Centre for Economic Policy Research

BANCO DE ESPAÑA
Eurosistema

European Summer Symposium in International Macroeconomics (ESSIM) 2008

Hosted by
Banco de España

Tarragona, Spain; 20-25 May 2008

Estimating DSGE Models with Unfiltered Data

Fabio Canova

We are grateful to the Banco de España for their financial and organizational support.

The views expressed in this paper are those of the author(s) and not those of the funding organization(s) or of CEPR, which takes no institutional policy positions.

Estimating DSGE models with unfiltered data

Fabio Canova
U-Bern, ICREA-UPF, AMeN and CEPR

May 19, 2008

Abstract

I propose a method to estimate cyclical DSGE with unfiltered data. The approach links the observables to the model conterparts via a flexible specification, in which the non-cyclical component can take various time series patterns, and do not assume that the cyclical component is solely located at business cycle frequencies. I show the fallacies of applying standard data transformation for structural estimation and explain the reasons for distortions they produce. The approach recovers the features of the cyclical component in selected experimental designs.

Very preliminary draft. Please, do not quote

JEL classification: E32, C32.

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

1 Introduction

There has been considerable development in the specification and estimation of DSGE models over the last 10 years. The original structure, featuring a single technological disturbance, has been enriched with the addition of shocks and frictions and our understanding of the propagation mechanism of important structural shocks considerably enhanced. Steps forward have also been made in the estimation of such models. While a few years ago it was standard to informally calibrate the structural parameters, now researchers routinely use limited and full information estimation procedures and, perhaps more importantly, this trend is shared by applied economists in both academic and policy circles (see, e.g., [23], [19], [25],[17], [22] among many others).

Despite recent developments, structural estimation of DSGE models is conceptually and practically difficult. For example, classical estimation is asymptotically justified only when the model is assumed to be the DGP of the actual data, up to a set of serially uncorrelated measurement errors, and standard validation exercises are meaningless without such an assumption (see [7] for an approach that does not require it). Furthermore, identification problems appear to be widespread (see [8]) and numerical difficulties often plague estimation exercises, making the outcomes of unrestricted analyses dubious. Finally, the vast majority of the models nowadays used in the literature is intended to explain only the cyclical portion of the fluctuations in the observables, rather than the observables themselves, and the latter contains many types of fluctuations, some of potentially displaying non-stationarities of various sorts.

There are several reasons why researchers prefer to work with models designed to explain only the cyclical fluctuations of the data. First, jointly accounting for cyclical and non-cyclical fluctuations is still an ambitious task and there is still plenty to be learned by focusing attention on cyclical fluctuations. Second, there are very few known theoretical mechanisms which are able to propagate for a sufficiently long time temporary shocks (we need e.g. R&D as in [10] or Schumpeterian creative destruction process as in [11]). Third, from the computation and the interpretation point of view, it is convenient to assume that the mechanism driving cyclical and non-cyclical fluctuations are distinct and orthogonal.

The fact that models are designed to explain only cyclical fluctuations but that the data contains much more than that seems to bother little developers of DSGE models but creates important headaches to applied investigators. Leaving aside issues of model singularities - there are typically less shocks than endogenous variables in the model - and of measurement errors - the variables in the model do not typically have an exact counterpart in the actual data - applied researchers typically proceed in one of the following two ways.

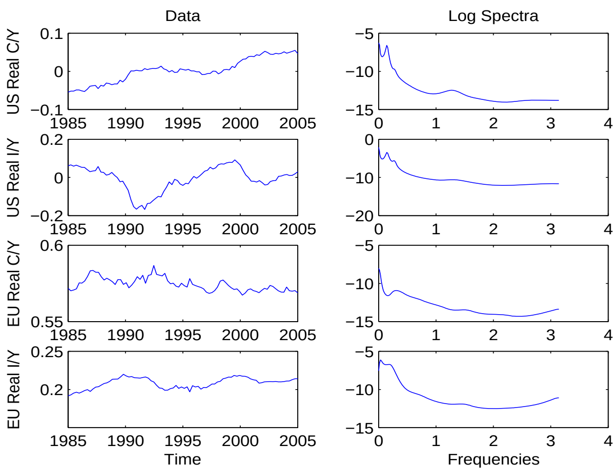

Figure 1: Real Great Ratios in US and EU and their spectral density

- Filter actual data and fit the model to the filtered data. Such an approach leaves open the questions of what are the cyclical features of the data, of whether univariate or multivariate methods should be used in extracting the relevant fluctuations, and of whether all or only to a subset of the variables should be filtered. In fact, many filters approximate what are typically called business cycle fluctuations, i.e. fluctuations with, say, 8-32 quarters average periodicity (see e.g. [4]). However, they all produce contaminated estimates of these fluctuations and it is hard to design criteria which effectively discriminate "good" from "bad" approaches in the samples macroeconomists typically work with. For example, empirical versions of Band Pass and Hodrick and Prescott filters only approximately capture the power of the spectrum at business cycle frequencies and taking growth rates greatly emphasize the high frequency content. Similarly, while it is typically to filter each series separately, there are both theoretical and empirical reasons to believe that the non-cyclical component should be similar across series and that therefore some multivariate procedure should be used to impose long run consistency. Along the same lines, one may wonder whether only real or all variables should be filtered. There are arguments in favour of both approaches: many models imply cyclical fluctuations for e.g., inflation and the nominal rate, even when most of the shocks are non-cyclical; however, not all actual fluctuations in nominal variables could be safely considered as cyclical. Finally, while usually not appreciated, the cyclical component produced by the majority of the filters can be represented as a symmetric, two-sided moving average of the raw data. Hence, the timing of information is altered by filtering, making standard dynamic analyses with structural parameter estimates difficult to interpret. In sum, rather than resolving the issue, such an approach multiplies the number of difficulties applied researchers face.

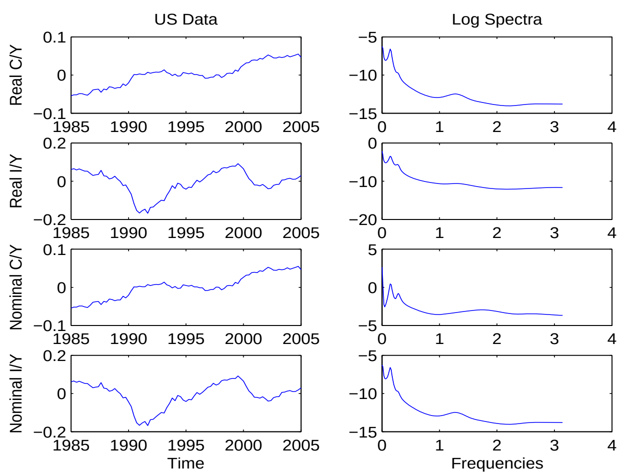

Figure 2: Real and Nominal Great ratios, US and their spectral density

- Take into account the dynamics of the non-cyclical component and use data transformations which are likely, at least in theory, to be void of non-cyclical components, e.g. consider the ratios C/Y, I/Y in the estimation (this is the suggestion contained in [9] and [20]) or their nominal counterpart, i.e. P_cC/P_yY or P_iI/P_yY, where P_j is the price of good j, if relative prices matter (this is the suggestion contained in [27]). While very few have exploited this kind of properties in structural estimation, both transformations are unlikely to resolve the mismatch issue because, as figures 1 and 2 shows, real and nominal "Great ratios" show visible trends over the last 30-40 years, both in the US and in the EU.

This paper shows that the data transformation one uses prior to estimation matters for estimates of the structural parameters and the interpretation of the transmission of structural shocks. Therefore, unless one takes a strong but unwarranted view of what the model is about, e.g. the model is a representation of HP filtered data, or, simulated BP filtered data correspond to actual BP filtered data, it is hard to select among various structural estimates and build knowledge on the economic phenomena on solid ground. There are two main reasons why the preliminary data transformation matters for estimation. First, as already mentioned, the data used in estimation is only a contaminated estimate of the cyclical fluctuations having the required average periodicity. Even an ideal band pass filter, once implemented on finite stretches of data, lead to considerable leakage and compression of the power that the spectrum displays at business cycle frequencies. Since the both amount and the properties of the resulting measurement errors are filter dependent and, as sample size grows, may vanish at different speeds, the association between the model and the estimated business cycle fluctuations is very imperfect. An approach to deal with this type of error which exploits ideas developed in [3]have been recently proposed in [6].

