data removes the trend specification errors, the incorrect structural model by assuming the stationarity of hours is spreading sizable distortions to the entire spectrum, corrupting all the estimates. Thus, while it is commonly perceived that model-based transformations are superior to statistical transformation, and this is clearly true if the underlying model is correctly specified, such a perception may be misleading when misspecification is present. Thus, it is incorrect to dismiss statistical filtering outright: there are situations like the one considered in this example, where they may provide a much better handle on the features of the cyclical component than model-based transformations.

What I would like to show is that, under misspecification, a one-step approach is superior to both types of two-step transformations.

1.2 The literature

Since the seminal paper of Cogley (2001), few papers have analyzed the impact of misspecifications of the time series components of the data on structural parameter estimates. Fukac and Pagan (2010) propose a limited information method to deal with this problem but their analysis is confined to a single equation framework. Gorodnichenko and Ng (2010) extend Cogley's analysis and propose a robust approach which exploits all the cross-equations restrictions of the model. For estimation they use the simulated method of moments, which is prone to severe identification problems (see Canova and Sala (2009)). I share with these papers an agnostic view about the properties of the non-cyclical component of the data, but my approach differs in three respects.

First, the proposed setup is flexible enough to permit model comparisons. Second, 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 macroeconomics - and link the model and the observables through flexible specifications. Third, I employ a structural times series approach and likelihood based methods, as in Canova (2009); this avoids any data transformation before or during estimation. While Canova focuses on a unique representation of the non-cyclical component that encompasses various low frequencies behavior, the proposed estimation strategy exploits the posterior weights to average across many potential specifications, making structural parameters robust to this form of uncertainty. Finally, Canova and Ferroni (2011) propose a multiple filters (MF) method where data filtered with alternative procedures are treated as contaminated proxy of the relevant model-based quantities and the estimation is jointly carried on structural and non-structural parameters. While with the MF method one can only make inference on the cyclical portion of the data, with a one-step approach, one can draw conclusions for levels of the variables.

2 Econometric Methodology

In this section, I develop the statistical framework I use to estimate the structural parameters of the model. The main idea of the one-step approach is to compute the likelihood of a system that embodies a reduced form representation for the non-cyclical component and a structural representation for the cyclical component.

I assume that we observe $y = \{y_t\}_{t=1}^T$, the log of a vector of times series. I postulate that the data is made up of a non-cyclical component, y_t^τ, and a cyclical component, y_t^c, so that

$$y_t = y_t^\tau + y_t^c \tag{1}$$

I also assume that the cyclical behavior of the data can be described by a DSGE model whose linear solution is given by

$$y_{t+1}^\dagger = \Phi(\theta^m) y_t^\dagger + \Psi(\theta^m)\nu_{t+1} \tag{2}$$

where $y_{t+1}^\dagger$ represents the variables in the DSGE model; Φ and Ψ are matrices which are functions of the structural parameters of the model, θ^m; ν_{t+1} are mutually uncorrelated innovations of the strucural model.

2.1 TWO-STEP APPROACH

With the two-step (2s) approach data is first filtered and then structural DSGE parameters estimated. Assume that $\mathcal{F}_\tau(y_t)$ is the filter that extracts the non-cyclical component, y_t^τ, from the data. Then, the resulting cyclical component is $y_t^c = y_t - \mathcal{F}_\tau(y_t)$. I consider three types of filters: a linear detrending filter, a first order difference filter, and the Hodrick-Prescott filter.
Once y_t^c is obtained, we link the dynamics of the DSGE model with the cycles through a state space representation, that is

$$\begin{aligned} y_{t+1}^\dagger &= \Phi(\theta^m) y_t^\dagger + \Psi(\theta^m)\nu_{t+1} \\ y_t^c &= S y_t^\dagger \end{aligned}$$

The second equation links the model-based quantities with data counterparts and S is a selection matrix which picks the variables that are observable or interesting from the point of view of the researcher.

It has been widely recognized that the filtering mechanism, $\mathcal{F}_\tau(y_t)$, has a strong influence on the properties, i.e. volatility and autocorrelation, of the detrended data (see Harvey (1985) and Canova (1998)). I acknowledge this problem in the context of DSGE models estimation. In a nutshell, different data transformations deliver distinct structural parameter estimates. As a consequence, the economic implications of DSGE models can vary substantially. The joint estimation of structural and filtering parameters allows to determine the best filtering technique via a Posterior Odds comparison.

2.2 ONE-STEP APPROACH

In the one-step approach (1s) the likelihood is computed directly from the observables, y_t, that is

$$\begin{aligned} y_t &= y_t^\tau + y_t^c \\ y_t^\tau &= \mathcal{F}_\tau(y_t) \\ y_t^c &= S y_t^\dagger \\ y_{t+1}^\dagger &= \Phi(\theta^m) y_t^\dagger + \Psi(\theta^m)\nu_{t+1} \end{aligned}$$

As long as the filter is linear, it is easy to cast the one-step setup into a state space representation (details are in the appendix).

2.2.1 Linear-Trend-DSGE setup

In this specification, I assume that the non-cyclical component of the data is

$$y_t^\tau = A + B * t + \eta_t \tag{3}$$

where A and B are column vectors, η_t is a white noise normally distributed with zero mean and variance covariance matrix, Σ_η. Therefore, the filter parameters to be estimated are $\theta^{lt} = [A, B, \Sigma_\eta]$. I assume that the covariance matrix is diagonal, but it is straightforward to specify a full matrix, and choose one specification or the other using the Posterior Odds ratio. Similarly, if there is a strong belief that a subset of data do not display any type of non-cyclical movements, some elements of the A and B vectors can be set to zero; the restricted and the unrestricted specifications can then be compared with a Posterior Odds ratio. I will refer to this set-up as lt-dsge setup.

2.2.2 First-Difference-DSGE setup

In this specification I assume that the non-cyclical component of the data can be represented as

$$y_t^\tau = \gamma + \Gamma y_{t-1}^\tau + \eta_t \tag{4}$$

where γ is the drift and Γ is a diagonal matrix, that have zeros or ones on the main diagonal, η_t is a white noise normally distributed with zero mean and a diagonal variance covariance matrix, Σ_η. Therefore, the filter parameters to be estimated are $\theta^{fd} = [\gamma, \Sigma_\eta]$. As before, various specifications for Σ_η could be considered and compared against each other. Similarly, one can consider different specifications of the matrix Γ; for example, if data contains hours worked, one could specify a version where they are stationary and one where they are not, and choose the best specification via Posterior Odds comparison. I will refer to this specification as fd-dsge setup.