Multiple filtering devices for the estimation of cyclical DSGE models

Fabio Canova *
ICREA-UPF, AMeN and CEPR

Filippo Ferroni
UPF

November 10, 2008

Abstract

We propose a method to estimate time invariant cyclical DSGE models using the information provided by a variety of filtering approaches. We treat data filtered with alternative procedures as contaminated proxy of the relevant model-based quantities and estimate structural and non-structural parameters jointly using an unobservable component structure. We employ simulated data to illustrate the properties of the procedure. We compare estimates with those obtained by an approach where the cyclical data produced by just one filter is used to estimate the structural parameters. We revisit the role of money in the transmission of monetary business cycles using the suggested technique.

JEL classification: E32, C32.

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

*I would like to thank the participants of the Central Bank workshop " Macroeconomic Modeling 2008", Cartagena, Colombia, of the CEMLA meeting of researchers, Lima, Peru, and of a seminar at the Czech National Bank for comments. The financial support of the Spanish Ministry of Education, through the grant SEJ-2004-21682-E and of the Barcelona Economic program (CREA) is gratefully acknowledged.

1 Introduction

DSGE models have become the paradigm for policy analyses in academic and policy circles over the last 10 years. Relative to earlier structures, current models are of larger scale and feature numerous frictions on the real and nominal side of the economy that help to closely replicate the dynamic responses that structural VARs produce. Also, while a few years ago it was standard to informally calibrate DSGE models, increased computing power, longer time series and recent developments in system-wide estimation methods allow researchers to routinely employ a variety of full information techniques in structural estimation exercises (see, e.g., Smets and Wouters (2003), Ireland (2004), Rabanal and Rubio Ramirez (2005) among many others).

Despite the increased popularity, structural parameter estimation faces important conceptual and numerical problems. For example, as emphasized by Canova (2007b), full information classical estimation makes sense only if the model is assumed to be the data generating process (DGP) of the observables, up to a set of serially uncorrelated measurement errors, an assumption which is hard to entrain unless the model is augmented with ad-hoc dynamics. Furthermore, there are abundant population identification problems (see Canova and Sala (2006)), numerical difficulties are widespread, singularities are often important (there are typically less shocks than endogenous variables in the model) and errors-in-variables are present (the variables in the model do not often have a direct counterpart in the data). Finally, the vast majority of the models used in the literature are time invariant and intended to explain only the cyclical portion of the data fluctuations while actual data includes, at a minimum, growth components, cyclical fluctuations and high frequency noise, all of which may be subject to breaks and other forms of slowly moving variations.

When faced with the problem of fitting stationary cyclical DSGE models to the data, applied investigators typically select a 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. Occasionally, one find authors, see e.g. Kehoe (2007), suggesting that filtering should be applied to both actual data and data simulated by the model but, to the best of our knowledge, such an approach has, so far, no followers in the estimation literature. Alternatively, a unit root in total factor productivity is assumed and the data filtered using a model-driven transformation.

Both statistical and model-based filtering are problematic. For example, while the profession largely agrees that a cyclical model should explain fluctuations with 8-32 quarters average periodicities, there is little agreement on how to obtain these fluctuations from the data and only a partial understanding of the consequences that incorrect or suboptimal filtering induce. For example, it is common to use linearly detrended or first differenced data as input in the estimation process, but such transformations do not extract fluctuations with the required periodicities (see e.g. Canova (1998)). A band pass (BP) filter which, with infinite amount of data can exactly isolate the fluctuations of interest, it is typically discarded in the estimation literature because its two-sided nature may change the timing of the data information - a similar argument is made also for Hodrick and Prescott (HP) filtered data. In addition, with samples of typical length, all filters induce considerable sampling errors in the estimates of the cyclical component which pile on top of population misspecification problems. On the other hand, model-driven filtering does not necessarily leave only cycles with 8-32 quarters average periodicity in the data (see Canova (2008)) and, lacking information on the sources of non-cyclical movements in the data, imposing a unit root in technology makes model-based filtering also subject to specification errors.

Two additional important issues should be mentioned. While researchers filter each series separately prior to estimation, theory suggests that there may be important commonalities in the non-cyclical component of the data (a balanced growth path is often used as working assumption). Hence, should economic theory or pragmatic considerations guide filtering? Moreover, while real variables typically show long run drifts, nominal variables just display low frequency fluctuations. Should we filter all the data or only real variables? Conversely, should we treat all the fluctuations present in, say, inflation as relevant for parameter estimation or not? Since different researchers choose different methods to filter a portion (or all) of the available data prior to estimation, and since measurement error with unknown properties is introduced regardless of the filtering approach one employs, the economic conclusions one draws from the analysis are likely to be distorted and the magnitude of the distortions function of the transformation employed (see Canova (2008)).

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

Our approach is advantageous in, at least, two respects. Since we do not have to arbitrarily choose one filtering method prior to the estimation, we avoid specification errors of various sorts. Moreover, cyclical data obtained with one-sided and two-sided filters of both univariate and multivariate nature can be used in estimation, as long as the list of filters is sufficiently rich. Second, if different filters have sufficiently different features, measurement error may have different time series properties. Since the implicit information averaging our procedure produces may reduce measurement error and eliminate its cyclicality, estimates of the cyclical components are more reliable and precise, making parameter estimates and inference, to a large extent, free of preliminary data transformation biases.

We investigate the properties of our approach using experimental data of the typical length used in macroeconomics. We show that estimating the structural parameters of the model with just one arbitrary filter typically induces large biases in the estimates and that these biases are considerably reduced with our approach. We also show that in an unconditional forecasting exercise, the one step ahead MSE produced by our approach is smaller than the MSE obtained with standard procedures and that the biases we have noticed in traditional approaches translate in conditional forecasts which are considerably distorted.

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

We want to be explicit for why we propose a procedure to improve parameters estimates of time invariant cyclical DSGE models, rather suggest to write and estimate DSGE models which are designed to jointly fit cyclical and non-cyclical fluctuations. In the long run, the latter should clearly the scope of model builders and of applied investigators. However, right now such a task appears unfeasible. In theory, little is known about mechanisms propagating cyclical shocks at longer frequencies (exceptions are Comin and Gertler (2006) or Canova et al. (2007)) or creating important cyclical implications from long run disturbances; moreover, it is also convenient, at least for policy purposes, to assume that the mechanisms driving cyclical and non-cyclical fluctuations are distinct and orthogonal. In practice, breaks of various sorts make data less informative about the relationships of interest than otherwise and samples are always helplessly short to get information about, say, medium term cycles.

The rest of the paper is organized as follows. Next section shows the problems one encounters using a single filtering method to estimate the parameters of DSGE models. Section 3 presents our approach and applies to experimental data. Section 4 examines the role of money in transmitting monetary business cycles. Section 5 concludes.

2 Filtering and structural estimation

To illustrate why the prevailing approach 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 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 log-linearized equilibrium condition, in deviation fro the steady states, are:

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

$$y_t = z_t + (1-\alpha)n_t \quad (2)$$

$$w_t = -\lambda_t + \sigma_n n_t \quad (3)$$

$$r_t = \rho_r r_{t-1} + (1-\rho_r)(\rho_\pi \pi_t + \rho_y y_t) + v_t \quad (4)$$

$$\lambda_t = E_t(\lambda_{t+1} + r_t - \pi_{t+1}) \quad (5)$$

$$\pi_t = k_p(w_t + n_t - y_t + \mu_t) + \beta E_t \pi_{t+1} \quad (6)$$

$$\chi_t = \rho_\chi \chi_{t-1} + \iota_t^\chi \quad (7)$$

$$z_t = \rho_z z_{t-1} + \iota_t^z \quad (8)$$

where $k_p = \frac{(1-\beta\zeta_p)(1-\zeta_p)}{\zeta_p}\frac{1-\alpha}{1-\alpha+\epsilon\alpha}$, λ_t is the Lagrangian on the consumer budget constraint, y_t is output, n_t is hours, w_t is the real wage and r_t the nominal interest rate; z_t is a technology shock, χ_t a preference shock, v_t is a monetary policy shock and μ_t a markup shock. The structural parameters of the model are β, the discount factor, σ_c the risk aversion coefficient, h the coefficient of consumption habit, $1-\alpha$ the share of labor in production, σ_n the inverse of Frish elasticity, ϵ the elasticity among consumption varieties, ζ_p the probability of changing prices, while ρ_π, ρ_y are parameters of the monetary policy rule. In addition, the parameter vector includes $\rho_r, \rho_\chi, \rho_z$ the autoregressive parameters, and $\sigma_i, i = \chi, z, \mu, v$, the standard deviation of the four shocks.

