# Bridging cyclical DSGE models and the raw data

January 27, 2011

**Abstract**

I propose a method to estimate cyclical DSGE models using the raw data. The approach links the observables to the model counterparts via a flexible specification which does not require that the cyclical component is solely located at business cycle frequencies and allows the non-cyclical component to take various time series patterns. Applying standard data transformation induces distortions in structural estimates and policy conclusions and explain the reasons for their emergence. The proposed approach recovers the features of the cyclical component in selected experimental designs.

JEL classification: E32, C32.

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

# 1 Introduction

There have been considerable developments in the specification of DSGE models over the last 10 years. The original structure, featuring a single technological disturbance, has been enriched with shocks and frictions and our understanding of the propagation mechanism of important shocks enhanced. Steps forward have also been made in the estimation of these models. While a few years ago it was standard to informally calibrate their structural parameters, now researchers routinely use limited and full information estimation procedures and, perhaps more importantly, this trend is common in academic and policy circles (see, e.g., [28], [11], [30],[22], [27] among many others).

Despite recent efforts, structural estimation of DSGE models is conceptually and practically difficult. For example, classical estimation is asymptotically justified only when the model is the generating process (DGP) of the actual data, up to a set of serially uncorrelated measurement errors, and standard validation exercises are meaningless without such an assumption. Identification problems (see [8]) and numerical difficulties are widespread. Finally, the majority of the models investigators use is intended to explain only the cyclical portion of observable fluctuations but macroeconomic data contains many types of fluctuations, some of which is hardly cyclical.

There are two reasons for why researchers prefer to work with "cyclical" models. Jointly accounting for all types of fluctuations is still an ambitious task since there are very few known theoretical mechanisms able to propagate temporary disturbances for a sufficiently long time (we need e.g. R&D as in [13] or a Schumpeterian creative destruction) and it is empirically difficult to measure non-cyclical fluctuations because samples are short and structural breaks likely to exist. In addition, it is convenient from the computation and the interpretation point of views, to assume that the mechanisms driving cyclical and non-cyclical fluctuations are distinct.

The mismatch between what models are designed to explain and what the data contains creates headaches for applied investigators. In the literature a number of approaches have been used, all of which are far from ideal:

• Ignore the existence of non-cyclical movements and fit the cyclical model to demeaned data (see [15]). Such an approach is likely to artificially increase the persistence of estimated shocks and, potentially, distort important economic mechanisms present in the model.

• Fit the cyclical model to data filtered with an arbitrary statistical device. Such an approach is popular, but problematic for at least three reasons. First, the cyclical component extracted by the majority of the statistical filters can be represented as a symmetric, two-sided moving average of the raw data. Thus, as suggested long ago in [21], the timing of the information is altered by filtering, and dynamic analyses become difficult to interpret. Second, while it is typical to filter each series separately and to remove non-cyclical fluctuations from real but not from nominal variables, there are consistency conditions that must hold - a resource constraint need not be satisfied if each variable is separately filtered - and many situations when not all the fluctuations in nominal variables can be safely considered as cyclical. Thus, specification errors can be present. Finally, even arbitrarily focusing attention on the so-called business cycle frequencies, i.e. fluctuations with 8-32 quarters average periodicity, filtering and contamination errors are likely to be present. For example, a Band Pass (BP) filter, when used with finite stretches of data, only very roughly captures the power of the spectrum at business cycle frequencies and taking growth rates greatly emphasizes the high frequency content of the data. In sum, rather than resolving the problems, statistical filtering increases the difficulties applied researchers face.

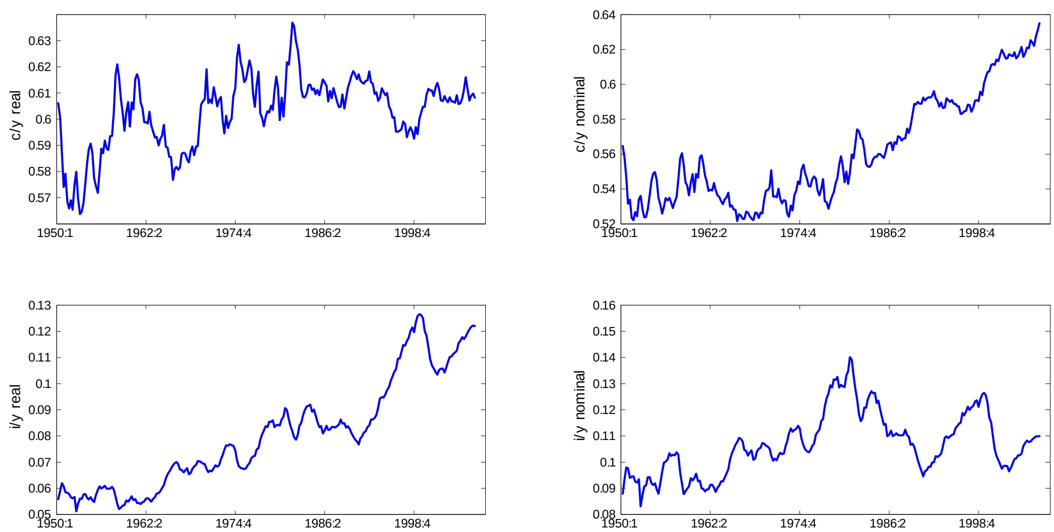


Figure 1: US real and nominal great ratios.

• Estimate the model using data transformations which, in theory, are likely to be void of non-cyclical fluctuations, e.g. consider real "great ratios" (as suggested in [12] and [26]) or nominal " great ratios", (as suggested in [32]). As Figure 1 shows, such transformations are unlikely to resolve the mismatch issue because many ratios still display important low frequency movements. In addition, since the number and the nature of the shocks driving non-cyclical fluctuations needs to be known, specification errors may be produced.

• Modify the cyclical model by arbitrarily adding a trend; detrend the model and the data with the same transformation; and fit the detrended model to the detrended data (see e.g. [17], [25] among many others). Such an approach imposes some consistency between the theory and the data, but also faces important problems. For example, since the choice of which shock is trending is often driven by computational rather than economic considerations, specification errors could be present. In addition, the estimated cyclical model depends on nuisance features, such as the shock which is assumed to be trending or the time series specification of the trend [1]. As shown in [12] and more recently in [20], misspecification of these nuisance features may lead to severe biases in the estimates of the structural parameters.

• Modify the cyclical model by arbitrarily adding a trend but use a frequency domain maximum likelihood approach (see e.g. [10]) to estimate the parameters of a detrended model. Here the frequency domain representation of the solution (in growth rates) is used to construct the likelihood function of the model and only particular frequencies are taken into account. This approach is also subject to filtering and specification errors since it inherits the trend misspecification problems of model-based detrending and the filtering problems of statistically based filtering approaches.

This paper provides a method to estimate cyclical DSGE models which avoids, to a large extent, specification and filtering problems. I proceed in three steps. First, I show that the transformation one takes to fit a cyclical model to the data matters for structural parameter estimation and for economic inference in general. Hence, unless one has a strong but unjustified view of what "cyclical" data the model should explain, one is left wandering how to credibly select among various structural estimates. Second, I argue that even if such a criteria could be found, estimates of the parameters of a cyclical model obtained with *any* preliminary data transformations should not be trusted. On the one hand, the presence of measurement error with low frequencies features distorts the conclusions applied investigators reach. An approach dealing with this type of contamination, has been recently proposed in [6].

On the other, the time series generated by a cyclical DSGE model have properties which are different from the output of a statistical filter or of theoretically based data transformations, even in large samples. All these approaches seek to isolate the power of a series in a band of frequencies. However, a cyclical DSGE model generates time series with power at all frequencies of the spectrum. Thus, any preliminary data transformation wipes out potentially important cyclical information. More generally, if a DSGE model features cyclical and non-cyclical shocks, the cyclical and the non-cyclical components jointly appear in any frequency band and it is not difficult to build examples where the non-cyclical component dominates at business cycle frequencies making inference based on any data transformations incorrect. [1] claimed that for Less Developed Countries (LDC) the non-cyclical component is important at business cycle frequencies. It turns out that the problem is relevant for structural estimation with the data of any country and that significant biases could emerge when the disturbances driving the two components have roughly similar variance.

In the last part of the paper, I propose to estimate the structural parameters of a "cyclical" DSGE model by creating a flexible link between the model and the raw data that allows the cyclical and the non-cyclical components to have power at all the frequencies of the spectrum. Since

[1] One can show, for example, that the estimable model one obtains assuming a stochastic linear trend in the technology of the model of section 2 has moving average terms that do not appear, for example, if a unit root is assumed and that, in case of unit root in preferences, a constant return to scale technology needs to be imposed, a conditions which is not needed if the unit root is in technology.

the specification I use encompasses, as special cases, situations where the non-cyclical component displays deterministic, stochastic or smooth features, the approach does not require researchers to take a stand on the non-cyclical features of the raw data and therefore shields the analysis from important specification errors. Furthermore, since the information present at all frequencies is used to estimate the parameters of the cyclical model, filtering distortions are eliminated and inefficiencies minimized. Finally, the hybrid model I construct takes the cyclical model as given and uses the flexible link to fill the rest of the spectrum of the raw data, thus avoiding the identification problems that model-based transformations produce. Using a simple experimental design I show that standard transformations produce distorted parameter estimates and I interpret the resulting biases using the decomposition of the likelihood function suggested in [21]. The suggested procedure instead captures the cyclical component of the DGP better and produces sensible estimates of crucial parameters when samples typically available in macroeconomics are used. Finally, I show that economic inference depends on the approach used to match the model to the data and that, with the suggested one, certain facts are much less stylized than one would like to assume for policy.

To focus attention on the issues of interest, I make a number of simplifying assumptions. In particular, I assume that (i) the estimated model is correctly specified; that is, there are no missing variables or omitted cyclical shocks; (ii) theoretical singularities are absent - the number of shocks equals the number of endogenous variables - and (iii) model variables have an exact counterpart in the data, i.e. no proxy error is present. While these issues are important in practice and semi-structural methods of the type suggested in [7] produce more robust inference when they are present, it is useful to keep them separate from the issue of estimating cyclical models with raw data because the problems I highlight occur regardless of whether (i)-(ii)-(iii) are present or not.

