Trend agnostic one step estimation of DSGE models

Ferroni, Filippo
Universitat Pompeu Fabra

04. April 2009

Trend agnostic one step estimation of DSGE models*

Filippo Ferroni†

April 5, 2009

Abstract

DSGE models are currently estimated with a two step approach: data is first filtered and then DSGE structural parameters are estimated. Two step procedures have problems, ranging from trend misspecification to wrong assumption about the correlation between trend and cycles. In this paper, I present a one step method, where DSGE structural parameters are jointly estimated with filtering parameters. I show that different data transformations imply different structural estimates; the two step approach lacks a statistical-based criterion to select among them. The one step approach allows to test hypothesis about the most likely trend specification for individual series and/or use the resulting information to construct robust estimates by Bayesian averaging. The role of investment shock as source of GDP volatility is reconsidered.

JEL Classification: C32, E32.

Keywords: DSGE models, Filters, Structural estimation, Business Cycles.

*I would like to thank Fabio Canova for many discussions and for his encouragement throughout this project. I acknowledge Jordi Galí, Pau Rabanal, Kristoffer Nimark, Thijs van Rens, Vasco Carvalho, Anton Braun and Francesco Caprioli for their suggestions; I thank also the macro seminar participants at Universitat Pompeu Fabra. All the remaining errors are mine.

†Department of Economics, Universitat Pompeu Fabra, Carrer Ramón Trias Fargas, 25-27, 08005 Barcelona, Spain. Office: 20.159, Jaume I building. Phone: 0034 93 542 2689. e-mail: filippo.ferroni@upf.edu

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

[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, Schoerfheide, Smets and Wouters (2007), Justiniano and Primiceri (2008), Rabanal (2006), among others.

(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.

In this paper, 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 (2007) 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 (2006)). 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 paper is organized as follows. Section 2 presents the econometric methodology with emphasis on the two approaches. In Section 3 the two procedures are confronted under various Monte Carlo experiments; results and biases are reported. Section 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 5 concludes.

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

$$y = y^\tau + y^c \tag{1}$$

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

$$y_t^c = RR(\theta^m)x_{t-1} + SS(\theta^m)z_t \tag{2}$$

$$x_t = PP(\theta^m)x_{t-1} + QQ(\theta^m)z_t \tag{3}$$

$$z_{t+1} = NN(\theta^m)z_t + \nu_{t+1} \tag{4}$$

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 of a model is usually computed using the Kalman filter after having defined a linear state space[3], of the form

$$Y_t = H(\theta)s_t + u_t \tag{5}$$

$$s_{t+1} = F(\theta)s_t + G(\theta)\omega_{t+1} \tag{6}$$

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 (5) is the measurement equation, which relates a set of observable variables, Y_t, to a set of (latent) state variables, s_t. State evolves along time according to equation (6).

2.1 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

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

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

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)-(4), fit the state space representation, (5) and (6), 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)).

2.2 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 (5)-(6). Details are reported in the appendix.

2.2.1 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.

$$y_t^{\tau} = A + B * t + \eta_t \tag{7}$$

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.

2.2.2 First-Difference-DSGE setup

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

$$y_t^{\tau} = \gamma + \Gamma y_{t-1} + \eta_t \tag{8}$$

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.

2.2.3 Hodrick-Prescott-DSGE setup

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

$$y_{t+1}^{\tau} = y_t^{\tau} + \mu_t \tag{9}$$

$$\mu_{t+1} = \mu_t + \zeta_{t+1} \quad (10)$$

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 (9) and (10). 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.

2.3 Estimation

Bayesian methods are employed to obtain the posterior distribution of the structural and non-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.

In the 2-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 1-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, $\mathbb{M} = \{\mathcal{M}, \mathcal{F}\}$, $Y = y$ and $\theta = (\theta^m, \theta^f)$, and posterior distribution of parameters is

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

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, equations (2)-(4), given $\theta_\dagger$; if indeterminacy or no-existence set $\mathcal{L}(Y | \theta_\dagger; \mathbb{M}) = 0$.[4]
3. evaluate the likelihood of the system of equations (5)-(6) 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). 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.

[4] In the one step approach, a candidate draw $\theta_\dagger = (\theta^m_\dagger, \theta^f_\dagger)$ is rejected, if $\theta^m_\dagger$ implies non-existence or indeterminacy for the system (2)-(4).

