Multiple filtering devices for the estimation of cyclical DSGE models

September 28, 2010

Abstract

We propose a method to estimate time invariant cyclical DSGE models using the information provided by a variety of filters. We treat data filtered with alternative procedures as contaminated proxies of the relevant model-based quantities and estimate structural and non-structural parameters jointly using a signal extraction approach. We employ simulated data to illustrate the properties of the procedure and compare our conclusions with those obtained when just one filter is used. We revisit the role of money in the transmission of monetary business cycles.

JEL classification: E32, C32.

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

1 Introduction

DSGE models have become the paradigm for business cycle and policy analyses in academic and policy circles. Relative to earlier structures, current models are of larger scale and feature numerous real and nominal frictions that help to closely replicate the dynamic responses that structural VARs produce. A few years ago it was standard to informally calibrate these models but today, increased computing power, and recent developments in system-wide estimation methods allow researchers to routinely employ full information techniques in structural estimation exercises.

Despite the increased popularity, structural estimation faces important conceptual and numerical problems. For example, as emphasized in Canova (2009), full information classical estimation makes sense only if the model is the data generating process (DGP) of the observables, up to a set of serially uncorrelated measurement errors. Since such an assumption is hard to entrain, unless the model is augmented with ad-hoc dynamics, Fukac and Pagan (2010) suggest to complement standard inference with a more robust limited information analysis. It is also well known that there are abundant population identification problems (see Canova and Sala (2009), Del Negro and Schorfheide (2008)), that numerical difficulties are widespread, and that errors-in-variables are present (the variables in the model do not often have a direct counterpart in the data). Finally, the vast majority of the models used in the literature are time invariant and intended to explain only the cyclical portion of the observable fluctuations while the actual data contains many types of fluctuations, all of which may be subject to breaks and other forms of slowly moving variations.

To fit stationary cyclical DSGE models to the data, applied investigators typically select a subsample where time invariance is more likely to hold, filter the raw data with an arbitrary statistical device, and treat the filtered data as the relevant measure of stationary cyclical fluctuations (see e.g. Smets and Wouters (2003), Ireland (2004)). Alternatively, one arbitrarily builds a non-cyclical component into the model (e.g. via a deterministic labor augmenting technology progress or unit roots in total factor productivity and/or the price of investment) and filters the raw data using a model-driven transformation (see e.g. Fernandez Villaverde and Rubio Ramirez (2007) or Justiniano et al. (2010)) or an arbitrary statistical device (see Smets and Wouters (2007)).

Both approaches are, in general, problematic. While the profession shares the idea that a cyclical model should explain fluctuations with an average periodicity of 8-32 quarters, there is little agreement on how to obtain these fluctuations from the data and only a partial understanding of the consequences that statistical filtering induce. For example, it is common to use linearly detrended or first differenced data as input in the estimation process, but such transformations do not isolate fluctuations with the required periodicity (see e.g. Canova (1998)). A band pass (BP) filter, which can potentially extract the fluctuations of interest with an infinite amount of data, it is typically discarded in the estimation literature because its two-sided nature alters the timing of the data information - a similar argument is also made for the Hodrick and Prescott (HP) filter. Moreover, while real variables typically show long run drifts, nominal variables just display low frequency fluctuations. Hence, should we filter all the data or only real variables? Investigators have taken both positions but it is not obvious which approach is preferable. Finally, since researchers filter each series separately, theoretically relevant constraints may not be satisfied with filtered data (for example, does a resource constraint holds with filtered data?).

Model-driven filtering also fails to extract cycles with the required periodicity. For example, when Total Factor Productivity (TFP) is trending, real variables share similar trends and appropriate linear combinations should be free of non-cyclical dynamics. However, as shown in Canova (2008), real and nominal "Great Ratios" display significant upward drifts and the portion of the variance of the transformed variables located outside the cyclical frequencies is generally large. Most problematic of all, model-based filtering requires knowledge of the number, the nature and the time series features of the shocks driving the non-cyclical component. Given our general ignorance on the subject, important specification errors may plague structural estimates.

Since solving this complex mismatch problem is difficult, this paper focuses on how to improve structural estimation of the parameters of a cyclical DSGE model when a statistical filtering approach is used to match the data to the model counterparts. We make three contributions to the existing literature. First, we show that a typical log-linearized DSGE model produces cyclical fluctuations which are not necessarily located at the so-called business cycle frequencies. Thus, standard filtering approaches induce measurement errors in the estimated cyclical components. Since these errors have important low frequency components, the true income and substitution effects are mismeasured leading to distortions in the estimates of important structural parameters. Second, we show how to design a statistical filter which captures the cyclical component of a DSGE model. This filter is model specific and the computational complexities involved make its practical implementation unfeasible on current computers. Third, we propose a method to estimate the structural parameters of a time invariant cyclical DSGE model which may potentially eliminate the biases that statistical filters produce. The approach borrows ideas from the recent data-rich environment literature (see Boivin and Giannoni (2005)). We set up a signal extraction framework where the cyclical DSGE is the unobservable factor; vectors of filtered data are contaminated observable proxies, and DSGE and non-structural parameters are jointly estimated.

Our approach is advantageous in, at least, two respects. Since we do not have to arbitrarily choose one filtering method prior to the estimation, nor to select which shock drives the non-cyclical component, we avoid important specification errors. Moreover, our method can be used with cyclical data obtained with one-sided and two-sided filters, of both univariate and multivariate nature, as long as the list of filters is sufficiently rich. For the approach to work properly, the list of filters should be carefully chosen and suggestions on how to do this in practice are provided.

We investigate the properties of our approach using experimental data of the typical length employed in macroeconomics and demonstrate that the biases obtained when just one filter is used are reduced with our approach. We also show that the unconditional one step ahead mean square error (MSE) produced by our approach is smaller than the MSE obtained with standard procedures and that conditional forecasts are better behaved.

To show that the biases are also economically relevant, we revisit the role of money in amplifying cyclical fluctuations. The recent literature has neglected the stock of money when studying monetary business cycles and Ireland (2004) demonstrates that such an approach is, by and large, appropriate using US data, standard filtering techniques and a maximum likelihood estimator. We show that when multiple filtered data is jointly used in the estimation, money balances matter for the transmission of cyclical fluctuations to output and inflation and the propagation of primitive shocks differs from the one obtained when only one data transformation is used.

We want to be clear for why we insist on working with time invariant cyclical models, rather than considering structures where cyclical and non-cyclical fluctuations are jointly accounted for. On one hand, constructing reasonable models with these features is hard: theory is largely silent on how cyclical shocks can be propagated at longer frequencies (exceptions are Comin and Gertler (2006) or Lopez Salido and Michelacci (2007)) or on how long run disturbances can produce important cyclical implications. Moreover, it is convenient for both policy and interpretation purposes to assume that the mechanisms driving cyclical and non-cyclical fluctuations are distinct and orthogonal. Finally, breaks make the data largely uninformative about the features of non-cyclical fluctuations.

The rest of the paper is organized as follows. The next section shows the problems one encounters using a single filter to estimate the parameters of DSGE models. Section 3 derives the features of an optimal filter. Section 4 presents our approach. Section 5 examines the role of money in transmitting monetary business cycles. Section 6 concludes.

2 Statistical filters and structural parameter estimates

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

$$
\begin{aligned}
0 &= \chi_t (C_t - hC_{t-1})^{-\sigma_c} - \mathcal{L}_t && (1)\\
0 &= N_t^{\sigma_n} - \mathcal{L}_t \frac{W_t}{P_t} && (2)\\
1 &= E_t \left[\beta \frac{\mathcal{L}_{t+1}}{\mathcal{L}_t} \frac{R_t}{\Pi_{t+1}} \right] && (3)\\
0 &= E_t \left(\sum_{k=0}^{\infty} \frac{P_{t+k}\mathcal{L}_{t+k}}{P_t \mathcal{L}_t} (\beta\zeta_p)^k \left[1 - \epsilon_{t+k} + MC^r_{t+k}\epsilon_{t+k} \left(\frac{\widetilde{P}_t}{P_{t+k}} \right)^{\frac{\alpha-1-\alpha\epsilon_{t+k}}{1-\alpha}} \right] Y_{t+k}(j) \right) && (4)\\
1 &= \left(\zeta_p \left(\frac{P_{t-1}}{P_t} \right)^{1-\epsilon_t} + (1-\zeta_p) \left(\frac{\widetilde{P}_t}{P_t} \right)^{1-\epsilon_t} \right)^{\frac{1}{1-\epsilon_t}} && (5)\\
Y_t &= C_t && (6)\\
N_t &= \left(\frac{Y_t}{Z_t} \right)^{\frac{1}{1-\alpha}} \int_0^1 \left(\frac{P_t(j)}{P_t} \right)^{\frac{-\epsilon_t}{1-\alpha}} dj && (7)\\
MC^r_t &= \frac{W_t}{P_t} \left(\frac{1}{Z_t} \right)^{\frac{1}{1-\alpha}} Y_t^{\frac{\alpha}{1-\alpha}} \int_0^1 \left(\frac{P_t(j)}{P_t} \right)^{\frac{-\epsilon_t \alpha}{1-\alpha}} dj && (8)\\
R_t &= R_{t-1}^{\rho_r} \pi_t^{(1-\rho_r)\rho_\pi} Y_t^{(1-\rho_r)\rho_y} v_t && (9)
\end{aligned}
$$

where h is the consumption habit coefficient, σ_c the risk aversion coefficient, $1/\sigma_n$ the Frisch elasticity, β the discount factor, $1-\alpha$ the share of labor in production, $1-\zeta_p$, the probability of changing prices, and ρ_π, ρ_y, ρ_r are the parameters of the monetary policy rule; $\mathcal{L}_t$ is the Lagrangian on the consumer budget constraint, Y_t aggregate output, $Y_t(j)$ output of good j, N_t aggregate hours, W_t the nominal wage, R_t the nominal interest rate, π_t the inflation rate, P_t the price level, $P_t(j)$ the price of good j, MC^r_t aggregate real marginal costs and $\widetilde{P}_t$ the optimal price; χ_t is a preference shock, Z_t a technology shock, ϵ_t a markup shock and v_t a monetary policy shock. The first equation equates the marginal utility of consumption to the Lagrangian; the second the intertemporal rate of substitution between leisure and consumption to the real wage and the third is a pricing relationship for one period real bonds. The next equation is a Phillips curve. Equation (5) describe the behavior of the aggregate price level. Equations (6), (7) and (8) define the resource constraints, aggregate hours and real marginal costs. The last equation is the policy rule of the central bank. A full description of the model and the log-linearized conditions are in appendix A.

For the sake of illustration, we consider two situations. In the first one, $\ln \chi_t = \rho_\chi \ln \chi_{t-1} + e_t$, where $e_t \sim N(0, \sigma_\chi^2)$; $\ln \epsilon_t = \epsilon + \frac{1-\epsilon}{\epsilon}\mu_t$, where $\mu_t \sim N(0, \sigma_\mu^2)$, $\ln v_t \sim N(0, \sigma_v^2)$ and $Z_t = Z_{t,c}Z_{t,T}$, where $\ln Z_{t,T} = \gamma t + e_{t,T}$ with $e_{t,T} \sim N(0, \sigma_{Z,T}^2)$ and $\ln Z_{t,c} = \rho_z \ln Z_{t-1,c} + e_{t,c}$ with $e_{t,c} \sim N(0, \sigma_{Z,c}^2)$ (DGP1). In the second case $\chi_t = \chi_{t,c}\chi_{t,T}$ where $\ln \chi_{t,c} = \rho_\chi \ln \chi_{t-1,c} + e_{t,c}$ with $e_{t,c} \sim N(0, \sigma_{\chi,c}^2)$; $\ln \chi_{t,T} = \ln \chi_{t-1,T} + e_{t,T}$ with $e_{t,T} \sim N(0, \sigma_{\chi,T}^2)$; $\ln \epsilon_t = \epsilon + \frac{1-\epsilon}{\epsilon}\mu_t$, where $\mu_t \sim N(0, \sigma_\mu^2)$, $\ln v_t \sim N(0, \sigma_v^2)$ and $\ln Z_t = \rho_z \ln Z_{t-1} + e_t$, where $e_t \sim N(0, \sigma_Z^2)$ (DGP2). Thus, in both specifications, there are four shocks driving cyclical (stationary) fluctuations and one shock driving non-cyclical (non-stationary) fluctuations. However, in DGP1 non-cyclical fluctuations are driven by a technology shock which is stochastic around a linear trend and in DGP2 they are driven by a preference shock displaying a unit root. For both DGPs, we set $\beta = 0.99; \sigma_c = 1.00; h = 0.70; \sigma_n = 0.70; \epsilon = 7.0; \rho_r = 0.2; \rho_\pi = 1.30; \rho_y = 0.05; \zeta_p = 0.8\ \rho_\chi = 0.5; \rho_z = 0.8; \sigma_v = 0.0012; \sigma_\mu = 0.2064$. In DGP1 we select $\alpha = 0.4, \sigma_\chi = 0.0112$, $\gamma = 0.002$, $\sigma_{Z,T} = 0.003$; $\sigma_{Z,c} = 0.0051$ and in DGP2 $\alpha = 0.0$; $\sigma_Z = 0.0051$, $\sigma_{\chi,c} = 0.0112$, and $\sigma_{\chi,T} = 0.0012$. None of the points we make, however, depends on the choice of these parameters.

		DGP1			DGP2		
Variable	Filter	St.dev.	AR1	corr(y,π)	St.dev.	AR1	corr(y,π)
Output	LT	0.0486	0.925	0.864	0.0123	0.911	-0.196
	HP	0.0366	0.876	0.834	0.0065	0.691	-0.288
	BP	0.0377	0.908	0.939	0.0060	0.859	-0.553
	FOD	0.0188	0.608	0.513	0.0052	0.100	-0.029
	True	0.0295	0.914	0.728	0.0082	0.811	-0.324
Inflation	LT	0.0043	0.703		0.0100	-0.005	
	HP	0.0037	0.602		0.0095	-0.083	
	BP	0.0034	0.873		0.0050	-0.810	
	FOD	0.0033	-0.138		0.0138	-0.495	
	True	0.0022	0.590		0.0098	0.005	

Table 1: Moments of filtered and true cyclical components; simulated data.