We assume that either the technology shock or the preference shock has two components (a stationary autoregressive and a unit root), while the monetary policy and the markup shocks are iid. In the simulations we set $\beta = 0.99; \sigma_c = 3.00; h = 0.70; \sigma_n = 0.70; \epsilon = 7.0; \alpha = 0.6; \rho_r = 0.2; \rho_\pi = 1.30; \rho_y = 0.05; \zeta_p = 0.8$, and $\rho_\chi = 0.5; \rho_z = 0.8; \sigma_\chi = 0.0112; \sigma_z = 0.0051; \sigma_v = 0.0010; \sigma_\mu = 0.2060$, while the standard deviation of the shock driving the unit root component is $\sigma_{z,nc} = 0.0021$ for the technology shock and $\sigma_{\chi,nc} = 0.0221$ for the preference shock. The exact magnitude of the standard deviations of these two shocks is relatively unimportant here because our parameters choice implies that the non-cyclical component has limited importance at business cycle frequencies.

We simulate 1300 data points for four observable variables (y_t, w_t, π_t, r_t), discard 1030 initial observations to eliminate the effect of initial conditions, and use the last 100 for forecasting exercises. This means that the sample size used in estimation is 170.

Figure 1 presents the log-spectrum of filtered output and of filtered and unfiltered inflation for the two DGPs when linear (LT), Hodrick and Prescott (HP), band pass (BP) and first difference filtering (FOD) are used. In each box, the two vertical bars isolate the frequencies corresponding to cycles of 8-32 quarters. Figure 2 reports the autocorrelation function of filtered output and of filtered and unfiltered inflation for the two DGPs.

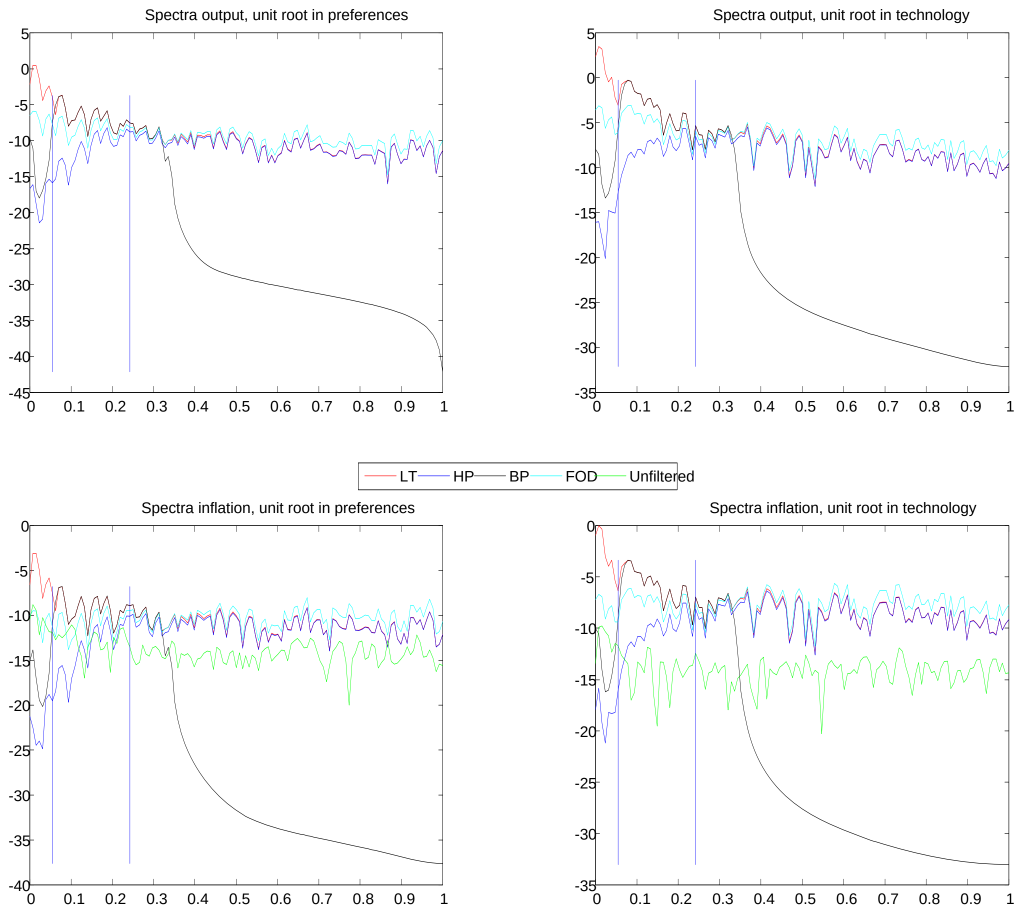

Figure 1: Log Spectra

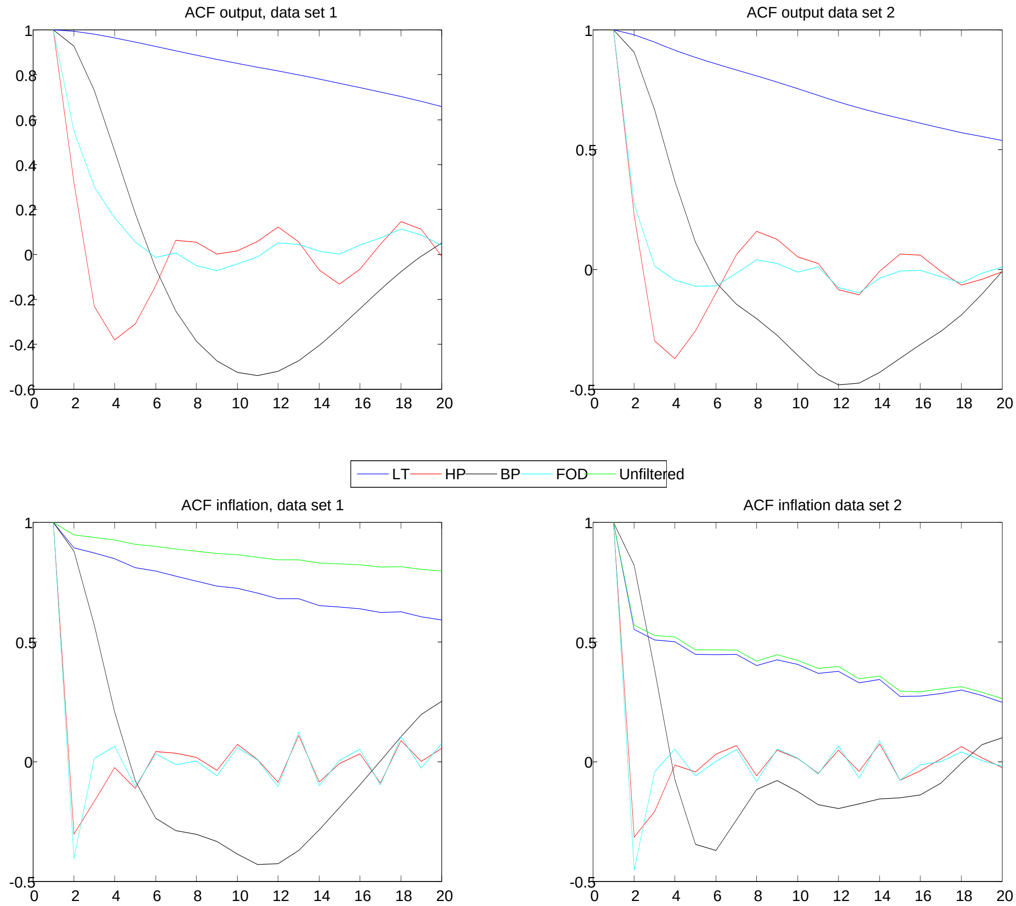

Figure 2: Autcorrelation functions

Figure 1 clearly indicates that all filters imperfectly isolate the power of the series at business cycles frequencies. This is due in part to the nature of the filters (in the case of LT, HP and FOD filtering) and in part to small sample distortions. In general, the estimated cyclical component will be contaminated by measurement errors no matter what filter is used and this error will not only be located in the high frequencies of the spectrum. Consequently, the persistence and the variability of the cyclical component is mismeasured, making estimates of income and substitution effects and of the structural parameters regulating preferences and technologies generally distorted. We show below that this is indeed the case. Figure 1 also shows that the spectral power of the measurement error depends on filter used. Different amounts of measurement errors, in particular the low frequencies of the spectrum, imply that the persistence of the cyclical component is mismeasured more with some filters than others (see figure 2) and this makes the magnitude of the distortions in the estimated structural parameters filter dependent. Finally, figures 1 and 2 demonstrate that inflation is a highly persistent but stationary variable, even when shocks have a unit root component, but the simulated series have about as much power at cyclical as non-cyclical frequencies. Hence, fitting the model to filtered or unfiltered inflation will make a difference for structural parameter estimation except, perhaps, after linear filtering.

