider the following alternative rule:

$$p_t^i = (\pi_{t-1})^{\gamma_p} (\pi_t^*)^{1-\gamma_p} p_{t-1}^i$$

where π_t^* is the implicit measure of trend inflation used by indexing firms. Clearly, if $\pi_t^* = \overline{\pi}$, we recover the traditional specification. But alternative measures are likely to generate less price distortions. For example, π_t^* could be the expected inflation rate of "naive" forecasters at date t (as considered in chapter 2), or be any statistical measure obtained from the data. Of course, in general, π_t^* is likely to be affected by the current economic situation. Thus, generally speaking, π_t^* should be considered as time-varying.

Let $X_{t,k}^p$ be an indexation factor, defined by

$$X_{t,k}^p = \left\{ \begin{array}{c} 1 \text{ for } i = 0 \\ \prod_{l=1}^{k} \left((\pi_{t+l-1})^{\gamma_p} \left(\pi_{t+l}^*\right)^{1-\gamma_p} \right) \text{ for } i = 1, ..., \infty \end{array} \right\}$$

This indexation rule implies that

$$p_{t+k}^i = X_{t,k}^p \widetilde{p_t^i}$$

where $\widetilde{p_t^i}$ is the initial price.

In this context, the optimal price $\widetilde{p_t^i}$ chosen by an optimizing firm in t is the solution of the following program:

$$\max_{\widetilde{p}_t^i} \quad \underset{t}{E} \sum_{k=0}^{\infty} \left(\beta\xi_p\right)^k \frac{\lambda_{t+k}}{\lambda_t} \left(\frac{\widetilde{p}_t^i X_{t,k}^p}{p_{t+k}} - mc_{t+k} \right) y_{t+k}^i$$

subject to

$$y_{t+k}^i = \left(\frac{\widetilde{p}_t^i X_{t,k}^p}{p_{t+k}} \right)^{-\theta} y_{t+k},$$

where $E_t \beta^k \lambda_{t+k}/\lambda_t$ is a stochastic discount factor, with β the subjective discount factor, λ_t the marginal utility of consumption and mc_t the firm's marginal cost.

The first-order condition associated with the above program leads to the following optimal pricing rule:

$$\frac{\widetilde{p}_t^i}{p_t} = \mu_p \frac{\underset{t}{E} \sum_{k=0}^{\infty} \left(\beta\xi_p\right)^k \lambda_{t+k} mc_{t+k} y_{t+k}^i}{\underset{t}{E} \sum_{k=0}^{\infty} \left(\beta\xi_p\right)^k \frac{p_t}{p_{t+k}} X_{t,k}^p \lambda_{t+k} y_{t+k}^i}, \tag{3.2.1}$$

where $\mu_p = \theta/(\theta-1)$ is the steady-state markup, with $\theta > 1$ the price-elasticity of sectoral demand.

The aggregate price index is given by

$$p_t = \left(\int_0^1 \left(p_t^i\right)^{1-\theta} \right)^{\frac{1}{1-\theta}} \tag{3.2.2}$$

Log linearization of (3.2.1) combined with the definition for the aggregate price index (3.2.2) leads to the following New Keynesian Phillips Curve:

$$\begin{aligned} \widehat{\pi}_t \quad = \quad & \frac{\left(1-\xi_p\right)\left(1-\beta\xi_p\right)}{\xi_p(1+\beta\gamma_p)} \widehat{mc}_t \\ & + \frac{\gamma_p}{(1+\beta\gamma_p)} \widehat{\pi}_{t-1} + \frac{\beta}{(1+\beta\gamma_p)} E_t \widehat{\pi}_{t+1} \\ & + \frac{\left(1-\gamma_p\right)}{(1+\beta\gamma_p)} \widehat{\pi}_t^* - \frac{\beta\left(1-\gamma_p\right)}{(1+\beta\gamma_p)} E_t \widehat{\pi}_{t+1}^* \end{aligned} \tag{3.2.3}$$

As expected, our generalized NKPC shows that current inflation depends, in addition to the traditional terms, on the current and expected value of the inflation target, which is generally time-varying. Note also that the NKPC results in reduced form coefficients that are time-invariant. This is the main difference with the Cogley and Sbordone (2008)'s NKPC.

For the purpose of the next subsection, we express the NKPC as a reduced form relation given by

$$\pi_t = \gamma_b \pi_{t-1} + \gamma_f E_t \pi_{t+1} + \lambda mc_t + \lambda_2 \pi_t^* + \lambda_3 E_t \pi_{t+1}^*, \tag{3.2.4}$$

where

$$\lambda = \frac{(1-\xi_p)(1-\beta\xi_p)}{\xi_p(1+\beta\gamma_p)},$$
$$\gamma_b = \frac{\gamma_p}{1+\beta\gamma_p},$$
$$\gamma_f = \frac{\beta}{1+\beta\gamma_p},$$
$$\lambda_2 = \frac{1-\gamma_p}{1+\beta\gamma_p},$$
$$\lambda_3 = -\beta\frac{1-\gamma_p}{1+\beta\gamma_p}.$$

3.3. Econometric methodology

This section describes the methodology that we shall use to test the NKPC. Given that the methodology is recent, we find it useful to describe it before estimating the NKPC.

3.3.1. Minimum distance

Magnusson and Mavroeidis (2010) consider the minimum distance estimation of the NKPC originated from Sbordone (2002) and Sbordone (2005). In our

context, the minimum distance estimation of the NKPC tries to minimize the distance between two dynamics of inflation. The first dynamics models inflation, marginal cost (proxied by labor share) and trend inflation as a VAR process of order p, while the second dynamic is the relation provided by the NKPC. In other words, the approach uses information coming from reduced-form models to estimate a structural model of inflation.

The approach is implemented as follows. Let $Z_t \equiv [\pi_t \ \pi_t^* \ mc_t]$ and assume that Z_t is determined by a VAR(p) process:

$$Z_t = \Phi_1 Z_{t-1} + \cdots + \Phi_p Z_{t-p} + \epsilon_{z,t} \tag{3.3.1}$$

where $E(\epsilon_{z,t}) = 0$ and $E(\epsilon_{z,t}\epsilon'_{z,t}) = \Omega$.

It is convenient to rewrite the VAR(p) in its companion form:

$$z_t = \Phi z_{t-1} + Q\epsilon_{z,t} \tag{3.3.2}$$

where $z_t = [z_t \ z_{t-1} \ z_{t-2} \ \cdots \ z_{t-p+1}]'$,

k is the number of variables in the VAR so that Φ is of dimension $kp \times kp$ and it contains k^2p unknown parameters, denoted by φ.

Let e_π, e_{π^*}, e_{mc} be appropriate selection vectors such that $\pi_t = e'_\pi z_t, \pi_t^* = e'_{\pi^*} z_t, mc_t = e'_{mc} z_t$.

Next, we link the reduced-form parameters φ to the structural parameters ϑ of the NKPC. We use the identifying restriction that $E_{t-1}\varepsilon_t^p = 0$. Taking expectations with respect to information in $t-1$ on both sides of equation 3.2.4, it follows that

$$E_{t-1}\pi_t = \gamma_b \pi_{t-1} + \gamma_f E_{t-1}\pi_{t+1} + \lambda_1 E_{t-1} mc_t + \lambda_2 E_{t-1}\pi_t^* + \lambda_3 E_{t-1}\pi_{t+1}^*. \tag{3.3.3}$$

From the VAR representation of the NKPC, we have $E_{t-1}\pi_{t+1} = e'_\pi \Phi(\varphi)^2 z_{t-1}$, $E_{t-1}\pi_{t+1}^* = e'_{\pi^*}\Phi(\varphi)^2 z_{t-1}$, $E_{t-1}\pi_t = e'_\pi \Phi(\varphi) z_{t-1}$, $E_{t-1} mc_t = e'_{mc}\Phi(\varphi) z_{t-1}$ and

$\pi_{t-1} = e'_{\pi} z_{t-1}$. Substituting them in (3.3.3) and transposing the resulting expression leads to a set of restrictions, $g(\varphi, \vartheta)$, where

$$g(\varphi, \vartheta) = \Phi(\varphi)' \left\{ e'_{\pi}[I - \gamma_f \Phi(\varphi)] - \lambda_1 e'_{mc} - e'_{\pi^*}[\lambda_2 I + \lambda_3 \Phi(\varphi)] \right\}' - \gamma_b e_{\pi}. \quad (3.3.4)$$

The estimation strategy proceeds in two steps. First, we estimate the unrestricted VAR to obtain $\widehat{\varphi}$ and an estimate of its variance, $\widehat{V}_{\varphi}$. The second step takes$\widehat{\varphi}$ as given and chooses the value ϑ that makes the empirical value of the function $g(\varphi, \vartheta)$ as close as possible to zero.

3.3.2. Identification robust tests

Let $\widehat{\varphi}$ denote a consistent and asymptotically normal estimator of φ, with asymptotic variance V_{φ}, and let $\widehat{V}_{\varphi}$ be a consistent estimator of V_{φ}. By the Delta method, the asymptotic variance of $g(\widehat{\varphi}, \vartheta)$ is $G_{\varphi}(\varphi, \vartheta)' V_{\varphi} G_{\varphi}(\varphi, \vartheta)$ where $G_{\varphi}(\varphi, \vartheta) \equiv \frac{\partial g(\varphi, \vartheta)}{\partial \varphi'}$. Efficient minimum distance estimation is based on the criterion function

$$Q(\vartheta) = g(\widehat{\varphi}, \vartheta)' \widehat{V}_{gg}(\overline{\vartheta})^{-1} g(\widehat{\varphi}, \vartheta) \quad (3.3.5)$$

where $\widehat{V}_{gg}(\vartheta) = G_{\varphi}(\widehat{\varphi}, \vartheta)' \widehat{V}_{\varphi} G_{\varphi}(\widehat{\varphi}, \vartheta)$ and $\overline{\vartheta}$ being a preliminary estimator of ϑ that could be inefficient. When $\overline{\vartheta} = \vartheta$, the criterion (3.3.5) is the continuous updating criterion provided in Hansen et al. (1996).

Under the assumption that the Jacobian matrix $G_{\varphi}(\varphi, \vartheta)$ has a full rank, $\widehat{\vartheta}$ is asymptotically normal and standard test statistics like the Wald and Lagrange Multiplier (LM) for hypotheses on ϑ are asymptotically chi-squared distributed. Consequently, inferences based on those statistic are reliable. However, when the Jacobian matrix is nearly of reduced or zero rank, i.e the NKPC is weakly identified, inferences based on the Wald or the the LM

statistics are unreliable because these statistics are not asymptotically chi-squared distributed.[1] Thus, it is important to consider test statistics that do not involve asymptotically the Jacobian matrix under the null hypothesis. All the tests are based on the continuous updating estimator (CUE) of the criterion (3.3.5), in which case the weighting matrix is continuously evaluated at the values of the structural parameters, i.e $\overline{\vartheta} = \vartheta$.[2] Inferences obtained from the test are robust to weak identification.

The first test statistic is the minimum distance version of the test proposed Anderson and Rubin (1949):

$$\mathcal{MD}.\mathcal{AR}(\vartheta_0) = Tg(\widehat{\varphi}, \vartheta_0)'\widehat{V}_{gg}(\overline{\vartheta})^{-1}g(\widehat{\varphi}, \vartheta_0), \tag{3.3.6}$$

where T is the sample size, ϑ_0 is the hypothesized value of the parameters. The $\mathcal{MD}.\mathcal{AR}$ is robust to weak identification issue since it does not involve the Jacobian matrix.

This test can be interpreted as a Wald test of the validity of the restrictions in (3.3.4) at ϑ_0.

The second statistic is the minimum distance version of the K statistic proposed by Kleibergen (2005):

$$\begin{aligned}\mathcal{MD}.\mathcal{K}(\vartheta_0) = {} & Tg(\widehat{\varphi}, \vartheta_0)'\widehat{V}_{gg}(\vartheta_0)^{-1}\widehat{D}(\vartheta_0)[\widehat{D}(\vartheta_0)\widehat{V}_{gg}(\vartheta_0)^{-1}\widehat{D}(\vartheta_0)]^{-1} \\ & \times (g(\widehat{\varphi}, \vartheta_0)'\widehat{V}_{gg}(\vartheta_0)^{-1}\widehat{D}(\vartheta_0))'.\end{aligned} \tag{3.3.7}$$

[1] In the framework of GMM estimation, Kleibergen (2005) shows that the asymptotic distribution of the LM statistic is not chi-squared since the average moment vector and the Jacobian estimator are correlated, thus adding nuisance parameters.

[2] Hansen et al. (1996) shows that, in the context of GMM, one advantage of the CU-GMM estimator relative to the two-step estimator is that the former is invariant to how the moment conditions are scaled. More importantly, Monte-Carlo experiments suggest that the CU-GMM estimator outperforms the traditional two-step GMM and the test for identifying restrictions is more reliable in many cases.

Since $\frac{\partial Q}{\partial \vartheta} = 2g(\widehat{\varphi}, \vartheta_0)'\widehat{V}_{gg}(\vartheta_0)^{-1}\widehat{D}(\vartheta_0)$ where $\widehat{D}(\vartheta_0)$ is an estimator of the Jacobian matrix, the minimum distance version of the K statistic is a quadratic form of the derivative of objective function in (3.3.5) with respect to its asymptotic variance $[\widehat{D}(\vartheta_0)\widehat{V}_{gg}(\vartheta_0)^{i-1}\widehat{D}(\vartheta_0)]^{-1}$.[3] It resembles the Lagrange Multiplier (LM) statistic. However, the key difference is that it does not depend on the Jacobian matrix (unlike the LM statistic), but on an estimated value of the Jacobian matrix. Actually, the appendix of Magnusson and Mavroeidis (2010) shows that, asymptotically, $\widehat{D}(\vartheta_0)$ is independent of the vector of restrictions $g(\widehat{\varphi}, \vartheta_0)$. It is this independence that makes the $\mathcal{MD.K}$ statistic robust to identification: conditional on $\widehat{D}(\vartheta_0)$, i.e treating $\widehat{D}(\vartheta_0)$ as a fixed matrix, the statistic $\mathcal{MD.K}$ is asymptotically chi-squared since the independence between D and $g(\widehat{\varphi}, \vartheta_0)$ does not involve additional nuisance parameters.

Details on the derivation of the matrix $\widehat{D}(\vartheta_0)$ can be found in Kleibergen (2005).

The $\mathcal{MD.K}(\vartheta_0)$ tests the null $H : \vartheta = \vartheta_0$, assuming that the identifying restrictions in (3.3.4) hold. Since the continuous updating estimator, which is the basis of all the test statistics, provides values for ϑ where the objective function is minimal, the identifying restrictions are violated around values of ϑ that maximize the objective function. Consequently, the $\mathcal{MDK}$ statistic provides spurious inference around values of ϑ that maximize the objective function. Therefore, Magnusson and Mavroeidis (2010) propose a third statistic that tests the identifying restrictions under the null. It is defined as

$$\mathcal{MD.J}(\vartheta_0) = \mathcal{AR} - \mathcal{MD.K}. \tag{3.3.8}$$

The joint test, i.e the test of the null $H : \vartheta = \vartheta_0$ *and* the validity of the restrictions, denoted $\mathcal{MD.KJ}(\vartheta_0)$, is constructed such that, given a significance level of α, the tested hypothesis is either rejected by an α_1 level $\mathcal{MD.K}$

[3] For a proof, see Kleibergen (2005).

test or by an α_2 level $\mathcal{MD}.\mathcal{J}(\vartheta_0)$ test, where $\alpha_1 + \alpha_2 = \alpha$. As our focus is on H, α_1 must be higher than α_2. Following Magnusson and Mavroeidis (2010), we choose $\alpha_1 = .8\alpha$ and $\alpha_2 = .2\alpha$.

Proposition 1 of Magnusson and Mavroeidis (2010) shows that the three statistics are asymptotically chi-square distributed under fairly general regularity conditions. Identification robust $(1 - \alpha)$ confidence sets are obtained by collecting all values of ϑ that are not rejected by the tests at the α level of significance.

3.4. Results

This section begins with the description of the data used to estimate the Phillips cuve. Then, we present estimates based on both non-robust and robust methods to identification for the reduced form and structural NKPC. Finally, we check the robustness of our results by estimating the NKPC by GMM.[4]

3.4.1. Data

Inflation is measured as the quarter to quarter percent change in the log GDP deflator. We use the labor share of Nonfarm Business sector as proxy for marginal cost. All data are obtained from the Fred Database. We restrict the sample to the period 1984:I-2008:III.

Various alternative measures for π_t^* could be considered. In this paper, we follow a two-step approach, which is in the same spirit of Aruoba and Schorfheide (2011). In the first step, we use a one-sided Hodrick-Prescott filter described for instance in Stock and Watson (1999), to extract the trend component of

[4] We thank Patrick Fève for this suggestion.

inflation with the smoothing parameter fixed to 1600.[5] In the second step, we assume that π_t^* partly results as medium-run inflation expectation and we combine the trend component of inflation with a measure of medium-run inflation expectation. Our measure of medium-run inflation expectation is the one-year ahead inflation expectation provided by the Federal Reserve Bank of Philadelphia. The two series are combined in order to extract common information they contain with respect to π_t^*. Typically, we estimate a state-space model with Bayesian methods and we use the Kalman filter to extract common factor between the two series.

In our state space model, the measurement equations are as follows: $\pi_t^{HP} = \pi_t^* + 0.25\epsilon_{1,t}$ and $\pi_t^{1y} = \pi_t^* + \epsilon_{2,t}$, where π_t^{HP} is the trend component of inflation, π_t^{1y} the (observable) one-year inflation expectation. $\epsilon_{1,t}$ and $\epsilon_{2,t}$ are measurement errors. As in Aruoba and Schorfheide (2011), we fix the standard error of the first measurement error to 0.25 percent to control the weight of the trend component of inflation in the combined series.

The transitions equations write $\pi_{t-1}^* = \rho_\pi \pi_{t-1}^* + \sigma_\pi \epsilon_{\pi,t}$ and $\epsilon_{2,t}, = \rho_2 \epsilon_{2,t-1} + \nu_t$, where $\epsilon_{2,t}$ and ν_t are i.i.d shocks. We assume that π_t^* is a stationary process, i.e $0 < \rho_\pi < 1$. This assumption comes from the fact that the sample considered in our estimation spans 1984:I to 2008:III, a period where key macroeconomic variables of the US economy have been particularly stable.

Once an estimate of ρ_π is available, the expected value of the inflation target can be straightforwardly computed, i.e $E_t \pi_{t+1}^* = \rho_\pi \pi_t^*$. Furthermore, for the purpose of the estimation, we add a cost-push shock, ε_t^p to the NKPC. Thus, we estimate the following specification of the NKPC:

$$\pi_t = \gamma_b \pi_{t-1} + \gamma_f E_t \pi_{t+1} + \lambda mc_t + \lambda_4 \pi_t^* + \varepsilon_t^p \tag{3.4.1}$$

[5] We thank Patrick for having pointed to us that the two-sided HP filter we have used in a previous version of this chapter could deliver spurious results.

Table 3.1: Estimates of the state space model

	Prior			Posterior		
Parameters	Distribution	para(1)	para(2)	Estimates	5%	95%
ρ_π	Beta	0.8	0.025	0.963	0.940	0.987
ρ_2	Beta	0.8	0.025	0.936	0.893	0.983
σ_π	Invgamma	0.1	2	0.191	0.168	0.213
σ_2	Invgamma	0.1	2	0.188	0.166	0.213

Notes: para(1) and para(2) list the means and the standard deviations for Beta distribution; the shape s and and the scale ν parameters for the Inverse Gamma distribution, where $p_{IG}(\sigma|\nu, s) \propto \sigma^{-\nu-1}\exp(-\nu s^2/2\sigma^2)$. Posterior estimates are obtained with the Metropolis algorithm, where a Markov chain of size 100000 has been simulate, with the first 30000 being discarded.

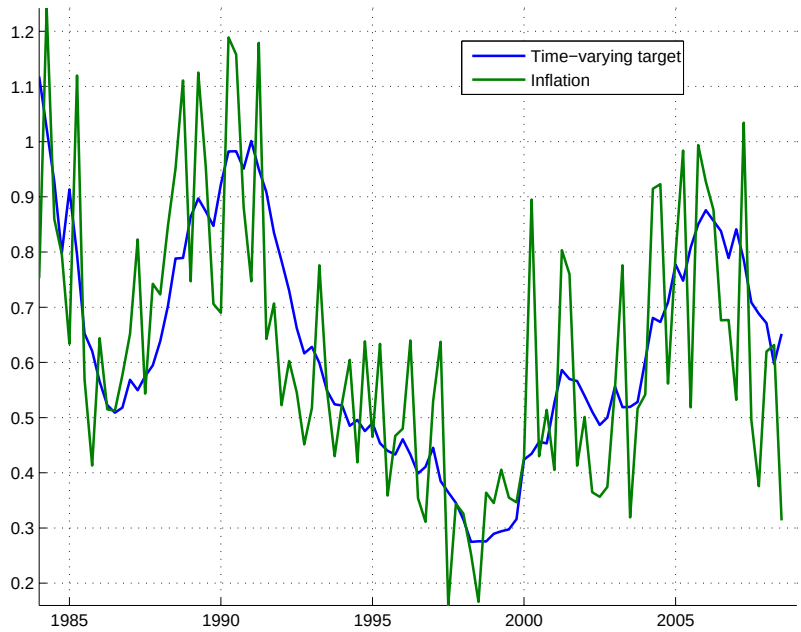

Figure 3.1: Inflation and time-varying inflation target.

where $\lambda_4 = \lambda_2 + \rho_\pi \lambda_3$.

We plot the time-varying inflation target in 3.2. The graph suggests a similar evolution of time-varying inflation target and inflation itself.

Table 3.2: Non-robust minimum distance estimation of the NKPC

	Modified NKPC		Baseline NKPC	
Parameters	Estimates	Std.err	Estimates	Std.err
γ_p	0.032	0.098	0.534	0.116
ξ_p	0.894	0.044	0.876	0.070
γ_b	0.031	0.092	0.349	0.050
γ_f	0.959	0.091	0.648	0.049
λ	0.013	0.011	0.012	0.014
λ_4	0.044	0.009		

Notes: Standard deviations for γ_p and ξ_p are computed by bootstrap. We use the delta method to compute standard deviation of reduced form parameters.

3.4.2. Estimates

This subsection begins with the discussion of (point) estimates obtained with non-robust minimum distance estimation methods. We then discuss our (confidence sets) estimates based on robust methods and show their consequences for the structural parameters of the NKPC. For each method, the discussion contrasts estimates obtained from both the baseline and the modified versions of the NKPC.

3.4.2.1. Non-robust methods: point estimates

We report the point estimates in Table 3.2. They are based on a VAR(3) for inflation and marginal cost for the baseline NKPC. For the modified NKPC, we also consider a VAR(3) for inflation, marginal cost and the time-varying inflation target.[6]

For the baseline NKPC, we note that the structural parameters γ_p and ξ_p are all significantly different from zero. The estimate of the degree of price

[6]The results are virtually the same when we consider four lags in the VAR.

stickiness, $\xi_p = 0.876$, implying that the duration of the price contract is about eight quarters. Such an implication seems inconsistent with evidence from microeconomic studies about the duration of the price contract, which report that price contracts last one to two quarters on average (see for instance Bils and Klenow (2004)). The indexation parameter, $\gamma_p = 0.534$, translates into an estimate of the backward-looking coefficient ($\gamma_b = 0.349$), broadly in line with results from other studies. For instance, Galí et al. (2005)'s estimate for the corresponding parameter is about 0.35 (see Table 1 in (Galí et al., 2005)).

Introducing a time-varying inflation target in the NKPC leads to two three main conclusions. First, estimate of the degree of price stickiness slightly increases ($\xi_p = 0.894$). Second, we cannot reject at the 5% level, the null that the indexation parameter is zero, a finding that is consistent with Cogley and Sbordone (2008) or with Ireland (2007). Not surprisingly, the estimate of the indexation parameter leads to a backward-looking term that is also not significant, again at 5% level ($\gamma_b = 0.031$). Third, the parameter λ_4, which assesses the role of the time-varying inflation target in the Phillips curve, is highly significant. Setting $\lambda_4 = 0$ in the modified NKPC leads to the baseline one. That is, our modified NKPC nests the baseline one. Thus, both specification should deliver similar results when the time-varying inflation target is not important in explaining the inflation dynamics. The fact λ_4 is highly significant tends to confirm our motivation in deriving the modified Phillips curve.

Estimates of λ, which takes into account the effect of marginal cost on inflation, are non significantly different from zero. This is in line with findings in Rudd and Whelan (2005), according to which empirical evidence on the role of the marginal cost, proxied by the labor share, is weak. By contrast, other studies, like Galí et al. (2005), do find that labor share is a key driver for

inflation.

Summarizing, Table 3.2 suggests the following conclusions: (i) the modified NKPC implies a much higher duration for the price contract than the baseline one, (ii) backward-looking component in the NKPC is not significant in the modified NKPC, while it is highly significant in the baseline one, (iii) time-varying inflation target is an important variable in explaining inflation.

To what extent are such conclusions reliable? In what follows, we provide answers based on estimates obtained with identification robust methods described in section 3.3.

3.4.2.2. Robust methods: confidence sets

Following Magnusson and Mavroeidis (2010), we compute confidence sets for the degree of price stickiness and the indexation parameter by grid search within the parameter space $\xi_p \in]0,1]$ and $\gamma_p \in [0,1]$.

Given that the $\mathcal{MD.K}$ test could deliver spurious inference, as stated in section (3.3), we will report only the results based on the $\mathcal{MD.AR}$ and $\mathcal{MD.KJ}$ Figure 3.2 reports the 90% and 95% for the $\mathcal{MD.K}$ test.

Confidence sets for the baseline NKPC of ξ_p and γ_p (Figure 3.2, top panel) suggest two conclusions. First, the indexation parameter γ_p lies roughly between 0.26 and 0.65. Thus, this parameter appears to be significantly different from zero. Second, the degree of price stickiness is significantly different from zero and lies between .8 and 1.

Turning to the modified NKPC, we note first that confidence sets for the degree of price stickiness are slightly wider than their counterpart in the baseline NKPC. In particular, they show that ξ_p lies between 0.7 and 1 for the $\mathcal{MD.AR}$ test statistic and 0.78 for the $\mathcal{MD.KJ}$, while in the baseline case, the corresponding coefficient lies between 0.8 and 1 for the two test statistics. This suggests that, introducing time-varying inflation target in the NKPC

and estimating it with identification-robust methods, deliver an estimate of the degree of price stickiness that is more reliable. For instance, we cannot reject at 5% level that $\xi_p = 0.7$. In particular, such an estimate is consistent with estimate of that parameter from DSGE models (see chapter 5). Second, both tests suggest that we cannot rule out the possibility that $\gamma_p = 0$, i.e the backward-looking term still appears to be insignificantly different from zero. However, the wide confidence set around γ_p under the $\mathcal{MD.AR}$ test statistic for the modified NKPC suggests that it is imprecisely estimated.

3.4.3. Robustness check

In this section, we provide point estimates of the baseline and modified specifications of the NKPC based on GMM. Our intention is to check our main conclusion about the indexation parameter γ_p, using alternative methods. Furthermore, this makes our results readily comparative to other studies, given that an important part of the literature uses the GMM methodology in estimating the Phillips curve.

Overall, results from the GMM estimation reported in Table 3.3 are in line with those from the minimum distance estimation. For instance, the indexation parameter, estimated from the modified NKPC, is not significantly different from zero ($\gamma_p = 0.058$ with a standard error of 0.05), whereas the same parameter is significantly different from zero under the baseline NKPC. Labor share does not drive inflation, while the forward-looking component in the NKPC is predominant.

The main difference between Table 3.2 and Table 3.3 is in the magnitude of the estimates. In particular, we estimate λ_4 to be 0.209 with a standard error of 0.07 while the same parameter is about 0.044, using minimum distance estimation. Thus, given that both estimates are significantly different from zero, our conclusion about the importance of the trend inflation for the dynamics

Table 3.3: GMM estimates of the new-Keynesian Phillips curve

	Modified NKPC		Baseline NKPC	
Parameters	Estimates	Std.err	Estimates	Std.err
γ_p	0.068	0.050	0.161	0.047
ξ_p	0.866	0.075	0.795	0.072
γ_f	0.725	0.084	0.869	0.037
γ_b	0.064	0.041	0.134	0.035
λ	0.052	0.122	0.039	0.032
λ_1	0.209	0.070		

Notes: Point estimates are derived using the CUE-GMM with Newey and West (1987) weighting matrix. Instrument are: a constant, three lags of inflation and marginal cost for the baseline NKPC while for the modified NKPC, we use additional three lags of the time-varying inflation target.

of inflation is not altered.

3.5. Concluding remarks

In this chapter, we consider the consequences of introducing time-varying inflation target in the New Keynesian Phillips curve. Our estimates lead to two main conclusions. First, the resulting NKPC is not time-varying, unlike in Cogley and Sbordone (2008). Second, ad hoc backward looking term does not matter (at least statistically). The second conclusion is found to be robust to weak identification issue. Third, time-varying inflation target is an important variable in explaining inflation dynamics, more important than labor share advocated by many empirical studies. Though these findings are encouraging, uncertainty associated with the way the time-varying inflation target is derived remains a concern. It is quite possible that other methods to extract time-varying inflation target could deliver different results. Hence, a fruitful avenue for future research is to estimate our NKPC for countries for

which time-varying inflation target is observed.

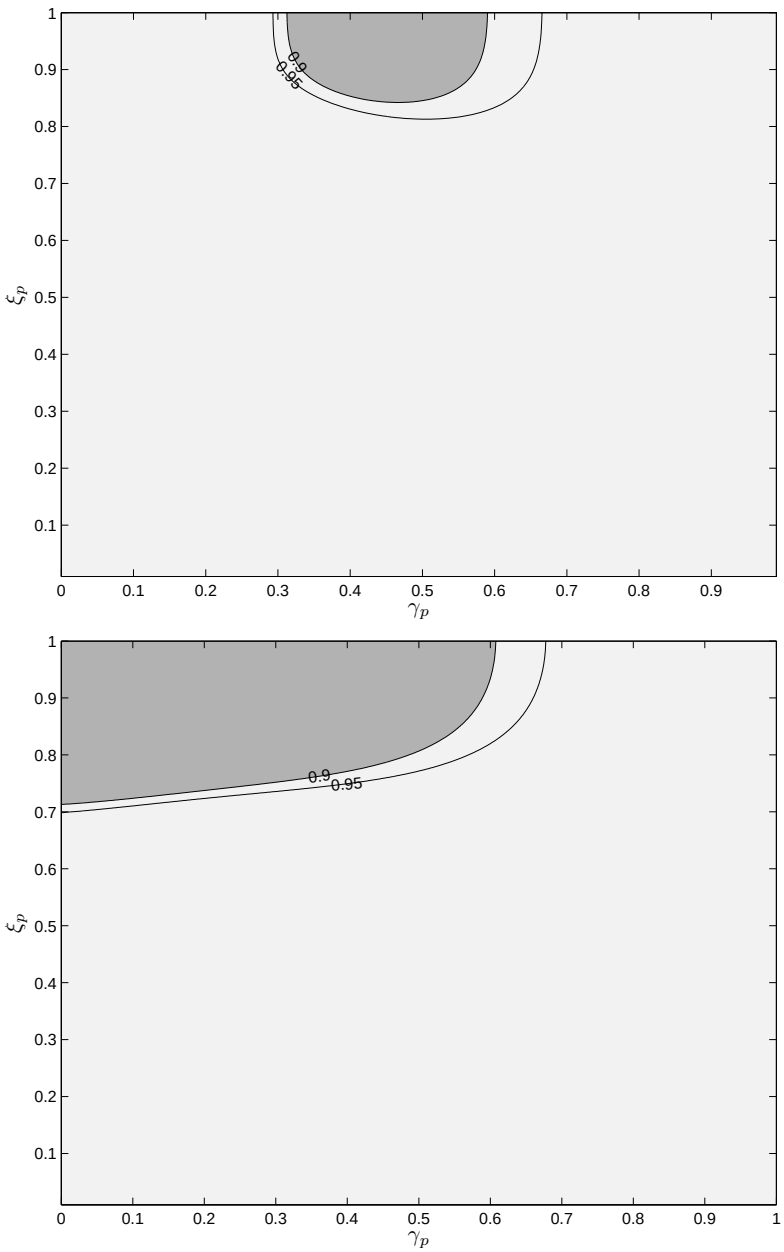

Figure 3.2: MD-AR 90% (gray) and 95% (dark) confidence sets for the baseline NKPC (top panel) and the modified NKPC (bottom panel).

