# Bridging DSGE models and the raw data

Fabio Canova *
EUI and CEPR

March 13, 2014

**Abstract**

A method to estimate DSGE models using the raw data is proposed. The approach links the observables to the model counterparts via a flexible specification which does not require the model-based component to be located solely at business cycle frequencies, allows the non model-based component to take various time series patterns, and permits certain types of model misspecification. Applying standard data transformations induce biases in structural estimates and distortions in the policy conclusions. The proposed approach recovers important model-based features in selected experimental designs. Two widely discussed issues are used to illustrate its practical use.

JEL classification: E3, C3.

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

---

*I would like to thank Urban Jermann and Bob King (the editors), two anonymous referees, Frank Schorfheide, Tim Cogley, Giorgio Primiceri, Tao Zha, Chris Sims, Harald Uhlig and the participants of seminars at the Bank of England, Bank of Italy, BIS, UCL, Yale, EUI, University of Modena, University of Sussex, Federal Reserve Bank of New York, University of Zurich, ESSIM, the Fed of Atlanta workshop on Methods of applications for DSGE models; the $I_t$ workshop in Time Series Econometrics, Zaragoza; and the conference Recent Development of Dynamic Analysis in Economics, Seoul; The Gent workshop in Macroeconomics for comments and suggestions. The financial support of the Spanish Ministry of Economy and Competitiveness, through the grants ECO2009-08556, ECO2012-33247, of the Barcelona Graduate School of Economics and of the European University Institute is gratefully acknowledged. Department of Economics, EUI, Via della Piazzuola 43, 50133 Firenze, Italy, fabio.canova@eui.eu.

# 1 Introduction

There have been considerable developments in the specification of DSGE models in the last few years. Steps forward have also been made in the estimation of these models. Despite recent efforts, structural estimation of DSGE models is conceptually and practically difficult. For example, classical estimation is asymptotically justified only when the model is the generating process (DGP) of the actual data, up to a set of serially uncorrelated measurement errors, and standard validation exercises are meaningless without such an assumption. Identification problems (see e.g. Canova and Sala, 2009) and numerical difficulties are widespread. Finally, while the majority of the models investigators use are intended to explain only the cyclical portion of observable fluctuations, both permanent and transitory shocks may produce cyclical fluctuations, and macroeconomic data contain many types of fluctuations, some of which are hardly cyclical.

The generic mismatch between what models want to explain and what the data contain creates headaches for applied investigators. A number of approaches, reflecting different identification assumptions, have been used:

• Fit a model driven by transitory shocks to the observables filtered with an arbitrary statistical device (see Smets and Wouters, 2003, Ireland, 2004a, Rubio and Rabanal, 2005, among others). Such an approach is problematic for at least three reasons. First, since the majority of statistical filters can be represented as a symmetric, two-sided moving average of the raw data, the timing of the information is altered and dynamic responses hard to interpret. Second, while it is typical to filter each real variable separately and to demean nominal variables, there are consistency conditions that must hold - a resource constraint need not be satisfied if each variable is separately filtered - and situations when not all nominal fluctuations are relevant. Thus, specification errors can be important. Finally, contamination errors could be present. For example, a Band Pass (BP) filter only roughly captures the power of the spectrum at the frequencies corresponding to cycles with 8-32 quarters average periodicity in small samples and taking growth rates greatly amplifies the high frequency content of the data. Thus, rather than solving the problem, such an approach adds to the difficulties faced by applied researchers.

• Fit a model driven by transitory shocks to transformations of the observables which, in theory, are void of non-cyclical fluctuations, e.g. consider real "great ratios" (as in Cogley, 2001, and McGrattan, 2010) or nominal "great ratios" (as in Whelan, 2005). As Figure 1 shows, such transformations may not solve the problem because many ratios still display low frequency movements. In addition, since the number and the nature of the shocks driving non-cyclical fluctuations needs to be a-priori known, specification errors may be produced.

• Construct a model driven by transitory and permanent shocks; scale the model by the assumed permanent shocks; fit the transformed model to the observables transformed in the same way (see e.g. Del Negro et al., 2006, Fernandez and Rubio, 2007, Justiniano et al., 2010, among others). Such an approach puts stronger faith in the model than previous ones, explicitly imposes a consistency condition between the theory and the observables, but it is not free of problems. For example, since the choice of which shock is permanent is often driven by computational rather than economic considerations, specification errors could be present. In addition, structural parameter estimates may depend on nuisance features, such as the shock which is assumed to be permanent and its time series characteristics. As Cogley (2001) and Gorodnichenko and Ng (2010) have shown, misspecification of these nuisance features may lead to biased estimates of the structural parameters.

• Construct a model driven by transitory and/or permanent shocks; estimate the structural parameters by fitting the transformed model to the transformed data over a particular frequency band (see e.g. Diebold et. al, 1998, Christiano and Vigfusson, 2003). This approach is also problematic since it inherits the misspecification problems of the previous approach and the filtering problems of the first approach.

The paper shows first that the approach one takes to match the model to the data matters for structural parameter estimation and for economic inference. Thus, unless one has a strong view about what the model is supposed to capture and with what type of shocks, it is difficult to credibly select among various structural estimates (see Canova, 1998). In general, all preliminary data transformations should be avoided if the observed data is assumed to be generated by rational agents maximizing under constraints in a stochastic environment. Statistical filtering does not take into account that cross equation restrictions can rarely be separated by frequency, that the data generated by a DSGE model has power at all frequencies and that, if permanent and transitory shocks are present, both the permanent and the transitory component of the data will appear at business cycle frequencies. Model based transformations impose tight restrictions on the long run properties of the data. Thus, any deviations from the imposed structure must be captured by the shocks driving the transformed model, potentially inducing parameter distortions.

As an alternative, one could estimate the structural parameters by creating a flexible non-structural link between the DSGE model and the raw data that allows model-based and non model-based components to have power at all frequencies. Since the non model-based component is intended to capture aspects of the data in which the investigator is not interested but which may affect inference, specification errors could be reduced. In addition, because the information present at all frequencies is used in the estimation, filtering distortions are eliminated and inefficiencies minimized. The methodology can be applied to models featuring transitory or transitory and permanent shocks and only requires that interesting features of the data are left out from the model - these could be low frequency movements of individual series, different long run dynamics of groups of series, etc.. The setup has two other advantages over competitors: structural estimates reflect the uncertainty present in the specification of non model-based features; what the model leaves out at interesting frequencies is quantifiable with R-squared type measures. Thus, one can "test" the structure and to evaluate the explanatory power of additional shocks.

The approach is related to earlier work of Altug (1989), McGrattan(1994) and Ireland (2004b). As in these papers, a non-structural part is added to a structural model prior to estimation, but here the non-structural part is not designed to eliminate singularity. More crucially, the approach does not substitute for theoretical efforts designed to strengthen the ability of DSGE models to account for all observable fluctuations. But it can fill the gap between what is nowadays available and such a worthy long run aspiration, giving researchers a rigorous tool with which to address policy questions.

Using a simple experimental design and two practically relevant cases, the paper documents the biases that standard transformations produce, interprets them using the tools developed in Hansen and Sargent (1993), and shows that crucial parameters are better estimated with the proposed procedure. To highlight how the approach can be used in practice, the paper examines finally two questions greatly discussed in macroeconomics: the time variations in the policy activism parameter and the sources of output and inflation fluctuations.

To focus attention on the issues of interest, two simplifying assumptions are made: (i) the estimated DSGE model features no missing variables or omitted shocks and (ii) the number of structural shocks equals the number of endogenous variables. While omitted variables and singularity issues are important, and the semi-structural methods suggested in Canova and Paustian (2011) produce more robust inference when they are present, I sidestep them because the problems discussed here occur regardless of whether (i)-(ii) are present or not.

The rest of the paper is organized as follows. The next section presents estimates of the structural parameters when a number of statistical and model based transformations are employed. Section 3 discusses the methodology. Section 4 compares approaches using a simple experimental design. Section 5 examines two economic questions. Section 6 concludes.

# 2 Estimation with transformed data

The purpose of this section is to show that estimates of the structural parameters and inference about the effect of certain shocks depend on the preliminary transformation employed to match a model to the data. Given the wide range of outcomes, we also argue that it is difficult to select a set of estimates for policy and interpretation purposes. We consider a textbook small scale New-Keynesian model, where agents face a labour-leisure choice, production is stochastic and requires labour, there is external habit in consumption, an exogenous probability of price adjustments, and monetary policy is conducted with a conventional Taylor rule. Details on the structure are in the on-line appendix.

The model features a technology disturbance $z_t$, a preference disturbance $\chi_t$, a monetary policy disturbance $\epsilon_t$, and a markup disturbance $\mu_t$. The latter two shocks are assumed to be iid. Depending on the specification $z_t, \chi_t$ are either both transitory, with persistence $\rho_z$ and $\rho_\chi$ respectively, or one of them is permanent. The structural parameters to be estimated are: $\sigma_c$, the risk aversion coefficient, $\sigma_n$, the inverse of the Frisch elasticity, $h$ the coefficient of consumption habit, $1-\alpha$, the share of labour in production, $\rho_r$, the degree of interest rate smoothing, $\rho_\pi$ and $\rho_y$, the parameters of the monetary policy rule, 1-$\zeta_p$, the probability of changing prices. The auxiliary parameters to be estimated are: $\rho_\chi, \rho_z$, the autoregressive parameters of transitory preference and technology shocks, and $\sigma_z, \sigma_\chi, \sigma_r, \sigma_\mu$, the standard deviations of the four structural shocks. The discount factor $\beta$ and the elasticity among varieties $\theta$ are not estimated since they are very weakly identified from the data.

Depending on the properties of the technology and of the preference shocks, the optimality conditions will have a log-linear representation around the steady state or a growth path, driven either by the technology or by the preference shock, see table 1. Four observable variables are used in the estimation. When the model is assumed to be driven by transitory shocks, parameter estimates are obtained i) applying four statistical filters (linear detrending (LT), Hodrick and Prescott filtering (HP), growth rate filtering (FOD) and band pass filtering (BP)) to output, the real wage, the nominal interest rate and inflation or ii) using three data transformations. In the first, the log of labour productivity, the log of real wages, the nominal rate and the inflation rate, all demeaned, are used as observables (Ratio 1). In the second the log ratio of output to the real wage, the log of hours worked, the nominal rate and the inflation rate, all demeaned, are used as observables (Ratio 2). In the third, the log of the labour share, the log ratio of real wages to output, the nominal interest rate and the inflation rate all demeaned, are used as observables (Ratio 3). When the model features a trending TFP (TFP), the linear stochastic specification $z_t = bt + \epsilon_t^z$, is used and, consistent with the theory, the observables for the transformed model are linearly detrended output, linearly detrended wages, demeaned inflation and demeaned interest rates. When the model features trending preferences shocks (Preferences), the unit root specification, $\chi_t = \chi_{t-1} + \epsilon_t^\chi$, is employed and the observables for the transformed model are the demeaned growth rate of output, demeaned log of real wages, demeaned inflation and demeaned interest rates. Finally, when the model features a trending TFP, the likelihood function of the transformed model is approximated as in Hansen and Sargent (1993) and only the information present at business cycle frequencies $(\frac{\pi}{32}, \frac{\pi}{8})$ is used in the estimation (TFP FD).

The data used comes from the FRED quarterly database at the Federal Reserve Bank of St. Louis and Bayesian estimation is employed. Since some of the statistical filters are two-sided, a recursive LT filter and a one-sided version of the HP filter have also been considered. The qualitative features of the results are unchanged by this refinement.

Table 2 shows that the posterior distribution of several parameters depends on the preliminary transformation used (see e.g. the risk aversion coefficient $\sigma_c$; the Frisch elasticity $\sigma_n^{-1}$; the interest smoothing coefficient $\rho_r$; persistence and the volatility of the shocks). Since posterior standard deviations are generally tight, differences across columns are a-posteriori significant. Posterior differences are also economically relevant. For example, the volatility of markup shocks in the LT, the Ratio 1 and the Preference economies is considerably larger and, perhaps unsurprisingly, risk aversion stronger. Note that, even within classes of transformations, differences are present. For example, comparing the Ratio 1 and Ratio 3 economies, it is clear that using the labour share and the ratio of real wages to output as observables considerably reduces the persistence of the technology shocks - rendering the Ratio 3 transformation more appropriate as far as stationarity of the observables is concerned - at the cost of making the risk aversion and habit coefficient very low.

Differences in the location of the posterior of the parameters translate into important differences in the transmission of shocks. As shown in Figure 2, the magnitude of the impact coefficient and of the persistence of the responses to technology shocks vary with the preliminary transformation and, for the first few horizons, differences are statistically significant. Furthermore, the sign of output and interest rate responses is affected.

Why are parameter estimates so different? The first four transformations only approximately isolate business cycle frequencies, leaving measurement errors in the transformed data. In addition, different approaches spread the measurement error across different frequencies: the LT transformation leaves both long and short cycles in the filtered data; the HP transformation leaves high frequencies variability unchanged; the FOD transformation emphasizes high frequency fluctuations and reduces the importance of cycles with business cycle periodicity; and even a BP transformation induces significant small sample approximation errors (see e.g. Canova, 2007). Since the magnitude of the measurement error and its frequency location is transformation dependent, differences in parameter estimates emerge. An approach which can reduce the problematic part of the measurement error is in Canova and Ferroni (2011). More importantly, filtering approaches neglect the fact that the spectral properties of a DSGE model are different from the output of a statistical filter. Data generated by a DSGE model driven by transitory shocks have power at all frequencies of the spectrum and if shocks are persistent most of the power will be in the low frequencies. Thus, concentrating on business cycles frequencies may lead to inefficiencies. When transitory and permanent shocks are present, the transitory and the permanent components of the model will jointly appear in any frequency band and it is not difficult to build examples where permanent shocks dominate the variability at business cycle frequencies (see Aguiar and Gopinath, 2007). Hence, the association between the solution of the model and the filtered observables generally leads to biases.

Implicit or explicit model-based transformations avoid these problems by specifying a permanent and a transitory component of the data with power at all frequencies of the spectrum. However, since specification problems are present (should we use a unit root process or a trend stationary process? Should we allow trending preferences or trending technology shocks?), nuisance parameters problems could be important (the model estimated with a trending TFP has MA components which do not appear when the preferences are trending, see table 1), and tight cointegration relationships are imposed on the observables, any deviation from the assumed structure leads to biases. Finally, frequency domain estimation may require a model-based transformation (in which case the problems discussed in the previous paragraph apply) and is generally inefficient, since most of the variability the model produces is in the low frequencies. In general, while frequency domain estimation can help to tone down the importance of aspects of the model researchers do not trust, see Hansen and Sargent (1993), it cannot reduce the importance of what the model leaves unexplained at business cycle frequencies.

# 3 The alternative methodology

Start from the assumption that the observable data has been generated by rational expectation agents, optimizing their objective functions under constraints in a stochastic environment. Suppose that the log of an $N \times 1$ demeaned vector of time series $x_t^d$ can be decomposed in two mutually orthogonal parts

$$x_t^d = z_t + x_t \tag{1}$$

Assume that the econometrician is confident about the process generating $x_t - u_t = x_t^m(\theta)$, where $\theta$ is a vector of structural parameters, and $u_t$ a vector of iid measurement errors but he/she is unsure about the process generating $z_t = z_t^m(\theta, \gamma)$, where $\gamma$ is another vector of structural parameters because she does not know the shocks which are driving $z_t$; because she does not feel confident about their propagation properties; or because she does not know how to model the relationship between $\theta$ and $\gamma$. In the context of section 2, $z_t$ is the permanent component and $x_t$ the transitory component of the data, and the researcher is unsure about the modelling of $z_t$ because it could be deterministic or stochastic, it could be driven by preference or technology shocks, and balance growth could hold or not. Still, she wants to employ $x_t^m(\theta)$ for inference because $z_t$ may be tangential to the issues she is interested in. Thus, she is aware that the model is misspecified in at least two senses: there are shocks missing from the model; and there are cross equation restrictions that are ignored.