2.4 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. 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 1s approach, the predictive density of the data, conditional on the DSGE model, $\mathcal{M}$, and on the trend specification, $\mathcal{F}$, is

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

where $\theta = (\theta^m, \theta^f)$. 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 2s 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.

2.5 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

$$y_t^\tau = A_t + B_t * t + \eta_t \tag{11}$$
$$A_{t+1} = A_t + \eta_{t+1}^A \tag{12}$$
$$B_{t+1} = B_t + \eta_{t+1}^B \tag{13}$$

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

$$y_t^\tau = \gamma_t + \Gamma y_{t-1} + \eta_t \tag{14}$$
$$\gamma_{t+1} = \gamma_t + \eta_{t+1}^\gamma \tag{15}$$

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

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 (2006).

3.1 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[5]

$$\lambda_t = \epsilon_t^{\chi} - \sigma_c c_t \tag{16}$$

$$y_t = \epsilon_t^{a} + n_t \tag{17}$$

$$mc_t = \omega_t - (y_t - n_t) \tag{18}$$

$$mrs_t = -\lambda_t + \sigma_n n_t \tag{19}$$

$$\omega_t = mrs_t \tag{20}$$

$$r_t = \rho_R r_{t-1} + (1-\rho_R)(\rho_\pi \pi_t + \rho_y y_t) + \epsilon_t^{r} \tag{21}$$

$$\lambda_t = E_t \lambda_{t+1} + r_t - E_t \pi_{t+1} \tag{22}$$

$$\pi_t = k_p(mc_t + \epsilon_t^{\mu}) + \beta E_t \pi_{t+1} \tag{23}$$

$$\epsilon_t^{\chi} = \rho_\chi \epsilon_{t-1}^{\chi} + \nu_t^{\chi} \tag{24}$$

$$\epsilon_t^{a} = \rho_a \epsilon_{t-1}^{a} + \nu_t^{a}. \tag{25}$$

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 (16) 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

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

to follow an AR(1) process, equation (24). Equation (17) 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 (25). The difference between real wage, ω_t, and the marginal product of labor, $y_t - n_t$, defines the marginal cost, mc_t, equation (18). 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 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 (21) is the monetary rule. Equation (22) 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 (23) 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)-(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,

$$y_{t+1}^{\tau} = y_t^{\tau} + \mu_t$$
$$\mu_{t+1} = \mu_t + \zeta_{t+1}$$

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

3.2 Prior Selection

Table 1 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.

3.3 Bias Computation

I generate data using four different population values, see Table 2. 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 2, 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 2.3 until convergence is achieved[6].
2. I then discard the first 300,000 draws and keep randomly one every 1,000 draws, θ_j^s, and compute
$$bias_\ell^s = \frac{1}{L}\sum_{j=1}^{L} |\frac{\theta_j^s - \theta_{true}^s}{\theta_{true}^s}|$$

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

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.

3.4 Bias in small samples

Tables 3 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 3), 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 3) 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 4 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 9), 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.

3.5 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 5 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 6, 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 3. 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

$$\eta_t = A_1 z_t + v_t \tag{26}$$

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

$$\zeta_t = A_1 z_t + v_t \tag{27}$$

As before, I first consider the bias in the estimates when data has deterministic trends and then when data has stochastic ones. Tables 7 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 (1)-(4), (7) and (26), and parameters are estimated assuming that the true DGP is given by (1)-(4), (7). 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 (26) 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 8 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 10 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.

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 3.1. 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.

4.1 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 1.

Table 11 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 1) 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 1), 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 2 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 11 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.

$$PO_{\mathcal{F}_k,\mathcal{F}_{lt}} = \frac{g(\mathcal{F}_k)}{g(\mathcal{F}_{lt})} \times \frac{p(y|\mathcal{M},\mathcal{F}_k)}{p(y|\mathcal{M},\mathcal{F}_{lt})}$$

for $k = hp, fd$. The first term is the ratio between prior filters probability, and the second term is the Bayes Factor; assuming equal ex ante probability to each filter, Posterior Odds and Bayes Factors coincide. 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 3 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.