Table 1 presents a few moments of filtered output and filtered inflation when linear (LT), Hodrick and Prescott (HP), band pass (BP) and first order difference (FOD) filtering are used together with

Filter	True	Prior	LT Median (s.e.)	HP Median (s.e.)	FOD Median (s.e.)	BP Median (s.e.)
σ_c	1.00	Γ(0.1,0.1) [1.00, 10.0]	3.77 (0.25)	4.38 (0.36)	2.21 (0.16)	5.23 (0.24)
σ_n	0.70	Γ(0.5,0.5) [1.00, 4.0]	0.28 (0.05)	0.13 (0.02)	0.04 (0.00)	0.06 (0.01)
h	0.70	B(10,3) [0.76, 0.11]	0.58 (0.03)	0.61 (0.06)	0.69 (0.03)	0.85 (0.05)
ϵ	7.00	N(6,0.5) [6.00, 0.50]	3.95 (0.13)	3.95 (0.13)	4.05 (0.13)	3.96 (0.13)
ρ_r	0.20	B(10,6) [0.71, 0.09]	0.30 (0.01)	0.27 (0.01)	0.39 (0.01)	0.59 (0.02)
ρ_π	1.30	N(1.5,0.2) [1.50,0.20]	1.71 (0.06)	1.60 (0.05)	1.79 (0.06)	1.50 (0.05)
ρ_y	0.05	N(0.4,0.2) [0.40, 0.20]	-0.03 (0.01)	-0.12 (0.03)	0.01 (0.00)	-0.04 (0.01)
ζ_p	0.80	B(6,6) [0.50, 0.14]	0.83 (0.03)	0.82 (0.03)	0.80 (0.03)	0.93 (0.03)
ρ_χ	0.50	B(10,6) [0.71, 0.09]	0.61 (0.03)	0.33 (0.02)	0.62 (0.05)	0.61 (0.04)
ρ_z	0.80	B(10,6) [0.71, 0.09]	0.72 (0.04)	0.54 (0.04)	0.24 (0.03)	0.70 (0.03)
$\sigma_{\chi,c}$	1.11	Γ^{-1}(10,20) [0.0056, 0.0020]	0.14 (0.02)	0.18 (0.16)	0.21 (0.05)	0.23 (0.43)
σ_z	0.51	Γ^{-1}(10,20) [0.0056, 0.0020]	0.15 (0.03)	0.27 (0.04)	3.87 (0.42)	1.72 (0.22)
σ_v	0.12	Γ^{-1}(10,20) [0.0056, 0.0020]	0.03 (0.00)	0.03 (0.00)	0.03 (0.00)	0.03 (0.00)
σ_μ	20.64	Γ^{-1}(10,20) [0.0056, 0.0020]	7.31 (0.35)	4.90 (0.39)	4.96 (0.19)	5.77 (0.23)

Table 2: Parameters estimates obtained using different filters; the DGP features a preference shock with two components, a stationary AR(1) and a unit root. All variables are filtered prior to estimation. The sample size is T=150. Γ stands for the gamma distribution, B for the beta distribution and N for the normal distribution. In square brackets are the mean and the standard deviation of the prior.

the moments of their true cyclical component, when T=1000 - this sample size effectively reduces small sample biases to zero. Clearly, regardless of the DGP, the variability, the serial and the cross correlation properties of the cyclical component of output and inflation are distorted. Also, although output displays a linear trend under DGP1 and a unit root under DGP2, LT filtering in DGP1 and FOD filtering in DGP2 are as biased as other arbitrary filtering approaches. Thus, misspecification of the non-cyclical component can not be reason for these distortions. Finally, although DGP2 features a unit root, the raw inflation series is persistent but stationary. Hence, it will matter for structural estimation whether the model is fitted to filtered or unfiltered inflation.

To show how filtering errors affect parameter estimation, we take the experimental data for output, real wages, interest rates and inflation constructed with DGP2 and estimate the structural parameters prefiltering the raw data with LT, HP, BP and FOD filters. Estimation is conducted with Bayesian methods: we choose relatively loose priors for all the parameters and, to give the best chance to the routine, start estimation at the true parameter values. Posterior estimates are obtained with a random walk Metropolis algorithm, where the jumping variable has a t-distribution with 5 degrees of freedom and the variance is tuned to have an acceptance rate of about 30 percent for each filtering approach. Half a million draws were made in each case; convergence was checked with a standard CUMSUM statistic and achieved after less than 250000 iterations. We keep one out of hundred of the last 100,000 draws to compute posterior statistics. Results obtained experimenting with a flat prior are available on request from the authors.

Table 2 reports the median and the standard deviation of the posterior of each structural parameter when all observables are independently filtered prior to estimation. Appendix B contains estimates for other relevant cases and other DGPs. There are important estimation biases in all cases and the magnitude of the bias can exceed 100 percent for some parameter. Interestingly, the parameters regulating the relative magnitude of income and substitution effects (the Frisch elasticity σ_n^{-1}, the habit parameter h, the policy parameter ρ_π and persistence of the shocks) are considerably distorted. Estimates of the structural parameters appear to be relatively similar across three of the columns but this outcome depends on the features of the DGP, in particular, on whether the non-cyclical component is driven by technology or preference disturbances, on the relative variability of the non-cyclical shocks, and on whether all observables or only a portion of them is filtered prior to estimation (see appendix B).

While we have chosen to perform estimation using 150 data points to mimic a realistic estimation situation, larger samples will not change the conclusions. Thus, distortions obtain because of "population" rather than "small sample" errors. Similarly, allowing for measurement errors in the estimation will not change the features of table 2: the variability of the structural shocks is altered but the magnitude and the direction of the biases in the estimates of important structural parameters is unchanged (for both exercises, see Appendix B).

To understand why distortions occur, it is useful to plot the spectral density of the cyclical component of output and inflation (obtained by setting $\gamma = \sigma_{z,T} = 0$ for GDP1 or $\sigma_{\chi,T} = 0$ for DGP2 in the simulations) together with the spectral density of the four filtered data when T=1000. If one filtering transformation recovers the true cyclical component, the difference between the two spectra will be zero at all frequencies. Imperfect isolation in certain frequency bands will be evident when the two spectra differ considerably in those bands. To facilitate the discussion, we divide the spectrum into low, business cycle, and high frequencies and, in figure 1, separate the frequencies corresponding to cycles of 8-32 quarters from the others with two vertical bars.

Two observations are immediate. First, the cyclical component produced by a DSGE model does not have power only at the so-called business cycle frequencies - in fact, its spectrum resembles the one of an AR(1) process . For the standard shock processes we have used, about half of the variability of the series is located at frequencies corresponding to cycles larger than 32 quarters. Thus, the idea that a statistical filter defines what is relevant for the analysis is incompatible with the assumption that a class of stationary DSGE models has generated the data. Moreover,

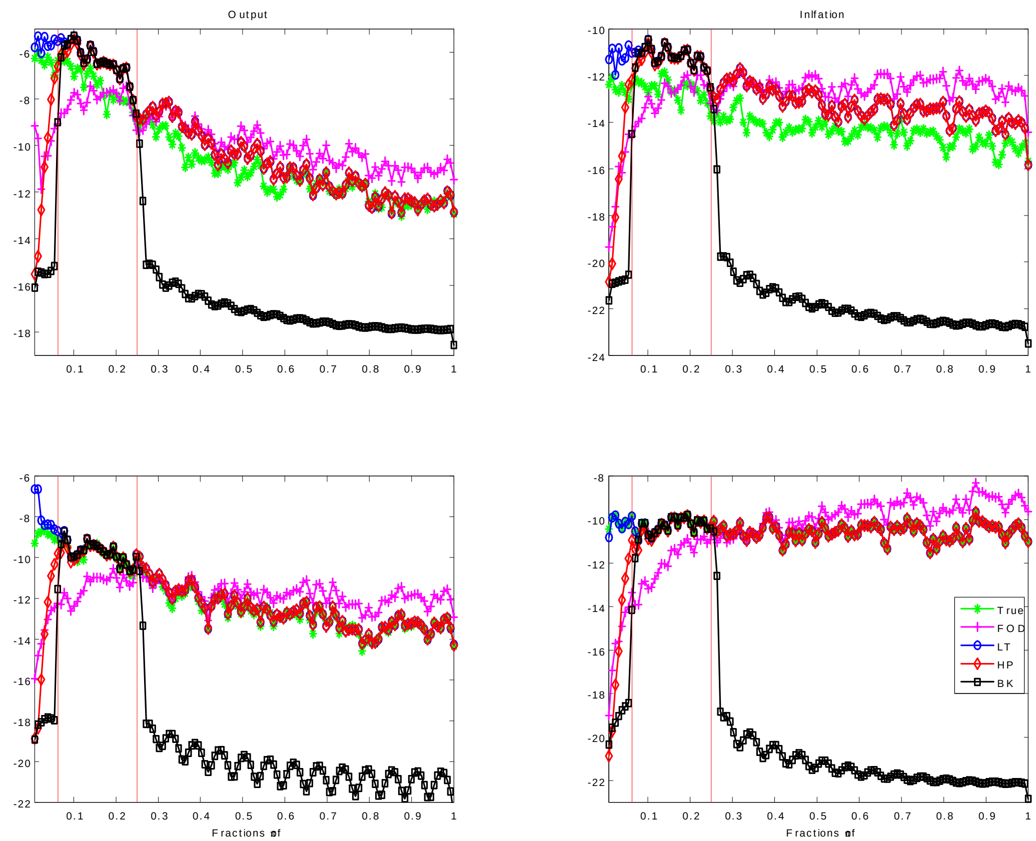

Figure 1: Log spectrum: true and estimated cyclical components. Top panel DGP1; bottom panel DGP2.

focusing on business cycle frequencies is restrictive and may bias the interpretation of the economic phenomena. Second, even with 1000 data points, all filters imperfectly capture the spectrum of the true cyclical component of output and inflation. More importantly, regardless of the DGP, the filtering error is not only located in the high frequencies and its frequency distribution is somewhat filter dependent. For example, LT filtered data has a stronger low frequency component and the other three filtered data a weaker low frequency component than the actual cyclical data. At business cycles frequencies, the cyclical component extracted with HP, BP and LT filters overestimates the true cyclical component of both variables while FOD filtered data grossly underestimates the variability of the true cyclical component.

Why are errors present? The statistical filters we consider look like "high pass" or "band pass" filters. Thus, they appropriately extract the cyclical component of the data if and only if the non-cyclical component of the model is solely located at those frequencies suppressed by the filters and the cyclical component is entirely located at the frequencies where the gain function of the filter is unity. Given that the cyclical component generated by a (log-linear) DSGE model will typically have power at all frequencies of the spectrum, filtering errors are created. In particular, since all the filters but LT attribute the power in low frequencies to the non-cyclical component, important downward distortions are created at these frequencies. For the LT filter instead, upward distortions are produced because the stochastic elements of the non-cyclical component are disregarded. Mismeasurement of the low frequencies portion of the cyclical fluctuations is particularly troublesome because the estimated income and substitution effects are different from the true income and substitution effects and this affects structural parameter estimates.

Knowing the DGP of the data is not a precondition for the above argument to hold. The (log-linear) solution of a stationary DSGE model has either a AR or an ARMA representation (depending on whether all or a subset of the endogenous variables is considered), regardless of its exact structure. If the shocks are persistent, as it is usually assumed, it will always be the case that the data simulated by the stationary solution will have power in the low frequencies of the spectrum . Furthermore, the proportion of the variability in those frequencies is an increasing function of the persistence of the shocks. Our point is also completely independent of the assumed process driving the non-cyclical component and of its exact location (compare, e.g., the plots obtained with DGP1 and DGP2).

It is common among practitioners to believe that different filters are simply different ways to capture what generates the non-cyclical component of the data. This perception is, in general, incorrect. The choice of filter has also implications for what we believe the cyclical component is. Incorrect filtering distorts both components and, as the discussion following table 1 demonstrates, misspecification of the cyclical component may have more severe consequences than misspecification of the non-cyclical component.

3 An ideal filter for DSGE models

To eliminate the distortions induced by imperfect filtering one should design a filter exploiting the information that the cyclical components of a DSGE has the features of an AR (ARMA) process.

For this purpose, suppose a time series y_t has two components: c_t, which carries relevant information about the parameters of the model, and T_t, a nuisance component, and suppose for simplicity that T_t and c_t are uncorrelated (which, in our context, means that they are driven by independent shocks). Suppose we have available a time invariant linear filter $g(L)$ and let $y_t^f = g(L)y_t$ be the filtered series. Under what conditions would $y_t^f = c_t$? For this to happen, we need $g(L)T_t = 0$ and $g(L)c_t = c_t$. In frequency domain, these two conditions imply that $g(\omega)(S_T(\omega) + S_c(\omega)) = S_c(\omega)$, where $S_i(\omega)$ is the spectral density of $i = T, c$ at frequency ω and $g(\omega)$ is the square modulus of the transfer function of $g(L)$. Then, if $0 \leq g^*(\omega) = \frac{S_c(\omega)}{S_T(\omega)+S_c(\omega)} \leq 1$, $y_t^f = c_t$. Thus, $g^*(\omega)$ needs to be large (small) at the frequencies where $S_c(\omega)$ is large (small).

In time series analysis, it is typical to assume that c_t has power only certain frequencies, say, $S_c(\omega) \neq 0,\ \forall\, \omega \in (\omega_1, \omega_2)$ and T_t has power at other frequencies, so that $S_T(\omega) = 0,\ \forall\, \omega \in (\omega_1, \omega_2)$. In this case, a band pass filter $g^*(\omega) = 1, \forall \omega \in (\omega_1, \omega_2)$ and $g^*(\omega) = 0$ otherwise, will make $y_t^f = c_t$. However, if c_t has also power for $\omega \in (0, \omega_1)$, $g^*(\omega) \neq 0,\ \forall\, \omega \in (0, \omega_1)$, a band pass filter fails to recover the true cyclical component.

As discussed, in log-linearized DSGE models c_t has, roughly, the structure of a persistent AR (or a persistent ARMA) process, meaning that $S_c(\omega) \neq 0, \forall \omega$ and $\frac{\partial S_c(\omega)}{\partial \omega} < 0, \forall \omega$ (or $\frac{\partial S_c(\omega)}{\partial \omega} \leq 0$ for $\omega < \omega_1$ and $\frac{\partial S_c(\omega)}{\partial \omega} \geq 0$ for $\omega \geq \omega_1$). Hence, regardless of the exact structure of T_t, band pass or high pass filters (such as the HP or FOD filters) will induce measurement error at some or all frequencies. High frequency measurement error affects the standard errors of the estimates but, in general, will not change the properties of point estimates (compare, for example, BP and HP estimates in table 2). On the other hand, low frequency measurement errors are problematic.

