FAQ: How do I extract the output gap?

Fabio Canova *
Norwegian Business School, CAMP and CEPR

February 7, 2020

Abstract

I study potentials and gaps, permanent and transitory fluctuations in macroeconomic variables using the Smets and Wouter (2007) model. Model-based gaps display low frequency variations; possess more than business cycle fluctuations; have similar frequency representation as potentials, and are correlated with them. Permanent and transitory fluctuations display similar features, but are uncorrelated. I use a number of filters to extract trends and cycles using simulated data. Gaps are best approximated with a polynomial filter; transitory fluctuations with a differencing approach, but distortions are large. Explanations for the results are given. I propose a filter which reduces the biases of existing procedures.

Key words: Gaps and potentials, permanent and transitory components, filtering, cyclical fluctuations, gain functions.

JEL Classification: C31, E27, E32.

*I thank Paolo Gelain, Jesper Linde' and the participants of seminars at the Riksbank, ECB, University of Cardiff, First MacroDolomiti Conference, Castelrotto , and the WGEM workshop, Paris for comments and suggestions.

1 Introduction

Since the 2008 financial crisis, academic economists and policymakers have been keenly interested in the level of certain latent variables (such the output gap, the natural rate of interest, the NAIRU, etc.) and in the changes that the crisis has brought about in nature of the cyclical fluctuations in inflation and unemployment. Unfortunately, while the profession agrees on the centrality of these issues, in practice, different users draw conclusions looking at different quantities. For example, the term output gap is interchangeably employed to refer to the difference between the actual output and its potential, defined as the level prevailing absent nominal frictions; between actual output and its permanent component; or between actual output and its (statistical long-run) trend. However, potential output may not be trending (in a statistical sense), and may feature both permanent and transitory swings. Similarly, the natural rate of interest is, at times, defined with reference to the frictions present in a model and, at other times, as the long run component of real interest rates. There are also a number of competing definition of cyclical fluctuations, see e.g. Pagan [2019], and different researchers use different approaches to extract them. Clearly, without a consensus on what the objects of interest are, the measurement of latent quantities becomes elusive. However, even if a consensus could be reached, the available tools are statistical in nature and do not generally employ, even in a reduced form sense, the information provided by the structural models economists use to discuss the features of latent variables. Hence, economic analyses and policy prescriptions may become whimsical.

Perhaps unsurprisingly, the lapse between theory and measurement creates confusion in the profession. In recent years macroeconomists have argued about what the data tells us about, e.g., potential output (see Coibon, Gorodnichenko, and Ulate [2018]), the long run properties of the natural rate of interest (see Laubach and William [2015]), or the dynamics of NAIRU (see Crump, Giannoni, and Sahini [2019]); what is the best tool to extract cyclical fluctuations (see Hamilton [2018]); whether permanent or transitory disturbances are responsible for macroeconomic fluctuations (see Schmitt-Grohe and Uribe [2019]); or which theory is consistent with business cycle facts (see Angeletos, Collard, and Dellas [2019]). Canova [1998] and Canova [1999] demonstrated that statistical methods used to separate one observable variable into two latent components (trend and cycle, for short) produce time series with different properties. Thus, without a firm stand on what the features of the objects of interest are and a reliable mapping between theoretical and statistical quantities, it is impossible to fruitfully select an extraction method for applied exercises. Canova [2014] reiterated the argument by showing that fitting stationary structural models to the output of standard statistical tools results in heterogeneous estimates of the structural parameters and in different dynamics in response to structural disturbances, both of which make inference difficult. More recently, Beaudry, Galizia, and Portier [2018] showed that there are interesting fluctuations in hours that standard filtering approaches disregard, see also Lubik, Matthes, and Verona [2019]. They also argue that these fluctuations could help researchers to understand better the type of models which are consistent with the data; see Kulish and Pagan [2019] for a critical view.

This paper attempts to shed light on the relationship between theoretical notions of gaps (transitory fluctuations) and the cycle one recovers with statistical approaches using a laboratory economy. I conduct a Monte Carlo exercise employing a version of the Smets and Wouters [2007] model as the data generating process (DGP). This model seems a natural starting point for the exercise for three reasons: it has a good fit to the data of many countries; it has been used to analyze policy trade-offs and optimal monetary policy decisions, see e.g. Justiniano, Primiceri, and Tambalotti [2013]; many policy institutions use versions of this model for their out-of-sample forecasting exercises.

In such a model, latent variables are well defined objects. Potentials are the equilibrium outcomes obtained eliminating nominal frictions, markup and monetary disturbances and the gaps are the deviations between the level variables and the potential outcomes. Similarly, the permanent component is what the model produces when certain disturbances have permanent features, while the transitory component is the difference between the level variables and the permanent component. With this taxonomy in mind, one can evaluate which statistical approach produces estimates which are close to those of model-based gaps (transitory components) and examine the reasons for why distortions occur. While the quantitative results I present are clearly model-dependent, the features responsible for the distortions are independent of the details of the lab economy. In particular the calibration of the experiment; the number or the properties of the disturbances; the presence of financial frictions, or the underlying principles used to construct the equations of the model do not affect the conclusions. I show instead that it is the interaction between the spectral properties of the latent variables and the characteristic of filters which accounts for the outcomes I obtain. In particular, the fact that gaps and potentials (transitory and permanent components) have similar spectral features, and similar distribution of variance by frequency renders standard approaches incapable of consistently estimating the latent variable of interest.

I employ numerous filters in the exercise covering well the procedures most commonly used in practice. Eight approaches are univariate (polynomial, Hodrick and Prescott, first order and long order differencing, unobservable component, band pass, wavelet, Hamilton local projection); two are multivariate (Beveridge and Nelson, Blanchard and Quah). Because many filters have free parameters, I also examine whether the ranking change when they are are set at different values.

The conclusions I obtain are somewhat surprising. When the object of investigation are the gaps, polynomial filtering minimizes the distortions relative to model-based quantities on average, across series and statistics. Because gaps and potentials display similar low frequencies components and similar distribution of spectral variability, polynomial filtering dominates because it leaves almost undistorted the distribution of spectral variability of the data while all other filters heavily change it. This conclusion is independent of the sample size and of the settings of the free parameters. When the transitory components are the object of investigation, the conclusions are less firm. Polynomial filtering is still performing well, on average across variables and statistics, in large but not in small samples. Still, simple approaches, such as first or long order differencing, are doing better than more elaborate local projection or unobservable component procedures, both in large and small samples. Also in this case, the assignment of low and business cycle frequency variations is crucial to explain the results. Procedures which simultaneously account for the fact that, at low frequency the transitory component matters and at business cycle frequencies the permanent component is important, and that the transitory component has more power at low than at business cycle frequencies come closest in capturing the dynamic features of model-based transitory components.

The presence of important low frequency fluctuations in the theory-based gaps (transitory components) and of business cycle fluctuations in the theory-based potentials (permanent components) indicates that the standard practice of defining cyclical those fluctuations with 2-8 years periodicity is bound to produce inferential distortions. Filters focusing on these fluctuations will overestimate gaps variance at business cycle frequencies and underestimating gaps variance at low frequencies, altering the sequence and the number of turning points, and the properties of expansions and recessions. It also suggests that, when disentangling gaps from potentials (transitory from permanent fluctuations), econometrically fancier procedures will not perform better than crude ones, because they are not designed to capture the theoretical features of my DGP.

Given the distortions that most procedures display, one may be tempted to go structural, estimate the model assumed to have generated the data, and with the parameter estimates construct model-based estimates of the latent components, much in the spirit of Justiniano et al. [2013] or Furlanetto, Gelain, and Taheri-Sanjani [2020]. If model misspecification is a concern, the composite posterior approach of Canova and Matthes [2018] could be used to render latent variables measurement more robust. While structural models are popular in academics, policy institutions still use a variety of statistical methods to extract gaps (transitory components), benchmarking the outcomes with an estimated model, see e.g. Croitorov, Hristov, McMorrow, Pfeiffer, Roeger, and Vandermuellen [2019].

For this reason, I design a filter displaying some optimality properties, given the features of the gaps I emphasize. I show that the filter can be rigged to produce gaps with important low frequency variations and potentials with interesting business cycle frequencies variations; can generate time paths that are close to those of the DGP; and can improve on the best procedures in terms of the Monte Carlo statistics I construct. While the gains are more moderate when transitory components are of interest, the filter is competitive with the best existing approaches even in this case.

The rest of the paper is organized as follows. The next section provides a refresher of the terminology used in the paper and definitions needed to understand what will come next; section 3 discusses the design of the experiment; section 4 lists the filtering procedures; section 5 presents the statistics; and section 6 summarizes of outcomes. Section 7 interprets the results and designs a filter which can capture relevant features of the data. Section 8 concludes.

2 A Refresher of the terminology and some definitions

A zero mean, stationary time series X_t can be equivalently characterized with the autocovariance function $\gamma(\tau) = E(X_t X'_{t-\tau})$ or with the spectral density $S(\omega) = \frac{1}{2\pi}\sum_{\tau=-\infty}^{\infty} e^{-i\omega\tau}\gamma(\tau), \omega = [-\pi, \pi]$, see e.g. Canova [2007]. While both functions are symmetric (around $\tau = 0$ or $\omega = 0$), the elements of the former are correlated while those of the latter are not. When the spectral density is evaluated at Fourier frequencies $\omega_j = \frac{2\pi j}{T}$, it is possible to associate spectral frequencies with periods of oscillations $p = \frac{2\pi}{\omega_j}$, and thus split the variance of the process (the area under the spectral density, $\sum_{\omega_j} S(\omega_j)$) into orthogonal regions comprising cycles with different periodicity.

In time series it is typical to associate low frequencies variations, i.e. fluctuations with long period of oscillations, with trends; medium frequency variations, i.e. fluctuations with medium period of oscillations, with business cycles; and high frequency variations, i.e. fluctuations with short period of oscillations, with irregular cycles. To analyze "trends" or the "business cycles' in isolation, one typically employs filters. Ideally, a filter eliminates fluctuations at "uninteresting" frequencies while leaving unchanged the fluctuations at "interesting" frequencies. Low pass filters, i.e. filters with the frequency representation $F(\omega_j) = 1$, if $\omega_j < \bar{\omega}_1$ and zero otherwise, and band pass filters, i.e. filters with the frequency representation $F(\omega_j) = 1$, if $\bar{\omega}_1 < \omega_j < \bar{\omega}_2$ and zero otherwise, have these features. However, with a finite amount of data, both types of filters generate distortions due leakages over other frequencies and compressions at the required frequencies - think of a badly tuned receiver: the signal from your favorite station will be weaker than otherwise; and signals from other stations will make the reception noisy.

Apart from low pass and band pass filters, there are a variety of statistical approaches one can use to extract trends or business cycles fluctuations. Generally speaking, the procedures fall into two classes: moving average and regression methods. Because all approaches will imperfectly extract the variance of the spectrum at the required frequencies, even in large samples, they should be considered approximations to ideal filters.

The class of moving average filters is of interest among applied investigators because, if the weights sum to one, they will detrend a series, in the sense that they produce a stationary output from a non-stationary input. Because not all filters can be represented via moving averages and because the weights do not necessarily sum one, detrending and filtering are distinct operations. Furthermore, because different approaches produce approximation errors with different properties, one should expect filtered data to display different time series profiles and different moments (see e.g. Canova [1998], Canova [1999]).

When X displays a unit root, $S(\omega_j = 0)$ goes to infinity, and the variance of X also goes to infinity. However, away from the zero frequency, one can still examine the spectral properties of non-stationary data. In particular, one can still analyze long cycles (say, those with periodicity 32 to, say, 64 quarters), provided that X_t is long enough to contain reliable information about them. While there are ways to analyze the very low frequency components of non-stationary series (for example, using the local spectrum or the growth rate of the data), when the DGP is non-stationary I will present results simply omitting the zero and neighboring frequencies (from 64 quarters to infinity).

The squared gain function provides a useful tool to understand the time series behavior of the filtered series and to interpret the distortions different filters generate. The function tells us, frequency by frequency, the proportion of the variance of the filtered series to the variance of the original series. For example, a unitary squared gain at ω_j means that a method has left untouched the variability of the original series, while a zero squared gain means that it has wiped out all the variability at ω_j. Squared gain values between zero and one, on the other hand, indicate the extent of the attenuation of the variability at frequency ω_j and values in excess of one provide evidence that the filter has amplified the variance of the original series at ω_j.

3 The design of the experiment

As DGP for the experiments, I use the standard closed economy new-Keynesian model popularized by Smets and Wouters [2007] with real frictions (habit in consumption, investment adjustment costs), nominal frictions (price and wage stickiness and indexation), and a Taylor rule for interest rate determination. In a new-Keynesian setup, a set of equations characterizing the optimality conditions for the potential economy - the economy without nominal frictions, markups and monetary disturbances - is added to the optimality conditions of the original problem (see appendix) and the solution for level and potential variables is jointly found. Gaps are obtained as the difference between the level and the potential for each endogenous variable. One can assume that all the exogenous disturbances are transitory or that some of them are permanent and others are transitory. In the latter case, apart from a potential-gap decomposition, one can separate the permanent and the transitory component of the endogenous variables, the former being the portion of the data driven by permanent disturbances, and the latter the portion of the data driven by transitory disturbances. As discussed below, when some of the disturbances are permanent, potentials and gaps will both display permanent and transitory features. Thus, gaps may not be interesting economic objects (gaps may "never close"), but the transitory components of the data are.

The model economy features seven structural disturbances (to TFP, to investment, to government expenditure, to the Taylor rule, to the price and wage markups, and to the risk premium). In the baseline specification, all disturbances are stationary. In this case, potentials and gaps are stationary, there is no long run trend (permanent path) to speak of. Thus, one may wonder what is the purpose of trend-cycle decompositions in this situation. When disturbances are stationary but persistent, important low frequency variations appears in the data. Thus, to focus attention on fluctuations with interesting periodicity, a researcher may want to purge the data of low frequency variations to highlight the object of interest. Hence, even in this case, the exercises I run are relevant.

Because in the baseline specification disturbances are persistent, gaps and potentials will both feature low and business cycle fluctuations, and their spectral power at low and business cycle frequencies will generally be similar. Furthermore, because TFP, investment and government expenditure disturbances affect both the gaps and potentials, the two components will be correlated.

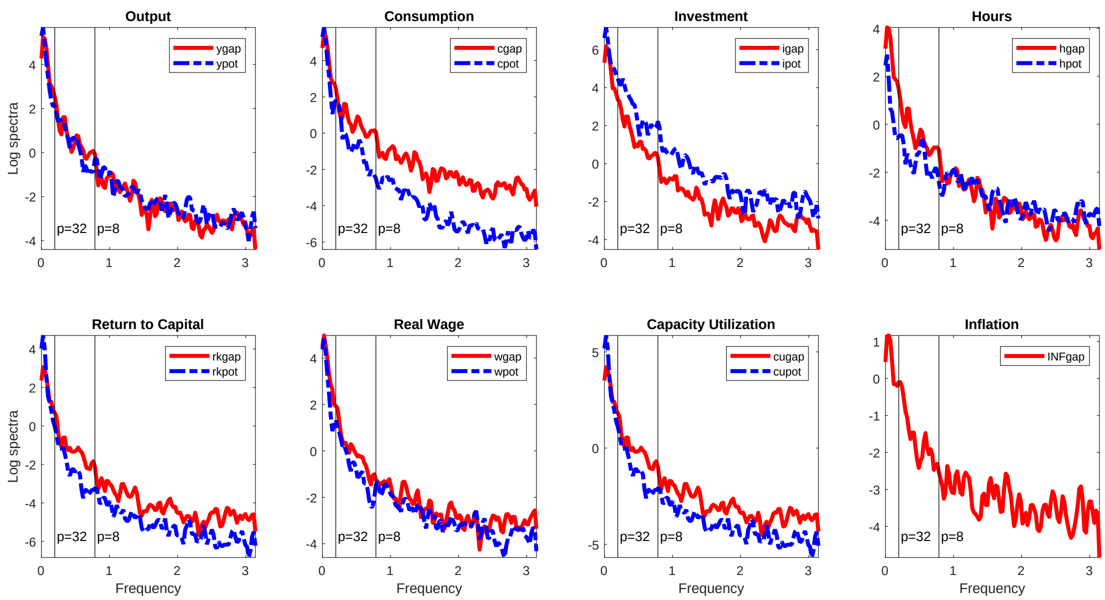

Figure 1: Log spectral densities of gaps and potentials: Basic SW.