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

Table 1: Prior Distribution for the structural parameters.

To show that indeed the measurement error produced by imperfect filtering distorts our ability to understand the features of the true economy, that the amount of distortions depends on the filter used and on whether some or all variables are filtered, we take the two sets of experimental data we have constructed and estimate the structural parameters after passing the raw data with LT, HP, BP and FOD filters. Estimation is conducted using Bayesian methods. We choose loose priors for all the parameters (see table 1) and, to give the best chance to the routine, start estimation at the true parameter values. Posterior estimates are obtained with a random walk Metropolis algorithm, where the jumping variable has a t-distribution with 5 degrees of freedom and variance is tuned up to have an acceptance rate of about 30 percent for each filtering approach. Half a million draws were made for each filtered/DGP combination convergence was checked with a standard CUMSUM statistic and achieved after about 250000 iterations. We keep one out of hundred of the last 100,000 draws to compute statistics of the posterior distribution.

Table 2 reports the median and the standard deviation of the posterior of each of the structural parameters. The top panel refers to the situation when all variables are independently filtered prior to estimation. The bottom panel to the case when only output and the real wage are independently filtered. We only show results obtained when the preferences shock has two components, since those obtained with the other DGP are roughly similar.

The table shows that there are important estimation biases and the magnitude of these biases can be as large as 100 percent. Since measurement error has, in general, important low frequency components, the persistence of the shocks is typically overestimated and the variability of the shocks is typically underestimated and, in some cases, by quite a lot (see e.g. the monetary and the markup shocks). Furthermore, as expected, there are parameters whose posterior distribution depend on the preliminary filtering one employs (see e.g. the coefficient on inflation in the Taylor rule ρ_π, and the share of labor in production α in the first panel). Finally, distortions are generally larger when only real variables are filtered. This happens because the combination of filtered and filtered data " unbalances" the likelihood - some equations become more misspecified than others. Since likelihood based methods produce parameters estimates which minimize the largest discrepancy between the model and the data, biases tend to be larger in this case.

How could eliminate the distortions induced by imperfect filtering? One option is to increase the sample size and consider only filters, which at least asymptotically isolate the frequencies of interest. However, even if longer samples were available, new problems will emerge because time invariance will be difficult to assume. Alternatively, one could choose a non-symmetric non-stationary version of a band pass filter (as suggested by Christiano and Fitzgerald (2003)), which is able to isolate much better the frequencies of interest, even in small samples. However, since band pass filters are two-sided, the properties of estimated parameters and the transmission of shocks could be distorted. Furthermore, the Christiano and Fitzgerald filter is asymmetric and may induce phase shifts, with unpredictable consequences on parameter estimates. One final possibility is to design one-sided filters which minimize the leakage and the compression at the frequencies of interest in small samples. While possible in theory, such an option is not currently available to the applied investigator. All in all, none of these alternatives seems viable.

3 The idea of the paper

Our suggestion is to use the information contained in the cyclical data obtained with different filters to try to average out the low frequency component of the measurement error. In other words, rather than arbitrarily selecting one filter and estimating the model with the resulting filtered data, we treat cyclical data extracted with various filtering methods as contaminated estimates of

Filter		LT	HP	FOD	BP	Factor$_u$	Factor$_r$
	true	Median (s.e.)	Median (s.e.)	Median (s.e.)	Median(s.e.)	Median (s.e.)	Median(s.e.)
				All filtered			
σ_c	3.00	1.63 (0.09)	1.70 (0.14)	1.98 (0.08)	1.87 (0.10)	1.58 (0.11)	2.08 (0.20)
σ_n	0.70	1.56 (0.07)	1.63 (0.10)	1.64 (0.07)	1.49 (0.06)	0.67 (0.16)	0.49 (0.09)
h	0.70	0.64 (0.02)	0.65 (0.02)	0.49 (0.02)	0.64 (0.02)	0.60 (0.04)	0.77 (0.10)
α	0.60	0.32 (0.02)	0.20 (0.04)	0.52 (0.02)	0.15 (0.02)	0.41 (0.03)	0.48 (0.02)
ϵ	7.00	3.99 (0.13)	4.09 (0.14)	4.07 (0.13)	3.85 (0.13)	6.21 (0.12)	6.38 (0.12)
ρ_r	0.20	0.44 (0.05)	0.32 (0.03)	0.53 (0.02)	0.44 (0.04)	0.31 (0.07)	0.26 (0.05)
ρ_π	1.30	2.01 (0.07)	2.05 (0.08)	1.68 (0.05)	2.03 (0.07)	1.50 (0.02)	1.50 (0.05)
ρ_y	0.05	0.11 (0.02)	0.15 (0.02)	0.11 (0.00)	0.18 (0.02)	0.43 (0.03)	0.21 (0.02)
ζ_p	0.80	0.92 (0.03)	0.93 (0.03)	0.87 (0.03)	0.94 (0.03)	0.81 (0.02)	0.79 (0.02)
ρ_χ	0.50	0.98 (0.03)	0.98 (0.03)	1.00 (0.03)	0.99 (0.03)	0.93 (0.01)	0.85 (0.01)
ρ_z	0.80	0.92 (0.03)	0.94 (0.03)	0.92 (0.03)	0.95 (0.03)	0.62 (0.02)	0.59 (0.04)
ω_χ	1.10	0.19 (0.03)	0.25 (0.04)	1.34 (0.45)	0.20 (0.04)	0.86 (0.16)	2.36 (0.48)
ω_z	0.57	0.64 (0.08)	0.59 (0.08)	3.67 (0.28)	0.19 (0.03)	0.66 (0.23)	0.47 (0.12)
ω_v	0.12	0.05 (0.01)	0.05 (0.01)	0.06 (0.01)	0.05 (0.01)	0.08 (0.01)	0.09 (0.01)
ω_μ	20.64	6.42 (0.36)	11.25 (0.82)	6.35 (0.23)	3.77(0.23)	5.17 (0.64)	6.38 (1.00)
				Real variables filtered			
σ_c	3.00	1.92 (0.07)	1.89 (0.07)	1.93 (0.07)	1.96 (0.09)	1.98 (0.22)	1.73 (0.26)
σ_n	0.70	2.10 (0.08)	2.11 (0.08)	2.04 (0.08)	2.03 (0.09)	0.68 (0.17)	0.69 (0.09)
h	0.70	0.58 (0.02)	0.58 (0.02)	0.52 (0.02)	0.62 (0.02)	0.68 (0.02)	0.76 (0.03)
α	0.60	0.48 (0.02)	0.47 (0.02)	0.52 (0.02)	0.43 (0.02)	0.66 (0.04)	0.57 (0.04)
ϵ	7.00	3.71 (0.13)	4.21 (0.15)	4.00 (0.13)	3.73 (0.15)	6.27 (0.10)	6.37 (0.12)
ρ_r	0.20	0.53 (0.04)	0.55 (0.05)	0.42 (0.01)	0.19 (0.03)	0.44 (0.03)	0.33 (0.09)
ρ_π	1.30	1.26 (0.05)	1.32 (0.05)	1.01 (0.03)	1.22 (0.04)	1.53 (0.04)	1.53 (0.05)
ρ_y	0.05	-0.17 (0.01)	-0.04 (0.02)	-0.00 (0.00)	-0.14 (0.02)	0.30 (0.06)	0.17 (0.05)
ζ_p	0.80	0.76 (0.02)	0.76 (0.03)	0.68 (0.02)	0.72 (0.02)	0.82 (0.05)	0.84 (0.04)
ρ_χ	0.50	1.00 (0.03)	1.00 (0.03)	1.00 (0.03)	1.00 (0.03)	0.96 (0.01)	0.96 (0.01)
ρ_z	0.80	0.84 (0.03)	0.90 (0.03)	0.87 (0.03)	0.85 (0.03)	0.95 (0.02)	0.95 (0.01)
σ_χ	1.10	0.09 (0.01)	0.21 (0.05)	3.07 (0.14)	0.16 (0.02)	1.07 (0.20)	1.66 (0.21)
σ_z	0.57	0.32 (0.04)	0.25 (0.03)	6.55 (0.24)	0.12 (0.01)	0.33 (0.07)	0.49 (0.21)
σ_{mp}	0.12	0.07 (0.01)	0.06 (0.01)	0.05 (0.01)	0.06 (0.01)	0.08 (0.01)	0.09 (0.01)
σ_μ	20.64	13.29 (0.64)	16.11 (1.03)	8.03 (0.26)	12.70 (0.63)	5.36 (0.88)	6.94 (1.02)

