Essay on Bayesian Estimation of DSGE models

Filippo Ferroni

TESI DOCTORAL UPF / 2009

DIRECTOR DE LA TESI

Prof. Fabio Canova (Departament d'Economia i Empresa)

UNIVERSITAT POMPEU FABRA

To My Family

Acknowledgments

I am indebted to my advisor, Fabio Canova, for his invaluable support and many discussions. During these years he guided me from the early stage of chaos and confusion to this final version. It a pleasure to express my appreciation for shaping my views about macro models and econometrics.
During my stay in Pompeu I also received comments, critiques and genuine encouragement from the following people: Jordi Gali, Thijs Van Rens, Vasco Carvalho, Kristoffer Nimark, Toni Braun, Albert Marcet and Michael Reiter.
Sincere thanks go to all of them.
I acknowledge Arturo Ormeño and Francesco Caprioli for many valuable conversations and discussions. I also would like to thank Marta Araque, Marta Aragay e Gemma Burballa for their constant support and precious help.

I owe my deepest gratitude to my parents and to my brother for their love and patience. Without them, all of this would not be possible.

Finally, for the good times spent together I feel to acknowledge also Basak, Zeynep, Mar, Andrea, Javi, Peter, Bea, Nacho, Shak, my office mates, the GPEFM Devils, la Base Nautica. There are many other people that I wish to thank, but the number would be too large to thank them individually. To make up for all of them, I thank Barcelona for making my stay intense and unforgettable.

Abstract

This thesis examines three different policy experiments using Bayesian estimates of DSGE models. First, we show that countercyclical fiscal policies are important to smooth fluctuations and that this is true regardless of how we specify the fiscal rule and several details of the model. Second, we show that the sources of output volatility obtained from a cyclical DSGE model crucially depend on whether estimation is done sequentially or jointly. In fact, while with a two step procedure, where the trend is first removed, nominal shocks drive output volatility, investment shocks dominate when structural and trend parameters are estimated jointly. Finally, we examine the role of money for business cycle fluctuations with a single and a multiple filtering approach, where information provided by different filters is jointly used to estimate DSGE parameters. In the former case, money has a marginal role for output and inflation fluctuations, while in the latter case is important to transmit cyclical fluctuations.

Resumen

Esta tesis presenta tres diferentes experimentos de política utilizando estimaciones Bayesianas de modelos DSGE. En la primera parte, se quiere demostrar que una política fiscal contracíclica es un instrumento importante para alcanzar la estabilidad macroeconómica. Este resultado es robusto a diferentes controles. En la segunda parte, se demuestra en forma cualitativa y cuantitativa las variaciones de las estimaciones de los parámetros estructurales según la descomposición ciclo-tendencia, si en uno o en dos estadios. Resulta que con un procedimiento a dos estadios la volatilidad del PIB es explicada mayormente por shocks nominales, mientras que con un procedimiento a un estadio por un shock a la inversión. Se argumenta que el procedimiento a un estadio proporciona una estructura probabilística más coherente. La tercera parte de la tesis propone una manera de estimar los parámetros estructurales utilizando la información procedente de distintos filtros. Mientras que con un tipo de estimación con un único filtro el dinero tiene poca influencia en las fluctuaciones de medio plazo, con un sistema de múltiples filtros el dinero tiene un papel importante en la transmisión de los shocks.

Foreword

Over the last 30 years, macroeconomic theory and applied macroeconomic research have changed dramatically, and there is a widespread consensus that these changes have led to a better understanding of economic fluctuations. While not confined to ivory tower, these improvements spilled over to policy-oriented institutions influencing central banks and government decisions.
To guide macroeconomic research one of the most influential works was the Lucas (1978) critique: Lucas argued that technology and preference are invariant to changes in policies, while decision rules of private agents are not. With a series of graphical examples, he showed that models that consider time invariant decision bring to undesirable policy implications. This motivated macroeconomic researchers to develop quantitative Dynamic Stochastic General Equilibrium (DSGE) models, in principle less vulnerable to the Lucas critique, where agents' decisions depend explicitly on policy choices and parameters of technology and preferences are argued and assumed to be invariant to policy.
Parallel advancements on statistical grounds came along to bridge DSGE models and reality. In this context, parameters and model uncertainty have been taken seriously and progresses in computer technology have allowed to process computationally intensive statistical problems. In this framework, Bayesian techniques have gained a predominant role for the DSGE models estimation, representing ideally the toolkit of every applied researcher. There are several reasons for that. First, the Bayesian paradigm provides a coherent framework to treat model uncertainty and to take decisions based on risk. Second, if a vector of times series is predicted by a function of its past values, then the resulting forecast error covariance matrix is non-singular. Hence, any DSGE model that generates a rank deficient covariance matrix is clearly at odds with the data. The Bayesian setup allows to avoid these problems by introducing so-called measurement errors. Moreover, while not treated as the 'true' data generating process, DSGE models are just considered as an approximation of the law of motion for the data. Third, Bayesian estimators have desirable property in small samples. Fourth, priors allow to introduce external information to the model and to consistently combine pre-sample information with the observed data.
The present thesis explores to what extent Bayesian estimates of DSGE models can be used to do policy analysis. In particular, we investigate to which degree point estimates for the DSGE model parameters are able to seize macroeconomic fluctuation and eventually be useful to draw policy implications.
A first policy experiment is conducted on tax policy. Among policy makers considerable interest has emerged to assess whether the tax policy in US has helped to reduce the volatility of GDP at business cycles frequency. A general equilibrium model with distorting taxation on household income is considered, where tax policies follow a 'simple' rule responding to the cyclical conditions of the economy, i.e. changes in employment

and output. Deep parameters governing preferences and technology are estimated along with the fiscal policy parameters using Post World War II US data. Results show that tax policies have shown a strong countercyclical patter, meaning that in bad (good) times taxes have been lower (higher). Counterfactual exercises show that the keynesian prescription of countercyclical fiscal policy is indeed a valuable instrument to reduce the volatility of aggregate consumption, investment and output. Several robustness checks are conducted to assess the validity of the exercise. Results look convincing and consistent in light of other related works with less structural approaches, such as Romer and Romer (2007) Cohen and Follette (2000), among others.

The following policy experiments are motivated by the relative silence of the treatment of trends in the literature of Bayesian DSGE estimation. While much effort has been devoted to reduce potential misspecifications on the cyclical equations by adding more frictions and shocks, little work has been done on developing coherent frameworks to handle trends in the data. In a seminal paper, Cogley (2001) showed that wrong trend specifications are likely to affect the estimates of technology and preference parameters, corrupting the policy conclusions to be drawn from the structural model. To address this issue, two different policy experiments are conducted.

First, we show how structural estimates might change if we detrend the data in a first stage and then estimate the structural parameters in a second stage or if we jointly estimate structural and trend parameters. The two approaches lead to different conclusions from the policy standpoint. In particular, the sources of GDP volatility change substantially if we take a two or one stage stand on the treatment of the trend. In fact, with a two step procedure we obtain that the mayor driving forces of output volatility are nominal shocks (i.e. price and wage markups), and with the joint estimate of structural and trend parameters the most likely contribution to the GDP volatility is a real shock, the investment specific technology shock. We show that the latter provides a more coherent probabilistic setup.

Second, in a joint work with Fabio Canova, we propose a method that allows to simultaneously process multiple information provided by a variety of filters and estimate the DSGE parameters. Relative to standard one filter estimates, estimates of DSGE parameters result more consistent and the Mean Square Error is smaller. Moreover, the multiple filtering approach produces not only statistically different results, but also divergent economic conclusions. In particular, Ireland (2004) has shown that for the US cycles the role of money stock is marginal in a likelihood based approach with linear detrended data. While with only one data transformation the role of money transmission is negligible, with the multiple filtering approach real balances statistically matter for the transmission of cyclical fluctuations to output and inflation.

In sum, the thesis presents three policy experiments using Bayesian estimates of DSGE models. First, we show that countercyclical fiscal policies are important to smooth fluctuations regardless of the exact fiscal rule specification. Second, we show that sources of output volatility obtained from a cyclical DSGE model crucially depend on whether estimation is done sequentially or jointly. In fact, while with a two step procedure nominal shocks drive output volatility, investment shocks dominate when structural and trend parameters are estimated jointly. Finally, we examine the role of money for business cycle fluctuations with a single and a multiple filtering approach. In the former case, money has a marginal role for output and inflation fluctuations, while in the latter case is important to transmit cyclical fluctuations.

The thesis is organized as follows: the first chapter presents the first policy experiment on tax policy. The second chapter proposes the agnostic one step procedure for the trend-cycles decomposition and the GDP variance decomposition exercise is presented at the end of the chapter. The third chapter presents the multiple filtering approach to estimate cyclical DSGE models; at the end of the chapter implications in terms of the role of money are considered.

Contents

1 Did Tax Policies mitigate US Business Cycles ?

1.1 Introduction

Can fiscal adjustment change the features of the Business Cycle ? Is it true that countercyclical fiscal policy helps to smooth fluctuations ? These topics have recently been at the center of public debate. A recent US Treasury Department study concluded that the absence of automatic stabilizers at the peak of the last US recession in 2001 would have added an additional 1.5 million people to the ranks of unemployed. In Europe, because of the creation of a single currency area and the disappearance of national monetary policy, the debate has focused on the role of national fiscal policy and the nature of the Stability and Growth Pact; Gali and Perotti (2004) ask whether after the Maastricht Treaty national fiscal policies have become less countercyclical, and they found no evidence to support this view. At the core of this debate there is the keynesian prescription that countercyclical fiscal policy has stabilizing effects that work through both automatic stabilizers and occasionally discretionary actions; many economists share this view[1]. At the opposite end, some recent studies question the stabilizing role of fiscal policies: Jones (2002) shows empirically that post war fiscal policy did not help the US economy to smooth Business Cycle fluctuations. From a theoretical standpoint, Gordon and Leeper (2005) highlight that countercyclical fiscal policy might amplify recessions through the policy expectations channel.
In this chapter, I revisit the issue of whether fiscal policy matters for business cycle fluctuations and ask whether US tax policy has been an important source for economic volatility, and (if so) which tax instrument is more important for controlling economic fluctuations. I assume that taxes respond to cyclical conditions of the economy by reacting to employment and GDP variations. I estimate the deep parameters of the model from a vector of time series using Bayesian techniques. I find that, consistent with the less structural analysis of Romer and Romer (2007) and Cohen and Follette (2000), US tax policies displayed a strong procyclical reaction to GDP and employment in the period considered. I find also that the countercyclical behavior of taxes is an important stabilizing device; indeed, the automatic response of the labor and capital tax to cyclical conditions can reduce the volatility of consumption, GDP, investment and hours worked. In particular, in the absence of the automatic response to cyclical conditions consumption volatility would increase by 28%, GDP volatility by 8%, employment by 28% and investment by 36%. Moreover, I find that labor tax can help to smooth BC more than capital tax. Finally, unexpected changes in taxes generate little economic volatility; indeed, I find that unexpected changes in tax policies accounts for less than 4% the volatility of GDP and consumption. The latter finding is consistent with what is found in the narrative approach of Romer and Romer (2007). They find

[1] Blinder and Solow (1973), Romer and Romer (1994), Romer (1999), Cohen and Follette (2000), Auerbach (2003).

that fiscal shocks explain 9% of GDP growth rate volatility.
For the purpose of this chapter, one of the most delicate issues is the definition of a policy rule that summarizes the evolution of tax over time. As explained in Gali and Perotti (2004), the fiscal response function can be seen as the combination of a 'cyclical' component and a 'structural' or 'discretionary' component; the first part comprises all the variations outside the direct control of the fiscal authority (like changes in the tax base). The second part should be interpreted as the part which is intentionally chosen by the policymaker, as in Fatás and Mihov (2001). Within this discretionary part, there is an 'endogenous' or 'systematic' response by which the policymaker automatically responds to cyclical economic conditions and there is a 'non-systematic' or 'exogenous' component. The former component arises from spending programs and tax cut that adjust systematically with economic conditions. The latter captures all the changes that do not correspond to systematic variations to cyclical conditions; we can interpret these exogenous changes as actions that are meant to sustain or fasten long run growth (Romer and Romer (2007)), or changes in the political process (Gali and Perotti (2004)). This work attempts to estimate these two components from the data.
There are different fiscal instruments, that could be taken into consideration. For instance, Fatas and Mihov (2006) consider government spending. Gali and Perotti (2004) or Auerbach (2003) instead consider the primary deficit. Jones (2002) considers average tax rates and government spending. In this work I focus on tax policies. Therefore, I deliberately ignore government spending as a fiscal instrument. The reason is that government expenditure is rather inflexible and while it can be easily increased it is very difficult to decrease it. Therefore, its role as a stabilization tool is doubtful. Another major limitation of the present analysis is that I am ignoring possible interactions between fiscal and monetary policy, see Canzoneri, Cumby and Diba (2002) or Taylor (2000). Thus, a possible extension to this work is to introduce a monetary authority and a monetary rule.
Several empirical studies have used VAR techniques to address the issue of interest in this chapter; Mountford and Uhlig (2005) studies the transmission mechanism of fiscal shocks. Canova and Pappa (2007) and Perotti (2002) shows that government spending has a significant output multiplier. Canova and Pappa (2006) look at the transmission properties of fiscal disturbances under different fiscal restrictions on government budget constraint.
There are endogenous feedbacks between economic activity and tax policies; on the one hand, taxes directly affect household consumption and labor decisions, and therefore economic activity. On the other hand, the fiscal authority sets tax policies by looking at the economic activity. Given the strong endogenous relations between household decisions and tax policy, I choose to employ a general equilibrium framework. There are several papers that look at fiscal policies in a general equilibrium framework. McGrattan (1994) estimates the fiscal response from a vector of autoregression considering a broad general class of fiscal responses; she concludes that a relevant portion of busi-

ness cycles fluctuations is due to fiscal instruments. Braun (1994) estimates in a GE model the reaction function of fiscal instruments, taxes and government spending; Jones (2002) estimates various fiscal policy rules form the US postwar data. The main limitation of their analysis is that they assume that the government runs a balanced budget constraint and that implicity debt plays no role. Giammarioli, Lambertini and Onorante (2007) measure the discrepancy of the US fiscal policy from the optimal fiscal policy derived from a Ramsey plan and find that the US fiscal policy has been 'too countercyclical'. They employ a GE framework with no capital accumulation and a tax on labor. My analysis considers capital accumulation and tries to disentangle the effect of the taxes on labor and on capital income.
A variety of techniques are available to estimate General Equilibrium models ranging from calibration, to Generalized Method of Moments (GMM) estimation to full information likelihood-based methods. This work innovates relative to the fiscal policy literature by using Bayesian estimation. Bayesian techniques have become very popular recently; being a likelihood-based approach they exploit cross equation restrictions, but differently from 'pure' maximum likelihood methods they can incorporate information external to the model in terms of priors. Moreover, with recent development of simulation algorithms it has become much easier to estimate densely parameterized models (see Smets and Wouters (2003, 2005)).
The chapter is organized as follows: Section 1.2 presents the model. The estimation procedure is described in Section 1.3. Section 1.4 shows the results and Section 1.5 draws policy implications. Section 1.6 provides robustness analysis and Section 1.7 concludes the chapter.

1.2 Model

I employ a prototype RBC model with no frictions and with a time varying 'wedge' on labor and capital accumulation, and an efficiency 'wedge' on the production side. Chari, Kehoe and McGrattan (2007) have shown that this prototype economy is equivalent to a large class of models with various types of frictions and can reasonably well account for the U.S. postwar Business Cycle fluctuations.
The model consists of a single representative firm, a representative household, and a government. The firm and the household behave in the standard fashion; the firm maximizes profit, and the household maximizes its discounted lifetime utility. The government, on the other hand, does not have an objective function and it has to finance an exogenous expenditure process using distortionary taxation and issuing real debt.
I assume that all exogenous processes are stationary. Business cycle theory has little to say about trend and when it is assumed (via unit root in the technology process) it implies stationary great ratios. The data do not support this implication, see Figure 1

panel 1 and 3.

The numeraire is final output Y_t, which is produced by a representative price-taking firm. The firm faces a Cobb-Douglas production function:

$$Y_t = A_t K_{t-1}^{\alpha} N_t^{1-\alpha}, \tag{1.1}$$

where K_{t-1} and N_t denote the capital and labor available at time t, respectively. A_t is the exogenous stochastic technology process. I assume

$$a_t = \rho a_{t-1} + \epsilon_t^a,$$

where $a_t = \ln A_t$, and ϵ_t^a is an i.i.d. shock with zero mean and variance σ_a^2. Each period the firm solves

$$\max_{\{N_t, K_t\}_{t=0}^{\infty}} A_t K_{t-1}^{\alpha} N_t^{1-\alpha} - r_t K_{t-1} - w_t N_t$$

where r_t is the cost of renting capital and w_t is the real wage. The firm optimality conditions are

$$w_t = (1-\alpha)\frac{Y_t}{N_t}$$

$$r_t = \alpha \frac{Y_t}{K_{t-1}}.$$

The economy is populated by a single representative household; there is no population growth. The household maximization problem is

$$\max_{\{C_t, N_t, I_t\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{C_t^{1-\eta} - 1}{1-\eta} - X_t N_t\right] \tag{1.2}$$

where β is the time discount factor and $1/\eta$ is the intertemporal elasticity of substitution; C_t, N_t are respectively consumption and hours worked at time t. The utility function is separable in consumption and leisure (see Hansen (1985)). All the variables are expressed in per capita terms. X_t is an exogenous preference shock which evolves according to

$$\chi_t = \rho_\chi \chi_{t-1} + \epsilon_t^\chi$$

where $\chi_t = \ln X_t$ and ϵ_t^χ is an i.i.d. shock with zero mean and variance σ_χ^2. The representative household faces a budget constraint,

$$I_t + C_t + B_t = (1-\tau_t^w) w_t N_t + (1-\tau_t^k) r_t K_{t-1} + (1+r_t^b) B_{t-1} \tag{1.3}$$

I indicate with τ_t^w and τ_t^k the taxes on labor and capital income, respectively; B_{t-1} is the real debt issued by the government at time $t-1$ which gives a net interest rate of r_t^b. New installed capital is the combination of undepreciatiated capital, $(1-\delta)K_{t-1}$, where δ is the rate of depreciation of capital, and investment

$$K_t = V_t I_t + (1-\delta)K_{t-1} \tag{1.4}$$

and V_t is an investment specific shock which follows an AR(1) process, i.e.

$$v_t = \rho_v v_{t-1} + \epsilon_t^v$$

where $v_t = \ln V_t$, and ϵ_t^v is an i.i.d. shock with zero mean and variance σ_v^2. The representative household problem is to maximize (1.2) subject to the budget constraint (1.3) and (1.4); the first order conditions for the household problem are

$$C_t^\eta = (1-\tau_t^w)(1-\alpha)\frac{Y_t}{X_t N_t}, \tag{1.5}$$

$$1 = \beta E_t\{V_t(\frac{C_t}{C_{t+1}})^\eta R_{t+1}\} \tag{1.6}$$

$$R_{t+1} = (1-\tau_{t+1}^k)\alpha\frac{Y_{t+1}}{K_t} + \frac{1-\delta}{V_{t+1}}. \tag{1.7}$$

$$1 = \beta E_t\{(\frac{C_t}{C_{t+1}})^\eta(1+r_{t+1}^b)\} \tag{1.8}$$

where we have substituted in the optimality conditions from the firm problem. Equation (1.5) is the intra-temporal optimality condition between consumption and leisure; equation (1.6) is the usual Euler equation and R_t is net depreciation after tax interest rate. Equation (1.8) is the intertemporal optimality condition for debt demand. In equilibrium, it must be the case that the no arbitrage condition between the after tax interest rate and the bond interest rate holds:

$$V_{t-1}[(1-\tau_t^k)r_t + \frac{1-\delta}{V_t}] = 1 + r_t^b \tag{1.9}$$

The government satisfies a period by period budget constraint,

$$G_t + (1+r_t^b)B_{t-1} = \tau_t^w w_t N_t + \tau_t^k r_t K_{t-1} + B_t \tag{1.10}$$

where G_t is government spending. The literature considers several specifications for the fiscal policy instruments. Some authors consider government spending as an instrument that stimulates private consumption (see Galí, López-Salido and Vallés (2007)). My goal here is to study the ability of taxes to affect the Business Cycle independently of public spending. Therefore, I use the simplifying assumption that the government spending evolves as AR(1) process,

$$g_t = \rho_g g_{t-1} + \epsilon_t^g$$

where $g_t = \ln G_t$ and again ϵ_t^g is an i.i.d. shock with zero mean and variance σ_g^2.
As mentioned, the fiscal literature defines the fiscal rule as a combination of two main elements; an endogenous automatic response to the economic conditions, through which the policymaker reacts automatically to cyclical conditions, such as employment or GDP variations, and an exogenous component meant to be an unexpected reply to economic

cycles. In line with this literature, I assume that the tax deviation from its steady state responds to the GDP and employment log deviations from their respective steady states and to the debt-GDP deviation; the latter variable is included in order to control for the explosiveness of the government debt. It is also common to include lagged value of the taxes to account for sluggish reaction of the fiscal instrument. Rules governing tax policies[2] take the following form:

$$\widetilde{\tau}_t^w = \varphi_{by}\widetilde{by}_t + \varphi_w \widetilde{\tau}_{t-1}^w + \varphi_n n_t + \varphi_y y_t + \xi_t^w \tag{1.11}$$

$$\widetilde{\tau}_t^k = \psi_{by}\widetilde{by}_t + \psi_k \widetilde{\tau}_{t-1}^k + \psi_n n_t + \psi_y y_t + \xi_t^k \tag{1.12}$$

where $\widetilde{\tau}_t^j$ is the tax j in deviation from its steady state, i.e. $\widetilde{\tau}_t^j = \tau_t^j - \tau^j$, for $j = w, k$, $\widetilde{by}_t$ is the deviation of the debt-GDP ratio from its steady state, i.e. $\widetilde{by}_t = B_t/Y_t - B/Y$, n_t is the log deviation of hours worked from its steady state, i.e. $n_t = \ln \frac{N_t}{N} \simeq \frac{N_t - N}{N}$, and y_t is the log deviation of the GDP from its steady state, i.e.$y_t = \ln \frac{Y_t}{Y} \simeq \frac{Y_t - Y}{Y}$. ξ_t^j are i.i.d policy shocks with zero mean and variance $\sigma^2_{\xi^j}$ for $j = w, k$.
Finally, total output is absorbed by consumption, investment and government spending, i.e.

$$Y_t = I_t + C_t + G_t \tag{1.13}$$

1.3 Econometric Methodology

Equations (1.5)-(1.13) and the exogenous processes form a system of rational expectations equations. This system has to be solved before the model can be estimated. In the context of likelihood-based DSGE model estimation, log-linear approximation is very popular because it leads to a linear state space representation that can be estimated using the Kalman filter. Hence, following An and Schorfheide (2007), DeJong, Ingram and Whiteman (2000), Smets and Wouters (2003) or Jones (2002), I log linearize the equilibrium conditions around a pivotal point, the steady state. Then, the solution of the system takes the form

$$Y_t = H(\theta^p)s_t + u_t \tag{1.14}$$

$$s_{t+1} = F(\theta^p)s_t + G(\theta^p)\omega_{t+1} \tag{1.15}$$

where ω_{t+1} is the set of structural innovation, s_t is the state vector and Y_t represent the control variable vector. From an econometric perspective, (1.14)-(1.15) can be seen as a linear state space model, where u_t and ω_{t+1} represent the measurement and the process noise, respectively. u_t and ω_{t+1} are uncorrelated and normally distributed with zero mean and constant covariance matrix. Equation (1.14) is the measurement equation, which relates a set of observable variables, Y_t, to a set of (latent) state variables,

[2]The tax policies specification I report here is the one that best replicates a set of statistics. See Section 1.6 for different specifications.

s_t. State evolves along time according to equation (1.15). The matrices $F(\theta^p)$, $G(\theta^p)$ and $H(\theta^p)$ are functions of the structural parameters of the DSGE model. The vector of structural parameters is $\theta^p = [\alpha, \eta, \delta, \frac{K}{Y}, \tau^w, \tau^k, \varphi_w, \varphi_{by}, \varphi_y, \varphi_n, \psi_k, \psi_{by}, \psi_y, \psi_n, \rho_a, \rho_g, \rho_\chi, \rho_v]$. Depending on the parametrization of the DSGE model there are three possibilities: no stable rational expectation solution exists, the stable solution is unique (determinacy), or there are multiple stable solutions (indeterminacy). I focus on the determinacy case and restrict the parameter space accordingly. I define the vector of endogenous state variables, as

$$s_t = [k_t, y_t, \widetilde{\tau}_t^w, \widetilde{\tau}_t^k, \widetilde{by}_t, g_t, a_t, \chi_t, v_t]'.$$

The specification is completed by defining a set of measurement equations that relate the elements of s_t to a set of observable variables, equation (1.14). I assume that the model time period corresponds to a quarter and that the following observations are available: logarithm of per capita real GDP, consumption, hours worked, investment and government spending, the capital and labor average tax rate, and debt-GDP ratio. Given that the model assumes a stationary behavior for the data and the data do not display a stationary behavior, we need to create a data analog for the model concept; usually, researchers use filters to transform non stationary time series into stationary ones[3]. Many techniques exists for detrending data; these include linear detrending, HP filter to Bandpass filter. The HP filter is popular because it removes the low frequencies of the time series spectral representation, i.e. the trend part of the time series. Quadratic detrending does a good job for filtering consumption, GDP, investment and hours worked, but it does not seem to remove the low frequency component from the debt and government spending times series, see Figure 2-3; in fact, it produces cycles with duration of about 100 quarters. The Bandpass filter removes low and high frequencies. For our purposes the HP filter is sufficient, given that we want to remove the non stationary part of the data, see Figure 4-5. Therefore, all the series are HP filtered[4]. Let

$$Y_t = [c_t, GDP_t, n_t, inv_t, tax_t^w, tax_t^k, debt_t/GDP_t, gov_t]$$

be the vector of observable filtered variables.

In order to avoid stochastic singularity, I add four measurement errors: one on GDP, one on each tax, one on debt-GDP ratio, and one on government spending. Therefore, the vector of innovations is

$$\omega_t = [\epsilon_t^a, \epsilon_t^\chi, \epsilon_t^v, \xi_t^w, \xi_t^k, \epsilon_t^g, \epsilon_t^{my}, \epsilon_t^{mw}, \epsilon_t^{mk}, \epsilon_t^{mby}, \epsilon_t^{mg}]$$

The parameters space is then augmented by a vector of auxiliary or non structural parameters, namely the variances for the orthogonal shocks, i.e. $\theta^{ns} = [\sigma_a, \sigma_\chi, \sigma_v, \sigma_{\xi^w},$

[3] I also tried to rewrite the model in terms of ratios in order to have stationary times series without using any filtering, but the series expressed in ratio did not display a stationary pattern, see Figure 1.

[4] HP filtering may change the timing of information in the data. We check that in our case the distortion induced is very limited.

$\sigma_{\xi^k}, \sigma_g, \sigma_{my}, \sigma_{mw}, \sigma_{mk}, \sigma_{mby}, \sigma_{mg}]$. Thus, the parameter vector θ is just the union of the two, i.e. $\theta = [\theta^p, \theta^{ns}]$.
Once defined the state space representation, the Kalman Filter is used to compute the likelihood.

a MCMC Methods

Bayesian methods are employed to obtain the posterior distribution of the structural parameters. For both approaches, posterior distributions are a combination of prior distribution of the parameters, and sample information, which is given by the likelihood of the model. In general, posterior distributions are computed using the Bayes theorem

$$g(\theta|Y;\mathbb{M}) = \frac{g(\theta)\mathcal{L}(Y|\theta;\mathbb{M})}{p(Y|\mathbb{M})} \propto g(\theta)\mathcal{L}(Y|\theta;\mathbb{M})$$

where $\mathcal{L}(Y|\theta;\mathbb{M})$ is the likelihood of the data, Y, given a model, $\mathbb{M}$; θ is the vector of parameters of the model and $g(\theta)$ is the prior distribution of the parameters.
Given the large number of parameters involved, we can not compute analytically the posterior distribution, and we need to use posterior simulators based on Monte Carlo Markov Chain (MCMC) methods. The main idea of MCMC simulators is to define a transition distribution for the parameters that induce an ergodic Markov chain. After a large number of iterations, draws obtained from the chain are draws from the limiting target distribution. Following Schorfheide (2000), I use the Random Walk Metropolis algorithm (RWM). Given Σ and prior $g(\theta)$, the algorithm is as follow. Starting from an initial value θ_0, for $\ell = 1, ..., L$

1. draw a candidate $\theta_\dagger = \theta_{\ell-1} + N(0, \Sigma)$
2. solve the linear expectations system given $\theta_\dagger$; if indeterminacy or no-existence set $\mathcal{L}(Y|\theta_\dagger;\mathbb{M}) = 0$.
3. evaluate the likelihood of the system of equations (1.14)-(1.15) given $\theta_\dagger$ with the Kalman filter, $\mathcal{L}(Y|\theta_\dagger;\mathbb{M})$.
4. compute $\breve{g}(\theta_\dagger|Y;\mathbb{M}) = g(\theta_\dagger)\mathcal{L}(Y|\theta_\dagger;\mathbb{M})$, and the ratio
$$R = \frac{\breve{g}(\theta_\dagger|Y;\mathbb{M})}{\breve{g}(\theta_{\ell-1}|Y;\mathbb{M})}$$
5. draw u from $U[0,1]$; if $R > u$ then we accept the draw and we set $\theta^\dagger = \theta_\ell$, otherwise set $\theta_{\ell-1} = \theta_\ell$

Iterated a large number of times, the RWM algorithm ensures that we get to the limiting distribution which is the target distribution that we need to sample from (for further details see also Canova (2007), Ch. 9).

b Prior Selection and Computation Details

I fix the depreciation rate to 0.025, which implies an annual depreciation rate of 10%, and estimate the remaining parameters. In the first three columns of Table 1, I report the priors. Following standard practice, I choose Beta distributions for those parameters that must lie within the unit interval, like capital share or steady state taxes. The persistence parameters of the AR(1) processes are assumed to follow a Beta distribution as well, with mean 0.7 and standard deviation 0.08. The Beta distribution covers the range between 0 and 1, but a small standard error was used to have a clearer separation between stationary and non stationary shocks, as in Smets and Wouters (2003). For the coefficient of relative risk aversion, η, I pick a Gamma distribution with mean 2, which implies an intertemporal elasticity of substitution, $1/\eta$, of 1/2: close to the RBC literature values. Standard deviations are assumed to be distributed as Inverse Gamma with mean 0.23 and a loose prior. The distribution guarantees a positive variance with a relatively large domain. The remaining parameters are normally distributed. For the fiscal policy parameters I choose Normal distributions centered at positive values, 0.2, with a large standard deviation, 0.1. This implies that a priori fiscal policies are countercyclical on average but there is a positive probability that the coefficients are negative. I run 1,000,000 draws and I tune up the RWM variance in order to achieve a 30%-40% acceptance rate.

1.4 Parameters Estimates and Moments

I use quarterly values for detrended real GDP, consumption, investment, hours worked, labor and capital tax, debt-GDP ratio and government spending from 1966:2 until 2006:2. The times series are from the FRED database of the Federal Reserve Bank of St. Louis; the taxes series are constructed using the BEA database (in the Appendix I report the details on taxes series construction). Figure 6 reports the evolution of the likelihood across draws; after roughly 100,000 draws, the marginal likelihood stabilizes around the value of 3000. Nevertheless, the parameters chain takes more time to converge; indeed, Figure 7 shows the Cumulative Sum Statistics of the 28 parameters, and indicates that convergence for all the parameters is achieved after 500,000 draws.

In addition to the prior distributions, Table 1 reports the mean, standard deviation and median of the 28 parameters estimates obtained with the RWM algorithm. Figures 8 and 9 summarize this information visually by plotting the prior distribution (dashed line) and the posterior one (solid line). Overall, most of the parameters are estimated to be significantly different form zero. This is true for all the fiscal policy parameters, with the exception of φ_w.

Analyzing, first, the estimated stochastic processes, it appears that the variance of the preference shock, σ_χ, is larger than the technology shock, σ_a, which is similar to what it is found by Smets and Wouters (2003, 2005). The standard deviations of the tax shocks are significantly different from zero. The standard deviations of the measurement errors are generally smaller than structural shocks ones, with the exception of the capital tax rate, σ_{mk}, which is quite large. The persistency of the structural shocks is large for the preference and the technology process; the remaining processes are weakly serially autocorrelated.

Turning to behavioral parameters, the overall picture is pretty much in line with what is available in the RBC literature. The intertemporal elasticity of substitution is close to 1, implying a logarithmic preference for consumption. The capital share in production has a median estimate close to 0.3, somehow smaller than the value commonly used, 33%. Given that I treat capital as an unobservable variable, it is difficult to pin down α precisely. As a by-product, the time discount factor, $\beta = 1/((1-\tau^k)\alpha[\frac{K}{Y}]^{-1}+1-\delta)$, has a mean value of 0.997, which implies a steady state interest rate of 0.3% for quarterly values.

Moving now to the fiscal policy parameters, $\theta^f = [\varphi_w, \varphi_{by}, \varphi_n, \varphi_y, \psi_k, \psi_{by}, \psi_n, \psi_y]$, we can notice that generally posterior standard deviations are smaller than the prior ones. The parameters controlling government debt, φ_{by} and ψ_{by}, are greater than zero as one would expect; since they are meant to avoid the explosiveness of public debt, the tax reaction should be positively correlated to debt changes. The two taxes are weakly correlated; ψ_k is negative and hard to identify (see also Section a). For our purpose the most interesting parameters estimates are φ_n, φ_y, ψ_n and ψ_y, which summarize the automatic response of the fiscal policy to cyclical conditions. The automatic response of the labor tax, φ_y and φ_n, are strictly positive, and we can rule out the possibility that these coefficients are zero or negative. This is clear when looking at plots of prior and posterior distributions (Figure 8 eight and ninth panels from top left); indeed, the left tail of the posterior assigns (almost) zero probability to the event $\varphi_y \leq 0$. This fact implies that the labor tax function is procyclical, corroborating the idea that fiscal policy has been countercyclical along the period considered. The story is similar for the capital tax, although the conclusions are less significant. Figure 8 indicates that in most of the cases the posteriors of the fiscal parameters do not overlap with the priors, meaning that the fiscal policy parameters are identifiable [5].

In the context of DSGE models, there is not a unique measure of model fit: some authors compare moments (see An and Schorfheide (2007)), others compare impulse responses. Since I am focusing on stabilization issues, I look at second moments. Model standard deviations are constructed simulating the model using a subset of accepted draws, and the values considered are the average standard deviations across simulations. In general, the model moments are computed as follows: let $h(Y)$ be a set of statistics constructed using the data set, $Y = \{y_t\}_{t=1}^T$. The average model statistic is given by

1. draw θ^ℓ from the posterior distribution, $g(\theta|Y;\mathbb{M})$.

[5] For further details on identification issues and tax policies rules specification I direct the reader to Section 1.6 for a more detailed discussion.

2. simulate a time series, $Y(\theta^\ell)$, of 500 observation length using equations (1.14) and (1.15),

3. discard the first $500 - T$ observations to eliminate the dependence of the first observation and compute the statistic of interest, $h(Y(\theta^\ell))$, using the simulated data, $Y(\theta^\ell) = \{y_t(\theta^\ell)\}_{t=1}^T$,

4. do 1) to 3) for $\ell = 1, ..., 100$ and compute

$$\bar{h}(\theta) = \frac{1}{100}\sum_{\ell=1}^{100} h(Y(\theta^\ell))$$

Table 2 presents the standard deviations of the observed data and of the model, respectively; standard deviations tend to be under-estimated. Nevertheless, the model is able to replicate the large difference between the investment standard deviation and the standard deviations of the other variables; in particular, in the data the investment standard deviation is around 4.4 percentage points, whereas the remaining standard deviation are around 0.85 and 1.60. In the model, this gap is preserved: investment standard deviation is about 2.2 percentage points, whereas the remaining standard deviations vary in the range of 0.15-0.45 percentage points.
I also compare contemporaneous correlations in the data and the correlations implied by the model. In Figures 10 and 11, I plot simulated correlations against actual. More precisely, the circles represent the correlation produced by the model and the vertical and horizontal lines represent the correlations in the data. I obtain nice results for the correlation between GDP and government consumption (seventh panel of Figure 10), between GDP and private consumption (first panel of Figure 10), the correlation between hours worked and investment (seventh panel of Figure 10), and between hours worked and government consumption (second panel of Figure 10). Nevertheless, it is clear that the model fails to match some other correlations. However, since I am trying to match many moments, it is predictable that the model does not replicate the data in all dimensions. The more relevant issue is whether the model provides an adequate framework for policy analysis. In this respect, the RBC model with time varying wedges and with tax rules does as well as most variants of DSGE models.
To gain further insights about the plausibility of the estimates it is instructive to look at the responses of variables of interest to structural shocks. Figures 12-15 plot the responses of consumption, GDP, hours worked, investment, labor and capital tax, debt-GDP ratio to a unitary increase in the technology, preference, labor tax and capital tax, respectively. Responses are constructed using the median value of the parameter estimates.
Figure 12 shows that, in response to a positive technology shock, output consumption and investment increase. Contrary to standard RBC models without a time varying 'wedge' and without tax rules, I obtain that the response of employment is negative

(third panel from top left); as pointed out by Galí (1999), the fall in employment is consistent with the estimated impulse responses of identified productivity shocks in the United States. Figure 13 shows the effects of one percent increase in the preference process on GDP, consumption, investment, taxes and debt-GDP ratio. A positive preference shock reduces the marginal rate of substitution and makes consumption fall. Leisure is more valuable and hours worked contract, making GDP drop as well; similar behavior can be found in Del Negro, Schoerfheide, Smets and Wouters (2007).
A labor tax shock, Figure 14, lowers the after tax wage inducing an increase in leisure; since hours worked drop, GDP contracts and so does consumption. The economy recovers very fast because of the countercyclicality of the fiscal policy. Since employment, GDP and debt-GDP ratio are below the steady state level and given that $\varphi_y > 0$ and $\varphi_n > 0$, the labor tax rate returns rapidly to the steady state level. A positive capital tax shock increases the average capital tax and decreases the average labor tax rate, Figure 15 fifth panel from top left. This produces a shift in the resources devoted to leisure to an increase in hours worked, which increases GDP, consumption and investment. The response to labor tax shocks are temporary and the effects die out within ten quarters. Finally, a labor tax shock produces considerably stronger responses relative the capital tax shock.