To illustrate these facts, I plot in figure 1, the log spectral density of the gaps and the potentials for eight variables of the model, using one realization of the shocks, when T=750. Note that the low frequency variability of the gaps (32-80 quarters) is generally larger than their business cycle variability (8-32 quarters); that gaps and potentials almost equally account for the variance of the observables at low and business cycle frequencies; and that the correlation between gaps and potentials is, on average across variables, equal to 0.86.

To some reader, the baseline specification may appear to be unrealistic: common wisdom suggests that certain real variables should display an upward trend. To insure that this is the case, I alternatively consider a setup which allows TFP disturbances to feature a unit root. In this case, the levels of output (Y), consumption (C), investment (I) and real wages (W) will be on a balanced growth path. Because TFP drives both the gaps and the potentials of these four variables, both latent components will display an unbounded peak at frequency zero (produced by the unit root) and will respond to transitory disturbances. Thus, as in the baseline specification, the gaps for these four variables will still feature significant low frequency variability; gaps and potentials will display similar spectral shape, and will be correlated. For the variables not affected by the unit root (hours (H), return to capital (RK), capacity utilization (CapU), inflation π), the baseline scenario applies.

It is also worth examining what happens to the spectral shape of the latent variables when, as in Aguiar and Gopinath [2007], "the trend is the cycle", i.e. the permanent TFP disturbance accounts for a large portion of the data variance at business cycle frequencies. In principle, none of the features I discussed should be altered, but the relative variability of the gaps at low and business cycle frequencies may be affected. To generate this setup, I decrease by 3/4 the persistence of the transitory disturbances.

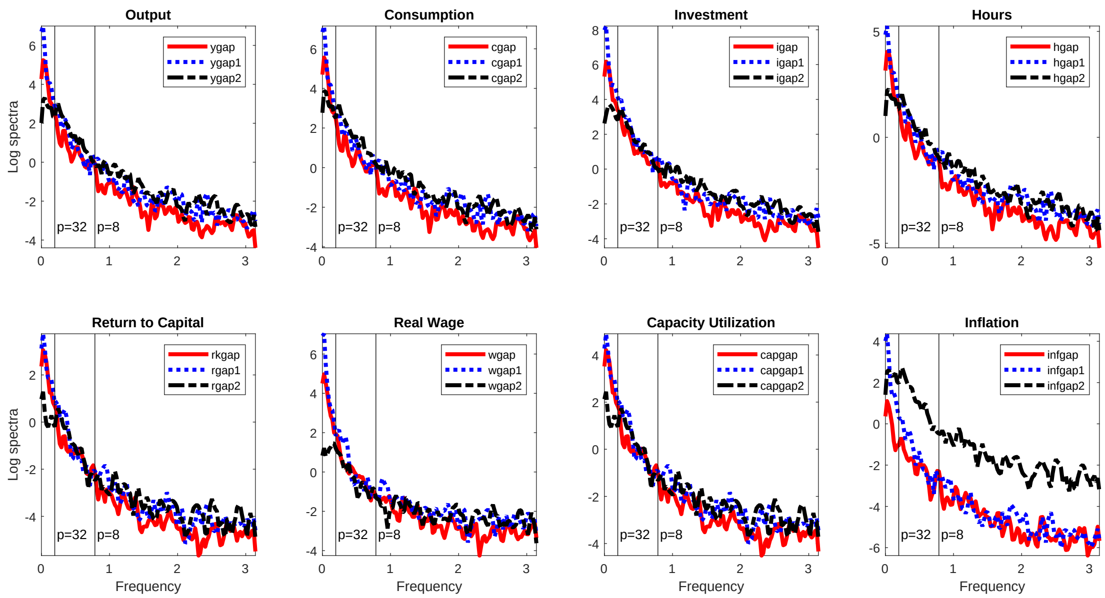

Figure 2: Spectral properties of gaps: Basic SW, SW with unit roots, SW 'trend is the cycle'.

Figure 2 which plots the gaps in the baseline, in the unit root and in the "trend is the cycle" setups confirms that the qualitative features of the log spectra of the gaps I emphasize remain unchanged. In particular, even in the last scenario, gaps will still display important low frequency variability.

Although the DGP I employ is popular, it does not account for financial frictions, nor it takes into account the relationship between the real and the financial side of the economy. Given the interest that the macro-financial links have generated since the financial crisis, one may be curious as to whether the presence of finance considerations and financial frictions alter the spectral properties of the gaps. In the models with financial frictions, I define the potentials in the same way I have done without them: they characterize the economy without nominal rigidities and with the markups, risk premium and inflation target disturbances set to zero. Thus, financial frictions (and risk shocks if present) affect the potential economy and the gaps.

Figure 3 compares the spectral properties of the gaps produced in the baseline specification (SW); in the baseline specification with financial frictions (SWFF), see Del Negro, Giannoni, and Schorfheide [2015]; and in the baseline specification with risky contracts (CMR), see Christiano, Motto, and Rostagno [2014]. The features of the gaps I emphasized are independent of the existence of financial frictions, except for the return to capital, which displays much larger business cycle and high frequency variations when financial frictions matter. If anything, the relative importance of low frequency components in the gaps grows larger. Quantitatively, the CMR model displays lower log

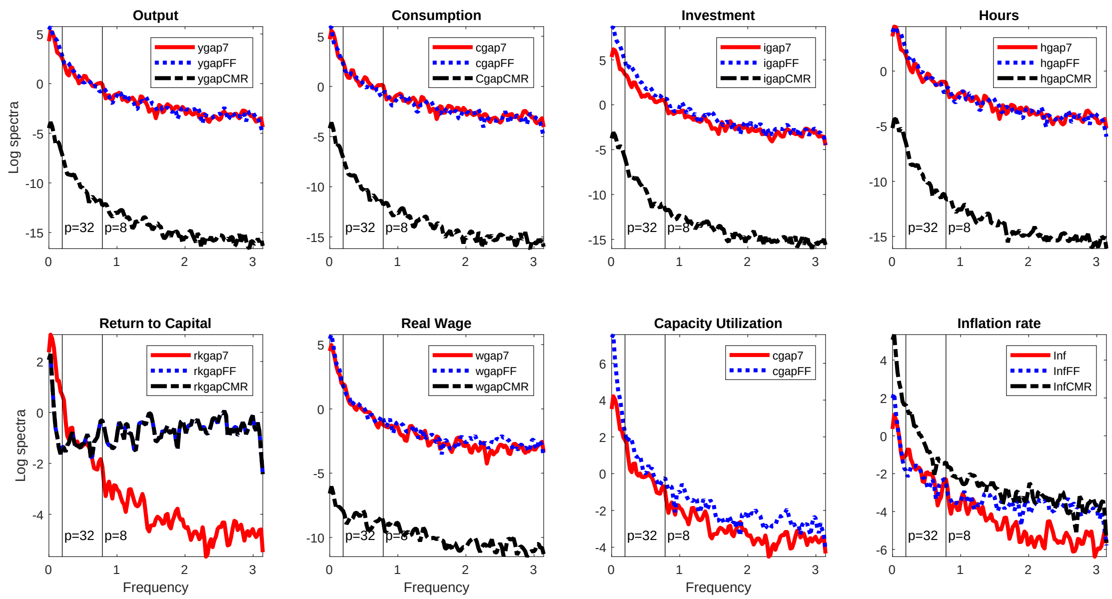

Figure 3: Spectral properties of gaps: Basic SW, SWFF and CMR.

spectral densities because the disturbances have smaller variability. Although not reported to save space, potentials either have unchanged spectral features (SWFF), or display magnified variance in very low frequency portion of the spectrum (CMR).

It is also useful to point out that policy models built according to different principles also produce gaps with the features I discuss. As an illustration, figure 4 plots the log spectra of the output gap and of the output potential produced by one simulation of ECB-Base, the Euro are version of US-FRB model. Clearly, there is an important portion of the gap variance in the low frequencies: over 40 percent of the total variance of the process is in this area; and it contributes to about 80 percent of the total output variance at those frequencies. Two features in this kind of models are however different: the potential (which is generated with a production function approach) is basically a unit root process and it contributes much less to the output variance away from the zero frequency; output gap and potential output are driven by independent disturbances.

One final exercise may provide useful information on the generality of the experiment. In the baseline specification, TFP, investment and government spending disturbances drive both potentials and the gaps. Given that these shocks are highly persistent, the conclusions that potentials and gaps have similar spectral features follows as a corollary. What happens if investment and government spending disturbances are absent? Would that make gaps and potentials different, given that markup, risk premium, and monetary policy disturbances affect only gaps? Figure 5 shows that this is not the case and none of the features I discuss is altered by this change [1].

In sum, in the class of models I consider, the persistence of the disturbances (rather than their

[1] I have also generated gaps and potentials in the model of Beaudry et al. [2018], where there is a TFP and preference shock (broadly interpreted as demand shock and interchangeable with a monetary policy shock). With an appropriately parameterization, the model is able to generate a peak the spectral density at low frequencies of the hours. However, because the endogenous mechanism generating this peak affects both gaps and potentials, gaps will display important low frequency components, gaps and potentials will have similar spectral shapes, and will be correlated.

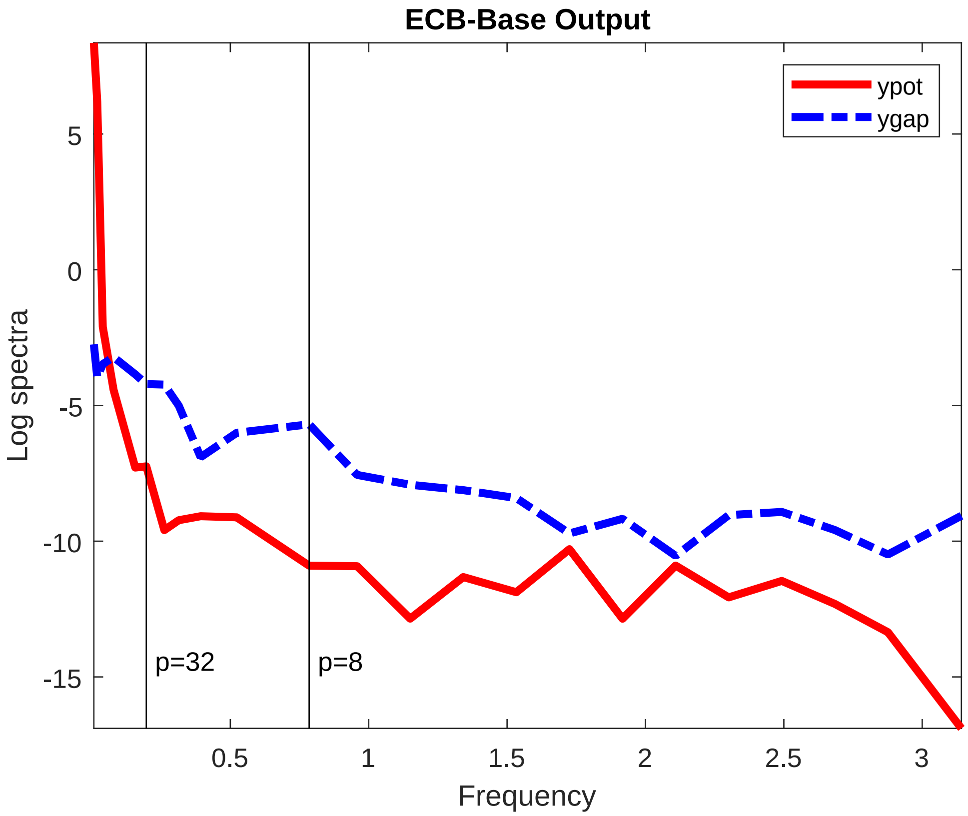

Figure 4: Spectral properties of output gap and potential: ECB-Base.

nature) determines the low frequencies properties of gaps. Any model featuring persistent TFP disturbances and the definition of potential I employ will make gaps and potentials correlated, induce similar spectral distribution of variance, and produce large low frequency variability in the gaps.

What are the spectral features of the transitory components the SW model generates? Figure 6 shows that when transitory shocks have standard persistence the distribution of the variance by frequencies of gaps and the transitory components is qualitatively similar (compare ”gaps” with ”tra1”). However, while gaps and potentials are correlated, transitory and permanent components are uncorrelated. Figure 6 also shows that the relative importance of transitory low frequency fluctuations decreases in ”the trend is the cycle” scenario (compare ”tra1” and ”tra2”). Thus, while this scenario comes closer to produce what certain filters assume, one should also be aware that the importance of the permanent component at business cycle frequencies is magnified relative to the baseline setup. Thus, the association between business cycle frequencies and transitory fluctuations will be poor.

I summarize the properties of output gaps (transitory components) for the DGPs I examined in table 1. All in all, there seems to be little loss of generality in taking the basic SW model as the DGP for the experiments. Furthermore, the setup with stationary but persistent disturbances provides a useful benchmark to estimate the distortions produced by standard filters for gap extraction. The scenario with a unit root in TFP, on the other hand, seems a reasonable alternative to examine the

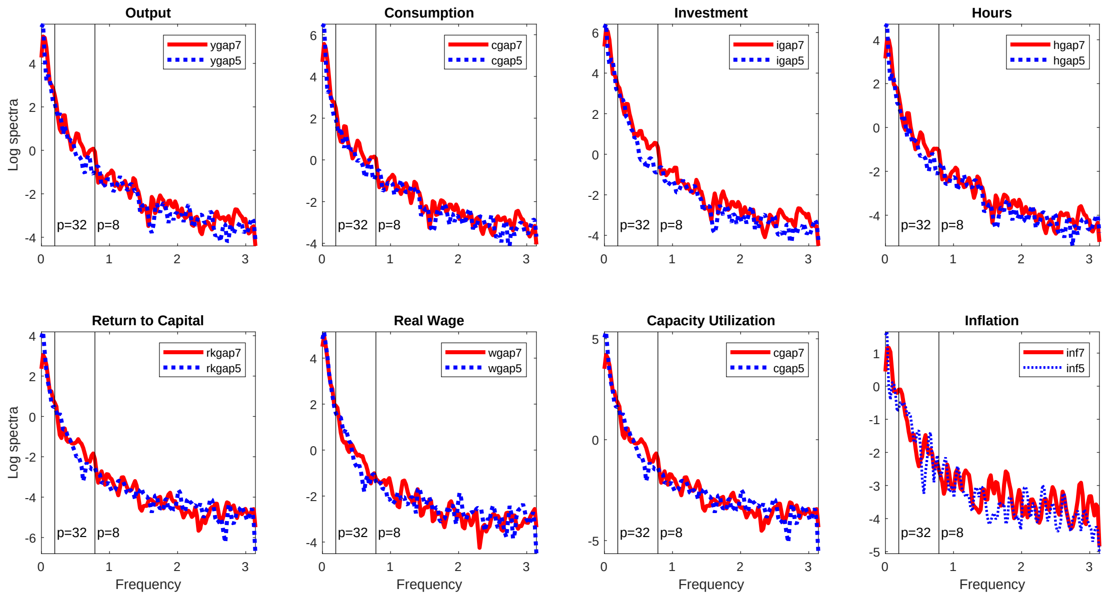

Figure 5: Spectral properties of gaps: Basic SW with 7 or 5 disturbances.