The second, and probably more important reason, is that the theoretical cyclical component that a DSGE model produces does not correspond to the cyclical component extracted by existing filters, even when the sample size is large. This is because the filters used by applied investigators implicitly assume that different components of the data have power exclusively concentrated at certain frequencies of the spectrum, e.g. the cyclical component has power only at business cycle frequencies, while both the cyclical and the non-cyclical component a model generates have power at all frequencies of the spectrum. This implies for example, that at the so-called business cycle frequencies, both components may matter and it is not very difficult to build examples, where most of the power at the business cycle frequencies is due to the former. Aguiar and Gopinath (2007) have argued that for LDC countries this is an important concern. What I show here is that the problem is general and is, in fact, relevant for the estimation of the structural parameters with the data on any country. The only condition that needs to be met is that the variance of the shocks driving the cyclical and the non-cyclical components are roughly of the same magnitude.

As an alternative, this paper proposes a methodology to estimate the structural parameters of a "cyclical" DSGE model using raw data by creating a flexible link between the two that allows both the non-cyclical and the cyclical components to have power at all the frequencies of the spectrum. Since the specification I use encompasses, as special cases, situations where the non-cyclical component of the data displays deterministic, stochastic or smooth features, the approach does not requires to take a stand on, for example, whether the raw data is trend-stationary or difference stationary (as is done e.g. in [9] and [14]) and therefore shields the empirical analysis from specification errors. Finally, while for expositional reasons, drivers of the cyclical and the non-cyclical components of the data are assumed to be orthogonal, there is no conceptual difficulties in considering cases, where shocks driving the two components are partially or perfectly correlated.

The approach uses a setup similar to the one employed in the unobservable component time series literature. In the specific case I consider, the cyclical component is what the DSGE model produces, given a set of structural parameters; the non-cyclical component is arbitrarily but flexibly modelled; high frequency noise is permitted; and the spectrum of the data is endogenously split in various components. This means that the procedure jointly estimates structural and non-structural parameters and this permits coherent inference and meaningful forecasts on both the components and the raw data.

I show using a simple experimental design that the procedure can effectively capture the cyclical components of the DGP when the data has more than cyclical fluctuations and lead to good estimates of the structural parameters and of the transmission of shocks, using samples of the size currently available in macroeconomics. I also show that the economic interpretation of the resulting transmission of structural shocks is different from those obtained with standard methods.

Throughout the paper I make the important simplifying assumption that the model is correctly specified; that is, there are no missing variables or omitted cyclical shocks. While this issue is important in practice and semi-structural methods of the type suggested by [7] are likely to produce more robust inference, I find it useful to keep this problem separate from the one of matching cyclical models to the raw data since the latter occurs independently of whether the model correctly represents the DGP of the data or not.

The rest of the paper is organized as follows. The next section presents a simple model, estimates its structural parameters after a number of preliminary data transformations and discusses why estimation and interpretation results differ. Section 3 presents the methodology. In section 4 experimental data is used to estimate the structural parameters with the proposed approach. Section 5 shows estimates the model with US data and the proposed methodology and compares estimates with those presented in section 2. Section 6 concludes. An appendix containz details about the model used and the estimation approach.

2 Structural estimation with transformed data

To show how parameters estimates of a "cyclical" DSGE model depend on the data transformation employed, I consider a rather standard New-Keynesian model where agents face a labor-leisure choice, there is external habit in consumption, production is carried out with labor, firms must confront an exogenous probability of price adjustments and monetary policy is conducted with a conventional Taylor rule. Details on the structure of the model are in the Appendix. The log linearized equilibrium conditions are:

$$\lambda_t = \chi_t - \frac{\sigma_c}{1-h}(y_t - hy_{t-1}) \tag{1}$$
$$y_t = z_t + (1-\alpha)n_t \tag{2}$$
$$mc_t = w_t + n_t - y_t \tag{3}$$
$$mrs_t = -\lambda_t + \sigma_n n_t \tag{4}$$
$$mrs_t = w_t \tag{5}$$
$$r_t = \rho_r r_{t-1} + (1-\rho_r)(\rho_\pi \pi_t + \rho_y y_t) + \epsilon_t^r \tag{6}$$
$$\lambda_t = E_t(\lambda_{t+1} + r_t - \pi_{t+1}) \tag{7}$$
$$\pi_t = k_p(mc_t + \epsilon_t^\mu) + \beta E_t \pi_{t+1} \tag{8}$$
$$\chi_t = \rho_\chi \chi_{t-1} + \epsilon_t^\chi \tag{9}$$
$$z_t = \rho_z z_{t-1} + \epsilon_t^z \tag{10}$$

where $k_p = \frac{(1-\beta\zeta_p)(1-\zeta_p)}{\zeta_p}\frac{1-\alpha}{1-\alpha+\varepsilon\alpha}$, λ is the Lagrangian on the consumer budget constraint, mc_t are marginal costs, mrs_t is the marginal rate of substitution between consumption and leisure, z_t is a technology shock, χ_t a preference shock, ϵ_t^r is an iid monetary policy shock and ϵ^μ an iid markup shock. The structural parameters are: σ_c the risk aversion coefficient, σ_n the elasticity of labor supply, h the coefficient of consumption habit, $1-\alpha$ the share of labor in production, ρ_r the degree of interest rate smoothing, ζ_p the probability of not changing prices, ε the elasticity among consumption varieties. The auxiliary parameters are: ρ_χ, ρ_z the autoregressive parameters, and the standard deviations σ_i of the four structural shocks.

I assume that there are four observable variables: output, the real wage, the inflation rate and the nominal interest rate (y_t, w_t, π_t, r_t), and I examine a variety of filtering approaches, applied to all or a subset of the variables. In the search, I have considered (i) a case where only real variables are independently filtered and nominal variables demeaned; (as in [21], [22] or [25]) (ii) a case where all variables are independently filtered (as in [17]); (iii) a case where all variables are jointly filtered with a multivariate deterministic approach, which takes into account potential commonalities in the non-cyclical component, and (iv) a case where real variables appear as ratios - $\log y_t - \log n_t, \log w_t - \log n_t$, where h is hours worked - - and all variables are demeaned. This transformation of real variables is selected because, according to the model, it should be void of non-cyclical fluctuations no matter what the time series properties of the shocks are. For the first two cases, I consider three different filtering approaches, which cover pretty much the range of filters used in the empirical DSGE literature and, as shown in [4]), produce cyclical components whose average periodicity is different (linear filtering (LT), HP filtering, and first difference (FOD) filtering). Since the first two approaches belong to the class of two-sided moving averages and may therefore alter the timing of the information of the data, I have also experimented with either recursive or one-sided versions of these filters. None of the results I report below are due to this nevertheless important problem.

Estimation is conducted using Bayesian methods. Details on the approach are in the Appendix. Posterior estimates are obtained with a Random Walk Metropolis algorithm, where the vector of jumping variables is t-distributed with 5 degrees of freedom, and the variance tuned to have an acceptance rate of roughly 30-35 percent for each data transformation considered. One million draws are made for each case-filter combination and convergence was checked using standard CUMSUM graphs. Since convergence to the erogodic distribution is rather slow in all cases and draws are highly serially correlated, I keep one every hundred of the last 100,000 draws to compute posterior statistics. The priors for the parameters are kept fixed in the exercise and are reported in the Appendix.