1.5 Tax Policies and Stabilization

The question addressed in this chapter is whether US tax policies helped to reduce business cycle fluctuations. As mentioned, there are two channels through which the fiscal authority can adjust tax rates: there is an endogenous or 'systematic' channel, by which the policymakers automatically respond to cyclical fluctuation, and I identified it with the four coefficients in the fiscal policy rule φ_y, φ_n, ψ_y and ψ_n. The second channel is through the orthogonal part of the fiscal policies rule. The estimated tax responses are

$$\begin{aligned}
\widetilde{\tau}_t^w &= 0.06\widetilde{\tau}_{t-1}^w + 0.47n_t + 0.66y_t + 0.38\widetilde{by}_t + \xi_t^w \\
\widetilde{\tau}_t^k &= -0.18\widetilde{\tau}_{t-1}^k + 0.26n_t + 0.29y_t + 0.37\widetilde{by}_t + \xi_t^k
\end{aligned}$$

The task here is to see whether they are important in influencing the amplitude of business cycles fluctuations, and if shutting them would affect the volatility of the macro variables. So, the question is: what happens to the standard deviations of GDP, consumption, investment and hours worked if I set these coefficients to zero, in particular, if $\varphi_y = \varphi_n = \psi_y = \psi_n = 0$ [6]. The importance of the automatic response is summarized in Table 3 and in Table 4. In both tables, the first column reports the (average) standard

[6]The coefficients on debt-GDP ratio can not be set to zero because they guarantee the stability of the DSGE solution.

deviations of the model when the two channels are operational. The second column in Table 3 displays the (average) standard deviations without the automatic response of the taxes, i.e. $\varphi_y = \varphi_n = \psi_y = \psi_n = 0$; these statistics are constructed using the parameter estimates of the unrestricted model setting the four fiscal policy parameters to zero. The third and fourth columns report the (average) standard deviations setting $\varphi_y = \varphi_n = 0$ and $\psi_y = \psi_n = 0$, respectively. The last column reports the (average) standard deviations without the exogenous channel, i.e. when fiscal shocks are set to zero.

Table 4 reports the same statistics using the estimates from a restricted model, with $\varphi_y = \varphi_n = \psi_y = \psi_n = 0$ in the second column, and with $\varphi_y = \varphi_n = 0$ and with $\psi_y = \psi_n = 0$ in third column and in the fourth column, respectively. Although both tables give the same information, I focus the discussion on restricted models. To test counterfactual hypothesis, most of the literature uses unrestricted models. As pointed out in Canova and Gambetti (forthcoming), such an approach could give misleading results since the coefficients of the reduced form are correlated. Therefore, when policy coefficients vary one should expect structural coefficients in other equations to change as well. This correlation structure is totally disregarded when unrestricted counterfactuals are performed. Since Table 4 takes this correlation structure into account, I will discuss the results of this table. Contrasting the first column with the second column, there is a consistent increase in the standard deviation of all macroeconomic variables; in particular, the consumption standard deviation increases by 27%; GDP standard deviation increases by 8%, employment standard deviation by 28% and investment standard deviation increases by 36%. Thus, without the procyclical behavior of taxes, the volatility of consumption, GDP, employment, investment would have increased, suggesting that the countercyclical reaction of fiscal policy is important for smoothing fluctuations. On the other hand, debt standard deviation decreases slightly without a countercyclical policy. Thus, with a countercyclical fiscal policy, the reduction in the amplitude of the fluctuations is compensated by an increase in the volatility of debt, from 0.26 without the automatic response to 0.43 with the automatic response. The intuition is the following: since the labor tax is procyclical (in recession the average tax decreases and in expansion increases), the debt needs to adjust more to satisfy the government budget constraint. The third and the fourth columns are meant to disentangle the impact on volatility of the two different taxes. The automatic response of the labor tax is more important in reducing the volatility than the capital tax one; indeed, if I omit the labor tax stabilizer (i.e. $\varphi_n = \varphi_y = 0$), the standard deviations of consumption, GDP, hours worked and investment increase more relatively to the case in which I omit the capital tax automatic stabilizer (i.e. $\psi_n = \psi_y = 0$). Therefore, labor tax policy is more prone to affect business cycle amplitudes. This fact is not surprising given the information present in the impulse responses; indeed, we saw that the labor tax shock produces a stronger pattern of responses than the one produced by the capital tax. The 'wrong' counterfactuals give the same results, although the differences are more pronounced.

Comparing the first column of Table 3 with the last one, the impact of fiscal policy shocks on the volatility of the macro variables is modest; without tax policy disturbances, consumption standard deviation decreases by 10%, GDP standard deviation by 4%, employment standard deviation by 7% and investment by standard deviation by 10%. The latter finding is consistent with what was found in the narrative approach of Romer and Romer (2007). They identified fiscal shocks by looking directly at the official documents (such as presidential speech, Congressional reports) and they found, with OLS regressions, that (twelve lags of) fiscal shocks explain 9% of the GDP growth rate volatility.
In sum, a stabilizing tax policy should respond in a procyclical way with respect to GDP and employment; when the economy expands labor tax should increase, an vice versa when the economy contracts. The pure discretionary channel does not help much in stabilizing the economy.

1.6 Robustness

Two delicate issues deserve separate attention.
First, there might be problems in identifying the tax policy parameters; especially, it could be that the magnitude or the sign of the estimated coefficients are not trustworthy. To avoid these problems, I did a number of robustness checks to verify whether the fiscal policy parameters were identified or not.
Second, I have already mentioned that fiscal policy rules are linear combinations of automatic stabilizers and unexpected changes. This broad definition leaves space for arbitrariness in writing down the exact fiscal policy function. For instance, Jones (2002) defines the (log deviation from the steady state) labor and capital tax as a linear combination of present and lagged values of GDP, hour worked and lagged values of taxes and government consumption. Davig and Leeper (2007) let the (deviation from the steady state) tax depend on the past level of debt-GDP ratio, current GPD, government consumption. In Gali and Perotti (2004), the primary deficit responds to the expected value of output gap, to debt, to the past level of the deficit and an orthogonal shock. The fiscal rule in Giammarioli et al. (2007) is a log linear function in which the tax is a linear combination of real public debt, government spending and the technology process. To account for this variety, I experiment with different possibilities.

a Identification

As we are aware, one should worry about weak identification of the tax policy parameters, in terms of magnitude and especially in terms of sign.
To verify the extent of these problems, Figures 16 and 17 plot the marginal likelihood against the fiscal policy parameters; in particular, the Figures represent on the x and y axis the tax policy parameters and on the the z axis the marginal likelihood. The

surface is computed leaving the other parameter free to vary. The likelihood is quite flat along some dimensions. Nevertheless, it is clear that the likelihood is in the positive quadrant of the (x, y) axes and the more we approach the origin the lower the likelihood becomes. This indicates that the signs of φ_n, φ_y, ψ_n and ψ_y are well identified, and that tax policies are countercyclical.

To further verify the extent of the identification problems, I augment and reduce the standard deviations of the priors and repeat the estimations. In Table 5, I report the fiscal policy parameter estimates, mean and standard deviations in parenthesis. The first column reports the fiscal policy parameters estimate with the baseline priors standard deviation and the following columns the estimate with the prior standard derivations increased and reduced by 20%, respectively. Overall, parameter estimates do not change much. First, when I increase the standard deviation by 20%, the estimated sign are preserved. Moreover, in most of the cases, the estimates fall into the 95% confidence interval of the baseline prior standard deviations. When I reduce the standard deviations, the tax policy parameter estimates seem to get closer to the priors mean (0.2). Also, there is a change in sign for the autoregressive component in the tax rule, ψ_k, and φ_w is statistically undistinguishable from zero. In sum, the core tax policy parameter estimates appear to be well estimated, and we can be confident about their sign and magnitude.

b Tax rule specifications

In this section, I examine different tax rules and I use a modified version of the Schorfheide's (2000) loss based method to select among them. The specifications that I consider are

- Taxes respond to GDP, $\mathcal{S}_1$:

$$\widetilde{\tau}_t^w = \varphi_w \widetilde{\tau}_{t-1}^w + \varphi_y y_t + \varphi_{by} \widetilde{by}_t + \xi_t^w$$
$$\widetilde{\tau}_t^k = \psi_k \widetilde{\tau}_{t-1}^k + \psi_y y_t + \psi_{by} \widetilde{by}_t + \xi_t^k$$

- Taxes respond to employment, $\mathcal{S}_2$:

$$\widetilde{\tau}_t^w = \varphi_w \widetilde{\tau}_{t-1}^w + \varphi_n n_t + \varphi_{by} \widetilde{by}_t + \xi_t^w$$
$$\widetilde{\tau}_t^k = \psi_k \widetilde{\tau}_{t-1}^k + \psi_n n_t + \psi_{by} \widetilde{by}_t + \xi_t^k$$

- Taxes respond to employment and GDP, $\mathcal{S}_3$:

$$\widetilde{\tau}_t^j = \varphi_w \widetilde{\tau}_{t-1}^j + \varphi_y y_t + \varphi_n n_t + \varphi_{by} \widetilde{by}_t + \xi_t^w$$
$$\widetilde{\tau}_t^k = \psi_k \widetilde{\tau}_{t-1}^k + \psi_y y_t + \psi_n n_t + \psi_{by} \widetilde{by}_t + \xi_t^k$$

- Taxes respond to expected GDP as in Gali and Perotti (2004), $\mathcal{S}_4$:

$$\begin{aligned}\widetilde{\tau}_t^w &= \varphi_w \widetilde{\tau}_{t-1}^w + \varphi_y E_t y_{t+1} + \varphi_{by} \widetilde{by}_t + \xi_t^w \\ \widetilde{\tau}_t^k &= \psi_k \widetilde{\tau}_{t-1}^k + \psi_y E_t y_{t+1} + \psi_{by} \widetilde{by}_t + \xi_t^k\end{aligned}$$

- As in $\mathcal{S}_2$ but with debt instead of debt-GDP ratio, $\mathcal{S}_5$:

$$\begin{aligned}\widetilde{\tau}_t^w &= \varphi_w \widetilde{\tau}_{t-1}^w + \varphi_n n_t + \varphi_y y_t + \varphi_b b_t + \xi_t^w \\ \widehat{\tau}_t^k &= \psi_k \widehat{\tau}_{t-1}^k + \psi_y y_t + \psi_n n_t + \psi_b b_t + \xi_t^k\end{aligned}$$

As mentioned, it is difficult to compare the 'likelihood' of different DSGE models. Although Bayes factor and posterior odds are commonly used to quantify the validity of a DSGE model relative to another, these comparisons may not be very informative about the quality of the approximation to the data. This happens in particular when one wishes to compare misspecified models, see Canova (2007).
The Posterior Odds ratio is constructed by comparing the Bayes Factor, which is the ratio of the predictive densities of the data conditional on different models, and prior odds, which is the ratio of prior probabilities associated to each model. The predictive density of the data, Y, conditional on the model, $\mathbb{M}$, for a given prior $g(\theta)$ is

$$p(Y|\mathbb{M}) = \int \mathcal{L}(Y|\theta;\mathbb{M})g(\theta)d\theta$$

Therefore, if one wishes to test different tax policy specifications S_i and S_j, the Posterior Odds are computed as

$$PO_{\mathcal{S}_i,\mathcal{S}_j} = \frac{g(\mathcal{S}_i)}{g(\mathcal{S}_j)} \times \frac{p(Y|\mathcal{S}_i)}{p(Y|\mathcal{S}_j)}$$

where $g(\mathcal{S}_i)$ and $g(\mathcal{S}_j)$ are prior probabilities of each trend specification.
Table 6 reports the marginal likelihood of the specifications and the log of the Posterior Odds ratio relative to $\mathcal{S}_3$. It is clear that the two nested specifications, $\mathcal{S}_1$ and $\mathcal{S}_2$, can be discarded relative to the specification $\mathcal{S}_3$. Moreover, if we compare $\mathcal{S}_1$ against $\mathcal{S}_4$ we do not gain much by adding an expectation term for GDP. Thus the two best specification are $\mathcal{S}_3$ and $\mathcal{S}_5$. Since marginal likelihood are very similar, to choose a specification among the two, I present a method in the spirit of Schorfheide (2000) that can be used to select among them. This method involves the following steps

1. Select a set of data statistics, $h_T = h(Y)$, where $Y = \{y_t\}_{t=1}^T$ are the observed data.

2. Draw θ_j^ℓ from $g(\theta|Y;\mathcal{S}_j)$ for j=1,...,5.

3. Simulate a time series, $Y(\theta_j^\ell)$, of 500 observation length using equations (1.14) and (1.15); discard the first $500 - T$ observations, $Y(\theta_j^\ell) = \{y_t(\theta_j^\ell)\}_{t=1}^T$, and compute the statistic of interest, $h_j^\ell = h(Y(\theta_j^\ell))$.

4. Construct a measure of discrepancy with the data (Loss function). I consider quadratic loss function, i.e.

$$L(h_T, h_j^\ell) = (h_T - h_j^\ell)' W (h_T - h_j^\ell)$$

   If one wants to give the same weight to each statistic, set $W = I$.

5. Do 2) to 4) L times and compute the average error in replicating the data statistics

$$R(h_j(\theta)) = \frac{1}{L} \sum_{\ell=1}^{L} L(h_T, h_j^\ell)$$

$R(h_j(\theta)|y)$ measures how specification $\mathcal{S}_j$ predicts a set of statistic of the data, h_T, on average; in particular, it evaluates the average risk in prediction. So, the lower is $R(h_j(\theta))$, the better the specification $\mathcal{S}_j$ replicates the data statistics. Table 6 presents the average risk in replicating standard deviations and correlations; in particular, the first line displays the average risk in replicating the data standard deviations, and the following lines show the average risk in replicating the correlations of a variable between the remaining ones; so, for example, $R(corr_c(\theta)|y)$ is the average risk in prediction the correlations between consumption and GDP, between consumption and hours worked, between consumption and investment, between consumption and the two taxes, between consumption and debt and between consumption and government spending. Even thought there is not a specification strictly preferred to the others, the third specification seems to have the lowest risk in replicating standard deviations and correlation at the same time. The main competitor is the fifth specification, $\mathcal{S}_5$, which is identical to $\mathcal{S}_3$ with the exception that in $\mathcal{S}_3$ I consider debt-GDP ratio and in $\mathcal{S}_5$ the debt. $\mathcal{S}_3$ does worse than $\mathcal{S}_5$ in replicating the correlation between labor tax and other variables. Nevertheless, along the remaining dimensions $\mathcal{S}_3$ has a lower risk in predicting moments. Hence, specification $\mathcal{S}_3$ where taxes respond to GDP, employment and debt-GDP ratio seems to be the best candidate for policy analysis.

1.7 Conclusions

In this chapter, I study whether US tax policies affected the volatility of the macro variables. There are endogenous feedbacks between economic activity and tax policies; on the one hand, the latter directly affects household decisions influencing consumption and labor choices, and therefore economic activity. On the other hand, the fiscal authority sets the tax policy by responding to cyclical situations of the economy. The task of this work was to estimate from the data the feedbacks between economic activity and tax policies, and in particular to test whether tax policies are useful in reducing economic volatility. To answer the question of interest, I chose to employ a General Equilibrium model that provides a theoretical framework to identify endogenous interactions. I found that tax policies helped to reduce economic volatility. In particular, the automatic response to cyclical conditions has been very important in shaping macroeconomic stability; indeed, if we assume that the labor and capital taxes do not respond to GDP and employment variations, the volatility of the main macro variables would increase. I also found that unexpected changes in the tax policy do not affect much the economic volatility. I also did a number of robustness checks to verify the estimates of the fiscal policy parameters and the tax policy specifications. The analysis I present has two major limitations. First, as mentioned before, I do not consider government spending as a possible fiscal instrument. I let the government follow an autoregressive process, but I do not assume that it has a direct impact an household decisions, and therefore I rule out the possible interaction between economic stability and government size, Galí (1994). Second, the policy implications are valid conditional on the absence of a monetary authority; I am ignoring possible interactions between fiscal and monetary policy. Thus, an interesting extension to this chapter could be re-estimate the model after introducing prices and a monetary authority into the economy.

2 Trend agnostic estimation of DSGE models

2.1 Introduction

Dynamic Stochastic General Equilibrium (DSGE) models are now considered the benchmark for macro analysis. Models are much more complex than in the past and in the last 10 years there has been considerable progress in estimating deep parameters of DSGE models. These improvements allow researchers to asses the degree of fit both in and out of sample, to test counterfactual hypotheses and to evaluate policy implications. In general, DSGE models are now considered trustworthy tools for policy analysis also because of a more rigorous econometric evaluation.

The vast majority of models nowadays is intended to capture cyclical fluctuations. This attitude is reflected in the relative number of parameters that seize cyclical and non-cyclical movements: indeed, in existing DSGE models almost all the parameters are meant to describe Business Cycles fluctuations, whereas none or rather few to explain non-cyclical movements. Since data contains fluctuations which do not need to be cyclical, preliminary data transformations are required when the model is estimated. In particular, applied researchers typically employ a 'two step' procedure to estimate structural DSGE parameters: in the first step, the cyclical component is extracted from the data; in the second, DSGE structural parameters are estimated using the transformed data. The first step involves either filtering the data[1] or defining a model-based concept of non stationary fluctuations and transforming the data accordingly[2]. In either cases, two step procedures have problems. First, an improper choice of trend affects structural parameters estimates. Cogley (2001) shows that a wrong trend specification leads to strong bias in parameter estimates with likelihood based methods. Even when the reduced form of the cyclical component is correctly specified, trend misspecification is likely to results in inconsistent estimates of 'deep' parameters. On the same track, Gorodnichenko and Ng (2007) show that estimates can be severely biased when the model concept of trend is inconsistent with data or detrended data are inconsistent with the model concept of stationarity. Second, wrong assumptions about the correlation between cyclical and non-cyclical components may bias structural parameters estimates. In two step approaches, the typical assumption used to identify trend and cycles is that the two are independent, but one can easily think of theoretical and practical reasons for making them correlated (see Comin and Gertler (2006) or Canova, Lopez-Salido and Michelacci (2007)). Third, unless one wants to take a strong stand on the property of the model, e.g. the model is a representation of HP filtered data, the uncertainty about the filter is likely to affect structural parameter estimates.

[1] See Smets and Wouters (2003), Rabanal and Rubio-Ramirez (2005), Rabanal (2007), Bouakez, Cardia and Ruge-Murcia (2005), Christensen and Dib (2008), among others.

[2] See Smets and Wouters (2007, 2005), Del Negro et al. (2007), Justiniano and Primiceri (2008), Rabanal (2006), among others.

In this chapter, I propose an alternative method to estimate DSGE models, where structural parameters are jointly estimated with trend parameters. The trend specification is flexible enough to capture various low frequency movements. I refer to this as the 'one step' approach. Among other things, the one step approach has two important by-products:

1. We can test trend specification.
   One could test the most likely trend specification for individual series or for a subset of them. Moreover, the setup is flexible enough to allow for potential instability in the trend parameters; if one suspects that a subsets of times series has experienced a change in its long run behavior, such a hypothesis can be tested.

2. We can construct robust structural parameters estimates via Bayesian averaging.
   Besides testing specifications, the one step approach is suitable to account for trend uncertainty. Given that we do not know the 'true' trend generating process, one can construct robust structural estimates by taking a weighted average of the estimates obtained with various trend specifications, with weights given by their posterior probability.

I show through Monte Carlo experiments that the one step approach has appealing properties in small samples. When trend is correctly specified, parameter bias is larger in the two step then in one step approach both in the deterministic and in the stochastic setup. The procedure displays also desirable features under misspecification. In particular, the one step estimates are robust to two types of misspecification: (a) when the trend specification is wrong, i.e. 'true' trend is deterministic and estimated as if it were stochastic (and viceversa), (b) when the assumption about the correlation between trend and cycles is wrong. The intuition for these results is as follows. The first of the two step involves the estimation of the trend parameters, and the residuals of the trend estimation are then the cycles. Thus, in the first step we are neglecting the information that the cycles have a specific structure, i.e. the solution of the DSGE model. The one step approach treats trend and cycles as unobserved states, and their parameters are jointly estimated; thus, all the information is jointly processed. Moreover, in almost all the cases the procedure is able to recover the true trend generating process through posterior weights.
When we apply the procedure to actual data interesting results emerge. First, since different data transformations imply different cycles (see Canova (1998)), data transformation affects the estimates of structural parameters. In this respect, the estimates of the exogenous processes (persistence and magnitude) mimic the duration and the amplitude of the cyclical component: indeed, the deeper are the cycles the larger the standard deviations are, and the longer are the cyclical fluctuations the more persistent the shocks are. Moreover, different structural parameter estimates produce different implications of the model, i.e. different impulse responses or distinct contributions of

the structural shocks to the volatility of the observable variables. While the two step procedure lacks a statistical-based criterion to select among them, the one step approach provides a natural benchmark to choose among different structural parameter estimates, and allows also to construct DSGE estimates robust to the trend uncertainty. Finally, applying the two approaches to a medium scale DSGE model different implications arise in terms of sources of GDP volatility at business cycles frequencies. I find that with a two step approach the main sources of GDP volatility are markup shocks, regardless of the type of filter employed. With a one step approach the GDP variance decomposition changes substantially according to trend specifications; I obtain that the most likely contribution to GDP volatility is given by investment-specific shocks.

Since the seminal paper of Cogley (2001), few papers have analyzed the impact of trend specification on structural parameter estimates. Fukac and Pagan (forthcoming) propose a limited information method to deal with the treatment of trend in DSGE estimations. While their analysis is confined to a single equation framework, Gorodnichenko and Ng (2007) extend the Cogley's analysis and propose a robust approach exploiting all the cross-equations restrictions of the DSGE model. They use simulated method of moments, which are prone to severe identification problems (see Canova and Sala (fortcoming)). Even though I share with them an 'agnostic' view about the non-cyclical properties of the data, my approach differs in two respects. First, I consider 'off-model' trends; this makes the structure able to capture not only linear deterministic and unit root trends, but also higher order integrated smooth trends. Moreover, the proposed setup is flexible enough to permit several hypothesis testing, such as testing for correlation among trends or for trend parameters instability. Second, I employ a structural times series approach and likelihood based methods, as in Canova (2008); this avoids any data transformation before or during estimation. While he focuses on a unique representation of the non-cyclical component that encompasses various low frequencies behavior, the proposed estimation strategy exploits the posterior weights of potentially many specifications, and by averaging across them structural parameters are robust to trend uncertainty.

The present chapter is organized as follows. Section 2.2 presents the econometric methodology with emphasis on the two approaches. In Section 2.3 the two procedures are confronted under various Monte Carlo experiments; results and biases are reported. Section 2.4 presents results and conclusions using actual data; two DSGE models are considered for estimation. A 'small' scale DSGE model is used to provide straight intuitions for the results and a more densely parameterized model is employed. Section 2.5 concludes the chapter.

2.2 Econometric Methodology

In this section, I develop the statistic framework I use to estimate the structural DSGE parameters. I first present the traditional two step approach, followed by the one step method I propose. The main idea of the one step approach is to compute the likelihood of a system that embodies a reduced form representation for the trend and a structural form for the cycles. More precisely, I assume that the linearized solution of the model provides a representation for the cyclical movements of the variables. These cyclical movements are combined with a parametric representation of non-cyclical fluctuations, and structural and non-structural parameters are jointly estimated. The general representation is flexible enough to allow as special cases various low frequency specifications of the trends.

I assume that we observe $y = \{y_t\}_{t=1}^{T}$, the log of a set of times series. As in Harvey, Trimbur and Dijk (2004), I assume that the data is made up of a non-stationary trend component, y^{τ}, and a cyclical component, y^{c}, so that

$$y = y^{\tau} + y^{c} \tag{2.1}$$

I also assume that the log-linear solution of the DSGE model represents the cyclical behavior of the data, i.e.

$$\begin{aligned} y_t^c &= RR(\theta^m)x_{t-1} + SS(\theta^m)z_t && (2.2)\\ x_t &= PP(\theta^m)x_{t-1} + QQ(\theta^m)z_t && (2.3)\\ z_{t+1} &= NN(\theta^m)z_t + \nu_{t+1} && (2.4) \end{aligned}$$

where PP, QQ, RR, SS are matrices which are functions of the structural parameters of the model, θ^m; x_{t-1} and z_t are the state vectors of the model, endogenous and exogenous respectively. ν_{t+1} are mutually uncorrelated zero mean innovations.

In a two step approach, the cyclical component is first extracted from the data. Then, the likelihood of the data, conditional on the DSGE model, $\mathcal{M}$, is computed

$$\mathcal{L}(y^c|\theta^m; \mathcal{M})$$

With the one step approach, we compute the likelihood of the observed data, given a system that embodies the solution of the model and a specification for the trend, i.e.

$$\mathcal{L}(y|\theta; \mathcal{M}, \mathcal{F})$$

where $\theta = (\theta^m, \theta^f)$ is the joint vector of structural and filtering parameters, and $\mathcal{F}$ is a functional specifications for the filter.

The likelihood is computed using the Kalman filter after having defined a linear state space[3], equations (1.14)-(1.15).

[3]Non linear state space can be found in Fernandez-Villaverde and Rubio-Ramirez (2005)

a Two step approach

With the two step (2s) approach data is first filtered and then structural DSGE parameters estimated.

- 1^{st} step:
  Assume that $\mathcal{F}(y_t;\tau,\mathcal{M})$ is the filter that extracts the trend y_t^τ from the data, given the model $\mathcal{M}$. Then, the cyclical component is

$$y_t^c = y_t - \mathcal{F}(y_t;\tau,\mathcal{M})$$

  Notice that when a statistical filter is used $\mathcal{F}(y_t;\tau,\mathcal{M}) = \mathcal{F}(y_t;\tau)$, while when a model-based filter is used $\mathcal{F}(y_t;\tau,\mathcal{M}) = \mathcal{F}(y_t;\mathcal{M})$. For example, a DSGE model with a unit root with drift in the technology process would imply real variables to grow at the same rate, the technology growth rate. Therefore, the model-based filter would require to take first difference on real variables data and leave unchanged the remaining ones.
  In both the one step and the two step approach, I consider only statistical filters, thus $\mathcal{F}(y_t;\tau,\mathcal{M}) = \mathcal{F}(y_t;\tau) \equiv \mathcal{F}_\tau(y_t)$. In particular, I consider three types of trends, τ: a linear trend, a unit root and a smooth integrated trend. Therefore, the appropriate filters are a linear detrending filter, a first order difference filter, and the unobserved component (Hodrick-Prescott) filter.

- 2^{nd} step:
  When y_t^c is obtained, the system of equation, (2.2)-(2.4), fit the state space representation, (1.14) and (1.15), by setting

$$\begin{aligned}
Y_t &= y_t^c \\
s_t &= \begin{pmatrix} x_{t-1} & z_t \end{pmatrix}' \\
F &= \begin{pmatrix} PP & QQ \\ 0 & NN \end{pmatrix} \\
G &= \begin{pmatrix} 0 & I \end{pmatrix}' \\
H &= \begin{pmatrix} RR & SS \end{pmatrix} \\
\omega_{t+1} &= \nu_{t+1}
\end{aligned}$$

The choice of the filter, $\mathcal{F}(y_t;\tau,\mathcal{M})$, affects the statistical properties of the cycles (see Canova (1998)), and consequently also the shape of the likelihood. This implies that the estimated structural parameters might be (statistically) different depending on the filter used (see Canova (2008)).

b One step approach

In the one step approach (1s) the likelihood is computed directly from the observables, y_t, that is

$$\begin{aligned} y_t &= y_t^\tau + y_t^c \\ y_t^\tau &= \mathcal{F}(y_t; \tau) \\ y_t^c &= RRx_{t-1} + SSz_t \\ x_t &= PPx_{t-1} + QQz_t \\ z_{t+1} &= NNz_t + \nu_{t+1} \end{aligned}$$

The following specifications fit the state space system, equations (1.14)-(1.15). Details are reported in the appendix.

Linear-Trend-DSGE setup

In this specification, I assume that the non-stationary component of the data is driven by a linear trend, i.e.

$$y_t^\tau = A + B * t + \eta_t \tag{2.5}$$

where A and B are column vectors. η_t is a white noise normally distributed with zero mean and variance covariance matrix, Σ_η. Therefore, the filter parameters to be estimated are $\theta^{lt} = [A, B, \Sigma_\eta]$. I will refer to this specification as lt-dsge setup.

First-Difference-DSGE setup

In this specification I assume that the data displays a unit root pattern, and that

$$y_t^\tau = \gamma + \Gamma y_{t-1} + \eta_t \tag{2.6}$$

where γ is the drift and Γ is a diagonal matrix, that have zeros or ones on the main diagonal. η_t is a white noise normally distributed with zero mean and variance covariance matrix, Σ_η. Therefore, the filter parameters to be estimated are $\theta^{fd} = [\gamma, \Sigma_\eta]$. I will refer to this specification as fd-dsge setup.

Hodrick-Prescott-DSGE setup

Here, I assume that the trend, y_t^τ, is an integrated random walk, i.e.

$$y_{t+1}^\tau = y_t^\tau + \mu_t \tag{2.7}$$

$$\mu_{t+1} = \mu_t + \zeta_{t+1} \tag{2.8}$$

where $\zeta_{t+1} \sim N(0, \Sigma_\zeta)$, and Σ_ζ is diagonal. Harvey and Jaeger (1993) have shown that the HP filter is the optimal trend extractor, when the trend, y_t^τ, is specified as in

(2.7) and (2.8). The set of shocks, ω_{t+1}, of the state space model is composed by the structural innovations of the model, ν_{t+1}, and the stochastic part in the trend, ζ_{t+1}. To make the link with the HP filter clearer, note that the ratio between the variance of innovations in trend and the variance of the cycles gives the smoothing parameter of the HP filter, λ. Usually, the smoothing parameter is set to 1'600 for quarterly values, but there is little reason for this choice. To account for the uncertainly in setting λ, Trimbur (2006) proposes a Bayesian HP filter where λ is estimated with a Gibbs sampler; he shows that depending on the times series λ can be statistically different from 1'600. In the hp-dsge set up the ratio of the variances is estimated along with the structural parameters of the DSGE model; this allows the statistical framework to be quite flexible. The filter parameters to be estimated are $\theta^{hp} = \Sigma_\zeta$. I assume that Σ_ζ is diagonal, but it is straightforward to consider a general matrix (allowing for correlation among trends), or a rank deficient one (so that the non stationary component is common across series). I will refer to this specification as hp-dsge estimates.

c Estimation

Bayesian methods are employed to obtain the posterior distribution of the structural and non-structural parameters. Recall that posterior distributions are computed using the Bayes theorem

$$g(\theta|Y;\mathbb{M}) = \frac{g(\theta)\mathcal{L}(Y|\theta;\mathbb{M})}{p(Y|\mathbb{M})} \propto g(\theta)\mathcal{L}(Y|\theta;\mathbb{M})$$

where $\mathcal{L}(Y|\theta;\mathbb{M})$ is the likelihood of the data, Y, given a model, $\mathbb{M}$; θ is the vector of parameters of the model and $g(\theta)$ is the prior distribution of the parameters.
In the two step approach, we compute the posterior distribution of the parameters conditional on filtered data, y^c, and on the DSGE model, $\mathcal{M}$. Thus, $\mathbb{M} = \mathcal{M}$, $Y = y^c$ and $\theta = \theta^m$, and the posterior distribution of parameters is

$$g(\theta^m|y^c;\mathcal{M}) \propto g(\theta^m)\mathcal{L}(y^c|\theta^m;\mathcal{M})$$

In the one step approach, we compute the posterior distribution of the parameters conditional on the raw data, on the DSGE model and on the trend specification, $\mathcal{F}$. Thus, $\mathbb{M} = \{\mathcal{M}, \mathcal{F}_\tau\}$, $Y = y$ and $\theta = (\theta^m, \theta^f)$, and posterior distribution of parameters is

$$g(\theta^m, \theta^f|y;\mathcal{M},\mathcal{F}) \propto g(\theta^m, \theta^f)\mathcal{L}(y|\theta^m, \theta^f;\mathcal{M},\mathcal{F}_\tau)$$

Given the large number of parameters involved, we can not compute analytically the posterior distribution, and we need to use posterior simulators based on Monte Carlo Markov Chain (MCMC) methods. Details on the MCMC methods are reported in Section 1.3 a).

Just a couple of technical remarks for the one step approach. First, a candidate draw $\theta_\dagger = (\theta_\dagger^m, \theta_\dagger^f)$ is rejected, if $\theta_\dagger^m$ implies non-existence or indeterminacy for the system (2.2)-(2.4). Sencond, since the state space generated by the hp-dsge setup is not stationary, we can not use unconditional moments to start the Kalman filter and we need to start from an arbitrary point. I picked $s_{1|0} = [y_1, \mathbf{0}, \mathbf{0}, \mathbf{0}]$ and $\Omega_{1|0} = 10 * I$, to account the uncertainty of my guess.

d Advantages of the one step approach

The advantage of having the joint posterior distribution of structural and filtering parameters, $\theta = (\theta^m, \theta^f)$, is twofold.
First, we can evaluate which trend specifications fits the data better by calculating the relative posterior support, i.e. Posterior Odds ratio, of various specifications. Recall that the Posterior Odds ratio is constructed by comparing the Bayes Factor, which is the ratio of the predictive densities of the data conditional on different models, and prior odds, which is the ratio of prior probabilities associated to each model. The predictive density of the data, Y, conditional on the model, $\mathbb{M}$, for a given prior $g(\theta)$ is

$$p(Y|\mathbb{M}) = \int \mathcal{L}(Y|\theta; \mathbb{M}) g(\theta) d\theta$$

In the one step approach, the predictive density of the data, conditional on the DSGE model, $\mathcal{M}$, and on the trend specification, $\mathcal{F}$, is

$$p(y|\mathcal{M}, \mathcal{F}_\tau) = \int \mathcal{L}(y|\theta; \mathcal{M}, \mathcal{F}_\tau) g(\theta) d\theta$$

where $\theta = (\theta^m, \theta^{f_\tau})$. Therefore, if one wishes to test different trend specifications (say a deterministic, $\mathcal{F}_0$, against a stochastic trend, $\mathcal{F}_1$), the 1s approach allows to compute the Posterior Odds,

$$PO_{\mathcal{F}_0, \mathcal{F}_1} = \frac{g(\mathcal{M}, \mathcal{F}_0)}{g(\mathcal{M}, \mathcal{F}_1)} \times \frac{p(y|\mathcal{M}, \mathcal{F}_0)}{p(y|\mathcal{M}, \mathcal{F}_1)} = \frac{g(\mathcal{F}_0)}{g(\mathcal{F}_1)} \times \frac{p(y|\mathcal{M}, \mathcal{F}_0)}{p(y|\mathcal{M}, \mathcal{F}_1)}$$

where $g(\mathcal{F}_0)$ and $g(\mathcal{F}_1)$ are prior probabilities of each trend specification. With the Posterior Odds ratio and a loss function, one can test trend specifications against each others. In the two step setup, the predictive density of the filtered data, y^c, is

$$p(y^c|\mathcal{M}) = \int \mathcal{L}(y^c|\theta; \mathcal{M}) g(\theta) d\theta$$

with $\theta = \theta^m$. Therefore, one can not test different trend specifications because the ratio between predictive density of data filtered in different way would be meaningless, since the likelihood is computed at different data point.
The second main advantage of this formulation is that we can construct estimates of

the structural parameters that are robust to trend uncertainty. Given that we do not know the 'true' data generating process, trend uncertainty can be accounted for by averaging across specifications. In particular, suppose that one does not know whether the non-stationary component of the data is driven by various trend specifications, $\mathcal{F}_1, \mathcal{F}_2, ..., \mathcal{F}_K$ (for example deterministic, stochastic, with correlation among trends, with common trend components, etc.). Then, one can compute

$$g(\theta^m|y, \mathcal{M}) = \sum_{j=1}^{K} \frac{p(y|\mathcal{M}, \mathcal{F}_j)}{\sum_{k=1}^{K} p(y|\mathcal{M}, \mathcal{F}_k)} \int g(\theta^m, \theta^{f_j}|y, \mathcal{M}, \mathcal{F}_j) d\theta^{f_j}$$

where the filtering parameters of each trend specification, θ^{f_j}, are intergraded out. The resulting structural parameters distribution, $g(\theta^m|y, \mathcal{M})$, is then robust to the trend uncertainty.

e Parameter drifts

One may suspect that, for a subset of times series, trends have changed over the sample. There is no conceptual difficulty in extending the setup we have used to allow trend parameters to be unstable. In the lt-dsge framework one could define the following specification

$$y_t^\tau = A_t + B_t * t + \eta_t \tag{2.9}$$
$$A_{t+1} = A_t + \eta_{t+1}^A \tag{2.10}$$
$$B_{t+1} = B_t + \eta_{t+1}^B \tag{2.11}$$

To test whether the trend has changed over time, one can compute the likelihood of the unstable system and compare it with the likelihood of the stable system using the Posterior Odds and a loss function.
Similarly, in fd-dsge setup we could set

$$y_t^\tau = \gamma_t + \Gamma y_{t-1} + \eta_t \tag{2.12}$$
$$\gamma_{t+1} = \gamma_t + \eta_{t+1}^\gamma \tag{2.13}$$

The likelihood can be computed and the stability of the trend parameters can then be tested.