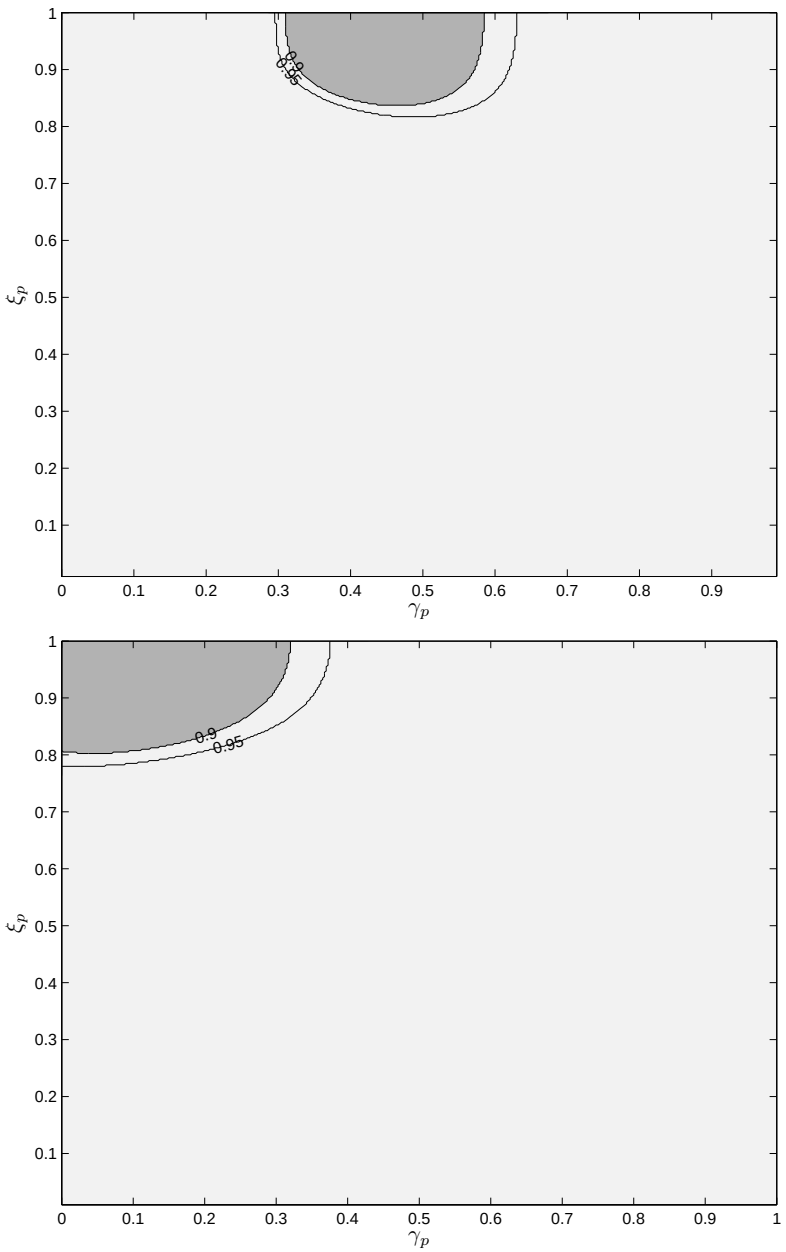

Figure 3.3: MD-KJ 90% (gray) and 95% (dark) confidence sets for the baseline NKPC (top panel) and the modified NKPC (bottom panel).

Chapter 4

A Review of Bayesian Analysis of DSGE models

Contents

4.1. Introduction

Dynamic Stochastic General Equilibrium (DSGE) models refer to micro-founded macroeconomic models that are used to address substantive economic topics. Two examples of such topics are: what is the propagation mechanism of a monetary policy shock? What shocks are mainly responsible for business cycle fluctuations? The usefulness of DSGE models for policy analysis is the primary reason for their recent diffusion in central bank and policy-making institutions.

A quick look at the different concepts forming the name "DSGE" should better clarify the structure of these models. First, these models are “dynamic” because they specify intertemporal constraints faced by agents. For instance, agents might choose to consume more today or to save more for tomorrow. Second, they are “stochastic” because they account for uncertainty in the economy. Thus, uncertainty is modelled as the occurrence of exogenous shocks, i.e deviations of some exogenous variables from their long run value that are unanticipated by agents. Finally, the concept “general equilibrium” mean that aggregate endogenous variables, such as consumption, output or price levels, are investigated within a whole economy.

Empirical properties of DSGE models have been studied using several econometric tools. The first of them was calibration, advocated by Kydland and Prescott (1982). A calibration exercise requires the following steps: first, researchers collect a set of stylized facts that the DSGE model is supposed to account for. Second, the model is parametrized (or calibrated) in order to account for a subset of these stylized facts. Conditional on the parameters value, the model is judged credible if it can account for the remaining stylized facts. The main shortcoming of a calibration exercise is that it does not attach any probabilistic measures of uncertainty to the quantitative statements

it generates (Schorfheide, 2006). To partially address this shortcoming, formal statistical tools have been considered. The generalized method of moments (GMM) (Hansen, 1982; Hansen and Singleton, 1982) and maximum likelihood estimation (Leeper and Sims, 1994) were among the first tools that researchers have used to estimate DSGE models.

However, the last decade has experienced an explosion in the number of papers using Bayesian methods. There are several reasons for this. First, Bayesian inference facilitates the incorporation of beliefs about the values of some structural parameters through the "prior". Second, medium scale DSGE are richly parametrized, while data used to estimate them are short (e.g. sixty years of US data, forty years of European data). Consequently, lack of data leads to imprecise estimates and quantitative statements generated with the estimated model. Bayesian inference helps to incorporate additional information, which leads to better estimation provided that such information comes from non-sample data, i.e data that are not used to compute the likelihood function. Another reason is the development of the so-called Markov Chain Monte Carlo (MCMC) methods in the 90's and after. Problems that were very difficult to address due to lack of mathematical tools and computer power are now easy to address, by means of MCMC.

In what follows, I give an overview of Bayesian analysis, where the focus is on DSGE models. This survey is far from exhaustive. My ultimate objective is to familiarize the non-versed reader in Bayesian econometrics with the main tools in the Bayesian analysis of DSGE models, because they are applied in Chapter 5.

The literature provides excellent surveys of Bayesian analysis of DSGE models. Frank Schorfheide, who was among the first to apply Bayesian analyis to DSGE model, offers with his coauthor An Sungbae one of the most cited papers in the field (An and Schorfheide, 2007). Fernández-Villaverde (2010)

focus on the history of the DSGE models both on a theoretical and empirical perspectives. More recently, Del Negro and Schorfheide (2010) discuss Bayesian methods applied to macroeconomics model beyond the DSGE framework. However, these surveys provide little (if any) information about Markov-Swiching DSGE models. Thus, my survey extends the previous one to include this recent development.

The remainder of the survey proceeds as follow. Section 4.2 provides the basic ideas of Bayesian econometrics and discusses the concepts of prior, likelihood and posterior. Section 4.3 presents the steps in the estimation procedure. In particular, we discuss how prior information is elicited. Then tools to compute the likelihood and summarize the posterior are presented. In section 4.4, I discuss tools to assess the model fit. Section 4.5 extends the DSGE model to feature Markov-switching process. In particular, I present how the steps in the estimation procedure are modified, due to the presence of Markov-Switches in the values of structural parameters. Section 4.6 discusses a recent alternative methodology to assess the model fit for Markov-Switching DSGE models, while section (4.7) briefly describes some convergence diagnostics, which aim to assess whether results obtained in the estimation steps are reliable. The last section briefly concludes the survey.

4.2. Basic ideas of Bayesian econometrics

Bayes's theorem is the central part of Bayesian econometrics. Before explaining this theorem, it is convenient to introduce some notations. Let $Y^T \in \mathbb{R}$ be the collection of data, Y_t the data at the period t, and $\mathcal{M}$ a model which is designed to explain some properties of the data. The model is composed of three ingredients: first, a parameter set, Θ, defining the admissible value of the parameters of the model. Second, a likelihood function $p(Y^T|\theta, \mathcal{M})$

which gives the density that the model assigns to data given some parameter values and third, a prior distribution $\pi(\theta|\mathcal{M})$ that captures information not contained in the sample used for estimation. Such information can be pre-sample beliefs, results from other studies or intuition about the sign and the magnitude of the parameters values. Bayes' theorem simply tells us that the posterior distribution of the parameters is given by:

$$\pi(\theta|Y^T, \mathcal{M}) = \frac{p(Y^T|\theta, \mathcal{M})\pi(\theta|\mathcal{M})}{p(Y^T)}. \tag{4.2.1}$$

where $p(Y^T) = \int p(Y^T|\theta, \mathcal{M})\pi(\theta|\mathcal{M})d\theta$.

In words, this result tells us how we should update our beliefs about parameter values after observing the data: combining our prior beliefs $\pi(\theta|\mathcal{M})$ with the sample information given by the likelihood function, we obtain a new set of posterior beliefs, $\pi(\theta|Y^T, \mathcal{M})$.

In (4.2.1), the quantity $p(Y_T)$ denotes the marginal data density. In the body of the paper, we will interchangeably use the terms marginal likelihood, marginal data density or marginal distribution of the data to refer to $p(Y^T)$.

4.3. The steps in the estimation procedure

Bayesian estimation of DSGE models requires five steps. Roughly speaking, these steps involve the "model preparation" and "data preparation". The model preparation requires that one specifies and solves the model, while the data preparation step implies the definition of the data that the model is supposed to match, transforming them in a way that is consistent to the model. For instance, a model where endogenous variables are stationary requires that we stationarize the data prior to the estimation stage. Following Smets and Wouters (2005), we summarize the steps in the Bayesian estimation of DSGE models as follow. The first step requires that we solve the model. Technically

speaking, a DSGE model is a non-linear system of expectational difference equations. The approach that is mostly used in the literature is to construct a linear approximation of the non-linear system around its well-defined steady state. Solving the model is no more than expressing endogenous variables as a (linear) function of their lagged values and some exogenous processes (typically the shocks). In the second step, the model solution is transformed in a state space representation. The third step exploits the state space representation to compute the likelihood function. Step four is devoted to prior elicitation. In the fifth step, prior and likelihood are combined to form the posterior, which is summarized using Markov Chain Monte Carlo (MCMC) tools. In what follows, I briefly describe each of these steps.

4.3.1. Model solutions

A DSGE model is a non linear rational expectations model system, which has the following form:

$$E_t(F(X_{t+1}, X_t, \epsilon_t) = 0, \tag{4.3.1}$$

where E_t denotes the expectation conditional on a set of information available in period t, X_t is a vector of endogenous variables of the model and ϵ_t is a vector of exogenous variables, typically the shocks of the DSGE model.

This rational expectation model has to be solved prior to estimation. The common practice is to consider a linear approximation of (4.3.1). The reason is that linear approximation methods lead to a state-space representation of the DSGE models, that is easy to analyse with filtering techniques.[1]

The solution of the rational expectations system takes the form

[1] The literature also offers non linear approximation methods such as projection methods, value-function iterations. However, I did not explored these alternative approximations in my research. Readers interested in those methods should refer to DeJong and Dave (2007).

$$X_t = \Phi(X_{t-1}, \epsilon_t; \theta) \tag{4.3.2}$$

where θ denotes the parameters of the model.

There exists a variety of numerical procedures that lead to (4.3.2). A non exhaustive list includes Blanchard and Kahn (1980), Sims (2002) and Uhlig (1997). The structure of the solution strongly depends on the parameterization adopted by the researcher: depending on the chosen parameterization, there may be cases where (i) the solution does not exist, (ii) the solution exists and is unique and stable (which is the determinacy case), (iii) multiple stable solutions exist (the indeterminacy case).

In what follows, I focus on the determinacy case in which there exists a unique and stable solution to 4.3.2.

4.3.2. State space representation

For ease of exposition, I assume that (4.3.2) is given by

$$X_t = \Phi X_{t-1} + R\epsilon_t \tag{4.3.3}$$

where the dependence on θ is dropped to simplify the notations. In the state space literature, equation (4.3.3) is called a transition equation.

Some variables in X_t are latent, i.e. they are not observable. Hence, estimating (4.3.3) directly is not possible. Instead, at time t, the researcher has an observable vector Y_t and links this vector with the state vector X_t through a set of measurement equations, i.e

$$Y_t = AX_t + D \tag{4.3.4}$$

where A is a selection matrix and D a vector of constants, which depends on

the model steady-state.

The transition and measurement equations form the state space representation of the DSGE model, which has to be estimated:

$$\begin{cases} X_t = & \Phi X_{t-1} + R\epsilon_t \\ Y_t = & AX_t + D \end{cases} \tag{4.3.5}$$

4.3.3. The likelihood function: Kalman filter

Since these observables are dependent, the likelihood function, which is functionally equivalent to the joint density of the observables, is given by

$$\begin{aligned} \ell(\theta|Y_T) \equiv p(Y_T|\theta) &= p(Y_1|\theta)\prod_{t=2}^{T} p(Y_t|Y^{t-1},\theta) \\ &= \int p(Y_1|X_1,\theta)dX_1 \prod_{t=2}^{T} \int p(Y_t|X_t,\theta)p(X_t|Y^{t-1},\theta)dX_t. \end{aligned} \tag{4.3.6}$$

An examination of equation (4.3.6) shows that we will need to compute the conditional densities $p(X_t|Y^{t-1},\theta)$, $p(Y_t|Y^{t-1},\theta)$ and $p(X_t|Y^t,\theta)$. All the densities require the knowledge of the initial distribution $p(X_1)$. The densities can be computed iteratively, using the following algorithm:

`Computation of conditional densities`

1. `Initialize the density` $p(X_1)$.
2. `Given the conditional density at period` $t-1$`, compute` $p(X_t|Y^{t-1}) = \int p(X_t|X_{t-1})p(X_{t-1}Y^{T-1})dX_{t-1}$.
3. `Compute the density of the observables as`
$$p(Y_t|Y^{t-1}) = \int p(Y_t|X_t)p(X_tY^{t-1})dX_t.$$

4. `Update the density` $p(X_t|X_{t-1})$ `with Bayes's rule, when a new series of observales is available:`

$$p(X_t|Y^t) = \frac{p(Y_t|X_t)p(X_tY^{t-1})}{p(Y_t|Y^{t-1})}.$$

5. `Repeat the steps 2-4 until the end of the sample size has been reached.`

Fortunately, when the state space representation of the model solution is linear and Gaussian, all these conditional densities are normal and the Kalman filter provides us with their means and variances at each iteration. Hence, to derive the Kalman filter, I assume that ϵ_t is normally distributed, with zero mean and variance Σ_ϵ . Define the linear projections $X_{t|t-1} = E(X_t|Y^{t-1})$ and $X_{t|t} = E(X_t|Y^t)$, where, as before, the notation $Y^t = \{Y_1, Y_2, \cdots, Y_t\}$ collects the data from the first period to the period t. Hence, $X_{t|t-1}$ is the conditional expectations based on data available until period $t-1$. In the same way, define the variance matrices $P_{t-1|t-1} = E(X_{t-1} - X_{t-1|t-1})(X_{t-1} - X_{t-1|t-1})'$ and $P_{t|t-1} = E(X_t - X_{t|t-1})(X_t - X_{t|t-1})'$.

At each iteration, the following computations are executed.

First, we compute a forecast of X_t:

$$X_{t|t-1} = \Phi X_{t-1|t-1}, \tag{4.3.7}$$

and the variance of the forecast error:

$$P_{t|t-1} = \Phi P_{t-1|t-1}\Phi' + R\Sigma_\epsilon R'. \tag{4.3.8}$$

Define the one-step ahead forecast error by

$$\nu_{t|t-1} = Y_t - Y_{t|t-1} = Y_t - A - BX_t$$

and its variance matrix by

$$F_{t|t-1} = BP_{t|t-1}B'.$$

The loglikelihood is already available at this step and reads

$$\log(\ell(Y_t|\theta) = -\frac{n}{2}\log 2\pi - \frac{1}{2}\log|F_{t|t-1}| - \frac{1}{2}\nu_t' F_{t|t-1}^{-1}\nu_t. \tag{4.3.9}$$

Next, we update the forecast of X_t, denoted by $X_{t|t}$. Hamilton (1994) shows that the updating forecast of X_t is given by

$$X_{t|t} = X_{t|t-1} + K\nu_{t|t-1}, \tag{4.3.10}$$

where the matrix K is the Kalman gain and reads

$$K = P_{t|t-1}B'(F_{t|t-1})^{-1}.$$

The updated variance of the forecast error reads

$$P_{t|t} = P_{t|t-1} - KBP_{t|t-1}.$$

This completes one iteration. Doing these computations until the end of the sample size delivers the log likelihood function.

4.3.4. Prior elicitation

Priors play a key role in the Bayesian estimation of DSGE models. The incorporation of prior information is perhaps the main point of disagreement between the frequentist and Bayesian approach. As emphasized by Del Negro and Schorfheide (2010), the use of fairly informative priors should not be interpreted as "cooking up desired results based on almost dogmatic prior". As

emphasized in the introduction, the philosophy behind the incorporation of prior information into the analysis is to use other sources of information not contained in the likelihood function. Thus, precise information is translated into fairly tight priors on the parameters. This has nothing to do with imposing dogmatic priors in order to produce results that match the analyst's desire.

However, as priors may affect the posterior estimates and model comparison, their specification requires some care and has to be precisely documented. Three sources of extraneous information are exploited when forming priors in the literature on DSGE models. The first is to use macroeconomic series to extract information, not contained in the likelihood, that the researcher finds useful given the empirical facts the model is supposed to match. For instance, if the model is fitted to data on, say, output growth, inflation and interest rate, data on labor share could be use to estimate the labor share of income in the model if one is interested in estimating such a parameter. Second, information from micro-economic studies can be used to shape the prior. For instance, there is microeconomic evidence that firms adjust their price every one to two quarters (Bils and Klenow, 2004). Such information could be used to form a prior on the probability for firms to adjust their price, under the Calvo price setting. Third, macroeconomic data, including those entering the likelihood function, can be used to shape the prior provided they come from a presample. For instance, if the researcher thinks that US monetary policy is well characterized by a Taylor-type rule and she wants to estimate a DSGE model based on post-1982 data, she can use pre-1982 data to form prior for the parameters of the Taylor-type rule. This is also true for parameters related to shock processes: those parameters can be chosen such that the implied dynamic of the model matches those of the presample data.

4.3.5. Posterior distribution

Once we compute the likelihood function and specify the prior, we can apply Bayes' theorem to compute the posterior. We will consider an algorithm to generate draws from the (non-normalized) posterior distribution of θ. From equation (4.2.1), we know that the posterior is given by

$$\pi(\theta|Y^T) = \frac{p(Y^T|\theta)\pi(\theta)}{\int p(Y^T|\theta)\pi(\theta)d\theta} \propto p(Y^T|\theta)\pi(\theta), \tag{4.3.11}$$

where the denominator is an integrating constant. This algorithm requires the evaluation of the likelihood times the prior, which is computed according to methods in sections (4.3.4) and (4.3.3). Because it is difficult to analytically the denominator in 4.3.11, so is the posterior. Thus, the best we can do is to proceed by simulations. It is achieved using a powerful tool known as Markov chain Monte Carlo (MCMC) methods. The aim of these methods consists in generating a Markov chain with ergodic distribution $\pi(\theta|Y^T)$. There are many alternatives for doing so. Following the literature, we will use a Random Walk Metropolis Algorithms, based on Schorfheide (2000).

I will omit deep and technical details about MCMC methods. The interested reader should have a look at Schorfheide (2000) or An and Schorfheide (2007) or any Bayesian econometrics textbook. In the following, I follow the excellent intuitive description found in Fernández-Villaverde (2010). We do not know what the whole posterior $\pi(\theta|y^T)$ is but we want to simulate from a Markov chain and approximate the whole posterior by the empirical distribution generated by the chain. Put in another way, we want to produce a Markov chain whose stationary distribution is $\pi(\theta|y^T)$. To do so, we require tools which allow us to construct a Markov chain. The Metropolis algorithm is one of the tools we have for doing so. Roughly speaking, this algorithm specifies a new proposed value of the parameter and evaluates whether it increases the

posterior, i.e if the posterior density evaluated at this new proposed value is greater than the posterior evaluated at the current value of θ. If it does, we accept it with probability one, and with some probability less than one if it does not. This allows us to go towards the higher regions of the posterior but we also travel with some probability, towards the lower regions. In doing so, all the parameter space is explored, thus avoiding getting "stuck" around specific values of the parameters. For the illustration, I will use the following implementation of the random walk version of this algorithm, which can be found in Schorfheide (2000) or more recently in An and Schorfheide (2007)[2].

`One-block Random Walk Metropolis Algorithm`

1. `Find the posterior mode of` $\ln p(\theta|y^T) + \ln \pi(\theta)$ `via a numerical optimization routine and denote it by` $\widetilde{\theta}$.

2. `Let` $\widetilde{\Sigma}$ `be the inverse of the Hessian computed at the posterior mode` $\widetilde{\theta}$.

3. `Draw` θ^0 `from` $\mathcal{N}(\widetilde{\theta}^0, c_0^2\widetilde{\Sigma})$ `or directly specify a starting value where` c_0 `is a scale parameter.`

4. `For` $s = 1, \cdots, n_{sim}$, `draw` ϑ `from the proposal density` $\mathcal{N}(\theta^{(s-1)}, c^2\widetilde{\Sigma})$, `with` c_2 `being a scale parameter.`
   `The jump from` $\theta^{(s-1)}$ `is accepted` $(\theta^{(s)} = \vartheta)$ `with probability` $\min\{1, r(\theta^{(s-1)}, \vartheta|y^T)\}$. `Here,`

$$r(\theta^{(s-1)}, \vartheta|y^T) = \frac{p(\vartheta|y^T)\pi(\vartheta)}{p(\theta^{(s-1)}|y^T)\pi(\theta^{(s-1)})}.$$

Under general regularity conditions, the posterior of θ will be asymptotically normal. Therefore, this algorithm constructs a Gaussian approximation

[2]The random walk denomination of this algorithm comes from the fact that the proposal density is specified as a random walk.

around the posterior mode $\widetilde{\theta}$ where a scaled negative inverse Hessian is used as the covariance for the proposal distribution. As the RWMH algorithm requires maximization of the posterior, we can face the problems inherent to pure maximum likelihood estimation if data are not informative. Hence steps 1 and 2 of the algorithm, although often useful, are not necessary since the algorithm can be initialized with values that would be retained for a calibration exercise.[3]

The draws generated from the posterior are used to obtain point estimates such as mean, variance, median and so on.

[3]Most of the empirical literature on Bayesian estimation of DSGE model uses the one-block RWMH algorithm. But this version of the algorithm is not always efficient : for instance, An and Schorfheide (2007) finds that when posterior distribution is bimodal, say a low and high mode, the RWMH algorithm is unable to jump from one mode to the another one. Chib and Ramamurthy (2010) proposes a new MCMC methods called Tailored-Randomized block Metropolis-Hastings algorithm, to estimate DSGE models. The motivation of this method is that the single block Metropolis-Hastings algorithm face difficulty to achieve convergence when the dimension of the vector of parameters is large. The paper proposes to cluster the DSGE model parameters in a random number of blocks at each iteration. Then, each block of parameters is updated with a tailored proposal density that mimics the target density of that block. One finding of the paper is that with this algorithm, jumping between low and high mode is possible, unlike with the RWMH algorithm. However, the approach is time-consuming. For a six equations model as the one estimated by An and Schorfheide (2007), it takes around 30 hours in an ordinary computer and "only" 3 hours with the RWMH algorithm (where the posterior mode is obtained after 24 hours). If the researcher has good reasons to start the RWMH algorithm in particular values other than the posterior mode, as aforementioned, investment in the TaRB-MH algorithm will be somewhat unnecessary.

4.4. Model evaluation

By model evaluation, we mean the assessment of the model fit. It can be done using measures of absolute fit such as posterior predictive checks, or the measure of fit relative to another model, either DSGE or VAR model. The measure of relative fit is done by either enlarging or restricting a model in some dimensions and assessing whether the data prefers or not such modification with respect to a benchmark case. Methods for doing so are presented in the next section.

Bayesian model comparison is conducted as follows. We assign prior probabilities to two competing models $\mathcal{M}_i$ and $\mathcal{M}_j$ and update the prior probabilities through the marginal likelihood ratios, according to

$$\frac{\Pr(\mathcal{M}_i|Y^T)}{\Pr(\mathcal{M}_j|Y^T)} = \frac{p(Y^T|\mathcal{M}_i)}{p(Y^T|\mathcal{M}_j)}\frac{\Pr(\mathcal{M}_i)}{\Pr(\mathcal{M}_j)}. \tag{4.4.1}$$

where $\Pr(\mathcal{M}_x|Y^T)$, $p(Y^T|\mathcal{M}_x)$ and $\Pr(\mathcal{M}_x), x = i, j$ are respectively the posterior model probability, the marginal data density and the prior model probability. When equal prior probability is assigned to the competing models, the conduct of model comparison reduces to the computation of the marginal likelihood ratio, also called the Bayes factor.

Computing the marginal data density is very challenging, given that it involves high-dimensional integral. There are various methods to approximate it. Throughout my research, I have used the modified harmonic mean (MHM) method proposed by Geweke (1999) and the Laplace approximation.

The method by Geweke (1999) relies on the harmonic mean to approximate the marginal likelihood:

$$p(Y^T|\mathcal{M}_i)^{-1} = \int \frac{h(\theta_{(i)}}{p(Y^T|\theta_{(i)}, \mathcal{M}_i)p(\theta_{(i)})} p(\theta_{(i)}|Y^T)d\theta_{(i)} \tag{4.4.2}$$

where $h(\theta)$ is a weighting function whose support has to be contained in the support of the posterior distribution. Geweke (1999) proposes the use the density of a truncated multivariate Gaussian distribution:

$$\begin{aligned} h(\theta) =& \frac{1}{p}(2\pi)^{-n/2}|\Sigma_\theta| \exp(-\frac{1}{2}(\theta-\overline{\theta})'\Sigma_\theta^{-1}(\theta-\overline{\theta})) \\ & \times \Pr[(\theta-\overline{\theta})'\Sigma_\theta^{-1}(\theta-\overline{\theta}) \leq F_{\chi_n^2}^{-1}(p)] \end{aligned} \tag{4.4.3}$$

where $p \in (0,1)$, Σ_θ is the posterior variance matrix obtained with the Metropolis algorithm, $\overline{\theta}$ is the posterior mean of θ and $F_{\chi_n^2}^{-1}$ is the cumulative density of a χ^2 distribution with n degree of freedom, where n is the dimension of θ.

Denote

$$m(\theta(i)) = \frac{h(\theta_{(i)})}{p(Y^T|\theta_{(i)}, \mathcal{M}_i)p(\theta_{(i)})}.$$

A numerical evaluation of the integral is achieved through Monte Carlo integration:

$$p(Y^T|\mathcal{M}_i)^{-1} = \frac{1}{N}\sum_{i=1}^{N} m(\theta_{(i)}) \tag{4.4.4}$$

where N is the number of simulations.

An alternative and straightforward way to approximate the marginal likelihood is the Laplace approximation. This approach is followed, for instance, by Schorfheide (2000). It is only valid when the likelihood function peaks at value around the posterior mode. Thus, the density kernel can be locally approximated by a multivariate Gaussian density:

$$\begin{aligned} \log p(\theta_{(i)})p(\theta_{(i)}|Y^T) \approx& \log p(\widehat{\theta}_{(i)})p(\widehat{\theta}_{(i)}|Y^T) \\ & + \frac{1}{2}(\theta_{(i)} - \widehat{\theta}_{(i)})'\Sigma_{\widehat{\theta}_{(i)}}(\theta_{(i)} - \widehat{\theta}_{(i)})' \end{aligned} \tag{4.4.5}$$

The estimator of the marginal likelihood is obtained by integration:

$$p(Y^T|\mathcal{M}_i) = (2\pi^{n/2}|\Sigma_{\widehat{\theta}_{(i)}}|^{1/2}p(\widehat{\theta}_{(i)})p(\widehat{\theta}_{(i)}|Y^T). \tag{4.4.6}$$

4.5. An extension: Markov-Switching DSGE models

There is a recent and growing literature on Markov-Switching DSGE (MS-DSGE) models, estimated with Bayesian methods. In this literature, a key paper that greatly improves our understanding of Markov-Switching DSGE models is Farmer et al. (2009b). The aim of Markov-Switching models is to capture sudden changes in the time-series dynamics of the data because "the world is changing" (Farmer et al., 2010). For instance, it is well known that the volatility of US macroeconomic series, typically inflation, output growth and interest rate have experienced dramatic decline during the post-Volcker era. This striking phenomenon is termed "Great Moderation" by economists and it is the object of chapter 5 of the thesis. A non exhaustive list of papers that estimate MS DSGE models with Bayesian methods are Schorfheide (2005), Bianchi (2011), Liu et al. (2011), Davig and Doh (2009) and Liu and Mumtaz (2010).