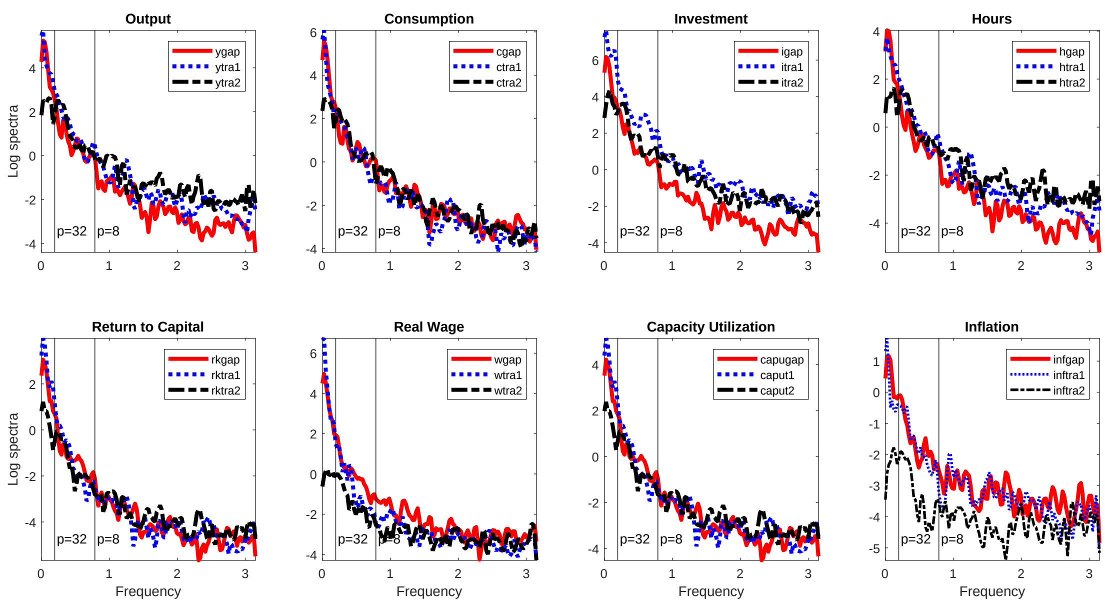

Figure 6: Spectral properties of gaps: stationary and non-stationary SW.

performance of filters when extracting transitory fluctuations [2].

[2]In policy circles, the idea that permanent disturbances can be represented with a random walk may be considered unreasonable, given that one often hears the view that "drivers in the long run are smooth". One could clearly use ARIMA processes for the disturbances, but this will not affect any of the features this section emphasize. The results obtained with ECB-Base model, which uses a production function approach to generate potentials confirms this. Also

Table 1: Relative variances

	All frequencies	Low frequencies	BC frequencies	Own variance low frequencies	Own variance BC frequencies
Gap (SW stationary)	0.58	0.66	0.50	0.16	0.07
Gap (SW 5 shocks)	0.64	0.65	0.91	0.10	0.04
Gap (SW-FF)	1.06	0.91	1.01	0.22	0.05
Gap (CMR)	< 0.01	< 0.01	< 0.01	0.15	0.02
Gap (ECB-Base)	0.01	0.35	0.27	0.33	0.40
Transitory (SW unitoot)	0.01	0.80	0.64	0.18	0.10
Transitory (SW trendiscycle)	< 0.01	0.61	0.36	0.22	0.44

Notes: SW is the standard model, SW-FF is the standard model with financial frictions; SW 5 shocks is the standard model without investment and government spending shocks. CMR is the model of Christiano et al. [2014]. Base is the Euro area version of the US-FRB model. The first three columns present the faction of the variance of the observable output due to the gap, over all, low (LOW) or business cycle (BC) frequencies. The last two columns report the fraction of the variance of the output gaps (transitory output components) at low (32-64 quarters) and at business cycles (8-32 quarters) frequencies. Sample T=750. Numbers may exceed 1 because latent components are correlated.

4 The filtering procedures

Given time series for the endogenous variables, I will be interested in analyzing which filter most accurately characterizes the gaps (or the transitory components) the model generates. There are numerous procedures a researcher can employ to extract two latent components from one observable variable. I focus on the most commonly used in the macroeconometric literature. I do not consider production function based procedures, because estimates of the long run values of inputs need demographics, participation rates, and other slow moving variables that can not be produced within my framework of analysis. The procedures I consider differ in many dimensions. Some are statistical and others have economic justification; some are univariate and other multivariate; some require parameter estimation and others do not. From my point of view, the two most important differences are the assumed properties of the trend; and the correlation between unobservables.

The first approach is the oldest and maintains that the trend is deterministic and uncorrelated with the cycle. Thus, the latter can obtained as the residual of a regression of the variable on a polynomial trend. I use a quadratic polynomial and run the regressions variable by variable, meaning that I do not exploit the fact that the trend will have common features in my dataset. The results obtained with this apporach are denoted by the acronym POLY in the text and the tables.

The second approach is the Hodrick and Prescott filter. Here the trend is assumed to be stochastic but smooth and uncorrelated with the cycle. The latter is the difference between the level of the series and the Hodrick and Prescott trend, which is obtained via the ridge estimator:

$$\tilde{y} = (H'H + \lambda Q'Q)^{-1}H'y \tag{1}$$

with ARIMA disturbances, potentials and gaps will still be correlated, will display similar frequency distribution of variance, and gaps will feature important low frequency variations. The use of ARIMA disturbances, on the other hand, complicates the extraction of transitory components because an important portion of the variability of the data at all frequencies is due to the permanent component.

where λ is a smoothing parameter, $y = (y_1, \dots y_t)$ the individual series, $\tilde{y} = (\tilde{y}_1, \dots, \tilde{y}_t, \tilde{y}_{t+1}, \tilde{y}_{t+2})$, the trend, $H = (I_{t\times t}, 0_{t\times 2})$ and $Q_{t\times(t+2)} = \begin{bmatrix} 1 & -2 & 1 & 0 & \dots & \dots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \dots & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & \dots & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \dots & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & \dots & \dots & 1 & -2 & 1 \end{bmatrix}$. I set $\lambda = 1600$ and denote the results with the acronym HP. In the sensitivity analysis, I consider $\lambda = 51200$ (acronym: HPa), a value close to the BIS recommendations, see Borio [2012]. With this value of λ, the trend has only minor variations and the cycle accounts for the majority of the fluctuations in the data.

The third approach assumes that the trend is stochastic, displays at least one unit root, and it is uncorrelated with the cycle. I consider two separate sub-cases: one where the cycle is obtained by short differencing (one quarter); and one where it is obtained by a long differencing (24 quarters) the observable variable. I denote the results in these two sub-cases with the acronyms FOD and LD. In sensitivity analysis, I also consider 4 and 16 quarters differencing operators (acronyms: FODa, LDa)

The fourth approach permits the trend to be stochastic and to display both stationary or non-stationary features, and assumes that its variability is entirely located in the low frequencies of the spectrum. To extract cycles with 8-32 quarters periodicity, I use the band pass filter implementation of Christiano and Fitzgerald [2003], which employs asymmetric and non-stationary (time-dependent) weights. I denote the results with the acronym BP. For sensitivity, I also examine the trigonometric version of the filter see Corbae, Ouliaris, and Phillips [2002] (acronym:Trigo), the stationary, truncated symmetric version suggested by Baxter and King [1999] (acronym: BK), and a version of the baseline filter, which extracts cycles with periodicities of 8-64 quarters (acronym: BPa)

As alternative, I have also considered the wavelet filtering approach, recently suggested by Lubik et al. [2019]. Wavelet are one-sided MA filters, where the length of the MA polynomial depends on the cycles being extracted. For example, to exact cycles with a 8-32 quarters periodicity, a MA(16) is used but to extract cycles with a 32-64 quarters periodicity, a MA(32) is employed. Wavelets have some intuitive advantages over band pass filters as they work in time domain and the number of MA terms is finite. I denote the results obtained with this approach with the acronym Wa.

The fifth approach is based on local projections and follows Hamilton [2018]. Here the trend is defined as the medium term predictable component of a variable and it is obtained by running a regression of each variable at $t+m$ on current and up to d lags of the variable. The cyclical component is assumed to be uncorrelated with the trend and obtained as the residual of the regression. I set m=8, d=4, and report results with this approach with the acronym Ham. In the sensitivity analysis, I consider the alternative of m=12 and d=2 (acronym: Hama)

The sixth approach is based on a state space formulation of the latent variable problem. It assumes that the trend is a random walk with drift; that the cycle is an AR(2) process, and allows the innovations in the trend and cycle to be correlated. No measurement error is included in the measurement equation. Using a flat prior on the parameters, I compute posterior distributions using a MCMC approach, as in Grant and Chan [2017], using 30000 burn-in draws and saving 5000 draws. The reported properties are computed averaging the resulting trends and cycles estimates over retained draws. I denote the results with the acronym UC. In the sensitivity analysis I also consider a bivariate UC filter with output and capacity utilizations as observables (acronym UCbiv).

The approaches so far described use univariate information to separate the two latent components. Given the general equilibrium nature of the DGP, these approaches are inefficient as they disregard, for example, the presence of balance growth, when there is a unit root in TFP, or the fact that cyclical components have similar features (since they are driven by the same disturbances). The next two procedures account for the possibility that commonalities may be present. The first, based on Beveridge and Nelson [1981]' decomposition, defines the trend as the predictable long run component of a vector of variables. The cycle is the difference between the vector of variables and the estimated trend. Here trend and cycles are driven by the same shocks (and thus perfectly correlated), which are the reduced form innovations of a vector autoregression on lags of the relevant variables. I run the decomposition unrestricted, that is, without the signal-to-noise prior restriction of Kamber, Morley, and Wong [2018], because the DGP is already a very low order VAR(p) polynomial.

The second procedure follows Blanchard and Quah [1989] and still uses a vector autoregression to compute the shocks to the variables. However, rather than the vector of innovations, it uses identified disturbances to separate the two latent components. Here, the trend is driven only by supply disturbances and the cycle by both supply and demand disturbances. In the implementation I use, the vector autoregression includes output growth and hours for both approaches. I denote the results with the acronyms BN and BQ. In the sensitivity analysis, I also consider a trivariate VARs with output growth, consumption to output, and investment to output ratios (acronyms: BNa, BQa). While there are other multivariate approaches one considered, the list of procedures I employ is sufficiently exhaustive and covers well what is available in the literature, thus making the comparison exercise meaningful, and the results informative.

4.1 What should we expect?

Given that different methods use different assumptions to identify the latent components, one should expect them to produce different outcomes, see Canova [1998]. Furthermore, since no procedure takes into account the features of the DGP I discussed in the previous section, biases might be expected.

To be specific, procedures which assume no correlation between the two latent components, should be relatively poor when extracting gaps but better endowed when extracting transitory fluctuations. A deterministic polynomial approach is likely to overestimate the volatility of both the gaps and of transitory fluctuations, given that both potentials and permanent components are stochastic; while methods which impose a unit root should be better suited to separate permanent and transitory components than potentials and gaps. Finally, methods designed to focus attention on cycles with particular periodicity, will distort the frequencies distribution of the variance of gaps and of transitory fluctuations. In fact, they will attribute the low frequency variations belonging to both components to the trend, and the cyclical (and high frequency) fluctuations belonging to both components to the cycle. Because model-based gaps and transitory components display variability at all frequencies, while such approaches carve the spectrum by frequencies, biases will be significant, see Hansen and Sargent [1979] for an earlier statement of this problem, and Canova [2014] for a recent one.

The presence of unit roots is unlikely to affect much conclusions obtained without and distortions will be generally present also in this case. Misspecification of the properties of the DGP is the reason for their existence. For example, since a number of procedures assume that the permanent components are random walks, while the growth of permanent components the model generates are highly serially correlated, estimates obtained with these methods will distort the properties of the transitory components. Perhaps more importantly, many approaches will twist the frequency distribution of transitory variability. Thus, turning point dating, measures of durations and amplitudes of expansions and recessions and cross-variable relationships will be generally distorted.

Hence, no procedure is close to the ideal given the DGP I consider. Moreover, because different procedures employ different assumptions to identify the latent variables, they will produce different biases. The question of interest is which one is least damaging and why. Note that because the distortions I discuss hold in population, small samples can add to the problems, especially for those procedures which require parameter estimation. Finally, while tinkering with the filters may lead to some quantitative improvements, it is unlikely that any modification will dramatically change the fact that no procedure is close to the ideal.

5 The statistics

To measure the performance of different approaches, I consider a number of standard statistics. These are computed averaging results using 100 data replications, to wash out simulation uncertainty. First, I compute the mean square error (MSE), calculated as the difference between the true gaps (the true transitory components) and the filtered series.

Second, I report the contemporaneous correlation between the true gaps (true transitory components) and the filtered series, the first order autocorrelation and the variability of the filtered series, benchmarking them with those of the true gaps (the true transitory components). I compute these four statistics for 9 series that the model generates: output, consumption, investment, return to capital, hours, real wages, capacity utilization, inflation and nominal rate. Note that the latter two series have no potential. I also extract a factor from the actual data, apply the filtering procedures, and compare MSE, variability, auto and contemporaneous correlations of the filtered series to the those of the factor computed using the true gaps (true transitory components).

Third, I compute turning points in the filtered data and compare their number, the average duration and the average amplitude of expansions and recessions with those present in the true gaps (the true transitory components). Because in the baseline case, the model is solved linearly around the steady state, duration and amplitudes are roughly symmetric across business cycle phases in the simulated gaps (transitory data). This may not necessarily be the case in the filtered series

Policymakers are interested in latent variables for two reasons. First, because the want to measure in real time the state of the economy. Second, because they want to use them to predict other variables, for example, inflation via a Phillips curve, or employment (hours) with a Okun law. For this reason, I also compute two additional set of statistics. The first measures the MSE in real time, focusing attention on the last 12 periods of each sample; the second compares the variance of the prediction error in the regressions implied by the true gap (transitory) data and those implied by the filtered data. Letting y^i_{t-j} be either the true output gap (true transitory output) or the filtered one and m=1,4, the predictive regressions take the form:

$$\pi_{t+m} = \alpha_0 + \alpha_1 \pi_t + \sum_{j=1}^{3} \beta_j y^i_{t-j} + e_{t+m} \quad (2)$$

$$h_{t+m} = \alpha_0 + \alpha_1 h_t + \sum_{j=1}^{3} \beta_j y^i_{t-j} + e_{t+m} \quad (3)$$

6 The results

Table 2 reports a summary of the results counting, for each statistics, the number of times a procedure is least distorting across variables. Counting measures assign a one to the best procedure (0.5 if there is a tie) and zero to the others. Totals are computed equally weighting all statistics. Tables 4-10 in the appendix give the details: for each variable (factor) and for each procedure, table 4 reports the average MSEs across replications; table 5 the average real time MSEs; table 6 the average contemporaneous correlation, table 7 the average AR1 coefficient; and table 8 the average variability; table 9 the average number of turning points, the average durations and amplitudes of recessions and expansions for output and the factor; and table 10 the average variance difference in the prediction error of the Phillips curve and Okun law regressions between each procedure and the true one.

Table 2: Summary results, T=750

Statistic	POLY	HP	FOD	LD	BP	Wa	Ham	UC	BN	BQ	BW
	Gaps										
MSE	5	3						1	0.5	0.5	8
Corr	9								0.5	0.5	8
AR1	4			3		3					6
Var	4			2		3	1				
TP	1.5	5	2	1.5							3
RT-MSE		1				3	2	3	0.5	0.5	8
PC	2										2
OL				1			1				
Total	25.5	9	2	8.5	0	9	4	4	1.5	1.5	35
	Transitory										
MSE			9					1			
Corr											
AR1	4					5.5		0.5			1
Var	3			6			1				1
TP	4	4		2							3
RT-MSE			4					6			
PC				1						1	
OL				2							
Total	11	4	13	11	0	5.5	1	7.5	0	1	5

Notes: POLY is polynomial detrending, HP is Hodrick and Prescott filtering, FOD is first order differencing, LD is 24 quarter differencing, UC is an unobservable component model, BP is band pass filtering, Wa a wavelet filter, Ham is local projection detrending, BN and BQ are bivariate Beveridge and Nelson and Blanchard and Quah decompositions, BW the Butterworth filter. MSE is the mean square error, Corr the contemporaneous correlation with the true series, AR1 the first autoregressive coefficient, Var the variability of the series, TP the number of turning points, the duration and the amplitude of expansions and recessions, RT-MSE is the real time MSE, PC is the Philipcs curve prediction, OL the Okun law prediction. In each row the ranking is over 9 series and one factor, except for TP where the ranking is for output and the factor. Numbers are computed summing the top ranks, equally weighting all variables; ties each get value of 0.5.