An investigator interested in estimating $\theta$ and conducting structural inference does not necessarily have to construct an estimate of $x_t^m(\theta)$, filtering out from the data what the model is unsuited to explain; add ad-hoc structural features hoping that $\tilde{z}_t^m \equiv D(\ell)(\theta,\gamma)\tilde{e}_{1t}$ is close to $x_t^d - x_t^m(\theta)$, where, as in section 2, $\tilde{e}_{1t}$ is a set of (permanent) shocks and $D(\ell)(\theta,\gamma)$ a model propagating $\tilde{e}_{1t}$, or transform the observables so that $z_t$ becomes a vector of iid random variables, as is commonly done. Instead, she can use the raw data $x_t^d$, the model $x_t^m(\theta)$, and build a non-structural link between the (misspecified) structural model and the raw data which is sufficiently flexible to capture what the model is unsuited to explain, and allows model-based and non model-based components to jointly appear at all frequencies of the spectrum.

As a referee has pointed out, the assumption of orthogonality of $z_t$ and $x_t$ is crucial for the procedure outlined below to be effective. When permanent drifts in the data occur because of drifting structure or drifting cyclical parameters rather than permanent shocks, alternative approaches need to be considered.

Let the (log)-linearized stationary solution of a DSGE model be of the form:

$$x_{2t} = A(\theta)x_{1t-1} + B(\theta)\epsilon_t \tag{2}$$

$$x_{1t} = C(\theta)x_{1t-1} + D(\theta)\epsilon_t \tag{3}$$

where $A(\theta), B(\theta), C(\theta), D(\theta)$ depend on the structural parameters $\theta$, $x_{1t} \equiv (\log \tilde{x}_{1t} - \log \bar{x}_{1t})$ includes exogenous and endogenous states, $x_{2t} = (\log \tilde{x}_{2t} - \log \bar{x}_{2t})$ all other endogenous variables, $\epsilon_t$ the shocks and $\bar{x}_{2t}, \bar{x}_{1t}$ are the long run paths of $\tilde{x}_{2t}$ and $\tilde{x}_{1t}$.

Let $x_t^m(\theta) = R[x_{1t}, x_{2t}]'$ be an $N \times 1$ vector, where $R$ is a selection matrix picking out of $x_{1t}$ and $x_{2t}$ variables which are observable and/or interesting from the point of view of the analysis and let $\bar{x}_t^m(\theta) = R[\bar{x}_{1t}, \bar{x}_{2t}]'$. Let $x_t^d = \log \tilde{x}_t^d - E(\log \tilde{x}_t^d)$ be the log demeaned $N \times 1$ vector of observable data. The specification for the raw data is:

$$x_t^d = c_t(\theta) + x_t^{nm} + x_t^m(\theta) + u_t \tag{4}$$

where $c_t(\theta) = \log \bar{x}_t^m(\theta) - E(\log \tilde{x}_t^d)$, $u_t$ is a iid $(0, \Sigma_u)$ (proxy) noise, $x_t^{nc}, x_t^m$ and $u_t$ are mutually orthogonal and $x_t^{nm}$ is given by:

$$\begin{aligned} x_t^{nm} &= \rho_1 x_{t-1}^{nm} + w_{t-1} + v_{1t} \qquad v_{1t} \sim \ iid\ (0, \Sigma_1) \\ w_t &= \rho_2 w_{t-1} + v_{2t} \qquad\qquad v_{2t} \sim \ iid\ (0, \Sigma_2) \end{aligned} \tag{5}$$

where $\rho_1 = diag(\rho_{11}, ...\rho_{1N}), \rho_2 = diag(\rho_{21}, ...\rho_{2N})$, $0 < \rho_{1i}, \rho_{2i} \leq 1, i = 1, ...N$ . To understand what (5) implies, notice that when $\rho_1 = \rho_2 = I$, and $v_{1t}, v_{2t}$ are uncorrelated $x_t^m(\theta)$ is the locally linear trend specification used in state space models, see e.g. Gomez (1999). On the other hand, if $\rho_1 = \rho_2 = I, \Sigma_1$ and $\Sigma_2$ are diagonal, $\Sigma_{1_i} = 0$, and $\Sigma_{2_i} > 0$, $\forall i$, $x_t^{nm}$ is a vector of I(1) processes while if $\Sigma_{1_i} = \Sigma_{2_i} = 0$, $\forall i$, $x_t^{mn}$ is deterministic. When instead $\rho_1 = \rho_2 = I$, and $\Sigma_{1_i}$ and $\Sigma_{2_i}$ are functions of $\Sigma_\epsilon$, (5) approximates the double exponential smoothing setup used in discounted least square estimation of state space models, see e.g. Delle Monache and Harvey (2010). Thus, if $\bar{x}_t^m(\theta) = \bar{x}^m(\theta), \forall t$, the observable $x_t^d$ can display any of the typical structures that motivate the use of the statistical filters. Furthermore, as emphasized by Delle Monache and Harvey (2010), (5) can capture several other types of structural model misspecification. For example, whenever $\Sigma_2$ is different from zero, the growth rate of the endogenous variables may display persistent deviations from their mean, a feature that characterizes many real macroeconomic variables, see e.g. Ireland (2012), even if the model is driven by transitory shocks. Finally, when $\bar{x}_t^m(\theta)$ is not constant, and $\rho_{1i}$ and $\rho_{2i}$ are complex conjugates for some i, the specification can account for residual low frequency variations with power at frequency $\omega$. To see this note that when N=1, (5) implies that $(1 - \rho_2 L)(1 - \rho_1 L)x_t^{nm} = (1 - \rho_2 L)v_{1t} + v_{2t-1} \equiv (1 - \psi L)\eta_t$. If the roots $\lambda_1^{-1}, \lambda_2^{-1}$of the polynomial $1 - (\rho_1 + \rho_2)z + \rho_1\rho_2 z^2 = 0$ are complex, they can be written as

$\lambda_1^{-1} = r(\cos\omega + i\sin\omega), \lambda_2^{-1} = r(\cos\omega - i\sin\omega)$, where $r = \sqrt{\rho_1\rho_2}$ and $\omega = \cos^{-1}[\frac{\rho_1+\rho_2}{2\sqrt{\rho_1\rho_2}}]$ and (5) is $x_t^{nm} = \sum_j r\frac{sin\omega(j+1)}{sin\omega}(1-\psi L)\eta_t$, whose period of oscillation is $p = \frac{2\pi}{\omega} = \frac{2\pi}{\cos^{-1}[\frac{\rho_1+\rho_2}{2\sqrt{\rho_1\rho_2}}]}$. Thus, given $r$ and $p$, there exists $\rho_1, \rho_2$ that produce $x_t^{nm}$ with the required properties.

Given (2)-(5), identification of the structural parameters is achieved via the cross-equation restrictions that the model imposes on the data. Estimates of the non-structural parameters are implicitly obtained from the portion of the data the model cannot explain.

## 3.1 Two special cases

Two special cases of the setup are of interest. Suppose that the model features only transitory shocks while the data may display common or idiosyncratic long run drifts, low frequency movements, and business cycle fluctuations. Here $\bar{x}_t^m(\theta) = \bar{x}^m(\theta), \forall t$, are the steady states of the model and, if the model is correctly specified on average, $c_t(\theta) = 0$. Assume that no proxy errors are present. Then (4) is

$$x_t^d = x_t^{nm} + x_t^m(\theta) \tag{6}$$

and $x_t^{nm}$ captures the features of $x_t^d$ that the stationary model does not explain. Depending on the specification of $\rho_1$ and $\rho_2$, these may include long run drifts, both of common and idiosyncratic type, and those idiosyncratic low and business cycle movements the model leaves unexplained. In this setup, $x_t^{nm}$ has two interpretations. As in Altug (1989), McGrattan (1994) and Ireland (2004b), it can be thought of as measurement error added to the structural model. However, rather than being iid or VAR(1), it has the richer representation (5) and it is present even when the number of structural shocks equals the number of endogenous variables. Alternatively, $x_t^{nm}$ can be thought of as a reduced form representation for the components of the data the investigator decides not to model. Thus, as in Del Negro et al. (2006), $x_t^{nm}$ relaxes certain cross equations restrictions that the DGP imposes on $x_t^d$.

Suppose, alternatively, that the model features transitory shocks and one or more permanent shocks. In this case $x_t^m(\theta)$ represents the (stationary) solution in deviation from a growth path and $\bar{x}_t^m(\theta)$ is the model-based component generated by the permanent shocks. Suppose again that there are no proxy errors. Then (4) is

$$x_t^d = c_t(\theta) + x_t^{*,nm} + x_t^m(\theta) \tag{7}$$

where $x_t^{*,nm}$ captures the features of $x_t^d$ which neither the transitory portion $x_t^m(\theta)$ nor the permanent portion $c_t(\theta)$ of the model explains. These may include idiosyncratic long run patterns (such as diverging trends), idiosyncratic low frequency movements, or unaccounted cyclical fluctuations. Comparing (6) and (7), one can see that $x_t^{nm} = c_t(\theta) + x_t^{*,nm}$. Thus, the setup can be used to measure how much of the data the model leaves unexplained and to evaluate whether the introduction of certain structural shocks reduces the discrepancy. To illustrate, suppose as in the application discussed in section 5.1, one starts from a model featuring a few transitory shocks and finds that the relative importance of $x_t^{nm}$ - measured, for example, by the variance decomposition at a particular set of frequencies - is large. Then, one could add a transitory shock or a permanent shock to the model and see how much the relative importance of $x_t^{nm}$ has fallen. By comparing the relative size of $x_t^{nm}$ in the various cases, one can then assess whether adding a permanent or a transitory shock is more beneficial for understanding the dynamics of $x_t^d$.

The same logic can be used to evaluate the model when, for example, the permanent shock takes the form of a stochastic linear trend, or of a unit root, or when all long run paths are left unmodelled. Hence, the approach provides a setup to judge the goodness of fit of a model; a constructive criteria to increase its complexity; and a framework to examine the sensitivity of the estimation results to the specification of nuisance features.

The specification has other advantages over existing approaches. As shown in Ferroni (2011), the setup can be used to find the most appropriate specification of the non model-based component, and to perform Bayesian averaging over different types of non model-based specifications, both of which are not possible in standard setups. Finally, since joint estimation is performed, structural parameter estimates reflect the uncertainty present in the specification of the non model-based component.

## 3.2 Estimation

Estimation of the parameters of the model can be carried out with both classical and Bayesian methods. (2)-(5) can be cast into the linear state space system:

$$\begin{aligned} s_{t+1} &= Fs_t + G\omega_{t+1} \qquad \omega_t \sim (0, \Sigma_\omega) && (8) \\ x_t^d &= c_t(\theta) + Hs_t && (9) \end{aligned}$$

where $s_t = \begin{pmatrix} x_t^{nm} & w_t & x_t^m(\theta) & u_t \end{pmatrix}'$, $\omega_{t+1} = (v_{1t+1}, v_{2t+1}, u_{t+1}, \epsilon_{t+1})'$, $H = \begin{pmatrix} I & 0 & I & I \end{pmatrix}$, $F = \begin{pmatrix} \rho_1 & I & 0 & 0 \\ 0 & \rho_2 & 0 & 0 \\ 0 & 0 & R[A\ C]' & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, $G = \begin{pmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & 0 & R[B\ D]' \\ 0 & 0 & I & 0 \end{pmatrix}$. Hence, the likelihood can be computed with a modified Kalman filter (accounting for the possibility of diffuse initial observations) given $\vartheta = (\theta, \rho_1, \rho_{2,}\Sigma_1, \Sigma_2, \Sigma_u)$ and maximized using standard tools.

When a Bayesian approach is preferred, one can obtain the non-normalized posterior of $\vartheta$, using standard MCMC tools. For example, the estimates I present are obtained with a Metropolis algorithm where, given a $\vartheta_{-1}$ and a prior $g(\vartheta)$, candidate draws are obtained from $\vartheta_* = \vartheta_{-1} + \upsilon$, where $\upsilon \sim t(0, \kappa * \Omega, 5)$ and $\kappa$ is a tuning parameter, and accepted if $\frac{\breve{g}(\vartheta_*|y)}{\breve{g}(\vartheta_{-1}|y)}$ exceeds the draw of a uniform random variable, where $\breve{g}(\vartheta_i|y) = g(\vartheta_i)\mathcal{L}(y|\vartheta_i)$, $i = *, -1$, and $\mathcal{L}(y|\vartheta_i)$ is the likelihood of $\vartheta_i$, . Iterated a large number of times, for $\kappa$ appropriately chosen, limiting distribution of $\vartheta$ is the target distribution (see e.g. Canova, 2007).

## 3.3 The relationship with the existing literature

Apart from the work of Altug (1989), McGrattan (1994), Ireland (2004b), and Del Negro et al. (2006) already mentioned, the procedure is related to a number of existing works.

First, the state space setup (8)-(9) is similar to the one of Harvey and Jeager (1993), even though these authors consider only univariate processes and do not use a structural model to explain the observables. It also shares important similarities with the one employed by

Cayen et al. (2009), who are interested in forecasting trends. Two are the most noticeable differences. First, these authors use a two-step estimation approach, conditioning on filtered estimates of the parameters of the DSGE model; here a one-step approach is employed. Second, all the deviations from the model are bundled up in the non-model specification while here it is possible to split them into interpretable and non-interpretable parts.

The contribution of the paper is also related to two distinct branches of the macroeconomic and macroeconometric literature. The first attempts to robustify inference when the trend properties of data are misspecified (see Cogley, 2001, and Gorodnichenko and Ng, 2010). I share with the first author the idea that economic theory may not have much to say about certain types of fluctuations but rather than distinguishing between trend-stationary and difference-stationary cycles, I design an estimation procedure which deals flexibly with the mismatch between theoretical and empirical concepts of fluctuations. The idea of jointly estimating structural and non-structural parameters without fully specifying the DGP is also present in Gorodnichenko and Ng. However, a likelihood based estimator, as opposed to a minimum distance estimator, is used here because it works regardless of the properties of the raw data. In addition, rather than assuming that the model is the DGP, the procedure assumes that the model is misspecified - a much more useful assumption in practice.

The second branch points out that variations in trend growth are as important as cyclical fluctuations in explaining the dynamics of macroeconomic variables (see Aguiar and Gopinath, 2007, and Andrle, 2008). While the first paper characterizes differences between emerging and developing economies, the latter is concerned with the misuse of models driven by transitory shocks in policy analyses for developing countries. The paper shows that the problems they highlight are generic and that policy analyses with misspecified models are possible without controversial assumptions on what the model is not designed to explain.