Table 2: Parameters estimates using different filters; preference shocks with two components.

an unobservable cyclical component and use the information provided by different filters jointly in the estimation of the structural parameters. As long as the measurement error is close to be idiosyncratic across filtering methods, less distortions should be present and more precise estimates of the cyclical features of the economy should be obtained. In this sense our approach builds on ideas of Boivin and Giannoni (2005), who suggest that a data rich environment can help to estimate the structural parameters of a DSGE model and more precisely forecast out-of-sample.

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

$$\begin{aligned} x_{1t} &= RR(\theta)x_{2t-1} + SS(\theta)x_{3t} && (9) \\ x_{2t} &= PP(\theta)x_{2t-1} + QQ(\theta)x_{3t} && (10) \\ x_{3t} &= NN(\theta)x_{3t-1} + \iota_t \quad iota_t sim(0, \Sigma(\theta)) && (11) \end{aligned}$$

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

Let x_t^i be the vector of filtered observable time series obtained with method $i = 1, 2, ...g$ and let $x_t^d = [x_t^1, x_t^2, \ldots, x_t^g]'$. Assume the following structure:

$$x_t^d = \lambda_0 + \lambda_1 x_t^m + u_t \tag{12}$$

where λ_0 is a $ng \times 1$ vector of constants, λ_1 a $ng \times n$ matrix of non-structural parameters and u_t is a $ng \times 1$ vector of measurement errors. For estimation purposes, we normalize the $n \times n$ block $\lambda_1^1 = I$ so that the remaining blocks of the matrix λ_1 can be interpreted as loadings relative to those of the first method. Joint estimation of the structural parameters θ and the non-structural parameters λ_j is now possible because (9)-(11)and (12) represent a state space system with the latter being a measurement equation and the former, state equations. Specifically, equations (9)-(11) and equation (12) can be cast into the state space system

$$\begin{aligned} s_{t+1} &= Fs_t + Ga_{t+1} && (13) \\ o_t &= Hs_t + \eta_t && (14) \end{aligned}$$

by setting

$$s_{t+1} = \begin{pmatrix} y_{1t} & y_{2t} & z_{t+1} \end{pmatrix}'$$

$$\begin{aligned}
F &= \begin{pmatrix} 0 & RR & SS \\ 0 & PP & QQ \\ 0 & 0 & NN \end{pmatrix} \\
G &= \begin{pmatrix} 0, & 0, & I \end{pmatrix}' \\
a_{t+1} &= \iota_{t+1} \\
o_t &= \left(x_t^i/\sigma_x^i - \lambda_0^i, i = 1, 2, \ldots, g \right)' \\
H &= \begin{pmatrix} \lambda_1 S & 0 \end{pmatrix} \\
\lambda_1 &= \left(\lambda_1^i, i = 1, 2, \ldots, g \right)' \\
\eta_t &= \left(u_t^i, i = 1, 2, \ldots, g \right)'
\end{aligned}$$

The likelihood of (13)-(14) can be computed with the Kalman filter. In our context the vector of parameters is ν of interest includes θ and $(\lambda_0^i, \lambda_1^{i+1}, \sigma_\eta^k,\ i = 1, 2, \ldots, g;\ k = 1, \ldots, ng$, which comprises both structural and non-structural parameters. If Bayesian estimation is preferred, the non-normalized posterior distribution of ν, can be obtained with Monte Carlo Markov Chain simulators. For example, one can employ the following algorithm, which appears to give reasonable results in estimation. Starting from an initial value $\nu^{\ell-1}$, given a Σ, and a prior $g(\nu)$:

1. Draw a shock vector υ from $t(0, \Sigma, 5)$ and construct a candidate $\nu^* = \nu^{\ell-1} + \upsilon$
2. Solve the model using ν^*; if the solution is indeterminate or no solution is found set $\mathcal{L}(\nu|o) = 0$. Otherwise, evaluate the likelihood of the observables o_t at ν^* $\mathcal{L}(\nu^*|o)$ with the Kalman filter.
3. Calculate $\breve{g}(\nu^*|o) = g(\nu^*)\mathcal{L}(\nu^*|o)$ and the ratio $MR^* = \frac{\breve{g}(\nu^*|o)}{\breve{g}(\nu^{\ell-1}|o)}$
4. Draw ς from $U[0, 1]$; if $MR^* > \varsigma$ set $\nu^\ell = \nu^*$, otherwise set $\nu^\ell = \nu^{\ell-1}$

Iterated a large number of times, the algorithm ensures that the sample $(\nu^{\bar{L}}, \nu^{\bar{L}+1}, \ldots,)$ for an appropriately chosen $\bar{L}$ is a draw from the target distribution that we need to sample from (for further details see Canova (2007a)).

In (12) different cyclical estimates x_t^i are treated as contaminated proxies of the true cyclical component. They are contaminated in two senses: they introduce fluctuations which are non-cyclical according to the definition we have used; they compress the power of the spectrum of the series at cyclical frequencies. The amount of information they contain for the model relevant concepts of cyclical fluctuations is measured by the vector λ_0 and the matrix λ_1. Ideally, λ_0 is a vector of zeros and λ_1 a matrix with the identity in each $n \times n$ block, so that each measure is an unbiased and perfectly correlated although noisy signal of the true cyclical component. In general, we expect either $\lambda_0^i \neq 0$ or $\lambda_1^{i+1} \neq I$, or both, for some or all i's. Since we have normalized λ_1^1, estimates of λ_1^{i+1} gives us an idea of the amount of correlation distortions each method displays relative to the first.

This setup is advantageous in, at least, two respects. First, since we do not have to arbitrarily choose one filtering approach prior to the estimation, we avoid specification errors. In addition, the output of one-sided and two-sided filters as well as the output of univariate and multivariate procedures can all used as observables for estimation, as long as the list of filters is sufficiently rich. Second, if different filters have sufficiently different features, measurement error may have different time series properties. Since the implicit information averaging that our procedure produces may reduce measurement error and eliminate part of its cyclicality, estimates of the cyclical components will be more reliable, structural parameter estimates better shielded from filtering errors and inference ore robust.

It is important to stress that, in this paper, we make two important assumptions. First, we assume that the model generating x_t^m is correctly specified; that is, there are no missing variables or shocks. When this is not the case, the interpretation of the λ's becomes more difficult. Second, we assume that the cyclical and the non-cyclical components are uncorrelated. While the majority of models used in the literature employ this assumption, the presence of such a correlation may introduce additional biases, which are neglected in this paper.

The literature is largely silent about the issues we address in this paper. Cogley (2001) and Gorodnichenko and Ng (2007) are concerned with the problem of estimating the structural parameters of a cyclical DSGE when the trend specification is incorrect, but they do not investigate what are the consequences of small sample filtering nor their implications for the structural estimates. Giannone et al. (2006) emphasize that if the variables of the model are measured with error, the solution has a natural factor structure and exploit this feature to compare responses obtained from VAR and factor models. Rather then considering a factor structure for the endogenous variables in terms of the states, we construct an estimable structure where vectors of filtered observable data have a factor structure in terms of a subset of the variables of the model. However as in Giannone et al., we emphasize that important measurement errors with low frequency components may exist. The paper which is closest in spirit to ours is Boivin and Giannoni (2005). Their main point is that the model variables do not have an exact counterpart in the real world and that some external indicators to the model may have important information for interesting variables. The point here is somewhat similar. The cyclical component of the model does not have an exact counterpart in the data because none of the existing filtering approaches is able to exactly extract fluctuations with 8-32 quarters average periodicity. Moreover, different cyclical vectors may have idiosyncratic error components and this error may be averaged out with our signal extraction approach.