Filter	LT	HP	FOD	Ratio
	Median (s.e.)	Median (s.e.)	Median (s.e.)	Median (s.e.)
σ_c	2.19 (0.10)	2.24 (0.13)	2.56 (0.14)	1.69 (0.11)
σ_n	1.81 (0.07)	1.59 (0.10)	1.93 (0.18)	2.16 (0.10)
h	0.67 (0.01)	0.60 (0.02)	0.43 (0.03)	0.64 (0.02)
α	0.18 (0.02)	0.12 (0.02)	0.11 (0.02)	0.13 (0.02)
ε	4.84 (0.16)	4.19 (0.11)	3.26 (0.06)	4.09 (0.12)
ρ_r	0.16 (0.04)	0.50 (0.03)	0.26 (0.04)	0.22 (0.04)
ρ_π	1.27 (0.04)	1.66 (0.09)	1.86 (0.06)	1.71 (0.05)
ρ_y	-0.15 (0.02)	0.15 (0.04)	0.12 (0.05)	-0.02 (0.01)
ζ_p	0.77 (0.01)	0.60 (0.03)	0.35 (0.03)	0.81 (0.01)
ρ_χ	0.73 (0.02)	0.58 (0.04)	0.28 (0.04)	0.81 (0.02)
ρ_z	0.95 (0.01)	0.59 (0.03)	0.86 (0.04)	0.92 (0.01)
σ_χ	0.22 (0.03)	0.34 (0.05)	0.23 (0.04)	0.95 (0.16)
σ_z	0.12 (0.02)	0.08 (0.01)	0.09 (0.01)	0.08 (0.01)
σ_r	0.11 (0.01)	0.08 (0.01)	0.12 (0.02)	0.12 (0.01)
σ_μ	18.88 (0.82)	1.15 (0.30)	0.17 (0.03)	34.70 (1.04)

Table 1: Posterior estimates using different filters. For LT, HP, FOD real variables detrended, nominal demeaned. For Ratio, real variables are in terms of hours and all variables demeaned.

Table 1 reports a summary of the estimation results. Here I report the median and the standard deviation of the posterior distributions for cases (i) and (iv). Results for cases (ii) -(iii) are similar and available on request from the author. Clearly, there are parameters whose posterior distribution is considerably affected by the preliminary data transformation (see e.g. ζ_p, the persistence and the volatility of the shocks, and the parameters of the monetary policy rule). Also, since posterior

standard errors ar
Posterior difference
where the agents li
shocks in the LT a
price stickiness mu

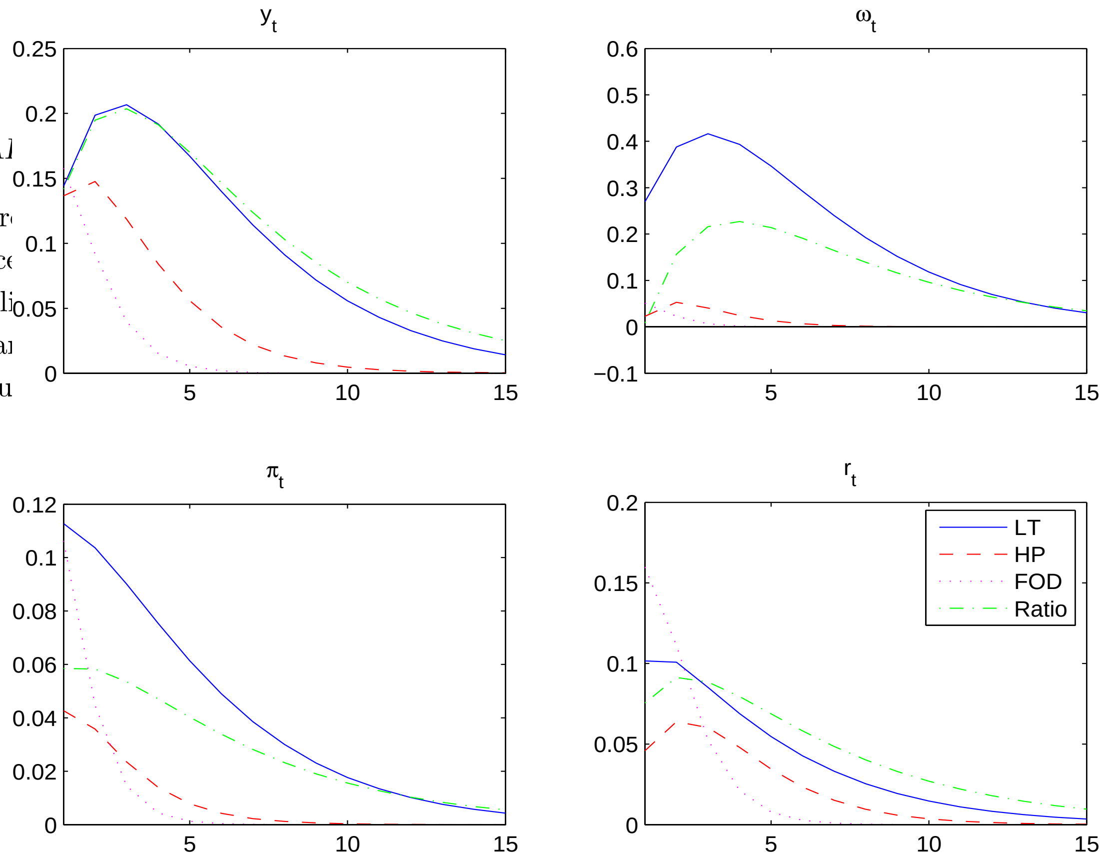

Figure 3: Responses to preference shocks, different filtering

The importance of the preliminary data transformations is not only limited to parameter estimates but also concerns functions of crucial interest in the literature. To illustrate this I have examined the transmission of structural shocks using the median estimates obtained in the four columns of table 1. Figures 3 and 4 present responses to preference and technology shocks. Clearly, not only the magnitude of the impact coefficients but also the persistence of the responses depend on the preliminary data transformation employed. Furthermore, at least in the case of technology shocks, responses have different signs depending on the preliminary data transformation used.

Responses obtaine
monetary and marl
they are statistical

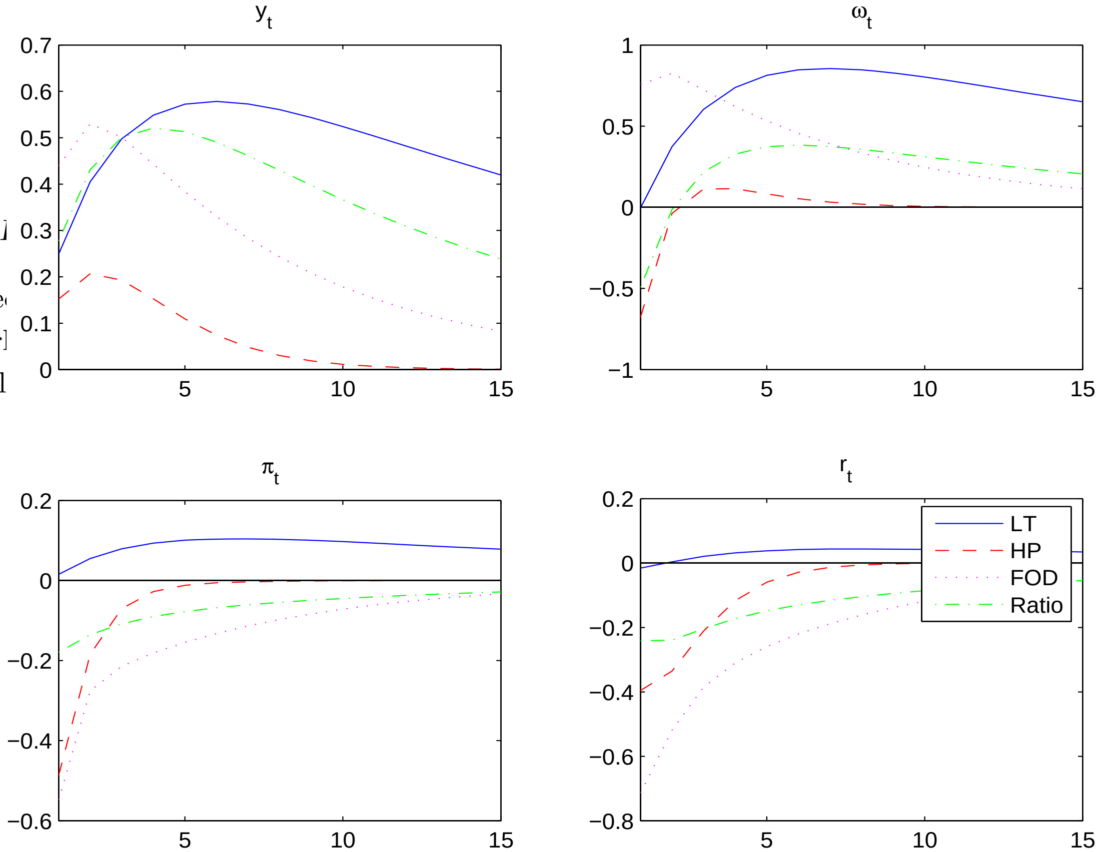

Figure 4: Responses to technology shocks, different filtering

Differences in estimation results should be expected; the growth rate of GDP is a very different object and has very different time series properties than, for example, a linearly detrended GDP. These differences would be inconsequential, if researchers would have a good reason to prefer one approach over the other. However, as emphasized in [4]), a good reason is hard to find since economists know little about the properties and the features of the non-cyclical component.