2.3 Simulated Data: Parameter Bias

The aim of this section is to compare performances of the two methods in a reasonable experimental design. Using simulated data, I compare the estimates of the structural parameters using 1s and 2s methods, and measure the bias induced by the two approaches in three different situations: (1) in small samples, (2) when the trend is misspecified, i.e. the 'true' trend is deterministic and the structural parameters estimated as if it were stochastic, and viceversa, (3) when the assumption about the correlation between trend and cycles is wrong. Overall, the results indicate that the one step approach gives estimates that are less biased on average than two step ones. Moreover, in most of the cases the one step approach is able to recover the true trend generating process. Remarkably, the structural paremeters bias is always statistically significant, meaning that in most of the cases 'deep' parameters are difficult to identify correctly, see Canova and Sala (fortcoming).

a The Data Generating Process

The model I use to generate the cyclical component of the data is the baseline version of the New Keynesian model where, as in Calvo (1983), producers face restrictions in the price setting process, households maximize a stream of future utility and a monetary authority sets the nominal interest rate following a simple Taylor rule. The equilibrium conditions of the prototype economy, where all variables are expressed in log deviations from the steady state, are[4]

$$\lambda_t = \epsilon_t^{\chi} - \sigma_c c_t \tag{2.14}$$

$$y_t = \epsilon_t^a + n_t \tag{2.15}$$

$$mc_t = \omega_t - (y_t - n_t) \tag{2.16}$$

$$mrs_t = -\lambda_t + \sigma_n n_t \tag{2.17}$$

$$\omega_t = mrs_t \tag{2.18}$$

$$r_t = \rho_R r_{t-1} + (1-\rho_R)(\rho_\pi \pi_t + \rho_y y_t) + \epsilon_t^r \tag{2.19}$$

$$\lambda_t = E_t \lambda_{t+1} + r_t - E_t \pi_{t+1} \tag{2.20}$$

$$\pi_t = k_p(mc_t + \epsilon_t^{\mu}) + \beta E_t \pi_{t+1} \tag{2.21}$$

$$\epsilon_t^{\chi} = \rho_\chi \epsilon_{t-1}^{\chi} + \nu_t^{\chi} \tag{2.22}$$

$$\epsilon_t^a = \rho_a \epsilon_{t-1}^a + \nu_t^a. \tag{2.23}$$

In this economy there is no capital accumulation nor government spending, thus output, y_t, is entirely absorbed by consumption, i.e. $c_t = y_t$. Equation (2.14) gives the value for the marginal utility of consumption, λ_t, which depends negatively on consumption since the elasticity of intertemporal substitution, σ_c, is positive. The shadow value of consumption is also hit by a preference shock, ϵ_t^{χ}, which I assume to follow an AR(1) process, equation (2.22). Equation (2.15) is the constant return to scale production function, by which output is produced with labor, n_t. Total factor productivity, ϵ_t^a, is assumed to be a stationary AR(1) process, see equation (2.23). The difference between real wage, ω_t, and the marginal product of labor, $y_t - n_t$, defines the marginal cost, mc_t, equation (2.16). Since labor market is perfectly competitive and frictionless, there is no wage markup and the marginal rate of substitution, mrs_t, is equal to the real wage. The

[4]For further details on the model see the Appendix.

marginal rate of substitution between working and consumption depends positively on hours worked, where σ_n is the inverse of the Frish elasticity of labor supply. Equation (2.19) is the monetary rule. Equation (2.20) is the standard Euler equation and β is the time discount factor. It states that current marginal utility of consumption depends positively on its future expected value and on the ex-ante real interest rate, $r_t - E_t\pi_{t+1}$. Equation (2.21) is the New Keynesian Phillips curve obtained from the forward looking behavior of the firms. The NKP curve is hit by a cost push shock, ϵ_t^μ. The cost push shock is determined by a stochastic parameter that determines the time varying markup in the goods market. The slope of the Phillips curve is $k_p = (1-\zeta_p)\frac{1-\beta\zeta_p}{\zeta_p}$, where ζ_p is the probability of keeping the price fixed. The four exogenous processes are driven by mutually uncorrelated, zero mean innovations, i.e. $\nu_t = [\nu_t^\chi, \nu_t^a, \nu_t^r, \nu_t^\mu]$.
I assume that the cyclical components of GDP, hours worked, real wages and inflation,

$$y_t^c = [y_t, n_t, \omega_t, \pi_t]$$

are determined by the solution to (2.2)-(2.4). The structural parameters of the model, θ^m, are

$$\theta^m = [\beta, \sigma_c, \sigma_n, \rho_R, \rho_\pi, \rho_y, \zeta_p, \rho_\chi, \rho_a, \sigma_\chi, \sigma_a, \sigma_r, \sigma_\mu]$$

and x_t is the vector of endogenous states,

$$x_t = [\lambda_t, mc_t, mrs_t, r_t].$$

Finally, the vectors of exogenous processes and of innovations are respectively

$$z_t = [\epsilon_t^\chi, \epsilon_t^a, \epsilon_t^r, \epsilon_t^\mu]$$

$$\nu_t = [\nu_t^\chi, \nu_t^a, \nu_t^r, \nu_t^\mu].$$

I specify two types of trends: a linear deterministic trend,

$$y_t^\tau = A + Bt + \eta_t$$

and a smooth integrated trend,

$$\begin{aligned} y_{t+1}^\tau &= y_t^\tau + \mu_t \\ \mu_{t+1} &= \mu_t + \zeta_{t+1} \end{aligned}$$

Therefore, the appropriate filters are a linear detrending filter and the unobserved component (Hodrick-Prescott) filter.

b Prior Selection

Table 7 reports the priors selection of the structural parameters. I assumed Beta distribution for those parameters that must lie in the 0-1 interval, like $\rho_R, \zeta_p, \rho_\chi, \rho_a$. I choose a prior mean close to 0.5 for the probability of keeping the prices fixed, whereas the autoregressive parameters in the exogenous processes have prior mean close to 0.7. I employ Gamma or Inverse Gamma distributions for the parameters that must be positive, like the elasticity of consumption and leisure (σ_c and σ_n). For the standard deviations, I use Inverse Gamma with mean close to 0.006 and standard deviation of 0.002. The remaining parameters have normal distributions.

c Bias Computation

I generate data using four different population values, see Table 8. I consider different persistency and volatility of the shocks: 'LP' stands for low persistence, 'HP' for high persistence, 'HV' stands for high volatility, 'LV' for low volatility. For each row of Table 8, I generate two data sets with the types of trend mentioned. Each data set is composed of a vector of four times series of 300 observations; I discarded the first 140 observations and keep last 160 for estimation, which represents 40 years of quarterly data observations. The bias is calculated according to the following algorithm

1. for each simulated dataset, $s = 1, ..., 8$, I run a RWM algorithm as specified in Section 1.3 a) until convergence is achieved[5].

2. I then discard the first 300,000 draws and keep randomly one every 1,000 draws, θ_j^s, and compute

$$bias_\ell^s = \frac{1}{L}\sum_{j=1}^{L} |\frac{\theta_j^s - \theta_{true}^s}{\theta_{true}^s}|$$

   with $L = (N - 300,000)/1,000$ and N is the number of iterations of the RWM.

3. I repeat 2. 100 times and take the average bias, i.e. $BIAS^s = \frac{1}{100}\sum_{\ell=1}^{100} bias_\ell^s$

I am interested only in the bias of the structural parameters estimates, θ^m. Throughout these simulations, the acceptance rate played a crucial role. I observed that the larger was the acceptance rate the larger was the bias; this is quite intuitive if we think that the acceptance rate is inversely related with the variance of the RWM algorithm. Indeed, with a small variance it becomes difficult for the algorithm to explore the entire parameters space and get close to the true values. I tried to keep the acceptance rate between 20% and 35%, as the literature suggests.

[5]Convergence is achieved for all the setups roughly alter 300,000 draws, and the number of iterations is set to 600'000.

d Bias in small samples

Tables 9 reports the bias of the 'deep' parameters estimates for the two methods with a deterministic trend. For the 2s estimates, in the first of the two steps I detrend the data with a linear trend. For the 1s step setup, I used the lt-dsge specification.
On average, the 1s method is superior to the 2s one in terms of parameter bias. In 29 cases out of 48, it turns out that the bias of the two step estimates is larger than the corresponding bias with the 1s setup. Looking at the average bias across DGP (last column of Table 9), one can notice that in 8 cases out of 12 parameters estimates are less biased in the 1s that in the 2s setup. In the 1s setup the most difficult parameters to estimate are the standard deviations, and the corresponding bias is larger for the 1s than for the 2s framework. Despite this, the average bias across parameters (last row of Table 9) is larger in the 2s setup in three cases out of four. When the trend is deterministic the superiority of the one step approach can be explained as follows. The first of the two steps involves OLS estimation of the trend parameters (slopes and intercepts), and the residual of the regression are the cycles. Small sample bias is absorbed by the cycles, and this distorts the structural parameters estimates. In the one step setup, cycles are treated as unobserved states and estimated optimally with the Kalman filter. This reduces the bias of the structural parameters estimates.
Table 10 reports the bias using the two methods, when data is generated with a stochastic trend. In the 2s setup, the first of the two steps uses the Hodrick-Prescott filter with a smoothing parameter of 1'600 to extract the stationary component of the data. The 1s approach seems to be better, in general. In 30 cases out of 48 the bias of the two step set up is larger than the hp-dsge one. Looking at the average bias across DGP, for eight parameters out of twelve the bias is smaller in the 1s than in the 2s setup. Moreover, the average bias across parameters is larger for the 2s in 3 cases out 4. The intuition for this result is straightforward: in the two step case, the ratio between the variance of innovations in trend and the variance of the cycles is fixed to 1'600, which may not reflect the 'true' ratio between trend and cycles variances. In the one step approach, the smoothing parameter is jointly estimated with other parameters. Hence, biases are reduced with a 1s approach.
The relative magnitude of the bias of the two approaches depends on the length of the sample: for larger samples, the differences in bias are smaller. For example, when I repeat the baseline exercise using times series of 500 and 1000 observations (see Table 15), I find that biases are reduced, but they do not disappear. In fact, asymptotic convergence is very slow. Note that, while relative biases are considerably reduced with a deterministic trend, they are still relevant in stochastic framework.

e Bias under misspecifications

One may wonder whether a wrong specification of the trend or incorrect assumptions about its correlation with the cycles could affect the bias of the parameter estimates obtained with the two approaches and in which direction. To examine these issues, I performed Monte Carlo experiments where a) the 'true' trend is deterministic and data are estimated as if it were stochastic (and viceversa), and b) the assumption about the correlation between trend and cycles is wrong. Two interesting results emerge. First, in the one step setup structural parameter estimates are robust regardless of the exact trend specification. Second, wrong assumption about correlation between trend and cycles affects strongly the two step estimates, whereas it leaves one step estimates roughly unchanged.

Table 11 reports the parameter bias when data has a deterministic trend and the one step approach has the 'wrong' trend specification. That is, data is linearly detrended in the first of the two step, whereas the hp-dsge setup is used to estimate parameters in the one step approach. Thus, the 2s setup has the correct trend specification, whereas the 1s framework is misspecified. Despite of this, one step estimates appear to be quite reasonable. In particular, more than half of the parameter estimates are more bias in the 2s approach than in the 1s one. This is due mainly to the fact that the hp-dsge setup a very flexible structure to capture smooth trends and includes as special case a linear trend specification, see Harvey and Jaeger (1993). Similarly, when data are simulated with a stochastic trend and the 1s approach employs a deterministic trend specification, structural parameters estimates do not seem affected much by the wrong trend specifications. Table 12, which reports the bias of both methods when trend is stochastic and 1s has a lt-dsge specification, indicates that in most of the cases parameter biases have not changed and are quite similar to Table 9. The reason for that is mainly due to the fact that the simulated data has clearly upward trend. This makes the linear deterministic trend a reasonable approximation.

As mentioned, the data I used is made of a cyclical and a non stationary component. To identify trend and cycles from the observables one typically assumes that the two are independent. Given that it is not known whether the two are independent or not, I simulate times series imposing a correlation structure between the two, and estimate the parameters as if they were uncorrelated. The aim of this exercise is to see how the procedure performs when there is misspecification in the identifying assumptions. To impose some correlation structure in the simulated data, I distinguish the case in which the trend is deterministic or stochastic. For deterministic trend, I assume that

$$\eta_t = A_1 z_t + v_t \tag{2.24}$$

where v_t is white noise ad A_1 is a non zero matrix. When the trend is stochastic

$$\zeta_t = A_1 z_t + v_t \tag{2.25}$$

As before, I first consider the bias in the estimates when data has deterministic trends and then when data has stochastic ones. Tables 13 reports the bias in the structural parameters estimates for the two methods. Misspecification strongly affects the estimates of the 2s procedure whereas for the 1s case the bias do not change much relative to baseline case. In this respect, the 2 step procedure produces huge bias in estimating σ_χ: in fact, on average the order of bias is 12 times larger in absolute value than the true parameter value. In 33 cases out of 48, 2s estimates are more biased than the corresponding value in the 1s approach. Moreover, notice that for three DGPs out of four the average bias across parameters in the 2s is double the corresponding value for the 1s. The intuition of this result is as follows. Data is generated with equations (2.1)-(2.4), (2.5) and (2.24), and parameters are estimated assuming that the true DGP is given by (2.1)-(2.4), (2.5). In the 2 step set up, we first regress the data on a linear trend and then with the residuals of the regression estimate the structural parameters. The residuals of the regression are stationary; thus, the OLS regression gives consistent estimates of A and B, the slope and the intercept of the linear trend. Hence, the error induced by the omission of equation (2.24) is absorbed by the residuals. This biases the structural parameters estimates. In the lt-dsge estimates, cycles are are treated as unobserved states and estimated jointly with the trend; thus, the bias is evenly split between filtering parameters and 'deep' parameters.
When data is simulated with a stochastic trend, the same conclusion applies. Table 14 suggests that in most cases parameter estimates are less biased in the 1s set up than in the two step one; in particular, in the hp-dsge setup only 16 parameters out of 48 are estimated with a larger bias that the corresponding values estimated with the 2s procedure. Once again the reason for this is that the ratio between the variances of the cycles and the trend is estimated along with the structural parameters in the one step approach.

Finally, it interesting to investigate the ability of the one step approach to recover the 'true' trend. Recall that this can be done using the Posterior Odds and a loss function. To this aim, Table 16 reports the difference between the logarithm of Posterior Odds between lt-dsge and hp-dsge specification, i.e.

$$\ln PO_{lt,hp} = \ln p(y|\mathcal{M}, \mathcal{F}_{lt}) - \ln p(y|\mathcal{M}, \mathcal{F}_{hp})$$

where I assume that the two specifications are equally ex ante probable. For all the setups considered I obtain positive values for $\ln p(y|\mathcal{M}, \mathcal{F}_j)$ with $j = lt, hp$. Thus, when the true trend is deterministic (stochastic), the log of Posterior Odds should be positive (negative). Except in one case (out of 16), the one step approach is able to recover the true trend generating process.

2.4 Actual Data: Parameters Estimates

In this section, I compare estimates of the two approaches using real data. I first present the parameter estimates I obtain for a 'small' New Keynesian model, presented in the section a. This gives us a better understanding of what the two procedures do to the data. In the next sections, I extend the analysis to a more densely parameterized model.

a 1s and 2s Estimates of a Small NK Model

I use quarterly values of GDP, real wages, hours worked and inflation from 1964:1 to 2007:2. Times series are from the FRED database of the Federal Reserve Bank of St. Louis. Hours worked are constructed by multiplying the average hours of production workers times the ratio of total employees over the civilian population. Inflation is calculated annualizing the quarterly growth rate of the producer price index. Prior selection is the one reported in Table 7.

Table 17 reports estimates of the 'deep' parameters using 2s and 1s approaches. In the 2s setup (columns (1),(3),(5)) many parameters estimates are statically different across different filtered data. These large differences are due to the filter used: indeed, each filter extracts cycles with properties statistically different from each other (Canova (1998)). Different cycles determines a different shape for the likelihood function, which implies statistically different estimates. For example, consider the estimates of the autoregressive coefficients. Looking at the cyclical component extracted by the filter, we can notice that linear detrended data are very persistent (see top row of Figure 18) compared to other data transformation. This occurs because a linear detrending filter do not remove entirely the low frequencies in the spectral density representation, and leaves in the spectrum a portion of fluctuations with periodicity larger than 32 quarters. This pushes upward the estimates of the persistence of the exogenous driving forces. At the same time, a persistent processes distorts the agents perception of the shocks of the economy and thus alters their optimality conditions; in particular, a persistent preference shock affects the estimates of the elasticities in the household's intra-temporal optimality condition. The direction of the contamination is not clear because behavioral parameters enters in an non-linear fashion during estimation. Similarly, a first difference filter extracts a very noisy cyclical component (bottom row of Figure 18), which pushes downward the estimates of the autoregressive coefficients and has effects upon the household's decision rules.

Moreover, the amplitude of the cycles affects the magnitude of the structural standard deviations. Comparing the three lines, we can notice that the deepest cycles are the ones given by a linear detrending filter, followed by a first difference filter and by the HP filter. The latter ranking implies a similar ordering in the magnitude of the estimates of structural standard deviations: in fact, the estimates of structural standard deviations are largest using linear detrended data, followed by the estimates obtained with first

difference filtered data and by the estimates obtained with HP filtered data.
Looking at columns (2),(4),(6) which contain 1s estimates, the first thing to notice is that large differences in the parameter estimates of exogenous process reduce. For example, in the 2s approach the range of the median estimates of the autoregressive parameters is 0.51-0.98 for ρ_χ and 0.38-0.98 for ρ_a. In the 1s approach, autoregressive coefficient median estimates vary from 0.57 to 0.79 for ρ_χ, and from 0.48 to 0.85 for ρ_a. In general, in the one step set up median estimates of structural parameters shrink across different trend specifications.
With different structural estimates policy implications are clearly different; for example, impulse responses look distinct. Figure 19 reports the response of GDP, employment, real wages and inflation to a one percent increase in the preference and technology shock using median estimates of the 2s approach. The solid (blue) line represents the response of a variable using linear detrended data (dotted lines give the 90% confidence interval), the dash dotted (green) line the response of a variable using first difference filtered data and the dotted (red) line the response of a variable with hp filtered data. Responses are statistically different: in most of the cases the the median values of the estimates with HP filtered or first differenced data do not fall in the 90% confidence interval of the estimates with linear detrended data. Moreover, notice that the effect of a positive demand shock to wages is completely different according to the filter used: in fact, it induces a positive reply with first difference data or a negative one with linear detrended data.
Given this outcome, which impulse responses should we choose ? Which estimates should we thrust ?
With the traditional two step method we can not answer this question, since do not have a statistical-based criterium to select among different DSGE estimates. The one step approach can easily deal with this question: one could either test trend specifications or construct robust estimates by averaging across trend specifications. The bottom part of Table 17 presents the priors, posterior densities and Posterior Odds for the three different specifications. Posterior Odds are computed with respect to the lt-dsge specification, i.e.

$$\ln PO_{k,lt} = \ln p(y|\mathcal{M}, \mathcal{F}_k) - \ln p(y|\mathcal{M}, \mathcal{F}_{lt})$$

for $k = hp, fd$, where we assume equal ex ante probability to each filter. Differences in posterior density of data are quite large across specifications. Data clearly prefers a specification with unit roots in the long run dynamics. The hp-dsge specification has the lowest posterior data density; in order to choose a smooth integrated trend over a linear trend, we need a prior probability $6.4 * 10^{13}$ (= exp(31.8)) times larger for the hp-dsge specification than the prior probability on lt-dsge setup. Comparing a linear deterministic with a unit root specification, the log of PO clearly reveals the preference of the unit root over a linear deterministic setup. In order to choose a linear over a unit root specification for the trend, we need a prior probability of $3.6 * 10^{42}$ time larger for the lt-dsge specification than the prior probability on fd-dsge setup; therefore, I conclude that the specification with unit root improves considerably the fit relative to a linear trend or a smooth integrated trend specification. Turning to the the question of interest, Figure 20 shows the effect of an increase in the a demand and a supply shock to the variables considered using a 1s approach. Notice that responses and dynamics look more similar across different data transformations in the 1s than in the 2s setup. This is due to the fact that median estimates shrink across trend specifications in the 1s procedure. Given the results in terms of PO, the most likely impulse responses are the ones given by the fd-dsge setup.

b An extension

The extension to a more densely parameterized model is easy to implement. To this aim, I borrow the model of Smets and Wouters (2007) (henceforth SW) with sticky price and wages and with price and wage indexation. Despite the fact that the model is almost identical, I depart from the SW model in two aspects. First, SW assume a labor augmenting deterministic growth rate, γ^t, in the production function, i.e.

$$Y_t(i) = \epsilon_t^a K_t(i)^\alpha [\gamma^t N_t(i)]^{1-\alpha}$$

This implies that the long run dynamics are entirely determined by the parameter γ, which makes GDP, real wages, capital, consumption and investment grow at the same rate in the model. I assume that $\gamma = 1$ and let the long run dynamics be determined by the trend specifications presented in Section b. Second, I consider a simpler version of the Taylor rule, i.e.

$$r_t = \rho_R r_{t-1} + (1 - \rho_R)(\rho_\pi \pi_t + \rho_y y_t) + \nu_t^r$$

The set of equations to be estimated are[6]

$$
\begin{aligned}
y_t &= \alpha\phi_p k_t + (1-\alpha)\phi_p n_t + \phi_p \epsilon_t^a && (2.26)\\
y_t &= \epsilon_t^g + c/yc_t + i/yi_t + r^k * k/yz_t && (2.27)\\
k_t &= k_{t-1}^s + z_t && (2.28)\\
k_t &= \omega_t + n_t - \frac{\psi}{1-\psi} z_t && (2.29)\\
mc_t &= \alpha\frac{\psi}{1-\psi} z_t + (1-\alpha)\omega_t - \epsilon_t^a && (2.30)\\
k_t^s &= (1-\delta)k_{t-1}^s + i/ki_t + i/k\varphi\epsilon_t^i && (2.31)\\
(1+\beta i_p)\pi_t &= \beta E_t\pi_{t+1} + i_p\pi_{t-1} + k_p mc_t + \epsilon_t^p && (2.32)\\
(1+\beta i_\omega)\omega_t &= \omega_{t-1} + \beta E_t(\omega_{t+1} + \pi_{t+1}) + i_\omega\pi_{t-1} + (1+\beta i_\omega)\pi_t - k_\omega\mu\omega_t + \epsilon_t^\omega && (2.33)\\
c_t &= \frac{1}{1+h}(E_t c_{t+1} - hc_{t-1}) + c_1(n_t - E_t n_{t+1}) - c_2(r_t - E_t\pi_{t+1}) + \epsilon_t^b && (2.34)\\
q_t &= -(r_t - E_t\pi_{t+1}) + \frac{\sigma_c(1+h)}{(1-h)}\epsilon_t^b + E_t(q_1 z_{t+1} + q_2 q_{t+1}) && (2.35)\\
i_t &= \frac{1}{1+\beta} i_{t-1} + \frac{\beta}{1+\beta} E_t i_{t+1} + \frac{1}{\varphi(1+\beta)} q_t + \epsilon_t^i && (2.36)
\end{aligned}
$$

Variables without the time subscript are steady state values and with time subscript are deviation from the steady state.
Equation (2.26) is linearized version of the production function, where output, y_t, is produced using capital k_t, and labor n_t; ϕ_p captures 1 plus the fixed cost in production, and α the capital share in the production. Total factor productivity, ϵ_t^a, is assumed to be an AR(1) exogenous technology process, i.e. $\epsilon_t^a = \rho_a\epsilon_{t-1}^a + \nu_t^a$. Equation (2.27) is the feasibility constraint of the economy: it says that the total output is assimilated by an exogenous government spending process, ϵ_t^g, investment, i_t, consumption, c_t, and by a function of the capital utilization rate, z_t. It is assumed that government spending follows an AR(1) process, i.e. $\epsilon_t^g = \rho_g\epsilon_{t-1}^g + \nu_t^g + \rho_{ga}\nu_t^a$. Current capital services used, k_t, are a function of the capital installed in the previous period, k_{t-1}^s, and the degree of capital utilization, equation (2.28). Equation (2.29) is derived from the firm cost minimization, which implies that the rental rate of capital is negatively related to capital-labor ratio and positively with the wage, i.e. $r_t^k = -(k_t - n_t) + \omega_t$. Moreover, the cost minimization by the household implies that the degree of capital utilization is a positive function of the rental rate of capital, i.e. $z_t = \frac{1-\psi}{\psi} r_t^k$. Equation (2.30) gives an expression for the marginal cost, mc_t; indeed, marginal cost is the sum of the real cost of the two factors in production, r_t^k and ω_t, with weights given by their respective share in production, net of the total factor productivity. New installed capital is formed by the flows of investment and the net of depreciation old capital, $(1-\delta)k_{t-1}^s$, equation

[6] Details on the model assumptions and its derivation can be found on the web page of the American Economic Review.

(2.31); moreover, the capital accumulation is hit by the investment-specific technology disturbance ϵ_t^i, which is assumed to follow an AR(1) process, i.e. $\epsilon_t^i = \rho_i \epsilon_{t-1}^i + \nu_t^i$. φ represents the steady state elasticity of the capital adjustment cost function. Equation (2.32) is the New Keynesian Phillips curve which states that current inflation depends positively on past and expected inflation, and on marginal cost. The NPK is also hit by a price markup disturbance, ϵ_t^p, which is assumed to follow an ARMA(1,1) process, i.e. $\epsilon_t^p = \rho_p \epsilon_{t-1}^p + \nu_t^p + \mu_p \nu_{t-1}^p$. The slope of the NKP curve is given by

$$k_p = \frac{(1-\beta\zeta_p)(1-\zeta_p)}{\zeta_p((\phi_p-1)e_p+1)},$$

where β is the time discount factor, ζ_p is the probability of keeping the prices fixed, e_p the curvature of the Kinball goods market aggregator, and the steady state markup, which in equilibrium is itself related to the share of fixed cost in production, $\phi_p - 1$, thought a zero profit condition. Equation (2.33) gives the dynamics of the real wage that moves sluggishly because of the wages stickiness and partial indexation assumption; wages responds to past and future expected real wage, to the (current, past and expected) movements of inflation. Real wage depends also on the wage markup, $\mu\omega_t$, with slope

$$k_\omega = \frac{(1-\zeta_\omega)(1-\zeta_\omega\beta)}{\zeta_\omega((\phi_\omega-1)e_\omega+1)},$$

where $(\phi_\omega - 1)$ is the steady state labor market markup, e_ω the curvature of the labor market Kinball aggregator. The wage markup is itself the difference between the real wage and the marginal rate of substitution between working and consumption, i.e.

$$\mu\omega_t = \omega_t - (\sigma_n n_t + \frac{1}{1+h}(c_t - hc_{t-1}))$$

Wage equation is hit by a wage markup disturbance which is assumed to follow an ARMA(1,1) process, i.e. $\epsilon_t^\omega = \rho_\omega \epsilon_{t-1}^\omega + \nu_t^\omega + \mu_\omega \nu_{t-1}^\omega$. Equation (2.34) is the Euler equation where $c_1 = \frac{(\sigma_c-1)\omega^h n/c}{\sigma_c(1+h)}$ and $c_2 = \frac{1-h}{\sigma_c*(1+h)}$. The Euler equation controls the dynamics of consumption, where current consumption depends on a weighted average of past and expected consumption, expected growth in hours worked, $n_t - E_t n_{t+1}$, and the ex-ante real interest rate, $r_t - E_t \pi_{t+1}$. The dependence on past consumption is controlled by the habit in consumption parameter, h. A disturbance term is assumed to hit the Euler equation and it should be interpreted as a wedge between the interest rate controlled by the central bank and return on asset held by household. Equation (2.35) is the Q equation that gives the value of capital stock, q_t, where $q_1 = \frac{r^k}{r^k+1-\delta}\frac{\psi}{1-\psi}$ and $q_1 = \frac{1-\delta}{r^k+1-\delta}$. It say that the current value of capital stock depends negatively on the real interest rate and positively on expected future value of the capital stock itself and of the real rental rate on capital, $E_t q_1 z_{t+1} = E_t \frac{r^k}{r^k+1-\delta} r_t^k$. Finally, the last equation is the investment equation, (2.36), by which current value of investment depends on past and expected future value of capital and on current value of the stock of capital.

Observables and priors

As in SW, I assume that we observe quarterly values for GDP, hours worked, consumption, investment, real wages, inflation and the nominal interest rate, i.e.

$$y_t = [GDP_t, N_t, C_t, I_t, W_t, \Pi_t, R_t]$$

The cyclical component, y_t^c, of the vector of observed times series evolves according to the system of equations (2.2)-(2.4) where the vector of endogenous state[7] is defined as

$$x_t = [k_t, z_t^k, k_t^p, c_t, i_t, mc_t, \omega_t, \pi_t, r_t, q_t]$$

The system is driven by vectors of exogenous processes and innovation, respectively

$$z_t = [\epsilon_t^a, \epsilon_t^g, \epsilon_t^i, \epsilon_t^r, \epsilon_t^p, \epsilon_t^\omega, \epsilon_t^b]$$
$$\nu_t = [\nu_t^a, \nu_t^g, \nu_t^i, \nu_t^r, \nu_t^p, \nu_t^\omega, \nu_t^b]$$

As in SW, I fix some parameters that might be difficult to identify: depreciation rate, δ, is fixed at 0.025, the exogenous government spending-GDP ratio is set at 18%. Three other parameters are hard to identify: the steady state markup in the labor market, ϕ_ω, which is set to 1.5 and the curvature of the Kinball aggregator in the goods and labor market, e_p and e_w, which are both fixed at 10. Remaining parameters are estimated. Table 18 shows the set of parameters to be estimated: 18 behavioral parameters, 10 autoregressive and moving average coefficients and 7 standard deviations. In additions, I also estimate a number of filtering parameters; 7 for the hp-dsge setup and 14 for the lt-dsge and fd-dsge setup. Priors selection is similar to SW with two exceptions. I assume a rather larger prior standard deviation for price and wage indexation, 0.28 instead of 0.15. Moreover, standard deviations priors have an Inverse Gamma with mean and standard deviation of 0.5.
I use the same database of SW, which is available on the American Economic Review website, and the sample estimated goes from 1966:1 to 2004:4. I run 1,000,000 draws and I tune up the RWM variance in order to achieve a 30%-40% acceptance rate. All the routines are in MATLAB, and it takes about 12 hours to obtain one million draws.

Model implications discussion

The main point I want to stress is that model implications are quite sensitive to data transformations. Figure 21 presents the effects of changes in the exogenous processes to output, hours worked, consumption and investment using 2s estimates. The response

[7] c_t, i_t and ω_t are included in the endogenous states vector because the Uhlig (1999) algorithm recognize as endogenous states variables all the variables that appear out of the expectation equations at time t and $t-1$.

of hours worked to technology shock (first row) has been lively debated. Galí (1999) argued that due to the presence of nominal price rigidities positive productivity shocks leads to an immediate fall in hours. Indeed, the immediate drop in hours is common across trend specifications[8], but dynamics are pretty different. While with first difference data hours need 25 quarters to revert to the steady state, with HP filter data hours worked almost immediately returns to the steady state and is positive for some quarters. The second row reports the responses to a government spending shock. The response of consumption confirms the difficulty of representative agent models[9] to replicate VAR results, where consumption increases after a positive fiscal shock (see Canova and Pappa (2007) or Mountford and Uhlig (2005)). Even thought the immediate reaction of consumption is similar across different data transformations, dynamics are different: in fact, with first difference data it seems that a positive government shock leads to a permanent drop in consumption, which is not the case for linear detrended data. Third raw displays the impulse responses to an investment shock. As in Justiniano and Primiceri (2008), investment specific shock produces positive co-movements of output, investment and hours worked and a complementary behavior of consumption. Even thought the signs of the response are common across data transformation, different dynamics are implied by different trend specifications. Finally, the last row presents the response to a positive preference shock. In this case, not only the dynamics also signs change completely.

Given this outcome, again the two step approach lacks a statistical-based criterion to choose, whereas the one steps approach provides posterior weights to each trend specification, which are reported in Table 19. As before, data strictly prefers a specification with unit root in the trend dynamics. Figure 23 reports the effects of changes in the exogenous processes with the one step approach. Notice that impulse responses are different from the 2s setup; this is because one and two step parameter estimates are different. Given the posterior densities, I conclude that responses with a unit root specification are by far the most likely.

One important implication of the estimated SW model is the little role of technology shocks as a driving forces of business cycles fluctuations. In general, various versions of estimated DSGE models tend to explain the volatility of output manly in terms of mark-up shocks[10], giving thus more importance to nominal innovation rather than real ones. However, in the case of investment-specific shocks a striking contradiction emerges. With a VAR with a long run restriction on the relative price of capital equipment, Fisher (2006) estimates that the investment-specific shock may explain 40-60% of the volatility of output. The two step estimates confirms the predominant role of markup shocks in explaining the GDP volatility. Figure 22 reports the k-step ahead forecast

[8]Consistent with SW, Francis and Ramey (2005), Gali and Rabanal (2004)

[9]Different results for heterogenous agents setups, see Galí et al. (2007)

[10]In this respect, the only exception is Justiniano, Primiceri and Tambalotti (2008) where they found that investment specific shock explain 50% of the unconditional volatility of GDP.

error of GDP in terms of structural shocks; clearly, either price or wage mark-up shocks are the driving forces of GDP fluctuations. This implication changes in the one step setup. According to the way in which we treat the data, the relative importance of structural shocks in explaining output volatility is distinct. In fact, Figure 24 shows that with the hp-dsge setup the main source of output volatility is given by the price mark-up, whereas with the fd-dsge specification shocks to total factor productivity and investment specific shocks explain almost all the variance od GDP. Since data strictly prefers a specification with a unit root in the long run dynamics, I conclude that it is more likely that the GDP volatility can be explained mainly by investment-specific shocks, in line with the recent finding of Justiniano et al. (2008).

2.5 Conclusion

In this chapter, I propose an alternative approach to estimate DSGE structural parameters. Current DSGE estimates involve a two step procedure, where the cyclical component is first extracted from the data and then structural parameters are estimated. The method combines a reduced form representation for the long run dynamics of the data and a structural representation for the cycles so that structural and non structural parameters are jointly estimated.

The methodology has been confronted with current 2 step procedure in reasonable Monte Carlo experiments. Simulation results indicated that the one step approach has desirable properties in small samples. Moreover, the procedure showed to be robust to two types of misspecifications: (a) when the trend specification is wrong, i.e. 'true' trend is deterministic and estimated as if it were stochastic (and viceversa), (b) when there is correlation among trend and cycles and structural parameters are estimated as if they were independent. Moreover, in almost all the cases the one step approach is able to recover the 'true' trend generating process.

When the two approaches are compared with real data, interesting results emerge. Structural parameters estimates and model implications are quite sensitive to the cyclical component extraction. While two step methods lacks a statistical-based criterion to select the most likely data transformation, the one step approach provides a natural benchmark to choose among different trend specification. Finally, applying the two approaches to a medium scale DSGE model different implications arise in terms of sources of GDP volatility at business cycles frequencies. I found that the most likely contribution to GDP volatility is given by investment-specific shocks.

3 Multiple filtering devices for the estimation of cyclical DSGE model

3.1 Introduction

DSGE models have become the paradigm for business cycle and policy analyses in academic and policy circles. Relative to earlier structures, current models are of larger scale and feature numerous frictions on the real and nominal side of the economy that help to closely replicate the dynamic responses that structural VARs produce. A few years ago it was standard to informally calibrate these models but today, increased computing power, longer time series and recent developments in system-wide estimation methods allow researchers to routinely employ a variety of full information techniques in structural estimation exercises (see, e.g., Smets and Wouters (2003)), Ireland (2004) and Rabanal and Rubio-Ramirez (2005) among many others).
Despite the increased popularity, structural estimation faces important conceptual and numerical problems. For example, as emphasized by Canova (forthcoming), full information classical estimation makes sense only if the model is assumed to be the data generating process (DGP) of the observables, up to a set of serially uncorrelated measurement errors. Since such an assumption is hard to entrain unless the model is augmented with ad-hoc dynamics, Fukac and Pagan (forthcoming) suggested to complement standard inference with a more robust limited information analysis. It is also well known that there are abundant population identification problems (see Canova and Sala (fortcoming)), that numerical difficulties are widespread, that singularities are often important (there are typically less shocks than endogenous variables in the model) and that errors-in-variables are present (the variables in the model do not often have a direct counterpart in the data). Finally, the vast majority of the models used in the literature are time invariant and intended to explain only the cyclical portion of the data fluctuations while the actual data contains, at a minimum, growth components, cyclical fluctuations and high frequency noise, all of which may be subject to breaks and other forms of slowly moving variations.
When faced with the problem of fitting stationary cyclical DSGE models to the data, applied investigators typically select a sub-sample where time invariance is more likely to hold, filter the raw data with an arbitrary statistical device, and treat the filtered data as the relevant measure of stationary cyclical fluctuations (see e.g. Smets and Wouters (2003), Ireland (2004)). Occasionally, one find authors, see e.g. Kehoe (2007), suggesting that filtering should be applied to both actual data and data simulated by the model but, to the best of our knowledge, such an approach has, so far, no followers in the estimation literature. Alternatively, a unit root in total factor productivity and/or the price of investment is assumed and the data is filtered using a model-driven transformation (see e.g. Fernandez Villaverde and Rubio Ramirez (forthcoming)).

Both statistical and model-based filtering are problematic. While the profession largely agrees that a cyclical model should explain fluctuations with an average periodicity of 8-32 quarters, there is little agreement on how to obtain these fluctuations from the data and only a partial understanding of the consequences that incorrect or suboptimal filtering induce. For example, it is common to use linearly detrended or first differenced data as input in the estimation process, but such transformations do not isolate fluctuations with the required periodicity (see e.g. Canova (1998)). A band pass (BP) filter which, with infinite amount of data, can exactly isolate the fluctuations of interest, it is typically discarded in the estimation literature because its two-sided nature may change the timing of the data information - a similar argument is made also for the Hodrick and Prescott (HP) filter. In addition, with samples of typical length, all filters induce considerable sampling errors in the estimates of the cyclical component which may compound population misspecification problems. Model-driven filtering, on the other hand, fails to extract cycles with the required periodicity (see Canova (2008)) and requires exact knowledge of the number and the nature of the shocks driving the non-cyclical component. Given our general ignorance on the subject, important specification errors are present also with this approach.

Two additional important issues should be mentioned. First, while researchers filter each series separately prior to estimation, a balanced growth path is often used as a working assumption in theoretical models. Hence, should economic theory or pragmatic considerations guide statistical filtering? Second, while real variables typically show long run drifts, nominal variables just display low frequency fluctuations. Should we filter all the data or only real variables? Conversely, should we treat all the fluctuations present in nominal variables as relevant for parameter estimation or not? Since different researchers choose different methods to filter a portion (or all) of the available data prior to estimation, and since measurement error with unknown properties is introduced regardless of the approach one employs, economic inference is likely to be distorted and the magnitude of the distortions may depend on the transformation employed (see Canova (2008)).