### 3.4 Setting the priors for $\Sigma_1$ and $\Sigma_2$

If the number of observables is small and the sample size large, one can estimate structural and non-structural parameters jointly from (8)-(9) in an unrestricted fashion. More realistically, when the number of observables and the sample size are moderate, unrestricted estimation is unfeasible - the model features $2N + 2N^2$ non-structural parameters - and weak identification problems may be important - variations in the level from variations in the growth rates of the observables are hard to distinguish. Thus, it is worth imposing some structure to reduce the number of estimated non-structural parameters. For example, one may assume that $\Sigma_1$ and $\Sigma_2$ are diagonal (the non model-based component is series specific) and of reduced rank (the non model-based component is common across [groups of] series). Alternatively, one may assume that these matrices have sparse non-zero elements on the diagonal (the non model-based component exists only in some observables) or that they are proportional to each other (shocks to the level and the growth rate are related). One may also want to make $\rho_{1i}$ and $\rho_{2i}$ common across certain variables. Restrictions of this type may be supported by plots or time series analysis of the observables.

Additional restrictions may be needed to make estimation meaningful in small samples because, given a DSGE structure, the decomposition in model based and non model-based components is indexed by the relative intensity of the shocks driving the two components. Given that it is difficult to estimate this intensity parameter unrestrictedly in small samples, a sensible smoothness prior for $\Sigma_1$ and $\Sigma_2$ may avoid that estimates of non model-based components feature undesirable high frequency variability. Harvey and Jeager (1993) have indicated that in univariate state space models, estimation of the cycle depends on the assumptions about the trend - in particular whether it is deterministic or stochastic. The problem we highlight here is different: given assumptions about the trend, different decompositions of the observables in model and non-model based components are implied by different estimates of the relative variance of the permanent shocks. Thus, for example, if we assume that the trend is driven by permanent shocks, different decompositions may be obtained if the relative magnitude of the shocks driving the two components is weak or strongly identified.

The specification I found to work best in practice, and it is employed in the two applications in section 5, involves making $\Sigma_1$ and $\Sigma_2$ a function of the structural shocks. As mentioned, it is possible to approximate the double exponential smoothing restrictions used in discounted least square estimation of state space models by selecting $\Sigma_{1_i} = \sqrt{\frac{\sigma_\epsilon^2}{\lambda}}$ and $\Sigma_{2_i} = \sqrt{\frac{\sigma_\epsilon^2}{(4\lambda)^2}}$, where $i$ indicates the non-zero elements of the matrices, $\epsilon_t$ is one of structural shocks and $\lambda$ a smoothing parameter. Thus, given a prior for $\epsilon_t$ and $\lambda$, a prior for all non-zero elements of $\Sigma_1$ and $\Sigma_2$ is automatically generated. Since $\lambda$ has the same interpretation as in the HP filter, an agnostic quarterly prior for $\lambda$ could be uniform over [4,6400], which allows for very smooth as well as relatively jagged non-model based components. It is worth noting that this specification is parsimonious and that selecting the signal to noise ratio $\lambda$ is less controversial than assuming a particular format for the drifts the data displays or selecting a shock driving them. Since a structural shock needs to be selected, one could experiment and choose the disturbance with the largest or the smallest variance. For the applications in section 5, the way the prior is scaled is irrelevant.

An alternative approach, suggested by one of the referees, would be to exploit the flexibility of (5) to perform sensitivity analysis to alternative specifications of $\rho_1, \rho_2, \Sigma_1, \Sigma_2$. Also in this case, restrictions to reduce the dimensionality of non-structural parameter space are generally needed to make estimation results sensible.

# 4 The procedure in a controlled experiment

To examine the properties of the procedure and to compare them to those of standard transformations, I use the same setup employed in section 2 and simulate 150 data points 50 times, assuming first that the preference shock has a transitory and a permanent component. Thus, $\chi_t = \chi_{1t} + \chi_{2t}, \chi_{1t} = \rho_\chi \chi_{1t-1} + \epsilon_t^{\chi T}$ and $\chi_{2t} = \chi_{2t-1} + \epsilon_t^{\chi P}$ where $\sigma_\chi^p / \sigma_\chi^T$ is uniformly distributed [1.1, 1.9]. Because $\chi_{2t}$ is orthogonal to all transitory shocks, the design fits the setup of section 3. The specification is chosen since Chang et al. (2007) have indicated that a model with permanent preference shocks can capture well low frequency variations in hours worked. In this setup, the data displays stationary fluctuations, driven by four transitory shocks (which are correctly captured with a model), and non-stationary fluctuations, driven by the permanent preference shock (which will be either filtered out, eliminated with certain data transformations, or accounted for with a non-model based component). The estimated model is misspecified since the permanent preference shock is left out, but all the other features are correctly represented. Since the contribution of the permanent component is of the same order of magnitude as the contribution of the transitory component at almost all frequencies, standard transformations will feature both filtering and specification errors. When the proposed approach is used, the non model-based component is restricted to having a double exponential smoothing format and, consistently with the DGP, is allowed to enter only in output and the real wage (see on-line appendix).

The second design features only transitory shocks, but measurement error is added to the data. The variability of the measurement error relative to the variability of the preference shock is uniformly distributed in the range [0.08, 0.12]. Here the model captures the dynamics of the data correctly, but (a constant) noise is present at all frequencies. The question of interest is whether the suggested specification will be able to recognize that there is no non model-based component or whether the non model-based component will absorb part of the model dynamics. Note that, since the signal to noise ratio differs in the two designs, we can also evaluate how our smoothness prior works in different situations.

The structural parameters will be estimated in the most ideal situations one could consider - these include priors centered at the true parameter vector (the same prior distributions displayed in table 2 are used) and initial conditions equal to the true parameter vector which is listed in the first column of table 3. The other columns report, for each of the six estimation procedures we consider, the mean square error (MSE) of each parameter separately, and two cumulative MSE measures, one for the structural parameters and one for all the parameters. The MSE is calculated using the posterior mean estimate in each replication.

In the first design, estimation with HP and BP filtered data produce MSEs that are larger than with LT or FOD data, in particular, for the inverse of the Frisch elasticity and the share of labour in production. Moreover, all filtering procedures have a hard time to pin down the value of the Taylor rule coefficient on output. Perhaps unsurprisingly, all transformations fail to capture both the absolute and the relative variability of the shocks. The ratio transformation is also poor and the cumulative MSEs are the largest of all. In comparison, the flexible approach does well in estimating structural parameters (the only exception is the consumption habit parameter) and captures the volatility and persistence of structural shocks much better than competitors.

The pattern of the results with the second design is similar, even though several transformations induce larger distortions in the estimates of the Frisch elasticity. The performance of the flexible approach is also good in this case. In particular, it does much better than other approaches in capturing the volatility and the persistence of the structural shocks.

The implications of these results for standard dynamic analyses are clear. For example, variance decomposition exercises are likely to be distorted if parameter estimates are obtained with standard procedures. This is much less the case when the flexible approach is employed. Furthermore, structural inference regarding, e.g. the sluggishness of the policy rate or its sensitivity to output gap fluctuations, is less likely to be biased when the approach suggested in the paper is used.

To highlight further the properties of the proposed approach, figure 3 compares the autocorrelation function and the spectral density of the true and estimated permanent and transitory components of output for first design, where the latter is obtained using the median estimates in one replication. The approach performs well: the rate of decay of the autocorrelation functions of the true and the estimated components is similar. As anticipated, the two estimated components have power at all frequencies, but at business cycle frequencies (indicated by the vertical bars in the last row of graphs) the permanent component is more important than the transitory component.

The conditional dynamics in response to transitory shocks with true and estimated parameters are in Figure 4. In general, the sign and the persistence of the responses are well matched. Magnitudes and shapes are occasionally imprecisely estimated (see e.g. the responses to technology shocks) but, overall, the approach does a reasonable job in reproducing the main qualitative features of the DGP.

To understand the nature of the distortions produced by standard transformations, note that the log-likelihood of the data can be represented as $L(\theta|y_t) = [A_1(\theta) + A_2(\theta) + A_3(\theta)|y]$, see Hansen and Sargent (1993), where $A_1(\theta) = \frac{1}{\pi}\sum_{\omega_j} \log\det G_\theta(\omega_j)$, $A_2(\theta) = \frac{1}{\pi}\sum_{\omega_j}$ trace $[G_\theta(\omega_j)^{-1}F(\omega_j)]$, $A_3(\theta) = (E(y) - \mu(\theta))G_\theta(\omega_0)^{-1}(E(y) - \mu(\theta))$, $\omega_j = \frac{\pi j}{T}, j = 0, 1, \ldots, T-1$. $G_\theta(\omega_j)$ is the model-based spectral density matrix of $y_t$, $\mu(\theta)$ the model-based mean of $y_t$, $F(\omega_j)$ is the data-based spectral density and $E(y)$ the unconditional mean of $y_t$. $A_2(\theta)$ and $A_3(\theta)$ are penalty functions: $A_2(\theta)$ sums deviations of the model-based from the data-based spectral density over frequencies; $A_3(\theta)$ weights deviations of model-based from data-based means with the spectral density matrix of the model at frequency zero.

Suppose the data is transformed so that the zero frequency is eliminated and the low frequencies de-emphasized. Then, the log-likelihood consists of $A_1(\theta)$ and of $A_2(\theta)^* = \frac{1}{\pi}\sum_{\omega_j}$ trace $[G_\theta(\omega_j)]^{-1}F(\omega_j)^*$, where $F(\omega_j)^* = F(\omega_j)I_{\omega_j}$, and $I_{\omega_j}$ is a function describing the effect of the filter at frequency $\omega_j$. Suppose that $I_\omega = I_{[\omega_1,\omega_2]}$, i.e. an indicator function for the business cycle frequencies, as in an ideal BP filter. Then $A_2(\theta)^*$ matters only at business cycle frequencies. Since at these frequencies $[G_\theta(\omega_j)] < F(\omega_j)^*$, $A_2(\theta)^*$ and $A_1(\theta)$ enter additively $L(\theta|y_t)$, two types of biases will be present. Since estimates $\hat{F}(\omega_j)^*$ only approximately capture the features of $F(\omega_j)^*$, $\hat{A}_2(\theta)^*$ has smaller values at business cycle frequencies and a nonzero value at non-business cycle ones. Moreover, in order to reduce the contribution of the penalty function to the log-likelihood, parameters are adjusted so that $[G_\theta(\omega_j)]$ is close to $\hat{F}(\omega_j)^*$ at those frequencies where $\hat{F}(\omega_j)^*$ is not zero. This is done by allowing fitting errors, (a larger $A_1(\theta)$), at frequencies where $\hat{F}(\omega_j)^*$ is zero - in particular, the low frequencies. Hence, the volatility of the structural shocks will be overestimated (this makes $G_\theta(\omega_j)$ close to $\hat{F}(\omega_j)^*$ at the relevant frequencies), in exchange for misspecifying their persistence. These distortions affect agents' decision rules: higher perceived volatility, for example, implies distortions in the Frisch elasticity. Inappropriate persistence estimates, on the other hand, imply that perceived substitution and income effects are distorted with the latter typically underestimated. When $I_\omega$ is not the indicator function, the derivation of the size and the direction of the distortions is more complicated but the same logic applies. Clearly, different $I_\omega$ produce different $\hat{F}(\omega_j)$ and thus different distortions.

Since estimates of $F(\omega_j)^*$ are imprecise, even for large $T$, there are only two situations when estimation biases are small. First, the permanent component has low power at business cycle frequencies - in this case, the distortions induced by the penalty function are limited. This occurs when transitory volatility dominates. Second, when Bayesian estimation is performed, the prior is selected to limit the distortions induced by the penalty function. This is very unlikely, however, since priors are not elicited with such a scope in mind.

If instead one fits a transformed version of the model to transformed data, as is done in model- based approaches, the log-likelihood is composed of $A_1(\theta)^* = \frac{1}{\pi}\sum_{\omega_j} \log|G_\theta(\omega_j) I_{\omega_j}|$ and $A_2(\theta)$ - since the actual and model data are filtered in the same way, the filter does not affect the penalty function. Suppose that $I_\omega = I_{[\omega_1,\omega_2]}$. Then $A_1(\theta)^*$ matters only at business cycle frequencies while the penalty function is present at all frequencies. Therefore, parameter estimates are adjusted so as to reduce the misspecification at all frequencies. Since the penalty function is generally more important at low frequencies, parameters are selected to make $[G_\theta(\omega_j)]$ close to $\hat{F}(\omega_j)$ at those frequencies and large fitting errors are permitted at medium and high frequencies. Consequently, the volatility of the shocks will be generally underestimated in exchange for overestimating their persistence - somewhat paradoxically, this procedure implies that the low frequency components of the data are those that matter most for estimation. Cross frequency distortions imply incorrect inference. For example since less noise is perceived, agents' decision rules imply a higher degree of data predictability, and higher perceived persistence implies that perceived substitution and income effects are distorted with the latter overestimated.

# 5 Two applications

This section shows how the proposed approach can be used to inform researchers about two questions which have received a lot of attention in the literature: the time variations in the policy activism parameter and the sources of output and inflation fluctuations. The first question is analyzed with the model presented in section 2. The second with a medium scale model, widely used in academic and policy circles.

## 5.1 The policy activism parameter

What are the features of the monetary policy rule in place during the "Great Inflation" of the 1970s and the return to norm of the 1980s and 1990s? This question has been extensively studied in the literature, following Clarida et al. (2000). One synthetic way to summarize the information contained in the data is to compute the policy activism parameter $\frac{\rho_y}{\rho_\pi - 1}$, which gives a sense of the relative importance of the output and the inflation stabilization objectives of the Central Bank. The conventional wisdom suggests that the absolute value of this parameter has declined over time, reflecting changes in the preferences of the monetary authorities, but most of the available evidence is obtained either with reduced form methods or, when structural methods are used, with filtered data. Are the results to be trusted? Is the characterization offered by the approach of this paper different? Figure 5 plots the posterior density of the policy activism parameter when the data is linearly detrended (top left box) or HP filtered (top right box) and when the approach of this paper is employed (lower left box) for the samples 1964:1-1979:4 and 1984:1-2007:4. The priors for the structural and auxiliary parameters are the same as in table 1. In the flexible approach, $\Sigma_e$ and $\Sigma_v$ are assumed to be diagonal, a common non model-based component is assumed for all the variables, the signal-to-noise ratio in the four series is captured by a single parameter $\lambda$, a-priori uniformly distributed over [100, 6400], $\rho_1 = \rho_2 = I$ and $u_t = 0, \forall t$.

The posterior density of the policy activism parameter shifts to the left in the second sample when HP filtered data is used and, for example, the posterior median moves from -0.23 in the first sample to -0.33 in the second. This left shift of the posterior density is absent when LT data is used and the median of the posterior in the second sample moves closer to zero (from -0.38 to 0.12) - care should be exercised here since the median is not a good estimator of the central tendency of the posterior for the 1984-2007 sample. In both cases, the Kolmogorov-Smirnov statistic rejects the null that the posterior distributions are the same in the two samples. Thus, standard approaches confirm the existence of a break in the conduct of monetary policy, although it is not clear in which direction the movement is: with HP filtered data, output gap considerations have become relatively more important; with LT filtered data, the opposite appears to be true.

When the approach of section 3 is used, the posterior density of $\frac{\rho_y}{\rho_\pi - 1}$ in the two samples overlaps considerably: both the location and the shape of the density in the two samples are very similar and the Kolmogorov-Smirnov statistic does not reject the null that the posterior distributions in the two samples are the same. Thus, evidence in favor of a structural break in the conduct of monetary policy is much weaker in this case.