3.1 How does the procedure fares with simulated data?

We estimate the structural parameters of the model using the suggested approach and the same experimental data used in section 2. As input in our procedure, we employ the vector of LT, HP and FOD filtered data. Although this is not the best choice, as LT and HP filtered data show considerable similarities, and the procedure works best if the vector of filtered data one selects are somewhat idiosyncratic in their spectral properties, it is sufficient to illustrate the features of the procedures, even in this less than ideal situation. Thus, the vector of observable variables is 12×1 (the model produces implications for four variables and there are three filtering methods). We employ the same Bayesian approach used in section 2, assuming the same priors on the structural parameters shown in table 1 and loose priors on the non-structural parameters entering (12). In particular, we assume that the prior for each element of λ_0 is normally distributed, centered at zero with variance equal to 0.5; the prior for the free diagonal elements of λ_1 is normal, centered at 1 with variance 0.5: and the prior for the standard deviation of the u_t's is inverted gamma with mean equal to 0.0056 and variance equal to 0.002.

We consider two specifications: one where the non-structural parameters are filter and series specific (in this case there are 32 non-structural parameters to be estimated) and another where the constants and the loadings in (12) are common across series for each filter (in this case, there are 17 non-structural parameters). We refer to the first specification as the unrestricted factor model; the second one to the restricted factor model.

The last two columns of table 2 present the posterior median and the posterior standard deviation for the structural parameters obtained with the two factor specifications, when the preference shock has two components. In general, the biases present in the first four columns of the table have been reduced: economic parameters are all better estimated with reasonably small standard deviation and auxiliary ones, although still biased, are closer to the true ones than those obtained with standard approaches. Note that, when all variables are filtered, the persistence of the stationary preference shocks is now estimated to be lower. When only real variables are filtered the reduction is negligible, primarily because inflation and the nominal rate, which are the most persistent series, enter three times among the observables unfiltered. The variability of the structural shocks is also better estimated in both frameworks except for the markup shocks, but this bias has more to do with the weak identification of this parameter than with the properties of both procedures.

To see how these estimates compare with the true ones and with those obtained with standard approaches in terms of economically meaningful statistics, we computed the unconditional autocorrelation function of the cyclical components of output, real wages, inflation, and the nominal rate, where by this we mean the component generated by the non-unit root shocks when the posterior median estimates of the parameters obtained when all variables are filtered are used. When we examine traditional approaches, we only report the autocorrelation of LT and FOD filtered data to make the plots readable - the plots produced by HP and BP are similar.

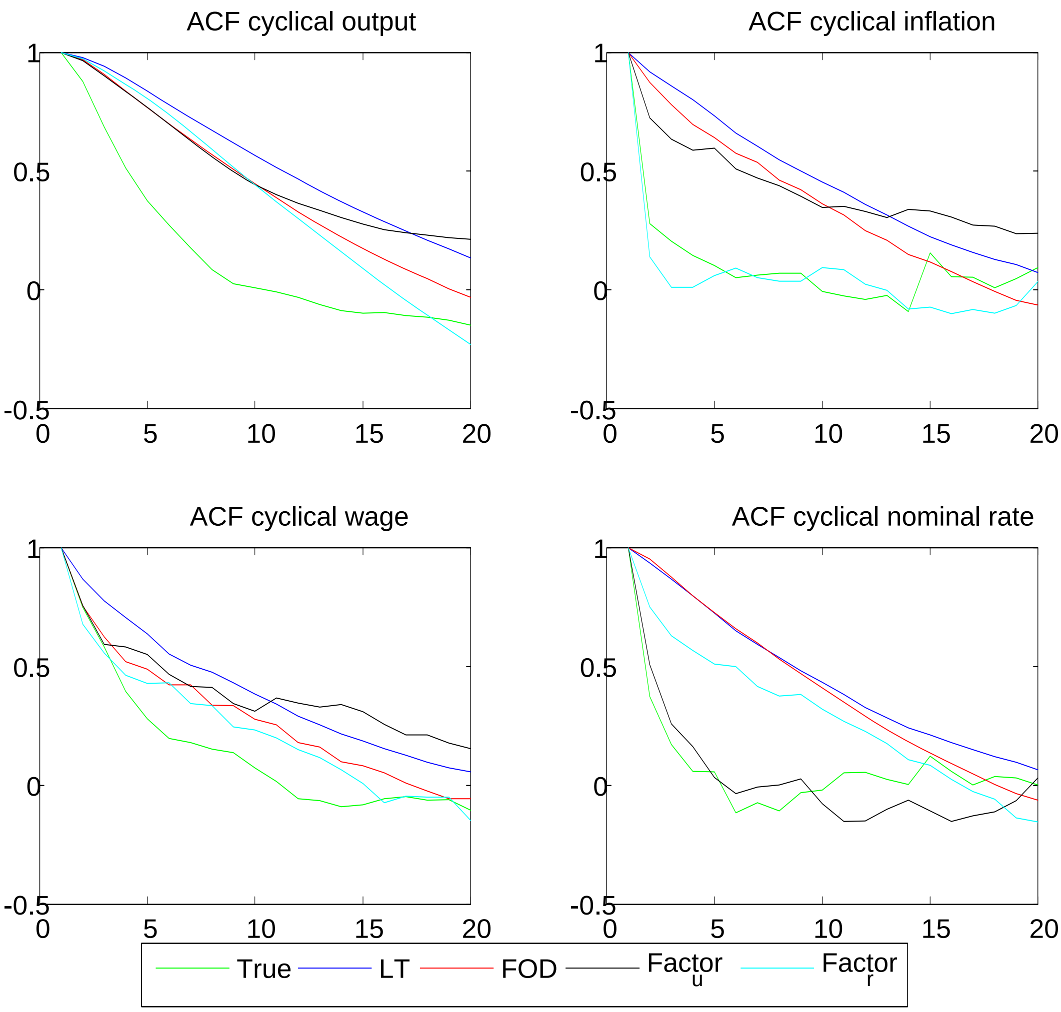

Figure 3: Autocorrelation of cyclical components

Figure 3, which plots the autocorrelation functions, reinforce previous conclusions. With both LT and FOD filtered data, the estimated autocorrelation functions are too persistent relative to the true ones. The same is true also for the autocorrelation function produced by the unrestricted factor model, except for the nominal rate. Here the large dimensionality of the parameter space relative to the size of data set plays an important role in determining the quality of the results. In contrast, the autocorrelation function produced by the mean estimates of the restricted factor model decays much faster for all series and, except for output, we can not reject the hypothesis that the true and the estimated autocorrelation functions are the same.

Our approach also helps in reproducing the unconditional standard deviations of the series better. Table 3, which reports these statistics, indicates that both factor models average out a good part of the measurement error that standard procedures introduce. Once again, the restricted factor model is preferable to the unrestricted one.

Series	True	LT	FOD	Factor_u	Factor_r
Output	0.25	0.54	3.70	0.33	0.32
Real wage	0.17	0.21	1.58	0.11	0.16
Inflation	0.11	0.13	1.23	0.11	0.06
Nominal rate	0.17	0.21	1.59	0.11	0.16

Table 3: Standard deviation of the cyclical components; simulated data.

The generally good performance of the restricted factor model is confirmed when looking at the responses of the endogenous variables to the four structural shocks. Figure 4 presents the responses produced with the true parameters and those generated with the posterior median estimates obtained with the restricted factor model. Overall, both the shape and the persistence of the conditional responses are well captured. There are a few cases where the impact sign is wrong (see e.g. the response of the real wage to technology shocks or the response of inflation to monetary shocks) but differences are not large a-posteriori. Also, in a few cases the magnitude of the responses is not well estimated but with 170 data points, this is far from surprising.