Each table has two parts. In the upper panel, I report the performance of different filters for gap extraction; in the second their ability to recover the transitory component of the data.

The Polynomial approach is, by far, the least distorting among the competitors when measuring gaps. The approach is superior as far MSE, contemporaneous correlation, AR1, variability, and Phillips curve regressions are concerned, while it lags behind in terms of real time MSE, turning points characterization, and Okun law regressions. The next approaches in the ranks are appropriate for some statistics only (the HP filter for TP detection; the Wavelet filter for TP detection, AR1 coefficient and variability) but seem to be much less suited than the Polynomial filter to capture features of the gap data. Three additional aspects of the top panel of table 2 are worth emphasizing. The long difference filter is superior to the Hamilton filter, suggesting that the horizon of the local projection is probably too short: Hamilton [2018] shows that the filter is close to a eight-period difference filter. The commonly used UC approach, on the other hand, is competitive only in terms of real time MSE; for other statistics, it is never among the top procedures. Finally, the performance of the bivariate BQ and BQ procedures is poor: they top other approaches for hours in terms of MSE, real time MSE, and contemporaneous correlation, but never rank first for output gaps.

Quantitatively speaking, the distortions in terms of time paths are large and biases relative to theory-based measures of gaps important. For example, the average MSEs are all larger than the average MSE produced by a random walk (which, e.g., for the output gap is 14.45), and the real time MSE are almost twice as large as the real time MSE produced by a forecast that uses $T-12$ value for all successive 12 periods. The magnitude of the distortions is also large in terms of volatility and for correlation measures; in the latter case, except for the Polynomial procedure. At the opposite extreme, biases in the persistence parameter are relatively small and, except for the FOD filter, all estimates are in close range of the true AR1 coefficient.

When it comes to characterizing the properties of transitory components, the FOD filter is the least distorting, closely followed by the Polynomial and LD filters. However, the superiority of the FOD filter is entirely concentrated in MSE measures, while the Polynomial and LD approaches come on top for a number of statistics. In general, the distribution of winners is more uniformly spread across procedures than with gap measures. Relative to gap extraction, one can notice a deterioration of the performance of HP and Hamilton filters and an improvement in the performance of the UC filter, which entirely comes from the real time MSEs, and involves output, consumption and investment. That is to say, the filter seems appropriate as far as predicting in real time the transitory component of output, consumption and investment up to 12 quarters ahead, but gives a poor characterization of their historical properties. BN and BQ performance is poor, despite the fact that they use more information to extract the latent transitory variables. While the inferiority of BN and BQ could have been expected in the case of gaps because they overdifference output data (the output series is stationary), it is less obvious why this is the case when extracting transitory components. Given the popularity of BQ procedures in the literature, the next section studies why this is the case. Finally, notice that the band pass filter is never the best for any statistics or series for extracting both gaps and transitory components and that the Wavelet filter does a reasonable job in comparison. Because both filters aim at capturing portions of the spectrum, the differential performance must due to the smaller approximation error of wavelet filter possesses.

The magnitude of the distortions when extracting transitory components is generally larger than when extracting gaps. As we have seen, the presence of a unit root changes the relative distribution of variances of the permanent and of the transitory components in the low frequencies (the former gets more weight) and this affects the performance of a number of approaches. Interestingly, almost all procedures have a hard time to capture the ups and downs of the transitory components, see figure 7 for examples. Thus, MSE errors tend to be large and the contemporaneous correlation between the true and filtered series is essentially zero for several filters.

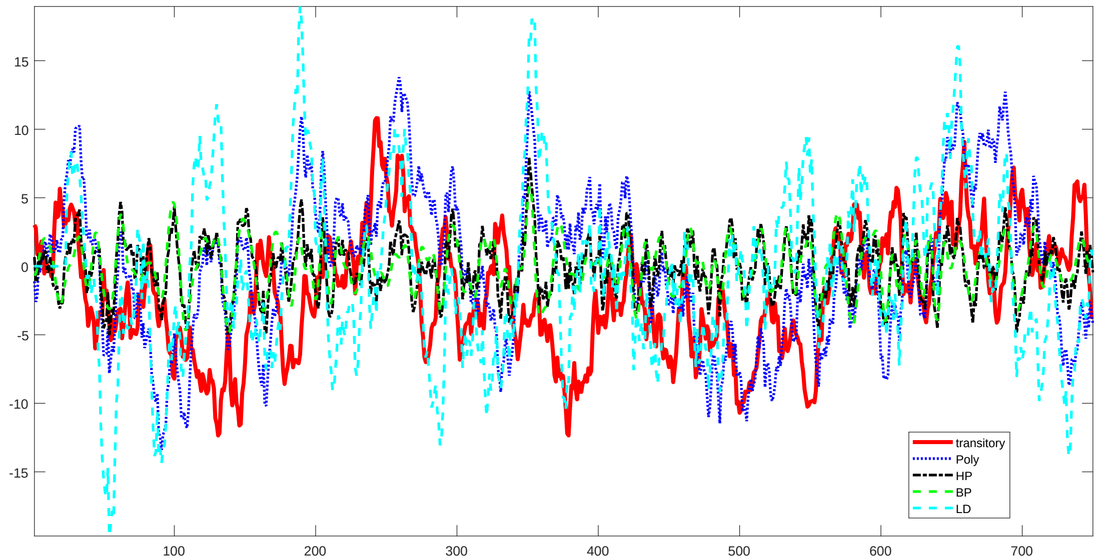

Figure 7: Time series: true and estimated transitory components. Various filters.

Many approaches have the tendency to underestimate the variability of the both the gaps and the transitory components. The exceptions are the Polynomial, the BN and BQ procedures. For the polynomial filter, this is expected given the nature of the squared gain function I discuss in section 7. Despite this obvious bias, it is remarkable that the procedure is still on the top for gaps and the second to the top for transitory components. The excess variability produced by BN and BQ procedures is less expected and, as shown in the next section, due to the fact that the approaches considerably distort the frequency distribution of the variability of the latent series and erroneously attribute low frequency variations present in the gap (transitory) data to the trend component.

To summarize, and somewhat unexpectedly, the Polynomial approach, the crudest and the oldest method existing in the literature, turns out to be the least damaging in characterizing the properties of both the gaps and the transitory fluctuations. A similarly crude first order (long) differencing filter ranks well when it comes to extract the transitory components of the data. Furthermore, using larger information sets or fancier econometrics does not seem to help.

Table 2 considers 9 series generated by the model and the factor characterizing their common dynamics. Because policymakers are mostly concerned with output gaps and output dynamics, is it worth zooming in the table and look at how different procedures characterize the latent output quantities. The HP filter is uniformly superior when measuring the number of turning points in the gap data, the duration and the amplitude of expansions and recessions. The result is striking, given the criticisms raised in the literature against the HP filter; see Hamilton [2018] for the latest installment. Next in the list are the Polynomial and the UC approaches. The next section shows why they do well for output gaps and, for the UC filter, why is not necessarily so for other series.

For transitory output, the Polynomial approach ranks first, and HP and FOD come close seconds. The FOD approach tops other procedures in terms of MSE measures, while the Polynomial and the HP approches are superior when measuring the number of turning points and the durations and amplitudes of expansions and recessions. Given that the Polynomial approach has also been discredited for leaving near non-stationary dynamics in filtered output - and referees often harp about filtering data with such a procedure - these results are surprising and deserve further discussion.

6.1 Sensitivity

I have repeated the Monte Carlo exercise varying the parameters of different procedures, tailoring certain filters to better capture the low frequency component of the data, or adding information to multivariate approaches. The performance of HP and BP filters can be generally improved by choosing a higher λ or lowering the limit of the frequency band (see tables A.1-A.7 in the on-line appendix) but, by and large, the conclusions I have presented hold.

In particular, it is still the case that the Polynomial procedure is least distorting when measuring gaps and none of the refinements is able to produce MSEs which are uniformly smaller, both for the gaps and the transitory components. The alternative HP and Hamilton filters produce estimated cycles whose frequency is more in line with those of gaps (transitory data) but this does not change the relative position of the filters in the ranks since the improvement is larger for those variables or statistics where the filters were the top approaches. The most significant changes are produced with the bivariate UC approach when computing the real time MSE; with the alternative BP filter when measuring the AR1 coefficients in gaps; and with the alternative Hamilton and LD filters when computing variabilies. Quantitatively speaking, the gains are small.

Perhaps more interesting are the results obtained with a sample of 150 observations. Table 2 used a large sample (T=750). In practice, estimation uncertainty matters, and given that some procedures require parameter estimation and others do not, some of the conclusions may be affected. A summary of the results for T=150 is in table 3 (tables A.8-A.14 in the on-line appendix provide the details). Overall, the Polynomial approach is still the best when measuring gaps, followed by HP and LD filters. Relatively speaking, a short sample worsens the performance of the Polynomial filter and improves the one of filters which do not require parameter estimation. Also, while there are statistics specific changes in the ranking, the snapshot of table 2 is broadly maintained.

When measuring the transitory components results are somewhat affected. The FOD filter, which was best in table 2, now looses its superiority and the LD filter becomes the least distorting procedure, and the UC approach lags third. The LD filter dominates when measuring the persistence and volatility of the transitory component, while the FOD filter still maintains its superiority in terms of MSE measures. The UC filter, on the other hand, does relatively well only in terms MSE measures. For the two regressions policymakers care about, the sample size makes little difference, and the LD filter is overall the most appropriate method to construct regressors in the predictions equations.

I have also conducted an additional experiment where the model is simulated using a second order solution. In this case procedures which assume a linear, parametric structure are penalized relative to procedures which non-parametrically split the data. In addition, the magnitude of the distortions is, on the whole, larger. However, the ranking among procedures is only minorly affected.

Table 3: Summary results, T=150

Statistic	POLY	HP	FOD	LD	BP	Wa	Ham	UC	BN	BQ	BW
	Gaps										
MSE	4	4						1	0.5	0.5	7
Corr	7						2		0.5	0.5	10
AR1	3			4.5		2.5					4
Var	5	1		3					0.5	0.5	2
TP	0.5	3.5	2	1		2	1				2
RT-MSE		3				4	2		0.5	0.5	8
PC								2			2
OL				2							
Total	19.5	11.5	2	10.5	0	8.5	6	4	1.5	1.5	33
	Transitory										
MSE			4		1			5			
Corr											
AR1	0.5			7		1.5			0.5	0.5	3
Var	3			7							7
TP	2	2		3	1.5			1.5			
RT-MSE			5		3			2			
PC				2							
OL			1	0		1					2
Total	5.5	2	10	19	5.5	2.5	0	8.5	0.5	0.5	13

Notes: POLY is polynomial detrending, HP is Hodrick and Prescott filtering, FOD is first order differencing, LD is 24 quarter differencing, UC is an unobservable component model, BP is band pass filtering, Wa a wavelet filter, Ham is local projection detrending, BN and BQ are bivariate Beveridge and Nelson and Blanchard and Quah decompositions, BW the Butterworth filter. MSE is the mean square error, Corr the contemporanoeus correlation with the true series,AR1 the first autoregressive coefficient, Var the variability of the series, TP the number of turning points, the duration and the amplitude of expansions and recessions, RT-MSE is the real time MSE, PC is the Philips curve prediction, OL the Okun law prediction. In each row the ranking is over 9 series and one factor, except for TP where the ranking is for output and the factor. Numbers are computed summing the top ranks, equally weighting all variables; ties each get value of 0.5.

7 Gain functions

To understand the pattern of results I obtain, I first examine the estimated squared gain function of each filter. As already mentioned, this function measures the reduction/amplification at each frequency of the variance of the observable series due to the filter. Since the theoretical squared gain function for output is different from zero, bounded above and, for gaps, close to half at all frequencies (see table 1), deviations of estimated squared gain functions from the ideal shape give a synthetic idea of the effects of each filter.

Figure 8 presents the estimated squared gain functions for 12 filters (Polynomial, HP, BP, Hamilton, FOD, LD, BN, BQ, UC, Wave, HPa and BPa) using one realization of the output series when all disturbances are stationary. Given the DGP, the spectral properties of output are similar across replications and only minor changes in the high frequency portion of the spectrum are visible. Thus, there is no loss of generality focusing on one realization. Furthermore, because the disturbances driving output and their propagation features are the same as those driving, e.g., consumption, investment or real wages, the output squared gain functions is sufficient to understand the qualitative distortions each filter produces.

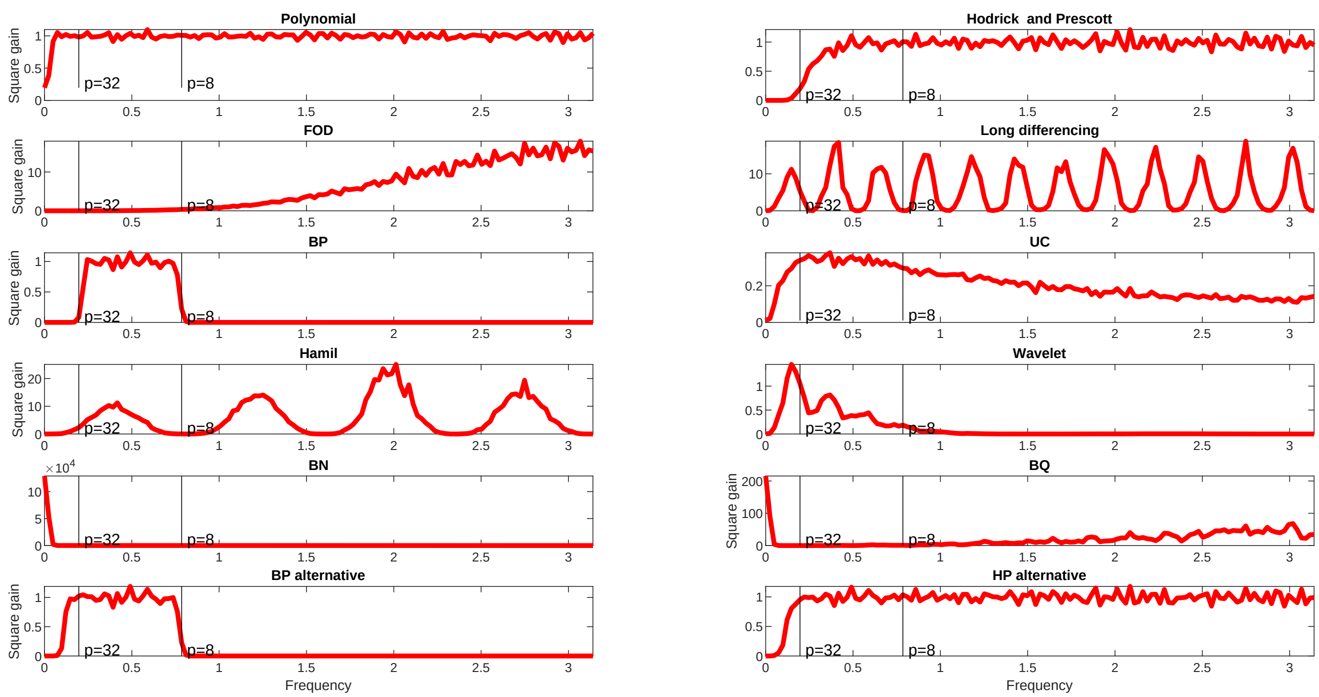

Figure 8: Estimated squared gain function for output: selected filters.

Figure 8 displays some known features and some less known ones. The squared gain of the polynomial approach is zero at the zero frequency and one everywhere else. Thus, the procedure removes very long run variability and leaves the rest of the spectrum of the original series untouched. In other words, the cycles the procedure generates have roughly the same features as in the original series and the frequency distribution of the variance is similar. The squared gain of the HP filter, on the other hand, shows the characteristics high-pass features of the approach. The filter eliminates low frequency variability, keeps the high frequency variability unchanged and, at business cycle frequencies, smoothly eliminates power, when moving from cycles of 2 to 8 years. The band pass filter knocks out low and high frequency variability and passes the business cycle frequencies almost unchanged. Because the sample is finite, the filter displays some compression, in the sense that at some of the selected frequencies, the squared gain is less than one.