Which of the three pictures should be trusted most? The Monte Carlo exercise of section 4 indicates that LT filtering may produce estimates of the two parameters entering the policy activism tradeoff with large MSEs for both DGPs. The picture is slightly better with the HP filter; still, the estimation of the output coefficient is poor. On the other hand, the MSE obtained by the flexible approach is small for both parameters and both DGPs. Thus, prima facie, the evidence provided by the flexible approach should be trusted more.

As mentioned, the non model-based component soaks up the features that the model is not designed to explain. Thus, in principle, it could absorb changes present in the endogenous variables. As a reality check, we examine whether estimates of the non-structural parameter suggest that this is true. It turns out that this is not the case: the median estimate of $\lambda$ is around 3200 in both samples, making the non model-based component quite smooth relative to the model based component (see the on-line appendix for plots of the two components of the four variables) and essentially time invariant. What happens instead is that structural non-policy parameters change to accommodate for the changes in the time series properties of inflation and the interest rate. Interestingly, the explanatory power of the model increases in the second sub-sample: on average, at business cycle frequencies, the model explains 40 per cent of output variations in the first sample and 55 per cent in the second sample. For inflation and interest rates, the increase is smaller (from 40 to 50 percent).

Since about 50 percent of the variability observables at business cycle frequencies is not captured by the model in both samples, it is worth investigating how the fit can be improved by altering its structure, keeping the number of observables and the estimation approach unchanged. To improve the fit of this kind of models the literature is now allowing a time varying inflation target in the policy rule, see e.g. Ireland (2007). The target is assumed to be driven by a permanent shock and enters only in the interest rate equation. Thus, the estimated specification moves from (6) to (7), where $c_t(\theta)$ now appears only in the interest rate equation. What would this modification do to the posterior distribution of the policy activism parameter?

The last box of figure 5 indicates that adding a time varying inflation target reduces the spread of the posterior distributions. Hence, the shift to the right in the posterior in the second sub-sample becomes statistically significant. Adding an inflation target improves the fit for the interest rate at business cycle frequencies (the proportion of the variance explained increase to 57 percent in the first sample and to 68 percent in the second); for inflation, instead, the explanatory power of the model is unchanged in the first sub-sample and worsens considerably in the second (the variance share explained at business cycle frequencies is now only 28 percent). Hence, adding a time varying inflation target does not seem to be a very promising way to improve our understanding of how inflation fluctuations are generated.

## 5.2 Sources of output and inflation fluctuations

The question of what drives output and inflation fluctuations has a long history in macroeconomics. In standard medium scale DSGE models, like the one employed by Smets and Wouters (2003) and (2007), output and inflation fluctuations tend to be primarily explained by markup shocks. Since these shocks are an unlikely source of cyclical fluctuations, Chari at al (2009) have argued that misspecification is likely to be present (see Justiniano et al., 2010, for an alternative interpretation). Researchers working in the area use filtering devices to fit the model to the data (as in Smets and Wouters (2003)), arbitrary data transformations (as in Smets and Wouters, 2007) or build a permanent component in the model (as in Justiniano et al., 2010) and use model-consistent data transformations to estimate the structural parameters. What would the approach of this paper tell us about sources of cyclical fluctuations in output and inflation? To answer this question, the same model and the same data set used in Smets and Wouters (2007) are employed but a more standard setup is used. In particular, no MA terms for the price and wage markup disturbances are assumed - all shocks have a standard AR(1) structure; the model is solved in deviations from the steady state, rather than in deviation from the flexible price equilibrium; and the policy rule does not include a term concerning output growth.

Table 4 reports results obtained eliminating a linear trend from the variables; taking growth rates of the real variables and demeaning nominal ones; and using the approach suggested in this paper. When a linear trend is removed, the forecast error variance decomposition of output at the five years horizon is indeed primarily driven by price markup shocks, with a considerably smaller contribution of investment-specific and preference shocks. For inflation, price markup shocks account for almost 90 percent of the forecast error variability at the five years horizon. When the model is instead fitted to growth rates, price markup shocks account for over 90 percent of the variability of both output and inflation at the five year horizon. Thus, even without some of the standard bells and whistles, the conclusion that markup shocks dominate remains. Why are price markup shocks important? Since, compared to other shocks, they are relatively unrestricted, they tend to absorb any misspecification the model has and any measurement error that the filters leave in the transformed data. Furthermore, since the combined specification and measurement errors are unlikely to be iid, the role of markup shocks is overestimated. When the bridge suggested in this paper is used, the non-model based component of real variables is restricted to having a common structure (there are only two parameters simultaneously controlling the non model-based component of output, consumption, investment), $\rho_1 = \rho_2 = I$ , and a proxy error is allowed in each equation, the picture is quite different. Output fluctuations at the five year horizon are driven almost entirely by preference disturbances, while inflation fluctuations are jointly accounted for by wage markup, TFP and price markup disturbances. Note that since the model explains only 20 percent of output and inflation fluctuations at business cycle frequencies, it seems premature to use it to evaluate policy alternatives.

It is useful to characterize the properties of the non model-based component to evaluate the theoretical modifications that are needed to capture what the current model leaves out. The non-model component is well represented by the specification employed and restrictions on the representation used assuming, for example, no or only one unit root are all rejected in formal testing (log Bayes factor exceeding 10 in both cases). Thus, if shocks are to be added to the model, it is important that they have permanent features and display persistent deviations from a balanced growth path. Ireland (2012) has suggested one such specification. Others, which allow both TFP and investment shocks to have these features, are also possible.

# 6 Conclusions

Estimating DSGE models with data that is statistically filtered or model-based transformed may lead researchers astray because the association between the output of the filter and the stationary solution of the model is generally incorrect and because model-based transformations impose tight restrictions which may be violated in the data. The consequences of these errors could be economically important because income and substitution effects could be distorted, the volatilities and persistence of the shocks over or underestimated and the decision rules of the agents, as perceived by the econometrician, altered.

The alternative methodology this paper proposes builds a flexible bridge between the model and the raw data. The procedure is applicable to a large class of models and i) takes into account the uncertainty in the specification of the non-model component when deriving estimates of the structural parameters; ii) provides a natural environment to judge the goodness of fit of a model; iii) gives researchers a framework to examine the sensitivity of the estimation results to the specification of nuisance features, and iv) it is easy to implement.

Unaccounted low frequency movements, such as those appearing in hours or labour productivity, or idiosyncratic trends, such as those present in relative prices, are hard to handle within standard DSGE models. Hence, certain shocks which are left somewhat unrestricted end up capturing these features. The approach this paper suggests is likely to be useful in these difficult situations because it helps researchers to distinguish what the model can explain and what it cannot.

Extensions of the setup used in the paper are easy to conceive. For example, structural breaks in the time series features of the observables could be handled either within the model-based (as in Eklund et al., 2008) or the non model-based components and the implications for structural parameters could be compared. Similarly, stochastic volatility could be captured in the model-based or non model-based components and the differences evaluated. The framework proposed in the paper requires small changes to capture these situations.

# 7 Tables

| | |
|---|---|
| **Model with transitory shocks** | |
| $w_t$ | $= (\frac{\sigma_n}{1-\alpha} + \frac{\sigma_c}{1-h})y_t - \frac{h\sigma_c}{1-h}y_{t-1} - \frac{\sigma_n}{1-\alpha}z_t - \chi_t$ |
| $y_t$ | $= E_t[\frac{1}{1+h}y_{t+1} + \frac{h}{1+h}y_{t-1} - \frac{1-h}{(1+h)\sigma_c}(\chi_{t+1} - \chi_t + r_t - \pi_{t+1})]$ |
| $\pi_t$ | $= \beta E_t \pi_{t+1} + \frac{1-\alpha}{1-\alpha+\alpha\theta}\frac{(1-\beta\zeta_p)(1-\zeta_p)}{\zeta_p}(\epsilon_t^\mu + w_t + \frac{\alpha}{1-\alpha}y_t - \frac{1}{1-\alpha}z_t)$ |
| $r_t$ | $= \rho_r r_{t-1} + (1-\rho_r)(\rho_y y_t + \rho_\pi \pi_t) + \epsilon_t^r$ |
| $n_t$ | $= \frac{1}{1-\alpha}(y_t - z_t)$ |
| **Model with stochastically trending TFP** | |
| $w_t$ | $= (\frac{\sigma_n}{1-\alpha} + \frac{1}{1-\bar{h}})y_t - \frac{\bar{h}}{1-\bar{h}}y_{t-1} - \chi_t - \frac{\bar{h}}{1-\bar{h}}(\epsilon_{t-1}^z - \epsilon_t^z)$ |
| $y_t$ | $= \frac{1}{1+\bar{h}}E_t(y_{t+1} + hy_{t-1} - (1-\bar{h})(\chi_{t+1} - \chi_t + r_t - \pi_{t+1}) + \bar{h}\epsilon_{t-1}^z + \epsilon_{t+1}^z - (1-\bar{h})\epsilon_t^z)$ |
| $\pi_t$ | $= \beta E_t \pi_{t+1} + \frac{1-\alpha}{1-\alpha+\alpha\theta}\frac{(1-\beta\zeta_p)(1-\zeta_p)}{\zeta_p}(\epsilon_t^\mu + w_t + \frac{\alpha}{1-\alpha}y_t)$ |
| $r_t$ | $= \rho_r r_{t-1} + (1-\rho_r)(\rho_y y_t + \rho_\pi \pi_t) + \epsilon_t^r$ |
| $n_t$ | $= \frac{1}{1-\alpha}y_t$ |
| **Model with unit roots in preferences** | |
| $w_t$ | $= (\sigma_n + \frac{1}{1-h})y_t - \frac{h}{1-h}y_{t-1} - \sigma_n z_t + \frac{h}{1-h}\epsilon_t^\chi)$ |
| $y_t$ | $= \frac{1}{1+h}E_t(y_{t+1} + hy_{t-1} - (1-h)(r_t - \pi_{t+1}) - (h\epsilon_t^\chi + ((1-h)\sigma_n - h)\epsilon_{t+1}^\chi))$ |
| $\pi_t$ | $= \beta E_t \pi_{t+1} + \frac{(1-\beta\zeta_p)(1-\zeta_p)}{\zeta_p}(\epsilon_t^\mu + w_t - z_t)$ |
| $r_t$ | $= \rho_r r_{t-1} + (1-\rho_r)(\rho_y y_t + \rho_\pi \pi_t) + \epsilon_t^r$ |
| $n_t$ | $= y_t - z_t$ |

Table 1: Log-linear optimality conditions, stationary model. All variables are expressed in percentage deviation from either the steady state or the balanced growth path. $\bar{h} = e^b h$ and $b$ is the slope of the stochastic trend. With trends $\sigma_c = 1$ and with unit roots in preferences also $\alpha = 0$. $z_t$ is a technology shock, $\chi_t$ a preference shock, $\epsilon_t^r$ a monetary policy shock and $\epsilon_t^\mu$ a markup shock. If $z_t$ and $\chi_t$ are transitory, $z_t = \rho_z z_{t-1} + \epsilon_t^z, \chi_t = \rho_\chi \chi_{t-1} + \epsilon_t^\chi$. When TFP is trending, $z_t = bt + \epsilon_t^z$, when preferences are trending $\chi_t = \chi_{t-1} + \epsilon_t^\chi$. In each panel the first equation defines the equilibrium real wage, the second is an Euler equation, the third a Phillips curve, the fourth a Taylor rule and the fifth a labor demand function.

| | Prior | LT | HP | FOD | BP | Ratio 1 | Ratio2 |
|---|---|---|---|---|---|---|---|
| | | Median (s.e.) | Median (s.e.) | Median (s.e.) | Median (s.e.) | Median(s.e.) | Median (s.e.) |
| $\sigma_c$ | $\Gamma(20, 0.1)$ | 1.90 (0.25) | 1.41 (0.21) | 0.04 (0.01) | 0.96 (0.11) | 2.33 (0.27) | 0.81 (0.15) |
| $\sigma_n$ | $\Gamma(20, 0.1)$ | 1.75 (0.16) | 1.37 (0.13) | 5.23 (0.08) | 1.19 (0.09) | 3.02 (0.24) | 2.68 (0.19) |
| $h$ | $B(6, 8)$ | 0.83 (0.02) | 0.88 (0.02) | 0.45 (0.01) | 0.96 (0.01) | 0.72 (0.05) | 0.88 (0.02) |
| $\alpha$ | $B(3, 8)$ | 0.07 (0.04) | 0.09 (0.05) | 0.42 (0.01) | 0.07 (0.03) | 0.05 (0.04) | 0.03 (0.01) |
| $\rho_r$ | $B(6, 6)$ | 0.19 (0.05) | 0.11 (0.04) | 0.62 (0.01) | 0.09 (0.02) | 0.38 (0.06) | 0.28 (0.04) |
| $\rho_\pi$ | $N(1.5, 0.1)$ | 1.33 (0.08) | 1.37 (0.05) | 1.53 (0.02) | 1.51(0.06) | 1.92 (0.06) | 1.80 (0.05) |
| $\rho_y$ | $N(0.4, 0.1)$ | -0.16 (0.03) | -0.18 (0.03) | 0.06 (0.00) | -0.22 (0.03) | 0.16 (0.02) | -0.03 (0.02) |
| $\zeta_p$ | $B(6, 6)$ | 0.82 (0.02) | 0.80 (0.03) | 0.63 (0.01) | 0.86 (0.01) | 0.82 (0.02) | 0.80 (0.02) |
| $\rho_\chi$ | $B(18, 8)$ | 0.69 (0.04) | 0.40 (0.05) | 0.52 (0.01) | 0.70(0.02) | 0.67 (0.03) | 0.66 (0.02) |
| $\rho_z$ | $B(18, 8)$ | 0.96 (0.02) | 0.95 (0.02) | 0.99 (0.01) | 0.97(0.01) | 0.97 (0.01) | 0.96 (0.01) |
| $\sigma_\chi$ | $\Gamma^{-1}(10, 20)$ | 0.53 (0.19) | 0.47 (0.11) | 4.96(0.13) | 0.23 (0.05) | 3.41 (0.74) | 0.97 (0.13) |
| $\sigma_z$ | $\Gamma^{-1}(10, 20)$ | 0.20 (0.04) | 0.23 (0.04) | 2.00 (0.22) | 0.19 (0.03) | 0.06 (0.01) | 0.06 (0.01) |
| $\sigma_r$ | $\Gamma^{-1}(10, 20)$ | 0.11 (0.01) | 0.08 (0.01) | 2.30(0.23) | 0.07 (0.01) | 0.10 (0.01) | 0.11 (0.18) |
| $\sigma_\mu$ | $\Gamma^{-1}(10, 20)$ | 25.06 (0.97) | 14.25 (0.93) | 7.17 (0.13) | 18.19 (0.66) | 22.89 (1.91) | 15.94 (0.49) |

| | Prior | Ratio 3 | TFP | Preferences | TFP FD | | |
|---|---|---|---|---|---|---|---|
| | | Median (s.e.) | Median (s.e.) | Median (s.e.) | Median (s.e.) | | |
| $\sigma_c$ | $\Gamma(20, 0.1)$ | 0.12 (0.03) | 1.0 | 1.0 | 1.0 | | |
| $\sigma_n$ | $\Gamma(20, 0.1)$ | 2.09 (0.14) | 2.24 (0.26) | 2.43 (0.20) | 0.50 (0.28) | | |
| $h$ | $B(6, 8)$ | 0.10 (0.03) | 0.08 (0.04) | 0.78 (0.03) | 0.54 (0.29) | | |
| $\alpha$ | $B(3, 8)$ | 0.03 (0.02) | 0.17 (0.03) | 1.0 | 0.49 (0.29) | | |
| $\rho_r$ | $B(6, 6)$ | 0.20 (0.06) | 0.30 (0.04) | 0.61 (0.02) | 0.49 (0.28) | | |
| $\rho_\pi$ | $N(1.5, 0.1)$ | 1.51 (0.07) | 1.74 (0.06) | 1.69 (0.05) | 1.69 (2.13) | | |
| $\rho_y$ | $N(0.4, 0.1)$ | 0.77 (0.04) | 0.49 (0.03) | 0.38 (0.07) | 0.25 (1.97) | | |
| $\zeta_p$ | $B(6, 6)$ | 0.81 (0.01) | 0.41 (0.03) | 0.84 (0.01) | 0.47 (0.29) | | |
| $\rho_\chi$ | $B(18, 8)$ | 0.75 (0.03) | 0.63 (0.03) | | 0.49 (0.28) | | |
| $\rho_z$ | $B(18, 8)$ | 0.62 (0.03) | | 0.59 (0.02) | | | |
| $\sigma_\chi$ | $\Gamma^{-1}(10, 20)$ | 0.26 (0.04) | 0.21 (0.03) | 0.06 (0.008) | 3.49(0.48) | | |
| $\sigma_z$ | $\Gamma^{-1}(10, 20)$ | 0.08 (0.01) | 0.05 (0.006) | 0.15 (0.02) | 2.09 (0.89) | | |
| $\sigma_r$ | $\Gamma^{-1}(10, 20)$ | 2.68 (0.27) | 0.10 (0.01) | 0.07 (0.007) | 0.79(0.55) | | |
| $\sigma_\mu$ | $\Gamma^{-1}(10, 20)$ | 15.98 (1.09) | 0.25 (0.04) | 36.68 (1.42) | 8.34(0.44) | | |