In the presence of Markov-Switching process, the steps outlined in estimating the DSGE models are still valid. However, in general the solution of the model is not straightforward and the computation of the likelihood function requires modifications of the standard Kalman filter. The number of papers that propose algorithms to solve these model is also growing. Farmer et al. (2008) consider a method to solve MS-DSGE model, which consists in rewriting the DSGE model into a fixed coefficient model. The advantage of their method is that with little modification, it can be solved with standard numerical procedures along the lines of Sims (2002). Furthermore, they show that

when a minimal state variable (MSV) solution to the new system exists, it is also a MSV solution to the original system. Farmer et al. (2010) propose a rather different algorithm, compared to previous versions of their paper. Their method is able to find all the solutions of the MS DSGE models, while, as they emphasize, other methods proposed by the literature (Davig and Leeper, 2007; Svensson and Williams, 2007) are not able to do so. We have tried to estimate the influential Smets and Wouters (2007) with the algorithm proposed by Farmer et al. (2010). Our experience suggests that the time taken by the algorithm to solve the model is reasonable when the model to be estimated is the standard three equations DSGE model documented in Woodford (2003)'s textbook. For a model of the kind of Smets and Wouters (2007), the algorithm is very time consuming. Consequently, I used an algorithm proposed by Dufourt (2011), which can be viewed as a generalization of the Svensson and Williams (2007) algorithm. The algorithm writes the model solution as

$$X_t = G_1(s_t)X_{t-1} + \Pi(s_t)\epsilon_t + L(s_t) \tag{4.5.1}$$

where $L(s_t)$ is a regime-switching constant and s_t denotes a Markov-Switching state. The main difference with (4.3.3) is that the matrices in (4.5.1) depend on the active regime s_t, which is defined through a probability transition matrix.

4.5.1. Kim's approximation of the likelihood

The monograph by Kim and Nelson (1999) provides the main tools I have used to estimate MS-DSGE models. To compute the likelihood, I adopt the algorithm of Kim and Nelson (1999) designed to cope with the presence of regime switching state s_t.

The precise reason for using this algorithm is that the number of trajectories to consider in a framework of regime-switching grows exponentially with time.

To understand why the standard Kalman filter cannot be used, consider a two regime DSGE models, i.e $s_t = 1, 2$. At the next iteration of the Kalman filter, each of the Gaussians will be propagated through 2 other Gaussians. Thus, at the next iteration, the distribution of X_t is a mixture of $4 = 2^2$ Gaussians. In general, at the tth iteration, the distribution of X_t is a mixture of 2^t Gaussians. For instance, when $t = 10$, the distribution of X_t is a mixture of more than 1000 Gaussians, making the Kalman filter inoperable. To deal with this exponential growth, the literature suggests various approximating mehods: collapsing some mixture components at the end of each operation or using a finite mixture components. In this paper, I mainly use Kim (1994)'s approximation described in details in Kim and Nelson (1999). To check the robustness of the results, we also use a finite mixture approximation. For the latter method, please refer to Schorfheide (2005).

The Kim algorithm works as follow. First, we compute the Kalman filter for every regime combination, according to section (4.3.3). That is, we run the following recursion:

1. Run the Kalman filter as follows:

$$
\begin{aligned}
X^{(i,j)}_{t|t-1} &= \Phi(j)X^{i}_{t-1|t-1} + L(j) \\
P^{(i,j)}_{t|t-1} &= \Phi_0(j)P^{(i)}_{t-1|t-1}\Phi_0'(j) + R(j)R\Sigma_j'(j) \\
\nu^{(i,j)}_{t|t-1} &= Y_t - D - AX^{(i,j)}_{t|t-1} \\
F^{(i,j)}_{t|t-1} &= AX^{(i,j)}_{t|t-1}A' \\
X^{(i,j)}_{t|t} &= X^{(i,j)}_{t|t-1} + P^{(i,j)}_{t|t-1}A'(F^{(i,j)}_{t|t-1})^{-1}\nu^{(i,j)}_{t|t-1} \\
P^{(i,j)}_{t|t} &= P^{(i,j)}_{t|t-1} - P^{(i,j)}_{t|t-1}A'(F^{(i,j)}_{t|t-1})^{-1}A'\nu^{(i,j)}_{t|t-1}.
\end{aligned}
$$

2. Run the Hamilton filter. This requires the computation of filtered prob-

abilities $\Pr(s_t, s_{t-1}|Y_t)$ and $\Pr(s_t|Y_t)$, for $i, j = 1, 2.$:

$$\begin{aligned}
\Pr(s_t, s_{t-1}|Y_{t-1}) &= \Pr(s_t|s_{t-1}) \Pr(s_{t-1}|Y_{t-1}) \\
f(Y_t|Y_{t-1}) &= \sum_{s_t} \sum_{s_{t-1}} f(Y_t|s_t, s_{t-1}, Y_{t-1}) \Pr(s_t, s_{t-1}|Y_{t-1}) \\
\Pr(s_t, s_{t-1}|Y_t) &= \frac{f(Y_t, s_t, s_{t-1})|Y_{t-1}}{f(Y_t|Y_{t-1})} = \frac{f(Y_t|s_t, s_{t-1}), Y_{t-1}) \Pr(s_t, s_{t-1}|Y_{t-1})}{f(Y_t|Y_{t-1})} \\
\Pr(s_t|Y_t) &= \sum_{s_{t-1}} \Pr(s_t, s_{t-1}|Y_t)
\end{aligned}$$

In words, we compute the filtered probabilities given the transition matrix and the initial probabilities $\Pr(s_{t-1}|Y_{t-1})$. The marginal density of Y_t (or the likelihood value at the the tth iteration) is given by $f(Y_t|Y_{t-1})$. Next, we compute the joint probability of the regime using the Baye's rule, which is then used to update the filtered probabilities when a new realization of Y_t is available.

3. Run the collapsing procedure. The likelihood approximation appears at this step. Use the probabilities in the previous steps to collapse 2×2 posteriors $X_{t|t}^{(i,j)}$ and $P_{t|t}^{(i,j)}$ into 2×1 with the following equations:

$$\begin{aligned}
X_{t|t}^{j} &= \frac{\sum_{i=1}^{2} \Pr(s_{t-1} = i, s_t = j|Y_t) X_{t|t}^{(i,j)}}{\Pr(s_t = j|Y_t)} \\
P_{t|t}^{j} &= \frac{\sum_{i=1}^{2} \Pr(s_{t-1} = i, s_t = j|Y_t)(X_{t|t}^{j} - X_{t|t}^{(i,j)})(X_{t|t}^{j} - X_{t|t}^{(i,j)})'}{\Pr(s_t = j|Y_t)}
\end{aligned}$$

Once the model is estimated, quantities such as smoothed and filtered probabilities are readily available.

4.6. Alternative methods to compute marginal data density

Sims et al. (2008) modify the method proposed by Geweke (1999) when it

is applied to Markov-Switching DSGE models. They point out that when the model parameters are time-varying, as it is the case here, the posterior distribution tends to be non Gaussian. They distinguish three aspects characterizing the non Gaussian shape of the posterior. First, when the posterior has multiple peaks, the density at the posterior mean can be very low. Second, the truncated weighting function used by Geweke (1999) tends to be a poor local approximation of the posterior density. Third, the posterior tends to be very close to zero in the interior points of the parameter space. To address these problems, Sims et al. (2008) propose a family of elliptical distributions as weighting function. Liu et al. (2011) also contain new methods to estimate the marginal data density.

4.7. Convergence diagnostics

It is important to verify that the posterior simulator converges to its ergodic distribution in order to ensure that the results obtained are reliable. For doing so, the literature offers both formal and informal methods. Formal methods we have used include the potential scale reduction factor (PSRF) proposed by Brooks and Gelman (1998), an the numerical standard error (NSE) and relative numerical efficiency (RNE) of Geweke (1992). Among informal methods, researchers typically run the MCMC with different starting values, leading to different chains and verify that each chain converges to the same distribution. It is of common practice to consider graphical methods, including the plot of the MCMC draws and the computation of recursive means.

4.8. Conclusion

This survey gives an overview of Bayesian methods used to estimate constant and regime switching DSGE models. The next chapter applies these methods to estimate a MS-DSGE model.

Chapter 5

Great Moderation and endogenous monetary policy switches

Contents

5.1. Introduction

From the mid-80s until the recent financial crisis, there has been a substantial decline in the volatility of various macroeconomic series, particularly inflation and output growth, both in the US and in other major industrial countries (including European countries). The literature has employed the special term "Great Moderation" to describe this striking phenomenon.

Economists have suggested three explanations of the Great Moderation. The first type suggests that structural changes in the economy, such as changes in economic institutions and in technology, have improved the ability of the economy to absorb shocks (Stock and Watson, 2003b), hence contributing to moderate economic fluctuations.

The second type relies on the so-called "Good policy" view according to which improved policy, particularly monetary policy, is the primary source of the Great Moderation. Proponents of this view observe that a reduction in volatility of inflation occurred simultaneously with a reduction in the volatility of output. Since there is a broad consensus that monetary policy plays a crucial role in stabilizing inflation, a reduction in the volatility of output may have been the result of better monetary policy (Bernanke, 2004; Lubik and Schorfheide, 2004; Clarida et al., 2000).

The third view argues that smaller exogenous shocks in the 80's might have helped the economy to become more stable. According to this view, the Great moderation is mainly the result of "Goood luck" rather than the result

of structural changes or improved policy (Stock and Watson, 2003b; Sims and Zha, 2006; Justiniano and Primiceri, 2008).

While each explanation contains elements of truth, as noted in Bernanke (2004), the "Good policy" hypothesis has been considered for a long time as the best explanation of the Great Moderation. Studies that support this view argue that during the 1960's and 1970's – The Great Inflation –, monetary policy has been insufficiently aggressive against inflation, while it has become more aggressive with the appointment of Volcker as Chairman at the Fed (Lubik and Schorfheide, 2004; Clarida et al., 2000). A monetary policy that is insufficiently aggressive leads to multiple equilibria, where some of these equilibria are characterized by large "sunspot" shocks, i.e, shocks that are unrelated to economic fundamentals. Large "sunspot" shocks lead to high variances of inflation and output. By contrast, when the monetary policy is sufficiently aggressive, rational agents understand that in response to an inflationary shock, the monetary authority will act aggressively to dampen its consequences on the economy. Thus, a direct consequence of the "Good policy" view is that the decline in volatility should persist as long as monetary policy continues to be "good".

However, besides the recent financial crisis that casts some doubts on the "Good policy" explanation, several recent papers, ranging from purely empirical (Sims and Zha, 2006) to more structural (DSGE) papers (Justiniano and Primiceri, 2008; Liu et al., 2011), have provided empirical evidence in favour of the "Good luck" hypothesis. For instance, Sims and Zha (2006) find no regime changes in US monetary policy. They conclude that the source of the Great Moderation is mainly the outcome of a reduction in the variance of shocks. They reach this conclusion using structural VAR. Unlike Sims and Zha (2006), Schorfheide (2005) considers the DSGE framework and reports considerable changes in US monetary policy. Schorfheide obtains such results

using a small scale DSGE model.

More recently, three papers have tried to address the scale effect problem in Schorfheide (2005). The first one is the paper by Liu et al. (2011). They consider a medium scale DSGE model along the lines of Christiano et al. (2005) and Smets and Wouters (2007), where monetary policy switches regime through the inflation targeted by the Fed. By means of counterfactual experiments, the authors find little evidence that change in the inflation target is the main driving force of the rise and fall in inflation. On the contrary, the role played by shocks processes is substantial. They conclude that "the shocks processes are more likely to be the main driving force of the rise and fall in inflation than changes in the inflation target". The second one is Bianchi (2011). He uses a model along the line of Christiano et al. (2005) and Justiniano et al. (2011), where both policy and variance parameters evolve as a Markov-switching process. His findings suggest, *inter alia*, that the 'Good policy' explanation is likely to be preferred by the data. Finally, Fernández-Villaverde et al. (2010) fit a non-linear medium scale DSGE model to U.S. data, where they seek to understand the role of stochastic volatility versus the role of changes in monetary policy rule in accounting for the Great Moderation in the U.S. aggregate data. They report strong evidence of changes in US monetary policy but such changes do not matter much for the Great Moderation. However, the time-consuming nature of their methodology forces researchers to estimate linearized instead of non-linear DSGE models.

This chapter proposes an alternative strategy, which is also developed in Dufourt and Jean-Baptiste (2011a,b).

We consider a medium-scale DSGE model along the lines of Smets and Wouters (2007), where policy and shock variance parameters switch regimes. In this sense, our strategy is similar to Bianchi (2011) and Liu et al. (2011). However, we emphasize some key differences. First, Liu et al. (2011) model

changes in policy as changes in the inflation target in the policy rule, while we allow both inflation target and reaction coefficients in the policy rule to switch regimes. Second, the model considered in Bianchi (2011) is a simplified version of the model we consider, where wage is flexible and there are no switches in inflation target. A third key difference with respect to Bianchi (2011) is that we consider synchronized as well as independent regimes in policy and the variance of shocks, while Bianchi (2011) considers only independent regimes. This consideration allows us to study to what extent changes in monetary policy in the U.S. and in the Euro Area are endogenous.

We estimate three specifications of our baseline DSGE model, with U.S. and Euro data. The first one considers only changes in the variance of shocks. The second one considers synchronized changes in both policy and shock variances. This means that, in periods of low volatility, the monetary authority is constrained to react strongly to deviations of inflation from its long run target level. We label this regime "the hawkish regime". On the contrary, a period of high volatility forces the monetary authority to conduct an "inflation accommodating" policy, which we call "the dovish regime". Finally, the third specification allows changes in policy and shock variances to be independent. That is, changes in policy regimes are independent of the current state of the economy. We define three criteria for changes in monetary policy regimes to be endogenous. First, the independent regimes and synchronized regimes specifications have similar fits. Second, changes in monetary policy occur (approximately) simultaneously with changes in the variance of shocks. Third, in the independent regimes specification, the following two regimes should almost never occur: (i) the hawkish regime in the presence of high volatility and (ii) the dovish regime in the presence of low volatility.

Our findings are the following. First, in terms of fit, the specification where policy parameters are allowed to switch regimes dominates the specification

where only shock variances switch regimes. This finding holds for both the U.S. and the Euro Area. Second, the specification with synchronized regime shifts in policy and shocks variance better fits the data than the independent regime shifts specification, again for both economies. Third, with respect to the question of endogenous monetary policy, only the first criterion is satisfied for the U.S., since the independent and synchronized regimes produce similar fits (see Table 5.3). Smoothed estimates of regime probabilities suggest that changes in U.S. monetary policy seem to be closely related to the personality of the Chairman, and not to the current state of the economy. That is, under the Martin era, the U.S. monetary policy has been hawkish against inflation, while the period under the Burns-Miller chairmanship has been a dovish regime. Finally, the Volcker-Greenspan-Bernanke era had been hawkish. We do find that, for the U.S. economy, the "hawkish-high volatility regime" occurred frequently in the 60's and the mid-80's, whereas the "dovish-low volatility" occurred in the 70's and the early 80's.

For the Euro Area, the same cannot be said. Our three criteria are satisfied. Specifically, in periods of low volatility we find that the Euro Area monetary policy is always hawkish. Inversely, the monetary policy is dovish when volatility is high. We thus conclude that the Euro Area monetary policy is endogenous.

The remainder of the chapter is organized as follows. In Section 5.2, we briefly present the model and discuss the solution in the presence of regime switching while additional details are provided in the Appendix. In Section 5.3, we present the econometric tools used to estimate the model. Section 5.4 presents the parameter estimates and their economic implications through impulse responses, variance decompositions and regime probabilities. Section 5.5 presents results for the Euro Area. Section 5.6 summarizes our results on the issue of monetary policy endogeneity . Finally, the last Section concludes.

In Appendix A.2, we provide additional results that are not reported in the main text due to space considerations.

5.2. The model

The model we consider is adapted from Smets and Wouters (2007) but extended to feature regime switches in the variance of shocks and in monetary policy. The model incorporates various nominal and real frictions such as monopolistic competition in goods and labor markets, sticky prices and wages, partial indexation of price and wages, habit persistence, investment adjustment costs, variable capacity utilization. We deviate from Smets and Wouters (2007) by letting the inflation target depend on the regime in place in the current period. This departure has some implications for the dynamic equations governing inflation, real wages and monetary policy.

In what follows, we present the log-linearised version of the model where we describe the aggregate demand side, the aggregate supply side and the monetary policy. Hatted variables denote percentage deviations with respect to steady state. Details of the model derivation are provided in appendix A.2 (see also (Smets and Wouters, 2007)).

5.2.1. The log-linearized model

The aggregate resource constraint is given by

$$\widehat{Y}_t - c_y\widehat{C}_t - i_y\widehat{I}_t - r_y\widehat{r}_t - \widehat{\varepsilon}_t^g = 0 \tag{5.2.1}$$

with $i_y = (\gamma_z - 1 + \delta)\, k_y$, $c_y = (1 - i_y - g_y)$ and $r_y = r^k k_y z_1$, where k_y is the steady-state capital to output ratio, r^k is the steady-state real rental rate of capital, $z_1 = (1-\psi)/\psi$, where ψ is a positive function of the elasticity

of the capital utilization adjustment cost function. Output, consumption, investment, capital-utilization rate and exogenous spending shock are denoted by $\widehat{Y}_t$, $\widehat{C}_t$,$\widehat{I}_t$, $\widehat{z}_t$ and $\widehat{\varepsilon}_t^g$., respectively.

The law of motion for the exogenous spending shock as well as for other shocks is defined later on.

The second equation describes the dynamics of consumption and is given by

$$\begin{aligned}\widehat{C}_t &= c_1\widehat{C}_{t-1} + (1-c_1)\,E_t\widehat{C}_{t+1} + c_2\left[\widehat{L}_t - E_t\widehat{L}_{t+1}\right] \\ &\quad -c_3\left[\widehat{R}_t - E_t\widehat{\pi}_{t+1} + \widehat{\varepsilon}_t^b\right],\end{aligned} \tag{5.2.2}$$

where

$$c_1 = [h/\gamma_z]\,/\,[1+h/\gamma_z]\,,$$

$$c_2 = [(\sigma_c - 1)\,wl/((1+\lambda_w)c)]\,/\,[\sigma_c\,(1+h/\gamma_z)]\,,$$

$$c_3 = [1-h/\gamma_z]\,/\,[\sigma_c\,(1+h/\gamma_z)]\,,$$

with h denoting an external habit formation, γ_z the steady-state growth rate, σ_c the elasticity of intertemporal substitution between labour and consumption, λ_w the steady-state wage mark-up, w, l, c, the steady states of wage ($\widehat{w}_t$), hours worked ($\widehat{L}_t$) and consumption. $\widehat{R}_t$, $\widehat{\pi}_t$ and $\widehat{\varepsilon}_t^b$ are the nominal interest rate, inflation rate and a risk-premium shock.

The dynamics of investment are given by

$$\widehat{I}_t = i_1\widehat{I}_{t-1} + (1-i_1)\,E_t\widehat{I}_{t+1} + i_2\widehat{Q}_t + \widehat{\varepsilon}_t^i, \tag{5.2.3}$$

with $i_1 = \left(1/1+\beta\,(\gamma_z)^{1-\sigma_c}\right)$ and $i_2 = \left(1/\left(1+\beta\,(\gamma_z)^{1-\sigma_c}\right)(\gamma_z)^2\,\varphi\right)$, where φ is the steady-state elasticity of the capital adjustment cost function, β is the discount factor applied to households and $\widehat{\varepsilon}_t^i$ is an investment-specific shock.

The Tobin's Q-equation is given by

$$\widehat{Q}_t = q_1E_t\widehat{Q}_{t+1} + (1-q_1)\,E_t\widehat{r}_{t+1} - \left(\widehat{R}_t - E_t\widehat{\pi}_{t+1} + \widehat{\varepsilon}_t^b\right), \tag{5.2.4}$$

with $q_1 = (1-\delta) / \left(1-\delta+r^k\right)$, where δ is the depreciation rate of capital.

Turning to the supply side, the aggregate production function is given by

$$\widehat{Y}_t = \left(1+\frac{\phi}{Y}\right)\left(\alpha\left(\widehat{K}_{t-1}+z_1\widehat{r}_t\right)+(1-\alpha)\widehat{L}_t+\widehat{\varepsilon}_t^A\right), \tag{5.2.5}$$

where ϕ/Y is the share of fixed-costs in production, α captures the share of capital in production and $\widehat{\varepsilon^A}_t$ is the total factor productivity.

The accumulation of capital $\widehat{K}_t$ is given by

$$\widehat{K}_t = k_1\widehat{K}_{t-1}+(1-k_1)\,\widehat{I}_t+k_2\widehat{\varepsilon}_t^i \tag{5.2.6}$$

with $k_1 = (1-\delta)/\gamma_z$ and $k_2 = (1-(1-\delta)/\gamma_z)\,(1+\beta\gamma_z^{1-\sigma_c})\,\gamma_z^2\varphi$.

The price mark-up $\widehat{\mu}_t^p$, corresponding to the difference between the average price and the nominal marginal cost, is given by

$$\widehat{\mu}_t^p = \alpha\left(\widehat{K}_{t-1}+z_1\widehat{r}_t-\widehat{L}_t\right)-\widehat{w}_t+\widehat{\varepsilon}_t^A. \tag{5.2.7}$$

Profit maximization by price-setting firms gives rise to the following New-Keynesian Phillips curve (NKPC):

$$\widehat{\pi}_t = \pi_1\left(\gamma_p\widehat{\pi}_{t-1}+\left(1-\gamma_p\right)\widehat{\pi}_t^*\right)+\pi_2\left(E_t\widehat{\pi}_{t+1}-\left(1-\gamma_p\right)E_t\widehat{\pi}_{t+1}^*\right)-\pi_3\widehat{\mu}_t^p+\widehat{\varepsilon}_t^p, \tag{5.2.8}$$

with $\pi_1 = 1/\left[(1+\beta\left(\gamma_z\right)^{1-\sigma_c}\gamma_p)\right]$, where γ_p is the degree of indexation to past inflation, and $\pi_2 = \left[\beta\left(\gamma_z\right)^{1-\sigma_c}\right]/\left[(1+\beta\left(\gamma_z\right)^{1-\sigma}\gamma_p)\right]$. The coefficient $\pi_3 = \left[\left(1-\xi_p\right)\left(1-\beta\left(\gamma_z\right)^{1-\sigma}\xi_p\right)\right]/\left[\xi_p(1+\beta\left(\gamma_z\right)^{1-\sigma}\gamma_p)\left(\left(\mu_p-1\right)\varsigma_p+1\right)\right]$ captures the role of real marginal cost in driving inflation. Here ξ_p is the degree of price stickiness, μ_p the steady-state price mark-up factor and ς_p is the curvature of the Kimball goods market aggregator, due to the fact that Smets and Wouters (2007) use the Kimball aggregator instead of the common Dixit-Stiglitz aggregator, as the former allows a more reasonable degree of price and wage stickiness. The term $\widehat{\varepsilon}_t^p$ can be interpreted as a cost-push shock or as a price mark-up shock. Thus, when the inflation target switches

regimes, the dynamics of inflation include the additional terms $(1-\gamma_p)\widehat{\pi}_t^*$ and $(1-\gamma_p)E_t(\widehat{\pi}_{t+1}^*)$, that are absent in the Smets-Wouters model.[1] We assume that the inflation target oscillates between high and low target.

Using the fact that the rental rate of capital is a positive function of the degree of capital utilization and is negatively related to the capital-labor ratio, as implied by cost minimization, we obtain the following equation

$$(1+z_1)\,\widehat{r}_t + \widehat{K}_{t-1} - \widehat{w}_t - \widehat{L}_t = 0. \tag{5.2.9}$$

The wage mark-up is given by

$$\widehat{\mu}_t^w = \widehat{w}_t - \sigma_L \widehat{L}_t - \frac{1}{1-h/\gamma_z}\widehat{C}_t + \frac{h/\gamma_z}{1-h/\gamma_z}\widehat{C}_{t-1} \tag{5.2.10}$$

while, due to nominal wage stickiness, the real wage dynamics reads

$$\begin{aligned}\widehat{w}_t \;=\;& w_1\left(\widehat{w}_{t-1} + \gamma_w\widehat{\pi}_{t-1} + (1-\gamma_w)\widehat{\pi}_t^*\right)\\ &+(1-w_1)\left(E_t\widehat{w}_{t+1} + E_t\widehat{\pi}_{t+1} - (1-\gamma_w)E_t\widehat{\pi}_{t+1}^*\right)\\ &-w_2\widehat{\pi}_t - w_4\widehat{\mu}_t^w + \widehat{\varepsilon}_t^w,\end{aligned} \tag{5.2.11}$$

with $w_1 = 1/\left[1+\beta\left(\gamma_z\right)^{1-\sigma_c}\right]$, $w_2 = \left[1+\beta\left(\gamma_z\right)^{1-\sigma_c}\gamma_w\right]/\left[1+\beta\left(\gamma_z\right)^{1-\sigma_c}\right]$, where γ_w is the degree of wage indexation to lagged inflation and $w_4 = \left[(1-\xi_w)\left(1-\beta\left(\gamma_z\right)^{1-\sigma_c}\xi_w\right)\right]/\left[\xi_w\left(1+\beta\left(\gamma_z\right)^{1-\sigma_c}\right)\left((\mu_w-1)\varsigma_w+1\right)\right]$, where ξ_w is the degree of wage stickiness, ς_w is the curvature of the Kimball goods market aggregator and $\widehat{\varepsilon}_t^w$ is a wage mark-up shock. The dynamics of real wages, as in the inflation equations, includes the additional terms $(1-\gamma_w)\pi_t^*$ and $(1-\gamma_w)E_t\pi_{t+1}^*$.

[1] In a related paper, **?** consider a time-varying inflation target, where inflation target evolves as a random walk without drift. This assumption implies that $E_t\widehat{\pi}_{t+1}^* = \widehat{\pi}_t^*$. Therefore, our approach turns out to be more general.

The policy block of the model is described by the following monetary policy reaction function:

$$\begin{aligned}\widehat{R}_t &= \rho_r(s_t)\widehat{R}_{t-1} + (1-\rho_r(s_t))\left[r_\pi(s_t)(\widehat{\pi}_t - \widehat{\pi}_t^*(s_t)) + r_y(s_t)\left(\widehat{Y}_t - \widehat{Y}_t^f\right)\right] \\ &\quad + r_{\Delta y}(s_t)\left[\left(\widehat{Y}_t - \widehat{Y}_t^f\right) - \left(\widehat{Y}_{t-1} - \widehat{Y}_{t-1}^f\right)\right] + \widehat{\varepsilon}_t^r \qquad (5.2.12)\end{aligned}$$

This is a generalized Taylor rule where the monetary authorities gradually adjust the policy rate ($\widehat{R}_t$) in response of inflation deviations from its targeted level, to the spread between actual and potential output ($\widehat{Y}_t - \widehat{Y}_t^f$), to the change in the output gap $(\widehat{Y}_t - \widehat{Y}_t^f) - (\widehat{Y}_{t-1} - \widehat{Y}_{t-1}^f)$. The potential output is the output that would prevail in the absence of price and wage stickiness and of the two mark-up shocks. Parameter ρ_r captures the degree of interest rate smoothing, $\widehat{\varepsilon}_t^r$ is a monetary policy shock and s_t is a dichotomous variable that captures the monetary policy regime in place at time t. It evolves according to the transition matrix $P = [p_{ij}]$ where $p_{ij} = \Pr[s_t = i | s_{t-1} = j]$.

Finally, we close the model by specifying the law of motion for various shocks that are considered in the paper:

$$\widehat{\varepsilon}_t^a = \rho_a \widehat{\varepsilon}_{t-1}^a + \sigma^a(v_t)\eta_t^a \qquad (5.2.13)$$

$$\widehat{\varepsilon}_t^g = \rho_g \widehat{\varepsilon}_{t-1}^g + \sigma^g(v_t)\eta_t^g + \rho_{ga}\sigma^a(v_t)\eta_t^a \qquad (5.2.14)$$

$$\widehat{\varepsilon}_t^b = \rho_b \widehat{\varepsilon}_{t-1}^b + \sigma^b(v_t)\eta_t^b \qquad (5.2.15)$$

$$\widehat{\varepsilon}_t^i = \rho_i \widehat{\varepsilon}_{t-1}^i + \sigma^i(v_t)\eta_t^i \qquad (5.2.16)$$

$$\widehat{\varepsilon}_t^p = \rho_p \widehat{\varepsilon}_{t-1}^p + \sigma^p(v_t)\eta_t^p - \upsilon_p \widetilde{\eta}_{t-1}^p \qquad (5.2.17)$$

$$\widehat{\varepsilon}_t^w = \rho_w \widehat{\varepsilon}_{t-1}^w + \sigma^w(v_t)\eta_t^w - \upsilon_w \widetilde{\eta}_{t-1}^w \qquad (5.2.18)$$

$$\widehat{\varepsilon}_t^r = \rho_r \widehat{\varepsilon}_{t-1}^r + \sigma^r(v_t)\eta_t^r \qquad (5.2.19)$$

Those shocks are shocks to total factor productivity, government spending, risk-premium, investment-specific, price mark-up, wage mark-up and mone-

tary policy, respectively . The wage-markup and price-markup disturbances are assumed to follow an ARMA(1,1), as in Smets and Wouters (2007), in order for the model to reproduce some of the high-frequency fluctuations in prices and wages. Finally, v_t is an unobservable dichotomous variable capturing heteroskedasticity in the shocks. It evolves according to the transition matrix $Q = [q_{ij}]$ where $q_{ij} = \Pr[v_t = i | v_{t-1} = j], i = 1, 2; j = 1, 2$. We consider cases where the regimes s_t and v_t are synchronized or independent. When s_t and v_t are independent, we consider a new state variable s^*, which indexes both regimes s_t and v_t. Thus, s^* is four-state variable whose transition matrix is $P^* = P \otimes Q$. The consideration of synchronized regime shifts is another key difference between our paper and Bianchi (2011), who considers only a specification where regimes in policy and the variance of shocks evolve independently.

Next, we solve the model using an iterative algorithm proposed by Dufourt (2011). As mentioned in Chapter 4, the main advantage of this algorithm is it economizes on computational time.

5.3. Estimation approach

This section begins with the description of the data used for the estimation. We then derive the state space representation of the model solution in order the compute to likelihood with the methods provided in Chapter 4. Finally, we describe the prior distribution of the parameters, which is combined with the likelihood to form the posterior distribution.

5.3.1. The data

We estimate the model using seven quarterly series for the U.S. economy and Euro area. The vector of observables is $obs_t =$

$[\Delta \log Y_t, \Delta \log C_t, \Delta \log I_t, \Delta \log \frac{W_t}{P_t}, \log H_t, \pi_t, R_t]$ where Δ is the first difference operator. We first present the data for the US economy. The sample spans 1954:III to 2009:I, where 1954:III is the first quarter where the Federal funds rate is available, while the end of our sample is the first period where the Federal funds rate hits a zero bound. All data come from the Fred Database. Per capita real GDP is constructed by dividing the nominal GDP by the US working age population and the GDP deflator.

Real series for consumption and investment are obtained in the same manner, where in contrast to Smets and Wouters (2007), investment is the sum of fixed private investment and durable consumptions, and consumption is the sum of non-durables and services.[2] Real wage is defined as the hourly compensation in the Nonfarm Business sector divided by the GDP deflator. Our measure of labor is the log of hours of all persons in the Nonfarm business sector divided by the working-age population. We measure inflation as the quarterly log difference in the GDP deflator. Finally, our measure for the nominal interest rate is the quarterly effective Federal funds rate.