4.2 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 2.2. 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[7]

$$y_t = \alpha\phi_p k_t + (1-\alpha)\phi_p n_t + \phi_p \epsilon_t^a \tag{28}$$

$$y_t = \epsilon_t^g + c/y c_t + i/y i_t + r^k * k/y z_t \tag{29}$$

$$k_t = k_{t-1}^s + z_t \tag{30}$$

$$k_t = \omega_t + n_t - \frac{\psi}{1-\psi} z_t \tag{31}$$

$$mc_t = \alpha \frac{\psi}{1-\psi} z_t + (1-\alpha)\omega_t - \epsilon_t^a \tag{32}$$

$$k_t^s = (1-\delta)k_{t-1}^s + i/k i_t + i/k \varphi \epsilon_t^i \tag{33}$$

$$(1+\beta i_p)\pi_t = \beta E_t \pi_{t+1} + i_p \pi_{t-1} + k_p mc_t + \epsilon_t^p \tag{34}$$

$$(1+\beta i_\omega)\omega_t = \omega_{t-1} + \beta E_t(\omega_{t+1} + \pi_{t+1}) + i_\omega \pi_{t-1} + (1+\beta i_\omega)\pi_t - k_\omega \mu \omega_t + \epsilon_t^\omega \tag{35}$$

$$c_t = \frac{1}{1+h}(E_t c_{t+1} - h c_{t-1}) + c_1(n_t - E_t n_{t+1}) - c_2(r_t - E_t \pi_{t+1}) + \epsilon_t^b \tag{36}$$

$$q_t = -(r_t - E_t \pi_{t+1}) + \frac{\sigma_c(1+h)}{(1-h)} \epsilon_t^b + E_t(q_1 z_{t+1} + q_2 q_{t+1}) \tag{37}$$

$$i_t = \frac{1}{1+\beta} i_{t-1} + \frac{\beta}{1+\beta} E_t i_{t+1} + \frac{1}{\varphi(1+\beta)} q_t + \epsilon_t^i \tag{38}$$

Variables without the time subscript are steady state values and with time subscript are deviation from the steady state.

Equation (28) 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 (29) 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 (30). Equation (31) 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 (32) gives an

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

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 (33); 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 (34) 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 (35) 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 (36) 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 (37) 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, (38), by which current value of investment depends on past and expected future value of capital and on current value of the stock of capital.

4.2.1 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)-(4) where the vector of endogenous state[8] 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 12 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;

[8] 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 $t-1$.

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.

4.2.2 Model implications discussion

The main point I want to stress is that model implications are quite sensitive to data transformations. Figure 4 presents the effects of changes in the exogenous processes to output, hours worked, consumption and investment using 2s estimates. The response 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[9], 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[10] 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 comovements 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

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

[10]Different results for heterogenous agents setups, see Galí, López-Salido and Vallés (2007)

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 13. As before, data strictly prefers a specification with unit root in the trend dynamics. Figure 6 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, estimated DSGE models tend to explain the volatility of output manly in terms of mark-up shocks[11], 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 mark-up shocks in explaining the GDP volatility. Figure 5 reports the k-step ahead forecast 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 7 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).

[11] 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.

5 Conclusion

In this paper, 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.

References

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.

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*.

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

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.: 2006, Back to square one: identification issues in DSGE models, *Computing in Economics and Finance 2006 196*, Society for Computational Economics.

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

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.

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

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.

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

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.: 2007, Limited Information estimation and Evaluation of DSGE models, *Economics Working Papers 6*, NCER.

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 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.

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

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.

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

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

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

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**.

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.

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 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).

A.1 Model

The representative household has a preference for variety: the consumption index is

$$C_t = \left(\int_0^1 C_t(j)^{\frac{\epsilon_t - 1}{\epsilon_t}} dj \right)^{\frac{\epsilon_t}{\epsilon_t - 1}} \tag{39}$$

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

$$C_t(j) = \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon_t} C_t \tag{40}$$

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.

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[X_t \frac{1}{1-\sigma_c} C_t^{1-\sigma_c} - \frac{1}{1+\sigma_n} N_t^{1+\sigma_n} \right] \tag{41}$$

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,

$$P_t C_t + b_t B_t = B_{t-1} + W_t N_t \tag{42}$$

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:

$$0 = X_t C_t^{-\sigma_c} - \mathcal{L}_t \tag{43}$$

$$0 = -N_t^{-\sigma_n} - \mathcal{L}_t \frac{W_t}{P_t} \tag{44}$$

$$1 = E_t \left[\ \beta \tfrac{\mathcal{L}_{t+1}}{\mathcal{L}_t} \tfrac{P_{t+1}}{P_t} R_t \ \right] \tag{45}$$

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,

$$Y_t(j) = A_t N_t(j) \tag{46}$$

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

$$\Pi_t^{1-\epsilon_t} = \zeta_p + (1 - \zeta_p)(P_t^*/P_{t-1})^{1-\epsilon_t} \tag{47}$$