Since $S_c(\omega) \neq 0, \forall \omega$ and $\frac{\partial S_c(\omega)}{\partial \omega} < 0, \forall \omega$, the ideal filter for a DSGE model must be such that $\log g^*(\omega)$ is increasing in ω, never vanishes over $(0, \pi)$ and approaches one only for $\omega = \pi$. If it is unique, it can be calculated with an iterative approach which we summarize next:

Algorithm 3.1 **1.** *Choose a θ^0 vector of structural parameters and compute $c_t(\theta^0)$ using the model.*

2. *Given an observable y_t, obtain $g(\omega)(\theta^0) = \frac{S_c(\omega,\theta^0)}{S_y(\omega)}$ and compute $y_t^f(\theta^0) = g(L)(\theta^0)y_t$ using the resulting $g(L)$ filter.*

3. *Use $y_t^f(\theta^0)$ to estimate the parameters of the model. Call the estimated vector θ^1.*

4. *Iterate on steps 1.-3. until $||g(L)(\theta^i)y_t - c_t(\theta^i)|| < \varepsilon$ or $||\theta^i - \theta^{i-1}|| < \varepsilon$, or both, $i = 2, \ldots$.*

where the metric in step 4. is chosen by the investigator and may be frequency specific. Under the assumption that the data have been generated by a Markov process which is irreducible, aperiodic and Harris recurrent, and that the metric used is the total variation norm over frequencies, adaptation of the results of Tierney (1994) will insure that convergence occurs as the number of iterations becomes large.

A few points about the algorithm are worth emphasizing. First, the optimal $g^*(\omega)$ does not necessarily generate a one-sided $g(L)$, nor weights g_j decaying fast to zero. Therefore, practical issues concerning alteration of the timing of the information and truncation need to be address. Second, the iterative procedure is time consuming since the model needs to be estimated numerous times before the fixed point is found. Given the computational costs of estimating the parameters of DSGE models by full information methods, this iterative approach is unfeasible on current computers. Finally, the ideal filter is model specific - it depends on what shocks drive c_t, their time series structure and the features of the internal propagation mechanism of shocks - and it is subject to standard specification errors if the cyclical component is misspecified.

Given these difficulties, rather than trying to construct the ideal filter for a particular model, we prefer to take another route to improve the quality of the estimates of the structural parameters of a cyclical DSGE, which is not model specific and is computationally feasible. Our idea is to use the information contained in the cyclical data generated by a number of filters in the estimation. In practice, we treat cyclical data extracted with various filtering methods as contaminated estimates of the unobservable model-based cyclical component and use the information provided by a carefully selected list of filters jointly in the estimation of the structural parameters. If the measurement error is close to be idiosyncratic across filtering methods in the low frequencies, our signal extraction approach will average it out. Thus, we obtain more precise estimates of the cyclical features of the economy and, hopefully, better estimates of the structural parameters are obtained.

How do we obtained improved estimates of the structural parameters? Let $g_i(L),\ i = 1, 2, \ldots, q$ to be different filters and y_t^i the resulting filtered data. Let $S_i(\omega), i = 1, 2, \ldots, q$ be the spectral density of the filtered data y_t^i and assume that $S_i(\omega) = S_c(\omega) + S_{u_i}(\omega)$. Then $\sum_i \varsigma_i(\omega) S_{u_i}(\omega) = 0$ for some set of weights $\varsigma_i(\omega)$ as long as $S_{u_i}(\omega)$ are sufficiently idiosyncratic across i. For weights $\varsigma_i(\omega)$ which are independent of ω, it may not be possible to set $\sum_i \varsigma_i S_{u_i}(\omega) = 0$ at all ω and one can choose them to reduce, e.g., low frequency measurement error. Note that the use of cross sectional information to identify the model-based cyclical component allows the u_i to be serially correlated. Given that q is finite here, we will not allow cross filter correlation in the u_i and this requires a careful selection of filtering methods to be used in the estimation.

4 An alternative framework

Let the log-linearized solution of a cyclical DSGE model be:

$$x_t = \Phi(\theta)x_{t-1} + \Psi(\theta)e_t \quad e_t \sim (0, \Sigma(\theta)) \tag{10}$$

where Ψ, Φ are time invariant functions of the vector of structural parameters $\theta = (\theta_1, \ldots, \theta_k)$, x_t are the endogenous variables and e_t the structural innovations. We let $x_t^m = Sx_t$, be a $n \times 1$ vector where S is a selection matrix picking the variables which are observable and interesting from the point of view of the analysis.

Let x_{it} be a vector of size $n \times 1$ of observable time series filtered with method $i = 1, 2, ...q$, and let $x_t = [x'_{1t}, x'_{2t}, \ldots, x'_{qt}]'$. Assume that the filtered observables are linked to the true cyclical component with the following structure:

$$x_t = \nu_0 + \nu_1 x_t^m + u_t \quad u_t \sim (0, \Sigma_u) \tag{11}$$

where $\nu_0 = [\nu_0^1, \nu_0^2, \ldots, \nu_0^q]'$ is a $nq \times 1$ vector of constants, $\nu_1 = [\nu_1^1, \nu_1^2, \ldots, \nu_1^q]'$ a $nq \times n$ matrix of non-structural parameters, ν_1^i is a $n \times n$ diagonal matrix each i, and $u_t = [u'_{1t}, u'_{2t}, \ldots, u'_{qt}]$ is a $nq \times 1$ vector of possibly serially correlated errors. For estimation purposes, we normalize $\nu_1^1 = I$.

Joint estimation of the structural parameters θ and the non-structural parameters (ν_0, ν_1, Σ_u) is now possible because (10) and (11) represent a state space system with the latter being a measurement equation and the former state equations. Thus, the likelihood of (10) and (11) can be computed with the Kalman filter. If Bayesian estimation is preferred, the posterior distribution for the parameters can be obtained with Monte Carlo Markov Chain simulators (see e.g. Canova (2007)). Note that identification of x_t^m is obtained from the cross section of filters under the conditions stated in Forni et al (2000).

In (11) different cyclical estimates x_{it} are treated as contaminated proxies of the true cyclical component x_t^m. They are contaminated because they alter the power spectrum of the true cyclical component at some or all frequencies. The information they contain for the model relevant concepts of cyclical fluctuations is measured by ν_0 and ν_1. Ideally, ν_0 is a vector of zeros and ν_1 a matrix with the identity in each $n \times n$ block, so that each set of filtered data is an unbiased and perfectly correlated although noisy signal of the true cyclical component. In general, we expect either $\nu_0 \neq 0$ or $\nu_1^i \neq I, i \neq 1$, or both, for some or all i . Since $\nu_1^1 = I$, estimates of ν_1^i for $i \neq 1$ give us the idea of the amount of correlation distortions each method displays relative to the first.

While we think of (10) and (11) as a way to correct for filtering biases, one could also think of our setup as a factor model, where the model concept of cyclical fluctuations is defined as the common factor to the noisy indicators produced by the various filters - we thank a referee for suggesting such an interpretation. This idea is appealing but disregards the information that the cyclical component of a DSGE model has a particular structure.

The signal extraction setup we use is advantageous in, at least, two respects. First, since we do not have to arbitrarily choose one filtering approach prior to the estimation or select which shock drives the non-cyclical component, we avoid specification errors. Second, our approach can use as observables the output of one-sided and two-sided filters, both of univariate and multivariate nature and of filters which assume that cyclical and non-cyclical components are correlated or not, as long as the list of filters is sufficiently rich.

We stress that our analysis is conditional on two important assumptions. First, we assume that the model generating x_t is correctly specified; that is, there are no missing variables or shocks. When this is not the case, the interpretation of the ν's becomes more difficult and there is no guaranteed that our signal extraction approach has better properties than any of the standard approaches. Second, we assume that the cyclical and the non-cyclical components are theoretically uncorrelated. While this simplifying assumption is common in the literature, the presence of a correlation among components adds misspecification and biases which are neglected in this paper.

4.1 Selecting the filters to be used in the estimation

We have mentioned that we need vectors of filtered data which are sufficiently idiosyncratic in their low frequency distortions. Hence, knowledge of features of various filters is necessary to create a list which effectively averages out the low frequency measurement errors induced by imperfect filtering.

We have also mentioned that, apart from LT, standard filters resemble high pass filters and thus tend to underestimate the low frequency contribution of the cyclical component. Therefore, it is important to use in the estimation filters which overestimate the low frequency contribution of the cyclical components. One class of filters with such a property is the cumulative operator $(1+L)^j, j = 1, 2, \ldots$. Notice that for $j = 1$, this filter has a square gain function which is the mirror image of the FOD filter. Low pass filters can also be considered - as long as the zero frequency is properly accounted for, for example, by requiring that the sum of the filter weights is zero. One can also consider Butterworth filters, where the two free parameters are chosen to let interesting frequencies (say, from 8 to 100 quarters) be passed with minor changes.

4.2 The relationship with the literature

The literature is largely silent about the issues we address in this paper. Cogley (2001) and Gorodnichenko and Ng (2010) are concerned with the problem of estimating the structural parameters of a cyclical DSGE when the trend specification is incorrect, but do not investigate what are the consequences that imperfect filtering has on the properties of the cyclical component nor their implications for structural estimates. Giannone et al. (2006) emphasize that if model variables are measured with error, the solution has a natural factor structure and exploit this feature to compare VAR and factor models impulse responses. Rather than considering a factor structure for the endogenous variables in terms of the states, we construct an estimable structure where vectors of filtered observable data have a factor structure in terms of the variables of the model. However, as in Giannone et al., we emphasize that low frequency measurement error may exist. Canova (2008) suggests to estimate cyclical DSGE models by specifying a flexible link between the model and the raw data - the approach is designed to deal with different sources of misspecification than those considered here. Ferroni (2009) provides a one-step approach which allows to test trend specifications. The paper closest in spirit to ours is Boivin and Giannoni (2005). Their main point is that the model variables do not have an exact observable counterpart and that some indicators external to the model may have important information for model variables. The idea here is somewhat similar. The cyclical component of the model does not have an exact counterpart in the data because none of the existing filters captures the time series features of the cyclical component produced by a DSGE. If different cyclical vectors have idiosyncratic error components, this error may be averaged out with our approach.

Commentators have noticed that the procedure resembles Bayesian averaging of outcomes. Two main differences set our approach apart from this procedure. First, in Bayesian averaging the weights are the posterior probabilities of each model, while here they capture the amount of information contained in the filtered data for the model based concept of cyclical fluctuations. Second, in Bayesian averaging the data is the same but the models are different. Here, there is a single model, but the data used to estimate it is different. Finally, our approach has the same flavour of multivariate unobservable component filtering (see e.g. Canova (2007)). The extraction problem applies here to vectors of filtered data rather than to a vector of raw data.

4.3 How does the procedure fare with simulated data?

To show the properties of our approach and to highlight the practical importance of appropriately choosing the list of filters, we estimate the structural parameters of the model of section 2 using the experimental data produced by DGP2. As input in our procedure, we employ either LT, HP and FOD filtered data (Factor 1) or HP and BP filtered data - in this case we use two smoothing constants $\lambda = 1600$ and $\lambda = 6400$ (Factor 2). As shown in figure 1, LT, HP and FOD filtered data display significant low frequency differences, while HP and BP filtered data have similar low frequency components. Thus, we expect a reduction of the parameter distortions in the first but not in the second case. Since the list of filters is short, biased will not be wiped out but improvements in the quality of the estimates could be significant.

We employ the same Bayesian approach used in section 2, assuming the same priors on the structural parameters and loose priors on the non-structural parameters entering (11). In particular, we assume that each element of ν_0 is normally distributed, with mean zero with standard deviation equal to 0.5; the prior for the non-normalized elements of ν_1 is normal, centered at 1 with standard deviation 0.5: and the prior for the standard deviation of the u_t's is inverted gamma with mean 0.0037 and standard deviation 0.0002.

Because the data set is short, we present results obtained when constants and the loadings in (11) are common across series for each filter (in this case, there are 17 non-structural parameters). It makes sense to restrict the model this way because the distortions we emphasize are independent of the series (see e.g. output and inflation in table 1). In an earlier version of the paper, we had also performed unrestricted estimation (which implies 32 non-structural parameters to be estimated): the direction of the changes was similar but the quality of the estimates worsened.

Table 3 presents the median and the standard error of posterior for the structural parameters when all variables are filtered. Results for other specifications are in appendix B. In general, the biases we noted in table 2 are reduced with the Factor 1 specification but not with the Factor 2 specification. For example, the habit and the risk aversion parameters are better estimated, and the inflation coefficient in the Taylor rule much closer to the true value with Factor 1. The variability of the structural shocks is still poorly estimated but for reasons distinct from those discussed here (these parameters are weakly identified, regardless of cyclical data used).

		Factor 1	Factor 2
	True	Median (s.e.)	Median (s.e.)
σ_c	1.00	1.10 (0.10)	2.18 (0.37)
σ_n	0.70	0.49 (0.05)	0.11 (0.03)
h	0.70	0.74 (0.11)	0.62 (0.02)
ϵ	7.00	6.28 (0.11)	6.27 (0.06)
ρ_r	0.20	0.30 (0.07)	0.28 (0.02)
ρ_π	1.30	1.46 (0.03)	1.53 (0.05)
ρ_y	0.05	0.06 (0.01)	0.32 (0.05)
ζ_p	0.80	0.85 (0.03)	0.89 (0.01)
ρ_χ	0.50	0.53 (0.05)	0.65 (0.04)
ρ_z	0.80	0.66 (0.03)	0.62 (0.02)
σ_χ	1.10	0.23 (0.07)	0.21 (0.07)
σ_z	0.57	0.19 (0.04)	0.49 (0.10)
σ_v	0.12	0.09 (0.01)	0.09 (0.01)
σ_μ	20.64	5.21 (0.52)	4.44 (0.73)

Table 3: Posterior parameters estimates. Factor 1 uses LT, HP and FOD filtered data; factor 2 HP($\lambda = 1600$), HP($\lambda = 6400$) and BP filtered data. The DGP features a preference shock with two components, a stationary AR(1) and a unit root. All variables are filtered prior to estimation. The sample size is T=150.

To see what Factor 1 estimates imply in terms of economically meaningful statistics, we first compute the autocorrelation function of the cyclical components of output and inflation when the posterior median estimates of the parameters are used and compare them with the true autocorrelation function and the autocorrelation function obtained with LT and FOD approaches (see figure 2). For output, the autocorrelation function obtained with our specifications is very close to the true one and it is different from the one obtained, for example, with the FOD filter. For inflation, the match is good but differences with standard methods are less dramatic, primarily because true inflation persistence is low.

The good performance of our approach is reinforced when we look at the responses of the endogenous variables to the four structural shocks. Figure 3 presents the responses produced with the true parameters, those generated with the posterior median estimates obtained with our model and with LT and FOD filtered data.