The rest of the paper is organized as follows. The next section presents a simple cyclical model, estimates its structural parameters using a number of preliminary data transformations and discusses why estimates should not be trusted. Section 3 presents the alternative methodology. Section 4 compares approaches using a simple experimental design. Section 5 compares estimates of functions of the parameters in two widely used cyclical DSGE models. Section 6 concludes. The appendix contains relevant additional material mentioned in the paper.

## 2 Cyclical estimation with transformed data

To show that estimates of the structural parameters a "cyclical" DSGE model depend on the preliminary data transformation employed, I consider a standard small scale New-Keynesian model where agents face a labor-leisure choice, there is external habit in consumption, production is stochastic and requires labour, there is an exogenous probability of price adjustments and monetary policy is conducted with a conventional Taylor rule. Details on the structure of this model are in the Appendix. The log-linearized equilibrium conditions are:

$$w_t = (\frac{\sigma_n}{1-\alpha} + \frac{\sigma_c}{1-h})y_t - \frac{h\sigma_c}{1-h})y_{t-1} - \frac{\sigma_n}{1-\alpha}z_t - \chi_t \quad (1)$$

$$y_t = E_t(\frac{1}{1+h}y_{t+1} - \frac{h}{1+h}y_{t-1} + \frac{1-h}{(1+h)\sigma_c}(\chi_{t+1} - \chi_t + r_t - \pi_{t+1}) \quad (2)$$

$$\pi_t = \beta E_t \pi_{t+1} + \kappa_p(\mu_t + w_t + \frac{\alpha}{1-\alpha}y_t - \frac{1}{1-\alpha}z_t) \quad (3)$$

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

$$n_t = \frac{1}{1-\alpha}(y_t - z_t) \quad (5)$$

where all variables are expressed in deviation from the steady states, $k_p = \frac{(1-\beta\zeta_p)(1-\zeta_p)}{\zeta_p}\frac{1-\alpha}{1-\alpha+\varepsilon\alpha}$, $z_t = \rho_z z_{t-1} + \epsilon_t^z$ is a technology disturbance, $\chi_t = \rho_\chi \chi_{t-1} + \epsilon_t^\chi$ a preference disturbance, $\epsilon_t^r$ an iid monetary policy disturbance and $\epsilon_t^\mu$ an iid markup disturbance. Equation (1) defines the equilibrium real wage, equation (2) is an Euler equation, equation (3) a Phillips curve, equation (4) a Taylor rule and equation (5) a labor demand function. The structural parameters to be estimated are: $\sigma_c$ the risk aversion coefficient, $\sigma_n$ the inverse of the Frisch elasticity, $h$ the coefficient of consumption habit, $1-\alpha$ the share of labor in production, $\rho_r$ the degree of interest rate smoothing, $\rho_\pi$ and $\rho_y$ the parameters of the policy rule, $\zeta_p$ the probability of not changing prices. The auxiliary parameters to be estimated are: $\rho_\chi, \rho_z$ the autoregressive parameters of preference and technology shocks, and $\sigma_z, \sigma_\chi, \sigma_r, \sigma_\mu$ the standard deviations of the four structural shocks. The discount factor $\beta = 0.99$ and the elasticity among varieties $\varepsilon = 7.00$ are kept fixed in the estimation exercises since they are very weakly identified from the data.

I assume that there are four observable variables and I have examined a variety of filtering and transformation approaches applied to all or a subset of the variables (as in [**?**], [27], [30]) or [22]). To illustrate the point of interest, I report parameter estimates and impulse responses obtained when output, the real wage, the nominal interest rate and inflation are transformed using four statistical filters (linear detrending (LT), Hodrick and Prescott filtering (HP), growth rate filtering (FOD)and band pass filtering (BP)). In addition, I present results obtained with two data transformations. The first uses the log of labour productivity, the log ratio of real wages to hours, the nominal rate and the inflation rate, all demeaned, as observables (Ratio 1). The second uses the log ratio of output to the real wage, hours worked, the nominal rate and the inflation rate, all demeaned, as observables (Ratio 2). The first transformation insures stationarity of the observables if the non-cyclical portion of the data is driven by a non-stationary preference shock; the second, insures stationarity of the observables if the non-cyclical component is driven by a non-stationary Total Factor Productivity (TFP) shock. The data comes from the FRED database at the Federal Reserve Bank of St. Louis and estimation is conducted using Bayesian methods. Results obtained applying the same transformations to real variables and simply demeaning nominal ones are in the appendix.

Since some of the filters belong to the class of two-sided moving averages, I have also experimented with a recursive LT filter and a one-sided version of the HP filter. The results presented below are unchanged by this refinement.

| | Prior | LT | HP | FOD | BP | Ratio 1 | Ratio2 |
|---|---|---|---|---|---|---|---|
| | | Median (s.e.) | Median (s.e.) | Median (s.e.) | Median (s.e.) | Median(s.e.) | Median (s.e.) |
| $\sigma_c$ | $\Gamma(20, 0.1)$ | 1.90 (0.25) | 1.41 (0.21) | 0.04 (0.01) | 0.96 (0.11) | 2.33 (0.27) | 0.81 (0.15) |
| $\sigma_n$ | $\Gamma(20, 0.1)$ | 1.75 (0.16) | 1.37 (0.13) | 5.23 (0.08) | 1.19 (0.09) | 3.02 (0.24) | 2.68 (0.19) |
| $h$ | $B(6, 8)$ | 0.83 (0.02) | 0.88 (0.02) | 0.45 (0.01) | 0.96 (0.01) | 0.72 (0.05) | 0.88 (0.02) |
| $\alpha$ | $B(3, 8)$ | 0.07 (0.04) | 0.09 (0.05) | 0.42 (0.01) | 0.07 (0.03) | 0.05 (0.04) | 0.03 (0.01) |
| $\rho_r$ | $B(6, 6)$ | 0.19 (0.05) | 0.11 (0.04) | 0.62 (0.01) | 0.09 (0.02) | 0.38 (0.06) | 0.28 (0.04) |
| $\rho_\pi$ | $N(1.5, 0.1)$ | 1.33 (0.08) | 1.37 (0.05) | 1.53 (0.02) | 1.51(0.06) | 1.92 (0.06) | 1.80 (0.05) |
| $\rho_y$ | $N(0.4, 0.1)$ | -0.16 (0.03) | -0.18 (0.03) | 0.06 (0.00) | -0.22 (0.03) | 0.16 (0.02) | -0.03 (0.02) |
| $\zeta_p$ | $B(6, 6)$ | 0.82 (0.02) | 0.80 (0.03) | 0.63 (0.01) | 0.86 (0.01) | 0.82 (0.02) | 0.80 (0.02) |
| $\rho_\chi$ | $B(18, 8)$ | 0.69 (0.04) | 0.40 (0.05) | 0.52 (0.01) | 0.70(0.02) | 0.67 (0.03) | 0.66 (0.02) |
| $\rho_z$ | $B(18, 8)$ | 0.96 (0.02) | 0.95 (0.02) | 0.99 (0.01) | 0.97(0.01) | 0.97 (0.01) | 0.96 (0.01) |
| $\sigma_\chi$ | $\Gamma^{-1}(10, 20)$ | 0.53 (0.19) | 0.47 (0.11) | 4.96(0.13) | 0.23 (0.05) | 3.41 (0.74) | 0.97 (0.13) |
| $\sigma_z$ | $\Gamma^{-1}(10, 20)$ | 0.20 (0.04) | 0.23 (0.04) | 2.00 (0.22) | 0.19 (0.03) | 0.06 (0.01) | 0.06 (0.01) |
| $\sigma_r$ | $\Gamma^{-1}(10, 20)$ | 0.11 (0.01) | 0.08 (0.01) | 2.30(0.23) | 0.07 (0.01) | 0.10 (0.01) | 0.11 (0.18) |
| $\sigma_\mu$ | $\Gamma^{-1}(10, 20)$ | 25.06 (0.97) | 14.25 (0.93) | 7.17 (0.13) | 18.19 (0.66) | 22.89 (1.91) | 15.94 (0.49) |
| $k_p$ | | 0.02 (0.001) | 0.03 (0.002) | 0.04 (0.001) | 0.01 (0.001) | 0.03 (0.002) | 0.04 (0.002) |

Table 1: Posterior estimates. LT refers to linearly detrended data, HP to Hodrick and Prescott filtered data, FOD to demeaned growth rates, BP to band pass filtered data. In these cases, all observables are filtered. For Ratio 1 the observables are $\log(y_t) - \log(n_t), \log(w_t) - \log(n_t), \pi_t, r_t$, all demeaned For Ratio 2 the observables are $\log(y_t) - \log(w_t), \log(n_t), \pi_t, r_t$, all demeaned. The sample is 1980:1-2007:4.

Table 1 contains the priors (which are the same for each transformation of the raw data) and the median and the standard deviation of the posterior distribution of the parameters of interest. Clearly, there are several parameters whose posterior is affected by the preliminary data transformation used (see e.g. the risk aversion coefficient $\sigma_c$, the Frisch elasticity of labor supply $\sigma_n^{-1}$, the interest smoothing coefficient $\rho_r$, and persistence and the volatility of the shocks). Since posterior standard deviations are tight, differences across columns are a-posteriori significant. Posterior differences are also economically relevant. For example, the volatility of markup shocks in the LT and Ratio 1 economies is considerably larger than in the other economies and, perhaps unsurprisingly, risk aversion stronger.

Differences in the location of the posterior of the parameters translate into important differences in the transmission of shocks. For example, as shown in Figure 2, the magnitude of the impact coefficient and of the persistence of the responses vary with the preliminary transformation employed. Furthermore, at least in the case of technology shocks, the sign of some of the responses is affected.