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

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

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 $\Pi_{t,t+k} \equiv P_{t+k}/P_t$

$$\sum_{k=0}^{\infty} \zeta_p^k E_t Q_{t,t+k} Y_{t+k|t} \left[\frac{P_t^*}{P_{t-1}} - \mathcal{M}_{t+k} MC_{t+k|t} \Pi_{t-1,t+k}) \right] = 0 \tag{48}$$

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

$$C_t = Y_t$$

A.2 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, (43)-(45), 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, (48), 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

$$p_t^* - p_{t-1} = \beta\zeta_p E_t(p_{t+1}^* - p_t) + (1-\beta\zeta_p)[\epsilon_t^{\mu} + mc_t] + \pi_t \tag{49}$$

The log linearization of law of motion of price, (47), leads to

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

Combining the latter equation with (49) 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 $\epsilon_t^r \sim N(0, \sigma_r^2)$

B State Spaces

Equations (1), (2)-(4) and (7) can be cast into the linear state space representation (5)-(6), 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 (1),(2)-(4) and (8) fit the state space representation (5)-(6) 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 (1),(2)-(4) and (9)-(10) can be cast into the linear state space representation (5) and (6), 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 (1),(2)-(4), (11)-(13) fit the state space representation,(5)-(6), 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 (1),(2)-(4), (14)-(15) 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^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 1: 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 2: Structural parameters: Population values. 'LP' stands for low persistence, 'HP' for high persistence, 'HV' stands for high volatility, 'LV' for low volatility.

TV	LP 2s	LP 1s	HP 2s	HP 1s	HV 2s	HV 1s	LV 2s	LV 1s	Average 2s	Average 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 3: Bias comparison between 2 step and 1 step. Data is generated with the population values of Table 2 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 4: Bias comparison between 2 step and 1 step. Data is generated with the population values of Table 2 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 5: Bias comparison under misspecification. Data is generated with the population values of Table 2 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 6: Bias comparison under misspecification. Data is generated with the population values of Table 2 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 7: Correlation Assumption. Data is generated with the population values of Table 2 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 8: Correlation Assumption. Data is generated with the population values of Table 2 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 9: 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 10: 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 11: 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 12: 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 13: Posterior Odds across specifications.