Perhaps more interesting is the squared gain of the Hamilton filter, which has not been yet explicitly described in the literature. Because the procedure uses y_{t+m} as dependent variable in the projection equation, the gain function is zero at $m/2$ separate frequencies. Between these frequencies the filter has a bell- shaped squared gain and, at the vertex, the height exceeds 10. Thus, while at certain frequencies the variability of the original series is eliminated, at others, the variability of the filtered series is 10 or more times larger than the variability of the original series. Apart from this severe alteration in the distribution of the variance by frequency, it is clear that the gain function does not resemble the one of a business cycle filter: it emphasizes frequencies of the spectrum not necessarily connected with meaningful cycles and creates excess variability at 'uninteresting' frequencies. Hamilton [2018] mentions that the cycles his approach produces are similar to those generated by a LD filter. Figure 8 confirms that the squared gain of the latter has, qualitatively, the same features as the Hamilton filter. Nevertheless, because I take a 24-quarter rather than an 8-quarter difference, the LD filter also emphasizes low frequency variability.

The FOD approach has a familiar squared gain function: it attributes all the variability of the original series in the low and business cycle frequencies to the trend while the cycle captures, primarily, very high frequency variability.

The BN and the BQ filters have qualitatively similar squared gain functions, despite the fact that they identify latent components with different assumptions (in the BN decomposition the two components are correlated; in the BQ they are not). Quantitatively, this source of misspecification matters for the results and the BN filter is, in general, worse. Because the squared gain function at the very low frequencies is large for both filters, the estimated cycles display strong very low frequency variability and considerable persistence.

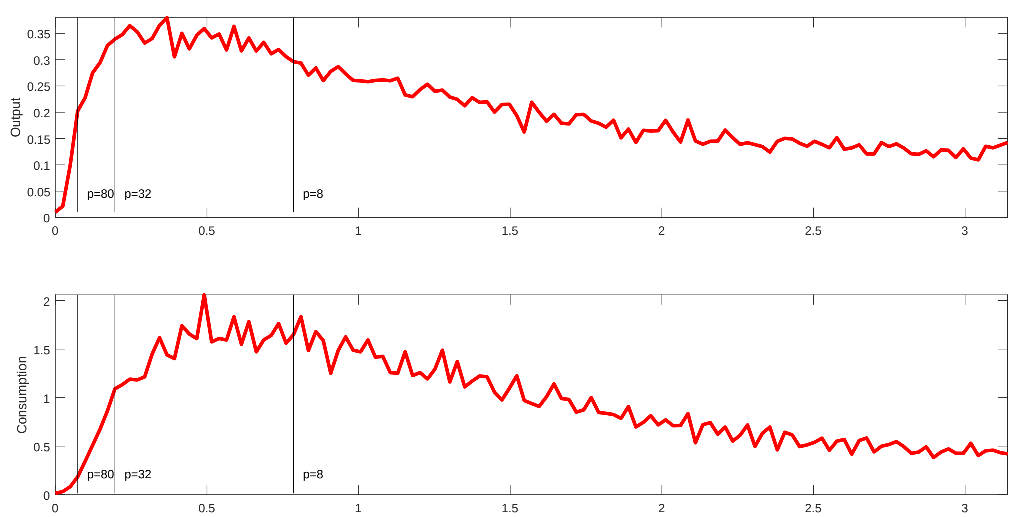

Figure 9: Squared Gain: UC filter

The UC and the Wavelet filters have gains functions that are different from the others. In particular, the squared gain of the UC filter is less than one at all frequencies and has a vertex of around 0.4 in the low frequencies. Thus, it seems to recognize that gaps display variability at all frequencies and possess substantial power in the low frequencies. However, this shape of the gain function is specific to the output series. As shown in figure 9, the squared gain for, e.g. consumption, produces a cyclical component with a large magnification in the variance at business cycle and at certain high frequencies. The Wavelet filter also has a squared gain function with no zeros except at the zero frequency. However, it magnifies the variability of the original series at low frequencies and this accounts for its somewhat mixed performance in the exercises I run.

Figure 8 also plots the gain function of the alternative HP and BP filters, both of which are designed to capture cycles from 8 to 64 quarters periodicity. Clearly, by changing the parameters of the filters, one can shifts the low frequency variations present in the observable data to the cycle. Still, these changes do not reduce the squared gain at business cycle or high frequencies. Thus, all the variability at these frequencies is still mistakenly attributed to the cycle and this accounts for the generally inferior performance of the alternative HP and BP approaches in the overall rankings.

Although the quantitative details change when the TFP disturbance has a unit root, the estimated squared gain functions are roughly similar (see figure 15 in the appendix). Noticeable differences occur for the BP filter, whose compression increases; the Wavelet filter, whose magnification in the low frequencies increases; and for BN and BQ filters, whose estimated squared gain function changes shape and it is now qualitatively similar to the one of the FOD filter. Thus, apart from the BN and BQ filter, none of the procedures sees any difference in the two exercises I run and fail to recognize that gaps and transitory components are objects with different statistical properties and different economic interpretation.

7.1 Discussion

The evidence I produced indicates that the Polynomial approach is superior to other filters when extracting the gaps and figure 8 explains why. Because the estimated squared gain is one at the majority of the frequencies, there is a general overestimation of the gaps variance but the persistence and variance share by frequency are roughly matched. Thus, the ups and downs in the filtered series and in the theory-based gaps are similar in terms of timing and durations. The same argument holds, in part, when extracting transitory components. However, because transitory components display a smaller amount of very low frequency variability, distortions are larger.

Because in the DGP potentials explain an important portion of the variance at business cycle frequencies and gaps explain a large fraction of the variance at low frequencies, HP and BP filters distort the frequency distribution of the variance of the latent variable of interest. In fact, most of the low frequency variations are attributed to the trend and all the business cycle variations are attributed to the cycle. This means that the persistence of the gap processes is underestimated and that the filtered series have patterns of ups and downs that do not generally match those of theory-based gaps in terms of timing, durations, and amplitudes. While the frequency distribution of the variance of the two latent components changes when TFP has a unit root is smaller, HP and BP filters fail to capture the fact that the low frequency transitory variability is significant. Thus, distortions are also large also in this case.

Since the HP filter is the leading procedure to extract cycles in international institutions (e.g. BIS or OECD), further discussion is warranted. The standard HP filter uses a smoothing parameter of $\lambda = 1600$, which typically is interpreted as indicating that the standard deviation of the cycle is 40 times larger than the standard deviation of the second difference of the trend. Hamilton [2018] criticizes this choice of λ suggesting that estimates of the ratio obtained in state space models that approximate a one-sided HP filter are much smaller. When I compute the range of theoretical λ values obtained by taking the variability of the gaps (transitory components) to the second difference of the potential (permanent components) across series I find that indeed they are much smaller than 1600 and in the range of [3,24]. However, as already mentioned, gaps display quite a lot of low frequency variations and only when $\lambda = 51200$ these variations become part of the estimated cycle. Note that such a λ value is close to the one typically used to extract financial cycles and that, with $\lambda = 51200$, the absolute performance of HP filter in the Monte Carlo exercise improves. Thus, when the two latent components have similar spectral properties, are potentially correlated, and the gaps are not iid, λ does not have the standard interpretation given in the literature. In other words, application of standard state space approaches to estimate the smoothing parameter would lead to important distortions when attributing low frequency components to trends or to cycles.[3]

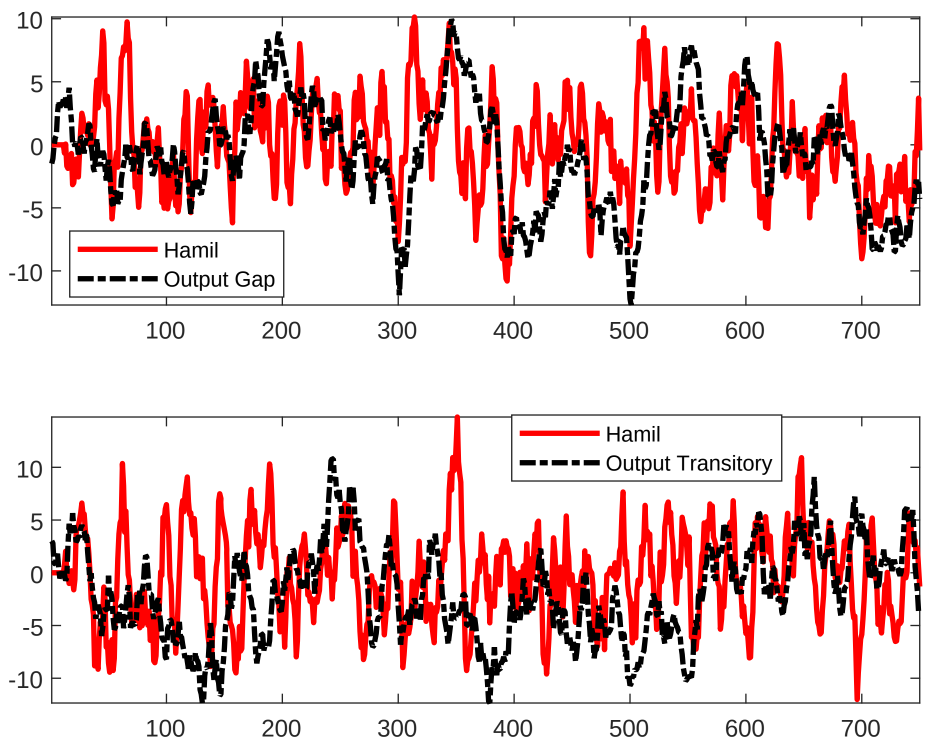

Figure 10: True output gap, true output transitory and estimated Hamilton cycles.

Hamilton [2018] suggested that the local projection filter is a reliable tool to extract cyclical fluctuations. My Monte Carlo exercise shows that when data is generated by standard models, the filter does poorly in extracting both gaps and transitory components. A few reasons may explain the outcomes. First, gaps and potentials are correlated while the projection equation used to separate them assumes that they are not. Second, even though the approach does not impose unit roots, unit roots are in fact removed (these are the zeros in the estimated squared gain function). Even when a unit root is present in the data, the approach overdifferences the data, thus creating spurious MA components in the estimated cycles. While theoretically important, these two reasons may not be crucial to explain the poor performance of the approach because the LD filter, which implicitly errs in the same direction, has a much better performance than the local projection filter. The main

[3]The UC setup I use is not appropriate to quantify the magnitude of the distortions in λ because the trend is correctly assumed to be a random walk.

reasons for the poor performance of approach are: the filter does not recognizes low frequency gap variations and attributes them to the trend; and it magnifies the importance of certain high frequency variations, which have little economic interpretation. Figure 10 illustrates the difficulties of the filter in replicating the dynamics of the output gap (transitory output) in one simulation.

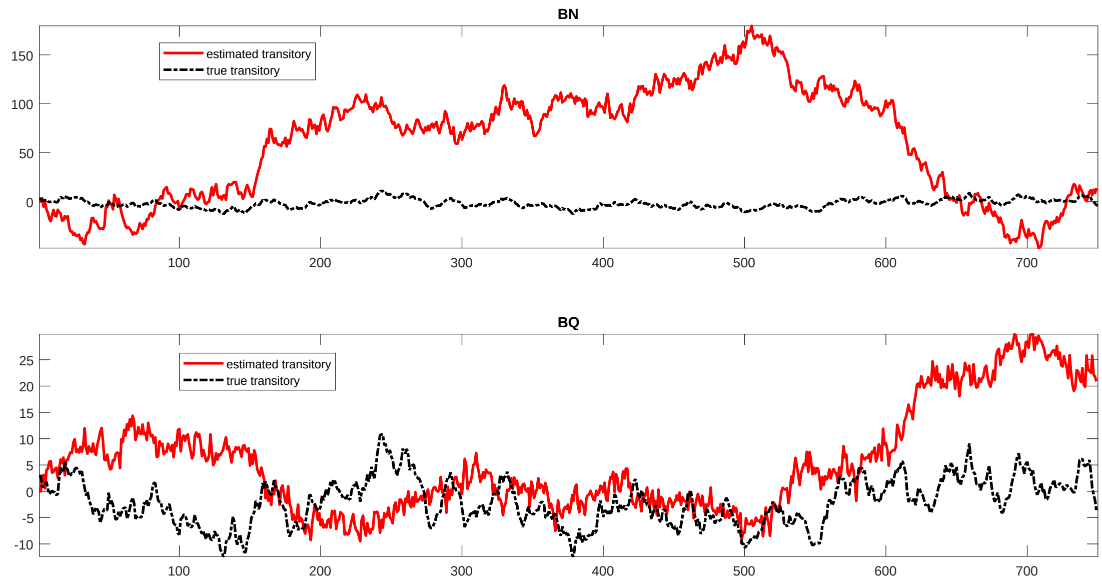

Figure 11: True and estimated transitory components: BN and BQ filters.

VAR-based decompositions perform well for hours but not for output and the distortions in terms of MSE, correlation with the true component, or variability are large, both in absolute and in relative terms. This is true when all shocks are stationary and when the TFP shock has a unit root. When shocks are stationary, the VAR used to extract latent variables is misspecified (since output is stationary, it is overdifferenced) this may be one of the reasons for the poor performance of BN and BQ approaches for output gap extraction.

When TFP has a unit root, the overdifferencing problem disappears. Still, both approaches are poor in extracting the transitory component of output. To illustrate the problem, figure 11 presents the estimated BN and BQ and true transitory output component for one simulation. Clearly, the estimated BN transitory component shows considerable low frequency variations which are absent from the true transitory component and due to the fact that the filter forces the transitory and the permanent components to be highly correlated. The BQ filter also produces a transitory component with low frequency variations, but also increases high frequency variations. Hence, in both cases some low frequency variations present in the theory-based permanent components are erroneously attributed to the transitory component.

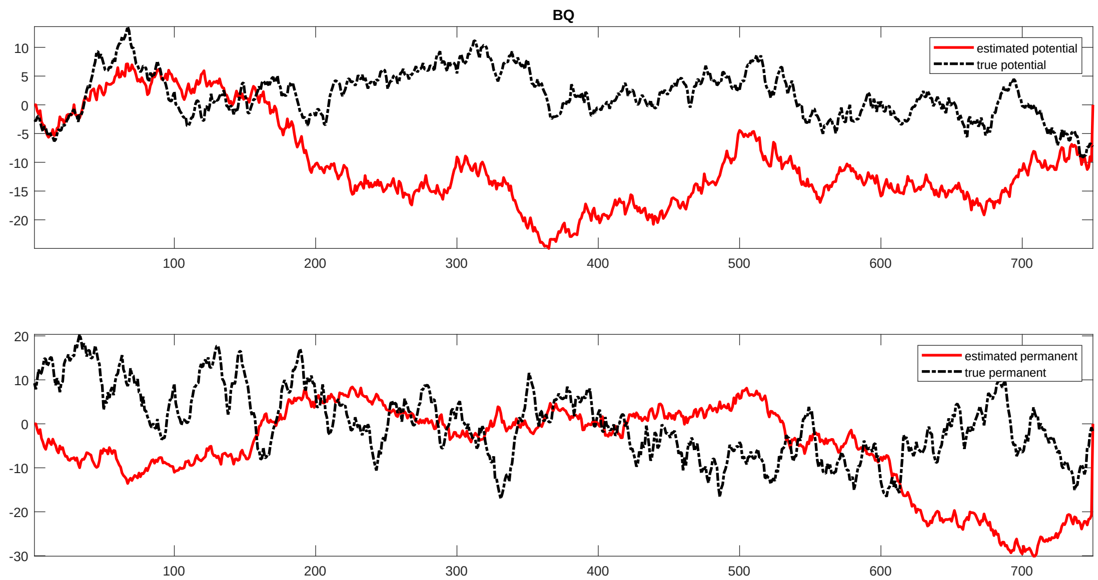

Figure 12: True and estimated potential and permanent components BQ filter.

Coibon et al. [2018] have argued that a BQ decomposition can be used to measure the dynamics of potential output. My results do not support their choice. Instead they indicate that, if the data has been generated by a DSGE model, the estimated permanent component will tend to overstate the dynamics of potential fluctuations. Figure 12, which plots potential and permanent output components in one simulation together with the estimated BQ permanent components, clearly shows the problem. The BQ permanent component displays too much low frequency variations (there are very long drifts in the ups and downs) and too little medium-business cycle variations.