This chapter proposes a method to estimate the structural parameters of a time invariant cyclical DSGE model using noisy and potentially mismeasured cyclical data. The approach borrows ideas from the recent data-rich environment literature (see Boivin and Giannoni (2005)) to set up an estimated structure where vectors of data filtered with alternative procedures are treated as contaminated estimates of the true cyclical component. We set up a signal extraction framework where the cyclical DSGE is the unobservable factor; vectors of filtered data are contaminated observable proxies, and the parameters of the DSGE model are jointly estimated together with the non-structural parameters. This work therefore complements those of Canova (2008), who study how to estimate DSGE models when the cyclical component is not solely located at business cycle frequencies and, conversely, the non-cyclical component may play an important role at business cycle frequencies, and of Ferroni (2008), who suggests ways to test trend specifications in DSGE models and compares the properties of one and two step estimators of its structural parameters.

Our approach is advantageous in, at least, three respects. Since we do not have to arbitrarily choose one filtering method prior to the estimation, or select which shock drives the non-cyclical component, we avoid specification errors. Moreover, our method can be used with cyclical data obtained with one-sided and two-sided filters, of both univariate and multivariate nature, as long as the list of filters is sufficiently rich. Finally, if the cyclical components obtained with different filters are relatively similar, their joint use in the estimation permits the elimination of small sample biases in parameter estimates. If, on the other hand, they are relatively different, measurement error may have different time series properties. Since our signal extraction procedure averages the cross-filter information, it may reduce this error and eliminate part of its cyclicality, making estimates of the cyclical components and of the structural parameters better shielded from filtering errors and inference more robust.

We investigate the properties of our approach using experimental data of the typical length employed in macroeconomics. We show that estimating the structural parameters with cyclical data produced by just one filter typically induces large biases. These biases are considerably reduced with our approach. We also show that the unconditional one step ahead mean square error (MSE) produced by our approach is smaller than the MSE obtained with a standard procedure and that the biases of the latter translates in conditional forecasts which are considerably distorted.

To show that the biases are not only statistically but also economically relevant, we revisit the role of money in transmitting monetary business cycles. The recent literature has largely neglected the stock of money when studying monetary business cycles and Ireland (2004) has shown that such an approach is, by and large, appropriate using US data, standard filtering techniques and a maximum likelihood estimation setup. We show that when the information produced by multiple filters is jointly used in the estimation, real balances statistically matter for the transmission of cyclical fluctuations to output and inflation, both directly and indirectly. Furthermore, we show that the propagation of primitive shocks in the estimated economy differs from the one obtained when only one data transformation is used.

We want to be clear for why we insist on working with time invariant cyclical models, rather than considering structures where cyclical and non-cyclical fluctuations are jointly accounted for. On one hand, writing down reasonable models with these features is hard. In theory, in fact, little is known about mechanisms propagating cyclical shocks at longer frequencies (exceptions are Comin and Gertler (2006) or Canova and Michelacci (2007)) or creating important cyclical implications from long run disturbances. On the other, it is convenient to assume that the mechanisms driving cyclical and non-cyclical fluctuations are distinct and orthogonal. Finally, breaks of various sorts make the data largely uninformative about the features of medium/longer term

cycles.
The rest of the chapter is organized as follows. The next section shows the problems one encounters using a single filtering method to estimate the parameters of DSGE models. Section 3.3 presents our approach and applies it to experimental data. Section 3.4 examines the role of money in transmitting monetary business cycles. Section 3.5 concludes.

3.2 Statistical filtering and structural estimation

To show why statistical filtering induces important measurement errors in the estimated cyclical components and to investigate how these errors affect structural estimates, we simulate data from a textbook New-Keynesian model (see Galí (2008)), where agents face a labor-leisure choice, production is carried out with labor, firms face an exogenous probability of price adjustments and monetary policy is represented with a conventional Taylor rule. The log-linearized equilibrium conditions, in deviation from the steady states, are:

$$\lambda_t = \chi_t - \frac{\sigma_c}{1-h}(y_t - hy_{t-1}) \tag{3.1}$$

$$y_t = z_t + (1-\alpha)n_t \tag{3.2}$$

$$w_t = -\lambda_t + \sigma_n n_t \tag{3.3}$$

$$r_t = \rho_r r_{t-1} + (1-\rho_r)(\rho_\pi \pi_t + \rho_y y_t) + v_t \tag{3.4}$$

$$\lambda_t = E_t(\lambda_{t+1} + r_t - \pi_{t+1}) \tag{3.5}$$

$$\pi_t = k_p(w_t + n_t - y_t + \mu_t) + \beta E_t \pi_{t+1} \tag{3.6}$$

where $k_p = \frac{(1-\beta\zeta_p)(1-\zeta_p)}{\zeta_p} s \frac{1-\alpha}{1-\alpha+\epsilon\alpha}$, λ_t is the Lagrangian on the consumer budget constraint, y_t is output, n_t is hours, w_t is the real wage and r_t the nominal interest rate; z_t is a technology shock, χ_t a preference shock, v_t is a monetary policy shock and μ_t a markup shock. The structural parameters of the model are β, the discount factor, σ_c the risk aversion coefficient, h the coefficient of consumption habit, $1-\alpha$ the share of labor in production, σ_n the inverse of Frisch elasticity, ϵ the elasticity among consumption varieties, ζ_p the probability of changing prices, while ρ_π, ρ_y, ρ_r are parameters of the monetary policy rule. In addition, the parameter vector includes the autoregressive parameters and the standard deviations of the shocks.
For the sake of illustration assume that the preference shock χ_t has two components (a stationary autoregressive and a unit root), the technology shock is a stationary AR(1), and the monetary policy and the markup shocks are iid. In the simulations we set $\beta = 0.99; \sigma_c = 3.00; h = 0.70; \sigma_n = 0.70; \epsilon = 7.0; \alpha = 0.6; \rho_r = 0.2; \rho_\pi = 1.30; \rho_y = 0.05; \zeta_p = 0.8$, and $\rho_\chi = 0.5; \rho_z = 0.8; \sigma_\chi = 0.0112; \sigma_z = 0.0051; \sigma_v = 0.0010; \sigma_\mu = 0.2060$, while the standard deviation of the shock driving the unit root component in preferences is $\sigma_{\chi,nc} = 0.0061$. None of the conclusions we highlight depend on whether the preference

or the technology shock has two components; on whether the non-cyclical component is driven by a unit root or a stochastic linear trend, and on the exact values of the standard deviations of the two components of the preference shocks, as long as $\sigma_{\chi,nc}$ is small relative to σ_{χ} (see Canova (2008)). We simulate 550 data points for four observable variables (y_t, w_t, π_t, r_t), discard 300 initial observations to eliminate the effect of initial conditions, and use the last 100 for forecasting exercises. This means that the sample we use has 150 observations.

Table 20 presents the variability of filtered output and filtered inflation when linear (LT), Hodrick and Prescott (HP), band pass (BP) and first order difference (FOD) filtering are used together with the variability of the true cyclical component of output and inflation. Figure 25 graphs the autocorrelation function of filtered output, of filtered and unfiltered inflation, and of the true cyclical component of the two variables. Clearly, the variability of the cyclical component of the two variables is generally mismeasured, and the serial correlation properties of the two variables are similarly distorted. Moreover, the direction and the magnitude of the distortions is filter and variable dependent - the HP and BP filters significantly underestimate both cyclical variabilities, while LT and FOD over or underestimate the cyclical variability of output only. Finally, despite the fact that the model features a unit root shock, unfiltered inflation is clearly stationary and the filtered and unfiltered inflation series have different persistence. Hence, it is generally going to matter for structural parameter estimation whether the model is fitted to filtered or unfiltered inflation.

A compact way to highlight what each of these filtering transformation do to the data is to plot the difference between log spectrum of filtered output and of filtered inflation and the log spectrum of the cyclical component of these two variables. If one filtering transformation exactly recovers the cyclical component, the difference will be zero at all frequencies. Imperfect isolation in certain bands of the spectrum will be evident when the plot is above or below zero at these frequencies. To facilitate the discussion, we separate frequencies into low, medium (business cycle), and high and, in Figure 26, we separate the frequencies corresponding to cycles of 8-32 quarters from the others with two vertical bars.

Figure 26 shows that all filters imperfectly capture the cyclical component of the data. Moreover, the measurement error is not only located in the high frequencies and its frequency distribution is filter dependent. For example, LT detrended data have much stronger low frequency components than the actual cyclical component while the other three have smaller low frequency components. In addition, all filtered series display significant compression at business cycle frequencies - none of the filter exactly extracts the cyclical component or the business cycle frequencies of the cyclical component. Distortions in the low frequencies of the spectrum are particularly concerning because the perceived income and substitution effects in the economy are different from the true income and substitution effects and this is likely to make estimates of the structural parameters generally distorted.

To show that indeed the errors produced by imperfect filtering distorts our ability to understand the features of the true economy and that the amount of distortions depends on whether some or all variables are filtered, we take the experimental data we have constructed and estimate the structural parameters prefiltering the raw data with LT, HP, BP and FOD filters. Estimation is conducted using Bayesian methods. We choose loose priors for all the parameters (see table 21) and, to give the best chance to the routine, start estimation at the true parameter values. Posterior estimates are obtained with a random walk Metropolis algorithm, where the jumping variable has a t-distribution with 5 degrees of freedom and variance is tuned up to have an acceptance rate of about 30 percent for each filtering approach. Half a million draws were made for each filtered/DGP combination; convergence was checked with a standard CUMSUM statistic and achieved after less than 250000 iterations. We keep one out of hundred of the last 100,000 draws to compute statistics of the posterior distribution.
Table 22 reports the median and the standard deviation of the posterior of each structural parameter. The top panel refers to the situation when all variables are independently filtered prior to estimation. The bottom panel to the case when only output and the real wage are independently filtered.
The table shows that there are important estimation biases and the magnitude of these biases can be larger than 100 percent for some important economic parameters. Since measurement error has important low frequency components, the parameters regulating the relative magnitude of income and substitution effects (the risk aversion coefficient σ_c, the inverse of the Frisch elasticity σ_n, and persistence of the shocks) are all distorted, implying a much stronger income effect that the true data displays. Furthermore, as expected, the distortions in the persistence parameters are larger with filtering procedures that eliminate "too much" low frequency components from the actual data. Finally, for the DGP we consider, distortions obtained when only real variables are filtered are comparable to those obtained when all variables are filtered. This depends on the fact that the cyclical component of inflation (and the nominal rate) is not very persistent. When cyclical inflation has higher persistence, the combination of filtered and filtered data " unbalances" the likelihood - some equations become more misspecified than others - and this makes estimates obtained filtering only part of the data more distorted. Since likelihood based methods produce parameters estimates which minimize the largest discrepancy between the model and the data, biases tend to be more relevant in this case.

How could one eliminate the distortions induced by imperfect filtering? With a fixed sample size, there is not much one can do other than attempt to design filters which are sufficiently flexible to adapt to the (unknown) features of the cyclical components. If the cyclical component of the data was truly located only at business cycle frequencies - Figure 26 shows that this is, in general, not the case - one could try to use filters which are capable to better isolate the frequencies of interest, even in small samples. Filters like these exist in the literature, see e.g. the one suggested by Christiano and Fitzgerald (2003), but their unconventional features (in particular their asymmetry and their time varying weights) may have important and uncontrollable consequences on parameter estimates.

3.3 The idea

Our suggestion is to use the information contained in the cyclical data obtained with different filters to reduce measurement error induced by imperfect filtering, in particular in the low frequencies of the spectrum. In other words, rather than arbitrarily selecting one filter and estimating the model with the resulting filtered data, we treat cyclical data extracted with various filtering methods as contaminated estimates of an unobservable cyclical component and use the information provided by different filters jointly in the estimation of the structural parameters. If the measurement error has similar features across filtering methods, the joint use of multiple cyclical data in the estimation reduces small sample biases in parameter estimates - this is the same idea that the literature on stochastic pooling has emphasized - and make inference more robust. If the measurement error is close to be idiosyncratic across filtering methods, less distortions should be present and more precise estimates of the cyclical features of the economy should be obtained. Hence, our approach also uses ideas of Boivin and Giannoni (2005), who suggest that a data rich environment can help to estimate the structural parameters of a DSGE model and more precisely forecast out-of-sample.
Let the log-linearized solution of a cyclical DSGE model be of the form:

$$x_{1t} = RR(\theta)x_{2t-1} + SS(\theta)x_{3t} \tag{3.7}$$

$$x_{2t} = PP(\theta)x_{2t-1} + QQ(\theta)x_{3t} \tag{3.8}$$

$$x_{3t+1} = NN(\theta)x_{3t} + \iota_t \quad \iota_t \sim (0, \Sigma(\theta)) \tag{3.9}$$

where PP, QQ, RR, SS are time invariant matrices which are functions of the vector of structural parameters $\theta = (\theta_1, \ldots, \theta_k)$, $x_{2t} = \tilde{x}_{2t} - \bar{x}_2$ includes predetermined states, $x_t = \tilde{x}_{1t} - \bar{x}_1$ the endogenous variables, x_{3t} the exogenous disturbances and $\bar{x}_i, i = 1, 2$ are the steady states of $\tilde{x}_{1t}$ and $\tilde{x}_{2t}$. We let $x_t^m = S[x_{1t}, x_{2t}, x_{3t+1}]'$, be a $n \times 1$ vector where S is a selection matrix picking those variables which are observable and interesting from the point of view of the analysis. Even though we suppress the dependence of x_t^m on θ, it should be understood that the data produced by the model is in fact conditional on the choice of θ.
Let x_t^i be the vector of filtered observable time series obtained with method $i = 1, 2, ...g$ and let $x_t^d = [x_t^1, x_t^2, \ldots, x_t^g]'$. Assume the following structure:

$$x_t^d = \lambda_0 + \lambda_1 x_t^m + u_t \tag{3.10}$$

where λ_0 is a $ng \times 1$ vector of constants, λ_1 a $ng \times n$ matrix of non-structural parameters and u_t is a $ng \times 1$ vector of possibly serially correlated errors. For estimation purposes, we normalize the $n \times n$ block $\lambda_1^1 = I$ so that the remaining blocks of the matrix λ_1 can be interpreted as loadings relative to those of the first method. Joint estimation of the structural parameters θ and the non-structural parameters $\lambda_j,\ j = 0, 1$ is now possible because (3.7)-(3.9)and (3.10) represent a state space system with the latter being a measurement equation and the former state equations. Specifically, these equations can be cast into the state space system

$$\begin{aligned} s_{t+1} &= Fs_t + G\omega_{t+1} && (3.11)\\ Y_t &= Hs_t + \eta_t && (3.12) \end{aligned}$$

by setting

$$\begin{aligned}
s_{t+1} &= \begin{pmatrix} x_{1t} & x_{2t} & x_{3t+1} \end{pmatrix}' \\
F &= \begin{pmatrix} 0 & RR & SS \\ 0 & PP & QQ \\ 0 & 0 & NN \end{pmatrix} \\
G &= \begin{pmatrix} 0, & 0, & I \end{pmatrix}' \\
\omega_{t+1} &= \iota_{t+1} \\
Y_t &= \left(x_t^i - \lambda_0^i, i = 1, 2, \ldots, g\right)' \\
H &= \lambda_1 S \\
\lambda_1 &= \text{diag } \left(\lambda_1^i, i = 1, 2, \ldots, g\right)',\ \lambda_1^1 = I \\
\eta_t &= u_t
\end{aligned}$$

The likelihood of (3.11)-(3.12) can be computed with the Kalman filter. In our context the vector of parameters of interest is ν, which includes the structural parameters θ and the non-structural parameters $(\lambda_0^i, \lambda_1^{i+1}, \sigma_\eta^k,\ i = 1, 2, \ldots, g;\ k = 1, \ldots, ng)$. If Bayesian estimation is preferred, the non-normalized posterior distribution of ν, can be obtained with Monte Carlo Markov Chain simulators, see Section 1.3 a.

In (3.10) different cyclical estimates x_t^i are treated as contaminated proxies of the true cyclical component. They are contaminated in two senses: they introduce fluctuations which are non-cyclical, i.e. their periodicity is outside the 8-32 quarter range; they compress the power of the spectrum of the series at cyclical frequencies. The amount of information they contain for the model relevant concepts of cyclical fluctuations is measured by the vector λ_0 and the matrix λ_1. Ideally, λ_0 is a vector of zeros and λ_1 a matrix with the identity in each $n \times n$ block, so that each measure is an unbiased and perfectly correlated although noisy signal of the true cyclical component. In general, we expect either $\lambda_0^i \neq 0$ or $\lambda_1^{i+1} \neq I$, or both, for some or all i's. Since we have normalized

λ_1^1, estimates of λ_1^{i+1} gives us an idea of the amount of correlation distortions each method displays relative to the first.

This setup is advantageous in, at least, three respects. First, since we do not have to arbitrarily choose one filtering approach prior to the estimation or select which shock drives the non-cyclical component, we avoid specification errors. Second, our approach can use as observables cyclical components obtained with one-sided and two-sided filters, both of univariate and multivariate nature and cyclical components obtained assuming that cyclical and non-cyclical components are correlated or not, as long as the list of filters is sufficiently rich. Third, if the cyclical components obtained with different filters are relatively similar, their joint use in the estimation reduces small sample biases in parameter estimates. If, on the other hand, they are relatively different, measurement error may have different time series properties. Since our signal extraction procedure averages the cross-filter information, it may reduce this error and eliminate part of its cyclicality, making estimates of the cyclical components and of the structural parameters better shielded from filtering errors and inference more robust.

It is important to stress that our analysis is conditional on two important assumptions. First, we assume that the model generating x_t^m is correctly specified; that is, there are no missing variables or shocks. When this is not the case, the interpretation of the λ's becomes more difficult and there is no guarantee that the signal extraction approach we describe has better properties than any of the standard approaches. Second, we assume that the cyclical and the non-cyclical components of the data are uncorrelated. While the majority of models used in the literature employs this simplifying assumption, the presence of such a correlation generates additional sources of misspecifications and biases which are neglected here.

The literature is largely silent about the issues we address in this chapter. Cogley (2001) and Gorodnichenko and Ng (2007) are concerned with the problem of estimating the structural parameters of a cyclical DSGE when the trend specification is incorrect, but they do not investigate what are the consequences of small sample filtering nor their implications for structural estimates. Giannone and Reichlin (2006) emphasize that if the variables of the model are measured with error, the solution has a natural factor structure and exploit this feature to compare responses obtained from VAR and factor models. Rather than considering a factor structure for the endogenous variables in terms of the states, we construct an estimable structure where vectors of filtered observable data have a factor structure in terms of a subset of the variables of the model. However, as in Giannone et al., we emphasize that important measurement error with low frequency components may exist. The paper which is closest in spirit to ours is Boivin and Giannoni (2005). Their main point is that the model variables do not have an exact counterpart in the real world and that some external indicators to the model may have important information for interesting variables. The point here is somewhat similar. The cyclical component of the model does not have an exact counterpart in the data because none of the existing filters is able to exactly extract the fluctuations of interest. Moreover, if different cyclical vectors have idiosyncratic error components, this error may be averaged out with our signal extraction approach.

a How does the procedure fare with simulated data?

We estimate the structural parameters of the model using the suggested approach and the same experimental data used in section 2. As input in our procedure, we employ the vector of LT, HP and FOD filtered data. Thus, the vector of observable variables o_t is 12×1 (the model produces implications for four variables and there are three filtering methods). As shown in Figure 26, these three set of data show considerable spectral similarities. Thus, the small sample bias reduction effect will probably be more important than the measurement error reduction effect. As mentioned, the latter will dominate if the vector of filtered data has sufficiently idiosyncratic spectral properties - an example of this situation will be considered in section 4. Nevertheless, since it is unlikely that applied investigators will resort to exotic filtering approaches to generate cyclical components with sufficiently idiosyncratic features, the results we present are sufficient to illustrate the advantages of using multiple filtering devices in the estimation of cyclical models and probably more relevant in practice than those obtained under ideal but impractical conditions.

We employ the same Bayesian approach used in section 2, assuming the same priors on the structural parameters shown in table 21 and loose priors on the non-structural parameters entering (3.10). In particular, we assume that the prior for each element of λ_0 is normally distributed, centered at zero with variance equal to 0.5; the prior for the free diagonal elements of λ_1 is normal, centered at 1 with variance 0.5: and the prior for the standard deviation of the u_t's is inverted gamma with mean equal to 0.0056 and variance equal to 0.002.

We report results obtained with two specifications: one where the non-structural parameters are filter and series specific (in this case there are 32 non-structural parameters to be estimated) and another where the constants and the loadings in (3.10) are common across series for each filter (in this case, there are 17 non-structural parameters). We refer to the first specification as the unrestricted factor model; the second one to the restricted factor model.

The last two columns of table 22 present the posterior median and the posterior standard deviation for the structural parameters obtained with these two specifications, when all variables are filtered (top panel) and when only real variables are filtered (bottom panel). In general, the biases present when only one set of cyclical data is used in the estimation are reduced or eliminated. For example, the risk aversion coefficient and the Frisch elasticity of labor supply are better estimated with reasonably tight standard deviations, and the persistence of the shocks much closer to the true values. The restricted specification appears to be slightly superior when all variables are filtered, but differences are small. Notice also, that the approach we suggest does reasonably well both when all data is filtered and when only real data is filtered. The variability of the structural shocks is still poorly estimated but this outcome is not very surprising since these parameters are weakly identified regardless of which set of cyclical data is used.

To see how these estimates compare with the true ones and with those obtained with standard approaches in terms of economically meaningful statistics, we first compute the unconditional autocorrelation function of the cyclical components of output and inflation, where by this we mean the component generated by the non-unit root shocks, when the posterior median estimates of the parameters obtained when all variables are filtered is used.

Figure 27 confirms the conclusion that parameter estimates obtained with our approach are superior to those obtained with just one set of cyclical data. For cyclical output, the autocorrelation function obtained with our two specifications is close to the true one: at short horizons they are practically identical; at longer horizons differences are small. One can also see that there is a significant difference with the autocorrelation function of cyclical output obtained with standard approaches. For cyclical inflation the improvement over standard methods is less dramatic and the match with the true autocorrelation not as impressive, primarily because true inflation persistence is low. Nevertheless, even in this case, both the restricted and the unrestricted factor approaches reproduce relatively well the first few terms of the true autocorrelation function of inflation.

The good unconditional performance of our approach is also confirmed when looking at the conditional responses of the endogenous variables to the four structural shocks. Figure 28 presents the responses produced with the true parameters, those generated with the posterior median estimates obtained with the restricted factor model and those obtained with FOD filtered data, which are the most appropriate terms of comparison, given the selected DGP. Few features of Figure 28 are worth commenting upon. First, both the shape and the persistence of the conditional responses are well captured by our setup. There is one case where the impact sign is wrong - the response of the real wage to preference shocks - but differences are not large a-posteriori. Second, in comparison, responses obtained with FOD estimates display several distortions. For example, the impact response of output to monetary and markup shocks is too large and the responses of the four variables to preference shocks are poorly captured in terms of sign (see e.g. inflation), magnitude (see e.g. real wages) and persistence (see e.g. nominal rate). Third, it is clear that, in relative terms, our approach trades off a reduced precision in capturing the responses to technology shocks for a better match in reproducing the responses to preference shocks - exactly the opposite of what FOD estimates do.

The statistics we have presented in Figures 27 and 28 are in-sample ones. Since DSGE models are nowadays used in policy institutions for out-of-sample unconditional and conditional forecasting exercises, it is worth examining the out-of-sample performance

of our setup relative to traditional ones. We conduct two forecasting exercises. In the first case, we compute the sequence of one step ahead forecast errors for output and inflation, when we take as parameter values the posterior median estimates obtained when all data are filtered, setting all the shocks in the forecasting period to zero. The MSE is computed over 100 forecasting periods, when no updating of the parameters in the forecasting sample is performed, and appears in table 23.
Figure 29 instead traces out the one-step ahead path of cyclical output and cyclical inflation that would have obtained with posterior median estimates of the parameters when monetary shocks were drawn so as to keep the nominal interest rate fixed over the forecasting path. That is, we allow the nominal interest rate to endogenously react to output and inflation but make sure that the monetary shocks we draw are such that the nominal rate is constant over the forecasting path and equal to the value taken at the date prior to the forecasting period (time 0 in the figure).
Table 23 and Figure 29 indicate that the differences in the estimates shown in table 22 lead to important differences in forecasting performance. Our two specifications are superior to traditional approaches in unconditionally forecasting one-step ahead cyclical output and better in unconditionally forecasting of cyclical inflation. For output the reduction in MSE exceeds 20 percent, for inflation is about 5 percent. Our two specifications are also much better in conditional forecasting. Figure 5 shows that the bias introduced by traditional procedures translates in conditional output forecasts which are consistently different from the true ones while the differences between the forecasts produced by our two specifications and the true ones are statistically insignificant. For inflation differences with standard methods are even more evident since the biases in estimated parameters produce counterintuitive fluctuations, which are absent from the true forecasts.
To conclude, the biases that a standard procedure induces in parameter estimates have important consequences for our understanding of both the unconditional autocorrelation properties of the cyclical component of the data and of the conditional responses to shocks. Overall, both statistics appear to be much better reproduced with the specifications we suggest. The out-of-sample forecasts produced by standard approaches inherit and magnify parameter biases providing a distorted picture of the cyclicality of the variables of interest. These problems are, to a large extent, eliminated when multiple sources of cyclical data are used in the estimation.

3.4 Does money matter in transmitting monetary business cycles?

To show that the additional information our procedure uses may be relevant for understanding important economic phenomena, we reconsider the role of money in transmitting monetary business cycles. The majority of the monetary models nowadays used in

the policy and academic literature attributes a minimal role to the stock of money. In most of the cases these models make no reference whatsoever to monetary aggregates, and when they do, they use a specification where a money demand function determines how much money needs to be supplied, given predetermined levels of output, inflation and the nominal rate. As a consequence, changes in the nominal (and real) quantity of money play no direct or indirect role in shaping the dynamics of output and inflation. Ireland (2004) has constructed a general specification in the class of textbook New Keynesian models where real balances may have a role in influencing the dynamics of output and inflation. He estimated the relevant parameters by likelihood techniques using post 1980 US data and found evidence supporting current theoretical practices. To construct the likelihood of this cyclical model, Ireland first transforms the actual data, taking away a separate linear trend from per-capita GDP and per-capita real balances and demeaning inflation and the nominal interest rate.
In this section, we conduct a similar exercise using post 1959 US data and the cyclical versions of real per-capita output, real per-capita money balances, inflation, and nominal rate series obtained with a number of filtering procedures. As a benchmark, we also estimate the model employing Ireland's preferred transformation over the same sample.

a The model economy

The model is close to the one employed by Ireland (2004), except that it also permits real balances to play an indirect role, via its effects on the interest rate. Relative to the model we have considered in section 2, we allow the real stock of money to potentially matter for the determination of the output and inflation; consider frictions in the form of adjustment costs to changing prices rather than with a Calvo staggered-price device; and set the habit parameter to zero.
Since the economy is quite standard, we only briefly describe its features. There is a representative household, a representative final good producing firm, a continuum of intermediate goods-producing firms supplying the differentiated commodity $i \in [0,1]$ and a monetary authority. At each t the representative household maximizes

$$E_t \sum_t \beta^t \chi_t [U(c_t, \frac{M_t}{p_t e_t}) - \eta n_t] \tag{3.13}$$

where $0 < \beta < 1,\ \eta > 0$, subject to the sequence of budget constraints

$$M_{t-1} + T_t + B_{t-1} + W_t h_t + D_t = P_t c_t + \frac{B_t}{R_t} + M_t \tag{3.14}$$

where c_t is consumption, n_t are hours worked, p_t is the price level, M_t are nominal balances, W_t is the nominal wage and B_t are one period nominal bonds with gross nominal interest rate R_t; T_t are lump sum nominal transfers made by the monetary authority at the beginning of each t, and D_t nominal dividends distributed by the intermediate firms. χ_t and e_t are disturbances to preferences and the money demand whose properties will be described below. Let $m_t \equiv \frac{M_t}{p_t}$ denote real balances and $\pi_t \equiv \frac{p_t}{p_{t-1}}$ the period t gross inflation rate.

The representative final good producing firm uses y_t^i units of intermediate good i, purchased at the price p_t^i to manufacture y_t units of final goods according to the constant return to scale technology $y_t = [\int_0^1 (y_t^i)^{(\epsilon-1)/\epsilon} di]^{\epsilon/(1-\epsilon)}$ with $\epsilon > 1$, where ϵ is the constant price elasticity of demand for each intermediate good. Profit maximization produces demand functions

$$y_t^i = (\frac{p_t^i}{p_t})^{-\epsilon} y_t \tag{3.15}$$

Competition within the sector implies that $p_t = (\int_0^1 (p_t^1)^{1-\epsilon} di)^{1/(1-\epsilon)}$

The intermediate good producing firm i, hires n_t^i units of labor from the representative household to produce y_t^i units of intermediate good i using the production function $y_t^i = z_t n_t^i$, where z_t is an aggregate productivity shock. Intermediate goods substitute imperfectly for one another in producing finished goods. Hence, intermediate firms can set the price of their good but must satisfy (3.15) at the chosen price. We assume a quadratic cost in adjusting prices, measured in finished goods, given by

$$\frac{\phi}{2}(\frac{p_t^i}{\pi^s p_{t-1}^i} - 1)^2 y_t \tag{3.16}$$

where $\phi > 0$ and π^s measures steady state inflation. Optimal prices are chosen to maximize

$$E \sum_t \beta^t \chi_t [U_1(c_t, \frac{M_t}{p_t e_t})](\frac{D_t^i}{p_t}) \tag{3.17}$$

subject to (3.15), where $\beta^t \chi_t U_1(c_t, \frac{M_t}{p_t e_t})$ measures the marginal utility value to the household of an additional unit of profits t and real dividends are

$$\frac{D_t^i}{p_t} = (\frac{p_t^i}{p_t})^{1-\epsilon} y_t - (\frac{p_t^i}{p_t})^{-\epsilon}(\frac{w_t y_t}{z_t}) - \frac{\phi}{2}(\frac{p_t^i}{\pi p_{t-1}^i} - 1)^2 y_t \tag{3.18}$$

The monetary authority sets the nominal interest rate according to

$$R_t = R_{t-1}^{\rho_r} y_{t-1}^{(1-\rho_r)\rho_y} \pi_{t-1}^{(1-\rho_r)\rho_\pi} \Delta M_t^{(1-\rho_r)\rho_m} v_t \tag{3.19}$$

where $\rho_r, \rho_y, \rho_\pi, \rho_m \geq 0$ are parameters and v_t is a monetary policy shock.

The law of motion of the disturbances of the model $d_t = (\chi_t, e_t, z_t, v_t)$ is $\log d_t = \bar{d} + N \log d_{t-1} + \iota_t$, where N is a diagonal matrix with entries $\rho_\chi, \rho_e, \rho_z, 0$, respectively. The covariance matrix of the structural shocks Σ is diagonal with entries $\sigma_\chi^2, \sigma_e^2, \sigma_z^2, \sigma_v^2$. In a symmetric equilibrium all firms make identical choices so $y_t^i = y_t, n_t^i = n_t, p_t^i = p_t, D_t^i = D_t$.

Log-linearizing the model around the steady state produces the following equilibrium conditions

$$\hat{y}_t = E_t\hat{y}_{t+1} - \omega_1((\hat{R}_t - E_t\hat{\pi}_{t+1}) - (\hat{\chi}_t - E_t\hat{\chi}_{t+1})) + \omega_2((\hat{m}_t - \hat{e}_t) - (E_t\hat{m}_{t+1} - E_t\hat{e}_{t+1})) \tag{3.20}$$

$$\hat{m}_t = \gamma_1\hat{y}_y - \gamma_2\hat{R}_t + (1 - (R^s - 1)\gamma_2)\hat{e}_t \tag{3.21}$$

$$\hat{\pi}_t = \beta E_t\hat{\pi}_{t+1} + \psi(\frac{1}{\omega_1}\hat{y}_t - \frac{\omega_2}{\omega_1}(\hat{m}_t - \hat{e}_t) - \hat{z}_t) \tag{3.22}$$

$$\hat{R}_t = \rho_r\hat{R}_{t-1} + (1-\rho_r)\rho_y\hat{y}_{t-1} + (1-\rho_r)\rho_\pi\hat{\pi}_{t-1} + (1-\rho_r)\rho_m(\Delta\hat{m}_t + \hat{\pi}) + \hat{v}_t \tag{3.23}$$

where

$$\omega_1 = -\frac{U_1(c_t, \frac{m_t}{e_t})}{y^s U_{11}(c_t, \frac{m_t}{p_t e_t})} \tag{3.24}$$

$$\omega_2 = -\frac{m^s}{e^s}\frac{U_{12}(c_t, \frac{m_t}{e_t})}{y^s U_{11}(c_t, \frac{m_t}{e_t})} \tag{3.25}$$

$$\gamma_1 = (R^s - 1 + \frac{y^s r^s \omega_2}{m^s})(\frac{\gamma_2}{\omega_1}) \tag{3.26}$$

$$\gamma_2 = \frac{R^s}{(R^s-1)(m^s/e^s)}(\frac{U_2(c_t, \frac{m_t}{e_t})}{(R^s-1)e^s U_{12}(c_t, \frac{m_t}{e_t}) - R^s U_{22}(c_t, \frac{m_t}{e_t})}) \tag{3.27}$$

$$\psi = \frac{\epsilon - 1}{\phi} \tag{3.28}$$

the superscript s denotes steady state values of the variables, U_j is the first derivative of U with respect to argument $j = 1, 2$ and U_{ij} is the second order derivative of the utility function, $i, j = 1, 2$.

The log-linearized Euler condition (equation (3.20)) includes terms involving real money balances and the money demand shocks. They drop out from the expression if and only if utility is separable in consumption and real balances, i.e. $U_{12} = 0$ (see equation (3.25)). Similarly, real balances play a role in the forward looking Phillips curve (equation (3.22)), as long as $\omega_2 \neq 0$, which in turn again implies that $U_{12} \neq 0$ is necessary for real balances to matter. Thus, real balances play a direct role in determining output and inflation if and only if real balances and consumption enter non-separably in the utility function. On the other hand, the posited policy rule implies that the growth rate of nominal balances may be an important determinant of output and inflation indirectly, via interest rate determination. When ω_2 and ρ_m are both zero real balances have no direct or indirect role in propagating cyclical fluctuations.

b Estimation

We estimate the model with quarterly US data spanning the period 1959:1-2008:2. All the data comes from the FRED data bank at the Federal Reserve Bank of Saint Louis and it is seasonally adjusted. For real GDP we take the GDPC96 series, which is a chain weighted real value of domestic production, convert it in per-capita terms dividing it by the civilian non-institutional population, age 16 and over (CNP16OV) and log it. For real balances, we use the stock of M2 (M2SL), divide it by the GDP deflator (GDPDEF), convert it into per-capita terms scaling it by the civilian non-institutional population, age 16 and over and log it. Inflation is calculated annualizing the quarterly growth rate of the GDP deflator and a three months T-bill (TB3M) is our measure of interest rates.

We employ 8 procedures to extract the cyclical component of the data. The first transformation (POLY) fits a second order deterministic polynomial to each series separately, allowing for a change in all the parameters at 1980:3. The cyclical component is the residual in the regression. The second transformation takes the first difference of all the series (FOD) as an estimate of the cyclical component. The third and the forth transformations are obtained with a HP filter with $\lambda = 1600$ and with a BP filter extracting cycles with 8 to 32 quarters periodicity. In this latter case we use the particular implementation of Baxter and King (1994). The fifth transformation is a univariate Beveridge and Nelson decomposition (BN) which fits an ARIMA(1,1,1) model to each series separately and takes as estimate of the cyclical component the difference between the original series and its model-based long run forecast. The sixth transformation is a multivariate version of this procedure (MBN) which fits a VAR with 6 lags to the four variables and takes as an estimate of the cyclical component, the difference between the level of the variables and their long run path implied by the model. The seventh transformation is a classical decomposition (CD) which assumes an additive representation of the components, fits a linear trend to the data and takes the residuals as the cyclical component. Finally, the last transformation employs an unobservable component (UC) decomposition which assumes that the non-cyclical component is a random walk and that the cyclical component has a trigonometric representation (see Canova (2007)). This implies that each of the series has an ARIMA(2,1,0) representation. The cyclical component is then estimated with the projected values of an AR(2) regression of the growth rate of each variable. Note that we have selected these procedures to introduce as much idiosyncratic features in the vectors of observables as possible. In fact, among the procedures we consider there are some where the non-cyclical component is deterministic, some where it is stochastic, and some where it is smooth; some use univariate and other multivariate information; some imply that cyclical and non-cyclical components are independent and some that they are correlated. Finally, some filtering procedures are two-sided and some one-sided.

We estimate the parameters of the model using Bayesian methods. The vector of observables is 32×1 (four series, 8 filtering methods) and the vector of states is 4×1. Since we set $\beta = 0.99$ and steady state inflation to 2 percent, there are 9 structural parameters $(\omega_1, \omega_2, \psi, \gamma_1, \gamma_2, \rho_r, \rho_p, \rho_y, \rho_m)$ - ϵ and ϕ are not separately identifiable - and seven auxiliary parameters $(\rho_\chi, \rho_e, \rho_z, \sigma_\chi, \sigma_e, \sigma_z, \sigma_v)$ to be estimated. We parameterize the link between the model and the cyclical data, allowing one intercept and one slope per filter, independent of the series, but allow the idiosyncratic term to be series and filter dependent. This implies that the intercept measures the average (across series and time) bias of each procedure in constructing the cyclical component and the slope measures the correlation between the data produced by each method and the model based quantities (again, on average across series). Since we normalize the slope of the first procedure to the identity, we have a total of 47 non-structural parameters to be estimated (8 intercepts, 7 slopes and 32 variances). We have also experimented with specifications which restrict the variances of the idiosyncratic component to be either series specific (independent of the filtering method) or filter specific (independent of the series) but discarded them because the model fit is relatively poor.