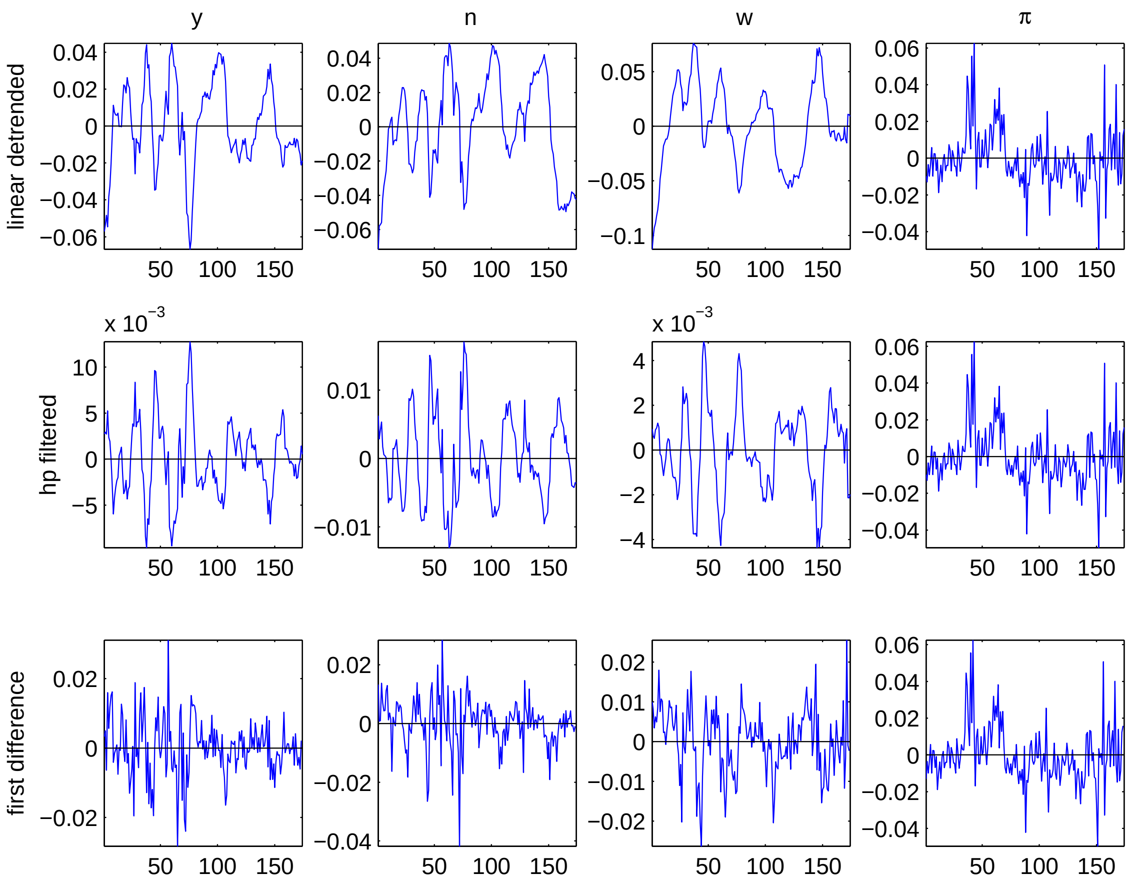

Figure 1: Plots of filtered data; from left to right GDP, hour worked, real wages and inflation. Form top, linear detrended data, hp filtered data and first differenced data.

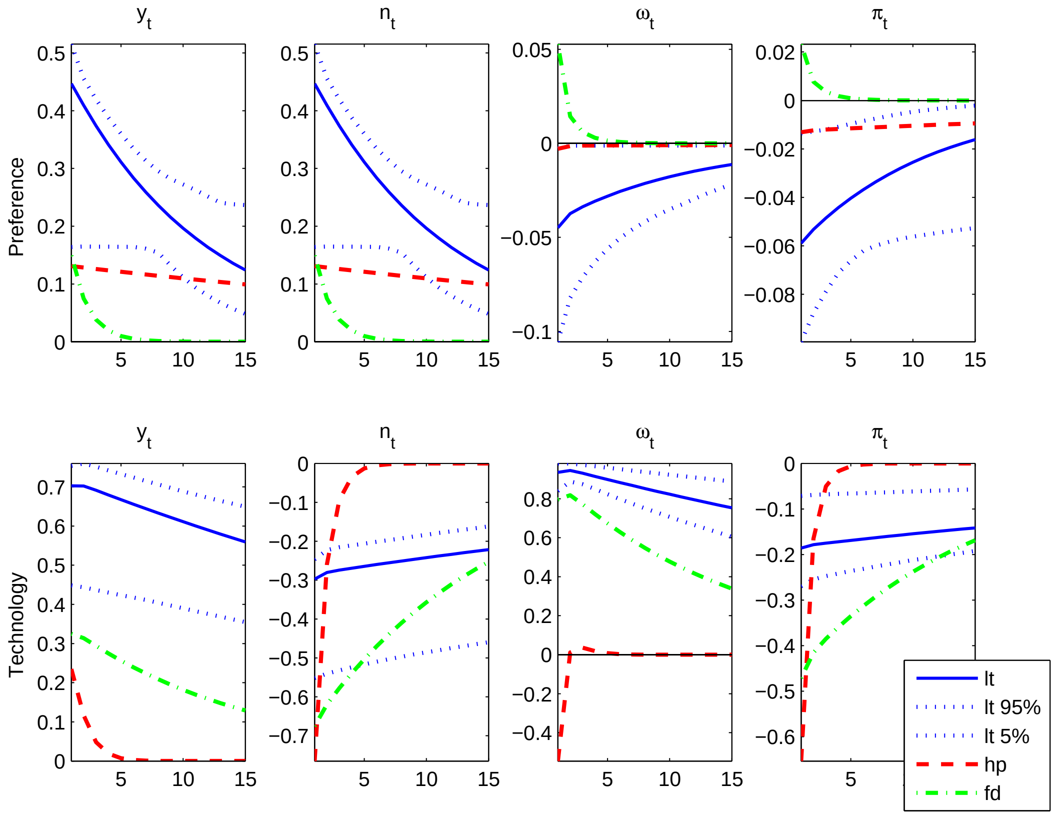

Figure 2: Impulse response of a 1 % increase in the preference (top line) and technology (bottom line) processes for GDP, hour worked, real wages and inflation with 2s approach. The solid blue line represents the response using the median values for the parameters estimates with linear detrended data,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