The solution favoured in the literature is to define the length of the cycles the model is assumed to explain, i.e. interesting fluctuations are those with an average periodicity between 8 and 32 quarters, and design procedures which allow to extract the information present at these frequencies in the actual data. It is not particularly difficult to show that the approaches I considered only approximately isolate cycles with the required periodicities: the linear trend specification leaves both long and short cycles in the filtered data; the HP filter has the features of a high pass filter and therefore leaves high frequencies variability unchanged; the growth transformation emphasizes high frequency noise and downweights the importance of cycles with a business cycle periodicity; great ratios leave important low frequency fluctuations in the data (see figures 1 and 2).

The ideal band pass filter is capable of isolating fluctuations at the frequencies of interest, but finite samples induce significant approximation errors also in this case (see e.g. [5], ch.3). Since the cyclical components extracted with different filters are, at best, contaminated estimates of the fluctuations with 8-32 quarters average periodicity and measurement errors play a different role at different frequencies, choosing the length of the cycles the model is assumed to explain does not solve the problem of which of the columns of table 1 one should use for inference. In [6] the fact that all filtered data represents a noisy measure of the fluctuations the literature in interested in analyzing is exploited to suggest an estimation approach which may reduce or eliminate measurement errors.

There is an additional and perhaps more important reason for why estimation results are transformation dependent. All the filters used in the estimation literature assume that, say, the cyclical and the non-cyclical fluctuations produced by a DSGE model are located at different frequencies of the spectrum and that the economic mechanism generating the two is distinct. Such an assumption is crucial to identify the frequencies corresponding to fluctuations with 8 to 32 quarters periodicity with the theoretical cycle and frequencies corresponding to fluctuations with periodicity larger than 32 quarters to the theoretical trend. However, such an association is problematic: the cyclical component of a DSGE model may have power at frequencies other than those corresponding to 8 to 32 quarters and the non-cyclical component may induce important fluctuations at the so-called business cycle frequencies. Time series econometricians have known this for a long time. For example, [24], [13] and [15] have all emphasized the fallacy of trying to estimate structural models using seasonally adjusted data, precisely for this reason. However, this fact is not well known or it is simply disregarded among applied macroeconomists. Therefore, it is worthwhile to illustrate the problem using the model represented by (1)-(10) assuming, for the sake of presentation, that the preference disturbance has unitary autoregressive coefficient. In this case, the theoretical non-cyclical component of the four observable variables corresponds to the fluctuations generated by this shock and the theoretical cyclical component corresponds to the fluctuations induced by the other three shocks. Using the parameter values reported in the last column of table 3, I have simulated a long sample (11,000 data points), discarded 1000 initial observations and passed the experimental data through LT, HP, FOD and BP filters. Figures 5 and 6 show the spectrum of the filtered output data, the contribution of the two theoretical components at each frequency, and their contribution to the autocorrelation function of filtered output.

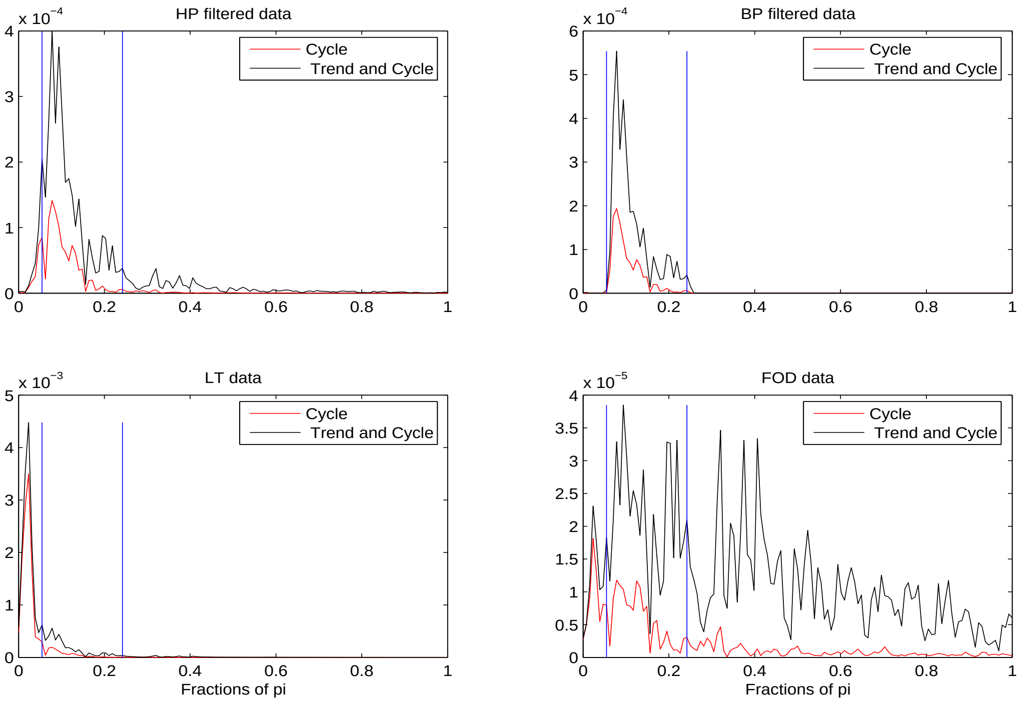

Figure 5: Model based output spectra, different filtering. Vertical bars indicate the frequencies where cycles with 8-32 quarters periodicities are located

Three features are clear from these figures. First, the two components have power at all frequencies of the spectrum of output. Second, at the so-called business cycle frequencies, the theoretical non-cyclical component plays an important role. This fact is apparent from the plot of the autocorrelation function - the autocorrelation induced by the two components is very similar - and becomes obvious when the BP filter is used - here the sample is sufficiently large that leakage and compression effects are minor. [1] and [2] have claimed that for LDC countries the trend is the cycle. Figures 5 and 6 show is that the problem potentially applies to the transformed data of any country. It is only required that the variability of the shock driving the two components is roughly of similar magnitude. The HP filter assumes that the variability of the two components is different with the trend beeing much smoother, but other filters, e.g. a Beveridge and Nelson one, imply similar variability of the shocks driving the two components and estimates reported in [4] tend to favour the latter assumption. Third, the noise induced by various filtering approaches is generalized and certain transformations (e.g. LT and FOD) produce cyclical components in which the role of the theoretical non-cyclical component is emphasized. In sum, if one is interested in estimating structural parameters and in conducting policy analyses, the use of filtered data is problematic. Biases in parameter estimates may be severe because the cyclical fluctuations produced by a DSGE model do not correspond with the business cycle fluctuations extracted by standard filters.