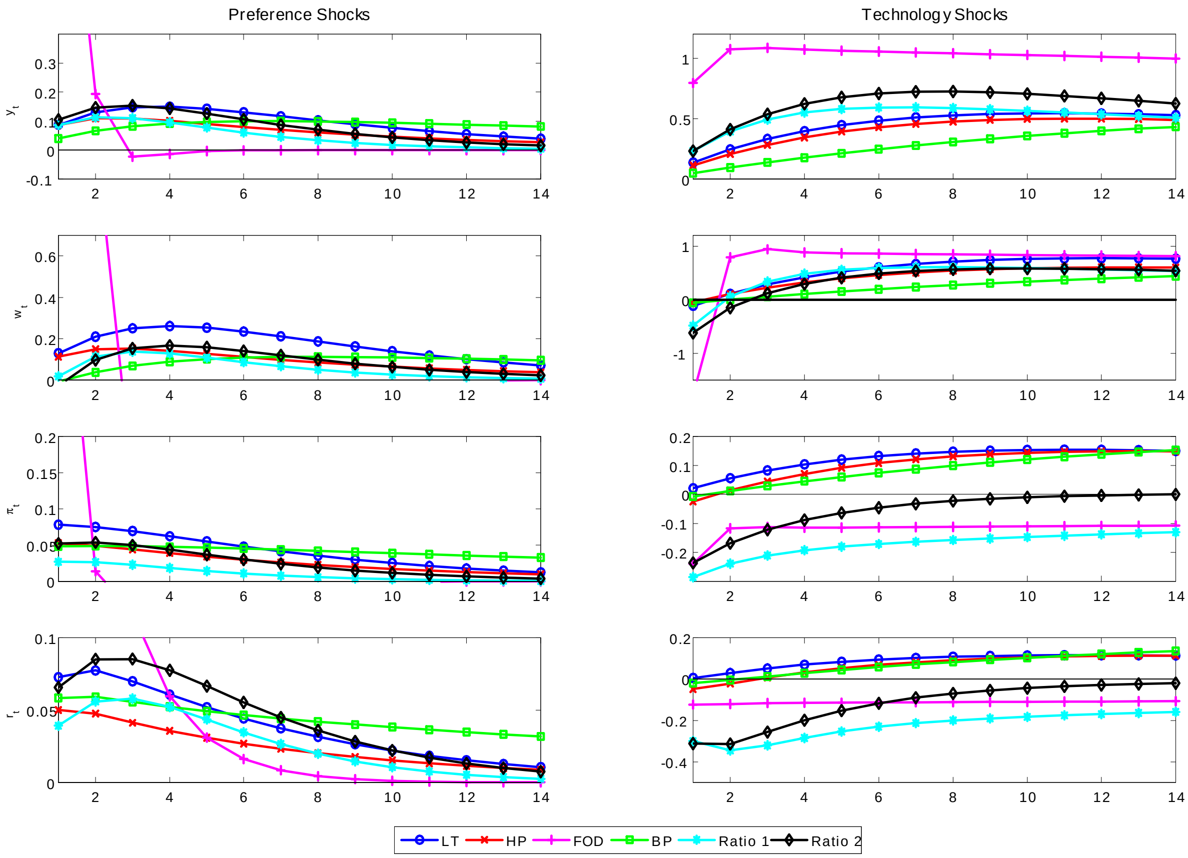


Figure 2: Impulse responses, all observables filtered, sample 1980:1-2007:4.

While it is tempting to sweep these differences under the rug, one should expect them to occur since the observables in different columns of table 1 have different time series properties. These differences would be inconsequential, if applied researchers had a good reason to prefer one column over another. As argued in [4], it is hard to design criteria to do so. But even if a criteria could be found, none of the columns of table 1 should be trusted for two reasons. First, all the transformations only approximately isolate business cycle frequencies: the LT transformation leaves both long and short cycles in the filtered data; the HP transformation leaves high frequencies variability unchanged; the FOD transformation emphasizes high frequency fluctuations and reduces the importance of cycles with business cycle periodicity; ratios leave important low frequency fluctuations in the transformed data; and even a BP transformation, once truncations due to finite samples are considered, induces significant approximation errors (see e.g. [5], ch.3). Since considerable measurement error is present in the transformed data, estimation distortions are likely to appear. In addition, since different approaches spread the measurement error across different frequencies, estimates are likely to depend on the preliminary transformation used. An estimation approach which can reduce the extent of these measurement errors is suggested in [6].

Second, all transformations assume that the cyclical and the non-cyclical components of the data are located at different frequencies of the spectrum and that the economic mechanism generating the two is distinct. Such an assumption is crucial, for example, when associating business cycle frequencies in the data with the output of the model. However, a cyclical DSGE model produces time series with power at frequencies other than those corresponding to 8 to 32 quarters and, on the other hand, the non-cyclical component may be important at business cycle frequencies. Time series macroeconometricians are generally aware that statistical transformations are unlikely to recover interesting economic objects. For example, [29], [19] and [21] have all emphasized the fallacy of estimating structural models using seasonally adjusted data, precisely for this reason.

To illustrate these points, I take the model generating (1)-(5) as log-linearized conditions as DGP but assume that the preference shock has two uncorrelated components: one with a unitary autoregressive (AR) coefficient and one with a AR coefficient equal to $|\rho_x| << 1$. Thus, the model possesses a theoretical non-cyclical component (induced by the disturbances to the non-stationary component of the preference shock) and a theoretical cyclical component (induced by the other four stationary disturbances). The exact decomposition of the four observables into the two components is in the appendix. With the parameterization reported in the first panel of table 4, I have simulated data from the model and passed the experimental data through LT, HP, FOD and BP filters and constructed the transformation corresponding to Ratio1. Figure 3 shows the log spectrum of the true cyclical and non-cyclical components of output (first box) and of the true and estimated cyclical components of output (boxes 2 to 6).

It is clear that all filters leave considerable power outside the business cycle frequencies (identified by the vertical bars). The problem is more evident with LT and FOD but leakages and compressions are present also with HP and BP filters and the Ratio1 transformation. In addition, both components have power at all frequencies and, with the chosen parameterization, the theoretical non-cyclical component is as important as the theoretical cyclical component at business cycle frequencies. [1] and [3] claimed that for LDC countries the trend is the cycle. Figure 3 shows that the problem may be generic. It is only required that the shocks driving the cyclical and non-cyclical components have similar variability. Note that the features we highlight do not depend on the assumption that the preference shock has two components; had I assumed that the technology shock has two components a similar picture would emerge.

Since all transformations partially or completely eliminate the power of the cyclical component in the low frequencies and force the cyclical model to explain the power of the non-cyclical component at business cycle and other frequencies, it is not hard to guess that estimates of the structural parameters will be distorted.

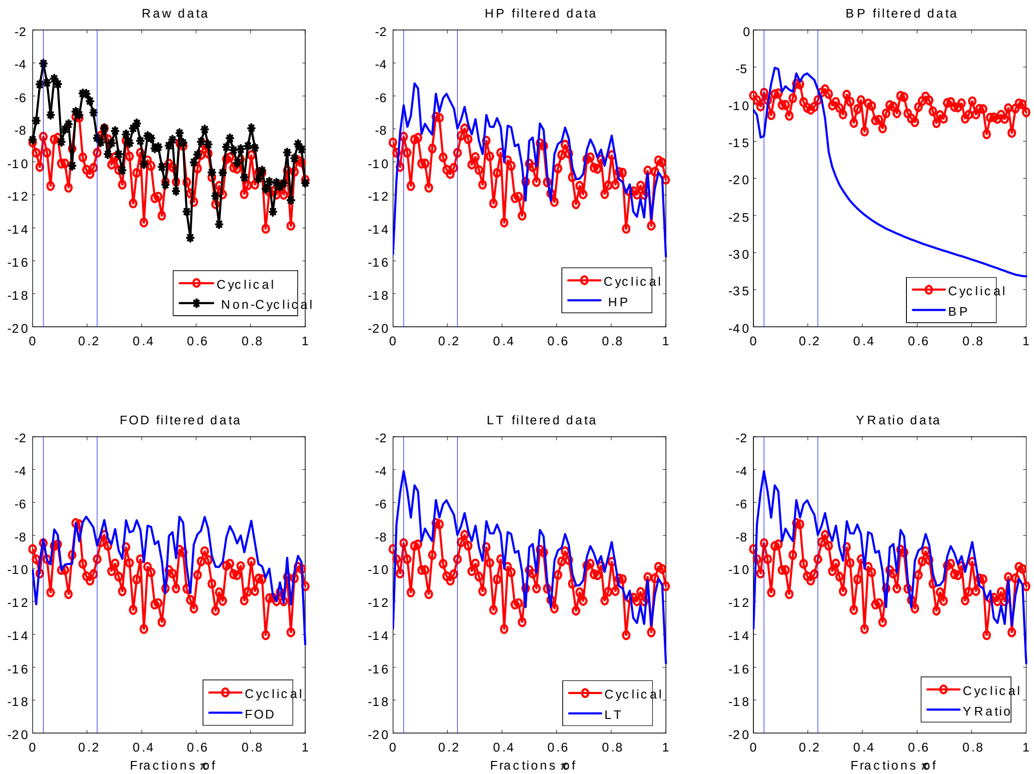


Figure 3: Output log-spectra. Vertical bars indicate the frequencies where cycles with 8-32 quarters periodicities are located. For Ratio1 the true and estimated log spectrum of the cyclical component of labor productivity is plotted. Simulated data.

## 3 The alternative methodology

One solution to the problems I have highlighted is to build a non-cyclical component directly into the DSGE model. I have mentioned in the introduction the reasons for why one may be reluctant to do so. There may also be practical concerns which may not make the approach viable (Should the non-cyclical component be deterministic or stochastic? Should it be correlated with the cyclical component or not? ). Progress in addressing these issues have been reported in [20] and [16].