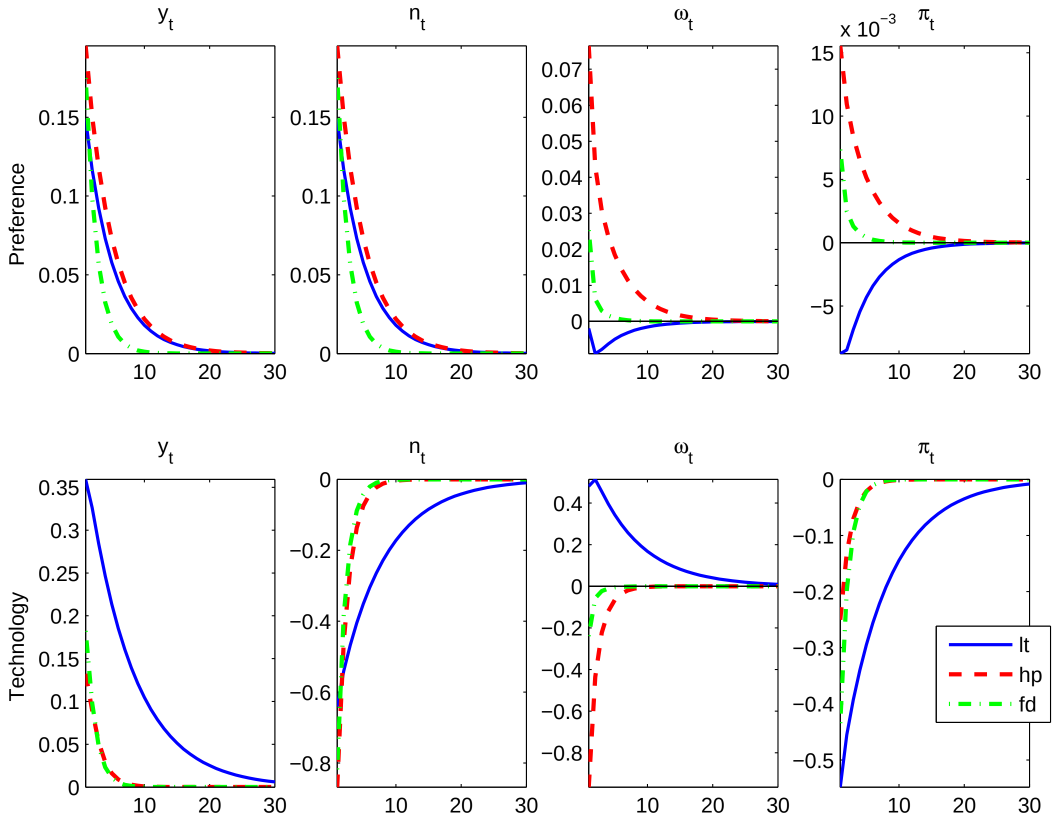

Figure 3: Impulse response of a 1 % increase in the preference (top line) and technology (bottom line) processes for GDP, hour worked, real wages and inflation with 1s approach. The solid blue line represents the response using the median values for the parameters estimates with linear detrended data, 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.