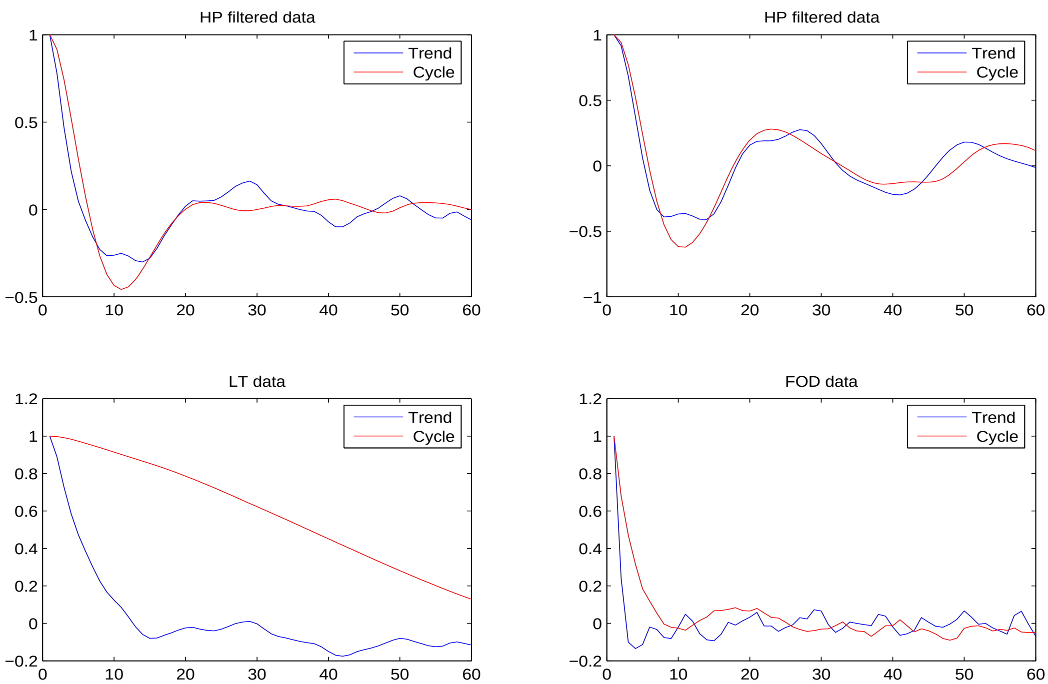

Figure 6: Autocorrelation function of filtered cyclical and non-cyclical component

To understand why biases appear and the direction they will take, consider the following decomposition. Using the Bayes theorem, it is easy to show that the posterior distribution of the structural parameters is proportional to log-likelihood of the parameters, given the data, multiplied by the prior. In turn, following [15], the log-likelihood of the parameters can be represented as the sum of three terms $L(\theta|y_t) = A_1(\theta) + A_2(\theta) + A_3(\theta)$ where $A_1(\theta) = \frac{1}{\pi}\sum_{\omega_j} \log\det G_\theta(\omega_j)$, $A_2(\theta) = \frac{1}{\pi}\sum_{\omega_j}$ trace $[G_\theta(\omega_j)]^{-1}F(\omega_j)$, $A_3(\theta) = (E(y) - \mu(\theta))G_\theta(\omega_0)^{-1}(E(y) - \mu(\theta))$, and where $\omega_j = \frac{\pi j}{T}, j = 0, 1, \ldots, T-1,$, $G_\theta(\omega_j)$ is the model based spectral density matrix of y_t, $\mu(\theta)$ the model based mean of y_t, $F(\omega_j)$ is the data based spectral density of y_t and $E(y)$ the unconditional mean of the data. The first term is the sum of the one-step ahead forecast error matrix across frequencies; the second a penalty function, which emphasizes deviations of the model-based from the data-based spectral density at various frequencies, and the third another penalty function, weighting deviations of model-based from data-based means, with the spectral density matrix of the model at frequency zero.

Suppose that the actual data is filtered so that frequency zero is eliminated and low frequencies deemphasized. Then the log-likelihood of the parameters is composed of $A_1(\theta)$ and of $A_2 * (\theta) = \frac{1}{\pi}\sum_{\omega_j}$ trace $[G_\theta(\omega_j)]^{-1}F(\omega_j)*$, where $F(\omega_j)* = F(\omega_j)I_\omega$ and I_ω is an indicator function describing the effect of the filter at various frequencies. Suppose that $I_\omega = I_{[\omega_1,\omega_2]}$, an indicator function for the business cycle frequencies, as in an ideal BP filter. Then the penalty $A_2 * (\theta)$ bites only at these frequencies and since here $[G_\theta(\omega_j)] < F(\omega_j)*$ (see figure 5), the penalty exceeds one. Since $A_2 * (\theta)$ and $A_1(\theta)$ enter additively in the log-likelihood function, two types of biases are present in estimates of θ. First, estimates $F_\theta * (\omega_j)$ only approximately capture the features of $F(\omega_j)*$ at the required frequencies - the sample version of $A_2 * (\theta)$ has a smaller values at business cycle frequencies and a nonzero value at non-business cycle ones. Second, in order to reduce the contribution of the penalty function to the log-likelihood, parameters are adjusted to make $[G_\theta(\omega_j)]$ as close as possible to $F(\omega_j)*$ at those frequencies where $F(\omega_j)*$ is not zero. This is done by allowing fitting errors in $A_1(\theta)$ large at frequencies $F(\omega_j)*$ is zero - in particular the low frequencies. Hence, the volatility of the structural shocks will be overestimated - this makes $[G_\theta(\omega_j)]$ close to $F(\omega_j)*$ at the relevant frequencies - in exchange for underestimating their persistence - this makes $G_\theta(\omega_j)$ small and the fitting error large at low frequencies. This distortion implies that the economy agents perceive has different features than the true economy and their decision rules will be altered. Higher perceived volatility, for example, implies distortions in the aversion to risk and a reduction in the internal amplification features of the model. Lower persistence, on the other hand, implies that perceived substitution and income effects are distorted with the latter typically underestimated relative to the former.

If we take for granted that estimates of $F(\omega_j)*$ are imprecise, even when T is large, there are only two situations when estimation distortions will be minor. First, the non-cyclical component has low power at the business cycle frequencies - in this case the distortions produced by the penalty function are limited. This occurs when the volatility of the non-cyclical component is considerably smaller than the volatility of the cyclical one. Second, the prior eliminates the distortions induced by the penalty function. While priors for DSGE parameters are typically tight and this reduces somewhat the distortions, it is very unlikely that the bias is entirely wiped out since priors are not designed with such a scope in mind.

While not so popular in the DSGE estimation literature, one could also conceive to fit the filtered version of the model to the filtered data. To understand what such an approach does to parameter estimates, note that now the log-likelihood is composed of $A_1 * (\theta) = \frac{1}{\pi}\sum_{\omega_j} \log\det G_\theta(\omega_j)I_\omega$ and $A_2(\theta)$. Suppose that $I_\omega = I_{[\omega_1,\omega_2]}$. Then $A_1 * (\theta)$ matters only at business cycle frequencies while the penalty function is present at all frequencies. Therefore, parameter estimates will be adjusted until misspecification is reduced at all frequencies of the spectrum. If the penalty function is more important in the low frequencies of the spectrum, parameters are adjusted to make $[G_\theta(\omega_j)]$ close to $F(\omega_j)$ at those frequencies. Hence, the log-likelihood is willing to incur in large fitting errors in $A_1(\theta)$ at frequencies where $F(\omega_j)$ does not differ much from $G_\theta(\omega_j)$ - in particular, the medium and high frequencies. Hence, the volatility of the shocks will be generally underestimated in exchange for overestimating their persistence - this will make differences between $G_\theta(\omega_j)$ and $F(\omega_j)$ small at low frequencies. Hence, contrary to what it is commonly thought, this procedure implies that the model is fitted to the low frequencies components of the data. If the low frequencies components are poorly measured, biases could be larger than in the previous case. Also here, cross frequencies distortions imply that agents think they are living in an economy with quite different features than the true one. For example, since less noise is perceived, agents decision rules will imply a higher degree of predictability of simulated time series and higher perceived persistence implies that perceived substitution and income effects are distorted with the latter overestimated.

3 The idea of the paper

One solution to the problems I have highlighted in section 2 is to build a non-cyclical component directly into the model. I have mentioned both theoretical and practical reasons for why researchers may be reluctant to do so. To these one can also add statistical concerns (What time series features should it have? Should it be deterministic or stochastic? Should it be correlated with the cyclical component or not? What economic mechanism may drive its fluctuations?) and specification issues (what happens to estimates if the structure of the non-cyclical component is misspecified? What if there are breaks?). Such issues have been recently addressed in [**?**] and [**?**] and will not be discussed here.

Rather than conditioning the analysis on an essentially arbitrary specification for the non-cyclical component, a choice which will make the analysis vulnerable to specification errors, I will use a flexible setup, in the spirit of [16], which allows the data to endogenously split the spectrum of the observables into cyclical and non-cyclical components and permits the two components to jointly appear at business cycle frequencies.