Rather than augmenting the model with an arbitrary non-cyclical component and making the analysis vulnerable to specification errors, I will use a flexible setup, in the spirit of [?], where the cyclical DSGE structure is fixed and a link is build from the cyclical model to the raw data which permits the cyclical and non-cyclical components to jointly appear at all frequencies of the spectrum and allow the non-cyclical component to potentially take a variety of patterns. Let the (log)-linearized solution of a DSGE model be of the form:

$$x_{2t} = RR(\theta)x_{1t-1} + SS(\theta)\epsilon_t \tag{6}$$

$$x_{1t} = PP(\theta)x_{1t-1} + QQ(\theta)\epsilon_t \tag{7}$$

where $PP(\theta), QQ(\theta), RR(\theta), SS(\theta)$ are functions of the structural parameters $\theta = (\theta_1, \ldots, \theta_k)$, $x_{1t} \equiv (\log \tilde{x}_{1t} - \log \bar{x}_1)$ includes the states and the predetermined variables, $x_{2t} = (\log \tilde{x}_{2t} - \log \bar{x}_2)$ all other endogenous variables, $\epsilon_t$ the shocks and $\bar{x}_2, \bar{x}_1$ are the steady states of $\tilde{x}_{2t}$ and $\tilde{x}_{1t}$.

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

$$x_t^d = c + x_t^{nc} + x_t^m(\theta) + u_t \tag{8}$$

where $c = \log \bar{x}^m(\theta) - E(\log \tilde{x}_t^d)$, $x_t^{nc}$ is the non-cyclical component, $u_t$ is a iid $(0, \Sigma_u)$ (measurement) noise and $x_t^{nc}, x_t^m$ and $u_t$ are mutually orthogonal. Furthermore, I assume that

$$\begin{aligned} x_t^{nc} &= x_{t-1}^{nc} + \bar{x}_{t-1} + e_t \qquad e_t \sim \; iid\ (0, \Sigma_e) \\ \bar{x}_t &= \bar{x}_{t-1} + v_t \qquad v_t \sim \; iid\ (0, \Sigma_v) \end{aligned} \tag{9}$$

Since the locally linear specification used in (9) may be considered controversial, some explanations for our choice may be in order. Note first that the specification in (9) is flexible and nests, as special cases, the structures which are typically thought to motivate the use of the statistical filters considered in the previous sections. For example, if both $\Sigma_e$ and $\Sigma_v$ are diagonal, $\Sigma_{v_i} > 0$ and $\Sigma_{e_i} = 0$, $\forall i$, $x_t^{nc}$ is a vector of I(2) processes while if $\Sigma_{v_i} = 0$, and $\Sigma_{e_i} > 0$, $\forall i$, $x_t^{nc}$ is a vector of I(1) processes. Furthermore, if $\Sigma_{v_i} = \Sigma_{e_i} = 0$, $\forall i$, $x_t^{nc}$ is deterministic. Finally, when $\Sigma_{v_i}$ and $\Sigma_{e_i}$ are functions of $\Sigma_\epsilon$, the setup approximates the double exponential smoothing restrictions used in discounted least square estimation of non-cyclical patterns (see e.g. [14]). Second, as [14] have shown, (9) is robust against several types of misspecification of the time series properties of the non-cyclical component. Finally, since $\Sigma_v$ is not constrained to be zero, the growth rates of the endogenous variables may display persistent deviations from their mean, a feature that characterizes many real macroeconomic variables (see e.g. [24]).

Identification of the structural parameters is achieved via the cross equation restrictions that the cyclical model imposes on the data. Since (8) and (9) can be thought as a sophisticated representation for what the cyclical model leaves out, estimates of the non-structural parameters are obtained from the portion of the data the model can not explain.

Given (6)-(7)-(8) and (9), 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 $\theta$. In (9) I have assumed that $\Sigma_v$ and $\Sigma_e$ are general matrices, but one can impose further structure by assuming that they are either diagonal (so that the non-cyclical component is series specific), that they are of reduced rank (so that the non-cyclical component is common across series), or that they have only sparse non-zero elements on the diagonal (so that a non-cyclical component is present only in a reduced number of observables).

The suggested specification has several advantages over existing approaches. On the one hand, 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. This shields researchers from important specification errors. In addition, as I will show below, the cyclical component extracted with this approach is not located at particular frequencies of the spectrum. In fact, by construction, all components in (8) may have power at each frequency. Furthermore, structural parameter estimates reflect the uncertainty present in the specification of the non-cyclical component. Finally, as shown in [18], statistical techniques can be used to find the most appropriate specification of the non-cyclical component and, if a researcher is interested in doing so, to perform Bayesian averaging over different types of non-structural specifications of the non-cyclical component, which is not possible in standard setups.

In (8) I have assumed that cyclical and non-cyclical fluctuations are driven by independent shocks. It is easy to show that the specification I select is observationally equivalent to one where the non-cyclical and the cyclical components are correlated. For example, the specification

$$x_t^d = x_t^{nc*} + x_t^{m*}(\theta) + u_t \tag{10}$$
$$x_t^{nc*} = x_t^{nc} - G(\theta)\bar{x}_t \tag{11}$$
$$x_t^{m*}(\theta) = x_t^{m}(\theta) + G(\theta)\bar{x}_t \tag{12}$$

for some $G(\theta)$, which makes $y_t^{nc*}$ and $y_t^{m*}(\theta)$ correlated, is indistinguishable from the point of view of the observed data from the specification I suggest. In practice, this means that if progress is made in the specification of features of the DSGE model which are now bundled up in the reduced form specification of $x_t^{nc}$, these can be eliminated from $x_t^{nc}$ and included in the model without altering the format of the estimated structure. For example, the setup provides an appropriate estimable structure if a researcher feels comfortable in specifying a unit root component for the TFP process, but still believes that important low frequency variations in e.g. labor productivity can not be explained with the specified set of shocks.

## 3.1 Estimation

Estimation of the hierarchical structure (7)-(6), (9) and (8) can be carried out with both classical and Bayesian methods. Equations (7)-(6), (9) and (8) can be cast into an (extended) state space system of the form

$$s_{t+1} = Fs_t + G\omega_{t+1} \quad (13)$$

$$y_t = Hs_t \quad (14)$$

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

$s_{t+1} = \begin{pmatrix} x_t^{nc} & \bar{x}_t & u_t & x_{t-1}^m(\theta) \end{pmatrix}$, $\omega_{t+1} = (e_t, v_t, u_t, \epsilon_{t+1})$ and $\Sigma_\omega$ is block diagonal. Hence, the likelihood of the system can be computed with a modified Kalman filter (taking into account the possibility of diffuse initial observations) for a given $\vartheta$ and maximized using standard tools.

When a Bayesian approach is preferred, one can obtain the non-normalized posterior distribution of $\vartheta$, using standard MCMC tools. For example, the estimates presented in this paper are obtained with a Random Walk Metropolis algorithm where, given initial $\vartheta_{-1}$, a $\Omega$, and a prior $g(\vartheta)$, candidate draws are obtained from $\vartheta_* = \vartheta_{-1} + \upsilon$ and $\upsilon$ is distributed $t(0, \kappa * \Omega, 5)$ where $\kappa$ is a tuning parameter and the draw accepted if the ratio $\chi_* = \frac{\breve{g}(\vartheta_*|y)}{\breve{g}(\vartheta_{-1}|y)}$, where $\breve{g}(\vartheta_i|y) = g(\vartheta_i)\mathcal{L}(y|\vartheta_i)$, $i = *, -1$, and $\mathcal{L}(y|\vartheta_i)$ is the likelihood of $\vartheta_i$, exceeds a uniform random variable. Iterated a large number of times, for $\kappa$ appropriately chosen, the algorithm ensures that the limiting distribution of $\vartheta$ is the target distribution (see e.g. [5], Ch. 9).

### 3.2 A comparison with existing literature

To the best of my knowledge, the literature has largely avoided from the problems I discuss in this paper. However, the paper relates to three existing branches of the macroeconomic and macroeconometric literature. The first branch attempts to robustify inference when the trend properties of data or of the model are misspecified (see [12] [20]). I share with [12] the point of view that economic theory has not much to say about non-cyclical fluctuations but rather than distinguishing between trend stationary or difference stationary non-cyclical fluctuations, and attempting to robustify inference, I am concerned with the generic mismatch between theoretical and empirical concepts of cyclical fluctuations and in designing a procedure which resolves the problem without taking a stand on the time series properties of the non-cyclical component. I share with [20] the idea of jointly estimating structural and auxiliary parameters without specifying the DGP of the data. However, I use a likelihood based estimator of the structural parameters, as opposed to a minimum distance estimator, which works regardless of the time series properties of the raw data. In addition, rather than assuming an arbitrary trend for one of the shocks, I assume that the DSGE model is build to explain only the cyclical component of the data - a much more common assumption in practice.

The second branch adds measurement errors to a cyclical model when estimation is performed (see e.g. [2] and [23]). I share with these authors the idea that there is more in the data than simply the fluctuations the model tries to capture. However, rather than using iid or VAR(1) noise, the setup I present allows a much richer structure for the measurement error while leaving the interpretation of the cyclical model unaffected.

The third branch points out that, in emerging markets, variations in trend growth are as important as cyclical fluctuations in explaining the dynamics of macroeconomic variables (see e.g. [1] and [3]). While the first paper is interested in characterizing differences between emerging and developing economies, the latter is concerned with the misuse of cyclical DSGE models in policy analyses for LDC countries. In this paper I show that the problems they highlight are generic and that one can conduct policy analyses in cyclical models without imposing controversial assumptions about the nature of the non-cyclical component.