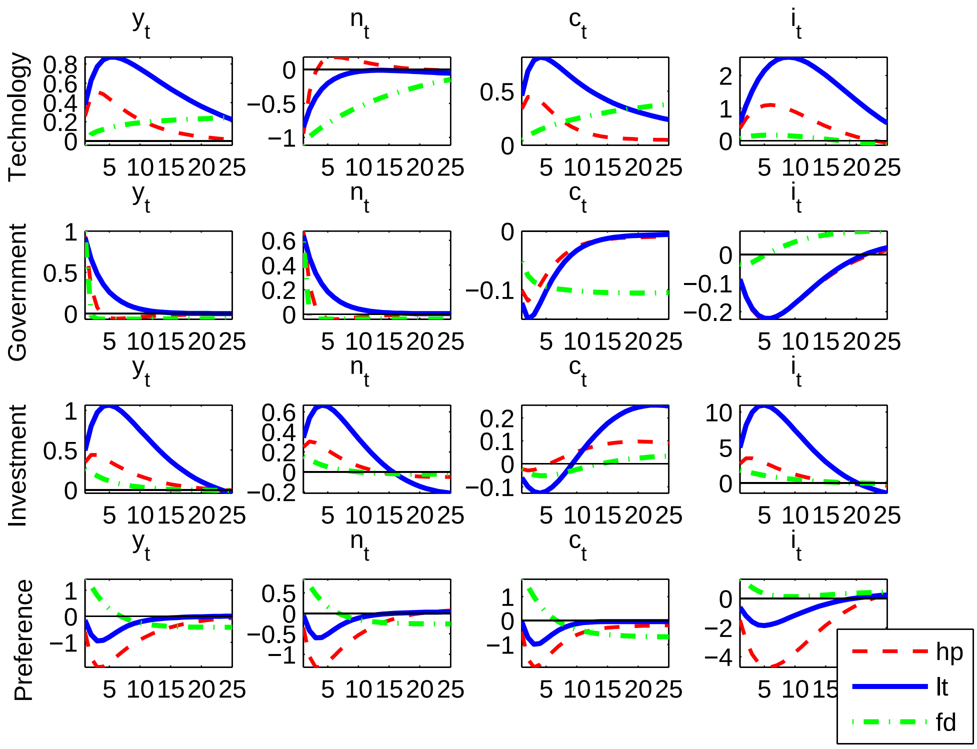

Figure 4: Impulse response of a 1 % increase in the exogenous processes for GDP, employment, consumption, investment with two step approach. The solid blue line represents the response using the median values for the parameters estimates with linear detrended data, the red dashed line the response using 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.

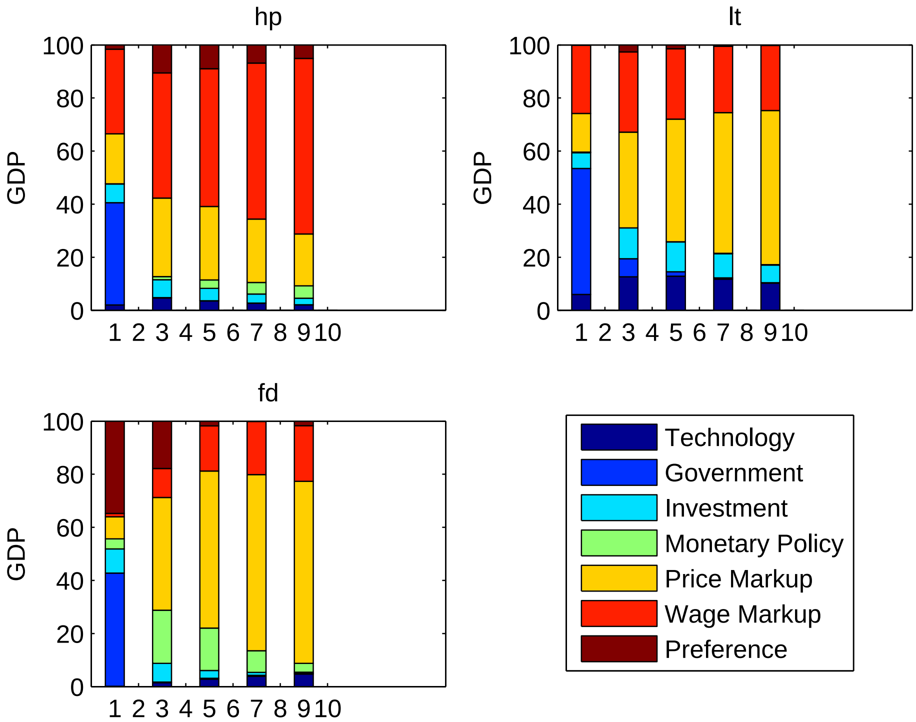

Figure 5: Variance decomposition of GDP in terms of the exogenous processes with the two step approach. The x-axis indicates the k-steps ahead error. The top left plot represents the decomposition using the median values for the parameters estimates with HP filtered data, the top right plot the decomposition 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.