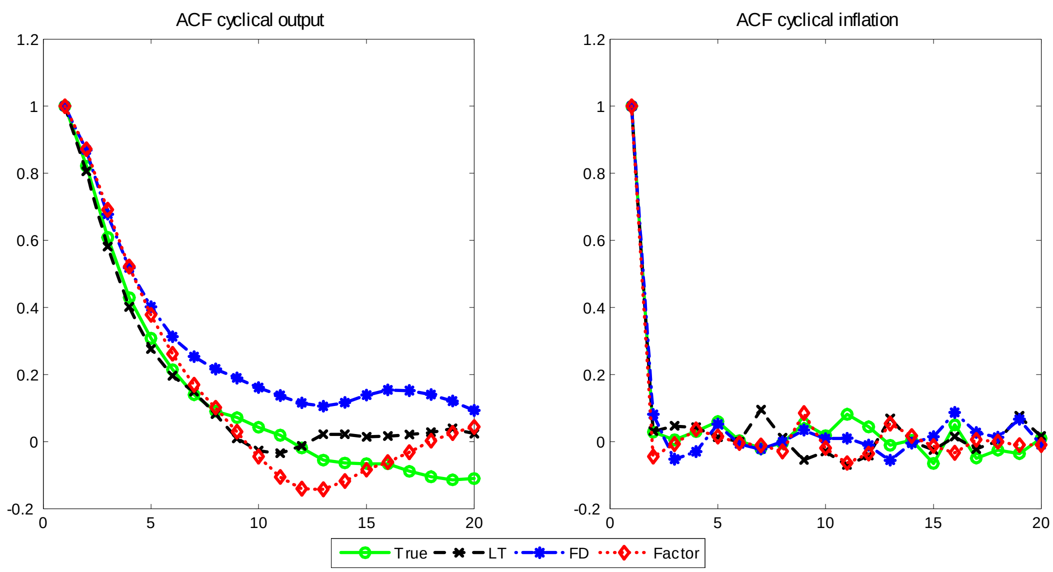

Figure 2: Autocorrelation functions of estimated and true cyclical components

Both the shape and the persistence of the conditional responses are reasonably captured by our setup. In addition, and contrary to what was happening with LT and FOD filters, the real wage response to technology has the right sign on impact. Finally, our estimates roughly replicate the magnitude of the responses to both preferences and technological disturbances while this is not the case with standard approaches.

Next, we examine the out-of-sample performance of our setup relative to traditional ones. We conduct two types of forecasting exercises. In the first, we compute the sequence of one step ahead forecast errors for output and inflation, when we take as parameter values the posterior median estimates, setting all the shocks in the forecasting period to zero. The MSE is computed over 150 forecasting periods, with no parameter updating in the forecasting sample, and appears in table 4.

Series	LT	FOD	Factor 1
Output	0.006	0.003	0.001
Inflation	0.030	0.031	0.029

Table 4: Mean square error of the unconditional forecasts; simulated data; scale 10^{-2}.

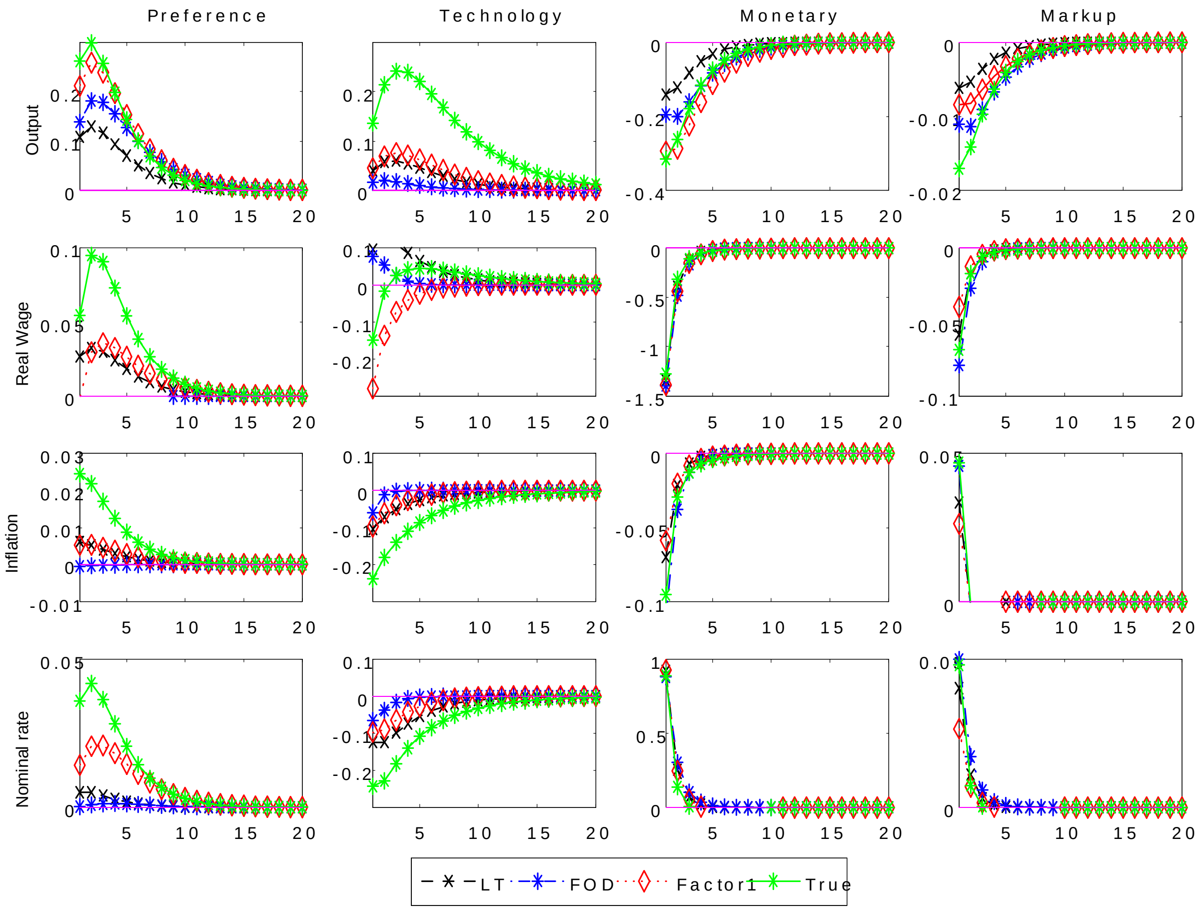

Figure 3: Impulse responses to shocks

Figure 4 traces out the one-step ahead path of cyclical output and cyclical inflation that would obtain with posterior median estimates when monetary shocks were drawn so as to keep the nominal interest rate fixed over the forecasting path - a standard assumption in policy projections. That is, we allow the nominal interest rate to endogenously react to output and inflation but make sure that the monetary shocks are such that the nominal rate is constant over the forecasting path and equal to the value taken prior to the forecasting period (time 0 in the figure).

Overall, our specification is superior to single filtering approaches in unconditionally forecasting one-step ahead cyclical output and cyclical inflation and for output, the reduction in MSE is considerable. Our specification does well also in conditional forecasting. The counterfactual path for output our specification produces is very close to the true one at all horizons and practically eliminates the systematic bias that LT and FOD filters generate. For inflation, the counterfactual path produced by our model is similar to the true path; it is significantly better than the one obtained with FOD estimates, but roughly comparable to the one produced by LT estimates.

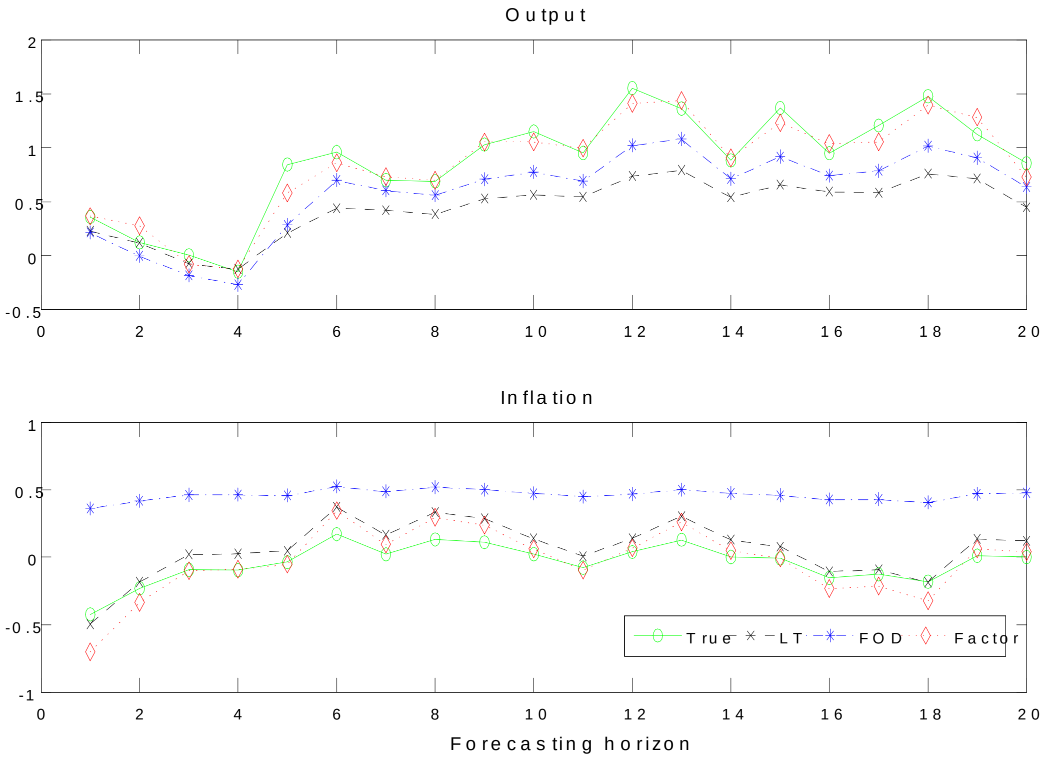

Figure 4: One step ahead forecasts, conditional on a constant interest rate path.

Since these conclusions hold also for alternative DGPs and combinations of filtered and unfiltered observables, the specification is effective in reducing low frequency measurement errors and can provide a more reliable picture of the cyclicality of the variables of interest.

5 Does money matter in transmitting monetary business cycles?

To show that our procedure may be relevant for understanding important economic phenomena, we reconsider the role of money in transmitting monetary business cycles. The majority of the monetary models nowadays used in the policy and academic literature attributes a minimal role to the stock of money. In most cases these models make no reference whatsoever to monetary aggregates, and when they do, they use a specification where a money demand function determines how much money needs to be supplied, given predetermined levels of output, inflation and the nominal rate. Ireland (2004) has constructed a specification in the class of New Keynesian models where real balances may have influence the dynamics of output and inflation. He estimated the relevant parameters by likelihood techniques using post 1980 US data and found that current theoretical practices are, by and large, appropriate. To construct the likelihood of his cyclical model, Ireland takes away a linear trend from per-capita GDP and per-capita real balances and demean inflation and the nominal interest rate. Here, we repeat Ireland's exercise using a number of filtering procedures.

5.1 The model economy

Since the economy is quite standard, we only briefly describe its features. At each t the representative household maximizes

$$E_t \sum_t \beta^t \chi_t [U(c_t, \frac{M_t}{p_t e_t}) - \eta n_t] \tag{12}$$

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

$$M_{t-1} + Tr_t + B_{t-1} + W_t n_t + D_t = p_t c_t + \frac{B_t}{R_t} + M_t \tag{13}$$

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

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

$$y_t^i = (\frac{p_t^i}{p_t})^{-\epsilon} y_t \tag{14}$$

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

The intermediate good producing firm $i \in [0,1]$, hires n_t^i units of labor from the representative household to produce y_t^i units of intermediate good i using the production function $y_t^i = z_t n_t^i$, where z_t is an aggregate productivity shock. Since intermediate goods substitute imperfectly for one another in producing finished goods, intermediate firms can set the price of their good but must satisfy (14) at the chosen price. We assume a quadratic cost in adjusting prices, measured in finished goods, given by $\frac{\phi}{2}(\frac{p_t^i}{\pi^s p_{t-1}^i}-1)^2 y_t$ where $\phi > 0$ and π^s measures steady state inflation. Optimal prices are chosen to maximize

$$E\sum_t \beta^t \chi_t [U_1(c_t, \frac{M_t}{p_t e_t})](\frac{D_t^i}{p_t}) \tag{15}$$

subject to (14), where $\beta^t \chi_t U_1(c_t, \frac{M_t}{p_t e_t})$ measures the marginal value to the household of an additional unit of profits t and real dividends are $\frac{D_t^i}{p_t} = (\frac{p_t^i}{p_t})^{1-\epsilon} y_t - (\frac{p_t^i}{p_t})^{-\epsilon}(\frac{w_t y_t}{z_t}) - \frac{\phi}{2}(\frac{p_t^i}{\pi p_{t-1}^i}-1)^2 y_t$.

The monetary authority sets the nominal interest rate according to

$$R_t = R_{t-1}^{\rho_r} y_{t-1}^{(1-\rho_r)\rho_y} \pi_{t-1}^{(1-\rho_r)\rho_\pi} \Delta M_t^{(1-\rho_r)\rho_m} v_t \tag{16}$$

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

The law of motion of the disturbances $d_t = (\chi_t, e_t, z_t, v_t)$ is $\log d_t = \bar{d} + H \log d_{t-1} + \iota_t$, where H is diagonal with entries $\rho_\chi, \rho_e, \rho_z, 0$, respectively. The covariance matrix of the structural shocks Σ is diagonal with entries $\sigma_\chi^2, \sigma_e^2, \sigma_z^2, \sigma_v^2$. In a symmetric equilibrium $y_t^i = y_t, n_t^i = n_t, p_t^i = p_t, D_t^i = D_t$. Log-linearizing the model around the steady state produces the following equilibrium conditions:

$$\hat{y}_t = E_t\hat{y}_{t+1} - \omega_1((\hat{R}_t - E_t\hat{\pi}_{t+1}) - (\hat{\chi}_t - E_t\hat{\chi}_{t+1})) + \omega_2((\hat{m}_t - \hat{e}_t) - (E_t\hat{m}_{t+1} - E_t\hat{e}_{t+1})) \tag{17}$$

$$\hat{m}_t = \gamma_1\hat{y}_y - \gamma_2\hat{R}_t + (1-(R^s-1)\gamma_2)\hat{e}_t \tag{18}$$

$$\hat{\pi}_t = \beta E_t\hat{\pi}_{t+1} + \psi(\frac{1}{\omega_1}\hat{y}_t - \frac{\omega_2}{\omega_1}(\hat{m}_t - \hat{e}_t) - \hat{z}_t) \tag{19}$$

$$\hat{R}_t = \rho_r\hat{R}_{t-1} + (1-\rho_r)\rho_y\hat{y}_{t-1} + (1-\rho_r)\rho_\pi\hat{\pi}_{t-1} + (1-\rho_r)\rho_m(\Delta\hat{m}_t + \hat{\pi}) + \hat{v}_t \tag{20}$$

where

$$\omega_1 = -\frac{U_1(c^s, \frac{ms}{e^s})}{y^s U_{11}(c^s, \frac{m^s}{p^s e^s})} \tag{21}$$

$$\omega_2 = -\frac{m^s}{e^s}\frac{U_{12}(c^s, \frac{m^s}{e^s})}{y^s U_{11}(c^s, \frac{m^s}{e^s})} \tag{22}$$

$$\gamma_1 = (R^s - 1 + \frac{y^s r^s \omega_2}{m^s})(\frac{\gamma_2}{\omega_1}) \tag{23}$$

$$\gamma_2 = \frac{R^s}{(R^s-1)(m^s/e^s)}(\frac{U_2(c^s, \frac{m^s}{e^s})}{(R^s-1)e^s U_{12}(c^s, \frac{m^s}{e^s}) - R^s U_{22}(c^s, \frac{m^s}{e^s})}) \tag{24}$$

$$\psi = \frac{\epsilon - 1}{\phi} \tag{25}$$

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

The log-linearized Euler condition (equation (17)) includes terms involving real money balances and the money demand shocks. They drop out if and only if utility is separable in consumption and real balances (see equation (22)). Similarly, real balances play a role in the forward looking Phillips curve (equation (19)), as long as $\omega_2 \neq 0$. Thus, real balances directly affect the determination of output and inflation if and only if real balances and consumption enter non-separably in the utility function. On the other hand, the posited policy rule implies that the growth rate of nominal balances may influence output and inflation indirectly, via interest rate determination. When $\omega_2 = \rho_m = 0$, real balances have no direct or indirect role in propagating cyclical fluctuations.