# 4 The procedure in a controlled experiment

To appreciate the estimation distortions that standard data transformations imply and to quantify the advantages of the suggested approach, I use the same setup employed in section 2 and simulate data again assuming that the preference shock has two components, a nonstationary one and a stationary one. I then estimate the structural parameters using a number of data transformations in the most ideal situations one could consider. These include a prior centred at the true parameter vector and initial conditions in the estimation equal to the true parameter vector. I also estimates the structural and non-structural parameters of the flexible model using the raw data and the specification presented in section 3. For the sake of parsimony, the non-cyclical component of the data is restricted to have a double exponential smoothing format. The true values and the common priors for the structural parameters are in table 2. I use prior distributions with relatively large standard deviations to allow the likelihood to explore a wide portion of the parameter space without being strongly downweighted by the prior. For comparison I consider two cases: one where the non-cyclical disturbance has low relative variability $\sigma_{\chi}^{nc} = 0.0015$ and one where the non-cyclical disturbance has high relative variability $\sigma_{\chi}^{nc} = 0.015$. In the first case, the contribution of the non-cyclical component to the spectrum of the series is everywhere small. Thus, estimation distortions are primarily driven by filtering errors. In the second case, the contribution of the non-cyclical component to the spectrum of the series is of the same order of magnitude as the contribution of the cyclical component at most frequencies. Thus, the distortions present in standard transformations are driven by both filtering and specification errors.

| $\sigma_{\chi}^{nc} = 1.50$ | | | | | | | |
|---|---|---|---|---|---|---|---|
| | True | LT | HP | FOD | BP | Ratio1 | Flexible |
| | | Median (s.e.) | Median (s.e.) | Median (s.e.) | Median(s.e.) | Median(s.e.) | Median(s.e.) |
| $\sigma_n$ | 0.50 | 0.12( 0.02) | 0.21( 0.02) | 1.30( 0.05) | 0.08( 0.01) | 1.00( 0.04) | 0.24( 0.03) |
| $h$ | 0.70 | 0.91( 0.03) | 0.74( 0.03) | 0.71( 0.03) | 0.88( 0.03) | 0.11( 0.04) | 0.76( 0.05) |
| $\alpha$ | 0.30 | 0.07( 0.02) | 0.06( 0.02) | 0.04( 0.02) | 0.16( 0.02) | 0.04( 0.02) | 0.20( 0.05) |
| $\rho_r$ | 0.70 | 0.39( 0.04) | 0.46( 0.04) | 0.74( 0.03) | 0.36( 0.02) | 0.47( 0.05) | 0.34( 0.03) |
| $\rho_\pi$ | 1.50 | 1.41( 0.06) | 1.60( 0.06) | 1.63( 0.06) | 1.36( 0.05) | 1.50( 0.08) | 1.59( 0.08) |
| $\rho_y$ | 0.40 | 0.01( 0.00) | 0.01( 0.01) | -0.01( 0.00) | -0.01( 0.00) | 0.55( 0.07) | -0.01( 0.01) |
| $\zeta_p$ | 0.75 | 0.88( 0.03) | 0.85( 0.03) | 0.88( 0.03) | 0.90( 0.03) | 0.89( 0.03) | 0.83( 0.03) |
| $\rho_\chi$ | 0.50 | 0.40( 0.03) | 0.36( 0.03) | 0.69( 0.03) | 0.73( 0.03) | 0.37( 0.03) | 0.51( 0.04) |
| $\rho_z$ | 0.80 | 0.68( 0.04) | 0.69( 0.04) | 0.99( 0.03) | 0.80( 0.03) | 0.64( 0.03) | 0.79( 0.04) |
| $\sigma_\chi$ | 1.20 | 3.38( 0.41) | 0.35( 0.06) | 0.26( 0.05) | 0.33( 0.12) | 0.24( 0.04) | 0.27( 0.07) |
| $\sigma_z$ | 0.50 | 0.50( 0.11) | 0.21( 0.04) | 0.62( 0.11) | 0.32( 0.06) | 0.09( 0.01) | 0.22( 0.04) |
| $\sigma_r$ | 0.10 | 0.06( 0.01) | 0.06( 0.01) | 0.07( 0.01) | 0.06( 0.01) | 0.07( 0.01) | 0.05( 0.00) |
| $\sigma_\mu$ | 1.60 | 5.97( 0.42) | 0.80( 0.28) | 5.60( 0.34) | 6.62( 0.25) | 12.33( 0.73) | 1.56( 0.53) |
| $\sigma_\chi$ | 1.20 | 3.38( 0.41) | 0.35( 0.06) | 0.26( 0.05) | 0.33( 0.12) | 0.24( 0.04) | 0.27( 0.07) |
| | | | | | | | |
| $\sigma_{\chi}^{nc} = 0.15$ | | | | | | | |
| | True | LT | HP | FOD | BP | Ratio1 | Flexible |
| | | Median (s.e.) | Median (s.e.) | Median (s.e.) | Median(s.e.) | Median(s.e.) | Median(s.e.) |
| $\sigma_n$ | 0.50 | 0.18( 0.03) | 0.35( 0.06) | 0.89( 0.03) | 0.31( 0.04) | 0.95( 0.04) | 0.14( 0.01) |
| $h$ | 0.70 | 0.92( 0.03) | 0.91( 0.03) | 0.90( 0.03) | 0.97( 0.03) | 0.13( 0.04) | 0.79( 0.03) |
| $\alpha$ | 0.30 | 0.05( 0.02) | 0.07( 0.04) | 0.23( 0.01) | 0.14( 0.02) | 0.03( 0.02) | 0.15( 0.01) |
| $\rho_r$ | 0.70 | 0.53( 0.03) | 0.51( 0.02) | 0.58( 0.02) | 0.50( 0.02) | 0.36( 0.04) | 0.50( 0.02) |
| $\rho_\pi$ | 1.50 | 1.75( 0.07) | 1.67( 0.06) | 1.59( 0.05) | 1.77( 0.06) | 1.53( 0.07) | 1.57( 0.05) |
| $\rho_y$ | 0.40 | -0.01( 0.01) | -0.03( 0.01) | -0.03( 0.00) | -0.03( 0.00) | 0.67( 0.09) | 0.34( 0.02) |
| $\zeta_p$ | 0.75 | 0.86( 0.03) | 0.89( 0.03) | 0.86( 0.03) | 0.93( 0.03) | 0.87( 0.03) | 0.83( 0.03) |
| $\rho_\chi$ | 0.50 | 0.27( 0.04) | 0.22( 0.04) | 0.66( 0.02) | 0.60( 0.03) | 0.27( 0.05) | 0.60( 0.03) |
| $\rho_z$ | 0.80 | 0.68( 0.04) | 0.87( 0.03) | 0.98( 0.03) | 0.92( 0.03) | 0.59( 0.05) | 0.67( 0.03) |
| $\sigma_\chi$ | 1.20 | 0.39( 0.11) | 0.31( 0.08) | 4.23( 0.18) | 0.30( 0.06) | 0.18( 0.03) | 0.85( 0.16) |
| $\sigma_z$ | 0.50 | 0.23( 0.05) | 0.22( 0.04) | 3.37( 0.22) | 0.17( 0.02) | 0.06( 0.01) | 0.22( 0.04) |
| $\sigma_r$ | 0.10 | 0.06( 0.01) | 0.06( 0.01) | 2.61( 0.17) | 0.06( 0.01) | 0.07( 0.01) | 0.07( 0.01) |
| $\sigma_\mu$ | 1.60 | 0.93( 0.29) | 1.97( 0.50) | 5.13( 0.18) | 6.11( 0.28) | 3.60( 0.37) | 0.93( 0.11) |

Table 2: Parameters estimates, simulated data. LT refers to linearly detrended data, HP to Hodrick and Prescott filtered data, FOD to demeaned growth rates, BP to band pass filtered data, Ratio1 to real variables scaled by hours, and Flexible to the approach suggested in the paper.

Clearly, the distortions produced by standard approaches are important. Apart from producing estimates of utility and technology parameters which are quite biased and very much filter dependent, the persistence of the preference and of the technology shocks $\rho_\chi, \rho_z$ and the standard deviations of the preference and the markup shocks $\sigma_\chi$ and $\sigma_\mu$ are highly distorted. For example, the standard deviation of markup shocks is overestimated by a factor larger than 3, except with the HP filter where it is underestimated. In comparison, estimates of utility and technology parameters reported in the column labelled "Flexible" appear to be closer to the true values. In addition, both the persistence and the standard deviations of the shocks are better captured. Matching the persistence and the volatility of the shocks is important since conditional and unconditional moments crucially depend on these two sets of parameters. Thus, one would expect the flexible approach to be better equipped in characterizing the dynamic behavior of the endogenous variables. Note also that while with standard transformations, estimates depend on how strong the non-cyclical signal is relative to the cyclical one, this is much less the case for the procedure I suggest.

To understand the nature of the distortions generated when standard transformations are used, it useful to recall that the log-likelihood of the data can be represented as the sum of three terms $L(\theta|y_t) = [A_1(\theta) + A_2(\theta) + A_3(\theta)|y]$, see [21], where $A_1(\theta) = \frac{1}{\pi}\sum_{\omega_j} \log\det G_\theta(\omega_j)$, $A_2(\theta) = \frac{1}{\pi}\sum_{\omega_j}$ trace $[G_\theta(\omega_j)^{-1}F(\omega_j)]$, $A_3(\theta) = (E(y)-\mu(\theta))G_\theta(\omega_0)^{-1}(E(y)-\mu(\theta))$, $\omega_j = \frac{\pi j}{T}, j = 0, 1, \ldots, T-1$. $G_\theta(\omega_j)$ is the model based spectral density matrix of $y_t$, $\mu(\theta)$ the model based mean of $y_t$, $F(\omega_j)$ is the data based spectral density of $y_t$ and $E(y)$ the unconditional mean of $y_t$. Note that $A_2(\theta)$ and $A_3(\theta)$ are penalty functions: $A_2(\theta)$ sums deviations of the model-based from the data-based spectral density at various frequencies; $A_3(\theta)$, weights deviations of model-based from data-based means, with the spectral density matrix of the model at frequency zero.

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

Since estimates of $F(\omega_j)^*$ are imprecise, even for large $T$, there are only two situations when estimation biases are small. First, the non-cyclical component has low power at business cycle frequencies - in this case, the distortions induced by the penalty function are limited. This occurs when the volatility of the shocks driving the cyclical component is considerably larger than the volatility of the shocks driving the non-cyclical component (compare the two panels in table 2). Second, when Bayesian estimation is performed, the prior is selected to limit the distortions induced by the penalty function. This is very unlikely, however, since priors are not elicited with such a scope in mind.