To gain additional evidence on the properties of the estimated cyclical components obtained with traditional methods and with the approach we suggest, we have also conducted two forecasting exercises: an unconditional and a conditional one. In the first case, we compute the sequence of one step ahead forecast errors for output and inflation, when we take as parameter values the posterior median estimates obtained using LT and FOD filtered data and our unrestricted and restricted factor approaches, setting all the shocks in the forecasting period to zero. The MSE is computed over 100 forecasting periods, when no updating of the parameters in the forecasting sample is performed, and appears in table 4.

Series	LT	FOD	Factor_u	Factor_r
Output	0.1049	0.2054	0.0340	0.0058
Inflation	0.3501	0.3814	0.5317	0.3407

Table 4: Mean Square Error of the unconditional forecasts; simulated data. Scale 10^{-3}.

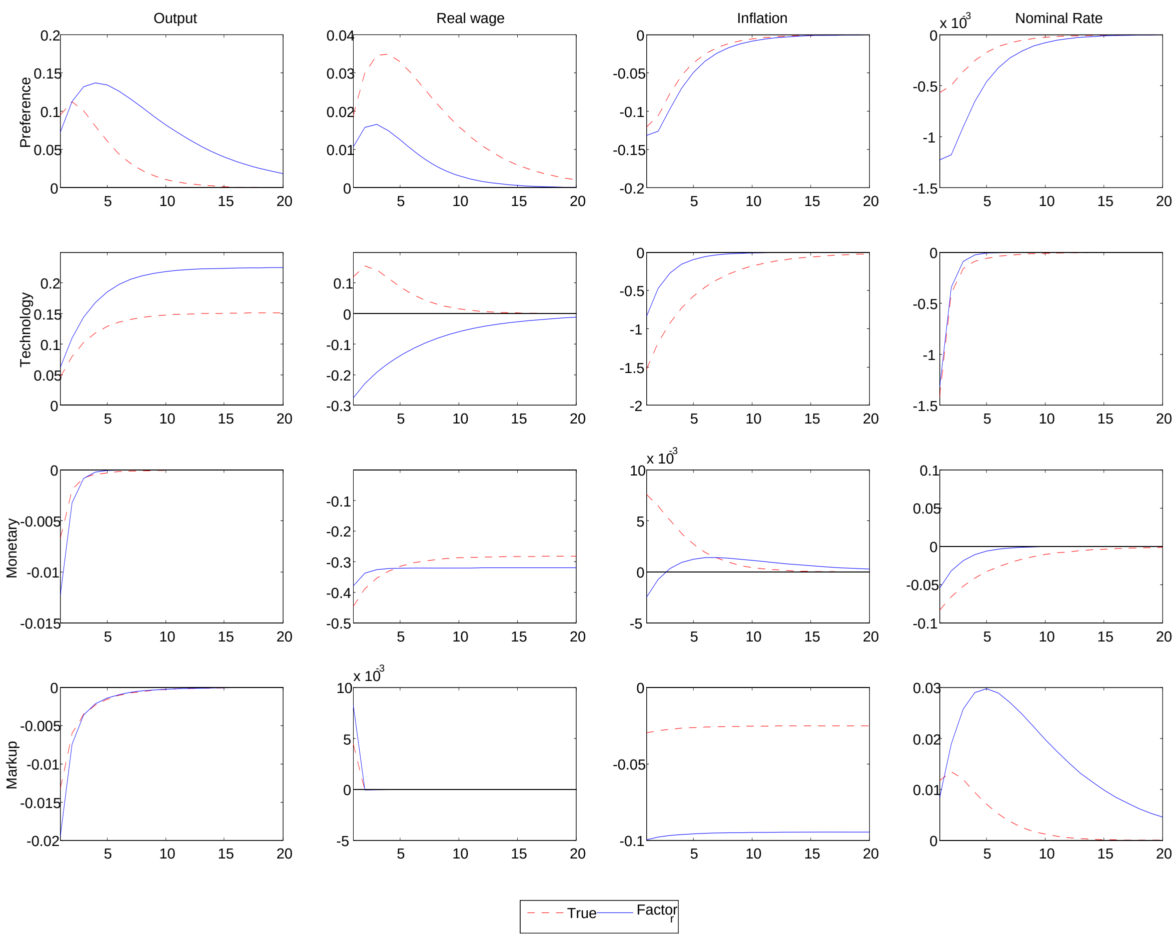

Figure 4: Impulse responses

Figure 5 instead traces out the one-step ahead path of output and inflation that would have obtained with posterior median estimates of the parameters when monetary shocks were drawn so as to keep the nominal interest rate fixed over the forecasting path. That is, we allow the nominal interest rate to endogenously react to output and inflation but make sure that the monetary shocks we draw are such that the nominal rate is constant over the forecasting path and equal to the value taken at the date prior to the forecasting period (time 0 in the figure).

Table 4 and the Figure 5 indicate that the differences in the estimates lead to important differences in forecasting performance. The restricted factor model unconditionally forecasts one-step ahead the cyclical component of output better than any standard approach and even the unrestricted specification is superior to procedures which first filter, then estimate the parameters and then forecast. For inflation, the results are less clear cut, but the restricted factor specification is at least as good as the LT and the FOD filtered specifications. The conditional forecasting exercise shows that the bias introduced by traditional procedures translates in conditional output forecasts which are consistently above those produced by our approach and the differences are statistically significant. For inflation differences with standard methods are less evident but, for example, one can clearly see that the cyclicality of the one step ahead conditional inflation series is smaller with the restricted factor specification than with standard approaches and the differences are, at times, large.

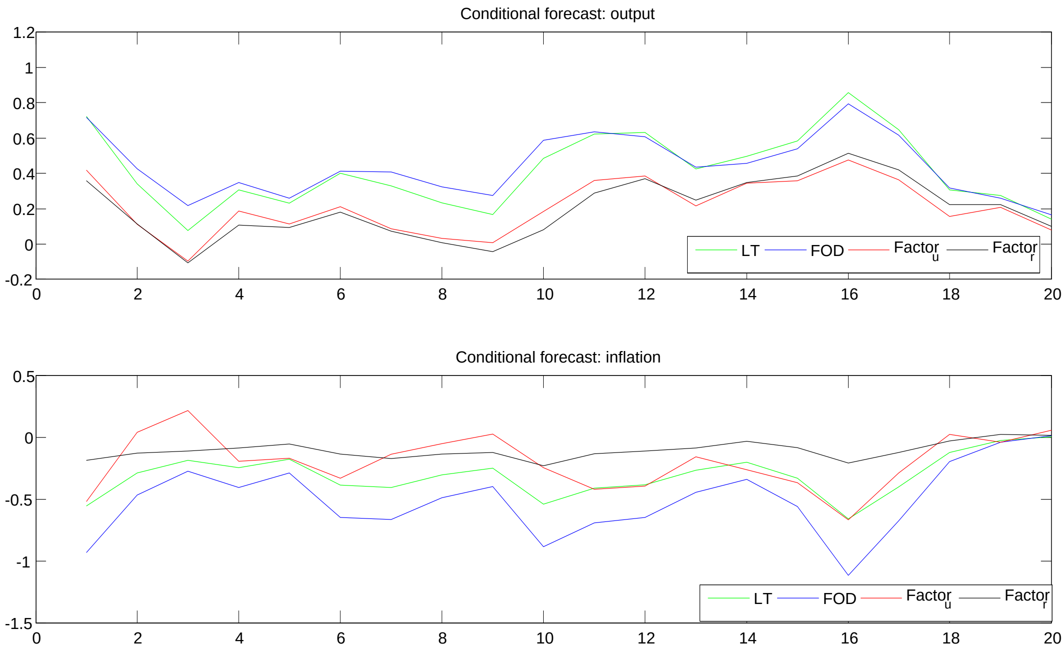

Figure 5: One step ahead conditional forecasts

In sum, the biases that a standard procedure induces in parameter estimates have important consequences for our understanding of both the autocovariance properties of the cyclical component of the data and of the conditional responses to shocks. Overall, both statistics appear to be much better reproduced with the specification we suggest The conditional and unconditional forecasts produced by standard approaches inherit and magnify parameter biases providing a distorted picture of the cyclicality of, e.g., actual inflation. These problems are either resolved or considerably reduced with our suggested approach.