5.2 Estimation

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

We employ 8 procedures to extract the cyclical component of all variables. The first (POLY) fits a second order polynomial to each series separately, allowing for a change in the parameters at 1980:3. The cyclical component is the residual of the regression. The second transformation takes the first difference of each series (FOD) as an estimate of the cyclical component. The third and the fourth transformations are obtained with a HP filter and $\lambda = 1600$ or $\lambda = 128000$ - the latter leaves almost unchanged cycles with 2 to 100 quarters periodicity. The fifth transformation takes the first cumulant of all series as an estimate of the cyclical component (CUM). The sixth transformation is a multivariate version of the Beveridge and Nelson decomposition (MBN) which fits a VAR with 6 lags to the growth rate of the four variables and takes as an estimate of the cyclical component, the difference between the level of the variables and their estimated long run values. The seventh transformation is a classical decomposition (CD) which assumes an additive representation of the components, fits a linear trend to the log data and takes the residuals as the cyclical component. The last transformation employs an unobservable component (UC) decomposition which assumes that the non-cyclical component is a random walk and that the cyclical component has a trigonometric representation (see Canova (2007)). Since each series has an ARIMA(2,1,0) representation, the cyclical component is estimated with the projected values of an AR(2) regression of the growth rate of each variable.

We have selected these procedures to introduce as much cross-sectional idiosyncrasies in the vectors of observables as possible. In fact, in some procedures the non-cyclical component is quasi-deterministic (CD, POLY), in some it is very volatile (FOD, UC, MBN), and in some it is stochastic but smooth (HP); most decompositions use univariate and one (MBN) multivariate information; most imply that cyclical and non-cyclical components are independent and one that they are correlated (MBN). Finally, some are two-sided (such as the HP filters) and some one-sided (such as the MBN or UC filters). Note that, as far as low frequency distortions are concerned, CD, POLY, CUM and HP128000 are likely to overestimate the low frequency variability of the cyclical component, while the other four are likely to underestimate it.

We estimate the parameters of the model by Bayesian methods. The priors are in appendix C. The vector of observables is 32×1 (four series, 8 filtering methods) and the vector of states is 4×1. Since we set $\beta = 0.99$ and steady state inflation to 2 percent, there are 9 structural parameters $(\omega_1, \omega_2, \psi, \gamma_1, \gamma_2, \rho_r, \rho_p, \rho_y, \rho_m)$ - ϵ and ϕ are not separately identifiable - and seven auxiliary parameters $(\rho_\chi, \rho_e, \rho_z, \sigma_\chi, \sigma_e, \sigma_z, \sigma_v)$ to be estimated. We parameterize the link between the model and the cyclical data with one intercept and one slope per filter, independent of the series, but we allow the idiosyncratic term to be series and filter dependent. Thus, the intercept measures the average (across series and time) bias of each procedure and the slope the average correlation between the data produced by each method and the relevant model-based quantities. Since we normalize the slope of the first procedure, we have a total of 47 non-structural parameters to be estimated (8 intercepts, 7 slopes and 32 variances) [1].

We also estimate the structural parameters of interest using Ireland's original transformation, but allow for measurement error in each of the four equations - since our approach has an idiosyncratic error built, this is the relevant setup for comparison. For both specifications we draw 500,000 elements of a MCMC chain; convergence was achieved in less than 100,000 draws, and posterior statistics are computed using one every 100 of the last 200,000 draws.

[1] We have also experimented with specifications which leaves all the intercepts and all the slopes free or which restricts the variances of the idiosyncratic component to be either series specific (independent of the filtering method) or filter specific (independent of the series) but discarded them because the model fit was relatively poor.

5.3 The results

Before presenting estimates of the relevant parameters, we briefly comment on the estimates of the non-structural parameters we have obtained. First, the vector of ν_0 is estimated to be zero with very small standard errors - level biases appear to be absent. Since steady state information is not used in the estimation, the mean of the data may be different from the steady state of the model at the estimated parameters. The fact that this does not happen is encouraging from an estimation point of view. Second, the loadings ν_1^i vary from 0.70 (with UC filtered data) to 0.86 (with CD filtered data). Thus, all filtered series are highly correlated with the respective model quantities. Finally, standard errors for each series vary across filtering methods, confirming the presence of sufficient idiosyncratic information in the vector of cyclical data we employ.

Specification	Marginal log Likelihood	ω_2	ρ_m
Basic	16274	0.44 (0.02)	0.48 (0.02)
$\omega_2 = 0$	16237	0	0.96(0.01)
$\rho_m = 0$	16212	0.43(0.02)	0
$\omega_2 = 0$, $\rho_m = 0$	16220	0	0
Ireland		0.03 (0.02)	0.04 (0.03)

Table 5: Marginal log likelihood and posterior estimates.

Table 5 presents the marginal likelihood of the basic specification, where both the direct and the indirect effects are allowed for, and for three restricted specifications, where either the direct effect is eliminated ($\omega_2 = 0$), the indirect effect is eliminated ($\rho_m = 0$), or both are eliminated and the estimates of ω_2 and ρ_m obtained in the various cases. For comparison, we also report estimates obtained with Ireland's filtering specification. The full set of estimates is in appendix C.

A model where both the direct and the indirect effects of money are present is preferable in terms of in-sample fit. Furthermore, restricting both $\rho_m = 0$ and $\omega_2 = 0$ is preferable to restricting only $\rho_m = 0$. Posterior estimates confirm this conclusions: both parameters are tightly estimated, a-posteriori different from zero and indicate that money has a moderate influence on output and inflation fluctuations. Estimates obtained with just one filter, on the other hand, imply that both the direct and the indirect effects of money are statistically small and economically unimportant.

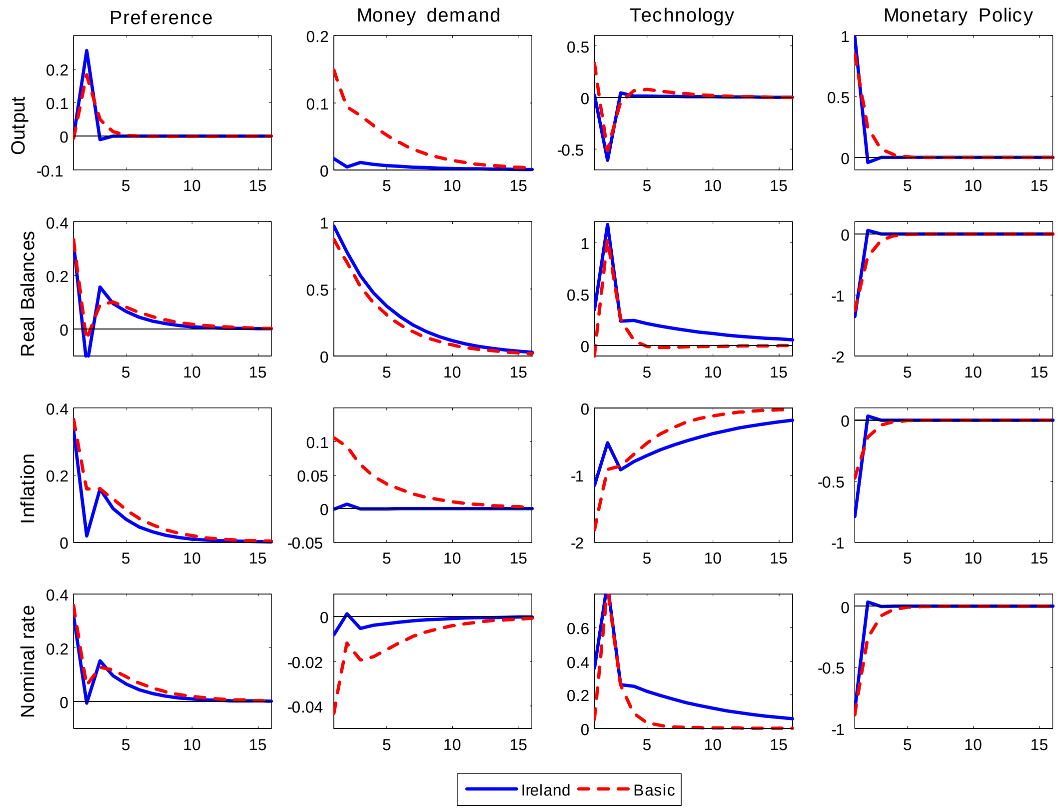

Figure 5: Impulse responses.

Figure 5 presents responses to unitary impulses in our basic specification and in Ireland's. Responses look qualitatively similar, but differences in the magnitude and the persistence of the responses to shocks are evident. In particular, when our approach is used, the persistence of the responses to technology shocks is reduced, and the responses to money demand shocks have different magnitude and persistence. Interestingly, both specifications produce a liquidity puzzle (expansionary monetary shocks decrease real balances rather than increasing them) and a price puzzle (expansionary monetary shocks decrease inflation rather than increasing it). We conjecture that with a more homogenous sample, say 1984-2008, both puzzles would disappear.

In sum, in our setup money plays a role in transmitting fluctuations to output and inflation while this is not the case when a standard single filtering approach is used. Since the list of filters we have used can average out low frequency measurement errors, the conclusions obtained with our approach appear to be more credible.

6 Conclusions

This paper has three parts. In the first, we show that standard filtering methods are unable to extract the cyclical component of a DSGE model and that measurement errors distorts estimates of the structural parameters. Biases obtain because a typical cyclical DSGE model produces time series with important low frequency components. These components are treated as non-cyclical by leading filtering approaches.

In the second part, we discuss how to construct a filter which takes into account the structure that a cyclical DSGE model imposes on the data. The derivation of this filter is theoretically straightforward, but it requires knowledge of the cyclical model generating the data. Furthermore, computational complexities make its implementation on existing computers unfeasible.

The third part proposes a method to estimate the structural parameters of a time invariant cyclical DSGE model which use multiple sources of cyclical information. The approach borrows ideas from the recent literature employing data-rich environments (see Boivin and Giannoni (2005)). We set up an estimation framework where the cyclical DSGE model is the unobservable factor; vectors of filtered data are contaminated observable proxies; and structural DSGE parameters are jointly estimated together with the non-structural parameters linking the model and the observables using signal extraction techniques.

Our approach is advantageous in, at least, two respects. Since we do not have to arbitrarily choose one filtering method prior to estimation, or select which shock drives the non-cyclical component, we avoid important specification errors. Moreover, our approach can be used with cyclical data obtained with one-sided and two-sided filters, of both univariate and multivariate nature, as long as the list of filters is sufficiently rich. When appropriate conditions are satisfied, low frequency errors can be averaged out making inference more reliable.

Using experimental data, we demonstrate that the biases obtained when just one filter is used are reduced, that the unconditional one step ahead mean square error (MSE) produced by our approach is smaller than the MSE obtained with a standard procedure and that conditional forecasts are better behaved. To show that the biases are also economically relevant, we revisit the role of money in transmitting monetary business cycles. We show that when the output of multiple filters is jointly used in the estimation, money balances statistically matter for the transmission of cyclical fluctuations to output and inflation and that the propagation of primitive shocks differs.

We want to reiterate two points which make alternatives to the procedure we present unpalatable. First, although nowadays popular, the approach of using model-based transformation to fit cyclical models is as problematic as any statistical filtering approach. Specification errors are likely to be important. Moreover, since we can solve models only when non-cyclical shocks affect the technology, the consumption/investment transformation frontier or preferences (see Chang, et al. (2007)), computational rather than economic considerations may drive model-based filtering. Thus, although some form of consistency between the model and the data is imposed, a great deal of arbitrariness is also present with this approach.

Second, the more appealing approach of employing (time varying) models to jointly explain the cyclical and the non-cyclical properties of the data is currently unfeasible. Many reasons make such a research program difficult to pursue. First, jointly modelling cyclical and non-cyclical fluctuations poses important theoretical challenges: there are few known mechanisms which are able to propagate temporary shocks for a long period of time (we need, for example, R&D, as in Comin and Gertler (2006) or Schumpeterian creative destruction, as in Lopez-Salido and Michelacci (2007)) or create important cyclical implications from long run disturbances. Second, to jointly account for both types of fluctuations we need to measure the features of non-cyclical dynamics. Relatively short reliable time series and breaks of various sorts make the data largely uninformative about these features. Third, although some progress in this respect has been reported by Fernandez Villaverde and Rubio Ramirez (2007), time varying structures are difficult to deal with in theory and hard to handle computationally.

Given these problems, this paper provides a simple setup where specification and measurement error biases could be reduced. In this sense, the paper constitutes a step forward in improving the reliability of inferential exercises in DSGE models.

References

Boivin, J. and Giannoni, M., (2005), DSGE estimation with data rich environments, University of Montreal, manuscript.

Canova, F., (1998), Detrending and Business Cycle Facts, *Journal of Monetary Economics*, 41, 475-512.

Canova, F., (2007), *Methods for Applied Macroeconomic Research*, Princeton University Press, Princeton, N.J.

Canova, F. , (2008), Bridging cyclical DSGE models and the raw data, UPF manuscript.

Canova, F., (2009), How much structure in empirical models?, in Mills, T. and Patterson, K. (eds.), *Palgrave Handbook of Applied Econometrics*, Palgrave Macmillan.

Canova, F. and Sala, L., (2009), Back to square one: identification issues in DSGE models, *Journal of Monetary Economics*, 56, 431-449.