We draw 500,000 elements of the MCMC chain using the algorithm described in section 3. Convergence was achieved in less than 100,000 draws for each model specification we present. Posterior statistics are computed using one every 10 of the last 200,000 draws. To compare our results, we estimate the parameters of interest using as vector of observables linearly detrending per-capita output and per-capita real balances and demeaned inflation and demeaned nominal rate, allowing for measurement error in each of the four equations. This is the right specification to employ for comparison purposes since our approach has idiosyncratic error built in (3.10).

c The results

Before presenting estimates of the relevant parameters, we briefly comment on the estimates of the non-structural parameters we have obtained. First, the vector of λ_0 is estimated to be zero with very small standard errors. Therefore, all filtered data do not display level biases relative to the cyclical components produced by the model. Second, the loadings parameters are estimated to be between 0.60 (with UC filtered data) and 0.86 (with CD filtered data). Therefore, there appears to be sufficient idiosyncratic information in the cyclical data obtained with the procedures we employ. Since posterior standard errors are tight, bilateral differences in the loadings are generally a-posteriori relevant. Third, the error u_t appears to have a highly idiosyncratic variance, both across series and across filtering methods. This reflects the fact that the variability of each individual series depends on the filtering approach and that vectors of filtered series contain different amount of cyclical information. This is the reason for why, for example, a restricted version of the setup we use, where only one parameter characterizes the variability across series or across filtering methods, produces a poor fit.

Table 24 presents the marginal likelihood of the unrestricted specification, where both direct and indirect effects of money are allowed, and for three restricted specifications, where either the direct effect is eliminated ($\omega_2 = 0$), the indirect effect is eliminated $\rho_m = 0$, or both are eliminated and the estimates of ω_2 and ρ_m obtained in the various cases.

A specification where both effects are present is preferable to the other specifications in terms of in-sample fit. Furthermore, restricting both $\rho_m = 0$ and $\omega_2 = 0$ is preferable to restricting only $\rho_m = 0$. Overall, at least in terms of in-sample fit, both the direct and indirect effects of money are important. This conclusion is confirmed when looking at location measures of the posterior of the two relevant parameters. Statistically, both parameters are estimated tightly and both are a-posteriori different from zero. Economically, our estimates imply that money has a moderate influence on output and inflation fluctuations. On the contrary, the standard specification implies that both the direct and the indirect effect of money are statistically quite small and economically unimportant.

How different are the implications of the model estimated with our procedure relative to the one estimated with a standard approach? Figure 30 presents responses to unitary impulses in the four shocks in the two specifications. Responses look qualitatively similar in the two cases, but there are important differences in the magnitude and the persistence of the responses to shocks. In particular, when our approach is used, the persistence of the responses to monetary shocks is reduced, the persistence of the responses to technology shocks is increased and the responses to money demand shocks have both different magnitude and different persistence.

In sum, with our estimation approach money plays a role in transmitting fluctuations to output and inflation while this is not the case when a standard procedure is used. Given the experimental evidence we have collected in the previous section, it is likely that our estimates display less biases than those obtained with a standard approach, as far as persistence of the shocks and measurement of the substitution and income effects are concerned, reinforcing the conclusion that leaving money out of the model induces important specification and measurement errors.

3.5 Conclusions

This work proposes a method to estimate the structural parameters of a time invariant cyclical DSGE model using multiple sources of cyclical information. The approach borrows ideas from the recent literature employing data-rich environments to estimate DSGE models (see Boivin and Giannoni (2005)), and uses vectors of filtered data obtained with alternative filtering procedures as potentially biased indicators of the true cyclical component. We set up an estimation framework where the cyclical DSGE model is the unobservable factor; vectors of filtered data are contaminated observable proxies; and the parameters of the DSGE model are jointly estimated together with non-structural parameters linking the DSGE model and the observables using signal extraction techniques.

Our approach is advantageous in, at least, two respects. First, since we do not have to arbitrarily choose one filtering approach prior to the estimation, we avoid specification errors. In fact, our approach can use as observables cyclical components obtained with one-sided and two-sided filters, both of univariate and multivariate nature, and cyclical components obtained assuming that cyclical and non-cyclical components are correlated or not, as long as the list of filters is sufficiently rich. The only constraint to the number of the vectors of filtered data used in the estimation is the RAM capacity of the computer and the ability of the investigator to limit the exponential proliferation of non-structural parameters. Second, if the cyclical components obtained with different filters are relatively similar, their joint use in the estimation permits the elimination of small sample biases in parameter estimates. On the other hand, if different filters have sufficiently different features, measurement error may have different time series properties. Since our signal extraction procedure averages the cross sectional information, it may reduce measurement error and eliminate part of its cyclicality, making estimates of the cyclical components more reliable, estimates of the structural parameter better shielded from filtering errors and inference more robust.

Using experimental data, we show that standard approaches employing just one arbitrary filter to estimate structural parameters typically induces large biases in the estimates and that these biases are considerably reduced with our approach. We also show that the estimates obtained with our procedure have superior properties in selected out-of-sample forecasting exercises.

To demonstrate that the biases induced by standard estimation approaches may have relevant economic implications, we revisit the role of money in the monetary business cycle. The literature has largely neglected the stock of money when studying monetary business cycles and Ireland (2004) has shown that such an approach is, by and large, appropriate using US data and a standard estimation setup. We show that when the cyclical information produced by alternative filters is jointly used in estimation, both the direct and the indirect channels through which money propagate fluctuations to output and inflation, are statistically important and economically significant and that the propagation of primitive shocks in the estimated economy is different from the one obtain if only one data transformation is used.

One may wonder why the literature uses time invariant cyclical models in the first place and does not, instead, employ (time varying) models which can jointly explain the cyclical and the non-cyclical properties of the data. We think there are three reasons for why such an approach is currently unfeasible. First, jointly modeling cyclical and non-cyclical fluctuations is an ambitious task since there are few theoretical mechanisms which are able to propagate temporary shocks for a long period of time (we need, for example, R&D, as in Comin and Gertler (2006) or Schumpeterian creative destruction, as in Canova and Michelacci (2007)) or create important cyclical implications from long run disturbances. Second, it is convenient to assume that the mechanism driving growth and cyclical fluctuations are distinct and orthogonal. Third, time varying structures are difficult to deal with in theory and hard to handle computationally (see e.g. Fernandez Villaverde and Rubio Ramirez (forthcoming)).

Given these problems, this chapter provides a simple setup where specification and measurement error biases in the estimates of the parameters of a cyclical DSGE model could be reduced, making inference less prone to arbitrary choices that researchers may make. In this sense, this work complements those of Canova (2008) and Ferroni (2008), who provided new methodologies to reduce specification and small sample errors in the estimation of cyclical DSGE models. Future work in the area will include studying the properties of the procedure using more complex experimental designs and the reconsideration of known puzzles in the macroeconomic literature using the approach proposed here.

Bibliography

An, S. and Schorfheide, F.: 2007, Bayesian Analysis of DSGE Models, *Econometric Reviews* **26**(2-4), 113–172.

Auerbach, A. J.: 2003, Fiscal policy, past and present, *Working Paper 10023*, National Bureau of Economic Research.

Baxter, M. and King, R.: 1994, Measuring Business Cycles: Approximate Band-Pass Filters For Economic Time Series, *The Review of Economics and Statistics* (81), 575–593.

Blinder, A. S. and Solow, R. M.: 1973, Does fiscal policy matter?, *Journal of Public Economics* **2**(4), 319–337.

Boivin, J. and Giannoni, M.: 2005, DSGE estimation with data rich environments, *NBER working paper* .

Bouakez, H., Cardia, E. and Ruge-Murcia, F. J.: 2005, Habit Formation and the Persistence of Monetary Shocks, *Journal of Monetary Economics* **52**, 1073–1088.

Braun, A. R.: 1994, Tax Disturbances and Real Economimc Activity in The postwar United State, *Journal of Monetary Economics* **33**(3), 441–462.

Canova, F.: 1998, Detrending and business cycles facts., *Journal of Monetary Economics* **41**, 475–512.

Canova, F.: 2007, *Applied Macroeconomic Research*, Princeton University Press, Princeton, New Jersey.

Canova, F.: 2008, DSGE estimation with unfiltered data, *Technical report*, manuscript.

Canova, F.: forthcoming, How much scructure in empirical models ?, *Palgrave*, Palgrave Handbook of Econometrics.

Canova, F. and Gambetti, L.: forthcoming, Structural Changes in the US economy: is there a role for monetary policy?, *Journal of Economic Dynamics and Control* .

Canova, F., L. S. D. and Michelacci, C.: 2007, The Labor Market Effects of Technology Shocks, *working paper 6365*, CEPR.

Canova, F., Lopez-Salido, D. and Michelacci, C.: 2007, Schumpeterian technology shocks, *Technical report*, manuscript.

Canova, F. and Pappa, E.: 2006, Does it cost to be Virtuous ? The macroeconomic effects of fiscal constraint, *The Journal of Public Economics* **90**(8), 1391–1414.

Canova, F. and Pappa, E.: 2007, Price Differentials in Monetary Unions: The Role of Fiscal Shocks, *The Economic Journal* **117**(520), 713–739.

Canova, F. and Sala, L.: fortcoming, Back to square one: identification issues in DSGE models, *Journal of Monetary Economics* .

Canzoneri, M. B., Cumby, R. E. and Diba, B. T.: 2002, Should the European Central Bank and the Federal Reserve be concerned about fiscal policy?, *Technical report*, working paper of the Federal Reserve Bank of Kansas City's Symposium on Rethinking Stabilisation Policy.

Chari, V. V., Kehoe, P. J. and McGrattan, E.: 2007, Business Cycle Accounting, *Econometrica* **75**(3), 781–836.

Christensen, I. and Dib, A.: 2008, The Financial Accelerator in an Estimated New Keynesian Model, *Review of Economic Dynamics* **11**(1), 155–178.

Christiano, L. and Fitzgerald, T.: 2003, The Band Pass Filter, *International Economic Review* (44), 435–465.

Cogley, T.: 2001, Estimating and testing rational expectations models when the trend specification is uncertain, *Journal of Economic Dynamics and Control* **25**(10), 1485–1525.

Cohen, D. and Follette, G.: 2000, The automatic fiscal stabilizers: quietly doing their thing, *Finance and Economics Discussion Series 1999-64*, Board of Governors of the Federal Reserve System (U.S.).

Comin, D. and Gertler, M.: 2006, Medium Term Business Cycles, *American Economic Review* **96**(2), 523–551.

Davig, T. and Leeper, E. M.: 2007, Fluctuating Macro Policies and the Fiscal Theory, *American Economic Review* **98**(2), 203–223.

DeJong, D. N., Ingram, B. F. and Whiteman, C. H.: 2000, A Bayesian approach to dynamic macroeconomics, *Journal of Econometrics* **98**(2), 203–223.

Del Negro, M., Schoerfheide, F., Smets, F. and Wouters, R.: 2007, On the fit and forecasting performancde of New Keynesian models, *Journal of Business and Economic Statistics* **25**, 123–143.

Fatas, A. and Mihov, I.: 2006, The macroeconomic effects of fiscal rules in the US states, *Journal of Public Economics* **90**(1-2), 101–117.

Fatás, A. and Mihov, I.: 2001, Government size and automatic stabilizer: International and intranational evidence, *Journal of International Economics* **55**(1), 3–28.

Fernandez Villaverde, J. and Rubio Ramirez, J.: forthcoming, How structural are structural parameter values?, *Macroeconomics annuals*, NBER.

Fernandez-Villaverde, J. and Rubio-Ramirez, J. F.: 2005, Estimating Dynamic Equilibrium Economies: Linear versus Nonlinear Likelihood, *Journal of Applied Econometrics* **20**, 891–910.

Ferroni, F.: 2008, Trend Agnostic Estimation od DSGE models, *Technical report*, Manuscript.

Fisher, J. D.: 2006, The Dynamic Effects of Neutral and Investment-Specific Technology Shocks, *Journal of Political Economy* **114**(3), 413–451.

Francis, N. and Ramey, V. A.: 2005, Is the Technology-Driven Real Business Cycle Hypothesis Dead?, *Journal of Monetary Economics* **52**(8), 1379–1399.

Fukac, M. and Pagan, A.: forthcoming, Limited Information estimation and Evaluation of DSGE models.

Galí, J.: 1994, Governement size and macroeconoimc stability., *European Economic Review* **38**, 117–132.

Galí, J.: 1999, Technology, Employment, and the Business Cycle: Do Technology Shocks Explain the Business Cycle, *American Economic Review* **89**(1), 249–271.

Galí, J.: 2008, *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*, Princeton University Press, Princeton, New Jersey.

Galí, J., López-Salido, J. D. and Vallés, J.: 2007, Understanding the Effects of Government Spending on Consumption, *Journal of the European Economic Association* **16**(3), 309–327.

Gali, J. and Perotti, R.: 2004, Fiscal Policy and Monetary Integration in Europe, *Economic Policy* **37**, 535–572.

Gali, J. and Rabanal, P.: 2004, Technology Shocks and Aggregate Fluctuations: How Well Does the RBS Model Fit Postwar U.S. Data?, *Working Paper 10636*, National Bureau of Economic Research.

Giammarioli, N., Lambertini, L. and Onorante, L.: 2007, Is U.S. Fiscal Policy Optimal?

Giannone, D., S.-L. and Reichlin, L.: 2006, VARs, common factors and the empirical validation of Equilibrium Business cycle models, *Journal of Econometrics* (127), 257–279.

Gordon, D. B. and Leeper, E. M.: 2005, Are Countercyclical Fiscal Policies Counterproductive?, *Nber working papers*, National Bureau of Economic Research, Inc.

Gorodnichenko, Y. and Ng, S.: 2007, Estimation of DSGE models when the data are persistent.

Hansen, G. D.: 1985, Indivisible labor and the business cycle, *Journal of Monetary Economics* **16**(3), 309–327.

Harvey, A. C. and Jaeger, A.: 1993, Detrending, Stylized Facts and the Business Cycle, *Journal of Applied Econometrics* **8**(3), 231–47.

Harvey, A., Trimbur, T. and Dijk, H. v.: 2004, Bayes estimates of the cyclical component in twentieth centruy US gross domestic product, *Technical report*, Basingstoke: Palgrave Macmillan.

Ireland, P.: 2004, Money's role in the Monetary Business cycle, *Journal of Money Credit and Banking* (36), 969–983.

Joines, D. H.: 1981, Estimates of Effective Marginal Tax Rates on Factor Incomes, *Journal of Business* **54**(2), 191–226.

Jones, J. B.: 2002, Has fiscal policy helped stabilize the postwar U.S. economy?, *Journal of Monetary Economics* **49**(4), 709–746.

Justiniano, A. and Primiceri, G. E.: 2008, The Time Varying Volatility of Macroeconomic Fluctuations, *American Economic Review* **98**.

Justiniano, A., Primiceri, G. E. and Tambalotti, A.: 2008, Investment Shocks and Business Cycles, *CEPR Discussion Paper 6739*.

Kehoe, P.: 2007, How to advance theory with structural VARs: the Sims-Cogley-Nason approach, *working paper*, NBER.

McGrattan, E. R.: 1994, The macroeconomic effects of distortionary taxation, *Journal of Monetary Economics* **33**(3), 573–601.

Mendoza, E. G., Razin, A. and Tesar, L. L.: 1994, Effective Tax Rates in Macroeconomics: Cross-Country Estimates of Tax Rates on Factor Incomes and Consumption, *Journal of Monetary Economics* **34**, 297–323.

Mountford, A. and Uhlig, H.: 2005, What are the effects of fiscal policy shocks?, *Discussion Paper 31*, Tilburg University, Center for Economic Research.

Perotti, R.: 2002, Estimating the effects of fiscal policy in OECD countries, *Economics Working Papers 015*, European Network of Economic Policy Research Institutes.

Rabanal, P.: 2006, Inflation Differentials in a Currency Union: A DSGE Perspective, *Working paper*, La Caixa.

Rabanal, P.: 2007, Does inflation increase after a monetary policy tightening? Answers based on an estimated DSGE model, *Journal of Economic Dynamic and Control* **31**, 906–937.

Rabanal, P. and Rubio-Ramirez, J. F.: 2005, Comparing New Keynesian models of the business cycle: A Bayesian approach., *Journal of Monetary Economic* **52**.

Romer, C. D.: 1999, Changes in Business Cycles: Evidence and Explanations, *Journal of Economic Perspectives* **13**(2), 23–44.

Romer, C. and Romer, D.: 1994, What ends recessions ?, *Economics working papers*, NBER Macroeconomics Annuals.

Romer, C. and Romer, D.: 2007, The Macroeconoimc Effects of Tax Change: Estimates Based in a New Measure of Fiscal Shocks, *Economics working papers*, Berckley University.

Schorfheide, F.: 2000, Loss function-based evaluation of DSGE models, *Journal of Applied Econometrics* **15**(6), 645–670.

Smets, F. and Wouters, R.: 2003, An estimated dynamic stochastic general equilibrium model of the Euro Area, *Journal of European Economic Association* **1**(5), 1123–1175.

Smets, F. and Wouters, R.: 2005, Comparing shocks and frictions in US and euro area business cycles: a Bayesian DSGE approach, *Journal of Applied Econometrics* **20**(2), 161–183.

Smets, F. and Wouters, R.: 2007, Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach, *American Economic Review* **97**(3), 586–606.

Taylor, J. B.: 2000, Reassessing Discretionary Fiscal Policy, *Journal of Economic Perspectives* **14**(3), 21–36.

Trimbur, T. M.: 2006, Detrending economic time series: a Bayesian generalization of the Hodrick-Prescott filter, *Journal of Forecasting* **25**(4), 247–273.

Uhlig, H.: 1999, A Toolkit for Analysing Nonlinear Dynamic Stochastic Models Easily, *Computational methods for the study of dynamic economies.*, in Marimon Scott Ed.

A Appendix to Chapter 1

Steady State Analysis

I shall indicate the variable without time subscript as the variable at the steady state. From the Euler equation, (1.6) and (1.7), we get that

$$\begin{aligned} 1 &= \beta R \\ R &= (1-\tau^k)\alpha\frac{Y}{K} + 1 - \delta \end{aligned}$$

Therefore, $1/\beta = (1-\tau^k)\alpha[\frac{K}{Y}]^{-1} + 1 - \delta$. From the production function, equation (2.15), we get

$$\frac{N}{Y} = [\frac{K}{Y}]^{\frac{\alpha}{1-\alpha}}$$

Moreover, from the intertemporal optimality condition, equation (1.5), we get that

$$C^\eta = (1-\tau^w)(1-\alpha)\frac{Y}{N}$$

At the non stochastic steady state, the exogenous process are identical to 1, $X = G = V = A = 1$. Thus the law of motion for capital, equation (1.4), becomes

$$I = \delta K$$

and the feasibility constraint, equation (1.13), is

$$\frac{1}{Y} = 1 - \delta\frac{K}{Y} - C\frac{1}{Y}$$

which can be rewritten as

$$Y = \frac{1+C}{1-\delta\frac{K}{Y}}$$

Finally, we can obtain the debt-gdp ratio form the government budget constraint, equation (1.10),

$$\frac{1}{Y} + \frac{B}{Y}(R-1) = (1-\alpha)\tau^w + \alpha\tau^k$$

Therefore,

$$\frac{B}{Y} = \frac{(1-\alpha)\tau^w + \alpha\tau^k - \frac{1}{Y}}{(1-\tau^k)\alpha[\frac{K}{Y}]^{-1} - \delta}$$

Log-linearization of the equilibrium conditions.

Except for some cases (taxes and interest rate), I denote the log deviation of a variable X_t from its steady state path, X (without time subscript), with small letter, i.e.

$$x_t = \ln(X_t/X).$$

The production function is given by

$$Y_t = A_t K_{t-1}^{\alpha} N_t^{1-\alpha}$$

At the non stochastic steady state we have that the variables are constant and the shock are zero; the log linear version is obtained by dividing the equation by its value at steady state and by taking logarithm; i.e.

$$y_t = a_t + \alpha k_{t-1} + (1-\alpha)n_t$$

F.O.C. of the household are equations (1.5)-(1.8). The F.O.C. calculated at the non stochastic steady state (where the shock are shut down to zero) are:

$$C^{\eta} = (1-\tau^w)(1-\alpha)\frac{Y}{N} \tag{A.1}$$

$$1 = \beta[\alpha(1-\tau^k)\frac{Y}{K} + 1 - \delta] \tag{A.2}$$

$$1 = \beta[1 + r^b] \tag{A.3}$$

To log linearize the system lets apply a first order Taylor expansion to equation (1.5), we get:

$$\eta C^{\eta-1} C c_t = -(1-\alpha)\frac{Y}{XN}\tau^w \widehat{\tau}_t^w + (1-\tau^w)\frac{1-\alpha}{XN} Y y_t -$$
$$(1-\tau^w)\frac{(1-\alpha)Y}{N^2 X} N n_t - (1-\tau^w)\frac{(1-\alpha)Y}{X^2 N} X \chi_t$$

where I denote with $\widehat{\tau}_t^w$ the log deviation of the tax on labor income from its steady state level, τ^w, i.e. $\widehat{\tau}_t^w = \ln \frac{\tau_t^w}{\tau^w}$. The latter equation can be rewritten as

$$\eta C^{\eta} c_t = -(1-\tau^w)(1-\alpha)\frac{Y}{XN}\frac{\tau^w}{1-\tau^w}\widehat{\tau}_t^w + (1-\tau^w)(1-\alpha)\frac{Y}{XN} y_t -$$

$$-(1-\tau^w)(1-\alpha)\frac{Y}{XN} n_t - (1-\tau^w)\frac{(1-\alpha)Y}{XN}\chi_t$$

using (A.1), the expression simplifies to

$$\eta c_t = -\frac{\tau^w}{1-\tau^w}\widehat{\tau}_t^w + y_t - n_t - \chi_t$$

The log linearized version of the Euler equation can be derived by applying a first order approximation to equation (1.6) and (1.7); we get

$$0 = \beta E_t\{\frac{C^\eta}{C^\eta} R v_t + \eta \frac{C^{\eta-1}}{C^\eta} RC c_t - \eta \frac{C^\eta}{C^{\eta+1}} RC c_{t+1} + \frac{C^\eta}{C^\eta} R \widehat{r}_{t+1}\}$$

where $\widehat{r}_t = \ln \frac{R_t}{R}$. Using the fact that $1 = \beta R$ the previous equation can be simplified to

$$0 = E_t\{v_t + \eta(c_t - c_{t+1}) + \widehat{r}_{t+1}\} \tag{A.4}$$

Next lets loglinearize the interest rate equation, (1.7)

$$R\widehat{r}_t = -\alpha \frac{Y}{K}(1-\tau^k)\frac{\tau^k}{1-\tau^k}\widehat{\tau}_t^w + (1-\tau^k)\alpha\frac{1}{K} Y y_t - (1-\tau^k)\alpha \frac{Y}{K^2} K k_{t-1} - (1-\delta) v_{t+1}$$

$$R\widehat{r}_t = \alpha(1-\tau^k)\frac{Y}{K}(y_t - k_{t-1} - \frac{\tau^k}{1-\tau^k}\widehat{\tau}_t^k) - (1-\delta)v_t$$

Using the steady state equation for the interest rate, i.e. $\mu = \alpha(1-\tau^k)\frac{Y}{K} = R - 1 + \delta$, we get

$$\widehat{r}_t = \frac{\mu}{\mu + 1 - \delta}(y_t - k_{t-1} - \frac{\tau^k}{1-\tau^k}\widehat{\tau}_t^k) - \frac{1-\delta}{\mu+1-\delta} v_t \tag{A.5}$$

and combining the two equations, i.e. (A.5) and (A.4), we get

$$0 = E_t\{v_t + \eta(c_t - c_{t+1}) + \frac{\mu}{\mu+1-\delta}(y_{t+1} - k_t - \frac{\tau^k}{1-\tau^k}\widehat{\tau}_{t+1}^k) - \frac{1-\delta}{\mu+1-\delta} v_{t+1}\}$$

Recall the government budget constraint

$$G_t + (1 + r_t^b) B_{t-1} = \tau_t^w w_t N_t + \tau_t^k r_t K_{t-1} + B_t$$

$$G_t + (1 + r_t^b) B_{t-1} = [(1-\alpha)\tau_t^w + \alpha\tau_t^k] Y_t + B_t \tag{A.6}$$

where the latter equality is obtained by substituting the optimality condition of the firm. Rewriting it in terms of the debt-GDP ratio B_t/Y_t

$$\frac{G_t}{Y_t} + (1 + r_t^b)\frac{B_{t-1}}{Y_{t-1}}\frac{Y_{t-1}}{Y_t} = (1-\alpha)\tau_t^w + \alpha\tau_t^k + \frac{B_t}{Y_t} \tag{A.7}$$

I shall indicate with by_t the logarithm deviation of debt-GDP ratio from its steady state at time t, i.e.

$$by_t = \ln \frac{B_t/Y_t}{B/Y}$$

The log linear version is derived as follows,

$$\frac{G}{Y} g_t - \frac{G}{Y} y_t + (1+r^b)\frac{B}{Y} y_{t-1} - (1+r^b)\frac{B}{Y} y_t + (1+r^b)\frac{B}{Y} by_{t-1} + \frac{B}{Y} r^b \widehat{r}_t^b =$$

$$= (1-\alpha)\tau^w \widehat{\tau}_t^w + \alpha\tau^k\widehat{\tau}_t^k + \frac{B}{Y} by_t$$

where $\widehat{r}_t^b = \log(\frac{r_t^b}{r^b})$. Moreover, by the no arbitrage condition we know that

$$1 + r_t^b = V_{t-1} R_t$$

The log linearized version of the no-arbitrage condition gives the following

$$r^b \widehat{r}_t^b = v_{t-1} + R\widehat{r}_t = v_{t-1} + \mu(y_t - k_{t-1} - \frac{\tau^k}{1-\tau^k}\widehat{\tau}_t^k) - (1-\delta)v_t \tag{A.8}$$

where the last equality follows from equation (A.5). We can get rid of the bond interest rate and express the government budget constraint in terms of GDP, debt-GDP, taxes and capital. The budget constraint becomes

$$\frac{G}{Y}g_t - \frac{G}{Y}y_t + (\mu+1-\delta)\frac{B}{Y}y_{t-1} - (\mu+1-\delta)\frac{B}{Y}y_t + (\mu+1-\delta)\frac{B}{Y}by_{t-1} + \frac{B}{Y}\mu(y_t - k_{t-1} - \frac{\tau^k}{1-\tau^k}\widehat{\tau}_t^k) =$$

$$= (1-\alpha)\tau^w\widehat{\tau}_t^w + \alpha\tau^k\widehat{\tau}_t^k + \frac{B}{Y}by_t + \frac{B}{Y}[(1-\delta)v_t - v_{t-1}]$$

Simplifying and rearranging the terms we obtain

$$\frac{G}{Y}g_t + \lambda_4 y_{t-1} + \lambda_4 by_{t-1} = \frac{B}{Y}by_t + (1-\alpha)\tau^w\widehat{\tau}_t^w + \lambda_1\widehat{\tau}_t^w + \lambda_2 y_t + \lambda_3 k_{t-1} + \frac{B}{Y}[(1-\delta)v_t - v_{t-1}]$$

where

$$\begin{aligned}
\lambda_1 &= \alpha\tau^k + \frac{B}{Y}\mu\frac{\tau^k}{1-\tau^k} \\
\lambda_2 &= \frac{G}{Y} + \frac{B}{Y}(1-\delta) \\
\lambda_3 &= \frac{B}{Y}\mu \\
\lambda_4 &= \frac{B}{Y}(\mu+1-\delta)
\end{aligned}$$

Alternatively we can define a fiscal policy rule with debt instead of the debt-gdp ratio, as the following specifications

$$\begin{aligned}
\tilde{\tau}_t^w &= \varphi_w\tilde{\tau}_{t-1}^w + \varphi_b\widehat{b}_t + \varphi_y y_t + \xi_t^w \\
\tilde{\tau}_t^k &= \psi_k\tilde{\tau}_{t-1}^k + \psi_b\widehat{b}_t + \psi_y y_t + \xi_t^k
\end{aligned}$$

Accordingly, we have to log linearize the government budget constraint in terms of debt, equation (A.6), and we get

$$\frac{G}{Y}g_t + (1+r^b)\frac{B}{Y}b_{t-1} + \frac{B}{Y}r^b\widehat{r}_t^b =$$

$$= (1-\alpha)\tau^w[\widehat{\tau}_t^w + y_t] + \alpha\tau^k[\widehat{\tau}_t^k + y_t] + \frac{B}{Y}by_t$$

Substituting equation (A.8) into the latter, we get

$$\frac{G}{Y}g_t + (1+r^b)\frac{B}{Y}b_{t-1} + \frac{B}{Y}v_{t-1} = (1-\alpha)\tau^w\widehat{\tau}_t^w + \alpha\tau^k\widehat{\tau}_t^k + \mu\frac{B}{Y}k_{t-1} + \lambda y_t + \frac{B}{Y}b_t + (1-\delta)\frac{B}{Y}v_t$$

Finally, the feasibility constraint is

$$Y_t = I_t + C_t + G_t$$

the log linear version

$$\begin{aligned} Yy_t &= Ii_t + Cc_t + Gg_t \\ y_t &= \frac{I}{Y}i_t + \frac{C}{Y}c_t + \frac{G}{Y}g_t. \end{aligned}$$

The law of motion for capital is

$$V_tI_t = K_t - (1-\delta)K_{t-1},$$

the log linear version is

$$IVi_t + IVv_t = Kk_t - (1-\delta)Kk_{t-1}. \tag{A.9}$$

At the non stochastic steady state the investment specific shock is shut down to zero, then $V = 1$.

Tax Series construction.

I use quarterly values for real series of gdp, consumption, hours worked, investment, government debt and government spending. Except taxes, all the times series are taken from the FRED database (Federal Reserve Bank of St. Louis http://research.stlouisfed.org/fred2/); average tax rate are constructed from the times series of the Bureau of Economic Analysis (www.bea.gov). The hours worked are constructed as follows: I took the average weekly hours (Average Weekly Hours of Production Worker) and I normalized to 1 unit measure. Then, I multiply the series for the level of employment (All Employees), and divide by the population (Total Population). The series of investment is the sum of Fixed Investment plus Durable Consumption. Government spending is the real Government Consumption series.

To calculate the average tax rates, I follow closely Jones (2002) and Mendoza, Razin and Tesar (1994). All these items are indexed by table and line number. I start with finding τ^p, the average personal income tax rate:

$$\tau^p = \frac{FIT + SIT}{W + PRI/2 + CI}$$

$$CI = PRI/2 + RI + CP + NI$$

where
FIT = Federal Income taxes (3.2: line 3);
SIT = State and local income taxes (3.3: line 3);
W = Wages and salaries (1.14: line 5);
CI = Capital income;
PRI = Proprietor's income (1.14: line 13);
RI = Rental income (1.14: line 17);
CP = Corporate profits (1.14: line 11);
NI = Net interest (1.14: line 25).

As discussed by Joines (1981), the division of proprietor's income into capital and income is somewhat arbitrary. Joines analyzes both cases, I follow Jones (2002), who takes 'a middle ground' and splits proprietor's income evenly between capital and labor income. The labor tax rate, tax^w, is then calculated as

$$tax^w = \frac{\tau^p(W + PRI/2) + CSI}{PRI/2 + EC}$$

where
CSI = Total contributions to social insurance (3.1: line 7);
EC = Total employee compensation (1.14: line 4).

In addition to wages and salaries, employee compensation includes contribution to social insurance and untaxed benefits. Tax capital rate is calculated as

$$tax^k = \frac{\tau^p CI + CT + PT}{PT + CI}$$

where
CT = Corporate taxes (3.1: line 5);
PT = Property taxes (3.3:line 9).

B Appendix to Chapter 2

New Keynesian Model

The model is a sticky price model where as in Calvo (1983) producers face restriction in the price setting process. An accurate description about price-setting assumption can be found in Smets and Wouters (2003), or for a comprehensive overview of New Keynesian models see Galí (2008).
The representative household has a preference for variety: the consumption index is

$$C_t = \left(\int_0^1 C_t(j)^{\frac{\epsilon_t - 1}{\epsilon_t}} dj \right)^{\frac{\epsilon_t}{\epsilon_t - 1}} \tag{B.1}$$

where $C_t(j)$ is the consumption of the good produced by firm j. As in Smets and Wouters (2003), we assume that ϵ_t is a stochastic parameter that determines the time varying markup in the goods market. Shock to this parameter will be interpreted as 'cost-push' shock to inflation equation. We assume that $\mathcal{M}_t \equiv \frac{\epsilon_t}{\epsilon_t - 1}$ is the price markup and

$$\mathcal{M}_t = \mu e^{\epsilon_t^\mu}$$

where $\epsilon_t^\mu \sim N(0, \sigma_\mu^2)$. The maximization of C_t w.r.t. $C_t(j)$ for a given total expenditure leads to a set of demand function of the type

$$C_t(j) = \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon_t} C_t \tag{B.2}$$

where $P_t(j)$ is the price of the good produced by firm j. Moreover, the appropriate price deflator is given by

$$P_t = \left(\int_0^1 P_t(j)^{1-\epsilon_t} dj \right)^{\frac{1}{1-\epsilon_t}}$$

Conditional on such optimal behavior, it will be true that $P_t C_t = [\int_0^1 P_t(j) C_t(j) dj]$. The representative household faces standard intertemporal decisions by choosing a stream of consumption and leisure.

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[X_t \frac{1}{1-\sigma_c} C_t^{1-\sigma_c} - \frac{1}{1+\sigma_n} N_t^{1+\sigma_n} \right] \tag{B.3}$$

A demand shifter is assumed: X_t affects the consumption-leisure intertemporal trade-off. We assume that the process is exogenous and (in logs) follows AR(1), i.e.

$$\epsilon_t^\chi = \rho_\chi \epsilon_{t-1}^\chi + \epsilon_t^\chi$$

where $\epsilon_t^\chi = \ln X_t$ and $\epsilon_t^\chi \sim N(0, \sigma_\chi^2)$. Household maximizes its objective function subject to the intertemporal budget constraint,

$$P_t C_t + b_t B_t = B_{t-1} + W_t N_t \tag{B.4}$$

Household holds its financial wealth in the form of bonds, B_t. Bonds are one period securities with price b_t. W_t is nominal wage and N_t is hour worked; current income is the sum of labor income and bond income. Current income can be either consumed either used to buy bonds. Once having transformed the nominal budget constraint into real terms (dividing by P_t), the first order conditions are:

$$0 = X_t C_t^{-\sigma_c} - \mathcal{L}_t \tag{B.5}$$

$$0 = -N_t^{-\sigma_n} - \mathcal{L}_t \frac{W_t}{P_t} \tag{B.6}$$

$$1 = E_t \left[\ \beta \tfrac{\mathcal{L}_{t+1}}{\mathcal{L}_t} \tfrac{P_{t+1}}{P_t} R_t \ \right] \tag{B.7}$$

where $\mathcal{L}_t$ is the lagrange multiplier associated to the budget constraint and R_t is the gross nominal rate of return on bonds ($R_t = 1 + i_t = 1/b_t$).