Except the series on population, all data for the Euro Area (EU-15) come from the synthetic Area Wide Model (AWM) dataset, first issued by Fagan et al. (2005). The dataset is publicly available from the Euro Area Business Cycle Network (EABCN). Series on the population come from the OECD. The AWM dataset lacks series on hours and no reliable measure of hours is available for the Euro Area. Following the literature, we use the employment series divided by the working-age population to express it in per capita terms. The estimation period spans 1970:I to 2009:IV, which is the period where most of the data is available.

We link the observables to the model variables by the following measurement

[2] In Smets and Wouters (2007), investment measured by the fixed private investment only, while consumption is measured as the sum of non-durables, durables and services.

equation

$$obs_t = D + Z\widetilde{X}_t \tag{5.3.1}$$

with $D = \left[\overline{\gamma}, \overline{\gamma}, \overline{\gamma}, \overline{\gamma}, \overline{l}, \pi^*, \overline{R}\right]$, where $\overline{\gamma}$ is the common quarterly trend growth rate of real GDP, consumption, investment and wages, π^* is the quarterly steady-state inflation rate, $\overline{R}$ is the steady-state nominal interest rate and $\overline{l}$ is the steady-state hours (or employment) per capita.

5.3.2. Prior distribution

Our prior follows closely the Smets and Wouters (2007) prior. We fix some parameters that are not identified or imprecisely estimated.

These parameters are: $\delta = 0.025, g_y = 0.18, \lambda_w = 1.5, \varsigma_p = 10, \varsigma_w = 10$. As noted by Smets and Wouters (2007), the first two parameters would be difficult to estimate unless investment and exogenous spending are used in the measurement equations, while the last three parameters are not identified. Several attempts to estimate the steady state log hours reveal that this parameter is imprecisely estimated. Hence, we choose $\overline{l} = 0$.

The prior distribution for the reaction coefficients to change in inflation are Gamma centered around Taylor (1993)'s values for both regimes. We also consider an alternative prior, which is consistent with the view that the Great Inflation of the 70's was the consequence of loose monetary policy. Specifically, we allow the Fed's reaction coefficients to changes in inflation relative to its target to be lower during the Great Inflation era and higher during the Great Moderation era. The reaction coefficients to both the output gap and the change in the output gap follow a Gaussian distribution, with mean 0.12 and standard deviation 0.05. The prior for the interest rate smoothing parameter follows the Beta distribution, with mean 0.5 and standard deviation 0.2. These

Table 5.1: Prior distribution of structural parameters for the synchronized and independent regimes models.

Parameters	Distr	Para (1)	Para (2) .
φ	Normal	4	1.5
σ_c	Normal	1.5	0.375
h	Beta	0.7	0.1
θ_w	Beta	0.5	0.1
σ_L	Normal	2	0.75
θ_p	Beta	0.5	0.1
γ_w	Beta	0.5	0.15
γ_p	Beta	0.5	0.15
ψ	Beta	0.5	0.15
μ_P	Normal	1.25	0.125
$r_{\pi,1}$	Gamma	1.5	0.25
$r_{\pi,2}$	Gamma	1.5	0.25
$\rho_{r,1}$	Beta	0.75	0.1
$\rho_{r,2}$	Beta	0.75	0.1
$r_{y,1}$	Gamma	0.12	0.05
$r_{y,2}$	Gamma	0.12	0.05
$r_{\Delta y,1}$	Gamma	0.12	0.05
$r_{\Delta y,2}$	Gamma	0.12	0.05
$100(1/\beta - 1)$	Gamma	0.25	0.1
$100\overline{\gamma}$	Normal	0.4	0.1
α	Beta	0.3	0.05
$100\pi^*(1)$	Gamma	0.62	0.1
$100\Delta\pi^*$	Gamma	0.5	0.1

Notes: Para(1) and Para(2) list the means and the standard deviations for Beta distribution; the shape s and and the scale ν parameters for the Inverse Gamma distribution, where $p_{IG}(\sigma|\nu, s) \propto \sigma^{-\nu-1}\exp(-\nu s^2/2\sigma^2)$.

prior values are identical across regimes and follow closely Smets and Wouters (2007).

The prior for the inflation target follows Schorfheide (2005). Instead of esti-

mating low and high targets, we estimate the low target $\pi^*(1)$ and the difference between low and high target $\Delta\pi^*$. This strategy allows us to assess whether the difference in the two targets is significant or not. The prior distributions for the low and high inflation targets are such that the annual low and high targets are 3% and 6%, respectively.

The priors for the transition probabilities are chosen to ensure that the regimes are persistent. Specifically, we assume a Beta distribution for the transition probabilities, with mean 0.9 and standard deviation 0.05.

The prior on the stochastic process follows closely Smets and Wouters (2007). The standard errors of all innovation follow an Inverse-Gamma distribution with mean 0.1 and standard deviation 2. The persistence coefficients of the shocks processes follows a Beta distribution with mean 0.5 and standard deviation 0.2. Additional details are available in Tables 5.1 and 5.2.

Table 5.2: Prior distribution of shocks processes parameters for the synchronized and independent regimes models.

Parameters	Distr	Para (1)	Para (2).
ρ_a	Beta	0.5	0.2
ρ_g	Beta	0.5	0.2
ρ_b	Beta	0.5	0.2
ρ_i	Beta	0.5	0.2
ρ_{ep}	Beta	0.5	0.2
ρ_w	Beta	0.5	0.2
ρ_r	Beta	0.5	0.2
μ_w	Beta	0.5	0.2
μ_p	Beta	0.5	0.2
ρ_{ga}	Beta	0.5	0.2
$100\sigma_{a,1}$	Inverse gamma	0.1	2
$100\sigma_{g,1}$	Inverse gamma	0.1	2
$100\sigma_{b,1}$	Inverse gamma	0.1	2
$100\sigma_{i,1}$	Inverse gamma	0.1	2
$100\sigma_{p,1}$	Inverse gamma	0.1	2
$100\sigma_{w,1}$	Inverse gamma	0.1	2
$100\sigma_{R,1}$	Inverse gamma	0.1	2
$100\sigma_{a,2}$	Inverse gamma	0.1	2
$100\sigma_{g,2}$	Inverse gamma	0.1	2
$100\sigma_{b,2}$	Inverse gamma	0.1	2
$100\sigma_{i,2}$	Inverse gamma	0.1	2
$100\sigma_{p,2}$	Inverse gamma	0.1	2
$100\sigma_{w,2}$	Inverse gamma	0.1	2
$100\sigma_{R,2}$	Inverse gamma	0.1	2
p_{11}	Beta	0.9	0.05
p_{22}	Beta	0.9	0.05
q_{11}	Beta	0.9	0.05
q_{22}	Beta	0.9	0.05

Notes: Para(1) and Para(2) list the means and the standard deviations for Beta distribution; the shape s and and the scale ν parameters for the Inverse Gamma distribution, where $p_{IG}(\sigma|\nu, s) \propto \sigma^{-\nu-1}\exp(-\nu s^2/2\sigma^2)$.

5.4. Empirical results: U.S.

This section presents our results for the US economy. We first present the model fit. Then we comment the estimates of the structural parameters and their economic implications through variance decomposition, impulse response analysis and the smoothed estimates of regime probabilities.

5.4.1. Model Fit

Table (5.3) reports measures for four specifications of our DSGE model. To assess the model fit, we compute the modified harmonic mean estimation of the marginal likelihood defined in Chapter 4. Several observations can be made from Table 5.3. First, the specifications where policy and variance switch regimes clearly dominate the other. Thus, our model fits better than the benchmark Smets and Wouters (2007) model, where policy and shock variance parameters are constant. This translates into a Bayes factor of exp(86) in favor of our model. Our model also fits better relative to Liu et al. (2011), where only regime shifts in variances and in inflation target are considered. Second, the two specifications where policy parameters switch regimes deliver a rather similar fit. Thus, this finding raises the issue about the endogeneity of regime switches in monetary policy.

5.4.2. Estimates of structural parameters and regime probabilities

In this section, we discuss the estimates of the best fit model. Posterior summary statistics for the best fit model (Posterior-Sync), such as the mean, the mode and confidence bands are obtained with the Metropolis-Hastings algorithm. We also report the posterior mode for the independent regime

Table 5.3: Model fit: (Log) Marginal Data Density

Specification	Modified harmonic mean
1.Constant parameters model	-1325.98
2.Switching variance only	-1240.10
3.Switching inflation target, policy rule and shock variances	-1222.20
4.Independent switching in inflation target, policy rule and shock variances	-1223.4

Notes: Log Marginal Data Density (MDD) computed for different specifications, with the modified harmonic mean estimator and the Laplace approximation.

specification (Posterior-Ind).

Table 5.4 and 5.5 report estimates of the best-fit model. Observing prior information and posterior summary statistics, we note that the data are informative about almost all parameters. Our estimates for the structural parameters fall within the range reported by the literature. The median estimate for φ, the steady-state elasticity of the capital adjustment cost is 5.03. This estimate lies on the error band reported by Smets and Wouters (2007)'s estimate for this parameter while it is higher than estimates by Liu et al. (2011) or Justiniano et al. (2010). Our estimate for σ_c, the intertemporal elasticity of substitution is 1.19, which is in line with the literature. The habit parameter is 0.90, much higher than estimates obtained in the literature. The labor elasticity is estimated to be 2.90. We estimate the share of fixed costs in the production function ($\mu_P - 1$) to be 0.425, somewhat lowerr than the value 0.60 obtained by Smets and Wouters (2007).

Table 5.4: Posterior distribution of structural parameters for the synchronized and independent regimes models

	Posterior-Sync				Posterior-Ind
Parameters	Mode	Mean	5%	95%	Mode
φ	5.0350	5.3152	3.7941	7.1853	5.0333
σ_c	1.1899	1.1885	1.1089	1.2723	1.1688
h	0.9092	0.9085	0.8764	0.9426	0.9158
θ_w	0.8514	0.8489	0.7984	0.8969	0.8550
σ_L	2.8995	2.9578	1.7948	4.1466	3.3043
θ_p	0.6947	0.7034	0.6294	0.7854	0.7096
γ_w	0.5018	0.4928	0.2974	0.7065	0.4679
γ_p	0.1836	0.2013	0.0762	0.3519	0.1804
ψ	0.4366	0.4860	0.3165	0.6592	0.5146
μ_P	1.4253	1.4338	1.2865	1.5786	1.4271
$r_{\pi,1}$	1.8748	1.9299	1.5958	2.3002	1.8095
$r_{\pi,2}$	1.5592	1.6320	1.2121	2.0836	1.5202
$\rho_{r,1}$	0.8855	0.8895	0.8624	0.9187	0.8819
$\rho_{r,2}$	0.7900	0.7770	0.6919	0.8572	0.7105
$r_{y,1}$	0.0463	0.0552	0.0242	0.0892	0.0342
$r_{y,2}$	0.1091	0.1176	0.0536	0.1920	0.1624
$r_{\Delta y,1}$	0.1124	0.1213	0.0843	0.1540	0.1145
$r_{\Delta y,2}$	0.1582	0.1645	0.0828	0.2461	0.1601
$100(1/\beta - 1)$	0.1187	0.1415	0.0519	0.2346	0.0873
$100\overline{\gamma}$	0.3770	0.3645	0.3081	0.4126	0.3543
α	0.2022	0.2055	0.1906	0.2203	0.2023
$100\pi^*(1)$	0.6862	0.6972	0.5125	0.8620	0.6331
$100\Delta\pi^*$	0.5189	0.5250	0.3414	0.7432	0.4646

Turning to the wage and price settings parameter, our estimates for the Calvo probabilities imply an average length of the wage contracts of six quarters and about three quarters for the price contract. These estimates are higher than the values reported in some papers (e.g (Liu et al., 2011)). The .95 error band estimates for the wage and price indexation suggest that they are precisely estimated and are very close to estimates reported by Smets and Wouters

(2007).

The shock processes suggest that productivity, spending, price markup and wage markup shocks are highly persistent. Investment shocks exhibit important persistence while preference and policy shocks exhibit low persistence. These findings are consistent with those of Smets and Wouters (2007).

Estimates of the standard deviations of shocks clearly indicate that shocks variances are effectively switching between regimes. Except for the wage markup shock, the second regime is the high volatility regime and it is slightly less persistent, as suggested by the transition probabilities ($p_{11} = 0.9461$, $p_{22} = 0.8904$) and Figure 5.1. Standard deviations for all shocks show drastic changes across regimes. In line with results reported in Justiniano and Primiceri (2008), monetary policy shock is the exogenous disturbance showing the largest degree of stochastic volatility when one compares the standard deviations for this shock in each regime: the ratio of the standard deviation for the monetary policy shock in regime 1 to the standard deviation in regime 2 is more than 350 percent. We find that the wage markup shock is relatively stable while the price markup shock has the smallest variance, with a standard deviation of this shock of 0.10 in the first regime and 0.20 in the second regime. The spending shock exhibits moderate variation while the productivity shock has the largest variance in absolute terms. Investment shock shows significant degree of variation across regime, much lower than variation for monetary policy shock. While our findings are in line with Justiniano and Primiceri (2008) about the patterns of monetary policy shocks, they are in sharp contrast with those reported in Liu et al. (2011), who find that monetary policy and technology shocks have the smallest variance.

Table 5.5: Posterior estimates of shock processes parameters.

	Posterior-Sync				Posterior-Ind
Parameters	Mode	Mean	5%	95%	Mode
ρ_a	0.9862	0.9853	0.9771	0.9940	0.9824
ρ_g	0.9848	0.9844	0.9727	0.9956	0.9837
ρ_b	0.3404	0.3563	0.2221	0.4896	0.3381
ρ_i	0.6820	0.6862	0.5733	0.7907	0.6743
ρ_{ep}	0.9718	0.9607	0.9230	0.9921	0.9469
ρ_w	0.9656	0.9600	0.9325	0.9868	0.9528
ρ_r	0.2359	0.2797	0.1600	0.4075	0.1747
μ_w	0.9362	0.9248	0.8803	0.9657	0.9173
μ_p	0.8557	0.8343	0.7222	0.9387	0.8334
ρ_{ga}	0.2793	0.2740	0.2115	0.3460	0.2784
$100\sigma_{a,1}$	0.4993	0.4968	0.4362	0.5625	0.4903
$100\sigma_{g,1}$	0.2746	0.2842	0.2421	0.3309	0.2739
$100\sigma_{b,1}$	0.1203	0.1204	0.0959	0.1440	0.1157
$100\sigma_{i,1}$	0.5367	0.5771	0.4435	0.7220	0.5211
$100\sigma_{p,1}$	0.1548	0.1581	0.1246	0.1899	0.1566
$100\sigma_{w,1}$	0.2660	0.2635	0.2290	0.2954	0.2641
$100\sigma_{R,1}$	0.0912	0.0980	0.0819	0.1144	0.0850
$100\sigma_{a,2}$	0.7444	0.7918	0.6504	0.9612	0.7531
$100\sigma_{g,2}$	0.3611	0.3696	0.2834	0.4499	0.3578
$100\sigma_{b,2}$	0.1819	0.1909	0.1367	0.2433	0.1898
$100\sigma_{i,2}$	1.4398	1.5307	1.1460	1.9232	1.4376
$100\sigma_{p,2}$	0.2474	0.2607	0.1880	0.3398	0.2592
$100\sigma_{w,2}$	0.2284	0.2225	0.1715	0.2748	0.2125
$100\sigma_{R,2}$	0.3344	0.3696	0.3029	0.4531	0.3012
p_{11}	0.9461	0.9383	0.8962	0.9762	0.9614
p_{22}	0.9110	0.8904	0.8148	0.9583	0.8991
q_{11}					0.9415
q_{22}					0.9037

Turning to the monetary policy rule, we note that the estimates of the coefficients response to changes in inflation exhibit significant variations across regimes. In the first regime (the hawkish regime), the response to changes

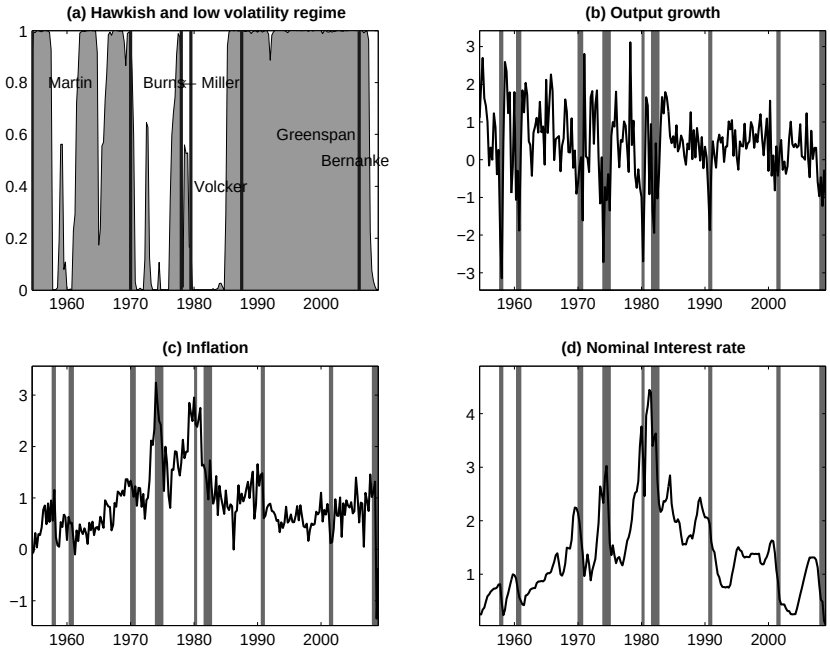

Figure 5.1: US: Posterior probabilities of the (synchronized) more hawkish-low volatility regime for the best fit model. Shaded bands in panel (b), (c) and (d) are the NBER recession and expansion dates. Vertical lines in panel (a) denote the appointment of the Chairmen.

in inflation is quite high ($r_{\pi,1} = 1.87$), and is in line with the one estimated by Smets and Wouters (2007). In the dovish regime, the same parameter is somewhat lower ($r_{\pi,2} = 1.55$). While there is significant difference between the reaction coefficients of the two regimes, our results do not support evidence reported in Clarida et al. (2000) or Lubik and Schorfheide (2004). These papers find that in the dovish regime, $r_\pi < 1$. In the dovish regime, posterior and prior modes for r_π are quite similar (1.55 and 1.5). However, one should not interpret this finding as reflecting the fact that data are uninformative about this parameter. Using an alternative prior that would imply indeterminacy in a constant coefficient model, we estimate posterior mode of $r_{\pi,2}$ to be 1.52, quite similar to the prior.[3]

[3]In a constant DSGE model, indeterminacy arises when the interest rate does not rise

Estimates of the interest rate smoothing coefficients suggest that the Fed strongly responds to the lagged interest rate in both regimes, though the response is higher in the hawkish regime (0.88) than in the second regime (0.77). Responses to output gap and changes in output gap are in line with the corresponding estimates by Smets and Wouters (2007), which suggest that the Fed responds more strongly to changes in the growth rate of output gap (0.1091) than in the output gap itself (0.0463). Finally, estimates of the inflation target imply that in the hawkish regime, the annual inflation target is about 2.7%, while it is about 5% in the dovish regime. Our estimates are somewhat lower than those of Schorfheide (2005), who estimates the inflation targets to be 2.8% and 8%, respectively.

Figure 5.1 depicts the smoothed posterior probabilities for the hawkish regime, as well as the series on output growth, inflation and interest rate.[4] The graph is consistent with the view that during much of the time in the 60's, the Fed was very hawkish against inflation, while it was dovish in the 70's. In the mid 80's until the recent financial crisis, the Fed was very hawkish against inflation. We note that these findings contrast those in Clarida et al. (2000), since their estimates suggest that pre-Volcker period was essentially a dove regime.

more than one for one in response to a change in inflation. In this case, the Taylor principle is violated and this violation can produce undesirable outcomes, such as large fluctuations in output and inflation, multiple equilibria where those variables respond to sunspot shocks, i.e. shocks that are unrelated to fundamentals of the economy but are the results of the beliefs of agents. For Markov Switching DSGE models of the kind we consider in this paper, however, there is no theoretical result for the existence of indeterminacy. Davig and Leeper (2007), Farmer et al. (2009b) and Farmer et al. (2009a) provide theoretical results for the indeterminacy/determinacy issue for forward-looking Markov-Switching DSGE models.

[4]Smoothed posterior probabilities are computed using methods provided in Kim and Nelson (1999).

5.4.3. Variance decomposition

This subsection seeks to understand what are the main driving forces of key macroeconomic variables of the model. Tables 5.6 and 5.7 show the variance decomposition for the best fit model computed at the posterior mode, where shocks are reported in column.

As one can see, under the two regimes, the main driving force of output fluctuations is the investment-specific shock. In the first regime, this shock accounts for more than 50 percent of the forecast error variance of output irrespective to the horizon considered. In the second regime, the result is even stronger, since more than 70 percent of the variance of output is explained by the this shock. In Smets and Wouters (2007), the investment-specific shock explains only about 20 percent of the variance of output. The role of investment shocks for the business cycle is documented in Justiniano et al. (2010), Justiniano and Primiceri (2008), Justiniano et al. (2011). Our finding is in line with their. Justiniano et al. (2010) criticize Smets and Wouters (2007) by showing that investment shocks are the main driving forces behind investment *and* output fluctuations when the definition of investment includes inventories and durables, or the observables include the relative price of investment.[5] As expected, investment shock is the main driving force behind investment fluctuations of. Fluctuations in hours are mostly explained by investment shock, while this shock explains an important part of fluctuations in nominal interest rate.

To some extent, spending shocks also explain an important part of the fluctuations. Price markup shocks, wage markup shocks and monetary policy

[5] We have estimated the Smets and Wouters (2007) model according to their definition of investment and consumption. Results not reported confirm the criticism of Justiniano et al. (2010). Furthermore, we have found that with the Smets-Wouters dataset, the main explanation of the Great Moderation is rather the Good luck hypothesis.

shocks play a very limited role in explaining output fluctuations.

As expected, the risk premium shock explains the largest part of the fluctuations in consumption, though this part is decreasing with the length of the forecast horizon. In the short run (up to one quarter), this shock explains a sizeable part of the nominal interest rate fluctuations. Otherwise, this shock is unimportant in explaining fluctuations for the other observables.

Price markup and wage markup shocks explain the largest part of fluctuations in inflation and real wages. Together, they account for more than 70 percent of the forecast variance of these series.

It is worth noting that the monetary policy shock plays a very limited role for series other than the interest rate. In the short run, the monetary policy shock explains a big part of nominal interest rate fluctuations. However, as the horizon lengthens, investment becomes the main driving force behind nominal interest rate fluctuations.

Summarizing, fluctuations in output, investment and hours are mostly due to investment shocks. Fluctuations in inflation are mostly explained by price markup and wage markup shocks, while fluctuations in nominal interest rate are mostly due to monetary policy and investment shocks.

Table 5.6: Variance decomposition for the best fit model (Regime 1)

Horizon	ϵ_a	ϵ_g	ϵ_b	ϵ_i	ϵ_p	ϵ_w	ϵ_r
				Output			
$1Q$	9.45	21.00	9.82	54.92	2.18	0.17	2.46
$4Q$	13.00	14.74	7.54	51.43	5.35	4.63	3.31
$8Q$	12.73	14.09	7.32	51.62	5.25	5.66	3.34
$24Q$	12.64	13.47	7.03	50.85	5.91	6.49	3.61
				Consumption			
$1Q$	0.72	0.06	96.07	0.57	0.09	0.11	2.38
$4Q$	8.08	0.52	69.57	6.52	3.85	6.76	4.70
$8Q$	9.78	0.66	64.11	6.20	4.62	10.35	4.28
$24Q$	10.27	0.76	59.51	6.11	5.18	13.47	4.69
				Investment			
$1Q$	2.99	0.16	0.78	92.43	1.54	0.34	1.76
$4Q$	6.45	0.36	0.87	84.08	3.67	2.62	1.94
$8Q$	6.13	0.34	0.83	84.39	3.53	2.70	2.07
$24Q$	6.43	0.36	0.90	82.51	4.13	3.50	2.18
				Real Wage			
$1Q$	2.23	0.00	0.21	0.49	27.82	69.20	0.05
$4Q$	5.37	0.01	0.24	2.13	28.71	63.21	0.33
$8Q$	5.80	0.02	0.25	2.11	28.48	63.01	0.33
$24Q$	5.90	0.03	0.27	2.31	28.83	62.26	0.41
				Hours			
$1Q$	44.96	13.13	5.91	33.64	0.40	0.48	1.49
$4Q$	12.36	6.21	4.23	51.23	8.77	10.59	6.61
$8Q$	9.12	5.21	3.28	41.09	13.29	21.35	6.66
$24Q$	4.98	3.62	1.70	21.92	14.10	49.86	3.82
				Inflation			
$1Q$	5.16	0.12	0.28	1.17	75.54	16.91	0.81
$4Q$	8.08	0.33	0.71	2.86	43.90	41.19	2.94
$8Q$	7.77	0.37	0.75	2.81	41.09	43.84	3.36
$24Q$	7.53	0.49	0.78	3.15	39.11	45.44	3.50
				Nominal interest rate			
$1Q$	11.27	0.72	15.88	6.75	14.53	6.04	44.82
$4Q$	15.02	1.24	6.70	25.44	14.50	24.47	12.63
$8Q$	14.12	1.31	5.79	25.32	13.25	30.08	10.13
$24Q$	13.13	1.74	5.14	23.45	11.88	36.08	8.59

Table 5.7: Variance decomposition for the best fit model (Regime 2)

Horizon	ϵ_a	ϵ_g	ϵ_b	ϵ_i	ϵ_p	ϵ_w	ϵ_r
				Output			
$1Q$	6.30	7.58	4.67	76.14	1.70	0.02	3.60
$4Q$	8.62	5.68	3.79	73.95	3.11	0.69	4.17
$8Q$	8.37	5.47	3.69	74.42	3.01	0.73	4.31
$24Q$	8.36	5.32	3.60	74.21	3.31	0.80	4.40
				Consumption			
$1Q$	1.35	0.08	90.46	0.59	0.36	0.11	7.06
$4Q$	9.22	0.46	64.54	11.41	3.18	1.95	9.23
$8Q$	10.52	0.55	62.13	11.95	3.32	2.50	9.03
$24Q$	10.86	0.61	59.99	12.47	3.62	2.89	9.55
				Investment			
$1Q$	1.52	0.06	0.44	95.37	0.81	0.07	1.72
$4Q$	3.05	0.12	0.51	92.66	1.53	0.29	1.84
$8Q$	2.90	0.12	0.49	92.80	1.49	0.28	1.93
$24Q$	3.03	0.12	0.52	92.39	1.66	0.35	1.93
				Real wage			
$1Q$	4.48	0.00	0.33	2.16	56.38	36.41	0.25
$4Q$	9.50	0.01	0.33	8.09	51.40	29.83	0.85
$8Q$	10.08	0.01	0.35	8.05	50.94	29.69	0.88
$24Q$	10.06	0.02	0.35	8.65	50.99	28.93	0.99
				Hours			
$1Q$	29.24	5.86	3.50	58.08	0.45	0.16	2.70
$4Q$	6.86	2.47	2.06	74.75	5.29	2.17	6.41
$8Q$	6.03	2.49	1.90	70.67	7.99	4.34	6.59
$24Q$	5.47	3.06	1.65	62.21	10.37	11.53	5.71
				Inflation			
$1Q$	5.22	0.07	0.20	2.49	85.16	5.08	1.78
$4Q$	9.75	0.21	0.62	6.33	62.69	15.68	4.72
$8Q$	9.64	0.25	0.68	6.29	60.38	17.83	4.92
$24Q$	9.25	0.39	0.74	7.33	56.72	21.01	4.56
				Nominal interest rate			
$1Q$	6.25	0.24	8.49	8.51	11.79	1.07	63.64
$4Q$	10.59	0.53	4.87	43.07	12.76	5.01	23.18
$8Q$	10.45	0.58	4.54	45.16	12.20	6.17	20.89
$24Q$	10.46	0.79	4.36	45.10	12.02	7.91	19.36

5.4.4. Impulse responses

In this section, we analyse the model's transmission mechanisms through the impulse responses. Following Smets and Wouters (2007), we report impulse response for output, hours, inflation and interest rate.

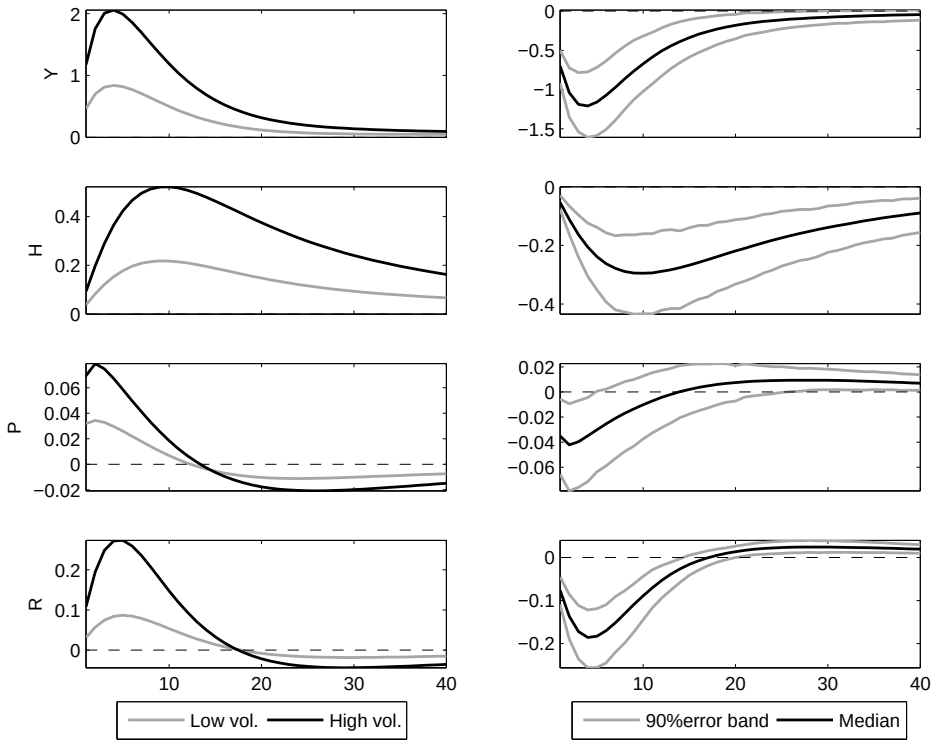

Figure 5.2: Impulse responses (median) to a investment shock. Left panels are impulse responses for the two regimes; right panels are the difference between impulse responses associated with their 90 percent credible intervals.

Table 5.6 and Table 5.7 clearly suggest that the main driving forces of fluctuations in output, inflation and interest rate are the investment shocks), price and markup shocks and monetary policy shocks, respectively. Thus, we report impulse responses to these shocks.