Let the linearized solution of a DSGE model be of the form:

$$y_t = RR(\theta)x_{t-1} + SS(\theta)z_t \tag{11}$$

$$x_t = PP(\theta)x_{t-1} + QQ(\theta)z_t \tag{12}$$

$$z_{t+1} = NN(\theta)z_t + \epsilon_{t+1} \tag{13}$$

where $PP(\theta), QQ(\theta), RR(\theta), SS(\theta)$ are functions of the vector of structural parameters $\theta = (\theta_1, \ldots, \theta_k)$, $x_t = \tilde{x}_t - \bar{x}$ includes the states and the predetermined variables, $y_t = \tilde{y}_t - \bar{y}$ all other endogenous variables and z_t the disturbances of the model and $\bar{y}, \bar{x}$ are the steady states of $\tilde{y}_t$ and $\tilde{x}_t$

Let $y_t^m(\theta) = S[y_t, x_t]'$, where S is a selection matrix picking out of y_t and x_t those variables which are observable and/or interesting from the point of view of the researcher. Let $y_t^d = \tilde{y}_t^d - E(\tilde{y}_t^d)$ be the log demeaned vector of observables. I assume that y_t^d can be decomposed as

$$y_t^d = c + y_t^T + y_t^m(\theta) + u_t \tag{14}$$

where $c = \bar{y} - E(\tilde{y}_t^d)$, y_t^T is the non-cyclical component of the data, u_t is a iid $(0, \Sigma_u)$ (measurement) noise and y_t^T, y_t^m and u_t are assumed to be mutually orthogonal. Without further restrictions θ would not be identifiable from y_t^d. Therefore, I assume that y_t^T can be represented as

$$y_t^T = y_{t-1}^T + \bar{y}_{t-1} + e_t \qquad e_t \sim \; iid\ (0, \Sigma_e^2) \tag{15}$$

$$\bar{y}_t = \bar{y}_{t-1} + v_t \qquad v_t \sim \; iid\ (0, \Sigma_v^2) \tag{16}$$

The specification in 16 is flexible and can account for several time series specification for y_t^T. For example, if $\Sigma_v^2 > 0$ and $\Sigma_e^2 = 0$, y_t^T is a vector of I(2) processes while if $\Sigma_v^2 = 0$, and $\Sigma_e^2 > 0$, y_t^T is a vector of I(1) processes. Finally, if $\Sigma_v^2 = \Sigma_e^2 = 0$, y_t^T is deterministic, while if both $\Sigma_v^2 > 0$ and $\Sigma_e^2 > 0$ and $\sigma_v^2 \sigma_e^2$ is large, y_t^t is "smooth" and nonlinear. Hence, in (16) are nested, as special cases, the structures which are typically thought to motivate the use of the three filters considered in the previous section.

Given this setup, I let the data endogenously select the specification for the non-cyclical component which is more appropriate for each series and this is done jointly with the estimation of the structural parameters of the model. While in (16) I have assumed that Σ_v^2 and Σ_e^2 that are general matrices one can impose some structure assuming that they are either diagonal (so that the non-cyclical component is series specific) or that they are of reduced rank (so that the non-cyclical component is common across series) and test various specifications using marginal likelihood comparisons (see [12]).

There are at least two advantages that the suggested specification has. First, it is not necessary to take a stand on the time series properties of the non-cyclical component and on the choice of filter to tone down its importance - such a choice may create distortions in the estimates of the structural parameters if specification errors are made. Second, as I will show below using experimental, the cyclical component extracted with the proposed approach is not localized at particular frequencies of the spectrum. In fact, at each frequency, all the three components may have power.

The specification I have assumed implies that cyclical, non-cyclical and measurement error fluctuations are driven by different shocks, a subset of which are structural, in the sense that they have economic interpretation and create cross equation restrictions. While such an assumption may appear restrictive, it is not for two reasons. First, if the non-cyclical component is driven by unmodelled structural shocks, it makes sense to assume that they these shocks are uncorrelated with structural shocks driving in the model. Second, the specification is observationally equivalent one where, for example, the non-cyclical and the cyclical components are correlated. Straightforward calculations in fact indicate that the specification

$$y_t^d = y_t^{T*} + y_t^{M*}(\theta) + u_t \tag{17}$$

$$y_t^{T*} = y_t^T + \bar{y}_t + y^M(\theta) \tag{18}$$

$$y_t^T = y_{t-1} + e_t \tag{19}$$

$$\bar{y}_t = \bar{y}_{t-1} + v_t \tag{20}$$

$$y_t^{M*}(\theta) = y_t^{M\dagger}(\theta) + A(\theta)C(\theta)\bar{y}_t \tag{21}$$

$$y_t^{M\dagger}(\theta) \equiv \tilde{y}_T^M(\theta) - y^M(\theta) - A(\theta)C(\theta)\bar{y}_t =$$

$$C(\theta)x_t^{M\dagger}(\theta) \equiv C(\theta)(\tilde{x}_T^M(\theta) - x^M(\theta) - A(\theta)\bar{y}_t) \tag{22}$$

$$x_t^{M\dagger}(\theta) = A(\theta)x_{t-1}^{M\dagger}(\theta) + B(\theta)\epsilon_t \tag{23}$$

which induces correlation in y_t^{T*} and $y_t^{M*}(\theta)$, is indistinguishable from the point of view of the observed data from the specification I suggest.

3.1 A comparison with existing literature

The empirical literature has mostly avoided to deal with the issue of the mismatch of DSGE models and the data. In a seminal paper, [9] considers estimating structural parameters when the data is trending and the model has little or nothing to say about the properties of the trends. He builds a useful taxonomy of cases, shows the distortions that incorrect assumptions have on the estimates of the parameters and investigates the properties of various estimation methods which can downsize the importance of specification errors. I share with Cogley the concern that economic theory has little to say about modelling non-cyclical components. However, rather than simply distinguishing between trend stationary or difference stationary fluctuations, and attempting to robustify inference, I am concerned with the generic mismatch between the empirical and the model-based concept of cyclical fluctuations and in designing a procedure which can be applied without taking a stand on whether the data is trend or difference stationary.

[14] extend Cogley study suggesting a particular transformation delivering robust estimates of the structural parameters when the trend specification one arbitrarily build into the model, is potentially misspecified. The suggested approach shares with Gorodnichenko and Ng the idea of joint estimating structural and auxiliary parameters without specifying the DGP of the data. Their study however differs from the one of this paper is several respect. First, they use minimum distance estimators of the structural and non-structural parameters while I use likelihood based estimators. Minimum distance estimators in DSGE models are subject to severe identification

problems which limits the credibility of the estimation results and of the inferential conclusions (see [8]). Second, rather than assuming an arbitrary trend for the technology, I assume that the model is built to explain only the cyclical component of the data - a much more common assumption in macroeconomics - and link model and observables through a flexible specification. Third, while with minimum distance estimators, data needs to stationary for estimates of the structural parameters to enjoy standard asymptotic properties, my approach does require stationary data.

[1] and [2] have recently pointed out that in emerging market economies variations in trend growth are as important as cyclical fluctuations in explaining the dynamics of macroeconomic variables. While the first is primarily interested in characterizing the difference between emerging and developing economies and find a common mechanism to explain the evidence, the latter explicitly suggests to build trends into the model to limit the distortions caused in policy analyses because of misspecified structural estimates. This paper shows that the problems they highlight are generic and that there is a way to link models to the data which does not require to take a stand on the time series properties of the non-cyclical component.

Finally, [18] investigate the distortions introduced by infrequent switches in trend growth on the estimated parameters of a DSGE model and propose an estimation approach which can deal with these breaks within a standard state space formulation.

3.2 An experiment

To show the properties of the approach, I simulated data from the model of section 2 assuming that the preference shock has two components, a nonstationary one and a stationary one (the properties of the other three shocks are unchanged), using the same parameters described in the first column of table 3 for two specifications of the variance of the non-cyclical shock relative to the variance of the other shocks: a large one and a small one.

I then estimate the parameters of the model and those of the flexible non-cyclical part of the data using the suggested specification and the same Bayesian approach I have used in section 2. The priors for the non-structural parameters are given in the appendix. With the true and the estimated parameters I then compute the model-based cyclical component, and calculate the autocorrelation function and the log spectrum of output after I have passed it through the same 4 filters employed in figures 5 and 6. The true and the estimated parameters are reported in table 2. The filtered log spectrum and the cyclical autocorrelation functions are in figures 7 and 8.