We assume a continuum of firms, indexed by $j \in [0,1]$, each of which produces a differentiated good. They all face the same technology,

$$Y_t(j) = A_t N_t(j) \tag{B.8}$$

where A_t is an exogenous technology process which (in logs) follows AR(1), i.e.

$$\epsilon_t^a = \rho_a \epsilon_{t-1}^a + \nu_t^a$$

where $\epsilon_t^a = \ln A_t$ and $\nu_t^a \sim N(0, \sigma_a^2)$. Following the formalism proposed by Calvo (1983), each firm may reset its price only with probability $1 - \zeta_p$ in any given period, independently of time elapsed since last adjustment. The above environment implies that the aggregate price dynamics are described by

$$\Pi_t^{1-\epsilon_t} = \zeta_p + (1 - \zeta_p)(P_t^* / P_{t-1})^{1-\epsilon_t} \tag{B.9}$$

The latter equation implies that in a zero inflation steady state ($\Pi = 1$), we must have $P_t^* = P_{t-1} = P_t$. A firm reoptimizing in period t will choose a price P_t^* that maximizes the current market value of the profits generated while that price remains effective. Formally, it solves

$$\max_{P_t^*} \sum_{k=0}^{\infty} \zeta_p^k E_t Q_{t,t+k} \left[\ P_t^* Y_{t+k|t} - TC_{t+k}(Y_{t+k|t}) \ \right]$$

subject to the sequence of demand constraints

$$Y_{t+k|t} = \left(\ \tfrac{P_t^*}{P_{t+k}} \ \right)^{-\epsilon_{t+k}} Y_{t+k}$$

for $k = 0, 1, 2, ...$ where $Q_{t,t+k} \equiv \beta^k (C_{t+k}/C_t)(P_t/P_{t+k})$ is the stochastic discount factor for nominal profits, $TC(.)$ is the total cost function, and $Y_{t+k|t}$ denotes output in period

$t+k$ for a firm that last reset its price in period t. The first order conditions associated with the above program is

$$\sum_{k=0}^{\infty} \zeta_p^k E_t Q_{t,t+k} Y_{t+k|t} \left[\; P_t^* - \mathcal{M}_{t+k} MC_{t+k|t}^n) \; \right] = 0$$

where $MC^n(.)$ is the nominal marginal cost; recall that ϵ_t is the elasticity of substitution between varieties and $\mathcal{M}_t \equiv \frac{\epsilon_t}{\epsilon_t - 1}$ is the price markup. Rewriting in real terms and with $\Pi_{t,t+k} \equiv P_{t+k}/P_t$

$$\sum_{k=0}^{\infty} \zeta_p^k E_t Q_{t,t+k} Y_{t+k|t} \left[\; \frac{P_t^*}{P_{t-1}} - \mathcal{M}_{t+k} MC_{t+k|t} \Pi_{t-1,t+k}) \; \right] = 0 \tag{B.10}$$

Market clearing conditions in the goods and labor market require

$$Y_t(j) = C_t(j)$$
$$N_t = \int_0^1 N_t(j) dj$$

Moreover, letting the aggregate output be defined as $Y_t \equiv \left(\int_0^1 Y_t(j)^{\frac{\epsilon_t - 1}{\epsilon_t}} dj \right)^{\frac{\epsilon_t}{\epsilon_t - 1}}$ we have that

$$C_t = Y_t$$

Log linearized Equilibrium Conditions

In what follows we shall denote small letter variables as log deviations from the steady state. The household's optimality conditions, (B.5)-(B.7), are linearized by taking first order Taylor approximation around the steady state.

$$0 = \epsilon_t^{\chi} - \sigma_c c_t - \lambda_t$$
$$0 = \omega_t + \sigma_n n_t - \lambda_t$$
$$0 = E_t[\lambda_{t+1} - \lambda_t + r_t - \pi_{t+1}]$$

where $\omega_t \equiv \ln W_t/P_t - \ln W/P$ is the log deviation of the real wage from its steady state. From the market clearing condition we have that $c_t = y_t$. The log linearization of the production function leads to

$$y_t = \epsilon_t^a + n_t$$

The firm's marginal cost is defined as the difference between the real wage and the marginal product of labor, $MPL_t = A_t = Y_t/N_t$

$$mc_t \equiv \omega_t - mpl_t = \omega_t - y_t + n_t$$

The log linearization of the optimal behavior of the firms, (B.10), leads to

$$p_t^* - p_{t-1} = (1-\beta\zeta_p)\sum_{k=0}^{\infty}(\beta\zeta_p)^k E_t[\epsilon_{t+k}^{\mu} + mc_{t+k|t} + p_{t+k} - p_{t-1}]$$

where we used the fact that at the steady state $\mathcal{M} = 1/MC$, $Q_{t+k,t} = \beta^k$, $\frac{P_t^*}{P_{t-1}} = 1$ and $\Pi_{t-1,t+k} = 1$. Since there are constant return to scale,

$$mc_{t+k|t} = mc_{t+k}$$

Plugging the latter and rearranging terms we obtain

$$p_t^* - p_{t-1} = \sum_{k=0}^{\infty}(\beta\zeta_p)^k E_t(1-\beta\zeta_p)[\epsilon_{t+k}^{\mu} + mc_{t+k}] + \pi_{t+k}$$

Given that $(1-\beta\zeta_p) < 1$, the latter equation can be rewritten as a difference equation

$$p_t^* - p_{t-1} = \beta\zeta_p E_t(p_{t+1}^* - p_t) + (1-\beta\zeta_p)[\epsilon_t^{\mu} + mc_t] + \pi_t \tag{B.11}$$

The log linearization of law of motion of price, (B.9), leads to

$$\pi_t = (1-\zeta_p)(p_t^* - p_{t-1})$$

Combining the latter equation with (B.11) we obtain the new Keynesian Phillips curve,

$$\pi_t = \beta E_t \pi_{t+1} + \kappa_p[mc_t + \epsilon_t^{\mu}]$$

where

$$\kappa_p = \frac{(1-\beta\zeta_p)(1-\zeta_p)}{\zeta_p}$$

Finally, we assume that there is a monetary authority that sets the nominal interest rate following a simple Taylor rule, i.e.

$$r_t = \rho_R r_{t-1} + (1-\rho_R)(\rho_\pi \pi_t + \rho_y y_t) + \epsilon_t^r$$

where $\epsilon_t^r \sim N(0, \sigma_r^2)$

State Spaces

Equations (2.1), (2.2)-(2.4) and (2.5) can be cast into the linear state space representation (1.14)-(1.15), by setting

$$
\begin{aligned}
Y_t &= y_t \\
s_t &= \begin{pmatrix} 1 & t & x_{t-1} & z_t \end{pmatrix}' \\
F &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & PP & QQ \\ 0 & 0 & 0 & NN \end{pmatrix} \\
G &= \begin{pmatrix} 0 & 0 & 0 & I \end{pmatrix}' \\
H &= \begin{pmatrix} A & B & RR & SS \end{pmatrix} \\
\omega_{t+1} &= \nu_{t+1}
\end{aligned}
$$

Equations (2.1),(2.2)-(2.4) and (2.6) fit the state space representation (1.14)-(1.15) by setting

$$
\begin{aligned}
Y_t &= y_t - \Gamma y_{t-1} \\
s_t &= \begin{pmatrix} 1 & x_{t-1} & z_t \end{pmatrix}' \\
F &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & PP & QQ \\ 0 & 0 & NN \end{pmatrix} \\
G &= \begin{pmatrix} 0 & 0 & I \end{pmatrix}' \\
H &= \begin{pmatrix} -\gamma & RR & SS \end{pmatrix} \\
\omega_{t+1} &= \nu_{t+1}
\end{aligned}
$$

Equations (2.1),(2.2)-(2.4) and (2.7)-(2.8) can be cast into the linear state space representation (1.14) and (1.15), by setting

$$
\begin{aligned}
Y_t &= y_t \\
s_t &= \begin{pmatrix} y_t^{\tau} & \mu_t & x_{t-1} & z_t \end{pmatrix} \\
F &= \begin{pmatrix} I & I & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & PP & QQ \\ 0 & 0 & 0 & NN \end{pmatrix} \\
G &= \begin{pmatrix} 0 & 0 \\ I & 0 \\ 0 & 0 \\ 0 & I \end{pmatrix} \\
H &= \begin{pmatrix} I & 0 & RR & SS \end{pmatrix} \\
\omega_{t+1} &= \begin{pmatrix} \zeta_{t+1} & \nu_{t+1} \end{pmatrix}
\end{aligned}
$$

Equations (2.1),(2.2)-(2.4), (2.9)-(2.11) fit the state space representation,(1.14)-(1.15), by setting

$$
\begin{aligned}
s_t &= \begin{pmatrix} A_t & B_t & x_{t-1} & z_t \end{pmatrix}' \\
F &= \begin{pmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & PP & QQ \\ 0 & 0 & 0 & NN \end{pmatrix} \\
G &= \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 0 \\ 0 & 0 & I \end{pmatrix} \\
H_t &= \begin{pmatrix} I & tI & RR & SS \end{pmatrix} \\
\omega_{t+1} &= \begin{pmatrix} \eta_{t+1}^{A} & \eta_{t+1}^{B} & \nu_{t+1} \end{pmatrix}
\end{aligned}
$$

Finally, equations (2.1),(2.2)-(2.4), (2.12)-(2.13) can be cast in a state space representation by setting

$$\begin{aligned}
Y_t &= y_t - \Gamma y_{t-1} \\
s_t &= \begin{pmatrix} \gamma_t & x_{t-1} & z_t \end{pmatrix}' \\
F &= \begin{pmatrix} I & 0 & 0 \\ 0 & PP & QQ \\ 0 & 0 & NN \end{pmatrix} \\
G &= \begin{pmatrix} I & 0 \\ 0 & 0 \\ 0 & I \end{pmatrix} \\
H &= \begin{pmatrix} I & RR & SS \end{pmatrix} \\
\omega_{t+1} &= \begin{pmatrix} \eta^{\gamma}_{t+1} & \nu_{t+1} \end{pmatrix}
\end{aligned}$$

θ	Prior F	mean	s.d.	Posterior mean	s.d.	median
α	$B(10, 28)$	0.26	0.07	0.3020	0.0072	0.2996
η	$\Gamma(2, 1.25)$	2.5	1.7	1.0514	0.0624	1.0601
K/Y	$N(2.5, 0.1)$	2.5	0.1	3.0991	0.1728	3.1772
τ^w	$B(4, 18)$	0.18	0.08	0.6435	0.0377	0.6428
τ^k	$B(2, 20)$	0.09	0.05	0.7154	0.0190	0.7158
φ_w	$N(0.2, 0.1)$	0.2	0.1	0.0560	0.0532	0.0537
φ_{by}	$N(0.2, 0.1)$	0.2	0.1	0.3813	0.0469	0.3812
φ_y	$N(0.2, 0.1)$	0.2	0.1	0.6578	0.0572	0.6722
φ_n	$N(0.2, 0.1)$	0.2	0.1	0.4722	0.0812	0.4961
ψ_k	$N(0.2, 0.1)$	0.2	0.1	-0.1854	0.0387	-0.1911
ψ_{by}	$N(0.2, 0.1)$	0.2	0.1	0.3724	0.0882	0.3735
ψ_y	$N(0.2, 0.1)$	0.2	0.1	0.2905	0.0421	0.2873
ψ_n	$N(0.2, 0.1)$	0.2	0.1	0.2638	0.0792	0.2752
ρ_a	$B(18, 8)$	0.69	0.08	0.6795	0.0315	0.6832
ρ_g	$B(18, 8)$	0.69	0.08	0.3977	0.0307	0.3975
ρ_χ	$B(18, 8)$	0.69	0.08	0.8022	0.0574	0.8017
ρ_v	$B(18, 8)$	0.69	0.08	0.0253	0.0063	0.0249
σ_a	$\Gamma^{-1}(6, 0.5)$	0.23	0.08	0.0738	0.0082	0.0734
σ_g	$\Gamma^{-1}(6, 0.5)$	0.23	0.08	0.1223	0.0180	0.1207
σ_χ	$\Gamma^{-1}(6, 0.5)$	0.23	0.08	0.2524	0.0562	0.2442
σ_v	$\Gamma^{-1}(6, 0.5)$	0.23	0.08	0.0924	0.0111	0.0917
σ_{ξ^w}	$\Gamma^{-1}(6, 0.5)$	0.23	0.08	0.1004	0.0119	0.0996
σ_{ξ^k}	$\Gamma^{-1}(6, 0.5)$	0.23	0.08	0.0871	0.0102	0.0864
σ_{my}	$\Gamma^{-1}(6, 0.5)$	0.23	0.08	0.0759	0.0085	0.0752
σ_{mw}	$\Gamma^{-1}(6, 0.5)$	0.23	0.08	0.0770	0.0085	0.0763
σ_{mk}	$\Gamma^{-1}(6, 0.5)$	0.23	0.08	0.1095	0.0133	0.1086
σ_{mby}	$\Gamma^{-1}(6, 0.5)$	0.23	0.08	0.0769	0.0088	0.0762
σ_{mg}	$\Gamma^{-1}(6, 0.5)$	0.23	0.08	0.0924	0.0110	0.0916

Table 1: Prior and Posterior statistics for the parameters

	$sd(c_t)$	$sd(y_t)$	$sd(n_t)$	$sd(i_t)$	$sd(\tau_t^w)$	$sd(\tau_t^k)$	$sd(by_t)$	$sd(g_t)$
Data	1.22	1.52	1.57	4.60	0.85	1.51	0.97	1.31
Model	0.40	0.23	0.28	2.22	0.15	0.21	0.43	0.16

Table 2: The first line displays the observed standard deviations of consumption, gdp, hours worked, investment, labor tax, capital tax, debt-gdp ratio and government spending. The second line displays the average standard deviations of the model simulated several times using a subset of accepted draws.

	Unrestr model	without τ_t^w & τ_t^k a.s.	without τ_t^w a.s.	without the τ_t^k a.s.	without FP shocks
$sd(c_t)$	0.40	0.67	0.68	0.40	0.36
$sd(y_t)$	0.23	0.80	0.78	0.30	0.22
$sd(n_t)$	0.28	1.16	1.13	0.40	0.24
$sd(i_t)$	2.22	9.04	8.80	3.15	1.98
$sd(\tau_t^w)$	0.15	0.33	0.32	0.17	0.14
$sd(\tau_t^k)$	0.21	0.31	0.47	0.40	0.22
$sd(by_t)$	0.43	0.65	0.65	0.71	0.68
$sd(g_t)$	0.16	0.16	0.16	0.15	0.16

Table 3: Standard Deviations with and without automatic stabilizers (a.s.).The first column shows the s.d. for the unrestricted model with the orthogonal tax policy shocks. The last column exhibits the s.d. for the model without the tax policy shocks. The second column contains the s.d. for the unrestricted model for which automatic stabilizers for the labor and capital tax are set to zero, i.e. $\psi_y = \psi_n = \varphi_y = \varphi_n = 0$. The third column shows the s.d. for the unrestricted model for which automatic stabilizers for the labor tax are set to zero, i.e. $\varphi_y = \varphi_n = 0$. The fourth column shows the s.d. for the unrestricted model for which automatic stabilizers for the capital tax are set to zero, i.e. $\psi_y = \psi_n = 0$.

	Unrestr model	Restr model without τ_t^w & τ_t^k a.s.	Restr model without τ_t^w a.s.	Restr model without τ_t^k a.s.
$sd(c_t)$	0.40	0.51	0.48	0.45
$sd(y_t)$	0.23	0.25	0.21	0.19
$sd(n_t)$	0.28	0.36	0.31	0.26
$sd(i_t)$	2.22	3.02	2.40	2.06
$sd(\tau_t^w)$	0.15	0.18	0.18	0.17
$sd(\tau_t^k)$	0.21	0.21	0.21	0.17
$sd(by_t)$	0.43	0.26	0.26	0.30
$sd(g_t)$	0.16	0.19	0.18	0.21

Table 4: Standard Deviations with and without automatic stabilizers (a.s.). The first column contains the s.d. for the unrestricted model with the orthogonal tax policy shocks. The second column displays the s.d. for the restricted model for which automatic stabilizers for the labor and capital tax are fixed to zero, i.e. $\psi_y = \psi_n = \varphi_y = \varphi_n = 0$. The third column exhibits the s.d. for the restricted model for which automatic stabilizers for the labor and capital tax are fixed to zero, i.e. $\varphi_y = \varphi_n = 0$. The fourth column shows the s.d. for the restricted model for which automatic stabilizers for the capital tax are fixed to zeros, i.e. $\psi_y = \psi_n = 0$.

FP parameter	baseline prior s.d.	s.d. increased by 20%	s.d. reduced by 20%
φ_w	0.06 (0.05)	0.08(0.10)	0.15(0.08)
φ_{by}	0.38(0.05)	0.43(0.06)	0.57(0.04)
φ_y	0.66(0.06)	0.74(0.06)	0.51(0.06)
φ_n	0.47(0.08)	0.54(0.07)	0.30(0.09)
ψ_k	-0.18(0.04)	-0.22(0.16)	0.05(0.07)
ψ_{by}	0.37(0.09)	0.31(0.07)	0.12(0.08)
ψ_y	0.29(0.04)	0.22(0.08)	0.20(0.06)
ψ_n	0.26(0.08)	0.26(0.08)	0.21(0.08)

Table 5: Mean Value for the fiscal policy parameters with the baseline standard deviation, then reducing and increasing the prior standard deviation of 20%, respectively. Posterior standard deviations are reported in parenthesis.

	S_1	S_2	S_3	S_4	S_5
$\ln p(Y\|\mathcal{S}_j)$	2270	2104	2409	2282	2397
Prior S_j Probability	0.2	0.2	0.2	0.2	0.2
log Posterior Odds w.r.t. S_3	-139	-305	0	-127	-2
$R(s.d.(\theta)\|y)$	0.05	0.09	0.01	0.01	0.01
$R(corr_c(\theta)\|y)$	1.91	2.81	1.76	3.12	2.28
$R(corr_y(\theta)\|y)$	1.63	2.27	0.87	2.21	1.43
$R(corr_n(\theta)\|y)$	4.68	3.64	1.19	4.54	1.47
$R(corr_i(\theta)\|y)$	1.28	4.85	1.10	3.31	1.64
$R(corr_{\tau^w}(\theta)\|y)$	4.26	5.59	2.83	4.34	1.79
$R(corr_{\tau^k}(\theta)\|y)$	2.91	3.34	1.19	4.06	1.41
$R(corr_{by}(\theta)\|y)$	3.08	1.92	1.90	4.17	2.06
$R(corr_g(\theta)\|y)$	2.56	2.17	2.57	1.74	3.20

Table 6: Posterior Likelihood and Loss Function for moments predictions across different Fiscal Policy specifications. The third line displays the Posterior Odd. The following lines display the average risk in predicting moments in the data, standard deviations and correlations. In particular, the fourth line displays the average loss in predicting the set of standard deviations; the fifth and following lines show the loss in predicting the correlation of a certain variable with the remaining variables.

θ^m	Description	Distribution	Mean	Standard Deviation
σ_c	elasticity of intertemporal substitution	$\Gamma(20, 0.1)$	2.00	0.45
σ_n	elasticity of labor supply	$\Gamma(30, 0.1)$	3.00	0.55
ρ_R	AR in the monetary rule	$B(6, 6)$	0.50	0.14
ρ_π	response to inflation in monetary rule	$N(1.5, 0.1)$	1.50	0.10
ρ_y	response to GDP in monetary rule	$N(0.4, 0.1)$	0.40	0.10
ζ_p	prob of keeping the price fixed	$B(6, 6)$	0.50	0.14
ρ_χ	AR in the preference process	$B(18, 8)$	0.69	0.09
ρ_a	AR in the technology process	$B(18, 8)$	0.69	0.09
σ_χ	sd preference	$\Gamma^{-1}(10, 0.05)$	0.0056	0.002
σ_a	sd technology	$\Gamma^{-1}(10, 0.05)$	0.0056	0.002
σ_r	sd monetary policy	$\Gamma^{-1}(10, 0.05)$	0.0056	0.002
σ_μ	sd markup	$\Gamma^{-1}(10, 0.05)$	0.0056	0.002
θ^{lt}				
A_j	intercept	$N(0, 0.09)$	0	0.09
B_j	slope	$N(0, 0.09)$	0	0.09
σ_j^η	trend sd	$\Gamma^{-1}(10, 0.05)$	0.0056	0.002
θ^{hp}				
σ_j^ζ	trend sd	$\Gamma^{-1}(10, 0.05)$	0.0056	0.002
θ^{fd}				
γ_j	drift	$N(0, 0.09)$	0	0.09
σ_j^η	trend sd	$\Gamma^{-1}(10, 0.05)$	0.0056	0.0020

Table 7: Prior Distribution for the parameters θ

θ	σ_c	σ_n	ρ_r	ρ_π	ρ_y	ζ_p	ρ_χ	ρ_z	σ_χ	σ_z	σ_r	σ_μ
LP	1.00	1.00	0.50	1.10	0.50	0.80	0.40	0.40	0.90	0.60	0.70	0.80
HP	3.00	2.00	0.40	1.70	0.33	0.61	0.90	0.70	0.78	0.54	0.20	0.57
HV	2.50	2.20	0.35	2.00	0.40	0.40	0.60	0.60	0.95	0.98	0.75	0.89
LV	3.00	3.00	0.40	2.20	0.30	0.70	0.80	0.70	0.85	0.56	0.21	0.38

Table 8: Structural parameters: Population values. 'LP' stands for low persistence, 'HP' for high persistence, 'HV' stands for high volatility, 'LV' for low volatility.

TV	LP		HP		HV		LV		Average	
	2s	1s	2s	1s	2s	1s	2s	1s	2s	1s
σ_c	439	381	161	64	172	87	124	47	224	145
	(0.11)	(0.14)	(0.03)	(0.05)	(0.05)	(0.06)	(0.03)	(0.05)		
σ_n	57	22	30	33	19	46	38	61	36	41
	(0.12)	(0.14)	(0.05)	(0.07)	(0.06)	(0.07)	(0.04)	(0.04)		
ρ_r	88	79	84	73	85	67	85	77	85	74
	(0.05)	(0.05)	(0.04)	(0.08)	(0.07)	(0.07)	(0.06)	(0.07)		
ρ_π	43	55	16	4	24	17	38	19	30	24
	(0.02)	(0.03)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)		
ρ_y	21	12	103	70	79	15	74	87	69	46
	(0.05)	(0.06)	(0.06)	(0.09)	(0.06)	(0.07)	(0.08)	(0.09)		
ζ_p	14	7	49	16	129	77	30	4	55	26
	(0.02)	(0.04)	(0.03)	(0.05)	(0.04)	(0.07)	(0.03)	(0.04)		
ρ_χ	10	33	41	37	46	15	47	26	36	28
	(0.06)	(0.07)	(0.02)	(0.03)	(0.04)	(0.05)	(0.03)	(0.03)		
ρ_a	69	58	29	24	10	21	18	20	32	31
	(0.06)	(0.07)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)		
σ_χ	70	13	36	228	44	262	39	48	47	138
	(0.19)	(0.50)	(0.34)	(2.84)	(0.77)	(2.94)	(0.37)	(1.05)		
σ_a	79	41	36	80	9	173	46	50	42	86
	(0.13)	(0.41)	(0.38)	(1.21)	(0.39)	(1.59)	(0.29)	(0.84)		
σ_r	71	21	28	113	9	184	20	239	32	139
	(0.18)	(0.56)	(0.49)	(1.45)	(0.40)	(1.72)	(0.72)	(2.01)		
σ_μ	35	118	332	411	932	566	925	728	556	456
	(0.91)	(2.38)	(2.57)	(4.03)	(6.78)	(5.32)	(6.37)	(8.21)		
Average	83	70	79	96	130	127	124	117	0	0

Table 9: Bias comparison between 2 step and 1 step. Data is generated with the population values of Table 8 and with a deterministic trend. The bias values are expressed in % terms, with the standard deviations in parenthesis in % as well.

TV	LP 2s	LP 1s	HP 2s	HP 1s	HV 2s	HV 1s	LV 2s	LV 1s	Average 2s	Average 1s
σ_c	336	248	120	42	132	98	145	50	184	110
	(0.13)	(0.11)	(0.04)	(0.03)	(0.05)	(0.04)	(0.04)	(0.04)		
σ_n	54	14	29	40	19	36	16	55	29	36
	(0.13)	(0.10)	(0.05)	(0.05)	(0.06)	(0.04)	(0.04)	(0.03)		
ρ_r	83	71	83	73	81	62	84	71	83	69
	(0.04)	(0.04)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.05)		
ρ_π	44	37	15	13	17	19	24	29	25	24
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)		
ρ_y	28	5	121	89	74	76	50	98	68	67
	(0.05)	(0.04)	(0.07)	(0.05)	(0.05)	(0.05)	(0.07)	(0.06)		
ζ_p	4	10	34	41	80	106	16	22	34	45
	(0.03)	(0.02)	(0.03)	(0.03)	(0.05)	(0.06)	(0.03)	(0.02)		
ρ_χ	10	83	23	8	6	9	13	8	13	27
	(0.07)	(0.05)	(0.02)	(0.02)	(0.04)	(0.03)	(0.03)	(0.02)		
ρ_a	28	118	24	26	42	48	6	27	25	55
	(0.06)	(0.05)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)		
σ_χ	75	77	89	74	89	77	89	76	86	76
	(0.15)	(0.23)	(0.06)	(0.22)	(0.07)	(0.20)	(0.06)	(0.20)		
σ_a	90	82	89	81	93	89	89	82	90	83
	(0.07)	(0.12)	(0.07)	(0.14)	(0.05)	(0.09)	(0.06)	(0.13)		
σ_r	90	86	70	53	76	62	70	56	77	64
	(0.06)	(0.10)	(0.16)	(0.33)	(0.13)	(0.24)	(0.21)	(0.31)		
σ_μ	76	30	51	22	71	45	20	46	55	36
	(0.18)	(0.44)	(0.30)	(0.54)	(0.19)	(0.42)	(0.45)	(0.88)		
Average	77	72	62	47	65	61	52	52	0	0

Table 10: Bias comparison between 2 step and 1 step. Data is generated with the population values of Table 8 and with a stochastic trend. The bias values are expressed in % terms, with the standard deviations in parenthesis in % as well.

TV	LP		HP		HV		LV		Average	
	2s	1s	2s	1s	2s	1s	2s	1s	2s	1s
σ_c	439	383	161	42	172	62	124	39	224	131
	(0.11)	(0.09)	(0.03)	(0.03)	(0.05)	(0.04)	(0.03)	(0.03)		
σ_n	57	41	30	42	19	47	38	65	36	49
	(0.12)	(0.08)	(0.05)	(0.04)	(0.06)	(0.04)	(0.04)	(0.03)		
ρ_r	88	78	84	73	85	59	85	64	85	69
	(0.05)	(0.04)	(0.04)	(0.05)	(0.07)	(0.05)	(0.06)	(0.04)		
ρ_π	43	32	16	22	24	26	38	35	30	29
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)		
ρ_y	21	15	103	88	79	60	74	95	69	64
	(0.05)	(0.04)	(0.06)	(0.05)	(0.06)	(0.05)	(0.08)	(0.06)		
ζ_p	14	10	49	40	129	118	30	26	55	48
	(0.02)	(0.02)	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)		
ρ_χ	10	96	41	36	46	18	47	4	36	38
	(0.06)	(0.05)	(0.02)	(0.02)	(0.04)	(0.03)	(0.03)	(0.02)		
ρ_a	69	122	29	14	10	47	18	31	32	53
	(0.06)	(0.04)	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.02)		
σ_χ	70	35	36	90	44	121	39	13	47	65
	(0.19)	(0.59)	(0.34)	(1.76)	(0.77)	(1.83)	(0.37)	(0.49)		
σ_a	79	63	36	13	9	71	46	11	42	40
	(0.13)	(0.27)	(0.38)	(0.57)	(0.39)	(1.13)	(0.29)	(0.41)		
σ_r	71	56	28	20	9	54	20	84	32	53
	(0.18)	(0.37)	(0.49)	(0.81)	(0.40)	(1.13)	(0.72)	(1.31)		
σ_μ	35	164	332	541	932	847	925	996	556	637
	(0.91)	(1.14)	(2.57)	(3.90)	(6.78)	(5.29)	(6.37)	(6.26)		
Average	83	91	79	85	130	128	124	122	0	0

Table 11: Bias comparison under misspecification. Data is generated with the population values of Table 8 and with a deterministic trend. For the 1s I consider the stochastic trend specification, for the 2s a linear trend specification is used. The bias values are expressed in % terms, with the standard deviations in parenthesis in % as well.

TV	LP		HP		HV		LV		Average	
	2s	1s	2s	1s	2s	1s	2s	1s	2s	1s
σ_c	336	357	120	49	132	94	145	67	184	142
	(0.13)	(0.19)	(0.04)	(0.06)	(0.05)	(0.07)	(0.04)	(0.06)		
σ_n	54	13	29	13	19	20	16	42	29	22
	(0.13)	(0.16)	(0.05)	(0.09)	(0.06)	(0.09)	(0.04)	(0.06)		
ρ_r	83	76	83	63	81	66	84	79	83	71
	(0.04)	(0.08)	(0.05)	(0.09)	(0.06)	(0.09)	(0.06)	(0.09)		
ρ_π	44	51	15	7	17	36	24	35	25	32
	(0.02)	(0.03)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)		
ρ_y	28	19	121	55	74	175	50	97	68	87
	(0.05)	(0.07)	(0.07)	(0.10)	(0.05)	(0.07)	(0.07)	(0.13)		
ζ_p	4	8	34	6	80	54	16	8	34	19
	(0.03)	(0.05)	(0.03)	(0.06)	(0.05)	(0.09)	(0.03)	(0.04)		
ρ_χ	10	60	23	2	6	63	13	20	13	36
	(0.07)	(0.09)	(0.02)	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)		
ρ_a	28	33	24	26	42	16	6	17	25	23
	(0.06)	(0.09)	(0.04)	(0.05)	(0.03)	(0.06)	(0.03)	(0.04)		
σ_χ	75	59	89	56	89	56	89	49	86	55
	(0.15)	(0.04)	(0.06)	(0.05)	(0.07)	(0.04)	(0.06)	(0.04)		
σ_a	90	69	89	69	93	82	89	63	90	71
	(0.07)	(0.06)	(0.07)	(0.06)	(0.05)	(0.04)	(0.06)	(0.06)		
σ_r	90	73	70	11	76	34	70	26	77	36
	(0.06)	(0.05)	(0.16)	(0.17)	(0.13)	(0.15)	(0.21)	(0.17)		
σ_μ	76	45	51	34	71	61	20	11	55	38
	(0.18)	(0.04)	(0.30)	(0.06)	(0.19)	(0.04)	(0.45)	(0.09)		
Average	77	72	62	33	65	63	52	43	0	0

Table 12: Bias comparison under misspecification. Data is generated with the population values of Table 8 and with a stochastic trend. For the 1s I consider the deterministic trend specification, for the 2s data is HP filtered. The bias values are expressed in % terms, with the standard deviations in parenthesis in % as well.

TV	LP		HP		HV		LV		Average	
	2s	1s	2s	1s	2s	1s	2s	1s	2s	1s
σ_c	224	261	20	22	52	34	35	16	83	83
	(0.11)	(0.17)	(0.00)	(0.06)	(0.05)	(0.06)	(0.04)	(0.05)		
σ_n	226	66	97	11	74	24	51	33	112	33
	(0.12)	(0.16)	(0.00)	(0.08)	(0.05)	(0.08)	(0.03)	(0.06)		
ρ_r	89	72	88	72	87	66	86	72	88	70
	(0.04)	(0.06)	(0.00)	(0.08)	(0.06)	(0.09)	(0.05)	(0.08)		
ρ_π	49	35	41	11	27	21	54	27	43	23
	(0.02)	(0.03)	(0.00)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)		
ρ_y	62	31	95	42	112	15	114	39	96	32
	(0.05)	(0.07)	(0.00)	(0.10)	(0.05)	(0.08)	(0.05)	(0.12)		
ζ_p	39	22	48	6	68	60	10	9	41	24
	(0.03)	(0.05)	(0.00)	(0.05)	(0.04)	(0.08)	(0.03)	(0.04)		
ρ_χ	19	58	10	19	32	9	17	18	20	26
	(0.06)	(0.08)	(0.00)	(0.03)	(0.03)	(0.06)	(0.02)	(0.04)		
ρ_a	32	18	27	22	14	19	11	31	21	22
	(0.06)	(0.08)	(0.00)	(0.06)	(0.03)	(0.05)	(0.03)	(0.05)		
σ_χ	355	24	1700	371	2672	351	330	153	1264	225
	(1.20)	(0.09)	(0.08)	(0.36)	(5.40)	(0.40)	(1.78)	(0.16)		
σ_a	79	35	51	93	9	202	46	59	46	97
	(0.15)	(0.11)	(0.06)	(0.35)	(0.35)	(0.64)	(0.34)	(0.29)		
σ_r	68	10	32	136	10	215	21	312	33	168
	(0.20)	(0.15)	(0.06)	(0.45)	(0.44)	(0.52)	(0.63)	(0.72)		
σ_μ	65	58	465	406	65	450	219	701	203	404
	(0.18)	(0.15)	(0.08)	(0.39)	(1.01)	(0.62)	(1.96)	(0.71)		
Average	109	58	223	101	269	122	83	123	0	0

Table 13: Correlation Assumption. Data is generated with the population values of Table 8 and with a deterministic trend and allowing for correlation between trend and cycles. The bias values are expressed in % terms, with the standard deviations in parenthesis in % as well.

TV	LP		HP		HV		LV		Average	
	2s	1s	2s	1s	2s	1s	2s	1s	2s	1s
σ_c	304	299	171	41	115	83	103	40	173	116
	(0.13)	(0.11)	(0.03)	(0.03)	(0.05)	(0.04)	(0.04)	(0.03)		
σ_n	52	26	133	33	36	31	5	51	56	35
	(0.12)	(0.10)	(0.06)	(0.05)	(0.05)	(0.05)	(0.03)	(0.03)		
ρ_r	85	70	84	65	84	62	84	71	84	67
	(0.04)	(0.04)	(0.06)	(0.05)	(0.06)	(0.06)	(0.05)	(0.05)		
ρ_π	49	29	10	9	23	24	21	32	26	23
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)		
ρ_y	19	24	63	40	78	17	52	79	53	40
	(0.05)	(0.04)	(0.07)	(0.06)	(0.06)	(0.05)	(0.08)	(0.07)		
ζ_p	4	8	21	41	83	102	2	24	27	44
	(0.02)	(0.03)	(0.03)	(0.03)	(0.05)	(0.05)	(0.03)	(0.03)		
ρ_χ	19	104	11	8	5	41	14	17	12	42
	(0.06)	(0.06)	(0.02)	(0.02)	(0.04)	(0.03)	(0.03)	(0.02)		
ρ_a	57	114	17	24	50	47	14	28	35	53
	(0.05)	(0.05)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)		
σ_χ	79	74	87	74	90	77	91	76	87	75
	(0.12)	(0.25)	(0.08)	(0.22)	(0.06)	(0.19)	(0.05)	(0.20)		
σ_a	90	82	89	81	93	89	89	82	90	84
	(0.06)	(0.14)	(0.08)	(0.14)	(0.04)	(0.08)	(0.06)	(0.12)		
σ_r	90	86	69	55	76	63	68	56	76	65
	(0.07)	(0.10)	(0.21)	(0.30)	(0.15)	(0.26)	(0.23)	(0.37)		
σ_μ	79	36	74	23	75	51	52	53	70	41
	(0.16)	(0.54)	(0.18)	(0.49)	(0.16)	(0.40)	(0.30)	(0.98)		
Average	77	79	69	41	67	57	50	51	0	0

Table 14: Correlation Assumption. Data is generated with the population values of Table 8 and with a stochastic trend and allowing for correlation between trend and cycles. The bias values are expressed in % terms, with the standard deviations in parenthesis in % as well.

θ	T=160		T=500		T=1000	
	2s	1s	2s	1s	2s	1s
Deterministic						
σ_c	439(0.11)	343(0.16)	69(0.02)	40(0.00)	16(0.00)	18(0.01)
σ_n	57(0.12)	14(0.13)	10(0.02)	40(0.00)	1(0.00)	7(0.01)
ρ_r	88(0.04)	81(0.05)	10(0.01)	6(0.00)	3(0.00)	2(0.00)
ρ_π	43(0.02)	47(0.03)	1(0.00)	0(0.00)	0(0.00)	0(0.00)
ρ_y	21(0.05)	17(0.06)	5(0.01)	4(0.00)	1(0.00)	1(0.00)
ζ_p	14(0.02)	10(0.03)	10(0.00)	2(0.00)	6(0.00)	3(0.00)
ρ_χ	10(0.06)	18(0.07)	5(0.01)	1(0.00)	5(0.00)	2(0.00)
Stochastic						
σ_c	336(0.13)	248(0.10)	80(0.03)	27(0.01)	14(0.01)	8(0.00)
σ_n	54(0.13)	14(0.10)	18(0.03)	4(0.01)	2(0.01)	1(0.00)
ρ_r	83(0.04)	71(0.04)	34(0.01)	14(0.01)	5(0.00)	1(0.00)
ρ_π	44(0.02)	37(0.02)	5(0.00)	0(0.00)	2(0.00)	1(0.00)
ρ_y	28(0.05)	5(0.04)	7(0.01)	2(0.01)	0(0.00)	1(0.00)
ζ_p	4(0.03)	10(0.02)	2(0.01)	3(0.00)	1(0.00)	2(0.00)
ρ_χ	10(0.07)	83(0.05)	2(0.01)	4(0.01)	1(0.00)	2(0.00)
ρ_z	28(0.06)	118(0.05)	5(0.01)	5(0.01)	3(0.00)	3(0.00)

Table 15: Bias comparison using different samples length. Data is simulated using the first population value.

True DGP	LP	HP	HV	LV
Deterministic with $corr(y_t^c, y_t^\tau) = 0$	72	81	71	77
Stochastic with $corr(y_t^c, y_t^\tau) = 0$	-8	11	-44	-38
Deterministic with $corr(y_t^c, y_t^\tau) \neq 0$	143	95	140	83
Stochastic with $corr(y_t^c, y_t^\tau) \neq 0$	-7	-147	-99	-73

Table 16: Difference between the (log) Posterior Odds of lt-dsge and hp-dsge specifications.

$\mathcal{F}$	lt		hp		fd	
	2s (1)	1s (2)	2s (3)	1s (4)	2s (5)	1s (6)
σ_c	1.57	3.82	5.28	4.78	5.56	4.55
	(1.12)	(0.48)	(0.80)	(0.83)	(1.09)	(0.50)
σ_n	0.57	1.69	2.34	1.93	1.49	1.31
	(0.67)	(0.28)	(0.33)	(0.38)	(0.38)	(0.26)
ρ_r	0.14	0.22	0.28	0.15	0.18	0.11
	(0.07)	(0.06)	(0.06)	(0.05)	(0.08)	(0.05)
ρ_π	1.77	1.63	1.73	1.64	1.53	1.71
	(0.17)	(0.12)	(0.16)	(0.12)	(0.16)	(0.15)
ρ_y	0.17	0.22	0.12	0.47	0.43	0.46
	(0.16)	(0.16)	(0.13)	(0.11)	(0.10)	(0.08)
ζ_p	0.69	0.78	0.58	0.65	0.59	0.65
	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)
ρ_χ	0.91	0.79	0.98	0.79	0.51	0.57
	(0.10)	(0.10)	(0.11)	(0.08)	(0.09)	(0.07)
ρ_a	0.98	0.55	0.38	0.87	0.93	0.48
	(0.14)	(0.11)	(0.08)	(0.08)	(0.13)	(0.08)
σ_χ	0.08	0.58	0.38	0.23	0.44	0.31
	(0.08)	(0.14)	(0.10)	(0.05)	(0.12)	(0.05)
σ_a	0.07	0.17	0.09	0.10	0.08	0.16
	(0.03)	(0.16)	(0.02)	(0.02)	(0.02)	(0.02)
σ_r	1.22	0.21	0.10	0.10	0.11	0.17
	(0.15)	(0.14)	(0.02)	(0.02)	(0.02)	(0.02)
σ_μ	1.44	0.78	0.25	0.27	0.74	0.34
	(0.55)	(0.15)	(0.15)	(0.09)	(0.22)	(0.07)
$g(\mathcal{F}_j)$	-	1/3	-	1/3	-	1/3
$\ln p(y\|\mathcal{M},\mathcal{F}_j)$	-	1203	-	1171	-	1301
$\ln(PO)$ w.r.t lt-dsge	-	0.00	-	-31.80	-	98.47

Table 17: Structural estimates comparison between 2 step and 1 step with real data. Median and standard deviations in parenthesis. Structural standard deviations are expressed in percentage terms.