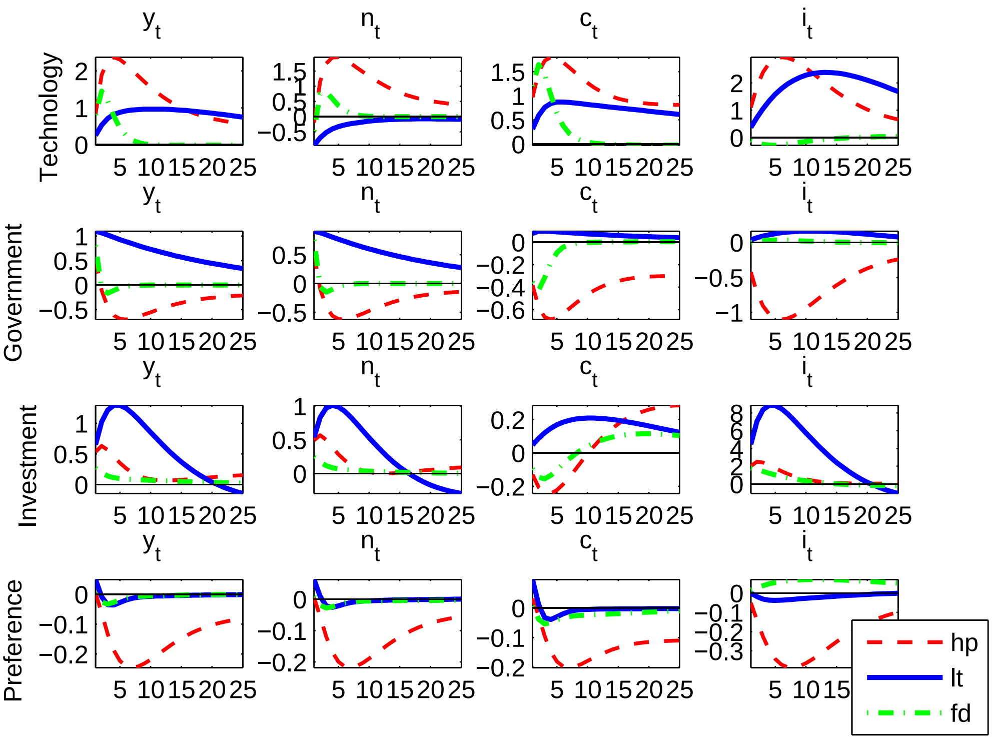

Figure 6: Impulse response of a 1 % increase in the exogenous processes for GDP, employment, consumption, investment with one step approach. The solid blue line represents the response using the median values for the parameters estimates with linear detrended data, the red dashed line the response using 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.

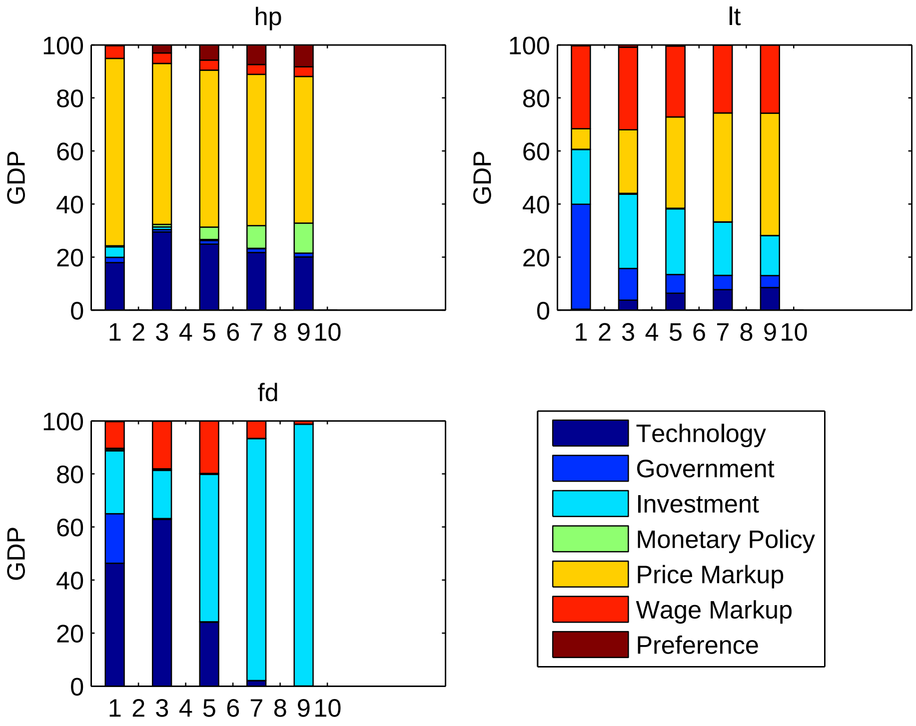

Figure 7: Variance decomposition of GDP in terms of the exogenous processes with the one step approach. The x-axis indicates the k-steps ahead error. The top left plot represents the decomposition using the median values for the parameters estimates with HP filtered data, the top right plot the decomposition 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.