Figure 5.2 depicts the posterior median impulse responses to a one standard deviation investment shock. On impact, output, hours, inflation and interest

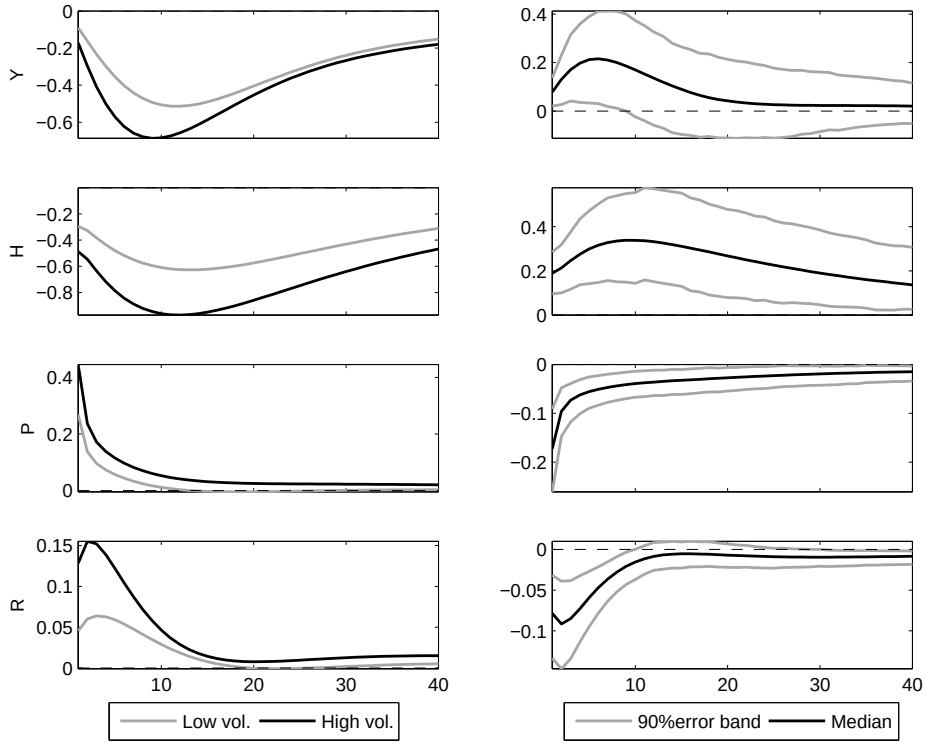

Figure 5.3: Impulse responses (median) to a price markup shock. Left panels are impulse responses for the two regimes; right panels are the difference between impulse responses associated with their 90 percent credible intervals .

rate increase. The difference in magnitude between regimes is significant, as suggested by the error bands around the difference of impulse responses for the two regimes (right panels). The transmission mechanism works as follow: investment rises while consumption declines (not reported). The increase in investment rises hours, which leads to a rise in output since firms are able to produce more. The rise in demand raises inflation to rise, so that the nominal interest rate increases (through the policy rule).

Figure 5.3 depicts the responses to a one standard deviation shock to the price shock. A positive price markup shock leads optimizing firms to increase their price. Consequently, inflation rises and output falls. The rise in inflation

implies that the nominal interest rate increases (through the policy rule) while aggregate output and hours fall.

Finally, following a monetary policy shock, the nominal interest rises while output, hours and inflation fall. This is depicted in Figure (5.4). The confidence bands show that there are important difference in the transmission mechanism across regimes.

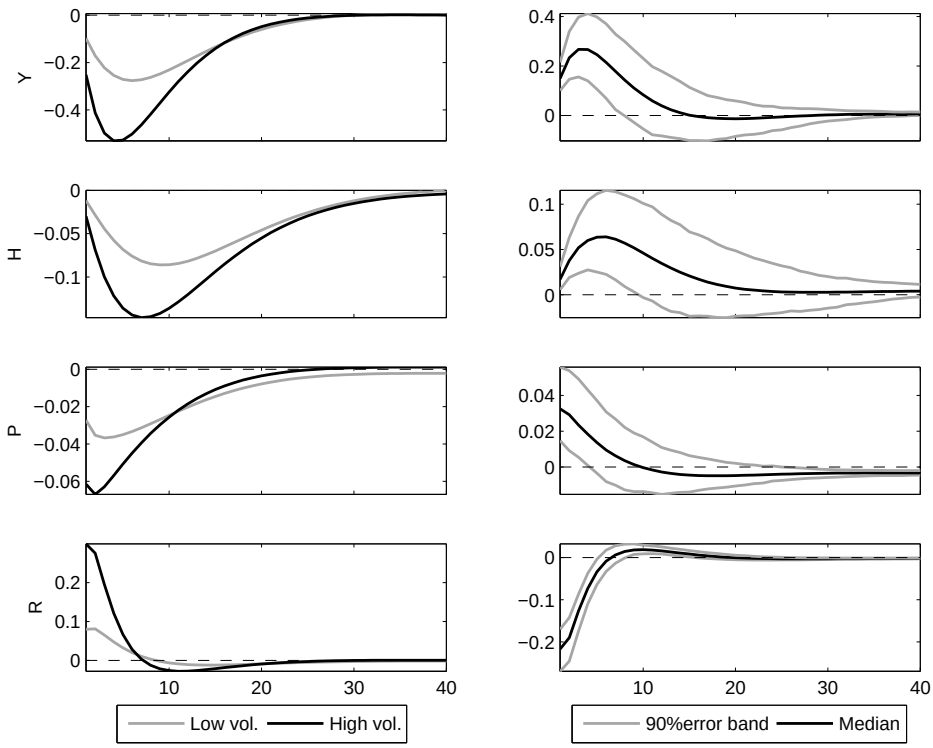

Figure 5.4: Impulse responses (median) to a monetary policy shock. Left panels are impulse responses for the two regimes; right panels are the difference between impulse responses associated with their 90 percent credible intervals.

5.5. Empirical results: Euro Area

This section presents the results for Euro Area. As for the US, we first assess the model fit, then the posterior summaries are analysed. Through the variance decomposition, we document what shocks are most important in explaining European business cycles. We then compute and interpret the impulse responses. Finally, we describe the regimes that the data lead to.

5.5.1. Model fit

While Euro Area and U.S. economies present important differences, we have found that such differences do not matter for the prior elicitation. That is, using the same prior as in Smets and Wouters (2003) and the prior used in Section 5.3.2 leads essentially to the same results. Thus, we comment our results with respect to prior information in section 5.3.2.

The model ranking is the same as for the US, as suggested in Table 5.8. The best fit model is the one where both policy and shocks variance switch regime. Note that our results stand in contrast with those of Rubio-Ramirez et al. (2005). According to their results, lower macroeconomic volatility observed in the Euro Area in the early 90's and after is due to smaller shocks to interest rate and inflation. They find little (if any) evidence about change in monetary policy. They obtain such a conclusion using Markov-Switching Vector autoregressive models. This might explain why our results are different from their.

5.5.2. Structural parameters and regime probabilities

Table 5.9 and 5.10 report estimates of the best-fit model. We focus on parameters characterizing price and wage stickiness, shocks processes and policy.

Table 5.8: Model fit: (Log) Marginal Data Density: Euro Area

Specification	Modified harmonic mean
1.Constant parameters model	-733.2301
2.Switching variance only	-694.0481
3.Switching inflation target, policy rule and shock variances	-672.4167
4.Independent switching in inflation target, policy rule and shock variances	-673.6352

Notes: Log Marginal Data Density (MDD) computed for different specifications, with the modified harmonic mean estimator and the Laplace approximation

The estimated price indexation γ_p is estimated to be 0.09. Such estimation comes with great uncertainty, as suggested by the large error band around this parameter. This implies that the backward looking component in the New Phillips curve is close to zero. For comparison, Smets and Wouters (2003)'s estimates is $\gamma_p = 0.46$.

We estimate the wage indexation closely to Smets and Wouters (2003). Estimates of the degree of price and wage stickiness suggest that price and wage contracts last three-and-a-half and four-and-a-half quarters, respectively. Such estimates are consistent with findings based on microeconomic studies.

Our estimates of policy parameters suggest that in the first regime, Euro Area monetary policy tends to be more aggressive against inflation than in the second regime. In particular, the estimates of the posterior mode for the second regime would imply indeterminacy in a constant parameter DSGE model since we estimate $r_{\pi,2} = 0.98$. However, such estimates should be interpreted with a bit of caution. Prior of 1999, the monetary policy in the Euro area was not unique. It could be the case that while some single economies would have had a monetary policy that leads to indeterminacy, the monetary policy followed by others may have implied determinacy. Both regimes exhibit considerable interest rate smoothing ($\rho_{r,1} = 0.90, \rho_{r,2} = 0.83$). Reaction to output gap is

slightly lower in the first regime ($r_{y,1} = 0.8$, $r_{y,2} = 0.9$). Estimates of the inflation target imply that in the first regime, the annual inflation target is 2.5% while it is 4.5% in the second regime.

Focusing on parameters characterizing the shock process, two remarks are in order. First, productivity, government spending, wage markup and price markup shocks are very persistent. Their autoregressive parameters are close to one. Smets and Wouters (2003) find a similar results. Second, estimates of standard deviations suggest that regime one is a regime of lower macroeconomic volatility. In fact, unlike the US case, the standard deviations of all shocks in the first regime are lower than their counterpart in the second regime.

Table 5.9: Posterior of Structural Parameters

	Posterior-Sync				Posterior-Ind
Parameters	Mode	Mean	5%	95%	Mode
φ	6.4280	6.5644	4.8020	8.2231	7.0427
σ_c	1.4128	1.4282	1.1538	1.7623	1.6704
h	0.8180	0.7932	0.7101	0.8696	0.4967
θ_w	0.7332	0.7275	0.6716	0.7842	0.6712
σ_L	2.6401	2.8212	1.8572	4.0339	2.5227
θ_p	0.7876	0.7757	0.7176	0.8289	0.7564
γ_w	0.2152	0.2515	0.0850	0.4350	0.1864
γ_p	0.0900	0.1108	0.0288	0.2091	0.0915
ψ	0.5013	0.5899	0.3748	0.7980	0.7702
μ_P	1.7921	1.7633	1.6163	1.9483	1.7246
$r_{\pi,1}$	1.3901	1.3255	0.9948	1.7128	1.6074
$r_{\pi,2}$	0.9878	1.0729	0.7980	1.3226	1.2201
$\rho_{r,1}$	0.9097	0.8907	0.8288	0.9461	0.9011
$\rho_{r,2}$	0.8320	0.8430	0.7744	0.9060	0.5773
$r_{y,1}$	0.0801	0.0819	0.0285	0.1412	0.0778
$r_{y,2}$	0.0923	0.1353	0.0585	0.2131	0.1215
$r_{\Delta y,1}$	0.1713	0.2014	0.1261	0.3029	0.2709
$r_{\Delta y,2}$	0.1205	0.1327	0.0657	0.2073	0.0899
$100(1/\beta - 1)$	0.1555	0.1563	0.0562	0.2584	0.1058
$100\overline{\gamma}$	0.2964	0.3029	0.2420	0.3601	0.3331
α	0.1859	0.1802	0.1474	0.2093	0.1877
$100\pi^*(1)$	0.6205	0.6190	0.4388	0.8081	0.6228
$100\Delta\pi^*$	0.4884	0.4941	0.3200	0.6624	0.4260

Table 5.10: Posterior of Structural Parameters

	Posterior-Sync				Posterior-Ind
Parameters	Mode	Mean	5%	95%	Mode
ρ_a	0.9974	0.9960	0.9928	0.9988	0.9932
ρ_g	0.9988	0.9974	0.9936	0.9999	0.9982
ρ_b	0.1372	0.2511	0.0640	0.4970	0.7976
ρ_i	0.5490	0.5063	0.3665	0.6435	0.4785
ρ_{ep}	0.7574	0.7461	0.5780	0.9064	0.7889
ρ_w	0.9798	0.9751	0.9645	0.9855	0.9735
ρ_r	0.4282	0.4298	0.3006	0.5743	0.3596
μ_w	0.8575	0.8332	0.7546	0.8993	0.7963
μ_p	0.6314	0.5900	0.3497	0.8070	0.6750
ρ_{ga}	0.4723	0.4507	0.3198	0.5982	0.4454
$100\sigma_{a,1}$	0.2451	0.2496	0.2109	0.2904	0.2586
$100\sigma_{g,1}$	0.2420	0.2431	0.2115	0.2802	0.2361
$100\sigma_{b,1}$	0.1428	0.1393	0.1007	0.1767	0.0540
$100\sigma_{i,1}$	0.3020	0.3686	0.2678	0.4617	0.4082
$100\sigma_{p,1}$	0.1154	0.1148	0.0839	0.1452	0.1260
$100\sigma_{w,1}$	0.0752	0.0852	0.0612	0.1124	0.0938
$100\sigma_{R,1}$	0.0934	0.1079	0.0835	0.1355	0.1106
$100\sigma_{a,2}$	0.3847	0.4226	0.3381	0.5192	0.4160
$100\sigma_{g,2}$	0.3560	0.3885	0.3096	0.4764	0.3934
$100\sigma_{b,2}$	0.2296	0.2188	0.1468	0.2801	0.0838
$100\sigma_{i,2}$	0.5712	0.6265	0.4799	0.7840	0.6280
$100\sigma_{p,2}$	0.2352	0.2438	0.1806	0.3101	0.2692
$100\sigma_{w,2}$	0.1927	0.2093	0.1547	0.2687	0.2296
$100\sigma_{R,2}$	0.2326	0.2563	0.2104	0.3057	0.2436
p_{11}	0.8993	0.9016	0.8453	0.9555	0.9411
p_{22}	0.9382	0.9141	0.8573	0.9672	0.9271
q_{11}					0.9462
q_{22}					0.9170

Figure 5.5 depicts the smoothed posterior estimates of regimes as well as the series on output growth, inflation and interest rate. The hawkish regime exhibits much of its persistence after 1993, which is the second important date

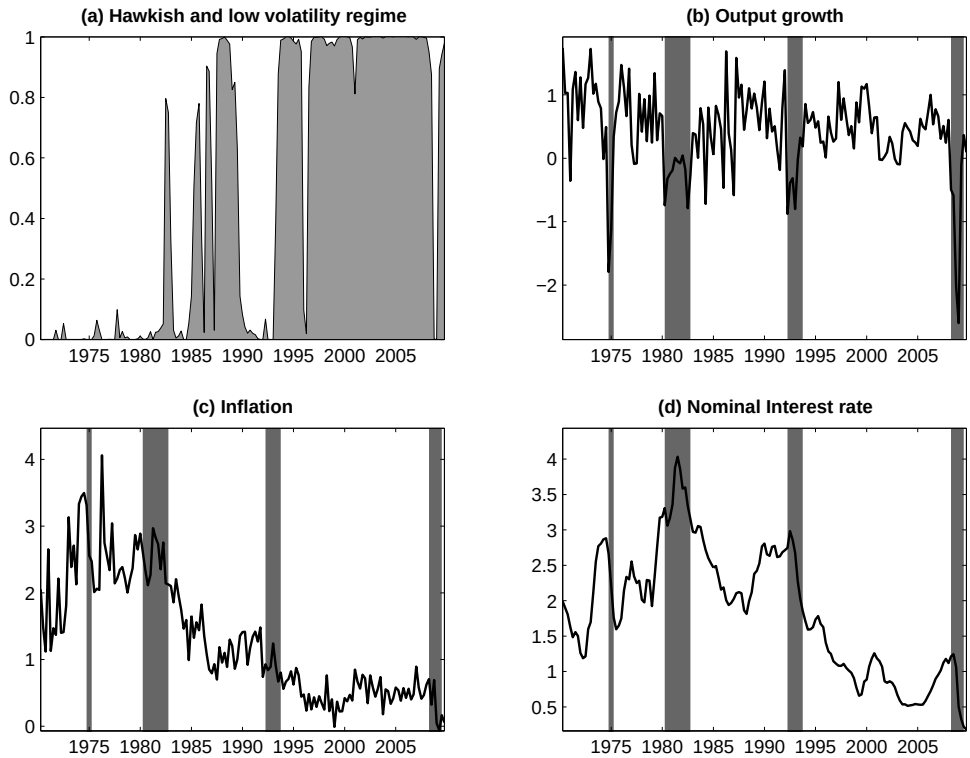

Figure 5.5: EU: Posterior probabilities of the (synchronized) hawkish-low volatility regime for the best fit model.

in the creation of the European Monetary Union (EMU). The dovish regime mainly occurs in the 70's, the early 80's and 90's. The high volatility regime in the early 90's can be explained by the German reunification as an exogenous shock. The creation of the German economic monetary union (GEMU) and the massive injection of money that followed are important factors behind the increase in inflation.

5.5.3. Variance decomposition

Table 5.11 and 5.12 report the contribution of the structural shocks to the forecast error variance for selected endogenous variables. Fluctuations in output are mainly driven by spending, preference, investment and monetary policy

shocks. Together, these shocks account for more than 70 percent of the output variation. Smets and Wouters (2003) find similar results for the role of monetary policy in contributing to output variations. They argue that such a role of monetary policy shocks is due to the disinflation period in the early 1980s and the ERM crisis. Unlike the US, the investment shock does not play a preponderant role. It only mainly contributes to investment variations. As expected, variations in consumption are mainly due to preference shocks. Note that monetary policy contributes importantly to consumption fluctuations. Price markup and wage markup shocks are the main drivers for variation in real wages, but monetary policy shocks still play an important role. Turning to inflation, we note that its variations are mainly due to price and wage markup shocks. However, as horizon lengthens, wage markup shocks become the main contributor. Finally, fluctuations in nominal interest rate are mainly determined by the wage markup shock and the monetary policy shock. In the very short term, the monetary policy shock mainly determines interest rate variations.

Table 5.11: Variance decomposition for the best fit model (Regime 1)

Horizon	ϵ_a	ϵ_g	ϵ_b	ϵ_i	ϵ_p	ϵ_w	ϵ_r
				Output			
$1Q$	12.38	28.87	23.16	21.60	0.35	1.84	11.80
$4Q$	13.02	23.54	20.34	20.47	1.01	2.40	19.23
$8Q$	12.91	23.05	20.02	20.55	1.01	3.04	19.41
$24Q$	12.49	22.22	19.32	20.14	1.19	3.85	20.80
				Consumption			
$1Q$	2.47	0.79	83.49	0.41	0.05	0.54	12.25
$4Q$	6.66	1.97	67.97	1.55	0.69	1.34	19.83
$8Q$	6.88	2.02	66.92	1.55	0.72	2.16	19.76
$24Q$	6.78	1.98	64.80	1.63	0.81	3.02	20.97
				Investment			
$1Q$	0.72	0.01	0.87	87.27	0.37	1.67	9.09
$4Q$	1.62	0.03	0.90	81.27	1.00	2.22	12.97
$8Q$	1.59	0.03	0.92	80.94	0.99	2.46	13.07
$24Q$	1.55	0.03	0.92	79.32	1.18	2.89	14.09
				Real Wage			
$1Q$	0.48	0.13	2.94	1.75	30.42	56.67	7.60
$4Q$	2.48	0.18	2.54	3.06	22.75	50.86	18.13
$8Q$	3.00	0.18	2.58	3.09	22.58	50.55	18.01
$24Q$	3.02	0.19	2.44	3.20	21.40	50.19	19.56
				Employment			
$1Q$	35.37	22.60	17.25	16.27	0.03	0.37	8.11
$4Q$	14.73	14.31	9.34	14.43	1.73	1.02	44.44
$8Q$	12.23	13.98	7.57	11.66	2.63	2.35	49.58
$24Q$	8.71	15.92	4.36	7.35	2.28	26.02	35.36
				Inflation			
$1Q$	0.94	0.17	0.21	0.03	72.29	22.48	3.88
$4Q$	1.67	0.45	0.47	0.05	29.58	57.63	10.15
$8Q$	1.64	0.49	0.46	0.07	24.30	62.49	10.53
$24Q$	1.98	0.70	0.40	0.43	17.76	69.72	9.02
				Nominal interest rate			
$1Q$	5.57	1.00	30.19	2.71	4.26	6.66	49.61
$4Q$	6.37	1.62	13.31	6.43	3.44	44.31	24.51
$8Q$	5.38	1.53	10.29	5.13	2.67	56.19	18.80
$24Q$	4.24	1.57	6.27	4.42	1.65	70.24	11.61

Table 5.12: Variance decomposition for the best fit model (Regime 2)

Horizon	ϵ_a	ϵ_g	ϵ_b	ϵ_i	ϵ_p	ϵ_w	ϵ_r
				Output			
$1Q$	11.12	24.07	19.61	21.04	0.67	3.51	19.97
$4Q$	11.35	18.67	16.69	19.24	1.51	4.30	28.24
$8Q$	11.04	17.90	16.11	19.08	1.53	5.03	29.30
$24Q$	10.63	17.16	15.47	18.64	1.81	5.52	30.77
				Consumption			
$1Q$	2.14	0.62	73.03	0.55	0.10	1.07	22.49
$4Q$	6.06	1.56	56.52	1.55	0.96	2.57	30.77
$8Q$	6.19	1.57	54.62	1.51	0.97	3.68	31.45
$24Q$	6.04	1.53	52.41	1.57	1.13	4.36	32.95
				Investment			
$1Q$	0.59	0.01	0.55	80.80	0.66	3.09	14.30
$4Q$	1.47	0.02	0.58	74.09	1.48	3.90	18.47
$8Q$	1.42	0.02	0.57	73.07	1.52	4.09	19.30
$24Q$	1.41	0.02	0.60	71.59	1.80	4.29	20.29
				Real Wages			
$1Q$	0.23	0.07	1.45	1.02	26.37	63.83	7.03
$4Q$	1.46	0.10	1.34	1.82	20.79	59.83	14.65
$8Q$	1.79	0.11	1.39	1.89	20.75	59.09	14.98
$24Q$	1.82	0.11	1.33	2.04	20.28	57.65	16.77
				Employment			
$1Q$	34.21	19.36	15.03	16.35	0.07	0.72	14.25
$4Q$	12.39	10.72	6.75	12.76	2.33	1.79	53.26
$8Q$	10.74	10.98	5.61	10.76	3.10	3.99	54.81
$24Q$	8.65	13.91	3.45	8.03	2.31	27.30	36.36
				Inflation			
$1Q$	0.58	0.10	0.12	0.02	67.53	28.18	3.48
$4Q$	0.89	0.24	0.18	0.03	25.61	66.37	6.69
$8Q$	0.85	0.25	0.16	0.07	20.77	71.76	6.14
$24Q$	1.01	0.35	0.11	0.33	13.76	80.07	4.36
				Nominal interest rate			
$1Q$	2.23	0.37	9.23	1.04	6.46	7.93	72.74
$4Q$	2.98	0.72	4.43	2.71	4.40	52.15	32.61
$8Q$	2.51	0.69	3.32	2.12	3.28	64.06	24.01
$24Q$	2.00	0.73	1.91	1.92	1.94	77.65	13.87

5.6. Endogenous monetary policy

In this section, we use our estimation results to shed new light on the following question: Are changes in monetary policy regime endogenous? More precisely, to what extent do changes in the conduct of monetary policy reflect the current state of the economy?

We provide insights to this question by jointly analyzing the estimates from the synchronized and independent regimes specifications. The fact that both versions fit the data equally well is interesting, given that the synchronized regimes specification is actually nested in the independent regimes specification.

To understand this point, assume for instance that the true data generating process features independent regime changes, and we estimate both versions of the model. Then, the specification with independent regime switches should clearly dominate. Assume now that the true Data Generating Process features *synchronized* regime changes. Because the synchronized version is nested in the independent regimes specification, both version should deliver roughly similar fits (at least, asymptotically).

In practice, things are complicated by the fact that there are additional parameters to estimate in the independent regime specification,[6] and that the number of observations is limited. But the finding of a similar fit for both versions of the model clearly points toward investigating the potential endogeneity of monetary policy regime changes.

In order to address this issue while taking into account the data limitation problem, we will consider a stricter diagnosis test for concluding that monetary

[6]For example, the specification with independent regime switches requires the estimation of four transition probabilities, while the version with synchronized regimes only requires 2.

policy is indeed endogenous. Specifically, we require that the following three criteria be *roughly* satisfied: (i) the "independent regimes" and "synchronized regimes" specifications produce similar fits ; (ii) in the "independent regimes" specification, changes in monetary policy occur (roughly) simultaneously with changes in the variance of shocks ; (iii) in the "independent regimes" specification, two of the four configurations possible *ex-ante* almost never occur. In particular, the "hawkish monetary policy – high volatility" and the "dovish monetary policy – low volatility" regimes should almost never occur.

5.6.1. Diagnosis test: U.S. economy

As mentioned above, the first criterion required for the conclusion that the US monetary policy has been endogenous is satisfied (see Table 5.3). However, the other two criteria remain far from being fulfilled.

Consider for example the smoothed regime probabilities for the US economy, depicted in Figure 5.6. The figure indicates that two switches in monetary policy occurred. The first switch, from the hawkish to the dovish regime, occurred roughly in the year 1970. The second switch, from the dovish to the hawkish regime, occurred in the early 80s. This timing does not concord well with the timing of switches in volatility regimes.

Consider now the smoothed probabilities associated with being in any of the four conceivable configurations in the independent regimes specification (see 5.7). Clearly, the figure indicates that the configuration of a high volatility regime associated with a hawkish monetary policy occurred quite frequently, especially during the late 50s – early 60s, and during the mid-80s. It also suggests that monetary policy has been dovish while volatility was low in the late 70s. Thus our criteria (ii) and (iii) are clearly not satisfied, and we cannot conclude that the US monetary policy has been endogenous.

This conclusion tends to be confirmed by another observation. Figure 5.6 sug-

gests, quite strikingly, that changes in monetary policy regimes are strongly related to the appointments of a new Chairman at the Federal Reserve. In particular, while the hawkish regime was prevailing in the Martin era, monetary policy apparently switched to the dovish regime with the appointment of Burns, and remained dovish under Miller. Then, according to the figure, a return to the hawkish regime occurred shortly after the appointment of Volcker as the Fed's Chairman. This hawkish regime continued to prevail during the Greenspan and Bernanke chairmanships. Thus, the personality of the Chairman in office appears to be a good indicator of the type of monetary policy conducted. This tends to confirm that monetary policy did not systematically change in response to changes in the economic situation.

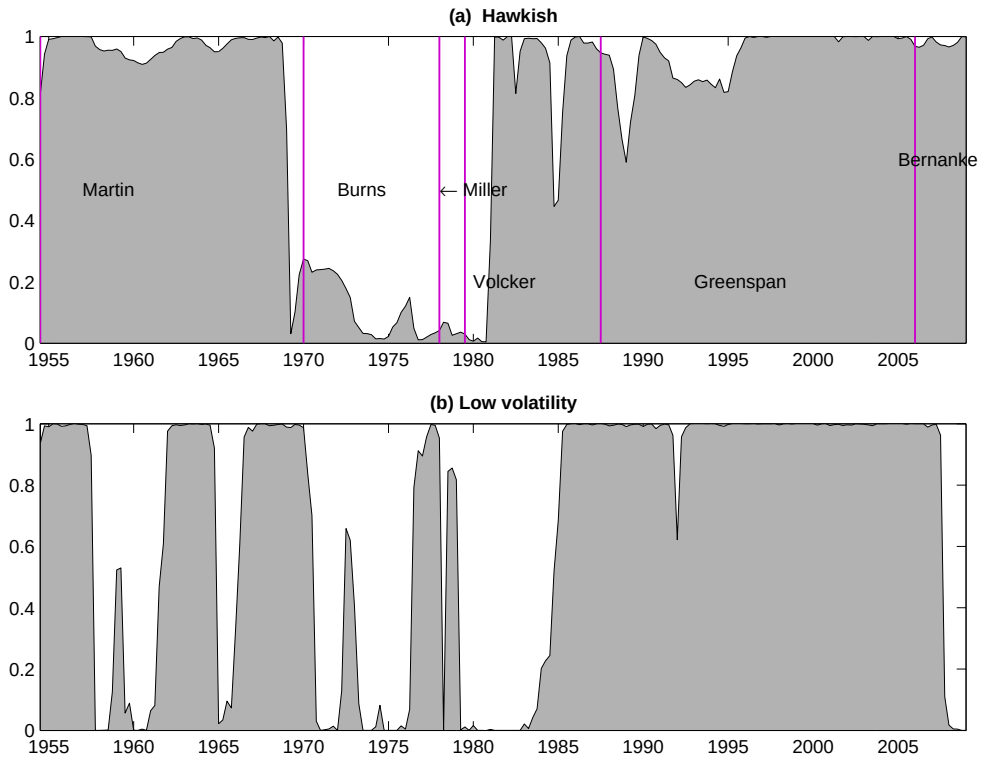

Figure 5.6: Posterior probabilities for the Hawkish regime, computed at the posterior mode estimates of the independent regime switching model. Vertical bars mark the chairmen appointment.

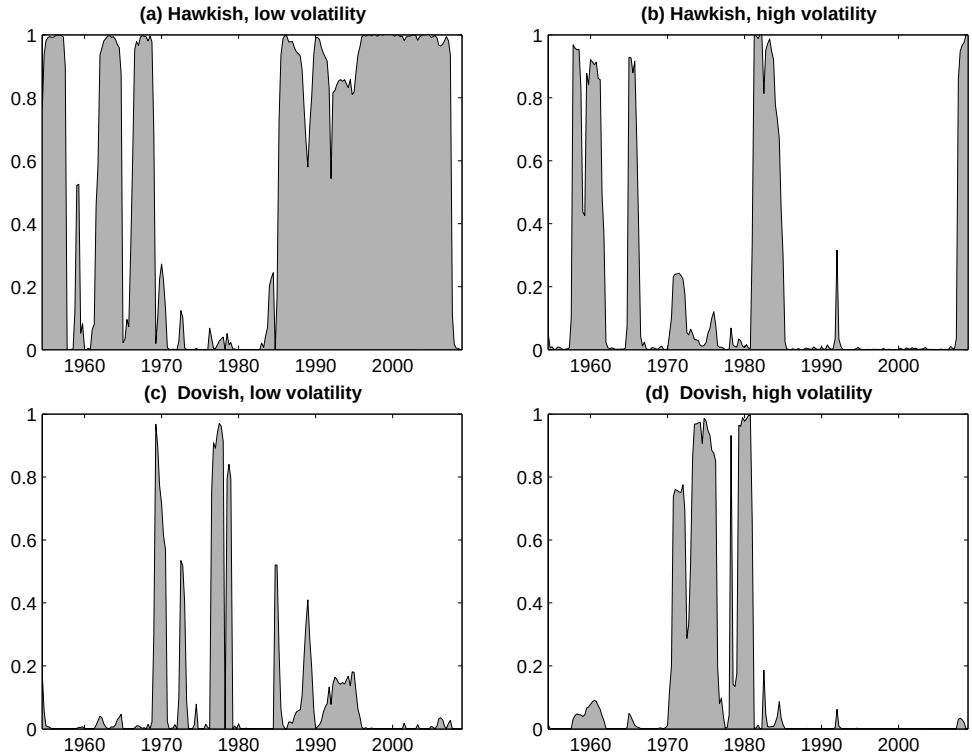

Figure 5.7: Posterior probabilities computed at the posterior mode estimates of the independent regime switching model switching model.