4 Does money matter in transmitting monetary business cycles?

To show that the additional information our procedure uses is 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 the majority of the cases these models make no reference whatsoever to monetary aggregates, and when they do, they use a specification where a money demand function determines how much money needs to supplied, given predetermined levels of output, inflation and the nominal rate. As a consequence, changes in the nominal (or real) quantity of money play no direct or indirect role in shaping the dynamics of output and inflation.

Ireland (2004) has constructed a general specification in the class of textbook New Keynesian models, where real balances may have a role in affecting the dynamics of output and inflation. He estimated the relevant parameters by likelihood techniques using post 1980 US data and found evidence supporting current theoretical practices. To construct the likelihood of his cyclical model, he first transforms the actual data, taking away a separate linear trend from per capita GDP and per-capita real balances and demeaning inflation and the nominal interest rate.

In this section, we conduct a similar exercise using post 1959 US data and the cyclical versions of real per-capita output, real par-capita money balances, inflation, and nominal rate series obtained with 8 different filtering procedures. As a benchmark, we also estimate the model employing Ireland's preferred transformation and the same post 1959 sample.

4.1 The model economy

The model we employ is similar to the one considered by Ireland (2004), except that it also permits real balances to play an indirect role, via its effects on interest rate determination. Relative to the model we have used in section 2, we allow the real stock of money to potentially matter for the determination of the output and inflation; consider pricing frictions in the form of adjustment costs to changing prices rather than with a standard Calvo setup; and set the habit parameter to zero.

Since the economy is quite standard we only briefly describe its features. There is a representative household, a representative final good producing firm, a continuum of intermediate goods-producing firms supplying the differentiated commodity $i \in [0,1]$ and a monetary authority. At each t the representative household maximizes

$$E_t \sum_t \beta^t \chi_t [U(c_t, \frac{M_t}{p_t e_t}) - \eta n_t] \tag{15}$$

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

$$M_{t-1} + T_t + B_{t-1} + W_t h_t + D_t = P_t c_t + \frac{B_t}{R_t} + M_t \tag{16}$$

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

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

$$y_t^i = (\frac{p_t^i}{p_t})^{-\epsilon} y_t \tag{17}$$

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

The intermediate good producing firm i, hires n_t^i unit of labor from the representative household to produce y_t^i units of intermediate good i using the production function $y_t^i = z_t n_t^i$, where z_t is an aggregate productivity shock. Intermediate goods substitute imperfectly for one another in producing finished goods. Hence, intermediate firms can set the price of their good but must satisfy (17) at the chosen price. We assume a quadratic cost in adjusting prices, measured in finished goods, given by

$$\frac{\phi}{2}(\frac{p_t^i}{\pi^s p_{t-1}^i} - 1)^2 y_t \tag{18}$$

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

$$E \sum_t \beta^t \chi_t [U_1(c_t, \frac{M_t}{p_t e_t})](\frac{D_t^i}{p_t}) \tag{19}$$

subject to (17), where $\beta^t \chi_t U_1(c_t, \frac{M_t}{p_t e_t})$ measures the marginal utility value to the household of an additional unit of profits t and real dividend are

$$\frac{D_t^i}{p_t} = (\frac{p_t^i}{p_t})^{1-\epsilon} y_t - (\frac{p_t^i}{p_t})^{-\epsilon}(\frac{w_t y_t}{z_t}) - \frac{\phi}{2}(\frac{p_t^i}{\pi p_{t-1}^i} - 1)^2 y_t \tag{20}$$

The monetary authority sets the nominal interest rate according to

$$R_t = R_{t-1}^{\rho_r} y_{t-1}^{(1-\rho_r)\rho_y} \pi_{t-1}^{(1-\rho_r)\rho_\pi} \Delta M_t^{(1-\rho_r)\rho_m} v_t \tag{21}$$

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

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

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

$$\hat{y}_t = E_t\hat{y}_{t+1} - \omega_1((\hat{R}_t - E_t\hat{\pi}_{t+1}) - (\hat{\chi}_t - E_t\hat{\chi}_{t+1})) + \omega_2((\hat{m}_t - \hat{e}_t) - (E_t\hat{m}_{t+1} - E_t\hat{e}_{t+1})) \quad (22)$$

$$\hat{m}_t = \gamma_1\hat{y}_y - \gamma_2\hat{R}_t + (1 - (R^s - 1)\gamma_2)\hat{e}_t \quad (23)$$

$$\hat{\pi}_t = \beta E_t\hat{\pi}_{t+1} + \psi(\frac{1}{\omega_1}\hat{y}_t - \frac{\omega_2}{\omega_1}(\hat{m}_t - \hat{e}_t) - \hat{z}_t) \quad (24)$$

$$\hat{R}_t = \rho_r\hat{R}_{t-1} + (1-\rho_r)\rho_y\hat{y}_{t-1} + (1-\rho_r)\rho_\pi\hat{\pi}_{t-1} + (1-\rho_r)\rho_m(\Delta\hat{m}_t + \hat{\pi}) + \hat{v}_t \quad (25)$$

where

$$\omega_1 = -\frac{U_1(c_t, \frac{m_t}{e_t})}{y^s U_{11}(c_t, \frac{m_t}{p_t e_t})} \quad (26)$$

$$\omega_2 = -\frac{m^s}{e^s}\frac{U_{12}(c_t, \frac{m_t}{e_t})}{y^s U_{11}(c_t, \frac{m_t}{e_t})} \quad (27)$$

$$\gamma_1 = (R^s - 1 + \frac{y^s r^s \omega_2}{m^s})(\frac{\gamma_2}{\omega_1}) \quad (28)$$

$$\gamma_2 = \frac{R^s}{(R^s-1)(m^s/e^s)}(\frac{U_2(c_t, \frac{m_t}{e_t})}{(R^s-1)e^s U_{12}(c_t, \frac{m_t}{e_t}) - R^s U_{22}(c_t, \frac{m_t}{e_t})}) \quad (29)$$

$$\psi = \frac{\epsilon - 1}{\phi} \quad (30)$$

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

The log-linearized Euler condition (equation (22)) includes terms involving real money balances and the money demand shocks. They drop out from the expression if and only if utility is separable in consumption and real balances, i.e $U_{12} = 0$ (see equation (27)). Similarly, real balances play a role in the forward looking Phillips curve (equation (24)), as long as $\omega_2 \neq 0$, which in turn again implies that $U_{12} \neq 0$ is necessary for real balances to matter. Hence, real balances play a direct role in determining output and inflation if and only if real balances and consumption enter non-separably in the utility function. On the other hand, the posited policy rule implies that the growth rate of nominal balances may be an important determinant of output and inflation indirectly, via interest rate determination. When ω_2 and ρ_m are zero real balances have no direct or indirect role in propagating cyclical fluctuations.

Since our scope here is illustrative, in the next subsection we focus attention on the estimates of ω_2 and ρ_m, only.

4.2 Estimation

We estimate the model with US data from 1959:1 to 2008:2. All the data comes from the FRED data bank at the Federal Reserve Bank of Saint Louis and it is seasonally adjusted. For real GDP we take the GDPC96 series, which is a chain weighted real value of domestic production, convert it in per-capita terms dividing it by the civilian non-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 series (TB3M) is our measure of interest rates.

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

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

We draw 500,000 elements of the MCMC chain using the algorithm described in section3. Convergence was achieve in less than 100,000 draws for each model specification we present. Posterior statistics are computed using one every 100 of the last 200,000 draws.

As a benchmark, we have estimated the parameters of interest using as vector of observables linearly detrending pre-capita output and per-capita real balances, and the demeaned value of inflation and the nominal rate. We present the results obtained when we allow for measurement error in each equation. This is the right specification to employ for comparison purposes since our approach has idiosyncratic error built in (12) even though, when DSGE models are estimated, measurement error is typically left out of the specification.

4.3 The results

Before we present the results of interest we briefly comment on the outcomes of the estimation of the non-structural parameters in our procedure. First, the vector of λ_0 is estimated, for all purposes, to be zero with very small standard errors. Therefore, all filtered data do not present level biases relative to the cyclical components produced by the model. Second, the loadings parameters are estimated to be between 0.60 (for the UC filtered data) up to 0.86 (with the CD filtered data). Therefore, there is sufficient idiosyncratic information in the data produced by the various procedures. Since posterior standard errors are tight, differences in the loadings between any of the procedures we consider are a-posteriori relevant. Third, the error u_t appears to have a highly idiosyncratic variance, both across series and across filtering methods. This clearly reflects the fact that the variability of an individual series depends on the filtering approach used and that filtered series display very different amounts of cyclical information. This is the reason for why, for example, a restricted version of the setup we use, where only one parameter characterizes the variability across series or across filtering methods, produces a poor fit.