Why is the BQ decomposition so poor? It turns out that the VAR used to extract latent variables is misspecified, even when TFP has a unit root, because of the deformation problems studied in Canova and Ferroni [2019]. In particular, the VAR model is bivariate (trivariate) while the DGP has seven shocks; not all the states of the DGP enter the empirical model. Thus, not only seven shocks are compressed into two (three) innovations; since states are omitted, some VAR innovations are serially correlated even when a generous lag length is used. Deformation problems increase the persistence of the estimated shocks and mess around the correlation structure between the estimated and actual shocks. Indeed, the estimated BQ transitory output series are more persistent than the true transitory output series and the correlation between the TFP disturbances and the estimated supply shocks is low (0.43) since the latter captures a number of stationary demand shocks present in the DGP. When the sample is short (T=150), standard problems estimating long term quantities with small samples are added, see Erceg, Guerrieri, and Gust [2005]. Since a long lag length is needed to reduce states omission in the VAR, parameter estimation may be further compromised, making inference about the latent components problematic.

Ramey and Zubairy [2018] have used estimates of trend output to scale down the variables prior to the computation government spending multipliers. While they use a polynomial approach and thus minimize the distortions if the DGP belongs the class of DSGE models I consider, the practice is dangerous since inference depends on the quality of the preliminary trend output estimates.

7.2 Optimal filters for DSGE-based gaps

None of the procedures I have employed approximates the ideal filter for the class of DGPs I consider, making inferential distortions potentially large. When the proportion of the variance of the two latent components at medium and business cycle frequencies is roughly similar, standard filters generate biases. Different filters carve the spectrum differently, but there is the tendency to attribute most of the low frequency variations to the trend and the majority of the business cycle variations to the cycle, muting the persistence and the dynamic properties of the estimated cycles, altering the sequence of turning points, and the properties of amplitudes and durations of business cycle phases.

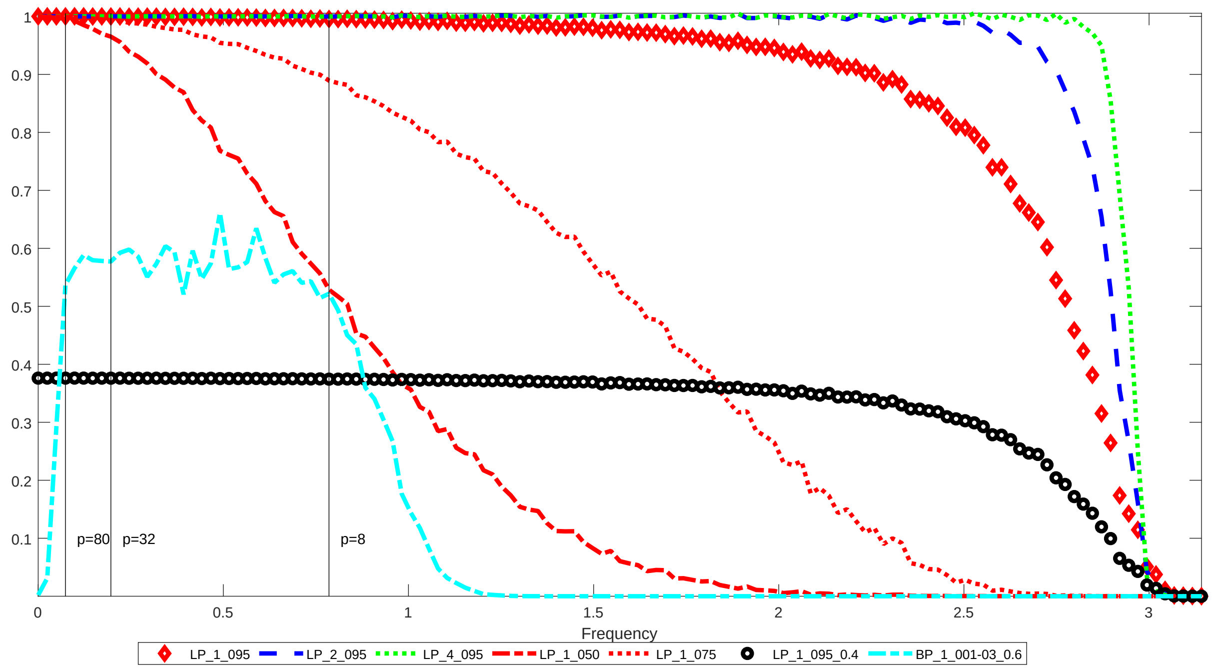

Figure 13: Squared gain functions, Butterworth filters.

How does one then proceeds in practice? One obvious way is to setup a structural model, estimate its parameters by conventional methods, and with mode estimates and some initial conditions, generate model-based latent quantities of interest. Examples of such an approach exist in the literature, see e.g. Justiniano et al. [2013] or Furlanetto et al. [2020]. Clearly, if the sample is short and the prior insufficiently tight, estimates of the latent variables will reflect the noise present in the parameter estimates. Furthermore, if the model is misspecified, biases will litter estimates of the latent quantities. While not much can be done about small samples, model misspecification can be taken care, in part, with the approach of Canova and Matthes [2018]. The setup is particularly useful because it forces consistency of parameter estimates across models with different features and this may help to robustify the measurement of latent quantities.

While the above solution is appealing, most researchers may still prefer to be agnostic about the process generating the data and, when selecting the filter to employ, only willing to take into consideration the frequency domain features I have emphasized in section 3. In this case, is there an alternative filter whose gap estimates are uniformly superior to those I consider?

Engineers extensively employ Butterworth filters in their analyses, because they are flexible, have uniform squared gain across the frequencies of interest, have a ARMA representation, and do not feature the side loops. In particular, Butterworth filters can be designed to make sure that the estimated cycle features significant low frequency variations and the estimated trend significant business cycle variations. Figure 13 shows the gain function of a number of such filters as function of the polynomial order (n) used to filter the data (reported are n=1,2,4); the cutoff point ω, where the squared gain declines (reported are 0.95π, 0.75π and 0.50π); and of the scale parameter G_0, determining the height of the squared gain (reported are $G_0 = 1, 0.4$). Clearly, by appropriately choosing the three parameters, one can give the function a variety of interesting shapes and mimic, e.g., low pass, high pass, or band pass filters. In fact, it can be shown the HP and BP filters I have used in this paper, are special Butterworth filters. For the purpose of gap estimation, the most relevant gain function is the black one (with circles). Note that has a uniform height of 0.4 up to $\omega = 2/3\pi$. Thus, it would attributed a portion of the low frequency variance of the data to the cycle and a portion of the business cycle frequency variance to the trend.

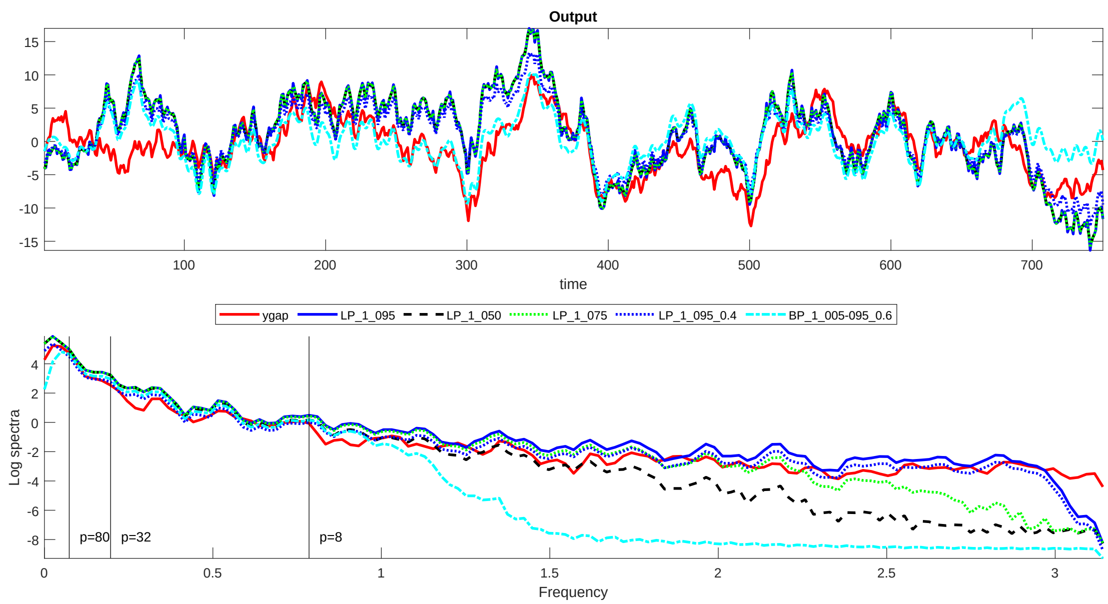

Figure 14: Output gap and Butterworth filtered series and log spectra.

Figure 14 reports the output gap generated by one simulation together with the estimated cycles obtained with different Butterworth filters and their log spectral density. The filtered series display the low frequency movements present in the true output gap and replicate quite well the distribution of variance of the process in the low and business cycle frequencies. When compared with the time paths reported, e.g. in figures 10 or 11, the match is clearly superior. The results are not specific to the output gap series. Figures A1-A4 in the on-line appendix present detailed information for 8 gaps series and 8 transitory series generated by one simulation of the model, using the same Butterworth filter for all series (for the gaps, a low pass filter, n=1,$\omega = 0.95\pi$, $G_0 = 0.4$; for the transitory components, a high pass filter, n=1, $\omega = 0.004$ and $G_0 = 1$).

The last column in tables 2 and 3 report the number of times Butterworth estimates are superior to those obtained with the best performing method for each statistics across series and simulations; the last column of tables 4-10 gives the average statistics across Monte Carlo simulation. It is remarkable that the Butterworth cycle is closer to the actual gap series than the previously selected best approach in 37 out of 66 cases when T=750 and in 33 out of 66 cases when T=150, and improves the average contemporaneous correlation with the actual gap series for all variables and both sample sizes. Thus, if one prefers a statistical to a model-based filter when measuring gaps, a careful design of Butterworth filters may help to solidify economic inference.

The performance when extracting the transitory component is less impressive, as it is difficult to twist a Butterworth filter to produce a cycle which simultaneously displays larger variability at low frequencies and smaller variability at business cycle frequencies relative to the permanent component (as reported in table 1). To do so, one would need a non-uniform gain function across the frequencies of interest Still, even in this less than ideal case, the approach does better or, at least, not worse than the previously selected best approach in over 1/3 of the cases the tables consider.

8 Conclusions

The punchline of the paper is simple. If the data has been generated by the class of models macroeconomists employ to interpret the dynamics of aggregated data and to provide out-of-sample predictions, then the toolkit of filters available to separate the observable data into latent components is inappropriate. If one has to choose among the lame horses that are available, the oldest (Polynomial) and the simplest (differencing) turn out to be the best.

These conclusions obtain because in theory, gaps and transitory fluctuations both have substantial low frequency variability; potentials and permanent fluctuations both have considerable business cycle variability, and the frequency distribution of the variance of gaps and potentials is similar. In this situation, filters that carve the spectrum by frequencies are unsuitable, because 'the trend is the cycle and the cycle is the trend', as are methods that fail to recognize the nature of low frequency variations, and that emphasize either high frequency variations, or fluctuations located at uninteresting frequencies. The polynomial approach produces the smallest distortions when measuring the gaps because, away from the zero frequency, it leaves the frequency distribution of variance of the level data unchanged. For the same reason, it is also close to the top when one wants to recover the transitory components. On the other hand, long differencing works better for transitory fluctuations, because it is able to attribute a portion of the low frequency variations to the cycle.

Because the performance of popular filters is far from satisfactory, one should be careful in using their output to evaluate the state of the business cycle, to forecast inflation or unemployment, or to provide policy recommendations. The warning is even more important when the sample is short, the filter requires parameter estimation and inference needs real time parameter estimates.

One could clearly take the position that the class of structural models I used to simulate data has nothing to do with the real world and therefore that the exercises I run have no information about the usefulness of existing procedures. I have shown that the features I emphasize also obtain in models with additional or different frictions and, even, with different organizing principles. Thus, unless one is willing to dismiss many of the currently available macroeconomic models, one must find a different way out of the conundrum I brought to light.

I have suggested two potential solutions. One solution is to go structural and to compute gaps (transitory components), conditional on a model and the estimated parameters. While straightforward, some researchers may be reluctant to follow such an approach, given that even complex models are not the DGP of the data and apparently innocent estimation choices may impair inference, see e.g. Canova [2014]. If robustness is the main concern, the approach of Canova and Matthes [2018] can be used to provide more reliable estimates of the latent components.

The other solution is to design methods that take into account the features of the DGP. A useful approach should allow low frequency variations to be present in the estimated cycle and business cycle frequency variations to be present in the estimated trend. I have described a class of filters that can be rigged to produce estimated latent components with these features and showed that they are uniformly superior to the available methods for gaps and competitive with the best approaches for transitory components. Clearly, a more detailed study of the properties of this class of filters is needed to strengthen user confidence about the approach.

Two additional implications of the results are worth emphasizing. While it is standard to think of economic and financial cycles as being distinct, in the sense that the largest share of the variances is located at different frequencies of the spectrum, see e.g. Borio [2012], the fact that models with or without financial constraints have similar features and that gaps and transitory components have considerable power in the low frequencies in both situations suggests that two phenomena may not be distinct. Perhaps, it is the insistence of macroeconomists on focusing attention on cycles of two to eight years that has given a misleading impression of what it is in macroeconomic data. I will explore this issue in future work.

There has been an industry over the last 30 years trying to collect stylized business cycle facts, both to inform the construction of realistic models and to test them, see Angeletos et al. [2019] for a recent example. These exercises generally focus attention on variations with 8 to 32 quarters periodicity and disregard the rest. Given the results I have presented, it is perhaps desirable to switch attention from 8 to 32 quarter to low frequency cycles, or at the minimum, take into consideration the fact that most of the variance in the data is not located at the so-called business cycles frequencies, see e.g. Kulish and Pagan [2019]. Paying more attentions at the spectral properties of the data will help researchers to better understand the kind of models which are consistent with the data.

References

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

Marios Angeletos, Fabrice Collard, and Harris Dellas. Business cycle anatomy. Technical report, MIT manuscript, 2019.

Marianne Baxter and Robert King. Measuring business cycles: Approximate band-pass filters for economic time series. *Review of Economics and Statistics*, 81:575–593, 1999.

Paul Beaudry, Dana Galizia, and Frank Portier. Putting the cycle back into business cycle analysis. Technical report, British Columbia, UCL and Carltron, 2018.

Stephen Beveridge and Charles R. Nelson. A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle. *Journal of Monetary Economics*, 7:151–174, 1981.

Olivier Jean Blanchard and Danny Quah. The dynamic effects of aggregate demand and supply disturbances. *American Economic Review*, 79:655–73, 1989.

Claudio Borio. The financial cycle and macroeconomics: what have we learned? Technical report, BIS working paper 2012, 2012.

Fabio Canova. Detrending and business cycle facts. *Journal of Monetary Economics*, 41:475–540, 1998.

Fabio Canova. Reference cycle and turning points: A sensitivity analysis to detrending and dating rules. *Economic Journal*, 109:126–150, 1999.

Fabio Canova. *Methods for Applied Macroeconomic Research*. Princeton University Press, 2007.

Fabio Canova. Bridging dsge models and the raw data. *Journal of Monetary Economics*, 67:1–15, 2014.

Fabio Canova and Filippo Ferroni. Mind the gap! stylized dynamic facts and structural models. Technical report, CEPR working paper 13948, 2019.

Fabio Canova and Christian Matthes. Dealing with misspecifaction in structural macroeconometric models. Technical report, Norwegian Business School, 2018.

Lawrence Christiano and Terry J. Fitzgerald. The band pass filter. *International Economic Review*, 44:435–465, 2003.

Lawrence Christiano, Roberto Motto, and Massimo Rostagno. Risk shocks. *American Economic Review*, 104:27–65, 2014.