While not very popular in estimation literature, one could also conceive to fit a transformed version of the model to transformed data, as it is done e.g. in [31]. To understand how parameter estimates are affected by this transformation note that, in this case, the log-likelihood is composed of $A_1(\theta)^* = \frac{1}{\pi}\sum_{\omega_j} \log|G_\theta(\omega_j)I_{\omega_j}|$ and $A_2(\theta)$ - since the actual and model data are filtered in the same way, the filter does not affect the penalty function. Suppose that $I_\omega = I_{[\omega_1,\omega_2]}$. Then $A_1(\theta)^*$ matters only at business cycle frequencies while the penalty function is present at all frequencies. Therefore, parameter estimates are adjusted so as to reduce the misspecification at all frequencies. Since the penalty function is more important at the low frequencies, parameters are adjusted to make $[G_\theta(\omega_j)]$ close to $\hat{F}(\omega_j)$ at those frequencies. Thus, the log-likelihood is willing to incur large fitting errors at frequencies where $\hat{F}(\omega_j)$ does not differ much from $G_\theta(\omega_j)$ - in particular, the medium and high frequencies. Consequently, the volatility of the shocks will be generally underestimated in exchange for overestimating their persistence - somewhat paradoxically, this procedure implies that the low frequency components of the data are those that matter most for estimation. Cross frequency distortions imply that agents think they are living in an economy which differs substantially from 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.

To highlight the properties of the flexible approach I perform two additional exercises. First, with the median estimates present in the top panel of table 4 I run the Kalman filter and construct time series estimates of the cyclical and non-cyclical component of output. Second, with the same median estimates, I simulate data from the flexible model (here $\mu_t$ is set to zero $\forall t$) and compare the autocorrelation function and the spectral density of the true and estimated components of output.

Figure 4 reports the results. The first two panels clearly show that the true and the estimated cyclical and non-cyclical components of output display similar time series and volatility properties. In addition, the rate of decay of the autocorrelation functions of the true and the estimated cyclical and non-cyclical components is practically identical. Finally, as anticipated in the previous section, the estimated cyclical and non-cyclical components have power at all frequencies of the spectrum as required by the theory, and at business cycle frequencies (indicated by the vertical bars in the two panels of the last row of graphs), the two components are equally important.

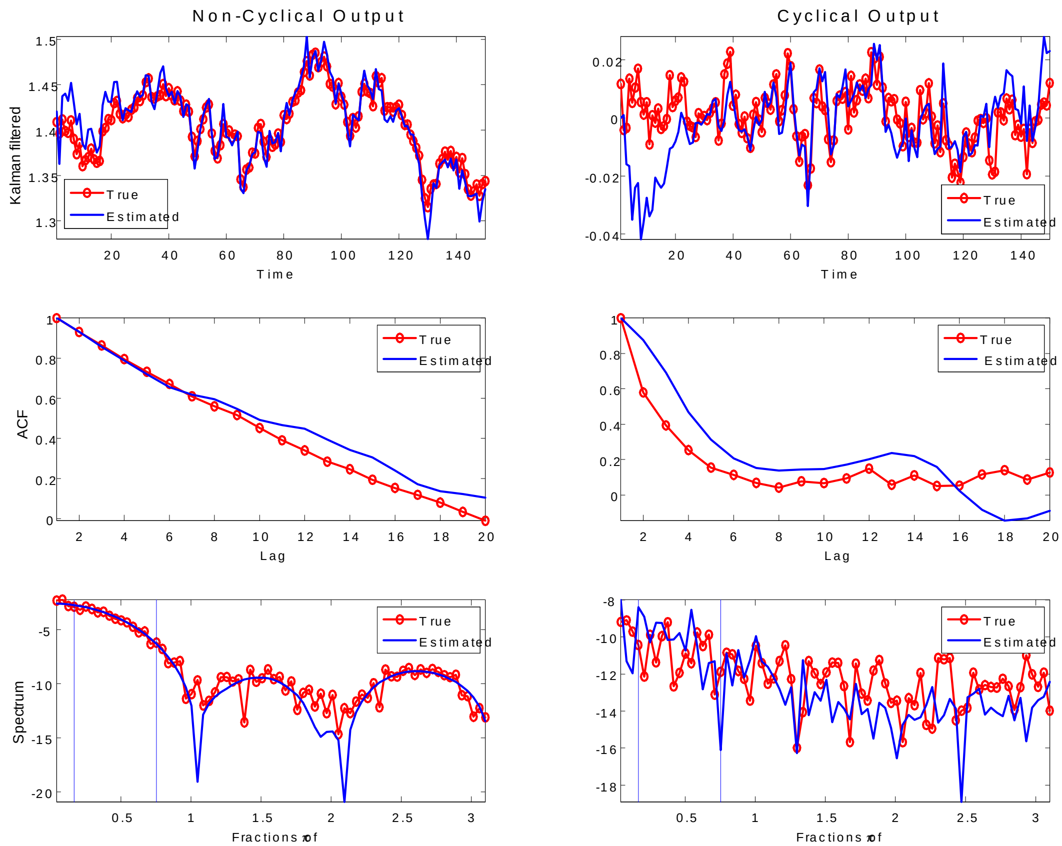


Figure 4: Output decompositions, true and estimated with a flexible approach. Vertical bars indicate the frequencies where cycles with 8-32 quarters periodicities are located.

Interestingly, not only the reduced form properties of the cyclical component of the endogenous variables are well captured by the flexible approach; the conditional dynamics in response to shocks are also well described. Figure 5, which presents impulse responses obtained from the cyclical model with true and estimated parameters, indicates that the methodology is able to capture both the sign and the persistence of the responses. Clearly, magnitudes are imprecisely estimated - this would be case even if we double the sample size. But overall, the approach does a good job in reproducing the main qualitative features of the DGP, even with a relatively small sample. Thus, not only structural estimates are more reliable; economic inference is likely to be less prone to "mismatch" distortions.

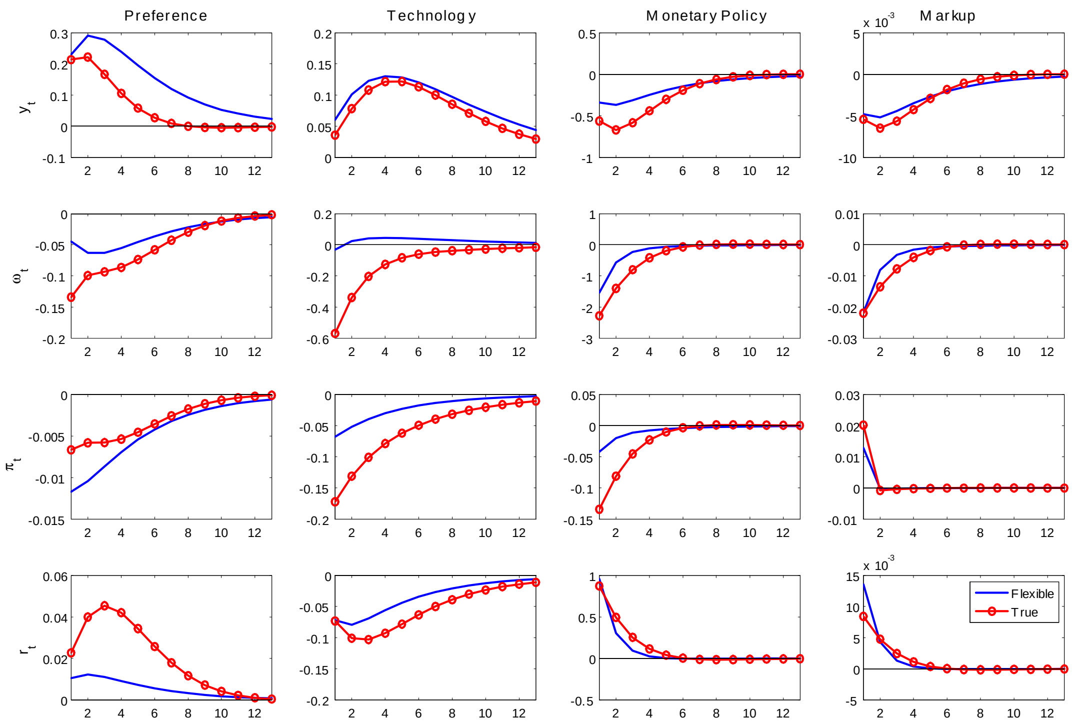


Figure 5: Model based impulse responses, true and estimated with flexible approach.

# 5 Does it make a difference for inference which approach one uses?

In this section I show that the approach I suggest not only captures the statistical properties of the data better than standard transformations; it also produces significantly different interpretations of the economic phenomena. To do this, I examine interesting functions of the parameters in the toy model I have been working with so far and in a more standard medium scale DSGE model, which has been widely used in the recent macroeconometric literature.

## 5.1 The slope of the Phillips curve and policy activism

Two objects which are often of interest in the policy discussion, are the slope of Phillips curve, $k_p$, which measures how much changes in marginal costs (real activity) are translated into inflation changes and the policy activism parameter $\frac{\rho_y}{\rho_\pi - 1}$, which gives a sense of how important the output stabilization objective is in the policy rule of the Central Bank relative to the inflation stabilization objective. Estimates of $\kappa_p$ are relatively similar across standard data transformations (see bottom of table 1) with values fluctuating from 0.01 to 0.06, but standard deviations are very tight, making even small posterior differences statistically significant. Applying the flexible approach I suggest to the same data set and imposing a joint smoothness restriction on the volatility of the non-cyclical component of output and real wages, yields a point estimate for $\kappa_p$ of 0.55 with a posterior standard error of 0.12. Thus, while the Phillips curve is almost vertical when parameters are estimated with standard transformations, there are important spillovers from marginal costs to inflation when the parameters are estimated with the flexible approach.