Chang, Y., Doh, T. and Schorfheide, F. (2007), Non stationary hours in a DSGE model, *Journal of Money Credit and Banking* , 39, 1357-1373.

Cogley, T., (2001), Estimating and testing Rational Expectations Models When the Trend Specification is Uncertain, *Journal of Economic Dynamics and Control*, 25, 1485-1525.

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

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

Fernandez Villaverde, J. and Rubio Ramirez, J., (2007), How structural are structural parameter values?, *NBER Macroeconomic Annual*, 83-137.

Forni, M., Hallin, M., Lippi, M. and Reichlin, L., (2000), The Generalized Dynamic Factor Model: Identification and Estimation, *The Review of Economic and Statistics*, 82, 540-554.

Ferroni, F., (2009), Trend agnostic one step estimation of DSGE models. MPRA paper 14550, Library of Munich, Germany.

Fukac, M. and Pagan, A., (2010), Limited Information Estimation and Evaluation of DSGE Models, *Journal of Applied Econometrics*, 25, 55-70.

Gali, J. (2008), *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*, Princeton University Press, N.J.

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

Gorodnichenko, V. and Ng, S., (2010), Estimation of DSGE models when data are persistent, *Journal of Monetary Economics*, 57, 325-340..

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

Justiniano, A., Primiceri, G. and Tambalotti, A., (2010), Investment shocks and the relative price of investments, forthcoming, *Review of Economic Dynamics*.

Lopez Salido, and Michelacci, C.(2007) Technology shocks and Job flows, *Review of Economic*

Studies, 74, 1195-1227.

Smets, F. and Wouters, R., (2003), An estimated Dynamic Stochastic General Equilibrium models of the Euro area, *Journal of the European Economic Association*, I, 1123-1175.

Smets, F. and Wouters, R., (2007), Shocks and Frictions in the US economy, *American Economic Review*, 97, 586-606.

Tierney, L.(1994), Markov chains for exploring posterior distributions, *Annals of Statistics*, 22, 1702-1762.

Appendix A: the NK model

The model we use is a version of a textbook New Keynesian model (e.g. Gali, 2008) with a few exceptions. We assume habit in consumption, a preference shock and, as in Smets an Wouters (2003, 2007), we assume that the elasticity of variety of goods is an exogenous stochastic process.

Households

The representative household prefers to consume a variety of goods: the consumption basket is

$$C_t = \left(\int_0^1 C_t(j)^{\frac{\epsilon_t - 1}{\epsilon_t}} dj \right)^{\frac{\epsilon_t}{\epsilon_t - 1}} \tag{26}$$

where $C_t(j)$ is the consumption of the good j. Maximization with respect to $C_t(j)$, for a given total expenditure, leads to a set of demand function of the type

$$C_t(j) = \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon_t} C_t \tag{27}$$

where $P_t(j)$ is the price of the good j. We let

$$\epsilon_t = \epsilon \exp \frac{1 - \epsilon}{\epsilon} \mu_t \qquad \epsilon > 1 \tag{28}$$

where μ_t is a i.i.d. normal shock. The appropriate price deflator for the consumption basket is

$$P_t = \left(\int_0^1 P_t(j)^{1-\epsilon_t} dj \right)^{\frac{1}{1-\epsilon_t}} \tag{29}$$

Conditional on the optimal consumer behavior, $P_t C_t = [\int_0^1 P_t(j) C_t(j) dj]$. The representative household chooses sequences for consumption, savings and leisure to maximize

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[\chi_t \frac{(C_t - hC_{t-1})^{1-\sigma_c}}{1 - \sigma_c} - \frac{N_t^{1+\sigma_n}}{1 + \sigma_n} \right] \tag{30}$$

where χ_t is an exogenous demand shifter. Household maximization is subject to the sequence of budget constraints:

$$P_t C_t + b_t B_t = B_{t-1} + W_t N_t \tag{31}$$

Thus, the household holds its financial wealth in the form of one period bonds, B_t with price b_t; W_t is the nominal wage and N_t is hours worked. The first order conditions of the problem are:

$$0 = \chi_t (C_t - hC_{t-1})^{-\sigma_c} - \mathcal{L}_t \tag{32}$$

$$0 = -N_t^{\sigma_n} + \mathcal{L}_t \frac{W_t}{P_t} \tag{33}$$

$$1 = E_t \left[\beta \frac{\mathcal{L}_{t+1}}{\mathcal{L}_t} \frac{P_t}{P_{t+1}} R_t \right] = E_t \left[\beta \frac{\mathcal{L}_{t+1}}{\mathcal{L}_t} \frac{R_t}{\Pi_{t+1}} \right] \tag{34}$$

where $\mathcal{L}_t$ is the Lagrangian multiplier associated to the budget constraint and R_t is the gross nominal rate of return on bonds ($R_t = 1 + r_t = 1/b_t$). In the non stochastic steady states:

$$w = W/P = N^{\sigma_n}(C - hC)^{\sigma_c}$$

$$1 = \beta R/\Pi$$

Firms

There is a continuum of firms, indexed by $j \in [0, 1]$, producing a differentiated good. They face the same technology:

$$Y_t(j) = Z_t N_t(j)^{1-\alpha} \tag{35}$$

where Z_t is an exogenous technology process. Firms pay a nominal wage W_t for every hour worked to the household. Following Calvo (1983), each firm may reset its price with probability $1 - \zeta_p$ in any given period, independently of time elapsed since last adjustment. Thus, a fraction $(1 - \zeta_p)$ chooses the price that maximizes nominal profits subject to a demand schedule [2], that is

$$\max_{P_t(j)} Pr_t = \max P_t(j)Y_t(j) - TC_t(j) = \max \left(P_t(j) - MC_t(j)\right) Y_t(j)$$

subject to $Y_t(j) = \left(\frac{P_t(j)}{P_t}\right)^{-\epsilon_t} Y_t$. The first order conditions imply that

$$\left(P_t(j) - \frac{\epsilon_t}{\epsilon_t - 1} MC_t(j) \right) Y_t(j) = 0$$

Thus, the optimal price exceeds the marginal cost since the elasticity of good variety exceeds 1.

For the fraction of firms ζ_p that can not reoptimize prices we assume

$$P_t(j) = P_{t-1}(j)$$

[2] Following Gali (2008), we assume that firms take the marginal cost as given and do not optimize subject to equation (38). Thus, the only constraint firms face is the demand schedule.

Let $\widetilde{V}_t$ be the value of a firm allowed to change prices at time t and let $V_t(P_{t-1}(i))$ be the value of a firm not allowed to change prices. Since the problem is identical for all firms of one type, they will choose the same optimal price. The value of a firm allowed to change the price is

$$\widetilde{V}_t = \max_{\widetilde{P}_t} \left[Pr_t(\widetilde{P}_t) + \beta E_t \frac{Q_{t+1}}{Q_t} \left((1-\zeta_p)\widetilde{V}_{t+1} + \zeta_p V_{t+1}(\widetilde{P}_t) \right) \right]$$

where $\frac{Q_{t+k}}{Q_t} = \frac{\mathcal{L}_{t+1}/P_{t+1}}{\mathcal{L}_t/P_t}$ is the stochastic discount factor. The value of the firm not allowed to change prices is

$$V_t(P_{t-1}) = Pr_t(P_{t-1}) + \beta E_t \frac{Q_{t+1}}{Q_t} \left((1-\zeta_p)\widetilde{V}_{t+1} + \zeta_p V_{t+1}(P_{t-1}) \right)$$

From the first order condition and the envelope theorem we have

$$\begin{aligned} 0 &= Pr_t'(\widetilde{P}_t) + \beta E_t \frac{Q_{t+1}}{Q_t} \zeta_p V_{t+1}'(\widetilde{P}_t) \qquad (36) \\ V_t'(P_{t-1}) &= Pr_t'(P_{t-1}) + \beta E_t \frac{Q_{t+1}}{Q_t} \zeta_p V_{t+1}'(P_{t-1}) \end{aligned}$$

Moving forward the latter equation, assuming $\widetilde{P}_t = P_t$ and iterating foreword we have

$$V_{t+1}'(\widetilde{P}_t) = Pr_{t+1}'(\widetilde{P}_t) + E_{t+1} \frac{Q_{t+2}}{Q_{t+1}} \beta\zeta_p \left(Pr_{t+2}'(\widetilde{P}_t) + E_{t+3} \frac{Q_{t+3}}{Q_{t+2}} \beta\zeta_p \left(Pr_{t+3}'(\widetilde{P}_t) + ... \right) \right)$$

multiplying by $\frac{Q_{t+1}}{Q_t}\beta\xi_p$ and taking expectations conditional on time t information we get

$$\begin{aligned} E_t \left(\frac{Q_{t+1}}{Q_t} \beta\zeta_p V_{t+1}'(\widetilde{P}_t) \right) &= E_t (\frac{Q_{t+1}}{Q_t} \beta\zeta_p Pr_{t+1}'(\widetilde{P}_t) + \frac{Q_{t+2}}{Q_t} (\beta\zeta_p)^2 Pr_{t+2}'(\widetilde{P}_t) \\ &+ \frac{Q_{t+3}}{Q_t} (\beta\zeta_p)^3 Pr_{t+3}'(\widetilde{P}_t) + ...) = \\ &= E_t \left(\sum_{k=1}^{\infty} \frac{Q_{t+k}}{Q_t} (\beta\zeta_p)^k Pr_{t+k}'(\widetilde{P}_t) \right) \end{aligned}$$

Substituting the latter into (36) we obtain

$$E_t \left(\sum_{k=0}^{\infty} \frac{Q_{t+k}}{Q_t} (\beta\zeta_p)^k Pr_{t+k}'(\widetilde{P}_t) \right) = 0$$

Substituting the first order condition for profit maximization we get

$$E_t \left(\sum_{k=0}^{\infty} \frac{Q_{t+k}}{Q_t} (\beta\zeta_p)^k \left(1 - \epsilon_{t+k} + \frac{MC_{t+k}(i)}{\widetilde{P}_t} \epsilon_{t+k} \right) Y_{t+k}(j) \right) = 0$$

Cost minimization implies that the marginal cost is equal to the average cost, so

$$
\begin{aligned}
MC_t(j) = TC_t(j)/Y_t(j) = \frac{W_t N_t(i)}{Y_t(j)} &= \frac{W_t}{Y_t(j)} \left(\frac{Y_t(j)}{Z_t}\right)^{\frac{1}{1-\alpha}} \\
&= W_t Y_t(j)^{\frac{\alpha}{1-\alpha}} Z_t^{-\frac{1}{1-\alpha}}
\end{aligned} \tag{37}
$$

Combining the marginal cost equation and the demand schedule we get

$$
\begin{aligned}
MC_t(i) = W_t Y_t(i)^{\frac{\alpha}{1-\alpha}} Z_t^{-\frac{1}{1-\alpha}} &= W_t \left(\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon_t} Y_t\right)^{\frac{\alpha}{1-\alpha}} Z_t^{-\frac{1}{1-\alpha}} \\
&= W_t Y_t^{\frac{\alpha}{1-\alpha}} Z_t^{-\frac{1}{1-\alpha}} \left(\frac{P_t(j)}{P_t}\right)^{\frac{-\alpha\epsilon_t}{1-\alpha}} = MC_t \left(\frac{P_t(j)}{P_t}\right)^{\frac{-\alpha\epsilon_t}{1-\alpha}}
\end{aligned} \tag{38}
$$

Thus the first order condition associated to the firm program is

$$
E_t \left(\sum_{k=0}^{\infty} \frac{Q_{t+k}}{Q_t} (\beta\zeta_p)^k \left[1 - \epsilon_{t+k} + MC^r_{t+k}\epsilon_{t+k} \left(\frac{\widetilde{P}_t}{P_{t+k}}\right)^{\frac{\alpha-1-\alpha\epsilon_{t+k}}{1-\alpha}} \right] Y_{t+k}(j) \right) = 0
$$

where $MC^r_{t+k} = \frac{MC_{t+k}}{P_{t+k}}$ is the real (aggregate) marginal cost. In the non stochastic steady state the latter equation is verified if an only if the term inside the square brackets is zero, thus

$$
1 - \epsilon + MC^r \, \epsilon \left(\frac{\widetilde{P}}{P}\right)^{\frac{\alpha-1-\alpha\epsilon}{1-\alpha}} = 0
$$

Recall that the price deflator is $P_t = \left(\int_0^1 P_t(j)^{1-\epsilon_t} dj\right)^{\frac{1}{1-\epsilon_t}}$. The law of motion of prices is

$$
1 = \left(\zeta_p \left(\frac{P_{t-1}}{P_t}\right)^{1-\epsilon_t} + (1-\zeta_p) \left(\frac{\widetilde{P}_t}{P_t}\right)^{1-\epsilon_t} \right)^{\frac{1}{1-\epsilon_t}}
$$

In the steady state

$$
1 = \left(\zeta_p \left(\frac{P}{P}\right)^{1-\epsilon} + (1-\zeta_p) \left(\frac{\widetilde{P}}{P}\right)^{1-\epsilon} \right)^{\frac{1}{1-\epsilon}} = \zeta_p + (1-\zeta_p) \left(\frac{\widetilde{P}}{P}\right)^{1-\epsilon}
$$

Thus, $\widetilde{P} = P$ and the real marginal cost in the steady state is $MC^r = \frac{\epsilon-1}{\epsilon}$.