Olivier Coibon, Yuriy Gorodnichenko, and Mauricio Ulate. The cyclical sensitivity in estimates of potential output. *Brookings Papers on Economic Activity*, Fall, 2018.

Dean Corbae, Sam Ouliaris, and C.B. Phillips, Peter. Band spectral regression with trending data. *Econometrica*, 70(3):1067–1109, 2002.

Olga Croitorov, Atanas Hristov, Kieran McMorrow, Phillipp Pfeiffer, Werner Roeger, and Valerie Vandermuellen. The eu cam and the relationship between output gaps, cyclical indicators, and inflation. Technical report, European Commission, DG ECFIN report, 2019.

Stefano Crump, Richard K.and Eusepi, Marc Giannoni, and Aysegul Sahini. A unified approach to measuring u*. Technical report, NBER working paper 25930, 2019.

Marco Del Negro, Mark Giannoni, and Frank Schorfheide. Inflation in the great recession and in new keynesian models. *American Economic Journals: Macroeconomics*, 7:168–196, 2015.

Christopher Erceg, Luca Guerrieri, and Christopher Gust. Can long run restrictions identify technology shocks? *Journal of the European Economic Association*, 3:1237–1278, 2005.

Francesco Furlanetto, Paolo Gelain, and Marzie Taheri-Sanjani. Output gap, monetary policy trade-offs, and financial frictions. *Review of Economic Dynamics*, page forthcoming, 2020.

Angelia Grant and Joshua Chan. A bayesian model comparison for trend-cycle decompositions of output. *Journal of Money Credit and Banking*, 49:525–552, 2017.

James Hamilton. Why you should never use the hodrick and prescott filter. *Review of Economics and Statistics*, 100:831–843, 2018.

Lars Hansen and Thomas Sargent. Formulating and estimating dynamic linear rational expectation models. *Journal of Economic Dynamics and Control*, 2:7–46, 1979.

Alejandro Justiniano, Giorgio Primiceri, and Andrea Tambalotti. Is there a trade-off between output and inflation stabilization. *American Economic Journals: Macroeconomics*, 5:1–31, 2013.

Gunes Kamber, James Morley, and Benjamin Wong. Intuitive and reliable estimates of the output gap from a beveridge and nelson filter. *Review of Economic and Statistics*, 100:550–566, 2018.

Mariano Kulish and Adrian Pagan. Turning points and oscillatory cycles. Technical report, University of Sydney, manuscript, 2019.

Thomas Laubach and John William. Measuring the natural rate of interest redux. Technical report, Brookings, working paper 15, 2015.

Thomas Lubik, Christian Matthes, and Fabio Verona. Assessing us output fluctuations across time and frequencies. Technical report, Fed of Richmond working paper 19-6, 2019.

Adrian Pagan. Business cycle issues: some reflection on a literature. Technical report, University of Sydney and CAMA., 2019.

Valerie Ramey and Sarah Zubairy. Goverment spending multipliers in good times and in bad: Evidence from us historical data. *Journal of Political Economy*, 126:850–901, 2018.

Stephanie Schmitt-Grohe and martin Uribe. Exchange rates and uncovered interest differentials: the role of permanent monetary shocks. Technical report, Colombia University, 2019.

Frank Smets and Rafael Wouters. Shocks and frictions in us business cycles: A bayesian dsge approach. *American Economic Review*, 97:585–606, 2007.

Appendix

Equations of the DGP

Flexible price economy

$$
\begin{aligned}
1 * a_t &= calfa * rkf_t + (1 - calfa) * wf_t - 0 * (1 - calfa) a_t && (4)\\
zcapf_t &= (1/(czcap/(1 - czcap))) * rkf_t && (5)\\
rkf_t &= wf_t + labf_t - kf_t && (6)\\
kf_t &= kpf_{t-1} + zcapf_t && (7)\\
invef_t &= \left(\frac{1}{1 + cbetabar * cgamma}\right) * (invef_{t-1} \\
&+ cbetabar * cgamma * invef_{t+1} + \left(\frac{1}{cgamma^2 * csadjcost}\right) * pkf_t) + qs_t && (8)\\
pkf_t &= -rrf_t - 0 * b_t + \left(\frac{1}{(1 - chabb/cgamma)/(csigma * (1 + chabb/cgamma))}\right) * b_t \\
&+ (crk/(crk + (1 - ctou))) * rkf_{t+1} + ((1 - ctou)/(crk + (1 - ctou))) * pkf_{t+1} && (9)\\
cf_t &= (chabb/cgamma)/(1 + chabb/cgamma) * cf_{t-1} + (1/(1 + chabb/cgamma)) * cf_{t+1} \\
&+ ((csigma - 1) * cwhlc/(csigma * (1 + chabb/cgamma))) * (labf_t - labf_{t+1}) \\
&- (1 - chabb/cgamma)/(csigma * (1 + chabb/cgamma)) * (rrf_t + 0 * b_t) + b_t && (10)\\
yf_t &= ccy * cf_t + ciy * invef_t + g_t + crkky * zcapf_t && (11)\\
yf_t &= cfc * (calfa * kf_t + (1 - calfa) * labf_t + a_t) && (12)\\
wf_t &= csigl * labf_t + (1/(1 - chabb/cgamma)) * cf_t - \frac{(chabb/cgamma)}{(1 - chabb/cgamma)} * cf_{t-1} && (13)\\
kpf_t &= (1 - cikbar) * kpf_{t-1} + cikbar * invef_t + cikbar * (cgamma^2 * csadjcost) * qs_t && (14)
\end{aligned}
$$

Sticky price - wage economy

$$
\begin{aligned}
mc_t &= calfa * rk + (1 - calfa) * w_t - 1 * a_t - 0 * (1 - calfa) * a_t && (15)\\
zcap_t &= (1/(czcap/(1 - czcap))) * rk_t && (16)\\
rk_t &= w_t + lab_t - k_t && (17)\\
k_t &= kp_{t-1} + zcap_t && (18)\\
inve_t &= (1/(1 + cbetabar * cgamma)) * (inve_{t-1} + cbetabar * cgamma * inve_{t+1} \\
&+ (1/(cgamma^2 * csadjcost)) * pk_t) + qs_t && (19)\\
pk_t &= -r_t + pinf_{t+1} - 0 * b_t + (1/((1 - chabb/cgamma)/(csigma * (1 + chabb/cgamma)))) * b_t \\
&+ (crk/(crk + (1 - ctou))) * rk_{t+1} + ((1 - ctou)/(crk + (1 - ctou))) * pk_{t+1} && (20)
\end{aligned}
$$

$$
\begin{aligned}
c_t &= (chabb/cgamma)/(1+chabb/cgamma)*c_{t-1}+(1/(1+chabb/cgamma))*c_{t+1} \\
&+ ((csigma-1)*cwhlc/(csigma*(1+chabb/cgamma)))*(lab_t-lab_{t+1}) \\
&- (1-chabb/cgamma)/(csigma*(1+chabb/cgamma))*(r_t-pinf_{t+1}+0*b_t)+b_t \quad (21)\\
y_t &= ccy*c_t+ciy*inve_t+g_t+1*crkky*zcap_t \quad (22)\\
y &= cfc*(calfa*k_t+(1-calfa)*lab_t+a_t) \quad (23)\\
pinf_t &= (1/(1+cbetabar*cgamma*cindp))*(cbetabar*cgamma*pinf_{t+1}+cindp*pinf_{t-1} \\
&+ \frac{((1-cprobp)*(1-cbetabar*cgamma*cprobp)/cprobp)}{((cfc-1)*curvp+1)}*mc_t)+spinf_t \quad (24)\\
w_t &= (1/(1+cbetabar*cgamma))*w_{t-1}+(cbetabar*cgamma/(1+cbetabar*cgamma))*w_{t+1} \\
&+ \frac{cindw}{(1+cbetabar*cgamma)}*pinf_{t-1}-\frac{(1+cbetabar*cgamma*cindw)}{(1+cbetabar*cgamma)}*pinf_t \\
&+ \frac{(cbetabar*cgamma)}{(1+cbetabar*cgamma)}*pinf_{t+1} \\
&+ \frac{(1-cprobw)*(1-cbetabar*cgamma*cprobw)}{((1+cbetabar*cgamma)*cprobw)}*\frac{1}{((clandaw-1)*curvw+1})\\
&* (csigl*lab+\frac{1}{(1-chabb/cgamma)}*c_t \\
&- ((chabb/cgamma)/(1-chabb/cgamma))*c_{t-1}-w_t)+1*sw_t \quad (25)\\
r_t &= crpi*(1-crr)*pinf_t+cry*(1-crr)*(y_t-yf_t) \\
&+ crdy*(y_t-yf_t-y_{t-1}+yf_{t-1})+crr*r_{t-1}+ms_t \quad (26)\\
& \quad (27)
\end{aligned}
$$

Law of motion of shocks

$$
\begin{aligned}
a_t &= crhoa*a_{t-1}+ea_t \quad (28)\\
b_t &= crhob*b_{t-1}+eb_t \quad (29)\\
g_t &= crhog*g_{t-1}+eg_t+cgy*ea_t \quad (30)\\
qs_t &= crhoqs*qs_{t-1}+eqs_t \quad (31)\\
ms_t &= crhoms*ms_{t-1}+em_t \quad (32)\\
spinf_t &= crhopinf*spinf_{t-1}+epinfma_t-cmap*epinfma_{t-1} \quad (33)\\
epinfma_t &= epinf_t \quad (34)\\
sw_t &= crhow*sw_{t-1}+ewma_t-cmaw*ewma_{t-1} \quad (35)\\
ewma_t &= ew_t \quad (36)\\
kp_t &= (1-cikbar)*kp_{t-1}+cikbar*inve_t+cikbar*cgamma^2*csadjcost*qs_t \quad (37)
\end{aligned}
$$

Table 4: Average MSE, T=750

Variable	POLY	HP	FOD	LD	BP	Wa	Ham	UC	BN	BQ	BW
	Gap										
Y	27.36	25.57	29.07	48.5	27.1	28.41	27.6	**23.14**	10852	184.5	18.74(*)
C	**29.46**	32.42	35.35	45.74	33.85	33.53	32.79	32.44	NaN	NaN	19.63(*)
I	125.37	**85.87**	86.83	240.06	90.35	114.44	109.71	106.12	NaN	NaN	83.33(*)
H	2.79	6.26	7.71	9.26	6.81	5.76	4.7	5.12	**1.94**	**1.94**	0.99(*)
RK	4.95	**2.36**	2.85	6.11	2.57	3.12	3.12	2.85	NaN	NaN	4.37
W	**17.7**	23.72	25.25	31.71	24.51	24.37	23.93	23.95	NaN	NaN	10.09(*)
CapU	15.67	**7.48**	9.03	19.37	8.15	9.89	9.88	9.07	NaN	NaN	13.86
π	**0.17**	0.45	0.65	0.65	0.53	0.45	0.26	0.52	NaN	NaN	0.03(*)
R	**0.16**	0.47	0.76	0.76	0.61	0.53	0.25	0.8	NaN	NaN	0.04(*)
Factor	**2.48**	4.61	5.07	4.64	4.83	4.02	4.15	5.02	NaN	NaN	1.83(*)
	Transitory										
Y	74.71	36.46	**33.96**	75.35	35.81	47.45	49.98	37.1	24543	302.32	56.11
C	92.58	41.55	**40.1**	80.13	41.03	53.65	51.61	42.91	NaN	NaN	63
I	307.1	200.36	**181.74**	385.41	198.38	252.35	268.54	238.4	NaN	NaN	283.76
H	16.01	10.37	**9.67**	18.38	10.12	12.3	13.62	9.78	521.27	521.27	14.28
RK	12.14	7.54	**7.31**	12.68	7.48	9.11	8.91	7.57	NaN	NaN	10.41
W	51.56	25.53	**25.08**	46.39	25.34	32.38	30.26	26.63	NaN	NaN	37.44
CapU	38.47	23.91	**23.17**	40.18	23.69	28.87	28.25	24.23	NaN	NaN	33
π	1.06	0.77	**0.69**	1.31	0.74	0.87	1	0.87	NaN	NaN	0.98
R	1.32	0.99	0.97	1.66	0.9	1.04	1.26	**0.9**	NaN	NaN	1.23
Factor	6.4	4.66	**4.57**	6.54	4.63	5.21	5.24	4.87	NaN	NaN	5.65

Notes: POLY is polynomial detrending, HP is Hodrick and Prescott filtering, FOD is first order differencing, LD is 24 quarter differencing, UC is an unobservable component model, BP is band pass filtering, Ham is local projection detrending, BN and BQ are bivariate Beveridge and Nelson and Blanchard and Quah decompositions, BW the Butterworth filter. Y is output, C consumption, I investment, H hours, RK the real rate, W the real wage, CapU capacity utilization, R the nominal rate, and π the inflation rate. Factor is the first principal component of the nine series. The MSE is computed averaging over 100 data replications. In bold the best approach. A (*) in the last column indicates that the BW filter improves over or is comparable to the best approach.

Table 5: Average real time MSE, T=750

Variable	POLY	HP	FOD	LD	BP	Wa	Ham	UC	BN	BQ	BW
	Gap										
Y	43.84	25.08	27.93	54.02	24.83	27.65	28.29	**21.26**	439.69	222.04	20.73(*)
C	53.13	**33.03**	35.1	52.69	33.4	36.33	35.48	34.36	NaN	NaN	18.44(*)
I	162.86	83.09	81.71	301.44	81.7	127.4	118.88	**74.94**	NaN	NaN	74.76(*)
H	5.35	6.55	7.71	9.1	6.51	5.36	4.91	6.32	**2.06**	**2.06**	1.04(*)
RK	6.7	3.35	3.75	6.1	3.27	**3.11**	3.84	3.19	NaN	NaN	3.95
W	39.03	26.31	28.01	38.73	26.33	27.07	27.63	**21.12**	NaN	NaN	10.72(*)
CapU	21.23	10.63	11.89	19.34	10.38	**9.85**	12.18	10.08	NaN	NaN	12.51
π	0.42	0.45	0.64	0.78	0.49	0.5	**0.26**	0.6	NaN	NaN	0.03(*)
R	0.4	0.49	0.76	0.81	0.61	0.51	**0.25**	0.93	NaN	NaN	0.04(*)
Factor	5.44	5.04	5.33	5.39	5.09	**4.37**	4.54	5.36	NaN	NaN	1.78(*)
	Transitory										
Y	104.69	38.47	**37.49**	69.19	38.86	47.51	49.89	33.64	995.6	481.3	49.1
C	134.98	45.2	**44.78**	85.81	44.92	57.57	53.08	36.17	NaN	NaN	53.99
I	383.51	219.49	**205.35**	415.73	219.43	283.34	302.13	206.23	NaN	NaN	302.89
H	21.14	11.53	11.35	17.27	11.64	13.28	15.29	**9.17**	563.41	563.41	13.96
RK	17.42	9.57	9.42	14.1	9.58	10.96	11.29	**9.03**	NaN	NaN	10.49
W	78.62	21.8	**21.52**	45.11	21.89	29.38	25.85	26.11	NaN	NaN	27.64
capU	55.19	30.34	29.86	44.68	30.37	34.74	35.78	**27.89**	NaN	NaN	33.24
π	1.21	0.78	0.74	1.32	0.75	0.89	1.11	**0.73**	NaN	NaN	0.91
R	1.66	1.07	1.1	1.87	0.97	1.18	1.52	**0.83**	NaN	NaN	1.24
Factor	9.81	5.85	5.87	8.2	5.83	6.7	6.61	**5.23**	NaN	NaN	6.17

Notes: POLY is polynomial detrending, HP is Hodrick and Prescott filtering, FOD is first order differencing, LD is 24 quarter differencing, UC is an unobservable component model, BP is band pass filtering, Ham is local projection detrending, BN and BQ are bivariate Beveridge and Nelson and Blanchard and Quah decompositions, BW the Butterworth filter. Y is output, C consumption, I investment, H hours, RK the real rate, W the real wage, CapU capacity utilization, R the nominal rate, and π the inflation rate. Factor is the first principal component of the nine series. The MSE is computed averaging over 100 data replications. In bold the best approach. A (*) in the last column indicates that the BW filter improves over or is comparable to the best approach.