Table 2: Posterior estimates. LT refers to linearly detrended data, HP to Hodrick and Prescott filtered data, FOD to demeaned growth rates, BP to band pass filtered data. For Ratio 1 the observables are $\log(y_t/n_t), \log(w_t), \pi_t, r_t$, all demeaned; for Ratio 2 they are $\log(y_t/w_t), \log(n_t), \pi_t, r_t$, all demeaned; for Ratio 3, the observables are $\log((w_t n_t)/y_t), \log(w_t/y_t), \pi_t, r_t$, all demeaned. For TFP trending, the observable are linearly detrending output and real wages and demeaned inflation and interest rates. For Preference trending, the observable are demeaned growth rate of output, demeaned log real wages, demeaned inflation and demeaned interest rates. When frequency domain estimation is used, only information in the band $(\frac{\pi}{32}, \frac{\pi}{8})$ is employed. The sample is 1980:1-2007:4.

| DGP1 | | | | | | | |
|---|---|---|---|---|---|---|---|
| | True value | LT | HP | FOD | BP | Ratio1 | Flexible |
| $\sigma_n$ | 0.50 | 0.04 | 0.08 | 0.00 | 0.11 | 0.05 | 0.04 |
| $h$ | 0.70 | 0.00 | 0.00 | 0.00 | 0.01 | 0.07 | 0.10 |
| $\alpha$ | 0.30 | 0.00 | 0.04 | 0.00 | 0.06 | 0.04 | 0.06 |
| $\rho_r$ | 0.70 | 0.05 | 0.05 | 0.01 | 0.06 | 0.13 | 0.01 |
| $\rho_\pi$ | 1.50 | 0.00 | 0.00 | 0.00 | 0.01 | 0.02 | 0.00 |
| $\rho_y$ | 0.40 | 0.17 | 0.20 | 0.17 | 0.19 | 0.15 | 0.00 |
| $\zeta_p$ | 0.75 | 0.03 | 0.04 | 0.03 | 0.03 | 0.02 | 0.03 |
| $\rho_\chi$ | 0.50 | 0.00 | 0.04 | 0.00 | 0.00 | 0.00 | 0.07 |
| $\rho_z$ | 0.80 | 0.03 | 0.05 | 0.00 | 0.05 | 0.00 | 0.05 |
| $\sigma_\chi$ | 1.12 | 1.60 | 0.45 | 3.89 | 0.64 | 8.79 | 1.00 |
| $\sigma_z$ | 0.50 | 1.47 | 0.01 | 3.18 | 0.03 | 0.02 | 0.16 |
| $\sigma_r$ | 0.10 | 1.37 | 0.03 | 3.75 | 0.03 | 0.00 | 0.00 |
| $\sigma_\mu$ | 1.60 | 13.14 | 18.81 | 17.68 | 38.52 | 38.36 | 1.94 |
| Total1 | | 0.30 | 0.40 | 0.21 | 0.48 | 0.49 | 0.24 |
| Total2 | | 17.91 | 19.79 | 28.71 | 39.75 | 47.66 | 3.45 |

| DGP2 | | | | | | | |
|---|---|---|---|---|---|---|---|
| | True value | LT | HP | FOD | BP | Ratio1 | Flexible |
| $\sigma_n$ | 0.50 | 0.04 | 0.11 | 0.17 | 0.12 | 0.12 | 0.06 |
| $h$ | 0.70 | 0.01 | 0.00 | 0.00 | 0.03 | 0.08 | 0.17 |
| $\alpha$ | 0.30 | 0.00 | 0.05 | 0.00 | 0.06 | 0.02 | 0.07 |
| $\rho_r$ | 0.70 | 0.05 | 0.05 | 0.04 | 0.05 | 0.13 | 0.02 |
| $\rho_\pi$ | 1.50 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 |
| $\rho_y$ | 0.40 | 0.16 | 0.21 | 0.08 | 0.19 | 0.15 | 0.00 |
| $\zeta_p$ | 0.75 | 0.03 | 0.04 | 0.02 | 0.05 | 0.04 | 0.03 |
| $\rho_\chi$ | 0.50 | 0.00 | 0.04 | 0.00 | 0.00 | 0.01 | 0.08 |
| $\rho_z$ | 0.80 | 0.04 | 0.05 | 0.03 | 0.03 | 0.00 | 0.06 |
| $\sigma_\chi$ | 1.12 | 10.41 | 0.87 | 2.80 | 0.69 | 9.43 | 0.97 |
| $\sigma_z$ | 0.50 | 9.15 | 0.06 | 1.91 | 0.06 | 0.01 | 0.17 |
| $\sigma_r$ | 0.10 | 9.35 | 0.00 | 1.05 | 0.03 | 0.00 | 0.00 |
| $\sigma_\mu$ | 1.60 | 10.41 | 20.72 | 20.33 | 57.03 | 40.17 | 1.90 |
| Total1 | | 0.29 | 0.46 | 0.32 | 0.51 | 0.55 | 0.35 |
| Total2 | | 39.65 | 22.20 | 26.44 | 58.34 | 50.17 | 3.54 |

Table 3: MSE. In DPG1 there is a unit root component to the preference shock and $\frac{\sigma^{nc}_{\chi}}{\sigma^{T}_{\chi}} = [1.1, 1.9]$. In DGP2 all shocks are stationary but there is measurement error and $\frac{\sigma_u}{\sigma^{T}_{\chi}} = [0.09, 0.11]$ The MSE is computed using 50 replications. LT refers to linearly detrended data, HP to Hodrick and Prescott filtered data, FOD to demeaned growth rates, BP to band pass filtered data, Ratio1 to real variables

scaled by hours, and Flexible to the approach suggested in the paper. Total1 is the total MSE for the first seven parameters; total2 the MSE for all 13 parameters.

| | LT | | FOD | | Flexible | |
|---|---|---|---|---|---|---|
| | Output | Inflation | Output | Inflation | Output | Inflation |
| TFP shocks | 0.01 | 0.04 | 0.00 | 0.01 | 0.01 | 0.21 |
| Gov. expenditure shocks | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 |
| Investment shocks | 0.08 | 0.00 | 0.00 | 0.00 | 0.00 | 0.05 |
| Monetary policy shocks | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 |
| Price markup shocks | 0.75(*) | 0.88(*) | 0.91(*) | 0.90(*) | 0.00 | 0.19 |
| Wage markup shocks | 0.00 | 0.01 | 0.08 | 0.08 | 0.03 | 0.49(*) |
| Preference shocks | 0.11 | 0.04 | 0.00 | 0.00 | 0.94(*) | 0.00 |

Table 4: Variance decomposition at the 5 years horizon. Estimates are obtained using the median of the posterior of the parameters. A (*) indicates that the 68 percent highest credible set is entirely above 0.10. The model and the data set are the same as in Smets and Wouters, 2007. LT refers to linearly detrended data, FOD to growth rates and Flexible to the approach this paper suggests.

# 8 Figures

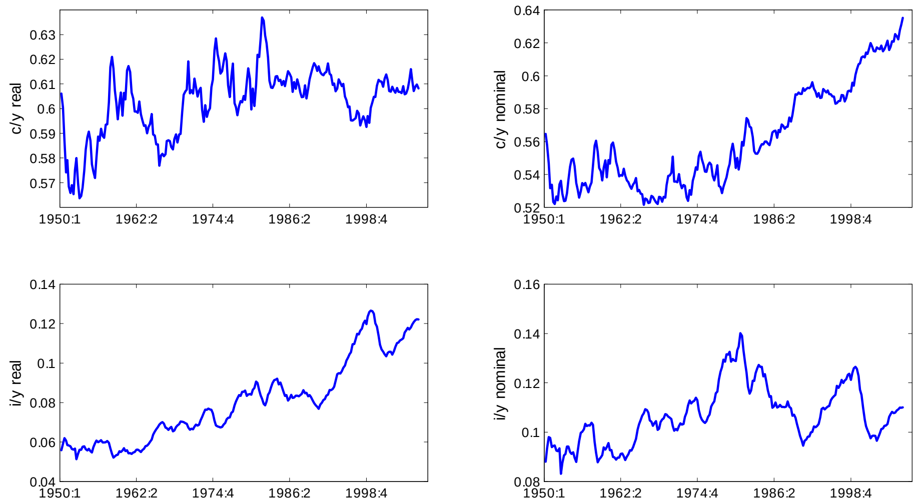


Figure 1: US real and nominal great ratios

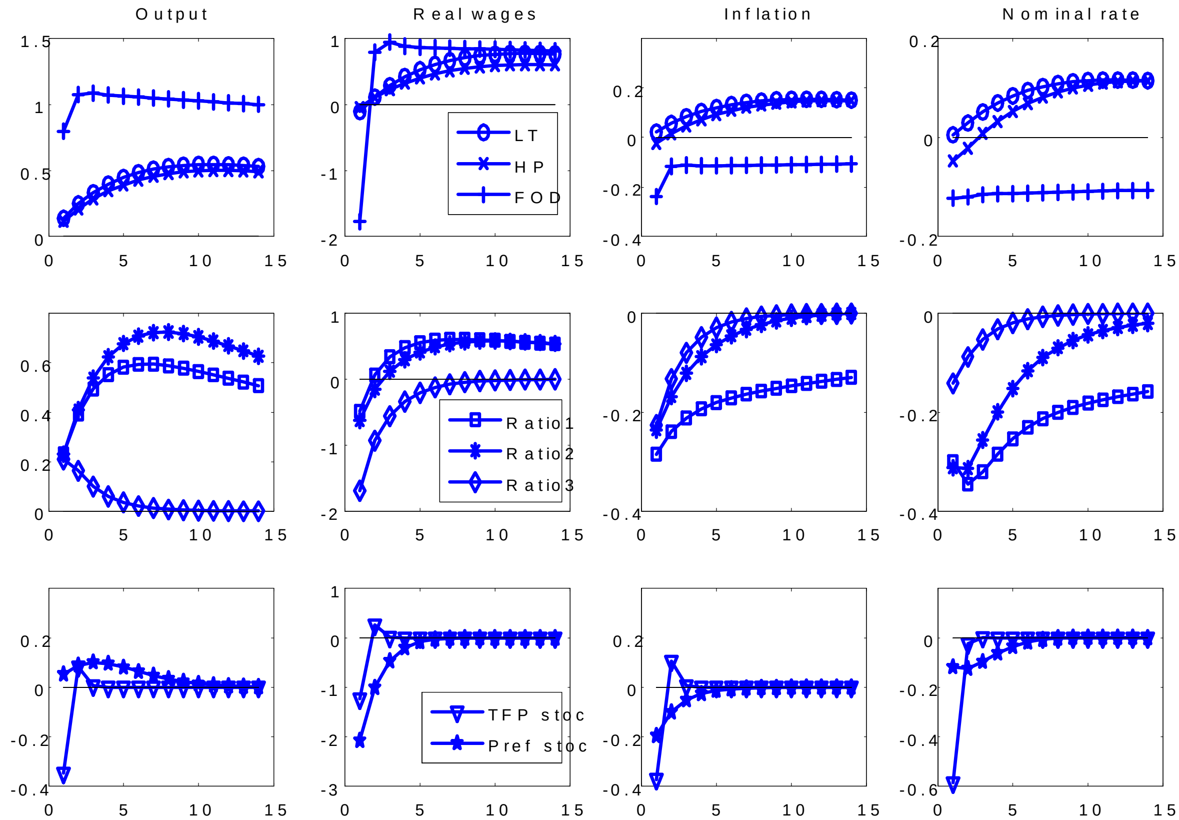


Figure 2: Impulse responses to technology shocks, sample 1980:1-2007:4

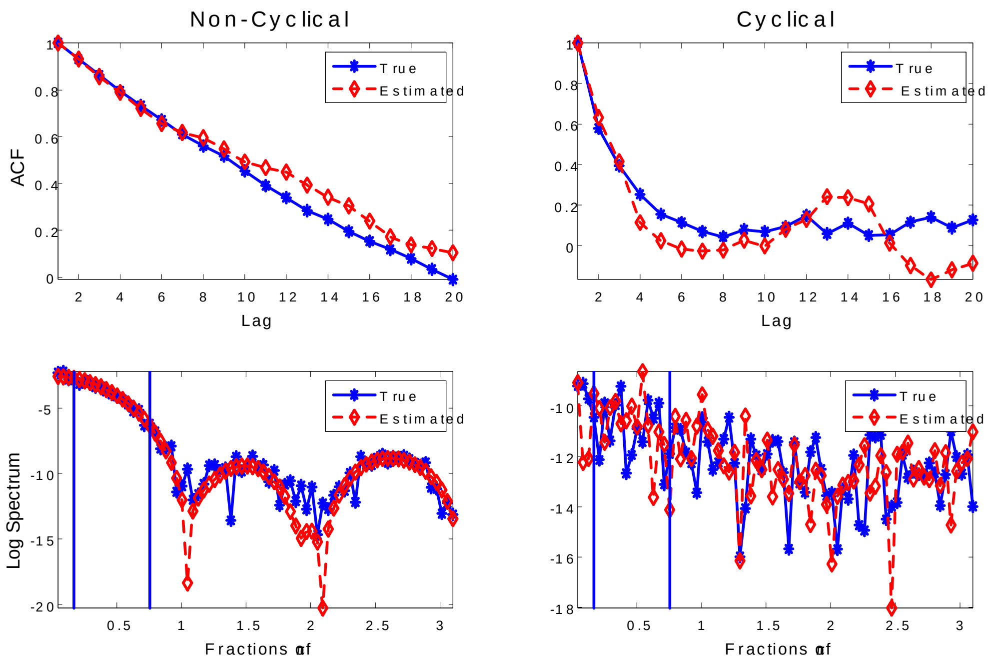


Figure 3: Output decomposition, true and estimated with a flexible approach. Vertical bars indicate business cycle frequencies