5.6.2. Diagnosis test: Euro Area

In contrast to the US, our results suggest that for the Euro Area monetary policy has been endogenous. Consider first the smoothed regime probabilities depicted in Figure 5.8. Clearly, periods during which the European monetary policy has been hawkish tend to correspond with periods of low volatility, and vice versa. Similarly, looking at the smoothed probability associated with being in any of the 4 conceivable situations ex-ante (see Figure 5.9), one clearly sees that the "hawkish monetary policy – high volatility" and the "dovish monetary policy – low volatility" regimes almost never occurred. The only exception is the period of the early 80s, where monetary policy remained hawkish while the Euro Area was experiencing a severe recession.

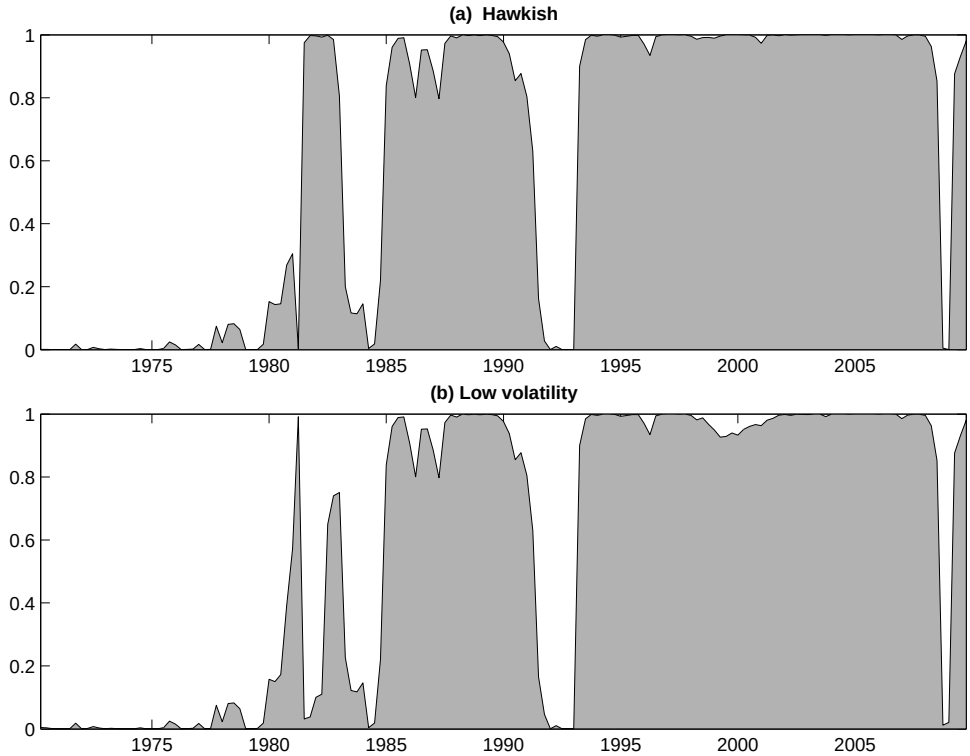

Figure 5.8: Posterior probabilities for the Hawkish regime, computed at the posterior mode estimates of the independent regime switching model. Vertical bars mark the three stages dates in the creation of the European Economic Union.

5.7. Concluding remarks

In this chapter, we have estimated the Smets and Wouters (2007) model in the presence of regime switches in both monetary policy and the shocks variance parameters. We have used the estimated model to shed new lights on the sources of the Great Moderation and the on issue related to the endogeneity of monetary policy. We find strong evidence in favor of regime switches, both in policy parameters and shock variances. Imposing synchronized regime switches in our model does not deteriorate its fit, as this version fits equally well than the version where regime switches are independent. Our last impor-

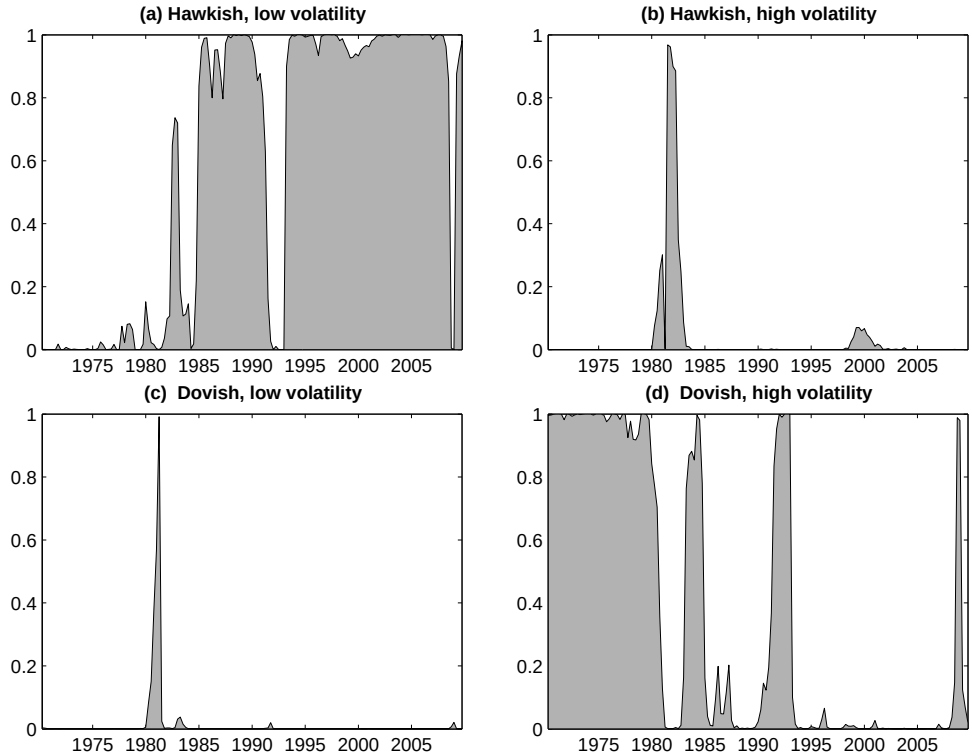

Figure 5.9: Posterior probabilities computed at the posterior mode estimates of the independent regime switching model switching model.

tant finding is the strong evidence that changes in European monetary policy had been endogenous, while for the US economy, changes in monetary policy are closely related to the personality of the Chairman in place.

The current version of the chapter lacks counterfactual experiments that would more deeply document the sources of the Great Moderation. Also, it will be useful to contrast our results with facts based on the European economy. We leave these two considerations for future research.

Chapter 6

Conclusion

Contents

6.1. Conclusion

Nowadays, dynamic stochastic general equilibrium models provide a new modeling framework for policy analysis, both in academic and policy-making spheres. There are several theoretical arguments that could explain the success of these models. In this thesis, we have focused instead on empirical rather than theoretical arguments. To do so, we have used recent econometrics tools, which have allowed us to provide a systematic confrontation of these models to the data.

6.1.1. Summary

Our econometric work has tried to evaluate DSGE models with respect to the data in three dimensions. Such dimensions are inflation forecasting performance, inflation persistence and the Great Moderation.

To study the predictions of DSGE models with respect to forecasting inflation, we have estimated the new-Keynesian Phillips curve, a key equation shared by every DSGE models featuring Keynesian ingredients. The key result from that estimation is that once direct observations from survey data on inflation expectations are used, the new-Keynesian Phillips curve helps to forecast inflation quite well.

With respect to the inflation persistence, we provide new results in the literature dealing with the fit of the new-Keynesian Phillips curve in matching inflation dynamics. We show that including a time-varying inflation target of the central bank in the new-Keynesian Phillips curve does not necessary lead to a time-varying new-Keynesian Phillips curve, unlike in studies such as Cogley and Sbordone (2008). However, like the authors, we do find that the introduction of a time-varying inflation target is sufficient to account for inflation persistence, instead of *ad hoc* backward-looking component in the new-Keynesian Phillips curve.

The last dimension of the data to which we have confronted the model is the Great Moderation, the idea that economic data before the mid 1980 are more stable than before. The model proves to be useful in replicating the Great Moderation both for the U.S. and for the Euro Area economies. Furthermore, the estimated model provides an interesting source of the Great Moderation: both the variance of the shocks and improved monetary policies are plausible explanations for the decline in macroeconomic data observed since the mid of 80's until the recent financial crisis.

Evidence of changes in monetary policy regime naturally raises the question of whether such changes are endogenous or exogenous. This question makes sense because the central bank is supposed *a priori* to react to the current state of the economy by adjusting its policy instrument, i.e. the nominal interest rate in the model. We have found that for the Euro Area economy, changes

in monetary policy regimes are endogenous precisely because the central bank sets its policy instrument with respect to the volatility of the economy. In particular, periods of low volatility coincide with those where the European monetary policy is aggressive towards inflation. On the contrary, changes in the U.S. monetary policy cannot be said to be endogenous, but appear to be closely connected to the personality of the Chairman in place, reflecting the "conservative central banker" view of the conduct of monetary policy (Rogoff, 1985).

6.1.2. Extensions

Despite the success of the DSGE models in matching interesting dimensions of the data, a number of recent papers have pointed out several limitations of the DSGE approach. Here, we focus on two of them.

The first one is the inability of DSGE models to take into account high and persistent unemployment found in the data. Such a limitation could potentially reduced the ability of the new-Keynesian to take into account some monetary phenomena. Indeed, Galí et al. (2011) show that the Smets and Wouters (2007) model can be reformulated to incorporate unemployment. Their results suggest that the model is able to reproduce observed unemployment fluctuations when it is estimated with data on unemployment, instead of data on hours worked, as we have done in the thesis. Thus, it remains interesting to see whether our main conclusions are robust with respect to this reformulations.

The second limitation has to do with the recent financial crisis. To model the monetary policy, we have considered a "conventional monetary policy", i.e. a policy where the Federal Reserve manipulates the Federal Funds rate in order to affect markets interest rates. However, the recent financial crisis started in August 2007 dramatically changes the environment, as it has led the Fed

to inject credit into private markets. In this sense, such a policy is termed "unconventional". Most of DSGE models considered in the literature specify a monetary policy in which the Federal Reserve acts conventionally. Thus, they are not truly useful to make predictions in crisis time where the central bank acts unconventionally.

There is a burgeoning literature trying to introduce the financial sector into DSGE models. The most representative paper of that literature is a paper provided by Gertler and Karadi (2011). The authors develop a quantitative monetary DSGE model that allows for a role for the financial sector through the financial intermediaries facing endogenous balance sheet constraints. While their model is not intended to model the sub-prime crisis, it tries to account for some key elements relevant to analyzing the unconventional monetary policy conducted by the Fed since August 2007 and particularly after the Lehman Brothers collapse. Hence, we see the Gertler and Karadi (2011) model as a good starting point to extend our work. The methodology developed in Chapters 4 and 5 could be used to estimate such a model.

Appendix A

A.1. Appendix to Chapter 3

A.1.1. The New Keynesian Phillips Curve

The first-order condition associated to the program of the firms is given by

$$\frac{\widetilde{p}_t^i}{p_t} \underset{t}{E} \sum_{k=0}^{\infty} \left(\beta\xi_p\right)^k \frac{p_t}{p_{t+k}} X_{t,k}^p \lambda_{t+k} y_{t+k}^i = \mu_p \underset{t}{E} \sum_{k=0}^{\infty} \left(\beta\xi_p\right)^k \lambda_{t+k} mc_{t+k} y_{t+k}^i.$$

Inserting the definition of y_{t+k}^i in the previous expression leads to

$$\frac{\widetilde{p}_t^i}{p_t} \underset{t}{E} \sum_{k=0}^{\infty} \left(\beta\xi_p\right)^k \left(X_{t,k}^p\right)^{1-\theta} \left(\frac{p_t}{p_{t+k}}\right)^{(1-\theta)} \lambda_{t+k} y_{t+k} = \mu_p \underset{t}{E} \sum_{k=0}^{\infty} \left(\beta\xi_p\right)^k \left(X_{t,k}^p\right)^{-\theta} \left(\frac{p_{t+k}}{p_t}\right)^{\theta} \lambda_{t+k} mc_{t+k} y_{t+k},$$

or

$$\begin{aligned} \frac{\widetilde{p}_t^i}{p_t} &= \mu_p \frac{\underset{t}{E} \sum_{k=0}^{\infty} \left(\beta\xi_p\right)^k \left(X_{t,k}^p\right)^{-\theta} \left(\frac{p_{t+k}}{p_t}\right)^{\theta} \lambda_{t+k} mc_{t+k} y_{t+k}}{\underset{t}{E} \sum_{k=0}^{\infty} \left(\beta\xi_p\right)^k \left(X_{t,k}^p\right)^{1-\theta} \left(\frac{p_{t+k}}{p_t}\right)^{\theta-1} \lambda_{t+k} y_{t+k}} \\ &= \mu_p \frac{p_t^N}{p_t^D}, \end{aligned}$$

where $\mu_p \equiv \theta/(\theta-1)$ and p_t^N and p_t^D are expressed recursively as

$$p_t^N = \lambda_t mc_t y_t + \beta\xi_p E_t \left\{ p_{t+1}^N \left(\frac{\pi_{t+1}}{X_{t,1}^p}\right)^{\theta} \right\},$$

and

$$p_t^D = \lambda_t y_t + \beta \xi_p \overline{\pi}^{(1-\gamma_p)(1-\theta)} E_t \left\{ p_{t+1}^D \left(\frac{\pi_{t+1}}{X_{t,1}^p} \right)^{\theta-1} \right\}.$$

Log linearization of p_t^D and p_t^N yields

$$\begin{aligned} \widehat{\left(\frac{\widetilde{p}_t^i}{p_t}\right)} &= \left(1-\beta\xi_p\right)\widehat{mc}_t + \beta\xi_p E_t \widehat{\left(\frac{\widetilde{p}_{t+1}^i}{p_t}\right)} + \beta\left(\gamma_z\right)^{1-\sigma}\xi_p E_t \widehat{\pi}_{t+1} \\ &\quad -\beta\xi_p\gamma_p\widehat{\pi}_t - \beta\xi_p\left(1-\gamma_p\right)E_t\widehat{\pi}_{t+1}^* \end{aligned}$$

From the definition of price index, we have

$$\widehat{\left(\frac{\widetilde{p}_t^i}{p_t}\right)} = \frac{\xi_p}{\left(1-\xi_p\right)}\left(\widehat{\pi}_t - \gamma_p\widehat{\pi}_{t-1} - \left(1-\gamma_p\right)\widehat{\pi}_t^*\right).$$

Equating the two previous expressions and solving for $\widehat{\pi}_t$ reads

$$\begin{aligned} \widehat{\pi}_t &= \frac{\left(1-\xi_p\right)\left(1-\beta\xi_p\right)}{\xi_p(1+\beta\gamma_p)}\widehat{mc}_t \\ &\quad + \frac{\gamma_p}{(1+\beta\gamma_p)}\widehat{\pi}_{t-1} + \frac{\beta}{(1+\beta\gamma_p)}E_t\widehat{\pi}_{t+1} \\ &\quad + \frac{\left(1-\gamma_p\right)}{(1+\beta\gamma_p)}\widehat{\pi}_t^* - \frac{\beta\left(1-\gamma_p\right)}{(1+\beta\gamma_p)}E_t\widehat{\pi}_{t+1}^*. \end{aligned} \tag{A.1.1}$$

A.1.2. The Jacobian Matrix

We derive the Jacobian matrix following the appendix of Magnusson and Mavroeidis (2010). Note that the companion matrix $\Phi(\varphi)$ can be written as

$$\Phi(\varphi) = BA + C$$

where $B = (I_k, 0, \cdots, 0)'$ is a $(kp) \times k$ matrix, $A = (\Phi_1, \cdots, \Phi_p)'$ is $k \times (kp)$ matrix of the VAR coefficients, and $C = (0, 0; I_{kp-1}, 0)$ is a $(kp) \times (kp)$ matrix.

Hence, $\Phi(\varphi)$ is linear in φ and the distance function is differentiable with respect to φ.

Since $\Phi(\varphi)' = A'B' + C'$, its derivative w.r.t φ is given by $I_{kp} \otimes B$. Using this last result and the properties of the Kronecker product, it is easy to show that

$$\begin{aligned} \frac{\partial g(\varphi, \vartheta)}{\partial \varphi'} &= I_{kp} \otimes \{[I - \gamma_f \Phi(\varphi)']e_\pi - \lambda_1 e_{mc}\}' B \\ &- \gamma_f [\Phi(\varphi)' \otimes e'_\pi B]. \end{aligned} \tag{A.1.2}$$

A.1.3. Derivatives

We transform our HNKPC to facilitate the computation of the derivatives. The vector of restrictions writes

$$\begin{aligned} g(\varphi, \vartheta) = \Phi(\varphi)' &\left\{ e'_\pi [I - \frac{\beta}{1+\beta\gamma_p}\Phi(\varphi)] - \frac{(1-\xi_p)(1-\beta\xi_p)}{\xi_p(1+\beta\gamma_p)} e'_{mc} \right\}' \\ &- \frac{\gamma_p}{1+\beta\gamma_p} e_\pi. \end{aligned} \tag{A.1.3}$$

Let $G_\varphi(\varphi, \vartheta) \equiv \frac{\partial g(\varphi,\vartheta)}{\partial \varphi'}$

$$\begin{aligned} G_\varphi(\varphi, \vartheta) = I_{kp} \otimes [I &- \frac{\beta}{1+\beta\gamma_p}\Phi(\varphi)']e_\pi - \frac{(1-\xi_p)(1-\beta\xi_p)}{\xi_p(1+\beta\gamma_p)} e_{mc} \\ &- \frac{\beta}{1+\beta\gamma_p}[\Phi(\varphi)' \otimes e'_\pi B]. \end{aligned} \tag{A.1.4}$$

The robust tests require the following derivatives:

$$\begin{aligned}\frac{\partial G_\varphi(\varphi,\vartheta)}{\partial \gamma_p} &= I_{kp}\otimes\left\{\frac{\beta^2}{(1+\beta\gamma_p)^2}\Phi(\varphi)'e_\pi\right.\\ &\left.+\frac{(1-\xi_p)(1-\beta\xi_p)\beta\xi_p}{[\xi_p(1+\beta\gamma_p)]^2}e_{mc}\right\}B'\\ &+\frac{\beta^2}{(1+\beta\gamma_p)^2}[\Phi(\varphi)'\otimes e_\pi' B],\end{aligned} \tag{A.1.5}$$

$$\frac{\partial G_\varphi(\varphi,\vartheta)}{\partial \xi_p} = -I_{kp}\otimes\left\{\frac{(1+\beta\gamma_p)(\beta\xi_p^2-1)}{[\xi_p(1+\beta\gamma_p)]^2}e_s\right\}'B,$$

$$\begin{aligned}\frac{\partial g(\varphi,\vartheta)}{\partial \gamma_p} &= \Phi(\varphi)'\left\{e_\pi'[\frac{\beta^2}{(1+\beta_1\gamma_p)^2}\Phi(\varphi)]\right.\\ &\left.-\frac{(1-\xi_p)(1-\beta\xi_p)\beta\xi_p}{[\xi_p(1+\beta\gamma_p)]^2}e_s'\right\}'\\ &-\frac{1}{(1+\beta\gamma_p)^2}e_\pi,\end{aligned} \tag{A.1.6}$$

$$\frac{\partial g(\varphi,\vartheta)}{\partial \xi_p} = -\Phi(\varphi)'\left\{e_\pi'\frac{(1+\beta\gamma_p)(\beta\xi_p^2-1)}{[\xi_p(1+\beta\gamma_p)]^2}\right\}',$$

$$\begin{aligned}\frac{\partial}{\partial\varphi'}\left(\frac{\partial g(\varphi,\vartheta)}{\partial\gamma_p}\right) &= I_{kp}\otimes\left\{e_\pi'[\frac{\beta^2}{(1+\beta\gamma_p)^2}\Phi(\varphi)]\right.\\ &\left.-\frac{(1-\xi_p)(1-\beta\xi_p)\beta\xi_p}{[\xi_p(1+\beta\gamma_p)]^2}e_s'\right\}'\\ &+\frac{\beta^2}{(1+\beta\gamma_p)^2}\Phi(\varphi)'I_{kp}\otimes(e_\pi'B)\end{aligned}, \tag{A.1.7}$$

$$\frac{\partial}{\partial\varphi'}\left(\frac{\partial g(\varphi,\vartheta)}{\partial\xi_p}\right) = -I_{kp}\otimes\left[e_\pi'\frac{(1+\beta\gamma_p)(\beta\xi_p^2-1)}{[\xi_p(1+\beta\gamma_p)]^2}\right]B. \tag{A.1.8}$$

A.2. Appendix to Chapter 5

The appendix reports a variant from the Smets and Wouters (2007) model with regime switching in shocks variances and in monetary policy.

A.2.1. Households

The representative household determines $\{c_t, B_t, K_t, I_t, z_t, l_t\}_{t=0}^{\infty}$ and its nominal wage $\widetilde{W}_0$ when it optimizes it, according to the indexation rule, which is defined in Section 2.2. In section 2.1, we derive how the household determines $\{c_t, B_t, K_t, I_t, z_t, l_t\}_{t=0}^{\infty}$ whereas the determination of wage is derived in section 2.2.

A.2.2. Standard problem

To determine the evolution of $\{c_t, B_t, K_t, I_t, z_t, l_t\}_{t=0}^{\infty}$, the household solves

$$\max_{\{c_t, B_t, K_t, I_t, z_t, l_t\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t U(c_t, l_t)$$

subject to

$$\begin{aligned}
\frac{B_{t-1}}{P_t} + \frac{\widetilde{W}_t}{P_t} l_t + (r_t z_t - \psi(z_t)) K_{t-1} + Div_t - c_t - I_t - \frac{1}{\varepsilon_{b,t}} \frac{B_t}{R_t P_t} &= 0 \\
K_t - \varepsilon_t^i (1 - S(I_t/I_{t-1})) I_t - (1-\delta) K_{t-1} &= 0 \\
\left(\frac{\widetilde{W}_t}{W_t}\right)^{-(1+\lambda_{w,t})/\lambda_{w,t}} L_t - l_t &= 0
\end{aligned}$$

and to the preference and investment shocks (whith $\widehat{\varepsilon}_t^x = \log \varepsilon_t^x$, $x = b, i$) :

$$\begin{aligned}
\widehat{\varepsilon}_t^b &= \rho \widehat{\varepsilon}_{t-1}^b + \sigma_t^b \eta_t^b, \\
\widehat{\varepsilon}_t^i &= \rho \widehat{\varepsilon}_{t-1}^i + \sigma_t^i \eta_t^i,
\end{aligned}$$

where, for $x = b, i$, $\eta_t^x \sim N(0,1)$ and the standard deviation σ_t^x is regime dependent.

The FOC write:

$$
\begin{aligned}
U^c(c_t, 1-l_t) &= \lambda_t \\
\frac{1}{\varepsilon_{b,t}}\frac{\lambda_t}{P_t} &= \beta R_t E_t \left\{ \frac{\lambda_{t+1}}{P_{t+1}} \right\}, \\
\nu_t &= \beta E_t \left\{ \lambda_{t+1} \left(r_{t+1} z_{t+1} - \psi\left(z_{t+1}\right) \right) + (1-\delta)\nu_{t+1} \right\}, \\
r_t &= \psi'\left(z_t\right), \\
\lambda_t &= \nu_t \varepsilon_t^i \left[(1 - S(I_t/I_{t-1}) - I_t/I_{t-1} S'(I_t/I_{t-1}) \right], \\
&\quad + \beta E_t \left\{ \nu_{t+1} \varepsilon_{t+1}^i \left(I_{t+1}/I_t\right)^2 S'(I_{t+1}/I_t) \right\}, \\
\chi_t &= U^l(c_t, l_t) + \lambda_t \frac{\widetilde{W}_t}{P_t}.
\end{aligned}
$$

Defining $\widetilde{\lambda}_t = \lambda_t z_t^{\sigma_c}$, $Q_t = \nu_t/\lambda_t$, $\widetilde{\chi}_t = \chi_t z_t^{\sigma_c - 1}$, we have

$$
\begin{aligned}
\widetilde{\lambda}_t &= U^c(\widetilde{c}_t, 1-l_t), && \text{(A.2.1)} \\
\widetilde{\lambda}_t &= \beta \left(\gamma_z\right)^{-\sigma_c} \varepsilon_{b,t} R_t E_t \left\{ \frac{\widetilde{\lambda}_{t+1}}{\pi_{t+1}} \right\}, && \text{(A.2.2)} \\
Q_t &= \beta \gamma_z^{-\sigma_c} E_t \left\{ \frac{\widetilde{\lambda}_{t+1}}{\widetilde{\lambda}_t} \left(r_{t+1} z_{t+1} - \psi\left(z_{t+1}\right) + (1-\delta) Q_{t+1} \right) \right\}, && \text{(A.2.3)} \\
r_t &= \psi'\left(z_t\right), && \text{(A.2.4)} \\
& Q_t \varepsilon_t^i \left[(1 - S(\gamma_z \widetilde{I}_t/\widetilde{I}_{t-1}) - \gamma_z \widetilde{I}_t/\widetilde{I}_{t-1} S'(\gamma_z \widetilde{I}_t/\widetilde{I}_{t-1}) \right] - 1 \\
&= -\beta \left(\gamma_z\right)^{-\sigma_c} E_t \left\{ \frac{\widetilde{\lambda}_{t+1}}{\widetilde{\lambda}_t} Q_{t+1} \varepsilon_{t+1}^i \left(\gamma_z \widetilde{I}_{t+1}/\widetilde{I}_t \right)^2 S'(\gamma_z \widetilde{I}_{t+1}/\widetilde{I}_t) \right\}, && \text{(A.2.5)} \\
\widetilde{\chi}_t &= U^l(c_t, l_t) z_t^{\sigma_c - 1} + \widetilde{\lambda}_t \frac{\widetilde{W}_t}{z_t P_t}. && \text{(A.2.6)}
\end{aligned}
$$

Following Smets and Wouters (2007), we specify

$$
U(c_t, l_t) = \frac{1}{1-\sigma_c} \left(c_t - h C_{t-1}\right)^{1-\sigma_c} \exp\left(\frac{\sigma_c - 1}{1+\sigma_L} l_t^{1+\sigma_L} \right),
$$

where C_{t-1} is the *aggregate consumption level* of the previous period. Thus,

$$\lambda_t = U^c(c_t, 1-l_t) = (c_t - hC_{t-1})^{-\sigma_c} \exp\left(\frac{\sigma_c - 1}{1+\sigma_L} l_t^{1+\sigma_L}\right),$$

which we stationarize as

$$\widetilde{\lambda}_t = \left(\widetilde{c}_t - \frac{h}{\gamma_z}\widetilde{C}_{t-1}\right)^{-\sigma_c} \exp\left(\frac{\sigma_c - 1}{1+\sigma_L} l_t^{1+\sigma_L}\right).$$

Define the marginal utility of leisure (with $\mathcal{L}_t = 1 - l_t$) as:

$$\begin{aligned} U^{\mathcal{L}}(c_t, \mathcal{L}_t) &\equiv -U^l(c_t, l_t) = (c_t - hC_{t-1})^{1-\sigma_c} \exp\left(\frac{\sigma_c - 1}{1+\sigma_L} (l_t)^{1+\sigma_L}\right) l_t^{\sigma_L}, \\ &= \lambda_t (c_t - hC_{t-1}) l_t^{\sigma_L}. \end{aligned}$$

which we stationarize as:

$$U^{\mathcal{L}}(c_t, \mathcal{L}_t) = \widetilde{\lambda}_t \left(\widetilde{c}_t - \frac{h}{\gamma_z}\widetilde{C}_{t-1}\right) l_t^{\sigma_L}.$$

Then, the marginal rate of substitution between consumption and leisure is:

$$mrs\,(\widetilde{c}_t, l_t) \equiv \frac{U^{\mathcal{L}}(\widetilde{c}_t, \mathcal{L}_t)}{U^c(\widetilde{c}_t, \mathcal{L}_t))} = \left(\widetilde{c}_t - \frac{h}{\gamma_z}\widetilde{C}_{t-1}\right) l_t^{\sigma_L}.$$

Linearizing it, we have

$$\widehat{mrs}_t = \sigma_L \widehat{l}_t + \frac{1}{1 - h/\gamma_z}\widehat{c}_t - \frac{h/\gamma_z}{1 - h/\gamma_z}\widehat{C}_{t-1}. \tag{A.2.7}$$

A.2.3. The determination of wage

For each period t, the wage, which cannot be optimized, is adjusted according to the rule

$$\widetilde{W}_{t+i} = \gamma_z (\pi_{t+i-1})^{\gamma_w} \left(\pi^*_{t+i}\right)^{1-\gamma_w} \widetilde{W}_{t+i-1}, \tag{A.2.8}$$

where π_t^* is the inflation target decided in period t by the central bank. Let the inflation factor be

$$X_{t,i}^w = \left\{ \begin{array}{c} 1 \text{ for } i = 0 \\ \gamma_z^i \left(\frac{P_{t+i-1}}{P_{t-1}}\right)^{\gamma_w} \left(\pi_{t+i}^* \pi_{t+i-1}^* ... \pi_{t+1}^*\right)^{1-\gamma_w} \text{ for } i = 1, ..., \infty \end{array} \right\},$$

or, equivalently:

$$X_{t,i}^w = \left\{ \begin{array}{c} 1 \text{ for } i = 0 \\ (\gamma_z)^i \prod_{l=1}^{i} \left((\pi_{t+l-1})^{\gamma_w} \left(\pi_{t+l}^*\right)^{1-\gamma_w} \right) \text{ for } i = 1, ..., \infty \end{array} \right\}.$$

This rule implies that

$$\widetilde{W}_{t+i} = X_{t,i}^w \widetilde{W}_t. \tag{A.2.9}$$

From the previous rule, we note that the wage is indexed on technological progress.

The wage optimizing household solves the following problem:

$$\max_{\widetilde{W}_t} E_t \sum_{i=0}^{\infty} (\beta \xi_w)^i U(c_{t+i}, l_{t+i}),$$

subject to

$$\begin{aligned} \frac{B_{t+i-1}}{P_{t+i}} + \frac{\widetilde{W}_{t+i}}{P_{t+i}} l_{t+i} + (r_{t+i} z_{t+i} - \psi(z_{t+i})) K_{t+i-1} + Div_{t+i} \\ -c_{t+i} - I_{t+i} - \frac{1}{\varepsilon_{b,t+i}} \frac{B_{t+i}}{R_{t+i} P_{t+i}} = 0, \\ \left(\frac{\widetilde{W}_{t+i}}{W_{t+i}}\right)^{-(1+\lambda_w)/\lambda_w} L_{t+i} - l_{t+i} = 0. \end{aligned}$$

and to the indexation rule (A.2.8).