What are the features of the monetary policy rule in place during the "Great Inflation" of the 1970s and the return to norm of the 1980s and 1990s? Is the characterization offered by the flexible approach different from the one provided by standard transformations? In figure 5 I plot the posterior density of the policy activism parameter obtained when the data is linearly detrended (left box) or HP filtered (central box) prior to estimation and when the flexible approach is employed to estimate the structural parameters (right box) for the samples 1964:1-1979:4 and 1984:1-2007:4. The posterior density of the policy activism parameter shifts to the left in the second sample when HP filtered data is used and, for example, the posterior median moves from -0.23 in the first sample to -0.33 in the second. This left shift of the posterior density is absent when LT data is used in the estimation and, if anything, the median of the posterior in the second sample increases from -0.38 to 0.12 - note that for the 1984-2007 sample the posterior is bimodal, suggesting the presence of two regimes within the second sample. In both cases the Kolmogorov-Smirnov statistic rejects the null that the posterior distributions are the same. Thus, there appears to be a structural break, even though it is not clear in which direction the movement is: with LT data, output gap considerations have become relatively less important; with HP data, the opposite appears to be true. When the flexible approach is used to estimate the structural parameters, the posterior densities of $\frac{\rho_y}{\rho_\pi - 1}$ are quite spread out in the two samples and considerably overlap. Interestingly, although the median of the posterior marginally decreases (from 0.22 to 0.l6) when we move from the first to the second sample, the change is statistically insignificant. Similarly, the Kolmogorov-Smirnov statistic does not reject the null that the distributions in the two samples are the same. Thus, the evidence in favour of a structural break in the conduct of monetary policy appears to be much weaker.

Why are the results so different? Apart from the distortions in parameter estimates mentioned in previous subsections, standard transformations fail to account for the uncertainty present in the specification of the non-cyclical component. Thus, by underestimating the spread of the posterior densities, they give a falsely sharper characterization of the changes the data display.

## 5.2 Sources of output and inflation fluctuations

In standard medium scale cyclical DSGE models, like the one employed in [30] and [31], important macroeconomic variables are primarily driven by markup shocks. Since these shocks are an unlikely source of cyclical fluctuations, [9] have argued that misspecification is likely to be present (see [25]

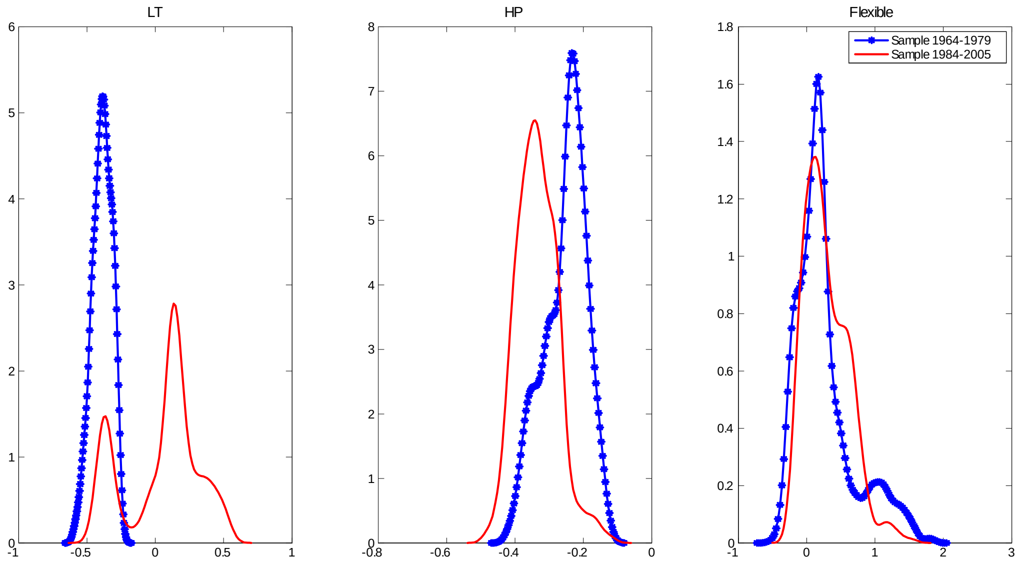


Figure 6: Posterior distributions of policy activism parameter, samples 1964:1-1979:4 and 1984:1-2007:4. LT refers to linearly detrended data, HP to Hodrick and Prescott filtered data and Flexible to the approach this paper suggests.

for an alternative interpretation of these shocks). Researchers working with this type of models use filtering devices to fit the model to the data (as in [30]), arbitrarily data transformations (as in [31]) or build a non-cyclical component in the model (as in [25]) and use model-consistent data transformations to estimate the structural parameters. What would the approach of this paper tell us about sources of cyclical fluctuations in output and inflation relative to standard transformations? To answer this question, I take the same model and the same data set used in [31] but I modify the setup in four ways. First, I do not allow MA terms in price and wage markup disturbances: all shocks have a standard AR(1) structure. Second, the model is solved in deviations from the steady state, rather than from the flexible price equilibrium, which is the most common setup. Third, no rescaling of the shocks is performed. Fourth, the policy rule does not include a term concerning output growth, again a more standard choice.

Table 3 reports results obtained eliminating a linear trend from the variables, taking growth rates of the real variables and demeaning nominal ones and using the flexible link suggested in the paper. When a linear trend is removed from the variables the forecast error variance decomposition of output at the five years horizon is indeed primarily driven by price markup shocks, with a considerably smaller contribution of investment specific and preference shocks. For inflation, price markup shocks account for almost 90 percent of the forecast error variability at the five years horizon. When the model is instead fitted to growth rates, price markup shocks account for over 90 percent of the variability of both output and inflation at the five years horizon. Thus, even without some of the standard bells and whistles used in the literature, the conclusion that markup shocks dominate remains. When the flexible bridge I suggest in this paper is used, the non-cyclical component of real variables is restricted to have a common structure (there are only two parameters simultaneously controlling the non-cyclical component of output, consumption, investment) and a measurement error is allowed in each equation, the picture is considerably different. Output fluctuations at the five year horizon are driven almost entirely by preference disturbances. On the other hand, inflation fluctuations are jointly accounted for by wage markup, TFP and price markup disturbances. Hence, while it is still true that the model is less structural than one would like - " black box" disturbances still dominate - the role of markup shocks is considerably reduced when the flexible approach proposed in this paper is used to estimate the parameters of the cyclical model.

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

Table 3: Variance decomposition at the 5 years horizon. Estimates are obtained using the median of the posterior of the parameters. A (*) indicates that the 68 percent highest credible set is entirely above 0.10. The model and the data set used are the same as in [31]. LT refers to linearly detrended data, FOD to growth rates and Flexible to the approach this paper suggests.

# 6 Conclusions

I have argued that estimating cyclical DSGE models with either transformed or filtered data is theoretically incorrect and may lead to serious distortions in the estimates of the structural parameters. There are two reasons for the distortions. First, the transformed/ filtered data imperfectly measures fluctuations appearing at frequencies corresponding to cycles of 8-32 quarters. Second, the cyclical component that a model produces is not entirely located at these frequencies and, viceversa, the non-cyclical component may have important power at business cycle frequencies. The consequences of these two specification errors could be important, both statistically and economically, because income and substitution effects are distorted, the volatilities and persistence of the shocks over or underestimated and the decision rules of the agents altered.

I propose an alternative methodology which allows researchers to estimate cyclical DSGE models using the raw data. The procedure has several advantages over competitors: there is no need to build an arbitrary non-cyclical component into a DSGE model nor to worry about its exact time series features; the procedure eliminates by construction the first source of measurement error and considerably reduces the second because the spectrum of the data is endogenously split into a cyclical and a non-cyclical part; it takes into account the uncertainty in the specification of the non-cyclical component when deriving estimates of the structural parameters; and it reduces identification problems that model-based transformations produce. Finally, the specification I suggest allows us to bring into the structural cyclical model aspects of the reduced form non-cyclical portion when this is deemed useful without any need to alter the estimable structure.

I have shown both the distortions induced by standard filtering approaches and the properties of the alternative methodology using data generated from a simple New Keynesian model and compared the posterior distribution of interesting economic quantities in the actual data. Clearly, the problems I highlight are general and could be important in a variety of situation. The flexible approach seems to work well both when problems are likely to be important and when they are likely to be minor.

## References


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

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

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

[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, Multiple filtering devices for the estimation of cyclical DSGE models, forthcoming, *Quantitative Economics*, 2011.

[7] Fabio Canova and Matthias Paustian, Measurement with some theory: using sign restrictions to evaluate business cycle models, 2010.

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

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

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

[11] Lawrence J. Christiano, Martin Eichenbaum, and Charles L. Evans, Nominal rigidities and the dynamic effects of a shock to monetary policy, *Journal of Political Economy*, 113, 1–43, 2005.

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

[13] Diego Comin and Mark Gertler, Medium term cycles, *American Economic Review*, 96, 523–551, 2006.

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

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

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

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

[18] Filippo Ferroni, Trend agnostic one step estimation of DSGE models, manuscript,2008.

[19] Eric Ghysels, A study toward a theory of dynamic seasonality in economic time series, *Journal of the American Statistical Association*, 82, 168-172,1988.

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

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

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

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

[24] Peter N. Ireland, Stochastic Growth in the US and the Euro Area, forthcoming, *Journal of the European Economic Association*, 2011.

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

[26] Ellen McGrattan, Measurement with minimal theory, manuscript, 2006.

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

[28] Julio Rotemberg and Mike Woodford, An optimization based econometric framework for the evaluation of monetary policy, *NBER Macroeconomic Annual*, 14, 297–346, 1997.

[29] Tom Sargent, Comment to "seasonal adjustment and multiple time series analysis" by K. Wallis in. *Seasonal Analysis of Economic time series, A. Zellner (ed.)*, pages 361–364, 1978.