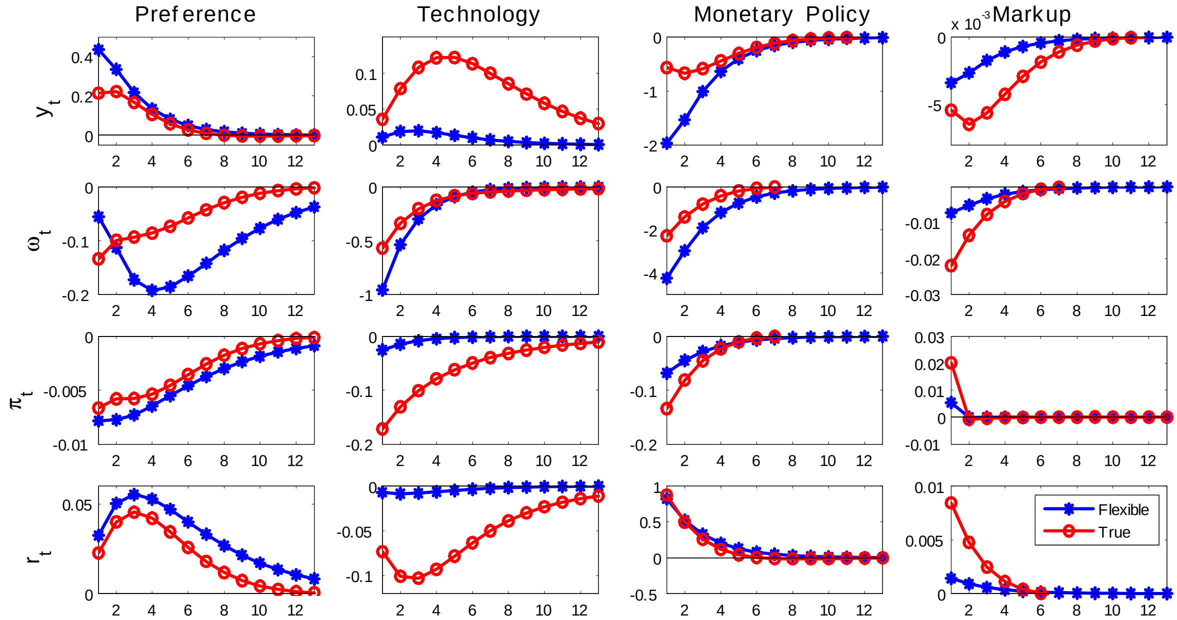


Figure 4: Impulse responses to transitory shocks, true and estimated with flexible approach.

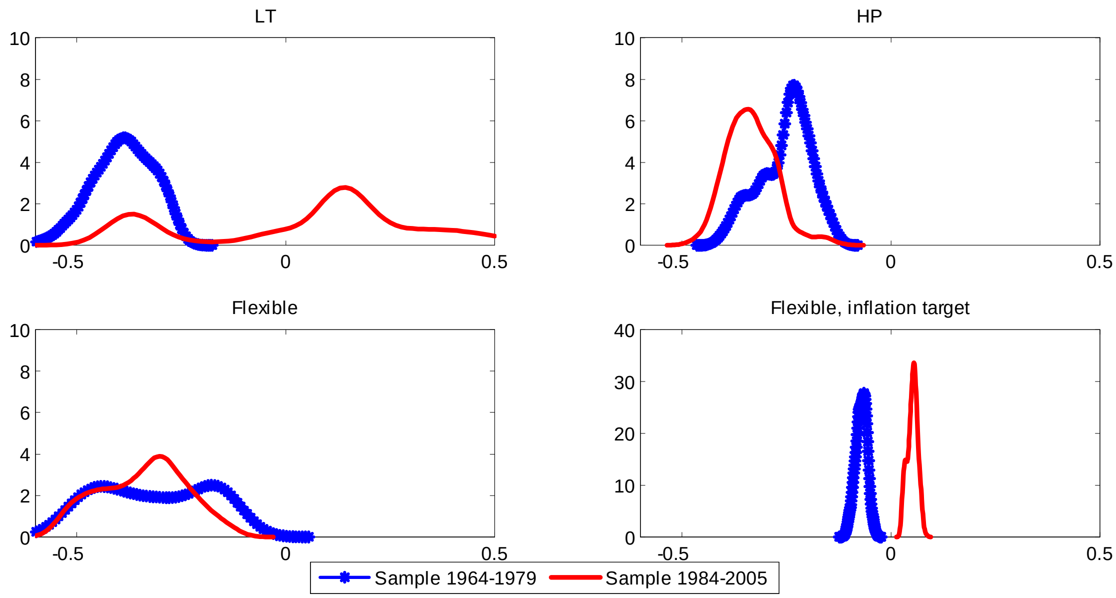


Figure 5: Posterior distributions of the policy activism parameter, samples 1964:1-1979:4 and 1984:1-2007:4. LT refers to linearly detrended data, HP to Hodrick and Prescott filtered data and Flexible to the approach the paper suggests

# References

[1] Mark Aguiar and Gita Gopinath, 2007. Emerging market business cycles: The cycle is the trend, *Journal of Political Economy*, 115, 69–102.

[2] Surmu Altug, 1989. Time to build and aggregate fluctuations: some new evidence, *International Economic Review*, 20, 883–920.

[3] Michal Andrle, 2008. The role of trends and detrending in DSGE model: Emerging countries need trendy models, manuscript.

[4] Fabio Canova, 1998. Detrending and business cycles facts, *Journal of Monetary Economics*, 41, 475–512.

[5] Fabio Canova, 2007. *Applied Macroeconomic Research*, Princeton University Press, Princeton, New Jersey.

[6] Fabio Canova and Filippo Ferroni, 2011. Multiple filtering devices for the estimation of cyclical DSGE models, *Quantitative Economics*, 2, 73-98.

[7] Fabio Canova and Matthias Paustian, 2011. Business cycle measurement with some theory, *Journal of Monetary Economics*, 48, 345-366.

[8] Fabio Canova and Luca Sala, 2009. Back to square one: identification issues in DSGE models, *Journal of Monetary Economics*, 56, 431–449.

[9] Jean Philipppe Cayen, Marc Andre Gosselin and Sharon Kozincki, 2009. Estimating DSGE models consistent trends for use in forecasting, Bank of Canada working paper 2009-35.

[10] Yongsung Chang, Taeyoung Doh, and Frank Schorfheide, 2007. Non stationary hours in a DSGE model, *Journal of Money Credit and Banking* , 39, 1357-1373.

[11] V. V. Chari, Patrick Kehoe and Ellen McGrattan, 2009. New Keynesian Models are not yet useful for policy analysis, *American Economic Journal: Macroeconomics*, 1, 242-266.

[12] Richard Clarida, Jordi Gali and Mark Gertler, 2000. Monetary policy rules and Macroeconomic Stability: Evidence and Some Theory *Quarterly Journal of Economics*, 113, 147–180.

[13] Lawrence J. Christiano and Robert Vigfusson, 2003. Maximum likelihood in the frequency domain: the importance of time to plan *Journal of Monetary Economics*, 50, 789–815.

[14] Timothy Cogley, 2001. Estimating and testing rational expectations models when the trend specification is uncertain, *Journal of Economic Dynamics and Control*, 25, 1485–1525.

[15] Davide Delle Monache and Andrew C. Harvey, 2010. The effect of misspecification in models for extracting trends and cycles, manuscript.

[16] Marco Del Negro, Frank Schorfheide, Frank Smets, and Rafael Wouters, 2006. On the fit of New Keynesian models, *Journal of Business and Economic Statistics*, 25, 123–143.

[17] Frank Diebold, Lee Ohanian and Jeremy Berkovitz, 1998. Dynamic equilibrium economies: a framework for comparing models and data, *Review of Economic Studies*, 65, 433-451.

[18] Jana Eklund, Richard Harrison, George Kapetanios, and Alasdair Scott, 2008. Breaks in DSGE models, manuscript.

[19] Jesus Fernandez Villaverde and Juan Rubio Ramirez, 2008. How structural are structural parameter values?, *NBER Macroeconomic Annual*, 24, 83–137.

[20] Filippo Ferroni, 2011. Trend agnostic one step estimation of DSGE models, *B.E. Journal in Macroeconomics*, 11(1), (Advances), Article 25.

[21] Victor Gomez, 1999. Three equivalent methods for filtering finite nonstationary time series, *Journal of Business and Economic Statistics*, 17, 109-116.

[22] Yuriy Gorodnichenko and Serena Ng, 2010. Estimation of DSGE models when the data are persistent, *Journal of Monetary Economics*, 57, 325–340.

[23] Lars Hansen and Tom Sargent, 1993. Seasonality and approximation errors in rational expectations models, *Journal of Econometrics*, 55, 21–55.

[24] Andrew C. Harvey and Andreas Jaeger, 1993. Detrending, stylized facts and the business cycle, *Journal of Applied Econometrics*, 8, 231–247.

[25] Peter N. Ireland, 2004. Technology shocks in the new Keynesian model, *The Review of Economics and Statistics*, 86, 923–936.

[26] Peter N. Ireland, 2004. A method for taking models to the data, *Journal of Economic Dynamics and Control*, 28, 1205–1226.

[27] Peter N. Ireland, 2007. Changes in the Federal Reserve´s Inflation target: causes and consequences, *Journal of Money Credit and Banking*, 38, 1851–1882.

[28] Peter N. Ireland, 2012. Stochastic Growth in the US and the Euro Area, *Journal of the European Economic Association*, 11, 1-24.

[29] Alejandro Justiniano, Giorgio Primiceri and Andrea Tambalotti, 2010. Investment shocks and the Business Cycle, *Journal of Monetary Economics*, 57, 132-145.

[30] Ellen McGrattan, 2010. Measurement with minimal theory, *Federal Reserve Bank of Minneapolis, Quarterly Review*, 33, 2-14.

[31] Ellen McGrattan, 1994, The macroeconomic effects of distortionary taxation, *Journal of Monetary Economics*, 33, 573-601.

[32] Pau Rabanal and Juan Rubio, 2005. Comparing new Keynesian models of the business cycle: A Bayesian approach, *Journal of Monetary Economics*, 25, 1151–1166.

[33] Frank Smets and Raf Wouters, 2003. An estimated dynamic stochastic general equilibrium model of the euro area, *Journal of European Economic Association*, 1, 1123–1175.

[34] Frank Smets and Raf Wouters, 2007. Shocks and frictions in the US Business Cycle: A Bayesian DSGE approach, *American Economic Review*, 97, 586–606.

[35] Karl Whelan, 2005. New evidence on balanced growth, stochastic trends and economic fluctuations, manuscript.

# On-line Appendix (not intended for publication)

## A. The basic DSGE model of section 2

The bundle of goods consumed by the representative household is

$$C_t = \left( \int_0^1 C_t(j)^{\frac{\epsilon_t - 1}{\epsilon_t}} dj \right)^{\frac{\epsilon_t}{\epsilon_t - 1}} \tag{10}$$

where $C_t(j)$ is the consumption of the good produced by firm $j$ and $\epsilon_t$ the elasticity of substitution between varieties. Maximization of the consumption bundle, given total expenditure, leads to

$$C_t(j) = \left( \frac{P_t(j)}{P_t} \right)^{-\epsilon_t} C_t \tag{11}$$

where $P_t(j)$ is the price of the good produced by firm $j$. Consequently, the price deflator is $P_t = \left( \int_0^1 P_t(j)^{1-\epsilon_t} dj \right)^{\frac{1}{1-\epsilon_t}}$ and $P_t C_t = [\int_0^1 P_t(j) C_t(j) dj]$.

The representative household chooses sequences for consumption and leisure to maximize

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[ X_t \frac{1}{1-\sigma_c} (C_t - hC_{t-1})^{1-\sigma_c} - \frac{1}{1+\sigma_n} N_t^{1+\sigma_n} \right] \tag{12}$$

where $X_t$ is an exogenous utility shifter following an AR(1) in logs:

$$\chi_t = \rho_\chi \chi_{t-1} + \epsilon_t^\chi \tag{13}$$

where $\chi_t = \ln X_t$ and $\epsilon_t^\chi \sim N(0, \sigma_\chi^2)$. The household budget constraint is

$$P_t C_t + b_t B_t = B_{t-1} + W_t N_t \tag{14}$$

where $B_t$ are one-period bonds with price $b_t$, $W_t$ is nominal wage and $N_t$ is hours worked.

There is a continuum of firms, indexed by $j \in [0, 1]$, each of which produces a differentiated good. The common technology is:

$$Y_t(j) = Z_t N_t(j)^{1-\alpha} \tag{15}$$

where $Z_t$ is an exogenous productivity disturbance following an AR(1) in log,

$$z_t = \rho_z z_{t-1} + \epsilon_t^z$$

where $z_t = \ln Z_t$ and $\epsilon_t^z \sim N(0, \sigma_z^2)$. Each firm resets its price with probability $1 - \zeta_p$ in any $t$, independently of time elapsed since the last adjustment. Therefore, aggregate price dynamics are

$$\Pi_t^{1-\epsilon_t} = \zeta_p + (1 - \zeta_p)(P_t^*/P_{t-1})^{1-\epsilon_t} \tag{16}$$

A reoptimizing firm chooses the $P_t^*$ that maximizes the current value of discounted profits

$$\max_{P_t^*} \sum_{k=0}^{\infty} \zeta_p^k E_t Q_{t,t+k} \left[ P_t^* Y_{t+k|t} - TC_{t+k}(Y_{t+k|t}) \right] \tag{17}$$

subject to the sequence of demand constraints

$$Y_{t+k|t} = \left( \frac{P_t^*}{P_{t+k}} \right)^{-\epsilon_{t+k}} Y_{t+k} \tag{18}$$

$k = 0, 1, 2, ...$ where $Q_{t,t+k} \equiv \beta^k (C_{t+k}/C_t)(P_t/P_{t+k})$, $TC(.)$ is the total cost function, and $Y_{t+k|t}$ denotes output in period $t + k$ for a firm that resets its price at $t$.

Finally, the monetary authority sets the nominal interest rate according to

$$r_t = \rho_r r_{t-1} + (1 - \rho_r)(\rho_\pi \pi_t + \rho_y gdp_t) + \epsilon_t^r \tag{19}$$

where $\epsilon_t^r \sim N(0, \sigma_{ms}^2)$.

The first order conditions of the optimization problems are:

$$\begin{aligned}
0 &= X_t (C_t - hC_{t-1})^{-\sigma_c} - \lambda_t && (20) \\
0 &= -N_t^{-\sigma_n} - \lambda_t \frac{W_t}{P_t} && (21) \\
1 &= E_t \left[ \beta \frac{\lambda_{t+1}}{\lambda_t} \frac{P_{t+1}}{P_t} R_t \right] && (22) \\
0 &= \sum_{k=0}^{\infty} \zeta_p^k E_t Q_{t,t+k} Y_{t+k|t} \left[ P_t^* - \mathcal{M}_{t+k} MC_{t+k|t}^n ) \right] && (23)
\end{aligned}$$

where $\lambda_t$ is the Lagrangian multiplier associated with the consumer budget constraint, $R_t \equiv 1 + i_t = 1/b_t$ is the gross nominal rate of return on bonds, $MC^n(.)$ are nominal marginal cost and

$$\mathcal{M}_t = \mu e^{\epsilon_t^\mu} \tag{24}$$

where $\epsilon_t^\mu \sim N(0, \sigma_\mu^2)$ and $\mu$ is the steady state markup.