The FOC writes

$$E_t \sum_{i=0}^{\infty} (\beta\xi_w)^i \left[\frac{\lambda_{t+i} l_{t+i} X_{t,i}^w}{P_{t+i}} - \frac{1+\lambda_w}{\lambda_w} \frac{\chi_{t+i} l_{t+i}}{\widetilde{W}_t} \right] = 0,$$

or, using (A.2.6)

$$E_t \sum_{i=0}^{\infty} (\beta\xi_w)^t \left[\frac{\lambda_{t+i} l_{t+i} X_{t,i}^w}{P_{t+i}} - \frac{1+\lambda_w}{\lambda_w} \left(\frac{U^l(c_{t+i}, l_{t+i}) l_{t+i}}{\widetilde{W}_t} + \frac{\lambda_{t+i} l_{t+i} X_{t,i}^w}{P_{t+i}} \right) \right] = 0.$$

Using the indexation rule $\widetilde{W}_t$ and replacing $U^l(c_{t+i}, l_{t+i})$ by $-U^{\mathcal{L}}(c_{t+i}, l_{t+i})$, we obtain

$$\frac{\widetilde{W}_t}{P_t} = (1+\lambda_w) \frac{E_t \sum_{i=0}^{\infty} (\beta\xi_w)^i U^{\mathcal{L}}(c_{t+i}, l_{t+i}) l_{t+i}}{E_t \sum_{i=0}^{\infty} (\beta\xi_w)^i \frac{P_t}{P_{t+i}} U^c(c_{t+i}, l_{t+i}) l_{t+i} X_{t,i}^w},$$

which can be stationarized as

$$\widetilde{w}_t \equiv \frac{\widetilde{W}_t}{z_t P_t} = (1+\lambda_w) \frac{E_t \sum_{i=0}^{\infty} \left(\beta (\gamma_z)^{1-\sigma_c} \xi_w\right)^i U^{\mathcal{L}}(c_{t+i}, l_{t+i}) l_{t+i}}{E_t \sum_{i=0}^{\infty} \left(\beta (\gamma_z)^{1-\sigma_c} \xi_w\right)^i \frac{P_t}{P_{t+i}} U^c(c_{t+i}, l_{t+i}) l_{t+i} \widetilde{X}_{t,i}^w},$$

or

$$\widetilde{w}_t = (1+\lambda_w) \frac{\widetilde{w}_t^N}{\widetilde{w}_t^D},$$

with

$$\widetilde{w}_t^N = E_t \sum_{i=0}^{\infty} \left(\beta (\gamma_z)^{1-\sigma_c} \xi_w\right)^i U^{\mathcal{L}}(\widetilde{c}_{t+i}, l_{t+i}) l_{t+i},$$

and

$$\widetilde{w}_t^D = E_t \sum_{i=0}^{\infty} \left(\beta (\gamma_z)^{1-\sigma_c} \xi_w\right)^i \frac{P_t}{P_{t+i}} U^c(\widetilde{c}_{t+i}, l_{t+i}) l_{t+i} \widetilde{X}_{t,i}^w.$$

Expressing $\widetilde{w}_t^N$ and $\widetilde{w}_t^D$ recursively, we have

$$\widetilde{w}_t^D = U^c(\widetilde{c}_t, l_t) l_t + \beta (\gamma_z)^{1-\sigma_c} \xi_w E_t \frac{\widetilde{X}_{t,1}^w}{\pi_{t+1}} \widetilde{w}_{t+1}^D.$$

Linearizing $\widetilde{w}_t^D$, we have

$$\begin{aligned}\widehat{\widetilde{w}_t^D} &= \left(1-\beta\left(\gamma_z\right)^{-\sigma_c}\xi_w\right)\left(\widehat{U_t^c}+\widehat{l}_t\right)\\ &\quad+\beta\left(\gamma_z\right)^{1-\sigma_c}\xi_w\left(E_t\widehat{\widetilde{w}_{t+1}^D}+\gamma_w\widehat{\pi}_t+(1-\gamma_w)E_t\widehat{\pi}_{t+1}^*-E_t\widehat{\pi}_{t+1}\right).\end{aligned}$$

In the same way, we have

$$\widetilde{w}_t^N = U^{\mathcal{L}}(\widetilde{c}_t, l_t)l_t + \beta\left(\gamma_z\right)^{1-\sigma_c}\xi_w E_t\widetilde{w}_{t+1}^N.$$

Linearizing $\widetilde{w}_t^N$, we have

$$\widehat{\widetilde{w}_t^N} = \left(1-\beta\left(\gamma_z\right)^{1-\sigma_c}\xi_w\right)\left(\widehat{U_t^{\mathcal{L}}}+\widehat{l}_t\right)+\beta\left(\gamma_z\right)^{1-\sigma_c}\xi_w E_t\widehat{\widetilde{w}_{t+1}^N}.$$

Linearization of $\widetilde{w}_t$ writes

$$\begin{aligned}\widehat{\widetilde{w}}_t &= \widehat{\widetilde{w}_t^N}-\widehat{\widetilde{w}_t^D}\\ &= \left(1-\beta\left(\gamma_z\right)^{1-\sigma_c}\xi_w\right)\left(\widehat{U_t^{\mathcal{L}}}-\widehat{U_t^c}\right)+\beta\left(\gamma_z\right)^{1-\sigma_c}\xi_w\\ &\quad\times\left(E_t\widehat{\widetilde{w}}_{t+1}-\gamma_w\widehat{\pi}_t-(1-\gamma_w)E_t\widehat{\pi}_{t+1}^*+E_t\widehat{\pi}_{t+1}\right). \qquad \text{(A.2.10)}\end{aligned}$$

From the definition of the wage index:

$$W_t = \left(\int_0^1 \left(W_t^i\right)^{-1/\lambda_{w,r}} di\right)^{-\lambda_{w,t}},$$

we deduce the stationarized real wage $w_t = W_t/z_tP_t$ as

$$w_t = \left(\left(1-\xi_w\right)\left(\widetilde{w}_t\right)^{-1/\lambda_{w,r}}+\xi_w\left(w_{t-1}\frac{\pi_{t-1}^{\gamma_w}\left(\pi_t^*\right)^{1-\gamma_w}}{\pi_t}\right)^{-1/\lambda_{w,r}}\right)^{-\lambda_{w,t}}.$$

Thus, w_t and $\widetilde{w}_t$ evolve at the same rate. Linearizing w_t, we obtain:

$$\widehat{w}_t = \left(1-\xi_w\right)\widehat{\widetilde{w}}_t+\xi_w\left(\widehat{w}_{t-1}+\gamma_w\widehat{\pi}_{t-1}+(1-\gamma_w)\widehat{\pi}_t^*-\widehat{\pi}_t\right) \qquad \text{(A.2.11)}$$

from which we have:

$$E_t\widehat{\widetilde{w}}_{t+1} = \frac{1}{(1-\xi_w)}E_t\widehat{w}_{t+1} - \frac{\xi_w}{(1-\xi_w)}\left(\widehat{w}_t + \gamma_w\widehat{\pi}_t + (1-\gamma_w)E_t\widehat{\pi}^*_{t+1} - E_t\widehat{\pi}_{t+1}\right). \tag{A.2.12}$$

Inserting (A.2.10) and (A.2.12) in (A.2.11), we have:

$$\begin{aligned}\widehat{w}_t &= (1-\xi_w)\left(1-\beta(\gamma_z)^{1-\sigma_c}\xi_w\right)\left(\widehat{U^{\mathcal{L}}_t} - \widehat{U^c_t}\right)\\ &+\beta(\gamma_z)^{1-\sigma_c}\xi_w\left(E_t\widehat{w}_{t+1} - \xi_w\widehat{w}_t - \gamma_w\widehat{\pi}_t - (1-\gamma_w)E_t\widehat{\pi}^*_{t+1} + E_t\widehat{\pi}_{t+1}\right)\\ &+\xi_w\left(\widehat{w}_{t-1} + \gamma_w\widehat{\pi}_{t-1} + (1-\gamma_w)\widehat{\pi}^*_t - \widehat{\pi}_t\right),\end{aligned}$$

or

$$\begin{aligned}\widehat{w}_t &= \frac{(1-\xi_w)\left(1-\beta(\gamma_z)^{1-\sigma_c}\xi_w\right)}{1+\beta(\gamma_z)^{1-\sigma_c}(\xi_w)^2}\left(\widehat{U^{\mathcal{L}}_t} - \widehat{U^c_t}\right) - \frac{(1+\beta(\gamma_z)^{1-\sigma_c}\gamma_w)\xi_w}{1+\beta(\gamma_z)^{1-\sigma_c}(\xi_w)^2}\widehat{\pi}_t\\ &+\frac{\xi_w}{1+\beta(\gamma_z)^{1-\sigma_c}(\xi_w)^2}\left(\widehat{w}_{t-1} + \gamma_w\widehat{\pi}_{t-1} + (1-\gamma_w)\widehat{\pi}^*_t\right)\\ &+\frac{\beta(\gamma_z)^{1-\sigma_c}\xi_w}{1+\beta(\gamma_z)^{1-\sigma_c}(\xi_w)^2}\left(E_t\widehat{w}_{t+1} - (1-\gamma_w)E_t\widehat{\pi}^*_{t+1} + E_t\widehat{\pi}_{t+1}\right).\end{aligned} \tag{A.2.13}$$

Smets and Wouters (2007) define the wage markup as the ratio between the real wage and the marginal rate of substitution between consumption and leisure: $\mu^w_t \equiv w_t/mrs_t$. Following them, we have

$$\widehat{mrs}_t = \widehat{w}_t - \widehat{\mu}^w_t, \tag{A.2.14}$$

with $\widehat{mrs}_t = \widehat{U^{\mathcal{L}}_t} - \widehat{U^c_t}$. Substituting this expression in (A.2.13) and factorizing terms in $\widehat{w}_t$, we have:

$$\begin{aligned}\widehat{w}_t &= -\frac{(1-\xi_w)\left(1-\beta(\gamma_z)^{1-\sigma_c}\xi_w\right)}{\xi_w\left(1+\beta(\gamma_z)^{1-\sigma_c}\right)}\widehat{\mu}_t^w - \frac{1+\beta(\gamma_z)^{1-\sigma_c}\gamma_w}{1+\beta(\gamma_z)^{1-\sigma_c}}\widehat{\pi}_t \\ &+\frac{1}{1+\beta(\gamma_z)^{1-\sigma_c}}\left(\widehat{w}_{t-1}+\gamma_w\widehat{\pi}_{t-1}+(1-\gamma_w)\widehat{\pi}_t^*\right) \\ &+\frac{\beta(\gamma_z)^{1-\sigma_c}}{1+\beta(\gamma_z)^{1-\sigma_c}}\left(E_t\widehat{w}_{t+1}-(1-\gamma_w)E_t\widehat{\pi}_{t+1}^*+E_t\widehat{\pi}_{t+1}\right).\end{aligned} \tag{A.2.15}$$

Using the Kimball aggregator instead, we obtain

$$\begin{aligned}\widehat{w}_t &= -\frac{(1-\xi_w)\left(1-\beta(\gamma_z)^{1-\sigma_c}\xi_w\right)}{\xi_w\left(1+\beta(\gamma_z)^{1-\sigma_c}\right)\left((\mu_w-1)\varsigma+1\right)}\widehat{\mu}_t^w - \frac{1+\beta(\gamma_z)^{1-\sigma_c}\gamma_w}{1+\beta(\gamma_z)^{1-\sigma_c}}\widehat{\pi}_t \\ &+\frac{1}{1+\beta(\gamma_z)^{1-\sigma_c}}\left(\widehat{w}_{t-1}+\gamma_w\widehat{\pi}_{t-1}+(1-\gamma_w)\widehat{\pi}_t^*\right) \\ &+\frac{\beta(\gamma_z)^{1-\sigma_c}}{1+\beta(\gamma_z)^{1-\sigma_c}}\left(E_t\widehat{w}_{t+1}-(1-\gamma_w)E_t\widehat{\pi}_{t+1}^*+E_t\widehat{\pi}_{t+1}\right),\end{aligned} \tag{A.2.16}$$

where $\mu_w \equiv 1+\lambda_w$.

A.2.4. Firms

The firm's program is described in two steps. First, for a given level of production, the firm determines the quantities of capital and labour that minimize its variable total cost $r_t z_t^i k_{t-1}^i + w_t L_t^i$, subject to the production constraint

$$y_t^i = A_t\left(z_t^i k_{t-1}^i\right)^\alpha\left((\gamma_z)^t L_t^i\right)^{1-\alpha} - (\gamma_z)^t\phi. \tag{A.2.17}$$

Here $\phi > 0$ is a fixed cost, A_t is an exogenous technological progress, following a known stochastic process given by :

$$A_t = f\left(A_{t-1}, \epsilon_{A,t}\right), \tag{A.2.18}$$

where $\epsilon_{A,t}$ is an i.i.d technological shock . The FOC writes:

$$z_t^i k_{t-1}^i = \left(\frac{\alpha}{1-\alpha}\right)^{1-\alpha} \left(\frac{w_t}{r_t}\right)^{1-\alpha} \left(y_t^i + (\gamma_z)^t \phi\right), \tag{A.2.19}$$

and

$$L_t^i = \left(\frac{\alpha}{1-\alpha}\right)^{-\alpha} \left(\frac{w_t}{r_t}\right)^{-\alpha} \left(y_t^i + (\gamma_z)^t \phi\right). \tag{A.2.20}$$

From the FOC, we deduce the total variable cost (for $y_t^i > 0$):

$$\begin{aligned} r_t z_t^i k_{t-1}^i + w_t L_t^i &= \Upsilon \frac{r_t^\alpha w_t^{1-\alpha}}{A_t} (y_t^i + (\gamma_z)^t \phi) && \text{(A.2.21)} \\ &\equiv s_t(y_t^i + \phi), && \text{(A.2.22)} \end{aligned}$$

where $s_t \equiv s(r_t, w_t, A_t)$ is the marginal cost of production and $\Upsilon = \alpha^{-\alpha}(1-\alpha)^{-(1-\alpha)}$.

Second, in each and every period, each firm faces a constant probability $1-\xi_p$ of being able to optimize its price $\widetilde{p_t^i}$. Otherwise, it determines its price according to the rule

$$p_{t+k}^i = (\pi_{t+k-1})^{\gamma_p} \left(\pi_{t+k}^*\right)^{1-\gamma_p} p_{t+k-1}^i,$$

where π_{t+k-1} is the past inflation factor and π_{t+k}^* is the inflation target. Define the indexation factor as

$$X_{t,k}^p = \left\{ \begin{array}{c} 1 \text{ for } k=0 \\ \left(\frac{P_{t+k-1}}{P_{t-1}}\right)^{\gamma_p} \left(\pi_{t+k}^* \pi_{t+k-1}^* ... \pi_{t+1}^*\right)^{1-\gamma_p} \text{ for } k=1,...,\infty \end{array} \right\},$$

or, equivalently:

$$X_{t,k}^p = \left\{ \begin{array}{c} 1 \text{ for } i=0 \\ \prod_{l=1}^{k} \left((\pi_{t+l-1})^{\gamma_p} \left(\pi_{t+l}^*\right)^{1-\gamma_p}\right) \text{ for } i=1,...,\infty \end{array} \right\}.$$

This rule implies

$$p_{t+k}^{i} = X_{t,k}^{p} \widetilde{p_t^i}, \tag{A.2.23}$$

In this context, the optimal price $\widetilde{p_t^i}$ chosen by an optimizing firm at t is the solution to the following program:

$$\max_{\widetilde{p}_t^i} \quad \underset{t}{E} \sum_{k=0}^{\infty} \left(\beta\xi_p\right)^k \frac{\lambda_{t+k}}{\lambda_t} \left(\frac{\widetilde{p}_t^i X_{t,k}^p}{p_{t+k}} - s_{t+k} \right) y_{t+k}^i$$

subject to

$$y_{t+k}^i = \left(\frac{\widetilde{p}_t^i X_{t,k}^p}{p_{t+k}} \right)^{-\theta} y_{t+k}$$

wherep_t is the aggregate price index and y_t is the aggregate output.

The FOC writes

$$\frac{\widetilde{p}_t^i}{p_t} \underset{t}{E} \sum_{k=0}^{\infty} \left(\beta\xi_p\right)^k \frac{p_t}{p_{t+k}} X_{t,k}^p \lambda_{t+k} y_{t+k}^i = \mu_p \underset{t}{E} \sum_{k=0}^{\infty} \left(\beta\xi_p\right)^k \lambda_{t+k} s_{t+k} y_{t+k}^i$$

or

$$\frac{\widetilde{p}_t^i}{p_t} \underset{t}{E} \sum_{k=0}^{\infty} \left(\beta\xi_p\right)^k \left(X_{t,k}^p\right)^{1-\theta} \left(\frac{p_t}{p_{t+k}} \right)^{(1-\theta)} \lambda_{t+k} y_{t+k} =$$
$$\mu_p \underset{t}{E} \sum_{k=0}^{\infty} \left(\beta\xi_p\right)^k \left(X_{t,k}^p\right)^{-\theta} \left(\frac{p_{t+k}}{p_t} \right)^{\theta} \lambda_{t+k} s_{t+k} y_{t+k},$$

or

$$\frac{\widetilde{p}_t^i}{p_t} = \mu_p \frac{\underset{t}{E} \sum_{k=0}^{\infty} \left(\beta\xi_p\right)^k \left(X_{t,k}^p\right)^{-\theta} \left(\frac{p_{t+k}}{p_t}\right)^{\theta} \lambda_{t+k} s_{t+k} y_{t+k}}{\underset{t}{E} \sum_{k=0}^{\infty} \left(\beta\xi_p\right)^k \left(X_{t,k}^p\right)^{1-\theta} \left(\frac{p_{t+k}}{p_t}\right)^{\theta-1} \lambda_{t+k} y_{t+k}},$$

where $\mu_p \equiv \theta/(\theta-1)$. Stationarizing it, we have

$$\frac{\widetilde{p}_t^i}{p_t} = \mu_p \frac{\underset{t}{E} \sum_{k=0}^{\infty} \left(\beta \left(\gamma_z\right)^{1-\sigma} \xi_p\right)^k \left(X_{t,k}^p\right)^{-\theta} \left(\frac{p_{t+k}}{p_t}\right)^{\theta} \widetilde{\lambda}_{t+k} s_{t+k} \widetilde{y}_{t+k}}{\underset{t}{E} \sum_{k=0}^{\infty} \left(\beta \left(\gamma_z\right)^{1-\sigma} \xi_p\right)^k \left(X_{t,k}^p\right)^{1-\theta} \left(\frac{p_{t+k}}{p_t}\right)^{\theta-1} \widetilde{\lambda}_{t+k} \widetilde{y}_{t+k}},$$

or

$$\frac{\widetilde{p}_t^i}{p_t} = \mu_p \frac{p_t^N}{p_t^D},$$

where p_t^N and p_t^D are expressed recursively as

$$p_t^N = \lambda_t s_t y_t + \beta \left(\gamma_z\right)^{1-\sigma} \xi_p E_t \left\{ p_{t+1}^N \left(\frac{\pi_{t+1}}{X_{t,1}^p}\right)^{\theta} \right\},$$

and

$$p_t^D = \lambda_t y_t + \beta \left(\gamma_z\right)^{1-\sigma} \xi_p \overline{\pi}^{\left(1-\gamma_p\right)(1-\theta)} E_t \left\{ p_{t+1}^D \left(\frac{\pi_{t+1}}{X_{t,1}^p}\right)^{\theta-1} \right\}.$$

Linearizing them, we have

$$\begin{aligned} \widehat{\left(\frac{\widetilde{p}_t^i}{p_t}\right)} &= \left(1 - \beta \left(\gamma_z\right)^{1-\sigma} \xi_p\right) \widehat{s}_t + \beta \left(\gamma_z\right)^{1-\sigma} \xi_p E_t \widehat{\left(\frac{\widetilde{p}_{t+1}^i}{p_t}\right)} + \beta \left(\gamma_z\right)^{1-\sigma} \xi_p E_t \widehat{\pi}_{t+1} \\ &\quad - \beta \left(\gamma_z\right)^{1-\sigma} \xi_p \gamma_p \widehat{\pi}_t - \beta \left(\gamma_z\right)^{1-\sigma} \xi_p \left(1 - \gamma_p\right) E_t \widehat{\pi}_{t+1}^* \end{aligned}$$

From the definition of the price index, we have

$$\widehat{\left(\frac{\widetilde{p}_t^i}{p_t}\right)} = \frac{\xi_p}{\left(1 - \xi_p\right)} \left(\widehat{\pi}_t - \gamma_p \widehat{\pi}_{t-1} - \left(1 - \gamma_p\right) \widehat{\pi}_t^*\right)$$

where:

$$\begin{aligned} &\frac{\xi_p}{\left(1 - \xi_p\right)} \left(\widehat{\pi}_t - \gamma_p \widehat{\pi}_{t-1} - \left(1 - \gamma_p\right) \widehat{\pi}_t^*\right) = \\ &\left(1 - \beta \left(\gamma_z\right)^{1-\sigma} \xi_p\right) \widehat{s}_t + \frac{\beta \left(\gamma_z\right)^{1-\sigma} \left(\xi_p\right)^2}{\left(1 - \xi_p\right)} \left(E_t \widehat{\pi}_{t+1} - \gamma_p \widehat{\pi}_t - \left(1 - \gamma_p\right) E_t \widehat{\pi}_{t+1}^*\right) \\ &\qquad + \beta \left(\gamma_z\right)^{1-\sigma} \xi_p \left(E_t \widehat{\pi}_{t+1} - \gamma_p \widehat{\pi}_t - \left(1 - \gamma_p\right) E_t \widehat{\pi}_{t+1}^*\right), \end{aligned}$$

i.e.

$$\begin{aligned}\widehat{\pi}_t &= \frac{\left(1-\xi_p\right)\left(1-\beta\left(\gamma_z\right)^{1-\sigma}\xi_p\right)}{\xi_p(1+\beta\left(\gamma_z\right)^{1-\sigma}\gamma_p)}\widehat{s}_t \\ &+\frac{1}{(1+\beta\left(\gamma_z\right)^{1-\sigma}\gamma_p)}\left(\gamma_p\widehat{\pi}_{t-1}+\left(1-\gamma_p\right)\widehat{\pi}_t^*\right) \\ &+\frac{\beta\left(\gamma_z\right)^{1-\sigma}}{(1+\beta\left(\gamma_z\right)^{1-\sigma}\gamma_p)}\left(E_t\widehat{\pi}_{t+1}-\left(1-\gamma_p\right)E_t\widehat{\pi}_{t+1}^*\right). \quad \text{(A.2.24)}\end{aligned}$$

Smets et Wouters (2007) define the price markup as as the ratio between marginla productivity of labour and real wage: $\mu_t^p \equiv mpl_t/w_t$.Following them, we have

$$\begin{aligned}\widehat{\mu}_t^p &= \widehat{A}_t+\alpha\left(\widehat{k}_{t-1}+\widehat{z}_{t-1}-\widehat{L}_t\right)-\widehat{w}_t \\ &= \widehat{A}_t-\alpha\widehat{r}_t-(1-\alpha)\widehat{w}_t \\ &= -\widehat{s}_t\end{aligned}$$

Substituting this in (A.2.24), we have:

$$\begin{aligned}\widehat{\pi}_t &= -\frac{\left(1-\xi_p\right)\left(1-\beta\left(\gamma_z\right)^{1-\sigma}\xi_p\right)}{\xi_p(1+\beta\left(\gamma_z\right)^{1-\sigma}\gamma_p)}\widehat{\mu}_t^p \\ &+\frac{1}{(1+\beta\left(\gamma_z\right)^{1-\sigma}\gamma_p)}\left(\gamma_p\widehat{\pi}_{t-1}+\left(1-\gamma_p\right)\widehat{\pi}_t^*\right) \\ &+\frac{\beta\left(\gamma_z\right)^{1-\sigma}}{(1+\beta\left(\gamma_z\right)^{1-\sigma}\gamma_p)}\left(E_t\widehat{\pi}_{t+1}-\left(1-\gamma_p\right)E_t\widehat{\pi}_{t+1}^*\right). \quad \text{(A.2.25)}\end{aligned}$$

Using the Kimball aggregator, we have (with $\mu_p \equiv 1+\lambda_p$) :

$$\begin{aligned}\widehat{\pi}_t &= -\left(\frac{\left(1-\xi_p\right)\left(1-\beta\left(\gamma_z\right)^{1-\sigma}\xi_p\right)}{\xi_p(1+\beta\left(\gamma_z\right)^{1-\sigma}\gamma_p)\left(\left(\mu_p-1\right)\varsigma_p+1\right)}\right)\widehat{\mu}_t^p \\ &+\frac{1}{(1+\beta\left(\gamma_z\right)^{1-\sigma}\gamma_p)}\left(\gamma_p\widehat{\pi}_{t-1}+\left(1-\gamma_p\right)\widehat{\pi}_t^*\right) \\ &+\frac{\beta\left(\gamma_z\right)^{1-\sigma}}{(1+\beta\left(\gamma_z\right)^{1-\sigma}\gamma_p)}\left(E_t\widehat{\pi}_{t+1}-\left(1-\gamma_p\right)E_t\widehat{\pi}_{t+1}^*\right). \quad \text{(A.2.26)}\end{aligned}$$

A.2.5. Monetary policy

As in Smets and Wouters (2007), the central bank adjusts the nominal interest rate according to the rule

$$\frac{R_t}{R_t^*}=\left(\frac{R_{t-1}}{R_t^*}\right)^{\rho(s_t)}\left[\left(\frac{\pi_t}{\pi_t^*}\right)^{r_\pi(s_t)}\left(\frac{Y_t}{Y_t^f}\right)^{r_y(s_t)}\right]^{1-\rho(s_t)}\left(\frac{Y_t/Y_{t-1}}{Y_t^f/Y_{t-1}^f}\right)^{r_{\Delta y}(s_t)}\varepsilon_t^R, \quad \text{(A.2.27)}$$

where $R_t^* \equiv (1/\beta\gamma^{-\sigma_c})\pi_t^*$ and $\pi_t^* = \pi^*(s_t)$ is the inflation target when the current regime is s_t. Unlike in Smets and Wouters (2007), the inflation target and the coefficients of the rule $(\rho(s_t), r_\pi(s_t), r_y(s_t), r_{\Delta y}(s_t))$ depend upon the regime in place. Using the fact that $\widehat{R}_t^* = \widehat{\pi}_t^*$, log-linearization of (A.2.27) leads to

$$\begin{aligned}(\widehat{R}_t-\widehat{\pi}_t^*) &= \rho(s_t)(\widehat{R}_{t-1}-\widehat{\pi}_t^*) \quad \text{(A.2.28)} \\ &+(1-\rho(s_t))[r_\pi(s_t)(\widehat{\pi}_t-\widehat{\pi}_t^*)+r_y(s_t)(\widehat{Y}_t-\widehat{Y}_t^f)] \\ &+r_{\Delta_y}(s_t)(\widehat{Y}_t-\widehat{Y}_t^f)-(\widehat{Y}_{t-1}-\widehat{Y}_{t-1}^f)]+\epsilon_t^r,\end{aligned}$$

where $\widehat{\pi}_t^* \equiv \ln(\pi_t^*/\pi)$, π denoting the long run inflation factor.

Practically, we assume as Schorfheide (2005) that the inflation target is computed from an annualized inflation rate $\ln(\pi_a^*(s_t))$ and evolves according to:

$$\ln(\pi_a^*(s_t) = \begin{cases} \ln \pi_a^L \text{ if } s_t = 1, \\ \ln \pi_a^H \text{ if } s_t = 2, \end{cases}$$

where $\pi_a^*(s_t)$ is the annualized inflation factor. With the transition matrix P, we can easily show that the annualized long run inflation rate is

$$\pi_a = \exp\left(\frac{1-p_{22}}{2-p_{11}-p_{22}} \ln \pi_a^L + \frac{1-p_{11}}{2-p_{11}-p_{22}} \ln \pi_a^H\right).$$

We express the π_a to a quarterly basis as in the model.

A.2.6. Steady State

$$\left(1+\frac{\phi}{Y}\right) = \mu_P,$$

$$r = \frac{1-\beta\gamma^{-\sigma_c}(1-\delta)}{\beta\gamma^{-\sigma_c}},$$

$$\frac{K}{L} = \left(\frac{\alpha}{r\mu_P}\right)^{\frac{1}{1-\alpha}},$$

$$w = \frac{1-\alpha}{\alpha} r \frac{K}{L},$$

$$\frac{Y}{L} = \frac{1}{\mu_P}\left(\frac{K}{L}\right)^{\alpha},$$

$$\frac{G}{L} = g_y \frac{Y}{L},$$

$$\frac{I}{L} = (\gamma - 1 + \delta)\frac{K}{L},$$

$$\frac{C}{L} = \frac{Y}{L} - \frac{I}{L} - \frac{G}{L},$$

. From household's optimization program, we have

$$\left(1-\frac{h}{\gamma}\right) L^{1+\sigma_L} = \frac{w}{(1+\lambda_w)(C/L)}.$$

which is a condition used to derive the linearized consumption dynamics.

$$R = \frac{\Pi}{\beta\gamma^{-\sigma_c}}$$

A.3. Log linearization of other equations

$\cdot\ \widetilde{\lambda}_t = \boldsymbol{\beta}\left(\gamma_z\right)^{-\sigma_c}\boldsymbol{\varepsilon}_{b,t}R_tE_t\left\{\frac{\widetilde{\lambda}_{t+1}}{\pi_{t+1}}\right\}$:

$$\widehat{\lambda}_t = E_t\widehat{\lambda}_{t+1} + \widehat{R}_t - E_t\widehat{\pi}_{t+1} + \widehat{\varepsilon}_t^b \tag{A.3.1}$$

$\cdot\ \widetilde{\lambda}_t = U^c(\widetilde{c}_t, 1-l_t) = (\widetilde{c}_t - h/\gamma_z C_{t-1})^{-\sigma_c}\exp((\sigma_c-1)/(1+\sigma_L)\,(l_t)^{1+\sigma_L})$.