Three features of table 2 are interesting. First, regardless of the variance of the non-cyclical component estimates of the structural parameters are roughly unchanged. Hence, the flexible link I have specified can adapt to capture different features of the non-cyclical component of the data. Second, parameter estimates are precise but the median is not necessarily on top of the true parameter value. This is particularly evident for those parameters are weakly identified from

	Small variance			Large variance		
	True	Median	(s.e)	True	Median	(s.e)
σ_c	3.00	2.77	(0.34)	3.00	2.60	(0.34)
σ_n	0.70	0.39	(0.08)	0.70	0.30	(0.06)
h	0.70	0.70	(0.03)	0.70	0.53	(0.03)
α	0.60	0.14	(0.02)	0.60	0.24	(0.02)
ϵ	7.00	5.95	(0.19)	7.00	6.03	(0.19)
ρ_r	0.20	0.24	(0.01)	0.20	0.26	(0.02)
ρ_π	1.30	1.58	(0.06)	1.30	1.52	(0.06)
ρ_y	0.05	0.31	(0.05)	0.05	0.30	(0.02)
ζ_p	0.80	0.79	(0.04)	0.80	0.73	(0.04)
ρ_χ	0.50	0.63	(0.04)	0.50	0.60	(0.02)
ρ_z	0.80	0.73	(0.04)	0.80	0.61	(0.02)
σ_χ	0.01	0.01	(0.001)	0.01	0.01	(0.001)
σ_z	0.01	0.01	(0.002)	0.01	0.01	(0.002)
σ_{mp}	0.006	0.01	(0.001)	0.006	0.01	(0.001)
σ_μ	0.21	0.13	(0.002)	0.21	0.12	(0.002)
σ_χ^{nc}	0.02			0.12		

Table 2: Parameters estimates using flexible trend specification, σ_χ^{nc} is the standard error of the shock to the non-cyclical component.

the data (such as α or σ_n). Third, both the relative magnitude of the various shocks and their persistence is well estimated. Hence, the estimated economy does not differ much from the true one and, as a consequence, decision rules are similar in the two cases.

This is quite evident from figures 7 and 8. The log spectrum and the autocorrelation function of the model-based cyclical component obtained with the true and estimated parameters are close regardless of the filter one applies to the data. Note also, as previously discussed, that both the true and the estimated cyclical components have power at all frequencies of the spectrum, contrary to what standard approaches imply.

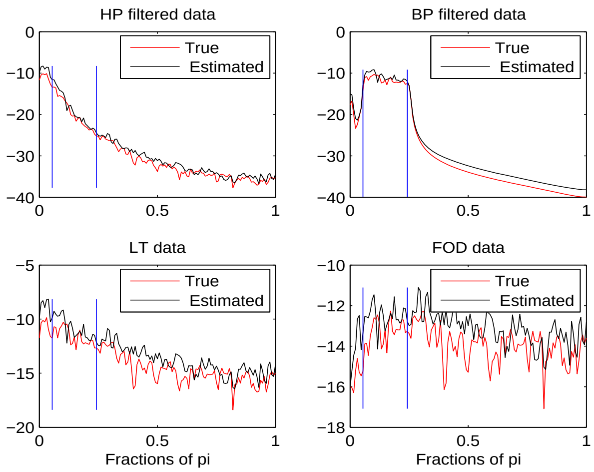

Figure 7: Model based output spectra, true and estimated, different filtering. Vertical bars indicate the frequencies where cycles with 8-32 quarters periodicities are located

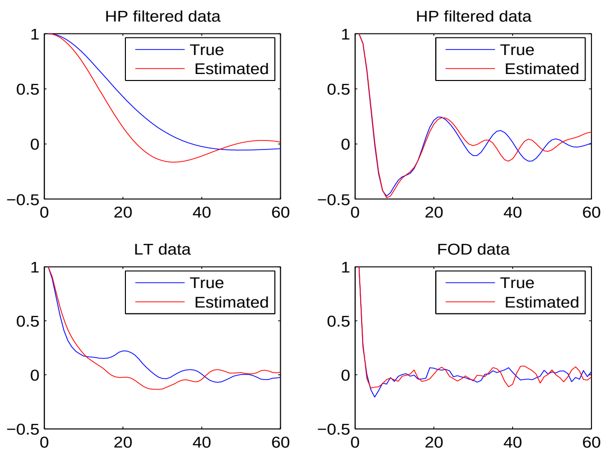

Figure 8: Autocorrelation function of filtered cyclical component, true and estimated

4 An application

References

[1] Mark Aguiar and Gita Gopinath. Emerging market business cycles: The cycle is the trend. *JOurnal of Political Economy*, 115(1):69–110, March 2007.

[2] Michal Andrle. The role of trends and detrending in dsge model: Emerging countries need trendy models. 2008.

[3] Jean Boivin and Mark Giannoni. Dsge estimation in data rich envirorments. 2005.

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

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

[6] Fabio Canova and Filippo Ferroni. Dsge estimation with many filters. 2008, in progress.

[7] Fabio Canova and Matthias Paustian. Measurement with some theory: using sign restrictions to evaluate business cycle models. 2007.

[8] Fabio Canova and Luca Sala. Back to square one: identification issues in dsge models. 2006.

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

[10] Diego Comin and Mark Gertler. Medium term cycles. *American Economic Review*, 96, 2006.

[11] David Lopez Salido Fabio Canova and Claudio Michelaci. Shumpeterian technology shocks. 2007.

[12] Filippo Ferroni. Testing trend specifications in dsge models. 2008.

[13] Eric Ghysels. A study toward a theory of dynamic seasonality ine conomic time series. *Journal of the American Statistical Association*, 83(401):168–172, 1988.

[14] Yuriy Gorodnichenko and Serena Ng. Estimation of dsge models when the data are persistent. 2007.

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

[16] A C Harvey and A Jaeger. Detrending, stylized facts and the business cycle. *Journal of Applied Econometrics*, 8(3):231–47, July-Sept 1993.

[17] Peter N. Ireland. Technology shocks in the new Keynesian model. *The Review of Economics and Statistics*, 86(4):923–936, 2004.

[18] George Kapetanios Jana Eklund, Richared Harrison and Alasdair Scott. Breaks in dsge models. 2008.

[19] Martin Eichenbaum Lawrence J. Christiano and Charles L. Evans. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, 113(1):1–45, 2005.

[20] Ellen McGrattan. Measurement with minimal theory. 2006.

[21] Pau Rabanal. Does inflation increase after a monetary policy tightening? answers based on an estimated dsge model. *Journal of Economic Dynamic and Control*, 31:906–937, 2007.

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

[23] Julio Rotemberg and Mike Woodford. An optimization based econometric framework for the evaluation of monetary policy. *NBER Macroeconomic Annual*, pages 297–346, 1997.

[24] Tom Sargent. Comment to ”seasonal adjustment and multiple time series analysis” by k. wallis. *Seasonal Analysis of Economic time series, A. Zellner (ed.)*, pages 361–364, 1978.

[25] Frank Smets and Raf Wouters. An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of European Economic Association*, 1(5):1123–1175, 2003.

[26] Harald Uhlig. A toolkit for analysing nonlinear dynamic stochastic models easily. Computational methods for the study of dynamic economies., in R. Marimon and A. Scott (Eds)., 1999.

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

Appendix

The baseline model

The model is rather standard. Therefore, it is only briefly summarized.