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

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

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

# Appendix (not intended for publication)

## A. The basic DSGE model

The bundle of goods consumed by the representative household is

$$C_t = \left( \int_0^1 C_t(j)^{\frac{\epsilon_t - 1}{\epsilon_t}} dj \right)^{\frac{\epsilon_t}{\epsilon_t - 1}} \tag{15}$$

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

$$C_t(j) = \left( \frac{P_t(j)}{P_t} \right)^{-\epsilon_t} C_t \tag{16}$$

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

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[ X_t \frac{1}{1-\sigma_c} (C_t - hC_{t-1})^{1-\sigma_c} - \frac{1}{1+\sigma_n} N_t^{1+\sigma_n} \right] \tag{17}$$

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

$$\chi_t = \rho_\chi \chi_{t-1} + \epsilon_t^\chi \tag{18}$$

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

$$P_t C_t + b_t B_t = B_{t-1} + W_t N_t \tag{19}$$

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:

$$Y_t(j) = Z_t N_t(j)^{1-\alpha} \tag{20}$$

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

$$z_t \quad = \rho_z z_{t-1} + \epsilon_t^z \tag{21}$$

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

$$\Pi_t^{1-\epsilon_t} = \zeta_p + (1-\zeta_p)(P_t^*/P_{t-1})^{1-\epsilon_t} \tag{22}$$

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

$$\max_{P_t^*} \sum_{k=0}^{\infty} \zeta_p^k E_t Q_{t,t+k} \left[ P_t^* Y_{t+k|t} - TC_{t+k}(Y_{t+k|t}) \right] \tag{23}$$

subject to the sequence of demand constraints

$$Y_{t+k|t} = \left( \frac{P_t^*}{P_{t+k}} \right)^{-\epsilon_{t+k}} Y_{t+k} \tag{24}$$

$k = 0, 1, 2, ...$ 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 sets the nominal interest rate according to

$$r_t = \rho_r r_{t-1} + (1 - \rho_r)(\rho_\pi \pi_t + \rho_y gdp_t) + \epsilon_t^{ms} \tag{25}$$

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

The first order conditions of the optimization problems are:

$$0 = X_t (C_t - hC_{t-1})^{-\sigma_c} - \lambda_t \tag{26}$$

$$0 = -N_t^{-\sigma_n} - \lambda_t \frac{W_t}{P_t} \tag{27}$$

$$1 = E_t \left[ \beta \frac{\lambda_{t+1}}{\lambda_t} \frac{P_{t+1}}{P_t} R_t \right] \tag{28}$$

$$0 = \sum_{k=0}^{\infty} \zeta_p^k E_t Q_{t,t+k} Y_{t+k|t} \left[ P_t^* - \mathcal{M}_{t+k} MC_{t+k|t}^n) \right] \tag{29}$$

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

$$\mathcal{M}_t = \mu e^{\epsilon_t^\mu} \tag{30}$$

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

Market clearing requires

$$Y_t(j) = C_t(j) \tag{31}$$

$$N_t = \int_0^1 N_t(j) dj \tag{32}$$

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

## B. The basic DSGE model with non-stationary preference shocks

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

$$Y_t^o = \nu(\theta^*, \vartheta^*) + H^{ns} x_t^{ns} + H^s x_t^s \tag{33}$$

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

$$x_{t+1}^s = \Phi(\theta, \vartheta) x_t^s + \Psi(\theta, \vartheta)\eta_{t+1} \qquad \eta_t \sim N(0, \Sigma(\theta, \vartheta)) \tag{34}$$

where $\eta_t$ is the vector of the structural innovations of the shock processes, $\eta_t = [\eta_t^{ns}, \eta_t^s]'$. It turns out that, for the particular model we have chosen, these equations are given by

$$w_t = \sigma_n(y_t - z_t) - \chi_t + \frac{1}{1-h}(y_t - hy_{t-1} - he_t^{X,P}) \tag{35}$$

$$r_t = \rho_r r_{t-1} + (1-\rho_r)(\rho_y y_t + \rho_\pi \pi_t) + v_t \tag{36}$$

$$\chi_t - \frac{1}{1-h}(y_t - hy_{t-1} - he_t^{X,P}) = E_t[\chi_{t+1} - \frac{1}{1-h}(y_{t+1} - hy_t - he_{t+1}^{X,P}) + r_t - \pi_{t+1} + \sigma_n e_{t+1}^{X,P}] \tag{37}$$

$$\pi_t = \beta E_t \pi_{t+1} + \kappa_p (w_t + z_t + \mu_t) \tag{38}$$

where $\kappa_p = \frac{1-\alpha}{1-\alpha+\alpha\epsilon}\frac{(1-\beta\xi_p)(1-\xi_p)}{\xi_p}$. The vector of non-stationary shock processes $\log X_t^P$is assumed to follows

$$\ln X_t^P = \ln X_{t-1}^P + e_t^{X,P} \tag{39}$$

while the vector of stationary shock processes is

$$\log z_t = \rho_z \log z_{t-1} + e_t^z \tag{40}$$

$$\log \chi_t = \rho_\chi log \chi_{t-1} + e_t^\chi \tag{41}$$

$$v_t = e_t^v \tag{42}$$

$$\mu_t = e_t^\mu \tag{43}$$

Thus:

$$x_t^s = [y_t, w_t, \pi_t, r_t, z_t, \chi_t]' \quad (44)$$

$$x_t^{ns} = \ln X_t^P \quad (45)$$

$$\eta_t^s = [e_t^z, e_t^\chi, v_t, \mu_t]' \quad (46)$$

$$\eta_t^{ns} = e_t^{X,P} \quad (47)$$

$$\nu(\theta, \vartheta) = [\ln y_s, \ln W_s, \ln \Pi_s, \ln R_s]' \quad (48)$$

$$H^{ns} = [1, 0, 0, 0]' \quad (49)$$

$$H^s = \begin{pmatrix} I_{4\times 4} & 0_{4\times 2} \\ & \end{pmatrix} \quad (50)$$

$$\theta = [h, \sigma_n, \rho_r, \rho_y, \rho_\pi, k_p, \rho_z, \rho_\chi, \sigma_z, \sigma_x, \sigma_r, \sigma_\mu] \quad (51)$$

$$\vartheta = \sigma_{X,P} \quad (52)$$

## C. The medium scale DSGE model

We only briefly sketch the log-linearized conditions used since the non-linear equations and their transformations are fully described in the appendices of [31] and [**?**].

(a): The variables of the model

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

(b): The parameters of the model

| Label | Definition |
|---|---|
| $\sigma_c$ | elasticity of intertemporal substitution |
| $\sigma_l$ | elasticity of labor supply with respect to real wages |
| $h$ | habit persistence parameter |
| $\delta$ | depreciation rate |
| $\phi_p - 1$ | share of fixed costs in production |
| $\chi$ | steady state elasticity of capital adjustment cost function |
| $\psi$ | positive function of the elasticity of capital utilization adjustment costs function. |
| $\alpha$ | share of capital services in production |
| $\gamma_p$ | price indexation parameter |
| $\zeta_p$ | price stickiness parameter |
| $\epsilon_p$ | curvature of good market aggregator |
| $\gamma_w$ | wage indexation parameter |
| $\zeta_w$ | wage stickiness parameter |
| $\epsilon_w$ | curvature of labor market aggregator |

| Label | Definition |
|---|---|
| $\lambda_r$ | interest smoothing parameter |
| $\lambda_\pi$ | inflation parameter |
| $\lambda_y$ | output parameter |
| $gy$ | government expenditure to output ratio |
| $ky$ | steady state capital output ratio |
| $r_* = \beta^{-1}$ | steady state rental rate |
| $w_*$ | steady state real wage rate |
| $N_*/C_*$ | steady state hours to consumption ratio |

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

| | |
|---|---|
| $y_t = (1 - gy - \delta\ ky)c_t + \delta\ ky\ i_t + r_*\ ky\ z_t + g_t$ | (C.1) |
| $c_t = \frac{h}{1+h}E_t c_{t+1} + \frac{h}{1+h}c_{t-1} - \frac{(\sigma_c - 1)w_* N_*/C_*}{(1+h)\sigma_c}(N_t - E_t N_{t+1}) - \frac{1-h}{(1+h)\sigma_c}(R_t - E_t\pi_{t+1} + e_t^b)$ | (C.2) |
| $i_t = \frac{\beta}{1+\beta}E_t i_{t+1} + \frac{1}{1+\beta}x_{t-1} + \frac{\chi^{-1}}{1+\beta}q_t + e_t^i$ | (C.3) |
| $q_t = \beta(1-\delta)E_t q_{t+1} + (1 - \beta(1-\delta))E_t r_{t+1} - (R_t - E_t\pi_{t+1} + e_t^b)$ | (C.4) |
| $y_t = \phi_p(\alpha k_t^s + (1-\alpha)N_t + e_t^a)$ | (C.5) |
| $k_t^s = k_{t-1} + z_t$ | (C.6) |
| $z_t = \frac{1-\psi}{\psi}r_t$ | (C.7) |
| $k_{t+1} = (1-\delta)\ k_t + \delta\ i_t + \delta\ (1+\beta)\ \psi\ e_t^i$ | (C.8) |
| $\mu_t^p = \alpha(k_t^s - N_t) + e_t^a - w_t$ | (C.9) |
| $\pi_t = \frac{\beta}{1+\beta\gamma_p}E_t\pi_{t+1} + \frac{\gamma_p}{1+\beta\gamma_p}\pi_{t-1} - T_p\mu_t^p + e_t^p$ | (C.10) |
| $r_t = -(k_t - N_t) + w_t$ | (E.11) |
| $\mu_t^w = w_t - (\sigma_l N_t + (1-h)^{-1}(c_t - hc_{t-1})$ | (C.12) |
| $w_t = \frac{1}{1+\beta}w_{t-1} + \frac{\beta}{1+\beta}(E_t\pi_{t+1} + E_t w_{t+1}) - \frac{1+\beta\gamma_w}{1+\beta}\pi_t + \frac{\gamma_w}{1+\beta}\pi_{t-1} - T_w\mu_t^w + e_t^w$ | (C.13) |
| $R_t = \lambda_r R_{t-1} + (1-\lambda_r)(\lambda_\pi\pi_t + \lambda_y y_t) + e_t^r$ | (C.14) |

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

(d): The process for the shocks

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

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

## D. Additional Tables

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

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