Here, tilded variables refers to their stationarized values.

$$\widehat{C}_t = \frac{h}{\gamma_z}\widehat{C}_{t-1} - \frac{1-h/\gamma_z}{\sigma_c}\widehat{\lambda}_t + \frac{(1-h/\gamma_z)\,(\sigma_c-1)\,l^{1+\sigma_L}}{\sigma_c}\widehat{l}_t, \tag{A.3.2}$$

from which we deduce

$$\widehat{\lambda}_t = -\frac{\sigma_c}{1-h/\gamma_z}\widehat{C}_t + \frac{h/\gamma_z\sigma_c}{1-h/\gamma_z}\widehat{C}_{t-1} + (\sigma_c-1)\,l^{1+\sigma_L}\widehat{l}_t.$$

Then,

$$E_t\widehat{\lambda}_{t+1} = \frac{h/\gamma_z\sigma_c}{1-h/\gamma_z}\widehat{C}_t - \frac{\sigma_c}{1-h/\gamma_z}E_t\widehat{C}_{t+1} + (\sigma_c-1)\,l^{1+\sigma_L}E_t\widehat{l}_{t+1}. \tag{A.3.3}$$

Inserting (A.3.3) and (A.3.1) in (A.3.2), we get:

$$\begin{aligned}\widehat{C}_t \;=\; & \frac{h/\gamma_z}{1+h/\gamma_z}\widehat{C}_{t-1} + \frac{1}{1+h/\gamma_z}E_t\widehat{C}_{t+1} + \frac{(1-h/\gamma_z)\,(\sigma_c-1)\,l^{1+\sigma_L}}{\sigma_c\,(1+h/\gamma_z)}\left[\widehat{l}_t - E_t\widehat{l}_{t+1}\right] \\ & -\frac{(1-h/\gamma_z)}{\sigma_c\,(1+h/\gamma_z)}\left[\widehat{R}_t - E_t\widehat{\pi}_{t+1} + \widehat{\varepsilon}_t^b\right].\end{aligned} \tag{A.3.4}$$

Since at the steady state $(1-h/\gamma_z)\, l^{1+\sigma_L} = wl/((1+\lambda_w)c)$, we obtain the dynamics of consumption as

$$\begin{aligned}\widehat{C}_t &= \frac{h/\gamma_z}{1+h/\gamma_z}\widehat{C}_{t-1} + \frac{1}{1+h/\gamma_z}E_t\widehat{C}_{t+1} + \frac{(\sigma_c-1)\, wl/((1+\lambda_w)c)}{\sigma_c\,(1+h/\gamma_z)}\left[\widehat{l}_t - E_t\widehat{l}_{t+1}\right] \\ &\quad -\frac{(1-h/\gamma_z)}{\sigma_c\,(1+h/\gamma_z)}\left[\widehat{R}_t - E_t\widehat{\pi}_{t+1} + \widehat{\varepsilon}_t^b\right]. \end{aligned} \tag{A.3.5}$$

$$\cdot\; Q_t = \beta\,(\gamma_z)^{-\sigma_c}\, E_t\left\{\tfrac{\widetilde{\lambda}_{t+1}}{\widetilde{\lambda}_t}\left(r_{t+1}z_{t+1} - \psi\,(z_{t+1}) + (1-\delta)Q_{t+1}\right)\right\}$$

At the steady state, we have $\overline{z}=1$, $\overline{Q}=1$, $r^k = \psi'\,(1)$ and $\psi\,(\overline{z}) = \psi\,(1) = 0$.
$1 = \beta\,(\gamma_z)^{-\sigma_c}\,(r^k + 1 - \delta)$, i.e $r^k = \left(1 - \beta\,(\gamma_z)^{-\sigma_c}\,(1-\delta)\right) / \left(\beta\,(\gamma_z)^{-\sigma_c}\right)$.
Then,

$$\begin{aligned}\widehat{Q}_t &= -\widehat{\lambda}_t + E_t\widehat{\lambda}_{t+1} + \beta\,(\gamma_z)^{-\sigma_c}\, r^k\,(E_t\widehat{r}_{t+1} + E_t\widehat{z}_{t+1}) - \beta\,(\gamma_z)^{-\sigma_c}\, r^k E_t\widehat{z}_{t+1} \\ &\quad +\beta\,(\gamma_z)^{-\sigma_c}\,(1-\delta)E_t\widehat{Q}_{t+1}.\end{aligned}$$

Using (A.3.1) and making some simplifications, we have:

$$\widehat{Q}_t = \frac{r^k}{1-\delta+r^k}E_t\widehat{r}_{t+1} + \frac{1-\delta}{1-\delta+r^k}E_t\widehat{Q}_{t+1} - \left(\widehat{R}_t - E_t\widehat{\pi}_{t+1} + \widehat{\varepsilon}_t^b\right). \tag{A.3.6}$$

$$\begin{aligned}&\cdot\; Q_t\varepsilon_t^i\left[(1 - S\left(\gamma_z\widetilde{I}_t/\widetilde{I}_{t-1}\right) - \gamma_z\widetilde{I}_t/\widetilde{I}_{t-1}S'\left(\gamma_z\widetilde{I}_t/\widetilde{I}_{t-1}\right)\right] - 1 \\ &= -\beta\,(\gamma_z)^{-\sigma_c}\, E_t\left\{\tfrac{\widetilde{\lambda}_{t+1}}{\widetilde{\lambda}_t}Q_{t+1}\varepsilon_{t+1}^i\left(\gamma_z\widetilde{I}_{t+1}/\widetilde{I}_t\right)^2 S'\left(\gamma_z\widetilde{I}_{t+1}/\widetilde{I}_t\right)\right\}\end{aligned}$$

Assumptions: $S(\gamma_z) = 0$; $S'(\gamma_z) = 0$; $S''(\gamma_z) = \varphi$.

We have:

$$\widehat{Q}_t + \widehat{\varepsilon}_t^i - (\gamma_z)^2\, S''(\gamma_z)\left[\widehat{I}_t - \widehat{I}_{t-1}\right] = -\beta\,(\gamma_z)^{-\sigma_c}\,(\gamma_z)^2\, S''(\gamma_z)\left[E_t\widehat{I}_{t+1} - \widehat{I}_t\right].$$

Then,

$$\widehat{I}_t = \frac{1}{\left(1+\beta\left(\gamma_z\right)^{1-\sigma_c}\right)\left(\gamma_z\right)^2\varphi}\widehat{Q}_t + \frac{1}{1+\beta\left(\gamma_z\right)^{1-\sigma_c}}\widehat{I}_{t-1} + \frac{\beta\left(\gamma_z\right)^{1-\sigma_c}}{1+\beta\left(\gamma_z\right)^{1-\sigma_c}}E_t\widehat{I}_{t+1} + \frac{1}{\left(1+\beta\left(\gamma_z\right)^{1-\sigma_c}\right)\left(\gamma_z\right)^2\varphi}\widehat{\varepsilon}_t^i. \quad \text{(A.3.7)}$$

Note that Smets and Wouters (2007) normalize the investment shock: $\widehat{\varepsilon}_t^{i'} = \widehat{\varepsilon}_t^i / \left(\left(1+\beta\left(\gamma_z\right)^{1-\sigma_c}\right)\left(\gamma_z\right)^2\varphi\right)$.

$\cdot\ r_t = \psi'\left(z_t\right)$

We have

$$\widehat{z}_t = z_1\widehat{r}_t,$$

where $z_1 \equiv \psi'\left(1\right)/\psi''\left(1\right)$ is the inverse of the elasticity of adjustment cost function.

$\cdot\ \widetilde{K}_t = (1-\delta)/\gamma_z\widetilde{K}_{t-1} + \varepsilon_t^i/\gamma_z\left[1 - S\left(\gamma_z I_t/I_{t-1}\right)\right]\widetilde{I}_t$

We have :

$$\widehat{K}_t = \frac{1-\delta}{\gamma_z}\widehat{K}_{t-1} + \left[1 - \frac{(1-\delta)}{\gamma_z}\right]\widehat{I}_t + \left[1 - \frac{(1-\delta)}{\gamma_z}\right]\widehat{\varepsilon}_t^i. \quad \text{(A.3.8)}$$

$\cdot\ Y_t = C_t + G_t + I_t + \psi\left(z_t\right)K_{t-1}$

We have :

$$Y\widehat{Y}_t = C\widehat{C}_t + G\widehat{G}_t + I\widehat{I}_t + \psi(1)K\widehat{K}_{t-1} + \psi'(1)K\widehat{z}_t$$

Since $\psi\left(1\right) = 0$, $\psi'\left(1\right) = r^k$ and $\widehat{z}_t = (1/\epsilon_{\psi'})\widehat{r}_t$,

$$\widehat{Y}_t = (1-(\gamma_z - 1 + \delta)\,k_y - g_y)\widehat{C}_t + g_y\widehat{G}_t + (\gamma_z - 1 + \delta)\,k_y\widehat{I}_t + r^k k_y z_1\widehat{r}_t \quad \text{(A.3.9)}$$

Note that Smets and Wouters (2007) normalize spending shock as $\epsilon_{g,t} \equiv g_y \widehat{G}_t$.

$\cdot$ $z_t K_{t-1}/L_t = \frac{\alpha}{1-\alpha} w_t/r_t$.

We have

$$\widehat{r}_t + \widehat{z}_t + \widehat{K}_{t-1} - \widehat{w}_t - \widehat{L}_t = 0.$$

.Using $\widehat{z}_t = z_1 \widehat{r}_t$, we obtain

$$(1 + z_1)\,\widehat{r}_t + \widehat{K}_{t-1} - \widehat{w}_t - \widehat{L}_t = 0. \tag{A.3.10}$$

$\cdot$ $Y_t = \widetilde{\mu}_t/\overline{\mu}_t\,(s_t)\left(A_t(z_t K_{t-1})^{\alpha} L_t^{1-\alpha} - A_t\phi\right)$

Aggregating individual productions, we obtain

$$\widehat{Y}_t = \frac{Y+\phi}{Y}\widehat{A}_t + \alpha\frac{Y+\phi}{Y}\left(\widehat{K}_{t-1} + z_1\widehat{r}_t\right) + (1-\alpha)\frac{Y+\phi}{Y}\widehat{L}_t. \tag{A.3.11}$$

A.3.1. Final system

$i_y = (\gamma_z - 1 + \delta)\,k_y$, $c_y = (1 - i_y - g_y)$, $r_y = r^k k_y z_1$,

$c_1 = [h/\gamma_z]\,/\,[1 + h/\gamma_z]\,$, $c_2 = [(\sigma_c - 1)\,wl/((1+\lambda_w)c)]\,/\,[\sigma_c\,(1 + h/\gamma_z)]$,
$c_3 = [1 - h/\gamma_z]\,/\,[\sigma_c\,(1 + h/\gamma_z)]\,$,

$i_1 = \left[1/1 + \beta\,(\gamma_z)^{1-\sigma_c}\right]$, $i_2 = \left[1/\left(1 + \beta\,(\gamma_z)^{1-\sigma_c}\right)(\gamma_z)^2\,\varphi\right]$,

$q_1 = [1-\delta]\,/\left[1 - \delta + r^k\right]$,

$k_1 = (1-\delta)\,/\gamma_z$, $k_2 = (1 - (1-\delta)\,/\gamma_z)\left(1 + \beta\,(\gamma_z)^{1-\sigma_c}\right)(\gamma_z)^2\,\varphi$,

$\pi_1 = 1/\left[(1 + \beta\,(\gamma_z)^{1-\sigma}\,\gamma_p)\right]$, $\pi_2 = \left[\beta\,(\gamma_z)^{1-\sigma}\right]/\left[(1 + \beta\,(\gamma_z)^{1-\sigma}\,\gamma_p)\right]$,

$$\pi_3 = \left[\left(1-\xi_p\right)\left(1-\beta\left(\gamma_z\right)^{1-\sigma}\xi_p\right)\right] / \left[\xi_p(1+\beta\left(\gamma_z\right)^{1-\sigma}\gamma_p)\left(\left(\mu_p-1\right)\varsigma_p+1\right)\right],$$

$$w_1 = 1/\left[1+\beta\left(\gamma_z\right)^{1-\sigma_c}\right],\ w_2 = \left[1+\beta\left(\gamma_z\right)^{1-\sigma_c}\gamma_w\right] / \left[1+\beta\left(\gamma_z\right)^{1-\sigma_c}\right],$$
$$w_4 = \left[(1-\xi_w)\left(1-\beta\left(\gamma_z\right)^{1-\sigma_c}\xi_w\right)\right] / \left[\xi_w\left(1+\beta\left(\gamma_z\right)^{1-\sigma_c}\right)\left(\left(\mu_w-1\right)\varsigma_w+1\right)\right].$$

We have:

$$\widehat{Y}_t - c_y\widehat{C}_t - i_y\widehat{I}_t - r_y\widehat{r}_t - \widehat{\varepsilon}_t^g = 0, \qquad \text{(Eq. 1)}$$

$$\begin{aligned}\widehat{C}_t &= c_1\widehat{C}_{t-1} + (1-c_1)\,E_t\widehat{C}_{t+1} + c_2\left[\widehat{L}_t - E_t\widehat{L}_{t+1}\right] \\ &\quad -c_3\left[\widehat{R}_t - E_t\widehat{\pi}_{t+1} + \widehat{\varepsilon}_t^b\right], \qquad \text{(Eq. 2)}\end{aligned}$$

$$\widehat{I}_t = i_1\widehat{I}_{t-1} + (1-i_1)\,E_t\widehat{I}_{t+1} + i_2\widehat{Q}_t + \widehat{\varepsilon}_t^i, \qquad \text{(Eq. 3)}$$

$$\widehat{Q}_t = q_1E_t\widehat{Q}_{t+1} + (1-q_1)\,E_t\widehat{r}_{t+1} - \left(\widehat{R}_t - E_t\widehat{\pi}_{t+1} + \widehat{\varepsilon}_t^b\right), \qquad \text{(Eq. 4)}$$

$$\widehat{Y}_t = \left(1+\frac{\phi}{Y}\right)\left(\alpha\left(\widehat{K}_{t-1} + z_1\widehat{r}_t\right) + (1-\alpha)\widehat{L}_t + \widehat{\varepsilon}_t^A\right), \qquad \text{(Eq. 5)}$$

$$\widehat{K}_t = k_1\widehat{K}_{t-1} + (1-k_1)\,\widehat{I}_t + k_2\widehat{\varepsilon}_t^i, \qquad \text{(Eq. 6)}$$

$$\widehat{\mu}_t^p = \alpha\left(\widehat{K}_{t-1} + z_1\widehat{r}_t - \widehat{L}_t\right) - \widehat{w}_t + \widehat{\varepsilon}_t^A, \qquad \text{(Eq. 7)}$$

$$\widehat{\pi}_t = \pi_1 \left(\gamma_p \widehat{\pi}_{t-1} + \left(1-\gamma_p\right) \widehat{\pi}_t^*\right) + \pi_2 \left(E_t \widehat{\pi}_{t+1} - \left(1-\gamma_p\right) E_t \widehat{\pi}_{t+1}^*\right) - \pi_3 \widehat{\mu}_t^p + \widehat{\varepsilon}_t^p, \tag{Eq. 8}$$

$$(1+z_1)\,\widehat{r}_t + \widehat{K}_{t-1} - \widehat{w}_t - \widehat{L}_t = 0, \tag{Eq. 9}$$

$$\widehat{\mu}_t^w = \widehat{w}_t - \sigma_L \widehat{L}_t - \frac{1}{1-h/\gamma_z}\widehat{C}_t + \frac{h/\gamma_z}{1-h/\gamma_z}\widehat{C}_{t-1}, \tag{Eq. 10}$$

$$\begin{aligned} \widehat{w}_t \;=\; & w_1 \left(\widehat{w}_{t-1} + \gamma_w \widehat{\pi}_{t-1} + (1-\gamma_w)\widehat{\pi}_t^*\right) \\ & + (1-w_1)\left(E_t \widehat{w}_{t+1} + E_t \widehat{\pi}_{t+1} - (1-\gamma_w) E_t \widehat{\pi}_{t+1}^*\right) \\ & - w_2 \widehat{\pi}_t - w_4 \widehat{\mu}_t^w + \widehat{\varepsilon}_t^w, \end{aligned} \tag{Eq. 11}$$

$$\begin{aligned} \widehat{R}_t \;=\; & \rho_r(s_t)\widehat{R}_{t-1} + (1-\rho_r(s_t))\left[r_\pi(s_t)\left(\widehat{\pi}_t - \widehat{\pi}_t^*\right) + r_y(s_t)\left(\widehat{Y}_t - \widehat{Y}_t^f\right)\right] \\ & + r_{\Delta y}(s_t)\left[\left(\widehat{Y}_t - \widehat{Y}_t^f\right) - \left(\widehat{Y}_{t-1} - \widehat{Y}_{t-1}^f\right)\right] + \widehat{\varepsilon}_t^r, \end{aligned} \tag{Eq. 12}$$

where $\widehat{Y}_t^f$ is potential output, defined by the same model but assuming flexible prices and wages.

Bibliography

Altug, S. (1989) "Time-to-build and aggregate fluctuations: some new evidence," *International Economic Review*, pp. 889–920.

An, S. and F. Schorfheide (2007) "Bayesian analysis of DSGE models," *Econometric Reviews*, Vol. 26, No. 2-4, pp. 113 – 172.

Anderson, T. W. and H. Rubin (1949) "Estimation of the parameters of a single equation in a complete system of stochastic equations," *The Annals of Mathematical Statistics*, Vol. 20, pp. 46–63.

Ang, A., G. Bekaert, and M. Wei (2007) "Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better?," *Journal of Monetary Economics*, Vol. 54, No. 4, pp. 1163–1212.

Aruoba, S. B. and F. Schorfheide (2011) "Sticky Prices versus Monetary Frictions: An Estimation of Policy Trade-Offs," *American Economic Journal: Macroeconomics*, Vol. 3, No. 1, pp. 60–90.

Bernanke, B. (2004) "The Great Moderation, remarks at the meetings of the Eastern Economic Association," *Washington, DC, February*, Vol. 20.

Bianchi, F. (2011) "Regime Switches, Agents' Beliefs, and Post-World War II US Macroeconomic Dynamics," *Forthcoming in Review of Economic Studies*.

Bils, M. and P. Klenow (2004) "Some evidence on the importance of price stickiness," *Journal of Political Economy*, Vol. 112, No. 5, pp. 947–986.

Blanchard, J. O. and C. M. Kahn (1980) "The Solution of Linear Difference Models under Rational Expectations," *Econometrica*, Vol. 48, pp. 1305–1312.

Blanchard, J. O. and J. Simon (2001) "The long and large decline in US output volatility," *Brookings Papers on Economic Activity*, Vol. 2001, No. 1, pp. 135–174.

Brooks, S. P. and A. Gelman (1998) "General methods for monitoring convergence of iterative simulations," *Journal of Computational and Graphical Statistics*, pp. 434–455.

Calvo, G. (1983) "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, Vol. 12, No. 3, pp. 383–98.

Carlson, J. A. and M. Parkin (1975) "Inflation Expectations," *Economica*, Vol. 42, No. 166, pp. 123–138.

Chib, S. and S. Ramamurthy (2010) "Tailored randomized block MCMC methods with application to DSGE models," *Journal of Econometrics*, Vol. 155, No. 1, pp. 19–38.

Christiano, L. J. and M. Eichenbaum (1992) "Current Real-Business-Cycle Theories and Aggregate Labor-Market Fluctuations," *American Economic Review*, Vol. 82, No. 3, pp. 430–50.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (1999) "Monetary policy shocks: what have we learned and to what end?," *Handbook of Macroeconomics*, Vol. 1, p. 65.

——— (2005) "Nominal rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, Vol. 113, No. 1, pp. 1–45.

Clarida, R., Galí J., and M. Gertler (2000) "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, Vol. 115, No. 1, pp. 147–180.

Cogley, T. and A. M. Sbordone (2008) "Trend inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve," *American Economic Review*, Vol. 98, No. 5, pp. 2101–2126.

Cooley, T. F. and G. D. Hansen (1995) "Money and the business cycle," in T. F. Cooley ed. *Frontiers in Business Cycle research*: Princeton University Press.

Davig, T. and T. Doh (2009) "Monetary policy regime shifts and inflation persistence," Research Working Paper, Federal Reserve Bank of Kansas City.

Davig, T. and E. M. Leeper (2007) "Generalizing the Taylor Principle," *American Economic Review*, Vol. 97, No. 3, pp. 607–635.

——— (2008) "Endogenous monetary policy regime change," in *NBER International Seminar on Macroeconomics 2006*, pp. 345–391, University of Chicago Press.

DeJong, D. N. and C. Dave (2007) *Structural Macroeconometrics*: Princeton University Press.

Del Negro, Marco and F. Schorfheide (2010) "Bayesian Macroeconometrics," in J. Geweke, G. Koop, and H.V. Dijk eds. *Hanbook of Bayesian Econometrics*: Oxford University Press.

Dufour, J. M., L. Khalaf, and M. Kichian (2010) "On the precision of Calvo parameter estimates in structural NKPC models," *Journal of Economic Dynamics and Control*, Vol. 34, No. 9, pp. 1582–1595.

Dufourt, F. (2011) "Solving Markov-Switching models with a fast iterative algorithm."

Dufourt, F. and F. Jean-Baptiste (2011a) "Are monetary policy regime switches endogenous? Evidence for Euro Area countries,"Technical report, Bureau d'Économie Théorique et Appliquée, Université de Strasbourg.

——— (2011b) "Good luck or Good Policy? Re-assessing the Great Moderation through an estimated Markov-Switching DSGE model.,"Technical report, Bureau d'Économie Théorique et Appliquée, Université de Strasbourg.

Fagan, G., J. Henry, and R. Mestre (2005) "An Area-Wide Model for the Euro Area," *Economic Modelling*, Vol. 22, No. 1, pp. 39–59.

Farmer, R. E. A., D. F. Waggoner, and T. Zha (2008) "Minimal state variable solutions to Markov-switching rational expectations models," *Federal Reserve Bank of Atlanta Working Papers*.

——— (2009a) "Indeterminacy in a forward-looking regime switching model," *International Journal of Economic Theory*, Vol. 5, No. 1, pp. 69–84.

——— (2009b) "Understanding Markov-switching rational expectations models," *Journal of Economic Theory*, Vol. 144, No. 5, pp. 1849–1867.

Farmer, R. E. A, D. F Waggoner, and T. Zha (2010) "Minimal state variable solutions to markov-switching rational expectations models," *Forthcoming in Journal of Economic Dynamics and Control*.

Fernández-Villaverde, J. (2010) "The econometrics of DSGE models," *Journal of the Spanish Economic Association*, Vol. 1, No. 1, pp. 3–49.

Fernández-Villaverde, J., P. A. Guerrón-Quintana, and J. Rubio-Ramirez (2010) "Fortune or virtue: Time-variant volatilities versus parameter drifting in us data."

Fischer, S. (1977) "Long-term contracts, rational expectations, and the optimal money supply rule," *The Journal of Political Economy*, Vol. 85, pp. 191–205.

Galí, J. (2008) *Monetary Policy, inflation, and the Business Cycle: An introduction to the new Keynesian Framework*: Princeton Univ Press.

Galí, J. and M. Gertler (1999) "Inflation Dynamics: A Structural Econometric Analysis," *Journal of monetary Economics*, Vol. 44, No. 2, pp. 195–222.

Galí, J., M. Gertler, and J. David Lopez-Salido (2005) "Robustness of the Estimates of the Hybrid New Keynesian Phillips Curve," *Journal of Monetary Economics*, Vol. 52, No. 6, pp. 1107–1118.

Galí, J., F. Smets, and R. Wouters (2011) "Unemployment in an estimated new Keynesian model," Economics Working Papers 1266.

Gertler, M. and P. Karadi (2011) "A model of unconventional monetary policy," *Journal of Monetary Economics*, Vol. 58, No. 1, pp. 17–34.

Geweke, J. (1992) "Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments," in J. Bernado, Berger J., Dawid A., and Smith A. eds. *Bayesian statistics 4*: Oxford: Calendron Press, pp. 641–649.

——— (1999) "Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication," *Econometric Reviews*, Vol. 18, No. 1, pp. 1–73.

Hamilton, J. D. (1994) *Time series analysis*: Cambridge University Press.

Hansen, L. P. (1982) "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, Vol. 50, No. 4, pp. 1029–54.

Hansen, L. P., J. Heaton, and A. Yaron (1996) "Finite-sample properties of some alternative GMM estimators," *Journal of Business & Economic Statistics*, Vol. 14, No. 3, pp. 262–280.

Hansen, L. P. and K. J. Singleton (1982) "Generalized instrumental variables estimation of nonlinear rational expectations models," *Econometrica*, Vol. 50, pp. 1269–1286.

Henzel, S. and T. Wollmershaeuser (2008) "The New Keynesian Phillips Curve and the Role of Expectations: Evidence from the CESifo World Economic Survey," *Economic Modelling*, Vol. 25, No. 5, pp. 811–832.

Ireland, P. N. (2007) "Changes in the Federal Reserve's inflation target: Causes and consequences," *Journal of Money, Credit and Banking*, Vol. 39, No. 8, pp. 1851–1882.

Jean-Baptiste, F. (2011) "Forecasting with the New Keynesian Phillips curve: Evidence from survey data," *Forthcoming in Economics Letters*.

Justiniano, A. and G. E. Primiceri (2008) "The time-varying volatility of macroeconomic fluctuations," *American Economic Review*, Vol. 98, No. 3, pp. 604–641.

Justiniano, A., G. E. Primiceri, and A. Tambalotti (2010) "Investment shocks and business cycles," *Journal of Monetary Economics*, Vol. 57, No. 2, pp. 132–145.

Justiniano, A., G. E Primiceri, and A. Tambalotti (2011) "Investment Shocks and the Relative Price of Investment," *Review of Economic Dynamics*, Vol. 14, No. 1, pp. 101–121.

Kim, C. J. (1994) "Dynamic linear models with Markov-switching," *Journal of Econometrics*, Vol. 60, No. 1-2, pp. 1–22.

Kim, C. J. and C. R. Nelson (1999) "State-space Models with Regime Switching: Classical and Gibbs-sampling Approaches with Applications," *MIT Press*.

King, R. G., C. I. Plosser, and S. T. Rebelo (1988) "Production, growth and business cycles:: I. The basic neoclassical model," *Journal of monetary Economics*, Vol. 21, No. 2-3, pp. 195–232.

Kleibergen, F. (2005) "Testing parameters in GMM without assuming that they are identified," *Econometrica*, Vol. 73, No. 4, pp. 1103–1123.

Kleibergen, F. and S. Mavroeidis (2009) "Weak instrument robust tests in GMM and the new Keynesian Phillips curve," *Journal of Business and Economic Statistics*, Vol. 27, No. 3, pp. 293–311.

Kydland, F. E. and E. C. Prescott (1982) "Time to build and aggregate fluctuations," *Econometrica*, Vol. 50, pp. 1345–1370.

Leeper, E. and C. Sims (1994) "Toward a modern macro model usable for policy analysis," *NBER Macroeconomics Annual*, pp. 81–117.

Liu, P. and H. Mumtaz (2010) "Evolving macroeconomic dynamics in a small open economy: an estimated Markov-switching DSGE model for the United Kingdom," *Bank of England working papers*.

Liu, Z., D. F. Waggoner, and T. Zha (2011) "Sources of Macroeconomic Fluctuations: A Regime-Switching DSGE Approach," *Quantitative Economics*, Vol. 2, No. 2, pp. 251–301.

Long, J. B. and C. I. Plosser (1983) "Real business cycles," *The Journal of Political Economy*, pp. 39–69.

Lubik, T. A. and F. Schorfheide (2004) "Testing for indeterminacy: an application to US monetary policy," *American Economic Review*, Vol. 94, No. 1, pp. 190–217.

Lucas, R. E. (1972) "Expectations and the Neutrality of Money," *Journal of Economic Theory*, Vol. 4, No. 2, pp. 103–124.

——— (1975) "An equilibrium model of the business cycle," *The Journal of Political Economy*, pp. 1113–1144.

Ma, A. (2002) "GMM estimation of the new Phillips curve," *Economics Letters*, Vol. 76, No. 3, pp. 411–417.

Magnusson, L. M. and S. Mavroeidis (2010) "Identification-Robust Minimum Distance Estimation of the New Keynesian Phillips Curve," *Journal of Money, Credit and Banking*, Vol. 42, No. 2-3, pp. 465–481.

Nason, J. M. and G. W. Smith (2008) "The New Keynesian Phillips Curve: Lessons from Single-Equation Econometric Estimation," *FRB Richmond Economic Quarterly*, Vol. 94, No. 4, pp. 361–395.

Newey, W. K. and K. D. West (1987) "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica*, Vol. 55, pp. 703–708.

Nielsen, H. (2003) "Inflation Expectations in the EU," *National Research Center 373, Discussion Paper No. 13.*

Nunes, R. (2010) "Inflation Dynamics: The role of expectations," *Journal of Money, Credit and Banking*, Vol. 42, No. 6, pp. 1161–1172.

Rogoff, K. (1985) "The Optimal Degree of Commitment to an Intermediate Monetary Target," *The Quarterly Journal of Economics*, Vol. 100, No. 4, pp. 1169–89.

Rubio-Ramirez, J., D. F. Waggoner, and T. Zha (2005) "Markov-switching structural vector autoregressions: Theory and application," *FRB of Atlanta Working Paper No. 2005-27.*

Rudd, J. and K. Whelan (2005) "Does Labor's Share Drive Inflation?" *Journal of Money, Credit and Banking*, Vol. 37, No. 2, pp. 297–312.

Sbordone, A. M. (2002) "Prices and unit labor costs: a new test of price stickiness," *Journal of Monetary Economics*, Vol. 49, No. 2, pp. 265–292.

——— (2005) "Do expected future marginal costs drive inflation dynamics?," *Journal of Monetary Economics*, Vol. 52, No. 6, pp. 1183–1197.

Schorfheide, F. (2000) "Loss function-based evaluation of DSGE models," *Journal of Applied Econometrics*, Vol. 15, No. 6, pp. 645–670.

——— (2005) "Learning and monetary policy shifts," *Review of Economic Dynamics*, Vol. 8, No. 2, pp. 392–419.

——— (2006) , Lecture notes for Estimation and Evaluation of DSGE models.

Sims, C. A. (2002) "Solving linear rational expectations models," *Computational Economics*, Vol. 20, No. 1, pp. 1–20.

Sims, C. A., D. F. Waggoner, and T. Zha (2008) "Methods for inference in large multiple-equation Markov-switching models," *Journal of Econometrics*, Vol. 146, No. 2, pp. 255–274.

Sims, C.A. and T. Zha (2006) "Were there regime switches in US monetary policy?," *American Economic Review*, Vol. 96, No. 1, pp. 54–81.

Smets, F. and R. Wouters (2003) "An estimated dynamic stochastic general equilibrium model of the euro area," *Journal of the European Economic Association*, Vol. 1, No. 5, pp. 1123–1175.

——— (2005) "Comparing shocks and frictions in US and euro area business cycles: a Bayesian DSGE approach," *Journal of Applied Econometrics*, Vol. 20, No. 2, pp. 161–183.

——— (2007) "Shocks and Frictions in US Business Cycles: a Bayesian DSGE Approach," *American Economic Review*, Vol. 97, No. 3, pp. 586–606.

Stock, J. H. and M. W. Watson (1999) "Forecasting inflation," *Journal of Monetary Economics*, Vol. 44, No. 2, pp. 293–335.

——— (2003a) "Forecasting Output and Inflation: The role of Asset prices," *Journal of Economic Literature*, pp. 788–829.

——— (2003b) "Has the Business Cycle Changed? Evidence and Explanations," in *Monetary Policy and Uncertainty: Adapting to a Changing Economy*, Jackson Hole, Wyoming, pp. 9–56.

Svensson, L.E.O. and N. Williams (2007) "Monetary policy with model uncertainty: distribution forecast targeting," *NBER working paper*.

Taylor, J. B. (1979) "Staggered wage setting in a macro model," *American Economic Review*, Vol. 69, No. 2, pp. 108–113.

——— (1993) "Discretion versus policy rules in practice," in *Carnegie-Rochester conference series on public policy*, Vol. 39, pp. 195–214, Elsevier.

Uhlig, H. (1997) "A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily," in Ramón Marimón and Andrew Scott eds. *Computational Methods for the Study of Dynamic Economies*: Oxford University Press.

Woodford, M. (2003) *Interest and prices: Foundations of a theory of monetary policy*: Princeton Univ Press.