Market clearing and Aggregation

Market clearing in the goods market requires $Y_t(j) = C_t(j)$. Letting the aggregate output be $Y_t \equiv \left(\int_0^1 Y_t(j)^{\frac{\epsilon_t-1}{\epsilon_t}} dj\right)^{\frac{\epsilon_t}{\epsilon_t-1}}$ we have $C_t = Y_t$. In the labor market we have that

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

Similarly, the aggregate real marginal cost is

$$\begin{aligned}
MC_t^r &= \int_0^1 MC_t^r(j)dj = \int_0^1 \frac{W_t/P_t N_t(j)}{Y_t(j)} dj \\
&= \frac{W_t}{P_t}\int_0^1 \frac{1}{Y_t(j)}\left(\frac{Y_t(j)}{Z_t}\right)^{\frac{1}{1-\alpha}} dj = \frac{W_t}{P_t}\left(\frac{1}{Z_t}\right)^{\frac{1}{1-\alpha}} \int_0^1 Y_t(j)^{\frac{\alpha}{1-\alpha}} dj = \\
&= \frac{W_t}{P_t}\left(\frac{1}{Z_t}\right)^{\frac{1}{1-\alpha}} Y_t^{\frac{\alpha}{1-\alpha}} \int_0^1 \left(\frac{P_t(j)}{P_t}\right)^{\frac{-\epsilon_t\alpha}{1-\alpha}} dj
\end{aligned}$$

To sum up the main equations of the model are

$$\begin{aligned}
0 &= \chi_t(C_t - hC_{t-1})^{-\sigma_c} - \mathcal{L}_t = \chi_t(C_t - hC_{t-1})^{-\sigma_c} - P_t Q_t \\
0 &= N_t^{\sigma_n} - \mathcal{L}_t \frac{W_t}{P_t} \\
1 &= E_t\left[\beta \frac{\mathcal{L}_{t+1}}{\mathcal{L}_t}\frac{R_t}{\Pi_{t+1}}\right] \\
0 &= E_t\left(\sum_{k=0}^{\infty}\frac{Q_{t+k}}{Q_t}(\beta\zeta_p)^k\left[1-\epsilon_{t+k} + MC_{t+k}^r\epsilon_{t+k}\left(\frac{\widetilde{P}_t}{P_{t+k}}\right)^{\frac{\alpha-1-\alpha\epsilon_{t+k}}{1-\alpha}}\right]Y_{t+k}(j)\right) \\
1 &= \left(\zeta_p\left(\frac{P_{t-1}}{P_t}\right)^{1-\epsilon_t} + (1-\zeta_p)\left(\frac{\widetilde{P}_t}{P_t}\right)^{1-\epsilon_t}\right)^{\frac{1}{1-\epsilon_t}} \\
Y_t &= C_t \\
N_t &= \left(\frac{Y_t}{Z_t}\right)^{\frac{1}{1-\alpha}} \int_0^1 \left(\frac{P_t(j)}{P_t}\right)^{\frac{-\epsilon_t}{1-\alpha}} dj \\
MC_t^r &= \frac{W_t}{P_t}\left(\frac{1}{Z_t}\right)^{\frac{1}{1-\alpha}} Y_t^{\frac{\alpha}{1-\alpha}} \int_0^1 \left(\frac{P_t(j)}{P_t}\right)^{\frac{-\epsilon_t\alpha}{1-\alpha}} dj
\end{aligned}$$

We now derive the log linearized conditions when either the technology process or the preference process have a non-stationary component.

Non stationary technology shock

Assume that the preference shock is $\ln \chi_t = \rho_\chi \ln \chi_{t-1} + \epsilon_t^\chi$ where $\epsilon_t^\chi \sim N(0, \sigma_\chi^2)$ and that technology process has two components, an autoregressive and a stochastic time trend, that is

$$Z_t = Z_t^c Z_t^T$$
$$\ln Z_t^T = bt + e_t^{Z,T}$$
$$\ln Z_t^c = \rho_z \ln Z_{t-1}^c + e_t^{Z,c}$$

The equilibrium conditions need to be rescaled by Z_t^T. Let $\widehat{Y}_t = \frac{Y_t}{Z_t^T}$, $\widehat{C}_t = \frac{C_t}{Z_t^T}$ and $\widehat{W}_t = \frac{W_t}{Z_t^T}$, $\widehat{\mathcal{L}}_t = \mathcal{L}_t (Z_t^T)^{\sigma_c}$, $\widehat{Q}_{t+k} = \widehat{\mathcal{L}}_{t+k} P_{t+k}$. Then:

$$N_t = \left(\frac{\widehat{Y}_t}{Z_t^c}\right)^{\frac{1}{1-\alpha}} \int_0^1 \left(\frac{P_t(j)}{P_t}\right)^{\frac{-\epsilon_t}{1-\alpha}} dj$$

$$MC_t^r = \frac{W_t}{P_t}\left(\frac{1}{Z_t}\right)^{\frac{1}{1-\alpha}} Y_t^{\frac{\alpha}{1-\alpha}} \int_0^1 \left(\frac{P_t(j)}{P_t}\right)^{\frac{-\epsilon_t \alpha}{1-\alpha}} dj$$

$$= \frac{W_t}{P_t}\left(\frac{1}{Z_t^c}\right)^{\frac{1}{1-\alpha}} \left(\frac{1}{Z_t^T}\right)^{\frac{1-\alpha+\alpha}{1-\alpha}} Y_t^{\frac{\alpha}{1-\alpha}} \int_0^1 \left(\frac{P_t(j)}{P_t}\right)^{\frac{-\epsilon_t \alpha}{1-\alpha}} dj$$

$$= \frac{\widehat{W}_t}{P_t}\left(\frac{1}{Z_t^c}\right)^{\frac{1}{1-\alpha}} \widehat{Y}_t^{\frac{\alpha}{1-\alpha}} \int_0^1 \left(\frac{P_t(j)}{P_t}\right)^{\frac{-\epsilon_t \alpha}{1-\alpha}} dj$$

$$\widehat{C}_t = \widehat{Y}_t$$

$$\mathcal{L}_t (1/Z_t^T)^{-\sigma_c} = \chi_t (\widehat{C}_t - h\widehat{C}_{t-1} \frac{Z_{t-1}^T}{Z_t^T})^{-\sigma_c}$$

$$\widehat{\mathcal{L}}_t = \chi_t (\widehat{C}_t - h\widehat{C}_{t-1} \exp\{-b + e_{t-1}^{Z,T} - e_t^{Z,T}\})^{-\sigma_c}$$

$$\widehat{N}_t^{-\sigma_n} = -\widehat{\mathcal{L}}_t \frac{\widehat{W}_t}{P_t}$$

where $\widehat{N}_t = \frac{N_t}{(Z_t^T)^{\frac{\sigma_c - 1}{\sigma_n}}}$. Thus, if $\sigma_c = 1$, hours worked is stationary and consistency is insured. The Euler equation becomes

$$1 = E_t \left[\beta \frac{\widehat{\mathcal{L}}_{t+1}}{\widehat{\mathcal{L}}_t} \exp\{-b - e_{t+1}^{Z,T} + e_t^{Z,T}\} \frac{R_t}{\Pi_{t+1}}\right]$$

The firm optimal condition when $\sigma_c = 1$ is

$$0 = E_t \left(\sum_{k=0}^{\infty} \frac{\widehat{Q}_{t+k}}{\widetilde{Q}_t} (\beta\zeta_p)^k \left[1 - \epsilon_{t+k} + MC^r_{t+k}\epsilon_{t+k} \left(\frac{\widehat{P}_t}{P_{t+k}} \right)^{\frac{\alpha-1-\alpha\epsilon_{t+k}}{1-\alpha}} \right] \widehat{Y}_{t+k}(j) \right)$$

Log linearization of the equilibrium conditions leads to

$$\begin{aligned}
\lambda_t &= \chi_t - \frac{1}{1-\overline{h}}(y_t - \overline{h}y_{t-1} - \overline{h}e^{Z,T}_{t-1} + \overline{h}e^{Z,T}_t) \\
w_t &= \sigma_n n_t - \lambda_t \\
y_t &= z_t + (1-\alpha)n_t \\
mc_t &= \omega_t + n_t - y_t \\
r_t &= \rho_r r_{t-1} + (1-\rho_r)(\rho_y y_t + \rho_\pi \pi_t) + v_t \\
\lambda_t &= E_t[\lambda_{t+1} + r_t - \pi_{t+1} - e^{z,T}_{t+1} + e^{z,T}_t] \\
\pi_t &= \beta E_t \pi_{t+1} + \kappa_p (\mu_t + mc_t)
\end{aligned}$$

where $\kappa_p = \frac{1-\alpha}{1-\alpha+\alpha\epsilon}\frac{(1-\beta\zeta_p)(1-\zeta_p)}{\zeta_p}$, $\overline{h} = e^{-b}h$ and variables in small letters are rescaled variables in log deviation from the steady state. Thus, in log deviations from the steady state:

$$\begin{aligned}
\ln Y_t &= bt + e^{Z,T}_t + y_t \\
\ln W_t &= bt + e^{Z,T}_t + w_t \\
\ln \Pi_t &= \pi_t \\
\ln R_t &= r_t
\end{aligned}$$

Non stationary preference shock

Assume that the technology shock is $\ln z_t = \rho_z \ln z_{t-1} + \epsilon^z_t$ where $\epsilon^z_t \sim N(0, \sigma^2_z)$ and that the preference process is

$$\begin{aligned}
\chi_t &= (\chi^T_t)^{1+\sigma_n}\chi^c_t \\
\ln \chi^T_t &= \ln \chi^T_{t-1} + e^{T,\chi}_t \\
\ln \chi^c_t &= \rho_\chi \ln \chi^c_{t-1} + e^{c,\chi}_t
\end{aligned}$$

where $e_t^{j,\chi} \sim N(0, \sigma_{j,\chi}^2)$ with $j = T, c$. Assume further that $\sigma_c = 1$ and $\alpha = 0$. Define $\widehat{C}_t = C_t/\chi_t^T$, $\widehat{Y}_t = Y_t/\chi_t^T$, $\widehat{N}_t = N_t/\chi_t^T$, $\widehat{\mathcal{L}}_t = \mathcal{L}_t(\chi_t^T)^{-\sigma_n}$, $\widehat{Q}_{t+k} = \widehat{\mathcal{L}}_{t+k}P_{t+k}$. The equilibrium conditions become

$$\begin{aligned}
\widehat{\mathcal{L}}_t &= \frac{\chi_t^c}{\widehat{C}_t - h\widehat{C}_{t-1}\exp(-e_t^{T,\chi})} \\
0 &= \widehat{N}_t^{\sigma_n} + \widehat{\mathcal{L}}_t \frac{W_t}{P_t} \\
1 &= \beta E_t \left[\frac{\widehat{\mathcal{L}}_{t+1}}{\widehat{\mathcal{L}}_t} R_t \frac{P_t}{P_{t+1}} \exp(\sigma_n \epsilon_{t+1}^{T,\chi}) \right]
\end{aligned}$$

$$\begin{aligned}
\widehat{N}_t &= \frac{\widehat{Y}_t}{Z_t} \int_0^1 \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon_t} dj \\
MC_t^r &= \frac{W_t}{P_t} \frac{1}{Z_t}
\end{aligned}$$

$$0 = E_t \sum_{k=0}^{\infty} \frac{\widehat{Q}_{t+k}}{\widehat{Q}_t} \exp\left[(1+\sigma_n) \sum_{j=0}^{k-1} \epsilon_{t+k-1}^{\chi,P} \right] (\beta\zeta_p)^k \left[1 - \epsilon_{t+k} + MC_{t+k}^r \epsilon_{t+k} \left(\frac{\widetilde{P}_t}{P_{t+k}} \right)^{\frac{\alpha - 1 - \alpha\epsilon_{t+k}}{1-\alpha}} \right] \widehat{Y}_{t+k}(j)$$

Log linearization leads to

$$\begin{aligned}
\lambda_t &= \chi_t^c - \frac{1}{1-h}(y_t - hy_{t-1} + he_t^{T,\chi}) \\
w_t &= \sigma_n n_t - \lambda_t \\
y_t &= z_t + n_t \\
mc_t &= \omega_t + n_t - y_t \\
r_t &= \rho_r r_{t-1} + (1-\rho_r)(\rho_y y_t + \rho_\pi \pi_t) + v_t \\
\lambda_t &= E_t[\lambda_{t+1} + r_t - \pi_{t+1} + \sigma_n e_{t+1}^{T,\chi}] \\
\pi_t &= \beta E_t \pi_{t+1} + \kappa_p (\mu_t + mc_t)
\end{aligned}$$

where variables in small letters are rescaled variables in log deviation from the steady state. Thus, in log deviations from the steady state:

$$\begin{aligned}
\ln Y_t &= \chi_t^T + y_t \\
\ln W_t &= w_t \\
\ln \Pi_t &= \pi_t \\
\ln R_t &= r_t
\end{aligned}$$

Appendix B

This appendix reports the estimation results mentioned in the paper for alternative specifications of the DGP of the non-cyclical component, for alternative combinations of filtered and unfiltered observables and for different sample sizes.

Filter	True	LT	HP	FOD	BP
		Median (s.e.)	Median (s.e.)	Median (s.e.)	Median (s.e.)
σ_c	1.0	1.13 (0.07)	1.15 (0.08)	1.07 (0.04)	1.07 (0.07)
σ_n	0.7	1.34 (0.06)	1.31 (0.06)	1.32 (0.05)	1.38 (0.06)
h	0.7	0.59 (0.03)	0.58 (0.03)	0.59 (0.02)	0.64 (0.02)
α	0.4	0.14 (0.02)	0.15 (0.02)	0.13 (0.01)	0.21 (0.02)
ϵ	7.0	3.85 (0.13)	4.51 (0.16)	4.19 (0.13)	3.81 (0.14)
ρ_r	0.2	0.74 (0.03)	0.73 (0.03)	0.67 (0.02)	0.68 (0.03)
ρ_π	1.3	1.45 (0.07)	1.53 (0.06)	1.59 (0.05)	1.54 (0.06)
ρ_y	0.05	0.48 (0.05)	0.46 (0.05)	-0.01 (0.00)	0.06 (0.02)
ζ_p	0.8	0.87 (0.03)	0.87 (0.03)	0.89 (0.03)	0.88 (0.03)
ρ_χ	0.5	0.74 (0.04)	0.76 (0.04)	0.42 (0.02)	0.99 (0.03)
ρ_z	0.8	0.40 (0.04)	0.46 (0.05)	0.99 (0.03)	0.57 (0.03)
σ_χ	1.12	0.19 (0.03)	0.19 (0.03)	0.16 (0.02)	0.07 (0.01)
$\sigma_{z,c}$	0.51	0.07 (0.01)	0.07 (0.01)	0.15 (0.02)	0.07 (0.01)
σ_{mp}	0.12	0.10 (0.01)	0.09 (0.01)	0.11 (0.01)	0.07 (0.01)
σ_μ	20.64	1.78 (0.33)	1.50 (0.20)	6.28 (0.25)	0.60 (0.08)

Table 6: Parameters estimates using different filters, all variables filtered, DGP1.

Filter	LT	HP	FOD	BP
	Median (s.e.)	Median (s.e.)	Median (s.e.)	Median(s.e.)
σ_c	1.14 (0.08)	1.15 (0.09)	1.27 (0.06)	1.21 (0.08)
σ_n	1.36 (0.07)	1.39 (0.08)	2.52 (0.11)	1.74 (0.11)
h	0.60 (0.03)	0.61 (0.03)	0.53 (0.03)	0.66 (0.03)
α	0.15 (0.02)	0.14 (0.02)	0.35 (0.03)	0.15 (0.03)
ϵ	3.98 (0.13)	3.37 (0.12)	4.10 (0.14)	4.19 (0.17)
ρ_r	0.75 (0.03)	0.75 (0.03)	0.71 (0.02)	0.66 (0.03)
ρ_π	1.66 (0.09)	1.66 (0.10)	1.61 (0.06)	1.45 (0.08)
ρ_y	0.49 (0.05)	0.58 (0.07)	-0.01 (0.00)	0.59 (0.06)
ζ_p	0.87 (0.03)	0.87 (0.03)	0.85 (0.03)	0.83 (0.03)
ρ_χ	0.73 (0.06)	0.78 (0.04)	0.30 (0.02)	0.82 (0.03)
ρ_z	0.45 (0.05)	0.39 (0.04)	0.99 (0.03)	0.24 (0.04)
σ_χ	0.19 (0.03)	0.23 (0.04)	0.83 (0.13)	0.48 (0.07)
$\sigma_{z,c}$	0.07 (0.01)	0.09 (0.01)	0.14 (0.02)	0.15 (0.02)
σ_{mp}	0.10 (0.01)	0.10 (0.01)	0.09 (0.01)	0.10 (0.01)
σ_μ	2.07 (0.31)	1.85 (0.27)	10.67 (0.49)	0.65 (0.14)

Table 7: Parameters estimates using different filters, real variables filtered, DGP1.