θ	Description	Prior	mean	sd
Behavioral				
$100(1/\beta - 1)$	β time discount factor	$\Gamma(6.25, 0.04)$	0.25	0.10
σ_c	intertemporal elasticity of substitution	$N(1.5, 0.27)$	1.50	0.27
σ_n	elasticity of labor supply	$N(2, 0.75)$	1.99	0.75
α	capital share	$N(0.3, 0.03)$	0.30	0.05
ϕ_p	1 plus the share of fixed cost in production	$N(1.25, 0.12)$	1.25	0.12
$100(\pi - 1)$	π steady state inflation	$\Gamma(38, 0.01)$	0.62	0.10
h	habit in consumption	$B(14, 6)$	0.70	0.10
ψ	elasticity capital utilization adjustment costs	$B(5.05, 5.05)$	0.50	0.15
φ	st. st. elasticity of capital adjustment costs	$N(4, 1.5)$	4.00	1.50
ζ_p	price stickiness	$B(12, 12)$	0.50	0.10
ζ_ω	wage stickiness	$B(12, 12)$	0.50	0.10
i_p	price indexation	$B(1, 1)$	0.50	0.29
i_ω	wage indexation	$B(1, 1)$	0.50	0.29
ρ_R	monetary policy autoregressive coeff.	$B(13, 4)$	0.75	0.10
ρ_π	monetary policy response to π	$N(1.5, 0.25)$	1.50	0.25
ρ_y	monetary policy response to y	$N(0.12, 0.25)$	0.12	0.05
AR Coeff				
ρ_a	technology autoregressive coeff.	$B(2.6, 2.6)$	0.50	0.20
ρ_g	gov spending autoregressive coeff.	$B(2.6, 2.6)$	0.50	0.20
ρ_i	investment autoregressive coeff.	$B(2.6, 2.6)$	0.50	0.20
ρ_r	monetary innovation autoregressive coeff.	$B(2.6, 2.6)$	0.50	0.20
ρ_p	price markup autoregressive coeff.	$B(2.6, 2.6)$	0.50	0.20
ρ_ω	wage markup autoregressive coeff.	$B(2.6, 2.6)$	0.50	0.20
ρ_b	risk premium autoregressive coeff.	$B(2.6, 2.6)$	0.50	0.20
ρ_{ga}	cross coefficient tech-gov	$B(2.6, 2.6)$	0.50	0.20
Sd				
σ_a	sd technology	$\Gamma^{-1}(3, 1)$	0.50	0.25
σ_g	sd gov spending	$\Gamma^{-1}(3, 1)$	0.50	0.25
σ_i	sd investment	$\Gamma^{-1}(3, 1)$	0.50	0.25
σ_r	sd mp	$\Gamma^{-1}(3, 1)$	0.50	0.25
σ_p	sd price markup	$\Gamma^{-1}(3, 1)$	0.50	0.25
σ_w	sd wage markup	$\Gamma^{-1}(3, 1)$	0.50	0.25
σ_b	sd preference	$\Gamma^{-1}(3, 1)$	0.50	0.25
MA Coeff				
μ_p	MA coeff. on price markup innovation	$B(2.6, 2.6)$	0.50	0.20
μ_ω	MA coeff. on wage markup innovation	$B(2.6, 2.6)$	0.50	0.20

Table 18: Parameters Description and Priors of the Smets and Wouters (2007) model.

	lt-dsge	hp-dsge	fd-dsge
$g(\mathcal{F}_j)$	1/3	1/3	1/3
$\ln p(y\|\mathcal{M}, \mathcal{F}_j)$	-1135	-1417	-1049
$\ln PO$ w.r.t lt-dsge	0.0	-282.3	85.8

Table 19: Posterior Odds across specifications.

	LT	HP	BP	FOD	True
Output	0.36	0.08	0.18	0.13	0.21
Inflation	0.12	0.07	0.07	0.13	0.12

Table 20: Standard deviation of filtered and true cyclical components; simulated data, scale 10^{-2}.

Parameter	Distribution	Mean	Standard Deviation
σ_c	$\Gamma(20, 0.1)$	2.00	0.45
σ_n	$\Gamma(20, 0.1)$	2.00	0.45
h	$B(10, 3)$	0.76	0.11
α	$B(3, 8)$	0.27	0.13
ϵ	$N(6, 0.5)$	6.00	0.50
ρ_r	$B(10, 6)$	0.71	0.09
ρ_π	$N(1.5, 0.2)$	1.50	0.20
ρ_y	$N(0.4, 0.2)$	0.40	0.20
ζ_p	$B(6, 6)$	0.50	0.14
ρ_χ	$B(10, 6)$	0.71	0.09
ρ_z	$B(10, 6)$	0.71	0.09
σ_χ	$\Gamma^{-1}(10, 20)$	0.0056	0.0020
σ_z	$\Gamma^{-1}(10, 20)$	0.0056	0.0020
σ_v	$\Gamma^{-1}(10, 20)$	0.0055	0.0020
σ_μ	$\Gamma^{-1}(10, 20)$	0.0056	0.0020

Table 21: Prior distributions for the structural parameters. Γ is the gamma distribution, B is the beta distribution, N the normal distribution.

Filter	True	LT	HP	FOD	BP	Factor_u	Factor_r
				All filtered			
σ_c	3.00	2.00(0.14)	1.96(0.12)	2.08(0.12)	1.95(0.11)	2.34(0.18)	2.99(0.33)
σ_n	0.70	1.24(0.06)	1.25(0.05)	1.25(0.05)	1.33(0.06)	0.10(0.03)	0.42(0.03)
h	0.70	0.47(0.04)	0.45(0.03)	0.43(0.04)	0.57(0.03)	0.73(0.04)	0.72(0.05)
α	0.60	0.14(0.02)	0.12(0.02)	0.14(0.02)	0.15(0.01)	0.49(0.02)	0.44(0.02)
ϵ	7.00	4.00(0.13)	4.70(0.17)	3.65(0.15)	3.89(0.13)	6.13(0.09)	6.34(0.08)
ρ_r	0.20	0.20(0.07)	0.14(0.06)	0.35(0.09)	0.23(0.05)	0.28(0.05)	0.26(0.12)
ρ_π	1.30	1.58(0.06)	1.66(0.09)	1.55(0.11)	1.56(0.06)	1.51(0.02)	1.56(0.02)
ρ_y	0.05	0.81(0.05)	0.84(0.07)	0.68(0.04)	0.71(0.04)	0.26(0.02)	0.25(0.05)
ζ_p	0.80	0.93(0.03)	0.93(0.03)	0.91(0.03)	0.95(0.03)	0.83(0.04)	0.83(0.03)
ρ_χ	0.50	0.93(0.03)	0.94(0.03)	0.91(0.03)	0.98(0.03)	0.72(0.05)	0.61(0.10)
ρ_z	0.80	0.87(0.03)	0.88(0.03)	0.85(0.04)	0.95(0.03)	0.68(0.03)	0.71(0.05)
σ_χ	1.10	0.12(0.02)	0.12(0.02)	0.12(0.02)	0.12(0.02)	0.21(0.04)	0.29(0.09)
σ_z	0.57	0.08(0.01)	0.09(0.01)	0.08(0.01)	0.08(0.01)	0.32(0.08)	0.37(0.09)
σ_{mp}	0.12	0.06(0.01)	0.06(0.01)	0.06(0.01)	0.06(0.01)	0.08(0.01)	0.08(0.01)
σ_μ	20.64	4.59(0.59)	4.30(0.52)	3.03(0.65)	3.28(0.52)	4.78(0.36)	5.45(0.73)
				Real filtered			
σ_c	3.00	1.99(0.14)	2.02(0.13)	1.90(0.14)	2.12(0.12)	2.40(0.36)	2.90(0.40)
σ_n	0.70	1.24(0.05)	1.25(0.05)	1.24(0.05)	1.43(0.08)	0.12(0.03)	0.45(0.03)
h	0.70	0.45(0.04)	0.49(0.03)	0.36(0.04)	0.61(0.03)	0.66(0.04)	0.62(0.04)
α	0.60	0.13(0.02)	0.14(0.02)	0.14(0.02)	0.15(0.03)	0.48(0.02)	0.36(0.04)
ϵ	7.00	3.63(0.14)	4.37(0.15)	3.77(0.16)	4.24(0.16)	6.17(0.09)	6.37(0.11)
ρ_r	0.20	0.16(0.05)	0.23(0.04)	0.38(0.08)	0.09(0.04)	0.36(0.05)	0.29(0.07)
ρ_π	1.30	1.60(0.09)	1.66(0.08)	1.63(0.10)	1.68(0.06)	1.48(0.02)	1.49(0.02)
ρ_y	0.05	0.83(0.05)	0.79(0.08)	0.77(0.06)	0.71(0.08)	0.33(0.02)	0.33(0.02)
ζ_p	0.80	0.93(0.03)	0.90(0.03)	0.94(0.03)	0.93(0.03)	0.78(0.05)	0.83(0.05)
ρ_χ	0.50	0.94(0.03)	0.93(0.03)	0.94(0.03)	0.97(0.03)	0.65(0.04)	0.68(0.06)
ρ_z	0.80	0.88(0.03)	0.86(0.03)	0.93(0.03)	0.94(0.03)	0.66(0.03)	0.66(0.05)
σ_χ	1.10	0.12(0.02)	0.11(0.02)	0.13(0.02)	0.12(0.02)	0.21(0.04)	0.26(0.08)
σ_z	0.57	0.08(0.01)	0.08(0.01)	0.09(0.01)	0.08(0.01)	0.29(0.08)	0.36(0.09)
σ_{mp}	0.12	0.06(0.01)	0.06(0.01)	0.06(0.01)	0.06(0.01)	0.08(0.01)	0.08(0.01)
σ_μ	20.64	6.02(0.52)	2.47(0.69)	11.72(0.94)	7.56(0.66)	4.88(0.36)	5.60(0.67)

Table 22: Parameters estimates obtained using different filters, median values and standard deviations in parenthesis; the DGP has a preference shock with two components, a stationary AR(1) and a unit root.

Series	LT	HP	FOD	BP	Factor_u	Factor_r
Output	0.614	0.659	0.590	0.684	0.485	0.084
Inflation	0.380	0.454	0.365	0.454	0.377	0.346

Table 23: Mean square error of the unconditional forecasts; simulated data; scale 10^{-4}.

Specification	Acceptance rate	Marginal log Likelihood	ω_2	ρ_m
Unrestricted	33.86	16274	0.44 (0.02)	0.48 (0.02)
$\omega_2 = 0$	33.64	16237	0	0.96(0.01)
$\rho_m = 0$	38.15	16212	0.43(0.02)	0
$\omega_2 = 0$, $\rho_m = 0$	33.77	16220	0	0
Standard			0.03 (0.02)	0.04 (0.03)

Table 24: Marginal likelihood and posterior estimates.

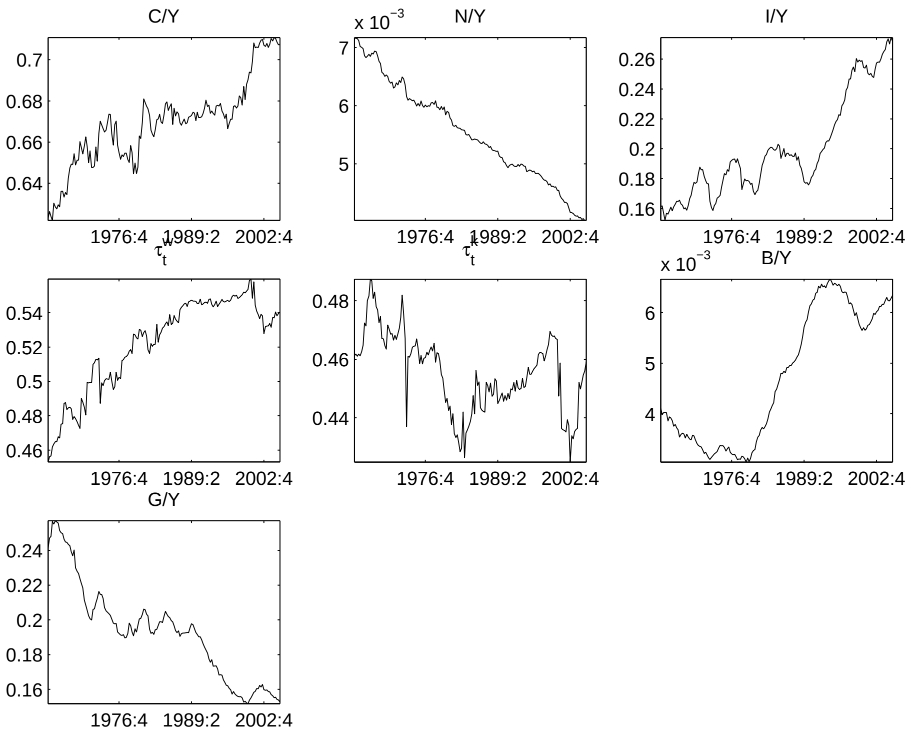

Figure 1: Time series plots, 1966:2 to 2006:2. From top left: consumption-GDP ratio, hours worked-gpd ratio, investment-GDP ratio, labor tax, capital tax, debt-GDP ratio and government spending-GDP ratio.

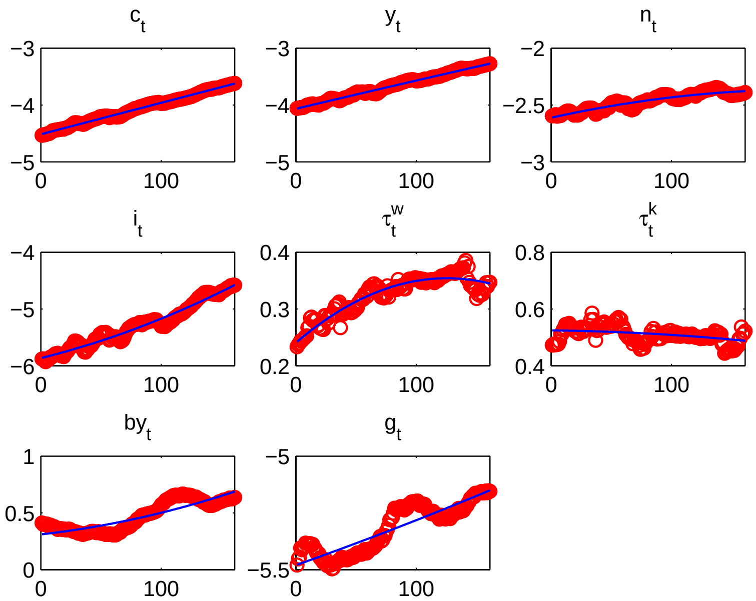

Figure 2: Plots of the time series, 1966:2 to 2006:2, using a quadratic trend; the circles represent the observations. From top left: consumption, GDP, hours worked, investment, labor tax, capital tax, debt-GDP ratio and government spending.

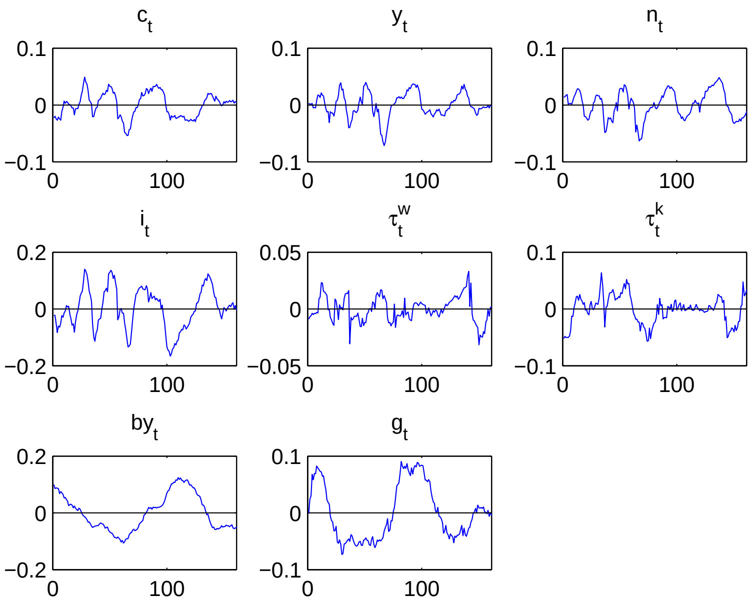

Figure 3: Plots of the extracted cycles of the times series, 1966:2 to 2006:2, using a quadratic trend. From top left: consumption, GDP, hours worked, investment, labor tax, capital tax, debt-GDP ratio and government spending.

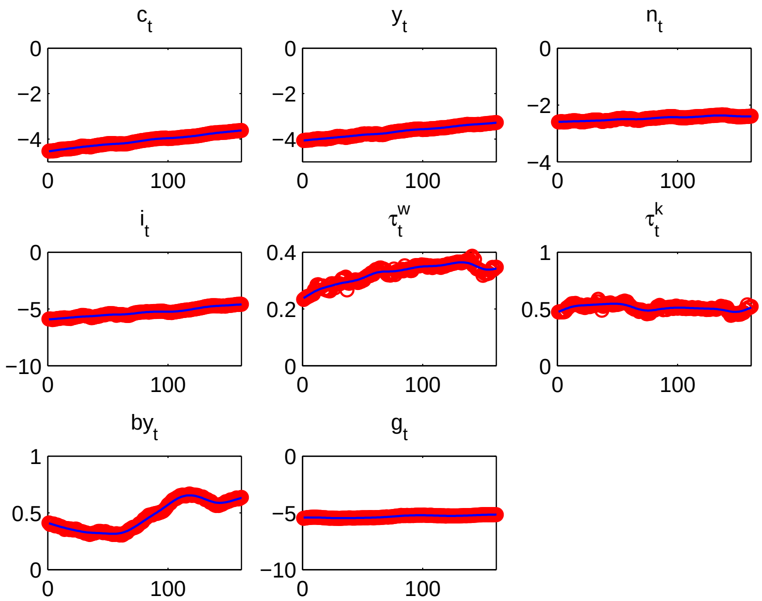

Figure 4: Plots of the time series, 1966:2 to 2006:2, using a HP filter; the circles represent the observations. From top left: consumption, GDP, hours worked, investment, labor tax, capital tax, debt-GDP ratio and government spending.

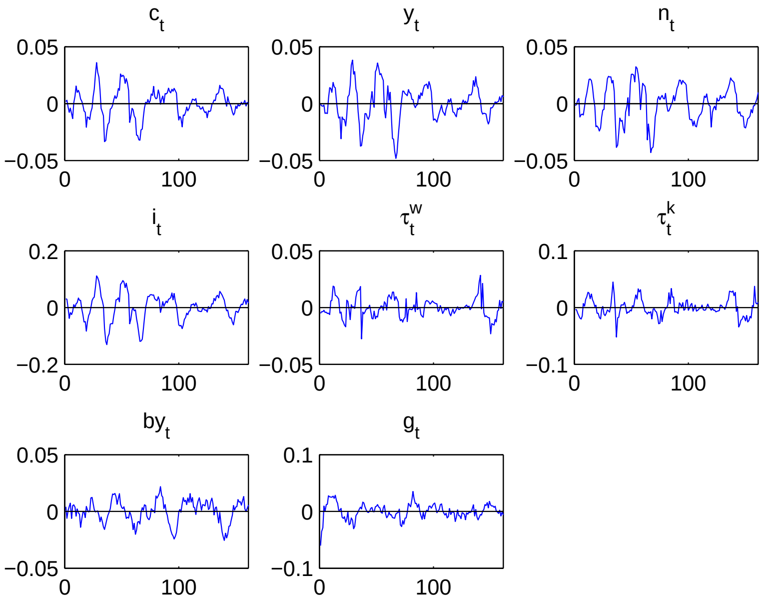

Figure 5: Plots of the extracted cycles of the times series, 1966:2 to 2006:2, using a HP filter. From top left: consumption, GDP, hours worked, investment, labor tax, capital tax, debt-GDP ratio and government spending.

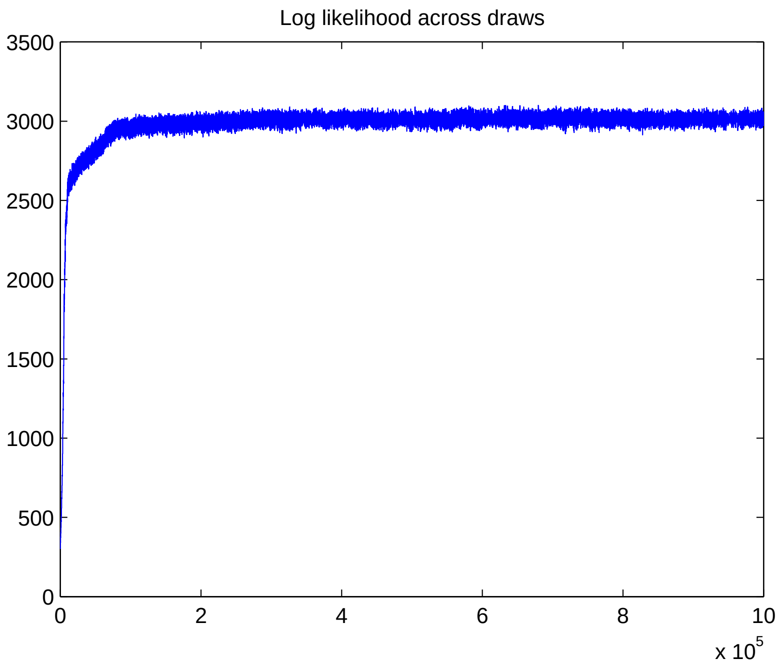

Figure 6: Log of the marginal likelihood across draws in the RWM algorithm

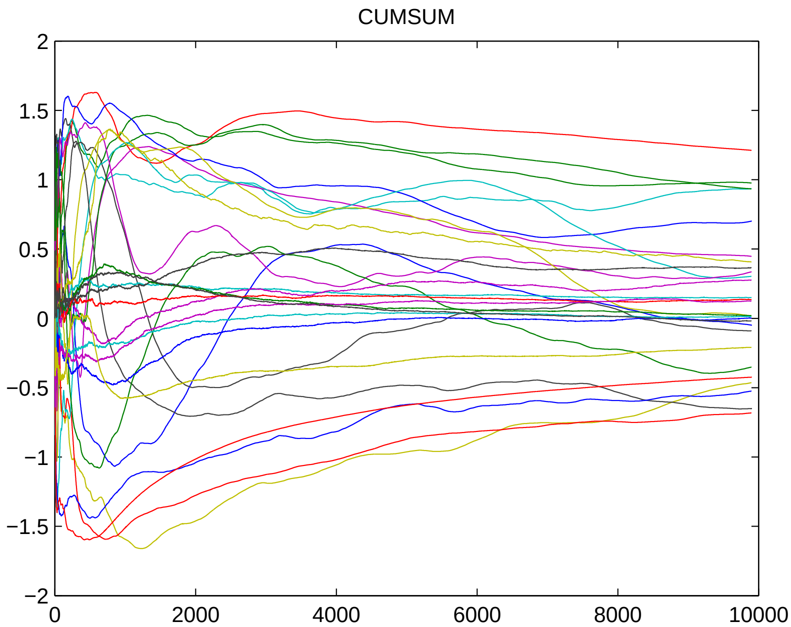

Figure 7: Convergence Statistics: cumulative sum of draws.

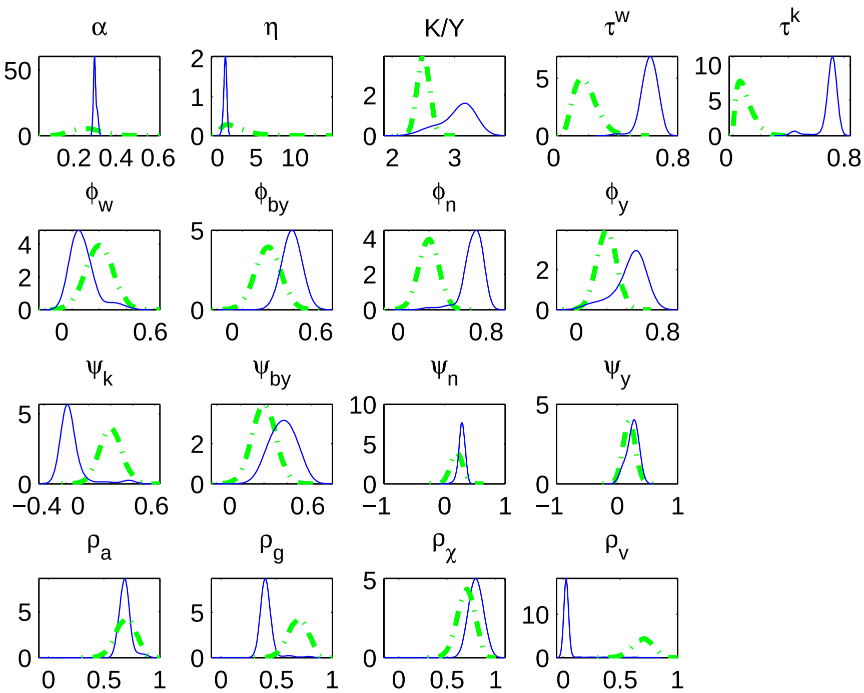

Figure 8: Prior and Posterior distributions for fiscal and autoregressive parameters, $\theta^p = [\alpha, \eta, \frac{K}{Y}, \tau^w, \tau^k, \varphi_w, \varphi_{by}, \varphi_n, \varphi_y, \psi_k, \psi_{by}, \psi_n, \psi_y, \rho_a, \rho_g, \rho_\chi, \rho_v]$. The solid lines represent the posterior and the dotted lines the prior.

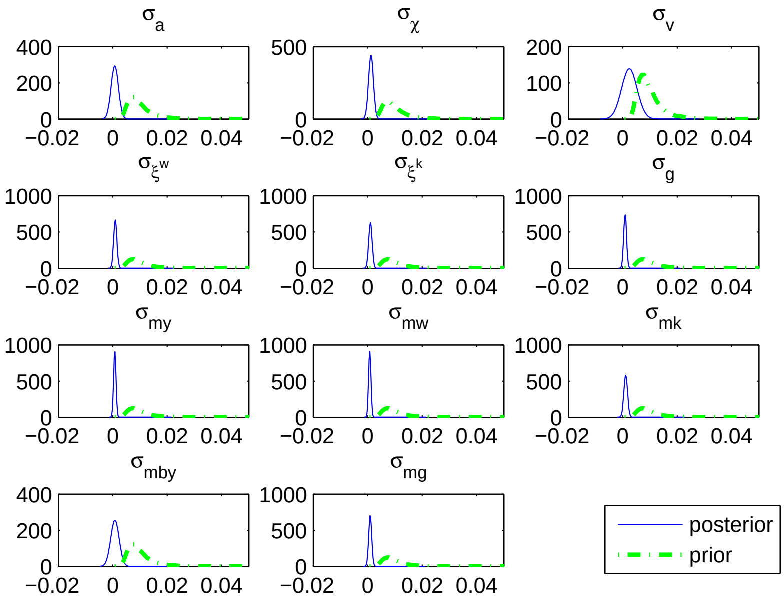

Figure 9: Prior and Posterior distributions for the auxiliary parameters, $\theta^a = [\sigma_a, \sigma_\chi, \sigma_v, \sigma_{\xi^w}, \sigma_{\xi^k}, \sigma_g, \sigma_{my}, \sigma_{mw}, \sigma_{mk}, \sigma_{mby}, \sigma_{mg}]$. The solid lines represent the posterior and the dotted lines the prior.

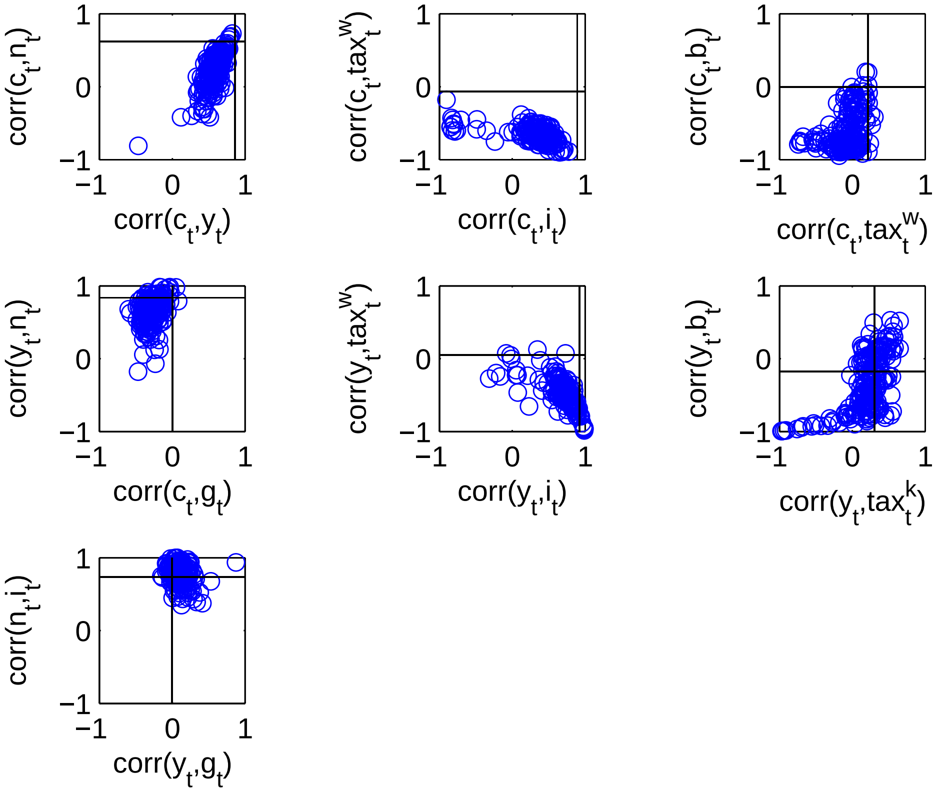

Figure 10: Posterior Scatter plots of the predicted values of the contemporaneous correlations and actual values. In particular, the circles represent the correlations implied by the model using a subset of accepted draws, whereas the vertical and horizontal lines represent the observed correlations in the data.

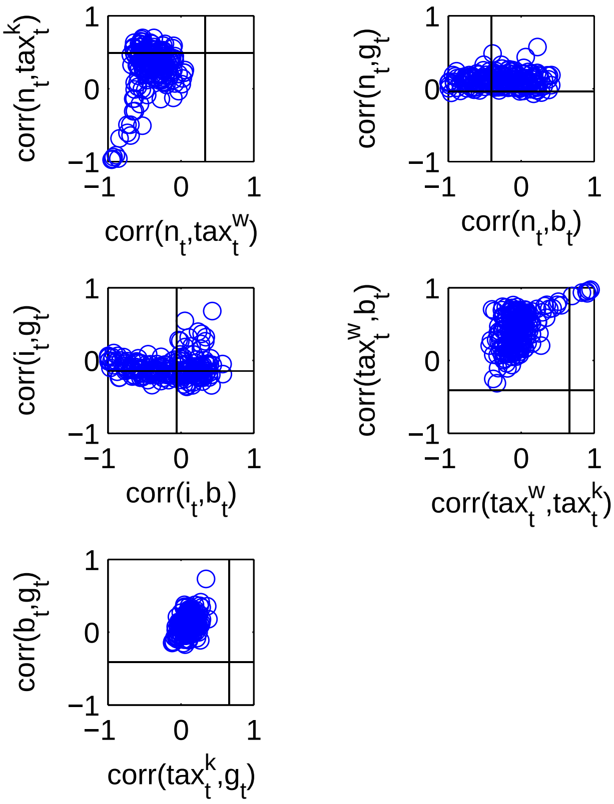

Figure 11: Posterior Scatter plots of the predicted values of the contemporaneous correlations and actual values. In particular, the circles represent the correlations implied by the model using a subset of accepted draws, whereas the vertical and horizontal lines represent the observed correlations in the data.

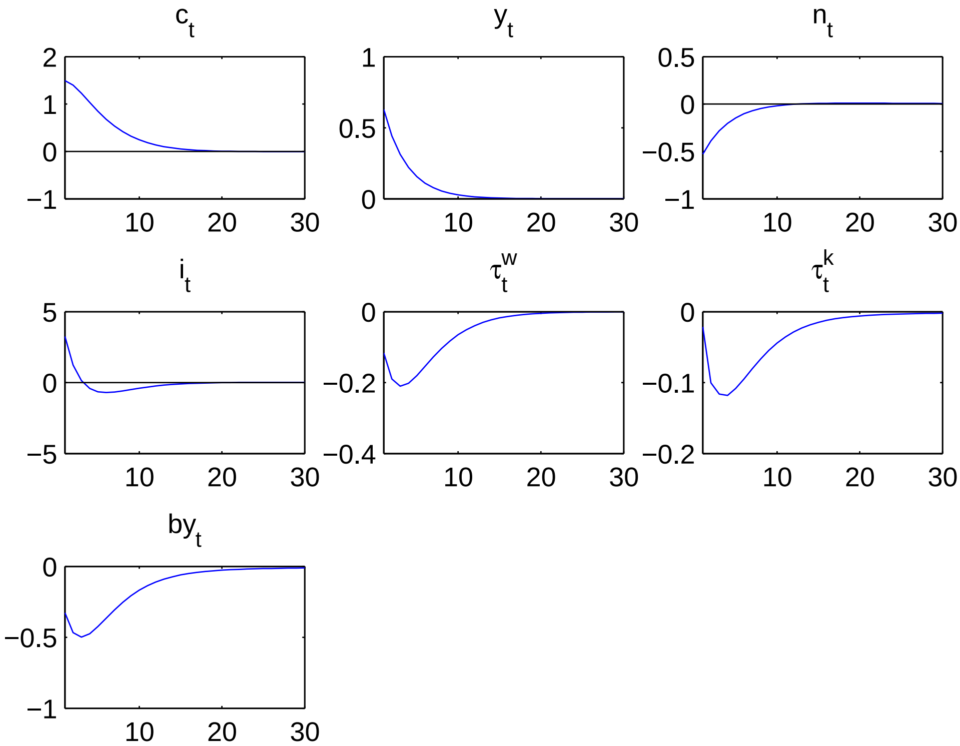

Figure 12: Impulse Response of consumption, GDP and hours worked, investment, labor and capital tax, debt-GDP ratio and government spending to a unitary increase in technology using the median estimates.

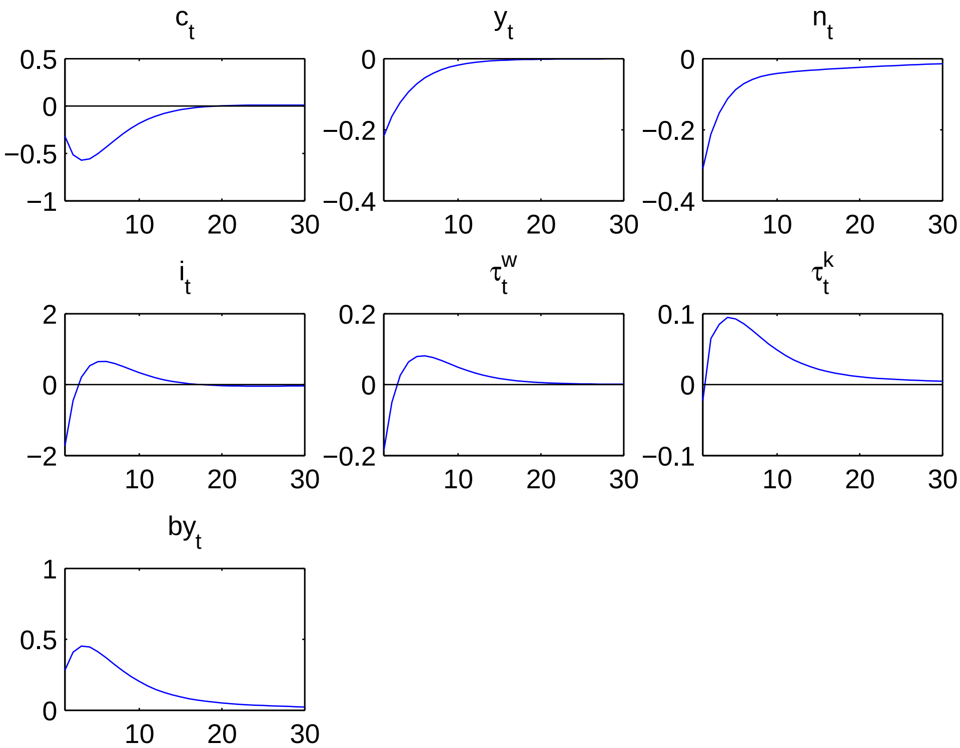

Figure 13: Impulse Response of consumption, GDP and hours worked, investment, labor and capital tax, debt-GDP ratio and government spending to a unitary increase in preference using the median estimates.

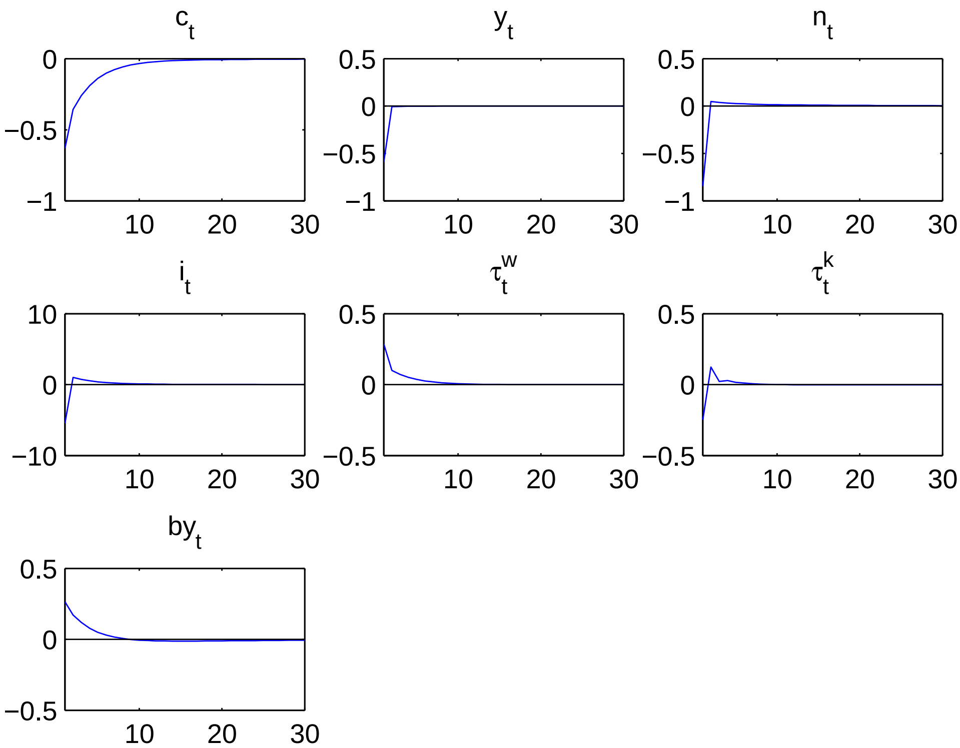

Figure 14: Impulse Response of consumption, GDP and hours worked, investment, labor and capital tax, debt-GDP ratio and government spending to a unitary increase in labor tax using the median estimates.