Market clearing requires

$$\begin{aligned} Y_t(j) &= C_t(j) && (25) \\ N_t &= \int_0^1 N_t(j)dj && (26) \end{aligned}$$

and letting the aggregate output be $GDP_t \equiv \left( \int_0^1 Y_t(j)^{\frac{\epsilon_t - 1}{\epsilon_t}} dj \right)^{\frac{\epsilon_t}{\epsilon_t - 1}}$ we have $C_t = GDP_t$.

The shocks driving the dynamics of the model are: a preference disturbance $\chi_t$, a technology disturbance $z_t$, a markup shock $\epsilon_t^\mu$ and a monetary shock $\epsilon_t^r$.

# B. The solution with transitory shocks

When all the shocks are transitory, the log-linearized equilibrium conditions are:

$$\begin{aligned} w_t &= \left(\frac{\sigma_n}{1-\alpha} + \frac{\sigma_c}{1-h}\right) y_t - \frac{h\sigma_c}{1-h} y_{t-1} - \frac{\sigma_n}{1-\alpha} z_t - \chi_t && (27) \\ y_t &= E_t\left[\frac{1}{1+h} y_{t+1} - \frac{h}{1+h} y_{t-1} + \frac{1-h}{(1+h)\sigma_c}(\chi_{t+1} - \chi_t + r_t - \pi_{t+1})\right] && (28) \\ \pi_t &= \beta E_t \pi_{t+1} + \kappa_p\left(\epsilon_t^\mu + w_t + \frac{\alpha}{1-\alpha} y_t - \frac{1}{1-\alpha} z_t\right) && (29) \\ r_t &= \rho_r r_{t-1} + (1-\rho_r)(\rho_y y_t + \rho_\pi \pi_t) + \epsilon_t^r && (30) \\ n_t &= \frac{1}{1-\alpha}(y_t - z_t) && (31) \end{aligned}$$

where all variables are expressed in deviation from the (constant) steady state, $k_p = \frac{(1-\beta\zeta_p)(1-\zeta_p)}{\zeta_p} \frac{1-\alpha}{1-\alpha+\psi\alpha}$, $z_t = \rho_z z_{t-1} + \epsilon_t^z$, $\chi_t = \rho_\chi \chi_{t-1} + \epsilon_t^\chi$, $\epsilon_t^r$ and $\epsilon_t^\mu$ are iid. Equation (27) defines the equilibrium real wage, (28) is an Euler equation, (29) a Phillips curve, (30) a Taylor rule and (31) a labour demand function.

This is the model fitted to filtered data (first four columns on the top part of table 2) and to transformed data (the next three columns of table 2).

# C. The solution with a stochastic trend in the technology

Assume that the technology has a stochastic linear trend, i.e. $z_t = bt + \epsilon_t^z$, while the other three shocks are assumed to be transitory. A log-linearized solution can be found only setting

$\sigma_c = 1$. Defining $\bar{h} = \exp(b)h$, the equations in this case are

$$
\begin{aligned}
w_t &= \left(\frac{\sigma_n}{1-\alpha} + \frac{1}{1-\bar{h}}\right)y_t - \frac{\bar{h}}{1-\bar{h}}y_{t-1} - \chi_t + \frac{\bar{h}}{1-\bar{h}}(\epsilon_{t-1}^{z,p} - \epsilon_t^{z,p}) && (32)\\
y_t &= \frac{1}{1+\bar{h}}E_t(y_t + hy_{t-1} - (1-\bar{h})(\chi_{t+1} - \chi_t + r_t - \pi_{t+1}) + \bar{h}\epsilon_{t-1}^{z,p} + \epsilon_{t+1}^{z,p} - (1-\bar{h})\epsilon_t^{z,p}) && (33)\\
\pi_t &= \beta E_t \pi_{t+1} + \frac{1-\alpha}{1-\alpha-\alpha\theta}\frac{(1-\beta\zeta_p)(1-\zeta_p)}{\zeta_p}(\epsilon_t^{\mu} + w_t + \frac{\alpha}{1-\alpha}y_t) && (34)\\
r_t &= \rho_r r_{t-1} + (1-\rho_r)(\rho_y y_t + \rho_\pi \pi_t) + \epsilon_t^r && (35)\\
n_t &= \frac{1}{1-\alpha}(y_t - z_t) && (36)
\end{aligned}
$$

where all variables are expressed in deviation from the (constant) steady state, $k_p = \frac{(1-\beta\zeta_p)(1-\zeta_p)}{\zeta_p}\frac{1-\alpha}{1-\alpha+\psi\alpha}$, $\chi_t = \rho_\chi \chi_{t-1} + \epsilon_t^\chi$, $\epsilon_t^r$ and $\epsilon_t^\mu$ are iid. Then

$$
\begin{aligned}
\ln Y_t - c_y - bt &= y_t + \epsilon_t^z && (37)\\
\ln W_t - c_w - bt &= w_t + \epsilon_t^z && (38)\\
\Pi_t - c_\pi &= \pi_t && (39)\\
R_t - c_r &= r_t && (40)
\end{aligned}
$$

where capital letters indicate the observable variables, lower case letters the model variables and $c_j$ are constants (the mean of each process). This is the model fitted to the data in columns 8 and 10 of the bottom part of table 2.

## D. The solution with non-stationary preference shocks

Assume that $\chi_t = \chi_{t-1} + \epsilon_t^\chi$. A log linearized solution can be found only setting $\sigma_c = 1.0$ and $\alpha = 0$. The log-linearized equilibrium conditions are

$$
\begin{aligned}
w_t &= (\sigma_n + \frac{1}{1-h})y_t - \frac{h}{1-h}y_{t-1} - \sigma_n z_t + \frac{h}{1-h}\epsilon_t^{\chi,p}) && (41)\\
y_t &= \frac{1}{1+h}E_t(y_{t+1} + hy_{t-1} - (1-h)(r_t - \pi_{t+1}) - (h\epsilon_t^{\chi,p} + ((1-h)\sigma_n - h)\epsilon_{t+1}^{\chi,p})) && (42)\\
\pi_t &= \beta E_t \pi_{t+1} + \frac{(1-\beta\zeta_p)(1-\zeta_p)}{\zeta_p}(\epsilon_t^{\mu} + w_t - z_t) && (43)\\
r_t &= \rho_r r_{t-1} + (1-\rho_r)(\rho_y y_t + \rho_\pi \pi_t) + \epsilon_t^r && (44)\\
n_t &= y_t - z_t && (45)
\end{aligned}
$$

where all variables are expressed in deviation from the (constant) steady state, $k_p = \frac{(1-\beta\zeta_p)(1-\zeta_p)}{\zeta_p}$, $z_t = \rho_z z_{t-1} + \epsilon_t^z$, $\epsilon_t^r$ and $\epsilon_t^{\mu}$ are iid. Then

$$
\begin{aligned}
\ln \Delta Y_t - c_y &= y_t + \epsilon_t^{\chi} && (46)\\
\ln W_t - c_w &= w_t && (47)\\
\Pi_t - c_\pi &= \pi_t && (48)\\
R_t - c_r &= r_t && (49)
\end{aligned}
$$

where capital letters indicate the observable variables, lower case letters the model variables and $c_j$ are constants (the mean of the process). This is the model fitted to the data in column 9 of table 2.

## E. Simulating data from a model with non-stationary preference shocks

Let $Y_t^o$ be a $N \times 1$ vector of observables and let:

$$Y_t^o = \nu(\theta^*, \vartheta^*) + H^{ns}x_t^{ns} + H^s x_t^s \tag{50}$$

where $x_t^s$ is $N_s \times 1$ vector containing the variables rescaled by the non-stationary preference shock in log deviations from the steady state, $\nu(\theta^*, \vartheta^*)$ is an $N \times 1$ vector of the logarithm of the (rescaled) variables at the steady state, and $x_t^{ns}$ is $N_{ns} \times 1$ vector containing the logarithm of the non-stationary preference shock. $H^{ns}$ is an $N \times N_{ns}$ selection matrix and $H^s$ is an $N \times N_s$ selection matrix. Finally, $\theta \in \Theta_s$ is the vector of structural parameters describing the stationary dynamics of the DSGE model and $\vartheta \in \Theta_{ns}$ is the vector of parameters that defines the non-stationary dynamics. Moreover, $\theta^* \in \Theta^*_s \subset \Theta_s$ and $\vartheta^* \in \Theta^*_{ns} \subset \Theta_{ns}$ are the vectors of parameters that affect the steady state values. Rescaled variables, $x^s_t$, evolve according to

$$x^s_{t+1} = \Phi(\theta, \vartheta) x^s_t + \Psi(\theta, \vartheta)\eta_{t+1} \qquad \eta_t \sim N(0, \Sigma(\theta, \vartheta)) \tag{51}$$

where $\eta_t$ is the vector of the structural innovations of the shock processes, $\eta_t = [\eta^{ns}_t, \eta^s_t]'$. It turns out that, for the particular model we have chosen, these equations are given (41)-(45) The vector of non-stationary shock processes $\log X^P_t$is assumed to follow

$$\ln X^P_t = \ln X^P_{t-1} + e^{X,P}_t \tag{52}$$

while the vector of transitory shock processes is

$$\begin{aligned}
\log z_t &= \rho_z \log z_{t-1} + e^z_t && (53) \\
\log \chi_t &= \rho_\chi log \chi_{t-1} + e^\chi_t && (54) \\
v_t &= e^v_t && (55) \\
\mu_t &= e^\mu_t && (56)
\end{aligned}$$

Thus:

$$x_t^s = [y_t, w_t, \pi_t, r_t, z_t, \chi_t]' \tag{57}$$

$$x_t^{ns} = \ln X_t^P \tag{58}$$

$$\eta_t^s = [e_t^z, e_t^\chi, v_t, \mu_t]' \tag{59}$$

$$\eta_t^{ns} = e_t^{X,P} \tag{60}$$

$$\nu(\theta^*, \vartheta^*) = [\ln y_s, \ln W_s, \ln \Pi_s, \ln R_s]' \tag{61}$$

$$H^{ns} = [1, 1, 0, 0]' \tag{62}$$

$$H^s = \begin{pmatrix} I_{4\times4} & 0_{4\times2} \\ & \end{pmatrix} \tag{63}$$

$$\theta = [h, \sigma_n, \rho_r, \rho_y, \rho_\pi, k_p, \rho_z, \rho_\chi, \sigma_z, \sigma_x, \sigma_r, \sigma_\mu] \tag{64}$$

$$\vartheta = \sigma_{X,P} \tag{65}$$

# F. The medium scale DSGE model used in section 5

(a): The variables of the model

| Label | Definition |
|---|---|
| $y_t$ | : output |
| $c_t$ | : consumption |
| $i_t$ | : investment |
| $q_t$ | : Tobin's $q$ |
| $k_t^s$ | : capital services |
| $k_t$ | : capital |
| $z_t$ | : capacity utilization |
| $r_t$ | : real rate |
| $\mu_t^p$ | : price markup |
| $\pi_t$ | : inflation rate |
| $\mu_t^w$ | : wage markup |
| $N_t$ | : total hours |
| $w_t$ | : real wage rate |
| $R_t$ | : nominal rate |

(b): The parameters of the model

| Label | Definition |
|---|---|
| $\sigma_c$ | elasticity of intertemporal substitution |
| $\sigma_l$ | elasticity of labour supply with respect to real wages |
| $h$ | habit persistence parameter |
| $\delta$ | depreciation rate |
| $\phi_p - 1$ | share of fixed costs in production |
| $\chi$ | steady state elasticity of capital adjustment cost function |
| $\psi$ | positive function of the elasticity of capital utilization adjustment costs function. |
| $\alpha$ | share of capital services in production |
| $\gamma_p$ | price indexation parameter |
| $\zeta_p$ | price stickiness parameter |
| $\epsilon_p$ | curvature of good market aggregator |
| $\gamma_w$ | wage indexation parameter |
| $\zeta_w$ | wage stickiness parameter |
| $\epsilon_w$ | curvature of labour market aggregator |
| Label | Definition |
| $\lambda_r$ | interest smoothing parameter |
| $\lambda_\pi$ | inflation parameter |
| $\lambda_y$ | output parameter |
| $gy$ | government expenditure to output ratio |
| $ky$ | steady state capital output ratio |
| $r_* = \beta^{-1}$ | steady state rental rate |
| $w_*$ | steady state real wage rate |
| $N_*/C_*$ | steady state hours to consumption ratio |

(c): The equations of the model (in deviation from steady states)

$$y_t = (1 - gy - \delta\ ky)c_t + \delta\ ky\ i_t + r_*\ ky\ z_t + g_t \tag{C.1}$$

$$c_t = \frac{h}{1+h}E_t c_{t+1} + \frac{h}{1+h}c_{t-1} - \frac{(\sigma_c - 1)w_* N_*/C_*}{(1+h)\sigma_c}(N_t - E_t N_{t+1}) - \frac{1-h}{(1+h)\sigma_c}(R_t - E_t \pi_{t+1} + e_t^b) \tag{C.2}$$

$$i_t = \frac{\beta}{1+\beta}E_t i_{t+1} + \frac{1}{1+\beta}x_{t-1} + \frac{\chi^{-1}}{1+\beta}q_t + e_t^i \tag{C.3}$$

$$q_t = \beta(1-\delta)E_t q_{t+1} + (1 - \beta(1-\delta))E_t r_{t+1} - (R_t - E_t \pi_{t+1} + e_t^b) \tag{C.4}$$

$$y_t = \phi_p(\alpha k_t^s + (1-\alpha)N_t + e_t^a) \tag{C.5}$$

$$k_t^s = k_{t-1} + z_t \tag{C.6}$$

$$z_t = \frac{1-\psi}{\psi}r_t \tag{C.7}$$

$$k_{t+1} = (1-\delta)\ k_t + \delta\ i_t + \delta\ (1+\beta)\ \psi\ e_t^i \tag{C.8}$$

$$\mu_t^p = \alpha(k_t^s - N_t) + e_t^a - w_t \tag{C.9}$$

$$\pi_t = \frac{\beta}{1+\beta\gamma_p}E_t \pi_{t+1} + \frac{\gamma_p}{1+\beta\gamma_p}\pi_{t-1} - T_p \mu_t^p + e_t^p \tag{C.10}$$

$$r_t = -(k_t - N_t) + w_t \tag{E.11}$$

$$\mu_t^w = w_t - (\sigma_l N_t + (1-h)^{-1}(c_t - h c_{t-1}) \tag{C.12}$$

$$w_t = \frac{1}{1+\beta}w_{t-1} + \frac{\beta}{1+\beta}(E_t \pi_{t+1} + E_t w_{t+1}) - \frac{1+\beta\gamma_w}{1+\beta}\pi_t + \frac{\gamma_w}{1+\beta}\pi_{t-1} - T_w \mu_t^w + e_t^w \tag{C.13}$$

$$R_t = \lambda_r R_{t-1} + (1-\lambda_r)(\lambda_\pi \pi_t + \lambda_y y_t) + e_t^r \tag{C.14}$$

The seven disturbances are: TFP shock ($e_t^a$); monetary policy shock ($e_t^r$); investment shock ($e_t^i$); price markup shock ($e_t^p$); wage markup shock ($e_t^w$); risk premium shock ($e_t^b$); government expenditure shock ($e_t^g$). The compound parameters in equation (C.11) and (C.13) are defined as: $T_p \equiv \frac{1}{1+\gamma_p}\frac{(1-\beta\zeta_p)(1-\zeta_p)}{((\phi_p - 1)\epsilon_p)\zeta_p}$ and $T_w \equiv \frac{1}{1+\beta}\frac{(1-\beta\zeta_w)(1-\zeta_w)}{((\phi_w - 1)\epsilon_w)\zeta_w}$.