The bundle of goods consumed by the representative household is

$$C_t = \left(\int_0^1 C_t(j)^{\frac{\epsilon_t - 1}{\epsilon_t}} dj \right)^{\frac{\epsilon_t}{\epsilon_t - 1}} \tag{24}$$

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

$$C_t(j) = \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon_t} C_t \tag{25}$$

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

The representative household chooses sequences for consumption and leisure to maximize

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[X_t \frac{1}{1-\sigma_c} (C_t - hC_{t-1})^{1-\sigma_c} - \frac{1}{1+\sigma_n} N_t^{1+\sigma_n} \right] \tag{26}$$

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

$$\chi_t = \rho_\chi \chi_{t-1} + \epsilon_t^\chi \tag{27}$$

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

$$P_t C_t + b_t B_t = B_{t-1} + W_t N_t \tag{28}$$

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

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

$$Y_t(j) = Z_t N_t(j)^{1-\alpha} \tag{29}$$

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

$$z_t \quad = \rho_z z_{t-1} + \epsilon_t^z \tag{30}$$

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

$$\Pi_t^{1-\epsilon_t} = \zeta_p + (1 - \zeta_p)(P_t^* / P_{t-1})^{1-\epsilon_t} \tag{31}$$

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

$$\max_{P_t^*} \sum_{k=0}^{\infty} \zeta_p^k E_t Q_{t,t+k} \left[P_t^* Y_{t+k|t} - TC_{t+k}(Y_{t+k|t}) \right] \tag{32}$$

subject to the sequence of demand constraints

$$Y_{t+k|t} = \left(\frac{P_t^*}{P_{t+k}} \right)^{-\epsilon_{t+k}} Y_{t+k} \tag{33}$$

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

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

$$r_t = \rho_r r_{t-1} + (1 - \rho_r)(\rho_\pi \pi_t + \rho_y gdp_t) + \epsilon_t^{ms} \tag{34}$$

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

The first order conditions of the optimization problems are:

$$0 = X_t (C_t - hC_{t-1})^{-\sigma_c} - \mathcal{L}_t \tag{35}$$

$$0 = -N_t^{-\sigma_n} - \mathcal{L}_t \frac{W_t}{P_t} \tag{36}$$

$$1 = E_t \left[\beta \frac{\mathcal{L}_{t+1}}{\mathcal{L}_t} \frac{P_{t+1}}{P_t} R_t \right] \tag{37}$$

$$0 = \sum_{k=0}^{\infty} \zeta_p^k E_t Q_{t,t+k} Y_{t+k|t} \left[P_t^* - \mathcal{M}_{t+k} MC_{t+k|t}^n) \right] \tag{38}$$

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

$$\mathcal{M}_t = \mu e^{\epsilon_t^\mu} \tag{39}$$

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

Market clearing requires

$$Y_t(j) = C_t(j) \tag{40}$$

$$N_t = \int_0^1 N_t(j) dj \tag{41}$$

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

Estimation methodology

I assume that a researcher observes a vector of times series y_t, some of which display non-cyclical fluctuations, which the model is interested in explaining. Here $y_t^d = [gdp_t, w_t, \pi_t, r_t]$ where gdp_t is the logarithm of demeaned real gdp, w_t the log of demeaned real wages, π_t the log of demeaned inflation and r_t the log of demeaned nominal interest rate. I assume that the model is intended to explain the cyclical properties of y_t^d and let the model's observable variables be $y_t^m(\theta)$, where θ are the structural parameters. The link between the model and the data is given in (14), which is the measurement equation and by (16) and(12)-(11)-(**??**) which represent the transition equations.

The vector of structural parameters θ is composed of

$$\theta = [\sigma_c, \sigma_n, h, \alpha, \epsilon, \rho_r, \rho_\pi, \rho_y, \zeta_p, \rho_\chi, \rho_z, \sigma_\chi, \sigma_z, \sigma_{mp}, \sigma_\mu] \tag{42}$$

and the vector of non-structural parameters is $\nu = (vec(\Sigma_e), \vec{(\Sigma_v)}, \vec{(\Sigma_u)})$.Let $\vartheta = (\theta, \nu)$.

x_t is the vector of endogenous states, which includes

$$x_t = [\lambda_t, gdp_t, n_t, mc_t, mrs_t, r_t]. \tag{43}$$

gdp_t, n_t and r_t enter the endogenous state because the algorithm in [26] recognizes as endogenous states all the variables dated at t and $t-1$ that appear in the expectation equations. To include output and the nominal interest rate in the vector of observables, I add to the equilibrium conditions,

$$gdp_t = \widetilde{gdp}_t \tag{44}$$

$$r_t = \widetilde{r}_t \tag{45}$$

Hence the vector of observable variables is

$$y_t = [\widetilde{gdp}_t, \omega_t, \pi_t, \widetilde{r}_t] \tag{46}$$

The vector of exogenous processes and exogenous innovations are

$$z_t = [\chi_t, z_t, \epsilon_t^{ms}, \epsilon_t^{\mu}] \tag{47}$$

$$\epsilon_t = [\epsilon_t^{\chi}, \epsilon_t^{z}, \epsilon_t^{ms}, \epsilon_t^{\mu}]. \tag{48}$$

Equations (12)-(13), (16) and (**??**) can be cast into the state space system

$$s_{t+1} = Fs_t + G\omega_{t+1} \tag{49}$$

$$y_t = Hs_t + u_t \tag{50}$$

where $F = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & PP & QQ \\ 0 & 0 & 0 & NN \end{pmatrix}$, $G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ & & \end{pmatrix}$, $H = \begin{pmatrix} 1 & 0 & RR & SS \end{pmatrix}$,

$s_{t+1} = \begin{pmatrix} y_t^T & \mu_t & y_{t-1}^m(\theta) & z_t \end{pmatrix}$, $\omega_{t+1} = (e_t, v_t, 0, \epsilon_{t+1})$ and Σ_ω is block diagonal. The likelihood of the system can be computed with the Kalman filter. The Kalman filter returns the optimal (latent) states estimates and the likelihood for the system for a given ϑ.

To obtain the non-normalized posterior distribution of (ϑ), I use a standard Random Walk Metropolis algorithm (RWM). The algorithm proceeds as follows. Given initial value ϑ_{l-1}, given a Ω, and a prior $g(\theta)$:

1. Draw a shock vector υ from $t(0, \kappa * \Omega, 5)$ and construct a candidate $\vartheta_* = \vartheta_{\ell-1} + \upsilon$.

2. Solve the model system given θ_*; if the solution is indeterminate or no solution is found set $\mathcal{L}(\vartheta^*|y) = 0$

3. Otherwise, evaluate the likelihood at ϑ_* with the Kalman filter, $\mathcal{L}(y|\vartheta_*)$

4. Calculate $\breve{g}(\vartheta_*|y) = g(\vartheta_*)\mathcal{L}(y|\vartheta_*)$ and the ratio $\chi_* = \frac{\breve{g}(\vartheta_*|y)}{\breve{g}(\vartheta_{\ell-1}|y)}$

5. Draw u from $U[0, 1]$; if $\chi_* > u$ set $\vartheta_\ell = \vartheta_*$, otherwise set $\vartheta_\ell = \vartheta_{\ell-1}$

Iterated a large number of times, for κ appropriately chosen, the algorithm ensures that the limiting distribution of the chain is the target distribution (see e.g. [5], Ch. 9).

The priors

Priors distributions for the structural and the non-structural parameters are in Table 3 For the non-structural parameters I assume inverted gamma distributions with large standard deviations to explore a wide region for the parameters without being downweighted by the prior.

Structural Parameters				
Parameters	Distribution	Mean	Standard Deviation	Calibration
σ_c	$\Gamma(20, 0.1)$	2.00	0.45	3.0
σ_n	$\Gamma(20, 0.1)$	2.00	0.45	0.7
h	$B(6, 8)$	0.43	0.13	0.7
α	$B(3, 8)$	0.27	0.13	0.3
ϵ	$N(6, 0.1)$	6.00	0.10	7.0
ρ_r	$B(6, 6)$	0.50	0.14	0.2
ρ_π	$N(1.5, 0.1)$	1.50	0.10	1.30
ρ_y	$N(0.4, 0.1)$	0.40	0.10	0.05
ζ_p	$B(6, 6)$	0.50	0.14	0.8
ρ_χ	$B(18, 8)$	0.69	0.09	0.5
ρ_z	$B(18, 8)$	0.69	0.09	1.0
σ_χ	$\Gamma^{-1}(10, 20)$	0.0056	0.0020	0.0051
σ_z	$\Gamma^{-1}(10, 20)$	0.0056	0.0020	0.0312
σ_{mp}	$\Gamma^{-1}(10, 20)$	0.0055	0.0020	0.0010
σ_μ	$\Gamma^{-1}(10, 20)$	0.0056	0.0020	0.2060
Non-Structural Parameters				
σ_{e_i}	$\Gamma^{-1}(10, 20)$	0.0056	0.0020	
σ_{v_i}	$\Gamma^{-1}(10, 20)$	0.0056	0.0020	
σ_{u_i}	$\Gamma^{-1}(10, 20)$	0.0055	0.0020	

Table 3: Prior Distribution for the parameters and calibrated values