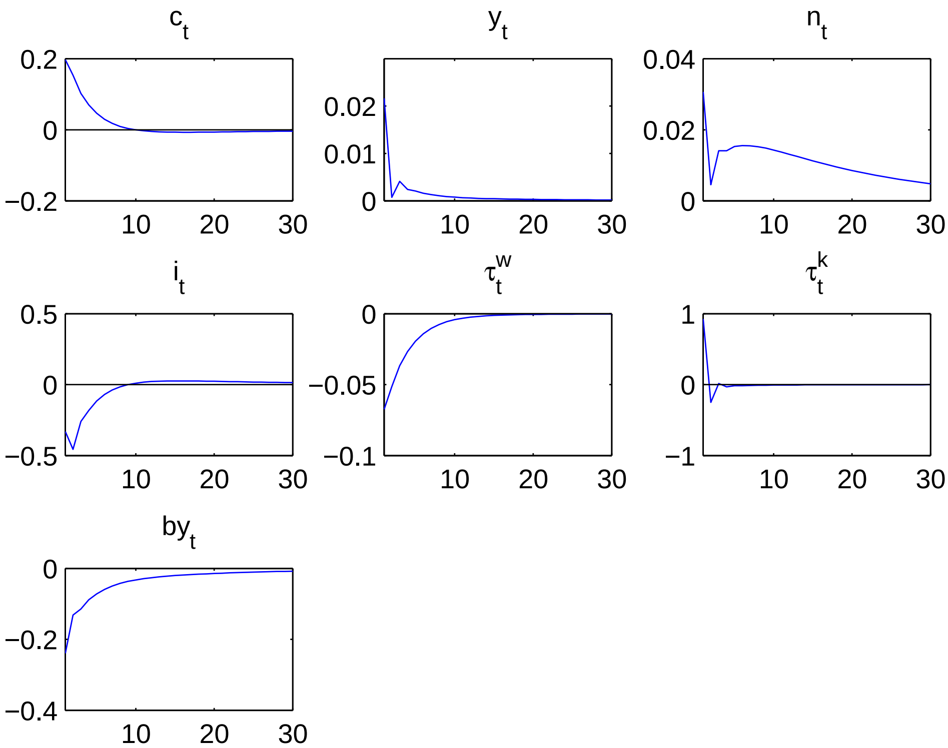

Figure 15: Impulse Response of consumption, GDP and hours worked, investment, labor and capital tax, debt-GDP ratio and government spending to a unitary increase in capital tax using the median estimates.

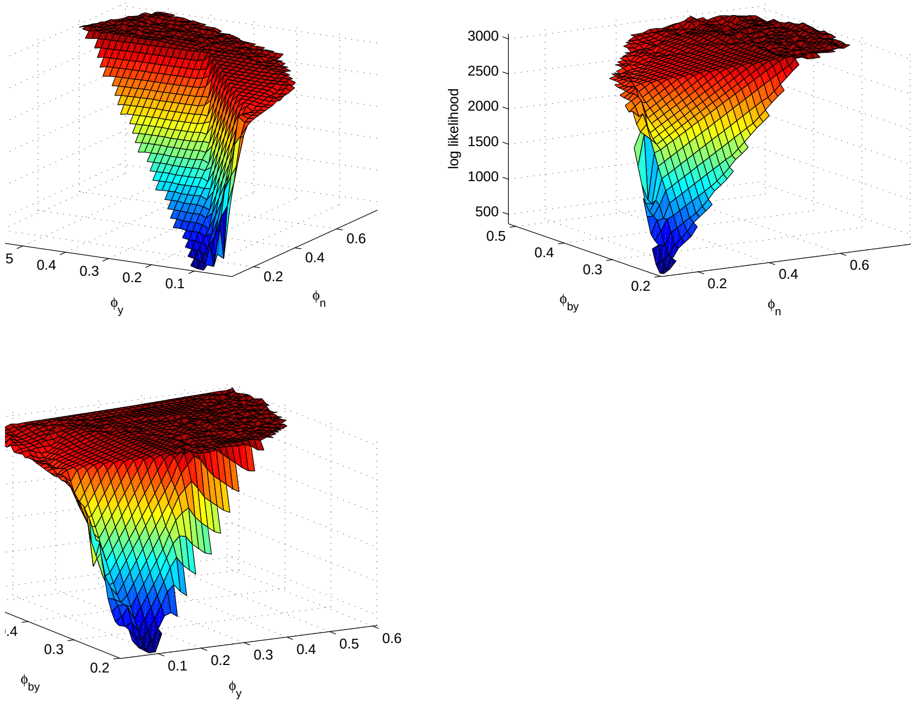

Figure 16: Marginal Log likelihood as function of labor tax policy parameters. In particular, from the top left panel, on the x and y axes are φ_y and φ_n, then φ_{by} and φ_n, then φ_{by} and φ_y. On the z axe there is the Marginal Log likelihood [These plots are constructed using the the draws of the posterior distribution leaving the other parameters free to vary].

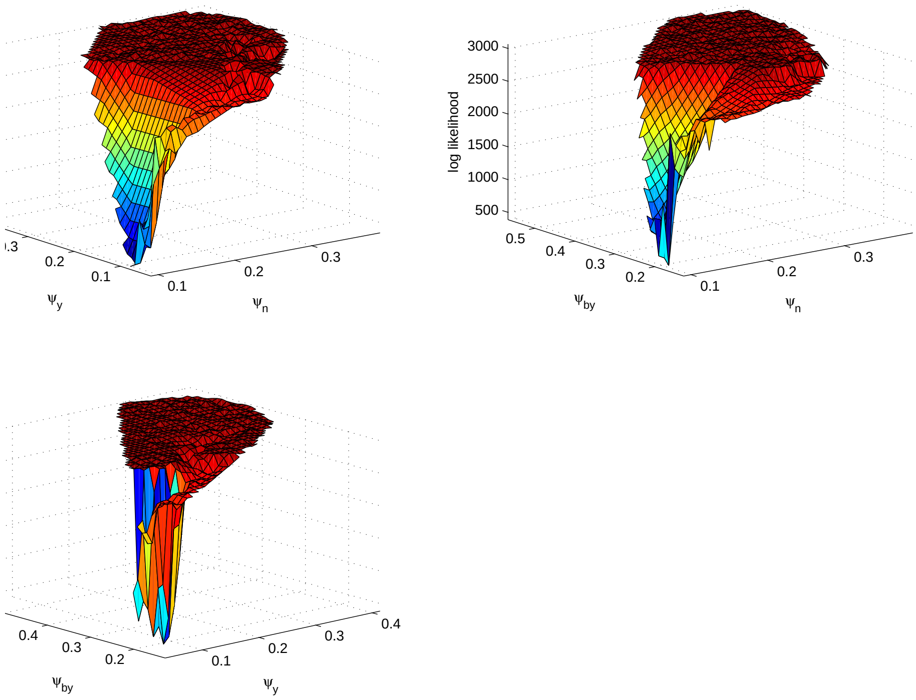

Figure 17: Marginal Log likelihood as function of labor tax policy parameters. In particular, from the top left panel, on the x and y axes there are ψ_{by} and φ_n, then ψ_{by} and ψ_y, then ψ_n and ψ_y. On the z axe there is the Marginal Log likelihood [These plots are constructed using the the draws of the posterior distribution leaving the other parameters free to vary].

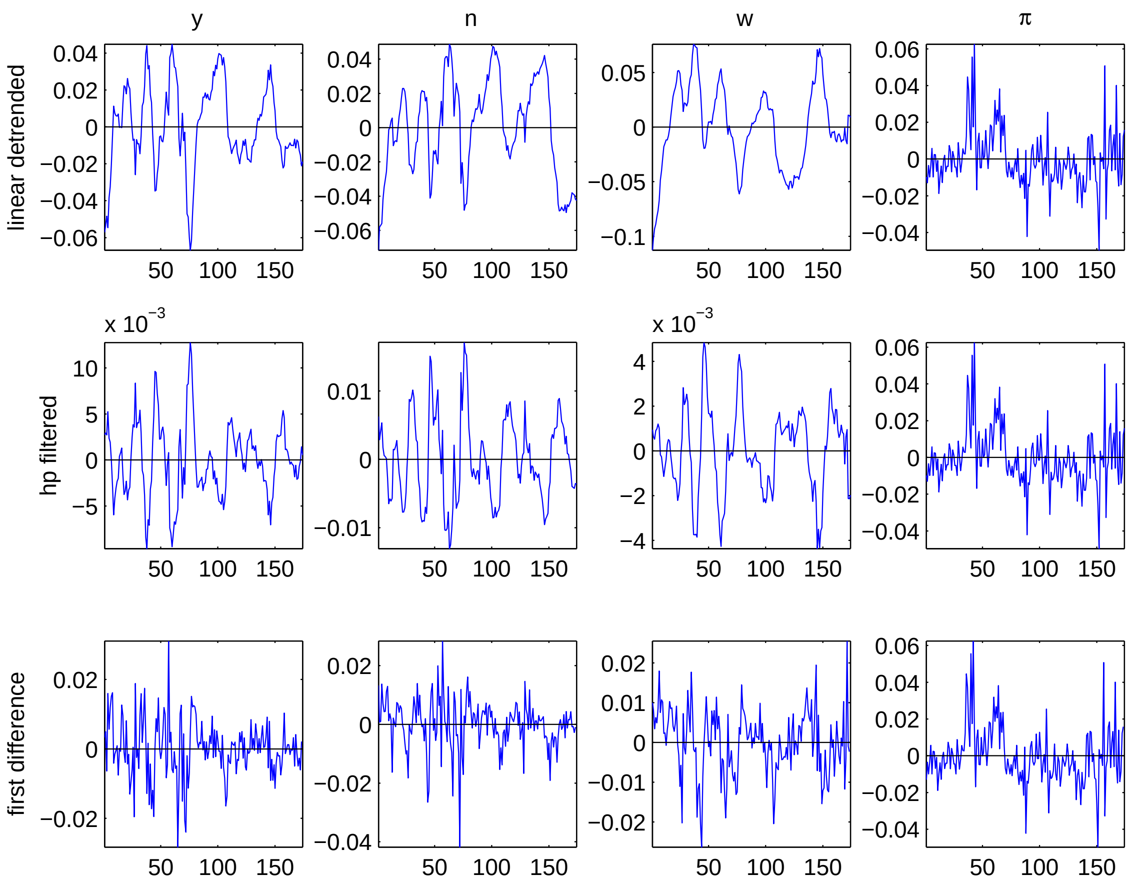

Figure 18: Plots of filtered data; from left to right GDP, hour worked, real wages and inflation. Form top, linear detrended data, hp filtered data and first differenced data.

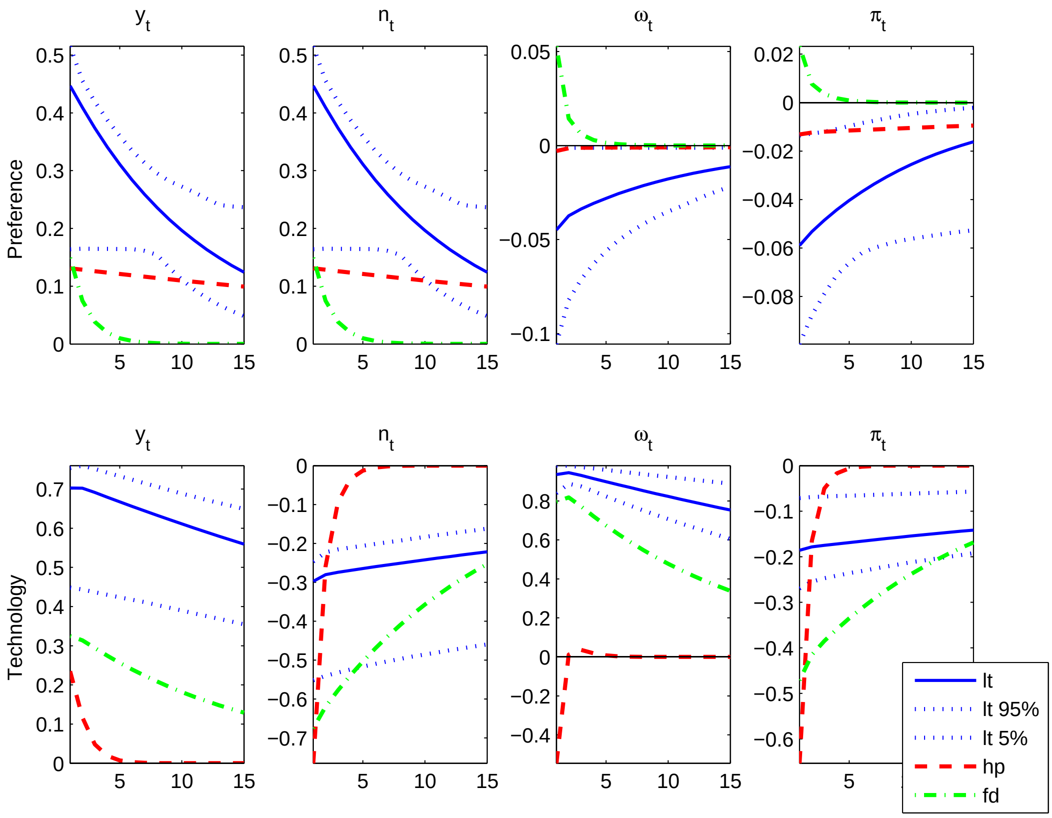

Figure 19: Impulse response of a 1 % increase in the preference (top line) and technology (bottom line) processes for GDP, hour worked, real wages and inflation with 2s approach. The solid blue line represents the response using the median values for the parameters estimates with linear detrended data, the red dashed line the response using the median values for the parameters estimates with hp filtered data, the green dash dotted line the response using the median values for the parameters estimates with first difference data.

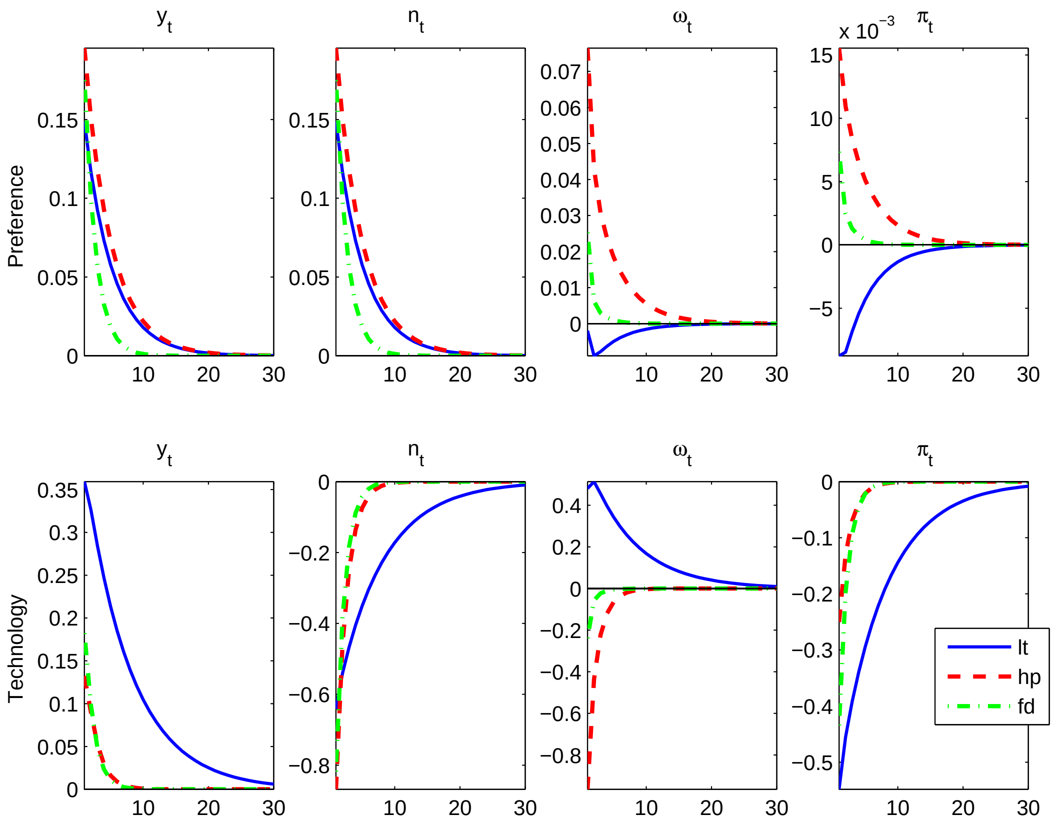

Figure 20: Impulse response of a 1 % increase in the preference (top line) and technology (bottom line) processes for GDP, hour worked, real wages and inflation with 1s approach. The solid blue line represents the response using the median values for the parameters estimates with linear detrended data, the red dashed line the response using the median values for the parameters estimates with hp filtered data, the green dash dotted line the response using the median values for the parameters estimates with first difference data.

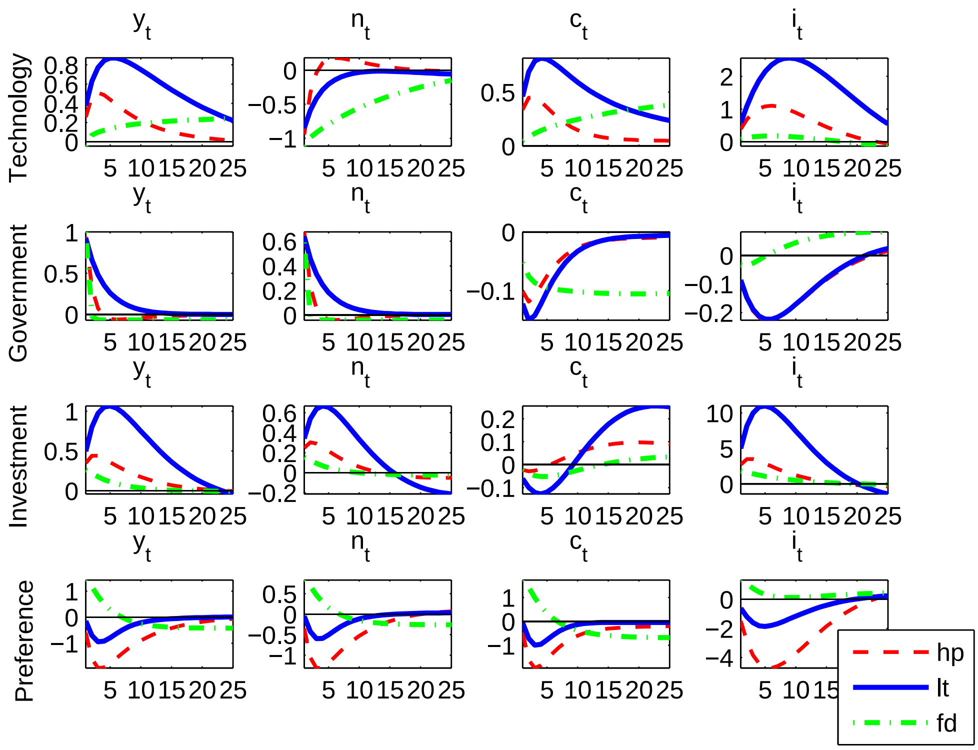

Figure 21: Impulse response of a 1 % increase in the exogenous processes for GDP, employment, consumption, investment with two step approach. The solid blue line represents the response using the median values for the parameters estimates with linear detrended data, the red dashed line the response using the median values for the parameters estimates with HP filtered data, the green dash dotted line the response using the median values for the parameters estimates with first difference data.

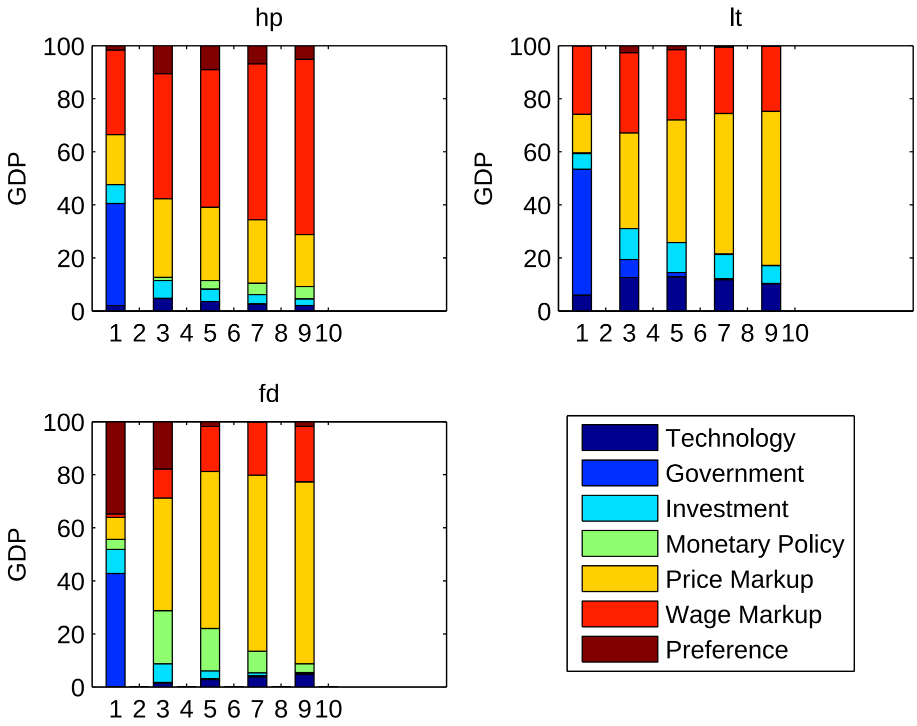

Figure 22: Variance decomposition of GDP in terms of the exogenous processes with the two step approach. The x-axis indicates the k-steps ahead error. The top left plot represents the decomposition using the median values for the parameters estimates with HP filtered data, the top right plot the decomposition using the median values for the parameters estimates with linear detrended data, the bottom plot the decomposition using the median values for the parameters estimates with first difference data.

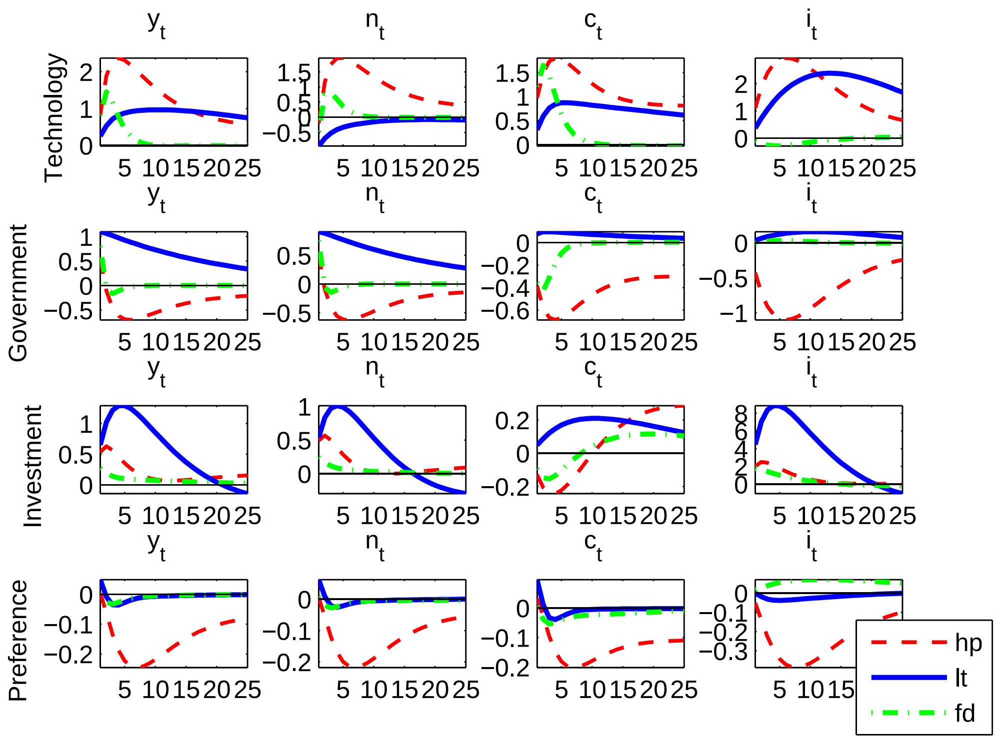

Figure 23: Impulse response of a 1 % increase in the exogenous processes for GDP, employment, consumption, investment with one step approach. The solid blue line represents the response using the median values for the parameters estimates with linear detrended data, the red dashed line the response using the median values for the parameters estimates with HP filtered data, the green dash dotted line the response using the median values for the parameters estimates with first difference data.

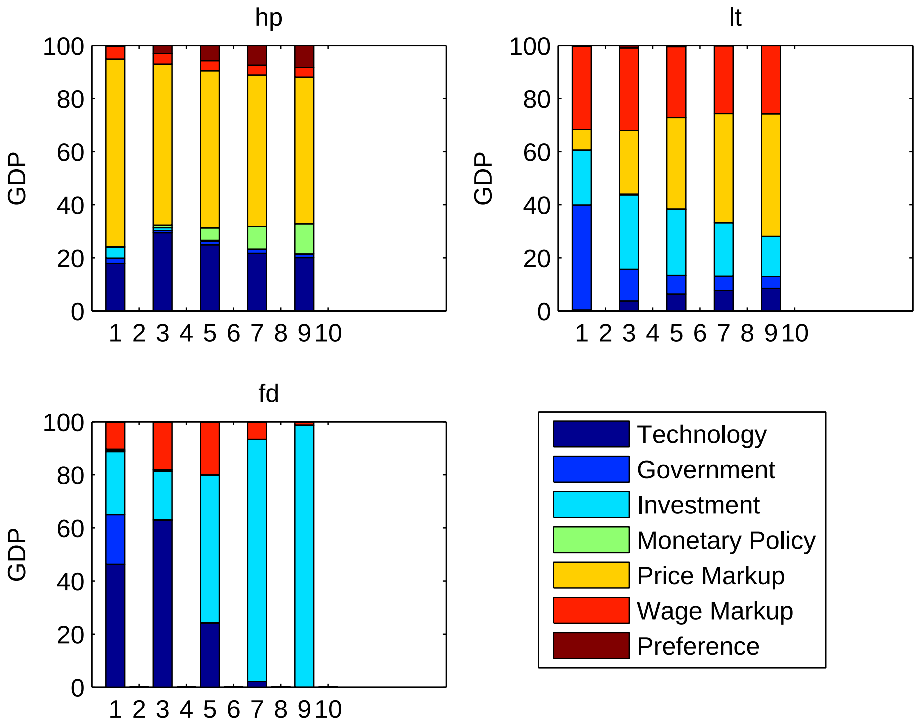

Figure 24: Variance decomposition of GDP in terms of the exogenous processes with the one step approach. The x-axis indicates the k-steps ahead error. The top left plot represents the decomposition using the median values for the parameters estimates with HP filtered data, the top right plot the decomposition using the median values for the parameters estimates with linear detrended data, the bottom plot the decomposition using the median values for the parameters estimates with first difference data.

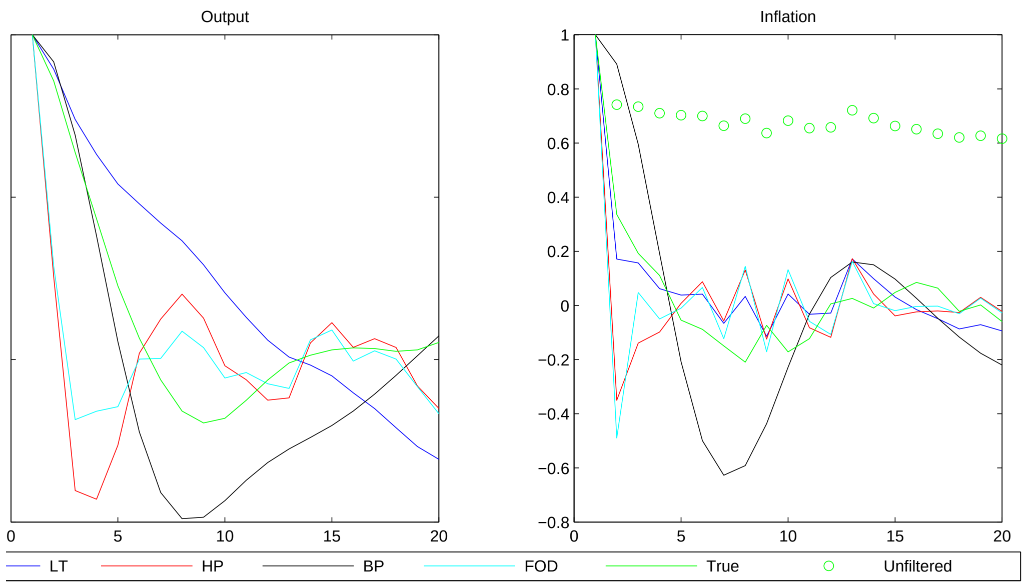

Figure 25: Autocorrelation functions of cyclical output and cyclical inflation

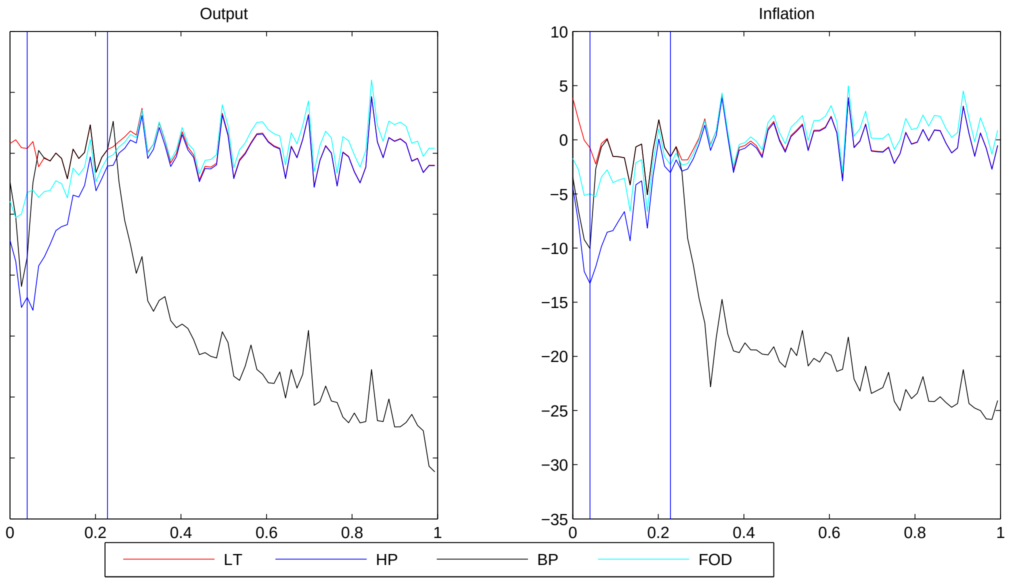

Figure 26: Log spectra of filtering error

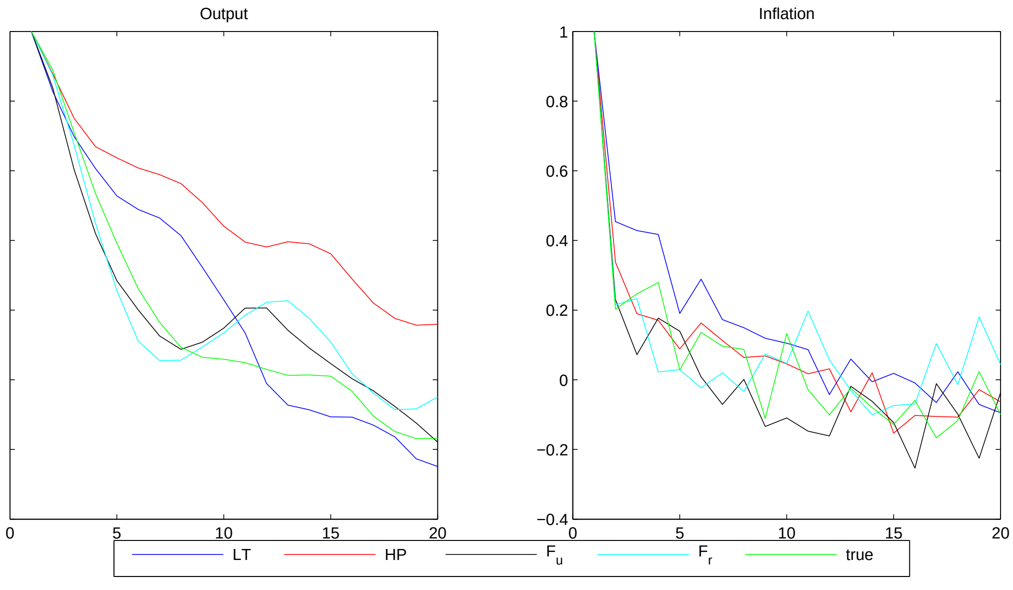

Figure 27: Autocorrelation functions of estimated and true cyclical components

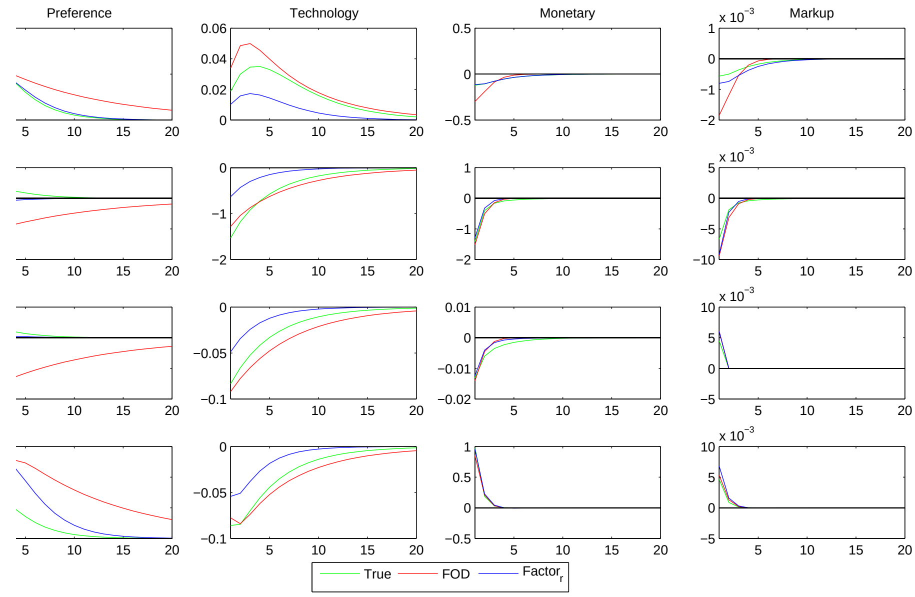

Figure 28: Impulse responses to shocks

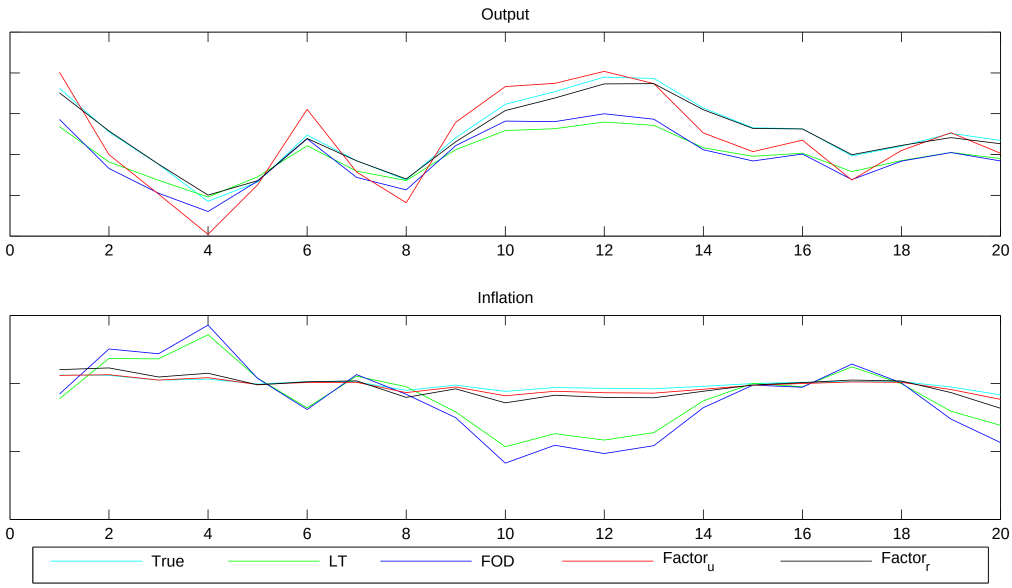

Figure 29: One step ahead forecasts, conditional on a constant interest rate path.

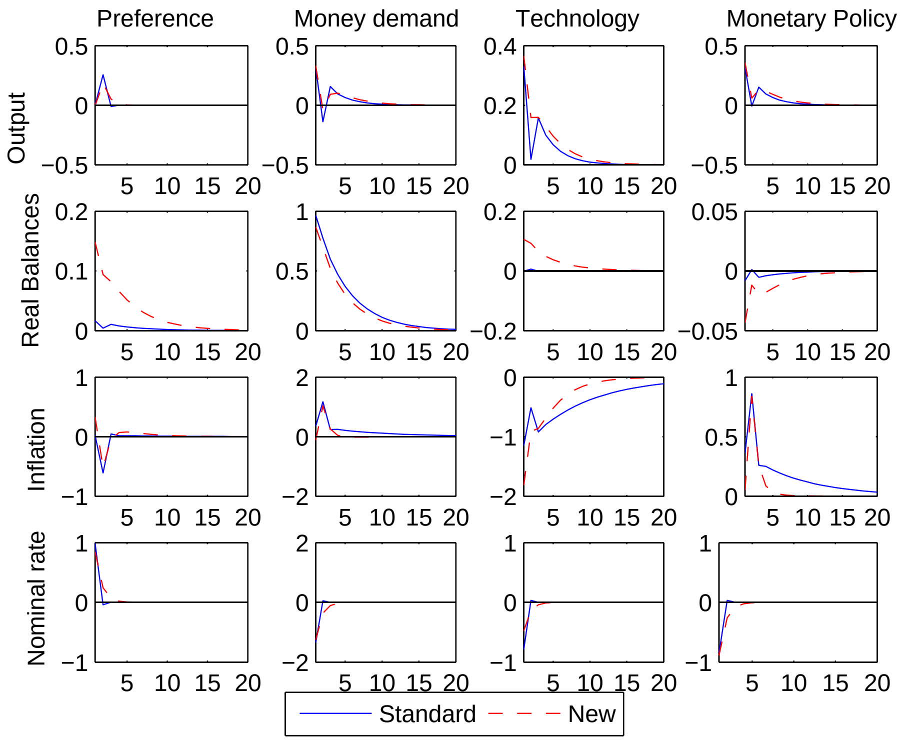

Figure 30: Impulse responses, standard and new approaches