Table 5 presents the marginal likelihood of the unrestricted specification, where both direct and indirect effects of money are allowed, and for three restricted specifications, where either the direct effect is eliminated ($\omega_2 = 0$), the indirect effect is eliminated $\rho_m = 0$, or both are eliminated.

Clearly, a specification where both effects are present is preferable to the other specifications

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

Table 5: Marginal likelihood and Posterior estimates, various specifications

and restricting both $\rho_m = 0$ and $\omega_2 = 0$ is preferable to restricting only $\rho_m = 0$. Overall, at least in terms of model fit, both the direct and indirect effects that money has in the model are important. This result is confirmed when looking at location measures of the posterior of the two parameters. Our specification implies a economically moderate direct and indirect effects of money on output and inflation fluctuations. Statistically, both parameters are estimated tightly and both are a-posteriori different from zero. A standard specification with measurement error, on the other hand, implies that both direct and indirect effects are quite small, and for all purposes, they can be set to zero.

How important are the differences in the point estimates with have in table 5? Figure 6 presents responses to the four shocks for our specification and a standard specification with measurement errors. Responses look qualitatively similar in the two cases, but there are differences in the magnitude and the persistence of the responses to shocks. In particular, the persistence of the responses to monetary shocks is reduced, the one of the technology shocks is increased and the responses to money demand shocks have both different magnitude and persistence relative to a standard approach.

In sum, our estimates are statistically significant. Given the experimental evidence we have collected in the previous section, it is also likely that they are less biased than those obtained with a standard approach, as far as persistence of the shocks and measurement of the substitution and income effects are concerned. From an economic point of view, our estimates suggest that to correctly understand how monetary business cycles are generated, money must be given a meaningful and intuitive way in the model.

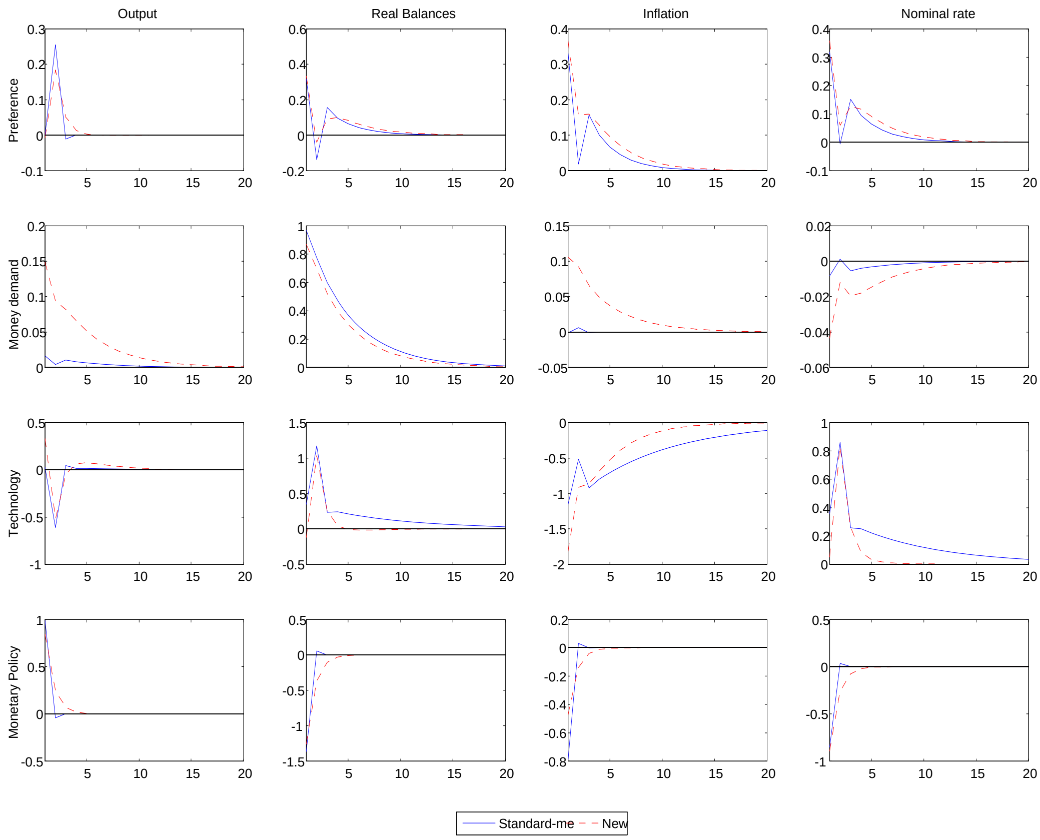

Figure 6: Impulse responses, standard and new approaches

5 Conclusions

This paper proposes a method to estimate the structural parameters of a time invariant cyclical DSGE model using multiple cyclical information. The approach borrows ideas from the recent literature employing data-rich environments to estimate DSGE models (see Boivin and Giannoni (2005)) to set up an estimated structure where vectors of filtered data obtained with alternative procedures are treated as contaminated estimates of the true cyclical component. Measurement error may have different features and different power at different frequencies, depending on the filtering approach used. We set up a signal extraction framework where the cyclical DSGE model is the unobservable factor, vectors of filtered data are contaminated observable proxies, and the parameters of the DSGE model are jointly estimated together with non-structural parameters 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 the estimation, we avoid specification errors of various sorts. Moreover, cyclical data obtained with one-sided and two-sided filters of both univariate and multivariate nature can be used in estimation, as long as the list of filters is sufficiently rich. Second, if different filters have sufficiently different features, measurement error may have different time series properties. Since the implicit information averaging our procedure produces may reduce measurement error and eliminate its cyclicality, estimates of the cyclical components are more reliable and precise, making parameter estimates and inference, to a large extent, free of preliminary data transformation biases. The only constraint to the number of the vectors of filtered data used in the estimation is the RAM capacity of the computer.

We investigate the properties of our approach using experimental data. We show that estimating the structural parameters of the model with just one arbitrary filter typically induces large biases in the estimates and that these biases are considerably reduced with our approach. We also show that, in an unconditional forecasting exercise, the one step ahead MSE produced by our approach is smaller than the MSE obtained with standard procedures and that the biases the latter possesses translate in conditional forecasts which are considerably distorted.

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

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

Given these problems, this paper offers a setup where specification and measurement error biases could be reduced making estimates of time invariant cyclical DSGE models more meaningful. In this sense, this work complements those of Canova (2008) and Ferroni (2008) who also provided new methodologies to reduce specification and small sample errors in the estimation of cyclical DSGE models. Future work in the area will include revisiting known puzzles in the macroeconomic literature, and investigating whether they can be solved with the approach proposed in this paper, and better understanding the properties of the procedure using interesting experimental designs.

References

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

Boivin, J. and Giannoni, M. (2005) "DSGE estimation with data rich environments", NBER working paper

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

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

Canova, F. (2007b) " How much structure in empirical models?", forthcoming in Mills, T. and Patterson, K. (eds.), *Palgrave Handbook of Econometrics*, Palgrave.

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

Canova, F. and Sala, L. (2006): Back to square one: identification issues in DSGE models, ECB working paper 544.

Canova, F., Lopez Salido, D. and Michelacci, C., (2007), The Labor Market Effects of Technology Shocks, CEPR working paper 6365.

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

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

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

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

Ferroni, F. (2008) One and two step estimation approaches for DSGE models: which one is more appropriate?, manuscript

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.

Gorodichenko, V. and Ng, S. (2007) "Estimation of DSGE models when data are persistent" Columbia University, manuscript.

Kehoe, P. (2007) " How to advance theory with structural VARs: the Sims-Cogely-Nason approach", NBER working paper 12575.

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

Rabanal, P., and J. Rubio-Ramirez (2005): "Comparing New Keynesian models of the business cycle: A Bayesian approach," *Journal of Monetary Economics*, 52, 1151-1166.

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.