Filter	LT	HP	FOD	BP
	Median (s.e.)	Median (s.e.)	Median (s.e.)	Median(s.e.)
σ_c	1.12 (0.05)	1.13 (0.05)	1.06 (0.04)	1.07 (0.05)
σ_n	1.33 (0.05)	1.34 (0.05)	1.35 (0.05)	1.34 (0.05)
h	0.60 (0.02)	0.60 (0.02)	0.61 (0.02)	0.63 (0.02)
α	0.13 (0.01)	0.14 (0.01)	0.13 (0.01)	0.18 (0.01)
ϵ	4.18 (0.13)	3.97 (0.13)	4.00 (0.13)	4.06 (0.13)
ρ_r	0.77 (0.03)	0.76 (0.03)	0.68 (0.02)	0.70 (0.02)
ρ_π	1.60 (0.05)	1.59 (0.09)	1.53 (0.05)	1.60 (0.06)
ρ_y	0.49 (0.03)	0.41 (0.04)	-0.01 (0.00)	0.08 (0.01)
ζ_p	0.88 (0.03)	0.88 (0.03)	0.89 (0.03)	0.87 (0.03)
ρ_χ	0.55 (0.06)	0.33 (0.02)	0.36 (0.03)	0.99 (0.03)
ρ_z	0.44 (0.04)	0.49 (0.03)	0.99 (0.03)	0.71 (0.03)
σ_χ	0.12 (0.02)	0.09 (0.01)	0.16 (0.01)	0.04 (0.00)
$\sigma_{z,c}$	0.04 (0.00)	0.04 (0.00)	0.11 (0.01)	0.04 (0.00)
σ_{mp}	0.07 (0.01)	0.06 (0.01)	0.08 (0.01)	0.04 (0.00)
σ_μ	2.31 (0.19)	2.10 (0.16)	7.20 (0.31)	0.59 (0.05)

Table 8: Parameters estimates using different filters, all variables filtered, DGP1, sample size is T=300.

Filter	LT	HP	FOD	BP
	Median (s.e.)	Median (s.e.)	Median (s.e.)	Median (s.e.)
σ_c	1.13 (0.07)	1.13 (0.08)	1.08 (0.04)	1.02 (0.07)
σ_n	1.34 (0.06)	1.32 (0.06)	1.30 (0.06)	1.38 (0.06)
h	0.59 (0.03)	0.58 (0.03)	0.58 (0.03)	0.65 (0.02)
α	0.13 (0.02)	0.14 (0.03)	0.13 (0.02)	0.19 (0.02)
ϵ	3.67 (0.14)	4.20 (0.14)	4.13 (0.13)	4.03 (0.13)
ρ_r	0.73 (0.03)	0.72 (0.03)	0.67 (0.02)	0.68 (0.03)
ρ_π	1.59 (0.12)	1.60 (0.10)	1.55 (0.05)	1.62 (0.06)
ρ_y	0.45 (0.04)	0.41 (0.05)	-0.01 (0.00)	0.06 (0.01)
ζ_p	0.88 (0.03)	0.87 (0.03)	0.89 (0.03)	0.88 (0.03)
ρ_χ	0.76 (0.04)	0.78 (0.04)	0.45 (0.02)	0.99 (0.03)
ρ_z	0.45 (0.05)	0.39 (0.06)	0.99 (0.03)	0.59 (0.04)
σ_χ	0.19 (0.03)	0.19 (0.03)	0.17 (0.02)	0.07 (0.01)
$\sigma_{z,c}$	0.07 (0.01)	0.07 (0.01)	0.15 (0.02)	0.07 (0.01)
σ_{mp}	0.10 (0.01)	0.09 (0.01)	0.11 (0.01)	0.07 (0.01)
σ_μ	1.87 (0.24)	1.44 (0.24)	6.32 (0.35)	0.58 (0.07)
σ_{me1}	0.61 (0.20)	0.68 (0.26)	0.51 (0.09)	0.70 (0.31)
σ_{me2}	0.64 (0.19)	0.62 (0.21)	0.75 (0.19)	0.58 (0.15)
σ_{me3}	0.68 (0.21)	0.66 (0.31)	0.89 (0.25)	0.56 (0.18)
σ_{me4}	0.56 (0.25)	0.68 (0.19)	0.64 (0.11)	0.68 (0.30)

Table 9: Parameters estimates using different filters, all variables filtered, DGP1, model with measurement error.

Filter	LT	HP	FOD	BP
	Median (s.e.)	Median (s.e.)	Median (s.e.)	Median (s.e.)
σ_c	3.90 (0.36)	4.71 (0.25)	3.23 (0.86)	5.22 (0.25)
σ_n	0.30 (0.05)	0.20 (0.02)	0.28 (0.03)	0.06 (0.03)
h	0.59 (0.03)	0.56 (0.03)	0.70 (0.02)	0.87 (0.03)
ϵ	4.20 (0.14)	4.00 (0.13)	4.10 (0.13)	4.02 (0.13)
ρ_r	0.30 (0.01)	0.29 (0.02)	0.56 (0.02)	0.16 (0.02)
ρ_π	1.75 (0.07)	1.67 (0.06)	1.56 (0.05)	1.48 (0.05)
ρ_y	-0.03 (0.01)	-0.08 (0.02)	0.03 (0.02)	-0.13 (0.01)
ζ_p	0.83 (0.03)	0.84 (0.03)	0.81 (0.03)	0.86 (0.03)
ρ_χ	0.62 (0.06)	0.39 (0.02)	0.48 (0.02)	0.79 (0.03)
ρ_z	0.72 (0.03)	0.67 (0.02)	0.48 (0.02)	0.38 (0.02)
$\sigma_{\chi,c}$	0.15 (0.03)	0.17 (0.02)	0.72 (0.38)	0.37 (0.07)
σ_z	0.15 (0.02)	0.20 (0.03)	0.38 (0.08)	5.15 (0.25)
σ_v	0.03 (0.00)	0.03 (0.00)	0.03 (0.00)	0.03 (0.00)
σ_μ	7.28 (0.49)	8.92 (0.46)	4.95 (0.23)	3.74 (0.33)

Table 10: Parameters estimates using different filters, real variables filtered, DGP2.

		Factor 1	Factor 2
	True	Median (s.e.)	Median (s.e.)
σ_c	1.00	0.87 (0.10)	1.72 (0.10)
σ_n	0.70	0.73 (0.06)	0.29 (0.09)
h	0.70	0.56 (0.10)	0.62 (0.03)
α	0.40	0.34 (0.04)	0.32 (0.03)
ϵ	7.00	6.29 (0.13)	6.45 (0.14)
ρ_r	0.20	0.67 (0.03)	0.80 (0.04)
ρ_π	1.30	1.61 (0.03)	1.51 (0.02)
ρ_y	0.05	0.40 (0.03)	0.31 (0.04)
ζ_p	0.80	0.85 (0.03)	0.85 (0.03)
ρ_χ	0.50	0.82 (0.06)	0.69 (0.08)
ρ_z	0.80	0.70 (0.03)	0.69 (0.02)
$\sigma_{\chi,c}$	1.10	0.22 (0.04)	0.21 (0.04)
σ_z	0.57	0.18 (0.03)	0.29 (0.08)
σ_v	0.12	0.13 (0.02)	0.10 (0.01)
σ_μ	20.64	6.51 (1.11)	6.07 (1.25)

Table 11: Posterior parameters estimates. Factor 1 uses LT, HP and FOD filtered data; factor 2 HP($\lambda =$ 1600), HP($\lambda = 6400$) and BP filtered data. The DGP features a technology shock with two components, a stationary AR(1) and a unit root. All variables are filtered prior to estimation. The sample size is T=150.

Appendix C: Full set of estimates, model with money

Parameter	Prior	Basic model median (s.e.)	Ireland's specification median (s.e.)
ω_1	$\Gamma(1, 0.1)$	1.03 (0.02)	0.98 (0.01)
ω_2	$\Gamma(1, 0.1)$	0.44 (0.02)	0.03 (0.01)
ψ	$N(1.0, 0.1)$	1.02 (0.02)	0.98 (0.03)
ρ_r	$B(2, 6)$	0.59 (0.01)	0.59 (0.01)
ρ_π	$N(1.5, 0.2)$	1.51 (0.02)	1.40 (0.01)
ρ_y	$N(0.2, 0.2)$	0.44 (0.01)	0.45 (0.01)
ρ_m	$N(1.0, 0.2)$	0.48 (0.02)	0.04 (0.02)
γ_1	$\Gamma(10, 0.1)$	0.92 (0.02)	1.00 (0.01)
γ_2	$\Gamma(5, 0.1)$	0.51 (0.01)	0.51 (0.01)
ρ_a	$B(8, 8)$	0.72 (0.01)	0.67 (0.01)
ρ_e	$B(8, 8)$	0.77 (0.01)	0.79 (0.03)
ρ_z	$B(22, 8)$	0.74 (0.04)	0.88 (0.01)
σ_a	$\Gamma^{-1}(5, 20)$	0.74 (0.10)	1.30 (0.07)
σ_e	$\Gamma^{-1}(5, 20)$	0.81 (0.08)	2.02 (0.16)
σ_z	$\Gamma^{-1}(5, 20)$	0.18 (0.03)	0.46 (0.40)
σ_v	$\Gamma^{-1}(5, 20)$	0.37 (0.06)	0.52 (0.15)

Table 12: Structural parameters estimates for model with money. Γ is the gamma distribution, B is the beta distribution, N is the normal distribution.

Parameter	Prior	Basic model median (s.e.)	Ireland's specification median (s.e.)
ν_0^{po}	N(0,0.1)	0.00 (0.00)	
ν_0^{fd}	N(0,0.1)	-0.00 (0.00)	
ν_0^{hp1}	N(0,0.1)	-0.00 (0.00)	
ν_0^{hp2}	N(0,0.1)	-0.00 (0.00)	
ν_0^{cum}	N(0,0.1)	0.00 (0.00)	
ν_0^{cd}	N(0,0.1)	0.00 (0.00)	
ν_0^{uc}	N(0,0.1)	0.00 (0.00)	
ν_0^{mbn}	N(0,0.1)	0.00 (0.00)	
ν_1^{fd}	N(1,0.5)	0.73 (0.01)	
ν_1^{hp1}	N(1,0.5)	0.82 (0.02)	
ν_1^{hp2}	N(1,0.5)	0.77 (0.01)	
ν_1^{cum}	N(1,0.5)	0.77 (0.02)	
ν_1^{cd}	N(1,0.5)	0.86 (0.02)	
ν_1^{uc}	N(1,0.5)	0.70 (0.05)	
ν_1^{mbn}	N(1,0.5)	0.78 (0.01)	
σ_y^{po}	$\Gamma^{-1}(10,30)$	0.12 (0.01)	0.12 (0.01)
σ_m^{po}	$\Gamma^{-1}(10,30)$	0.24 (0.03)	0.06 (0.01)
σ_π^{po}	$\Gamma^{-1}(10,30)$	0.06 (0.01)	0.09 (0.01)
σ_r^{po}	$\Gamma^{-1}(10,30)$	0.06 (0.01)	0.07 (0.01)
σ_y^{fd}	$\Gamma^{-1}(10,30)$	0.03 (0.00)	
σ_w^{fd}	$\Gamma^{-1}(10,30)$	0.04 (0.00)	
σ_π^{fd}	$\Gamma^{-1}(10,30)$	0.03 (0.00)	
σ_r^{fd}	$\Gamma^{-1}(10,30)$	0.03 (0.00)	
σ_y^{hp1}	$\Gamma^{-1}(10,30)$	0.05 (0.01)	
σ_w^{hp1}	$\Gamma^{-1}(10,30)$	0.05 (0.01)	
σ_π^{hp1}	$\Gamma^{-1}(10,30)$	0.04 (0.00)	
σ_r^{hp1}	$\Gamma^{-1}(10,30)$	0.04 (0.00)	
σ_y^{hp2}	$\Gamma^{-1}(10,30)$	0.05 (0.00)	
σ_w^{hp2}	$\Gamma^{-1}(10,30)$	0.05 (0.01)	
σ_π^{hp2}	$\Gamma^{-1}(10,30)$	0.04 (0.00)	
σ_r^{hp2}	$\Gamma^{-1}(10,30)$	0.04 (0.00)	
σ_y^{cum}	$\Gamma^{-1}(10,30)$	0.04 (0.00)	
σ_w^{cum}	$\Gamma^{-1}(10,30)$	0.07 (0.01)	
σ_π^{cum}	$\Gamma^{-1}(10,30)$	0.04 (0.00)	
σ_r^{cum}	$\Gamma^{-1}(10,30)$	0.03 (0.00)	
σ_y^{cd}	$\Gamma^{-1}(10,30)$	0.12 (0.01)	
σ_w^{cd}	$\Gamma^{-1}(10,30)$	0.25 (0.03)	
σ_π^{cd}	$\Gamma^{-1}(10,30)$	0.06 (0.01)	
σ_r^{cd}	$\Gamma^{-1}(10,30)$	0.06 (0.01)	
σ_y^{uc}	$\Gamma^{-1}(10,30)$	0.03 (0.00)	
σ_w^{uc}	$\Gamma^{-1}(10,30)$	0.04 (0.00)	
σ_π^{uc}	$\Gamma^{-1}(10,30)$	0.04 (0.00)	
σ_r^{uc}	$\Gamma^{-1}(10,30)$	0.03 (0.00)	
σ_y^{mbn}	$\Gamma^{-1}(10,30)$	0.06 (0.01)	
σ_w^{mbn}	$\Gamma^{-1}(10,30)$	0.07 (0.01)	
σ_π^{mbn}	$\Gamma^{-1}(10,30)$	0.04 (0.00)	
σ_r^{mbn}	$\Gamma^{-1}(10,30)$	0.04 (0.00)	

Table 13: Additional parameters estimates for a model with money. Γ is the gamma distribution, B is the beta distribution, N is the normal distribution.