(d): The process for the shocks

$$e_t = (e_t^a, e_t^r, e_t^i, e_t^p, e_t^w, e_t^b, e_t^g)$$
$$e_t = \rho e_{t-1} + \eta_t$$

where both $\rho$ and $\Sigma = E_t \eta_t \eta_t'$ are diagonal.

# G. Additional Tables and Graphs

| | LT | HP | FOD | BP |
|---|---|---|---|---|
| | Median (s.e.) | Median (s.e.) | Median (s.e.) | Median (s.e.) |
| $\sigma_c$ | 1.68 (0.30) | 1.53 (0.26) | 0.04 (0.01) | 2.98 (0.49) |
| $\sigma_n$ | 1.73 (0.15) | 1.62 (0.12) | 5.28 (0.07) | 0.55 (0.06) |
| h | 0.85 (0.03) | 0.87 (0.03) | 0.40 (0.01) | 0.89 (0.02) |
| $\alpha$ | 0.05 (0.02) | 0.08 (0.03) | 0.41 (0.01) | 0.04 (0.02) |
| $\rho_r$ | 0.18 (0.06) | 0.16 (0.05) | 0.64 (0.01) | 0.13 (0.03) |
| $\rho_\pi$ | 1.36 (0.07) | 1.36 (0.08) | 1.48 (0.02) | 1.42 (0.06) |
| $\rho_y$ | -0.17 (0.03) | -0.17 (0.04) | 0.05 (0.00) | -0.11 (0.03) |
| $\zeta_p$ | 0.82 (0.01) | 0.82 (0.02) | 0.64 (0.01) | 0.83 (0.01) |
| $\rho_\chi$ | 0.66 (0.04) | 0.67 (0.04) | 0.54 (0.01) | 0.81 (0.03) |
| $\rho_z$ | 0.97 (0.02) | 0.97 (0.01) | 0.99 (0.01) | 0.76 (0.02) |
| $\sigma_\chi$ | 0.63 (0.18) | 0.65 (0.21) | 4.63 (0.07) | 0.45 (0.12) |
| $\sigma_z$ | 0.19 (0.04) | 0.23 (0.05) | 2.89 (0.19) | 0.14 (0.02) |
| $\sigma_{mp}$ | 0.11 (0.01) | 0.11 (0.01) | 2.69 (0.14) | 0.12 (0.01) |
| $\sigma_\mu$ | 23.13 (1.99) | 29.07 (0.94) | 7.63 (0.10) | 30.22 (1.12) |

Table G.1 Parameters estimates obtained with standard transformations; real variables filtered, nominal variables demeaned.

| DGP1 | | | | | | | |
|---|---|---|---|---|---|---|---|
| | True | LT | HP | FOD | BP | Ratio1 | Flexible |
| | | Mean (s.e.) | Mean (s.e.) | Mean (s.e.) | Mean (s.e.) | Mean (s.e.) | Mean (s.e.) |
| $\sigma_n$ | 0.50 | 0.71 ( 0.14) | 0.22 (0.18) | 0.54 (0.33) | 0.16 ( 0.15) | 0.28 ( 0.54) | 0.73 ( 0.15) |
| $h$ | 0.70 | 0.71 ( 0.09) | 0.72 (0.08) | 0.69 (0.68) | 0.82 ( 0.88) | 0.97 ( 0.04) | 0.36 ( 0.18) |
| $\alpha$ | 0.30 | 0.29 ( 0.10) | 0.10 (0.09) | 0.29 (0.32) | 0.05 ( 0.02) | 0.09 ( 0.11) | 0.05 ( 0.03) |
| $\rho_r$ | 0.70 | 0.47 ( 0.04) | 0.48 (0.10) | 0.61 (0.59) | 0.45 (0.48) | 0.33 ( 0.13) | 0.80 ( 0.08) |
| $\rho_\pi$ | 1.50 | 1.57 ( 0.05) | 1.52 (0.06) | 1.55 (1.55) | 1.42 (1.49) | 1.62 ( 0.08) | 1.55 ( 0.08) |
| $\rho_y$ | 0.40 | -0.01 (0.02) | -0.05 (0.10) | -0.01 (0.07) | -0.04 ( 0.04) | 0.01 ( 0.08) | 0.40 ( 0.16) |
| $\zeta_p$ | 0.75 | 0.92 (0.02) | 0.94 (0.02) | 0.91 (0.91) | 0.91 ( 0.97) | 0.91 ( 0.21) | 0.92 ( 0.01) |
| $\rho_\chi$ | 0.50 | 0.45 (0.07) | 0.31 (0.06) | 0.52 ( 0.51) | 0.50 ( 0.48) | 0.52 ( 0.17) | 0.76 ( 0.04) |
| $\rho_z$ | 0.80 | 0.98 (0.09) | 0.58 (0.14) | 0.80 (0.88) | 0.58 ( 0.59) | 0.79 ( 0.15) | 0.59 ( 0.16) |
| $\sigma_\chi$ | 1.12 | 2.47 (1.68) | 0.53 (0.85) | 3.17 (3.48) | 0.40 ( 0.19) | 4.17 ( 0.90) | 0.20 ( 0.03) |
| $\sigma_z$ | 0.50 | 1.71 (1.60) | 0.39 ( 0.87) | 2.28 (2.85) | 0.32 ( 0.11) | 0.37 ( 0.42) | 0.10 ( 0.02) |
| $\sigma_r$ | 0.10 | 1.27 (1.69) | 0.28 (0.96) | 2.04 (2.72) | 0.27 ( 0.06) | 0.07 ( 0.00) | 0.06 ( 0.00) |
| $\sigma_\mu$ | 1.60 | 5.22 (0.79) | 5.94 (1.00) | 5.81 (5.48) | 7.81 ( 7.92) | 7.79 ( 1.76) | 0.21 ( 0.03) |

| DGP2 | | | | | | | |
|---|---|---|---|---|---|---|---|
| | True | LT | HP | FOD | BP | Ratio1 | Flexible |
| | | Mean (s.e.) | Mean (s.e.) | Mean (s.e.) | Mean (s.e.) | Mean (s.e.) | Mean (s.e.) |
| $\sigma_n$ | 0.50 | 0.70 ( 0.05) | 0.17 ( 0.05) | 0.92 ( 0.37) | 0.15 ( 0.11) | 0.16 ( 0.20) | 0.78 ( 0.10) |
| $h$ | 0.70 | 0.60 ( 0.07) | 0.70 ( 0.07) | 0.67 ( 0.67) | 0.87 ( 0.88) | 0.98 ( 0.04) | 0.30 ( 0.02) |
| $\alpha$ | 0.30 | 0.33 ( 0.04) | 0.09 ( 0.06) | 0.29 ( 0.32) | 0.05 ( 0.03) | 0.16 ( 0.13) | 0.04 ( 0.01) |
| $\rho_r$ | 0.70 | 0.49 ( 0.03) | 0.48 ( 0.08) | 0.51 ( 0.51) | 0.47 ( 0.48) | 0.34 ( 0.14) | 0.83 ( 0.04) |
| $\rho_\pi$ | 1.50 | 1.55 ( 0.04) | 1.55 ( 0.05) | 1.57 ( 1.58) | 1.52 ( 1.52) | 1.62 ( 0.10) | 1.53 ( 0.09) |
| $\rho_y$ | 0.40 | -0.00 ( 0.00) | -0.06 ( 0.09) | 0.12 ( 0.12) | -0.04 ( 0.04) | 0.01 ( 0.03) | 0.42 ( 0.11) |
| $\zeta_p$ | 0.75 | 0.91 ( 0.01) | 0.94 ( 0.01) | 0.90 ( 0.91) | 0.97 ( 0.97) | 0.95 ( 0.00) | 0.92 ( 0.00) |
| $\rho_\chi$ | 0.50 | 0.52 ( 0.06) | 0.30 ( 0.05) | 0.55 ( 0.53) | 0.49 ( 0.47) | 0.58 ( 0.13) | 0.78 ( 0.05) |
| $\rho_z$ | 0.80 | 1.00 ( 0.00) | 0.59 ( 0.05) | 0.62 ( 0.82) | 0.63 ( 0.61) | 0.80 ( 0.11) | 0.55 ( 0.06) |
| $\sigma_\chi$ | 1.12 | 4.43 ( 1.19) | 0.27 ( 0.16) | 2.87 ( 3.07) | 0.37 ( 0.20) | 4.27 ( 0.98) | 0.21 ( 0.02) |
| $\sigma_z$ | 0.50 | 3.53 ( 1.01) | 0.25 ( 0.58) | 1.88 ( 1.50) | 0.26 ( 0.11) | 0.41 ( 0.59) | 0.09 ( 0.00) |
| $\sigma_r$ | 0.10 | 3.16 ( 1.16) | 0.11 ( 0.23) | 1.12 ( 0.07) | 0.26 ( 0.06) | 0.07 ( 0.00) | 0.06 ( 0.00) |
| $\sigma_\mu$ | 1.60 | 4.83 ( 0.39) | 6.15 ( 0.87) | 6.11 ( 5.60) | 9.15 ( 8.51) | 7.94 ( 1.36) | 0.22 ( 0.02) |

Table G.2:Average Posterior mean estimates and dispersions across replications. In DPG1 there is a unit root component to the preference shock and $\frac{\sigma_\chi^{nc}}{\sigma_\chi^T} = [1.11.9]$. In DGP2 all shocks are stationary but there is measurement error in each equation and $\frac{\sigma_u}{\sigma_\chi^T} = [0.090.11]$. The MSE is computed using 50 replications. LT refers to linearly detrended data, HP to Hodrick and Prescott filtered data, FOD to demeaned growth rates, BP to band pass filtered data, Ratio1 to real variables scaled by hours, and Flexible to the approach suggested in the paper.

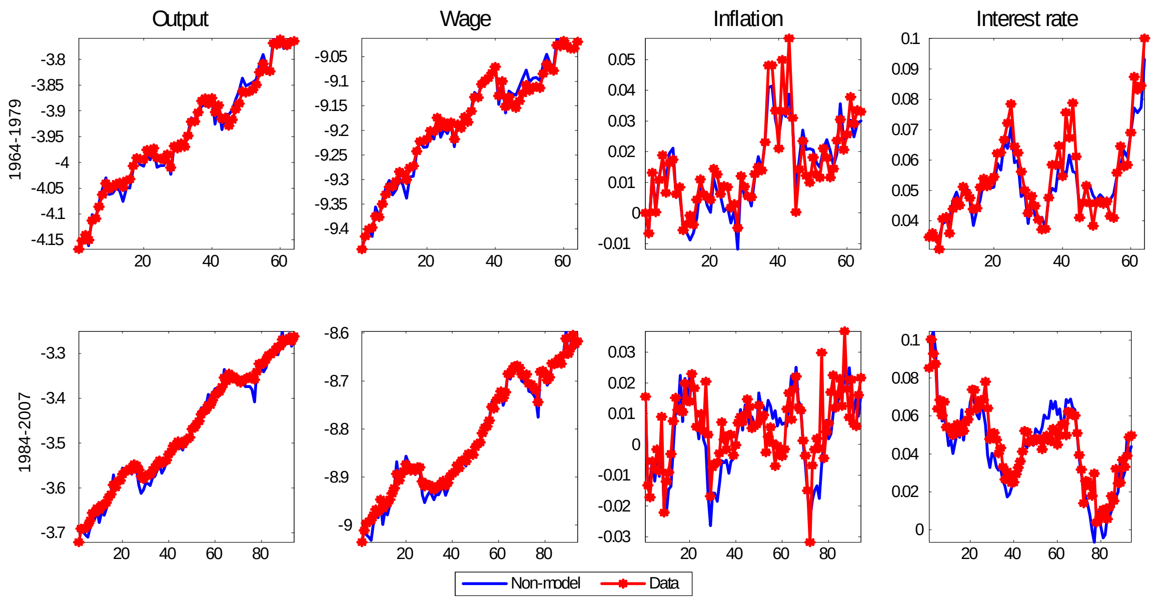


Figure G.2: Data and estimated non-model based components, samples 1964:1-1979:4 and 1984:1-2007:4, flexible approach

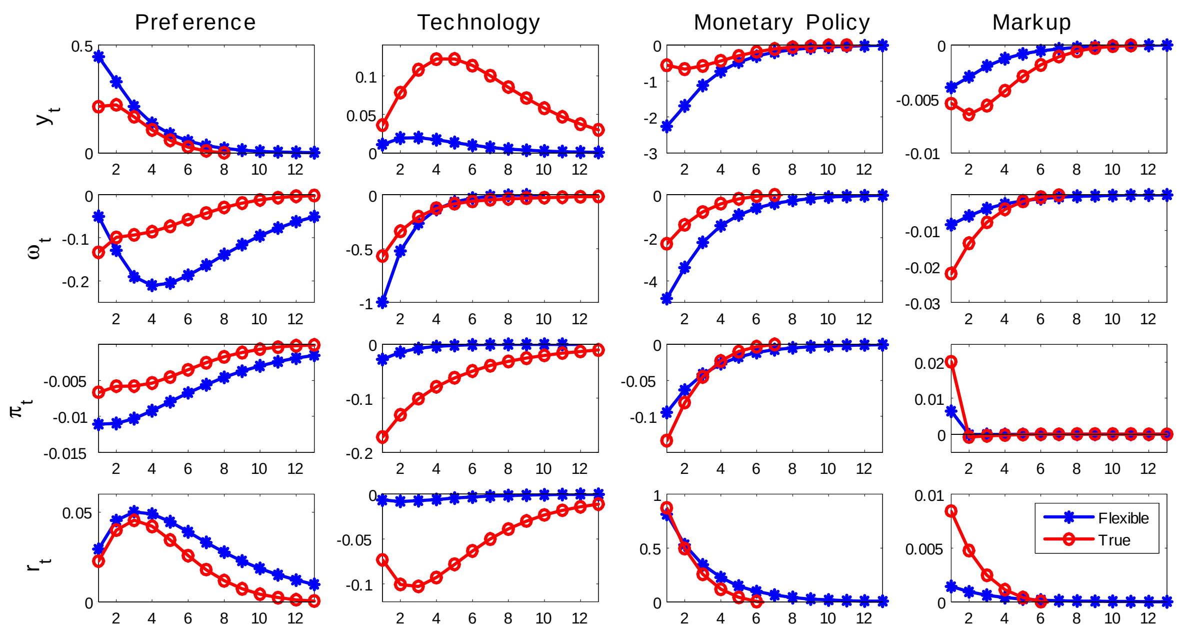


Figure G3: Impulse responses to transitory shocks, true and estimated with flexible approach, no permanenent shocks