Table 6: Contemporaneous correlations, T=750

Variable	POLY	HP	FOD	LD	BP	Wa	Ham	UC	BN	BQ	BW
	Gaps and Filtered variables										
Y	**0.65**	0.34	0.03	0.38	0.25	0.41	0.44	0.44	-0.33	0.25	0.71(*)
C	**0.65**	0.33	0.06	0.36	0.23	0.37	0.38	0.31	NaN	NaN	0.72(*)
I	**0.59**	0.27	0.04	0.34	0.19	0.37	0.44	0.43	NaN	NaN	0.65(*)
H	0.86	0.42	0.06	0.5	0.31	0.54	0.68	0.67	**0.93**	**0.93**	0.93(*)
RK	**0.53**	0.41	0.04	0.38	0.29	0.4	0.36	0.22	NaN	NaN	0.55(*)
W	**0.67**	0.26	0.01	0.31	0.16	0.32	0.3	0.06	NaN	NaN	0.77(*)
CapU	**0.53**	0.41	0.04	0.38	0.29	0.4	0.36	0.16	NaN	NaN	0.55(*)
π	**0.93**	0.61	0.17	0.57	0.48	0.61	0.84	0.57	NaN	NaN	1.00(*)
R	**0.94**	0.67	0.29	0.59	0.48	0.6	0.88	0.36	NaN	NaN	1.00(*)
Factor	**0.71**	0.31	0.04	0.43	0.2	0.45	0.41	0.16	NaN	NaN	0.79(*)
	Transitory and Filtered variables										
Y	0.01	0	0	0.01	0	0.01	-0.01	0.01	0.05	0	0.01
C	0.01	0	-0.01	-0.01	0	-0.02	-0.02	0.01	NaN	NaN	0
I	0.01	0	-0.01	0	0	0	-0.01	0	NaN	NaN	0
H	-0.01	0	0	0.01	0	0.01	0	0	-0.02	-0.02	0
RK	0.01	0	-0.01	-0.01	0	-0.01	-0.01	0	NaN	NaN	0.01
W	-0.02	0.01	-0.01	-0.02	0	-0.02	-0.02	0	NaN	NaN	-0.01
CapU	0.01	0	-0.01	-0.01	0	-0.01	-0.01	-0.01	NaN	NaN	0.01
π	0.03	0	0	0.01	0	0.01	0.01	0.01	NaN	NaN	0.01
R	0.01	0	0	0	0	0	0	0	NaN	NaN	0
Factor	0.03	0	0	0	0	0	0	0.01	NaN	NaN	0.01

Notes: POLY is polynomial detrending, HP is Hodrick and Prescott filtering, FOD is first order differencing, LD is 24 quarter differencing, UC is an unobservable component model, BP is band pass filtering, Ham is local projection detrending, BN and BQ are bivariate Beveridge and Nelson and Blanchard and Quah decompositions, BW the butterworth filter. Y is output, C consumption, I investment, H hours, RK the real rate, W the real wage, CapU capacity utilization, R the nominal rate, and π the inflation rate. Factor is the first principal component of the nine series. Correlations are computed averaging over 100 data replications. In bold the best approach. A (*) in the last column indicates that the BW filter improves over or is comparable to the best approach.

Table 7: AR1 coefficient, T=750

Variable	True	POLY	HP	FOD	LD	BP	Wa	Ham	UC	BN	BQ	BW
		Gaps and filtered variables										
Y	0.98	**0.98**	0.83	0.21	0.96	0.93	**0.98**	0.89	0.89	1	0.97	0.98(*)
C	0.98	**0.99**	0.83	0.23	**0.97**	0.93	**0.99**	0.89	0.88	NaN	NaN	0.98(*)
I	0.99	**0.98**	0.89	0.51	0.97	0.93	**0.98**	0.91	0.98	NaN	NaN	0.98(*)
H	0.98	0.95	0.78	0.11	0.93	0.93	**0.97**	0.89	0.9	0.96	0.96	0.96
RK	0.97	0.99	0.83	0.21	**0.97**	0.94	0.99	0.89	0.93	NaN	NaN	0.99
W	0.98	0.99	0.82	0.12	**0.98**	0.94	0.99	0.87	0.94	NaN	NaN	0.99
CapU	0.97	0.99	0.83	0.21	**0.97**	0.94	0.99	0.89	0.91	NaN	NaN	0.99
π	0.94	**0.93**	0.78	0.16	0.9	0.92	0.96	0.88	0.83	NaN	NaN	0.94(*)
R	0.83	**0.80**	0.54	-0.21	0.76	0.91	0.95	0.77	0.47	NaN	NaN	0.85(*)
Factor	0.98	**0.99**	0.8	0.14	0.97	0.93	**0.99**	0.89	0.93	NaN	NaN	0.99(*)
		Transitory and filtered variables										
Y	0.97	**0.98**	0.83	0.2	0.96	0.93	**0.98**	0.89	0.89	0.99	0.98	0.97(*)
C	0.99	**0.99**	0.84	0.26	0.98	0.93	**0.99**	0.89	0.87	NaN	NaN	0.98
I	0.98	0.97	0.89	0.5	0.97	0.93	**0.98**	0.91	**0.98**	NaN	NaN	0.97
H	0.96	**0.95**	0.78	0.1	0.93	0.93	**0.97**	0.89	0.9	0.96	0.96	0.94
RK	0.99	0.98	0.83	0.19	0.97	0.94	**0.99**	0.89	0.93	NaN	NaN	0.98
W	0.99	**0.99**	0.82	0.14	0.98	0.94	**0.99**	0.87	0.94	NaN	NaN	0.98
CapU	0.99	0.98	0.83	0.19	0.97	0.94	**0.99**	0.89	0.93	NaN	NaN	0.98
π	0.94	**0.93**	0.78	0.15	0.91	0.92	0.97	0.88	0.84	NaN	NaN	0.91
R	0.84	**0.81**	0.54	-0.21	0.76	0.91	0.95	0.76	0.41	NaN	NaN	0.77
Factor	0.99	0.98	0.8	0.12	0.96	0.93	**0.99**	0.88	0.93	NaN	NaN	0.97

Notes: POLY is polynomial detrending, HP is Hodrick and Prescott filtering, FOD is first order differencing, LD is 24 quarter differencing, UC is an unobservable component model, BP is band pass filtering, Ham is local projection detrending, BN and BQ are bivariate Beveridge and Nelson and Blanchard and Quah decompositions, BW the Butterworth filter. Y is output, C consumption, I investment, H hours, RK the real rate, W the real wage, CapU capacity utilization, R the nominal rate, and π the inflation rate. Factor is the first principal component of the nine series. The AR1 coefficient is computed averaging over 100 data replications. In bold the best approach. A (*) in the last column indicates that the BW filter improves over or is comparable to the best approach.

Table 8: Variability, T=750

Variable	True	POLY	HP	FOD	LD	BP	Wa	Ham	UC	BN	BQ	BW
		Gaps and filtered variables										
Y	23.62	38.81	4.07	1.56	42.88	3.45	15	**16.62**	6.83	4632.6	81.43	31.08
C	28.77	36.04	2.14	0.82	**29.97**	1.78	10.41	9.93	4.14	NaN	NaN	31.38
I	69.37	171.48	26.61	6.85	235.27	24.01	**84.51**	98.16	139.45	NaN	NaN	129.97
H	6.43	**6.82**	1.39	0.65	9.39	1.15	3.28	4.46	1.26	8.06	8.06	4.98
RK	2.48	6.3	0.4	0.15	6.05	0.33	**2.1**	1.88	0.3	NaN	NaN	5.21
W	20.85	**20.83**	0.9	0.38	17.78	0.68	6.13	4.67	0.6	NaN	NaN	18.04
CapU	7.86	19.98	1.25	0.48	19.17	1.03	**6.64**	5.96	0.8	NaN	NaN	16.52
π	0.58	0.47	0.14	0.07	**0.68**	0.12	0.24	0.39	0.25	NaN	NaN	0.35
R	0.69	**0.59**	0.24	0.22	0.88	0.15	0.27	0.51	0.17	NaN	NaN	0.42
Factor	5.05	**3.13**	0.21	0.09	2.8	0.17	0.97	0.93	0.39	NaN	NaN	2.7
		Transitory and filtered variables										
Y	28.94	42.79	4.07	1.56	42.04	3.38	14.72	**17.11**	6.16	7649.1	140.59	24.2(*)
C	33.58	54.45	2.32	0.86	**38.29**	1.86	13.22	11.65	3.99	NaN	NaN	23.65
I	160.43	**137.45**	25.7	6.62	203.33	23.23	73.7	89.49	100.49	NaN	NaN	110.4
H	8.29	6.96	1.35	0.64	**9.19**	1.1	3.21	4.5	1.15	12.03	12.03	5.25
RK	6.32	5.14	0.38	0.15	**5.33**	0.31	1.85	1.72	0.3	NaN	NaN	3.37
W	20.43	26.62	0.93	0.39	**19.92**	0.7	6.85	5.06	0.74	NaN	NaN	12.45
CapU	20.02	16.3	1.2	0.46	**16.88**	0.97	5.86	5.45	1.01	NaN	NaN	10.69
π	0.57	0.46	0.14	0.06	**0.67**	0.11	0.23	0.37	0.25	NaN	NaN	0.37
R	0.69	**0.58**	0.23	0.21	0.87	0.15	0.27	0.5	0.15	NaN	NaN	0.49
Factor	4.52	**2.11**	0.17	0.08	1.98	0.14	0.69	0.72	0.44	NaN	NaN	1.18

Notes: POLY is polynomial detrending, HP is Hodrick and Prescott filtering, FOD is first order differencing, LD is 24 quarter differencing, UC is an unobservable component model, BP is band pass filtering, Ham is local projection detrending, BN and BQ are bivariate Beveridge and Nelson and Blanchard and Quah decompositions, BW the Butterworth filter. Y is output, C consumption, I investment, H hours, RR the real rate, W the real wage, CapU capacity utilization, R the nominal rate, and π the inflation rate. Factor is the first principal component of the nine series. The variability is computed averaging over 100 data replications. In bold the best approach. A (*) in the last column indicates that the BW filter improves over or is comparable to the best approach.

Table 9: Number of turning points, average durations, average amplitudes, T=750

Variable	True	POLY	HP	FOD	LD	BP	Wa	Ham	UC	BN	BQ	BW
		Output Gap and filtered output										
number TP	125.87	123.4	**124.78**	109.11	124.51	129.74	135.95	112.77	130.24	147.91	111.78	128
DurE	5.67	5.8	**5.73**	6.68	5.79	5.5	5.22	6.39	5.52	4.5	6.49	5.62(*)
DurR	5.71	5.81	**5.75**	6.49	5.75	5.5	5.27	6.32	5.49	5.19	6.36	5.57
AmpE	-2.49	-3.62	**-3.57**	-4.69	-5.09	-1.43	-1.47	-5.67	-1.48	-16.91	-6.47	-2.67(*)
AmpR	2.49	3.62	**3.57**	4.69	5.09	1.43	1.47	5.67	1.48	16.91	6.47	2.67(*)
		Factor Gap and filtered factor										
Number TP	120.07	**120.44**	122.63	103.61	120.51	130.47	135.31	113.42	130.78	NaN	NaN	124.97
DurE	6.01	5.96	5.84	6.87	**5.98**	5.48	6.28	5.25	5.46	NaN	NaN	5.71
DurR	5.92	**5.93**	5.84	7.05	**5.93**	5.46	6.39	5.31	5.63	NaN	NaN	5.75
AmpE	-1.08	-0.83	-0.82	**-1.12**	-1.17	-0.3	-1.36	-0.32	-0.49	NaN	NaN	-0.61
AmpR	1.08	0.83	0.82	**1.12**	1.17	0.3	1.36	0.32	0.49	NaN	NaN	0.61
		Output Transitory and filtered output										
number TP	125.57	123.94	**125.69**	109.32	125.35	130.63	137.32	112.3	130.72	150.61	111.23	123.88
DurE	5.66	5.76	**5.72**	6.56	5.73	5.44	5.19	6.35	5.6	4.18	6.52	5.76
DurR	5.75	**5.79**	5.66	6.55	5.68	5.49	5.2	6.45	5.37	5.32	6.38	5.79(*)
AmpE	-3.4	**-3.62**	-3.58	-4.72	-5.11	-1.41	-1.47	-5.74	-1.44	-22.65	-5.59	-3.62(*)
AmpR	3.4	**3.61**	3.57	4.71	5.1	1.41	1.47	5.73	1.44	22.65	5.59	3
		Factor transitory and filtered factor										
Number TP	121.2	120.79	**121.56**	105.4	121.8	131.03	113.62	135.26	126.26	NaN	NaN	120.69
DurE	5.9	5.96	**5.87**	6.87	5.93	5.45	6.25	5.25	5.72	NaN	NaN	5.96
DurR	5.92	**5.92**	5.94	6.76	5.86	5.44	6.38	5.32	5.76	NaN	NaN	5.92(*)
AmpE	-0.98	-0.74	-0.73	-1	**-1.05**	-0.27	-1.17	-0.29	-0.52	NaN	NaN	-0.74
AmpR	0.98	0.74	0.73	1	**1.05**	0.27	1.17	0.29	0.52	NaN	NaN	0.74

Notes: POLY is polynomial detrending, HP is Hodrick and Prescott filtering, FOD is first order differencing, LD is 24 quarter differencing, UC is an unobservable component model, BP is band pass filtering, Ham is local projection detrending, BN and BQ are bivariate Beveridge and Nelson and Blanchard and Quah decompositions, BW is the Butterworth filter. Y is output and Factor is the first principal component of the nine series. DurE and DurR are the durations of expansions and recessions; AmpE and AmpR the amplitude of expansions and recessions. Statistics are computed averaging over 100 data replications. In bold is the best approach. A (*) in the last column indicates that the BW filter improves over or is comparable to the best approach.

Table 10: Phillips curve and Okun law predictions, T=750

Step ahead	POLY	HP	FOD	LD	BP	Wa	Ham	UC	BN	BQ	BW
	Phillips curve prediction: Output Gap										
1 step	**0.25**	0.26	0.27	0.37	0.26	0.32	0.36	0.27	0.36	0.4	0.21(*)
4 steps	**1.41**	1.5	1.67	1.91	1.67	1.55	1.87	1.53	2.14	2.23	1.04(*)
	Phillips curve prediction: Transitory Output										
1 step	0.21	0.45	0.19	0.14	0.43	0.15	0.14	0.19	0.13	**0.1**	0.21
4 steps	0.79	0.93	0.6	**0.49**	2.89	0.7	0.52	0.77	0.52	0.58	0.68
	Okun law prediction: Output Gap										
1 step	0.24	1.24	0.24	0.19	3.79	0.21	**0.19**	1.21	0.21	0.26	0.44
4 steps	3.67	33.84	2.49	**1.75**	45.94	2.74	2.23	10.88	2.43	2.04	3.13
	Okun law prediction: transitory Output										
1 step	0.32	0.31	0.33	**0.30**	0.31	0.32	0.31	0.37	0.34	0.33	0.32
4 steps	3.02	2.81	2.76	**2.68**	3.05	2.73	2.85	3.27	3.21	2.89	2.93

Notes: POLY is polynomial detrending, HP is Hodrick and Prescott filtering, FOD is first order differencing, LD is 24 quarter differencing, UC is an unobservable component model, BP is band pass filtering, Ham is local projection detrending, BN and BQ are bivariate Beveridge and Nelson and Blanchard and Quah decompositions, BW is the Butterworth filter. The Phillips curve and the Okun law predictions are regression of the form: $x_{t+m} = \alpha_0 + \alpha_1 x_t + \sum_{j=1}^{3} \beta_j y_{t-j}$ where y_{t-j} is the true gap (transitory) or the estimated one, $x_t = \pi_t$ or h_t and m=1,4. Reported the difference in variance of the prediction error between each procedure and the true prediction error, averaged over 100 data replications. In bold the best approach. A (*) in the last column indicates that the BW filter improves over or is comparable to the best approach.

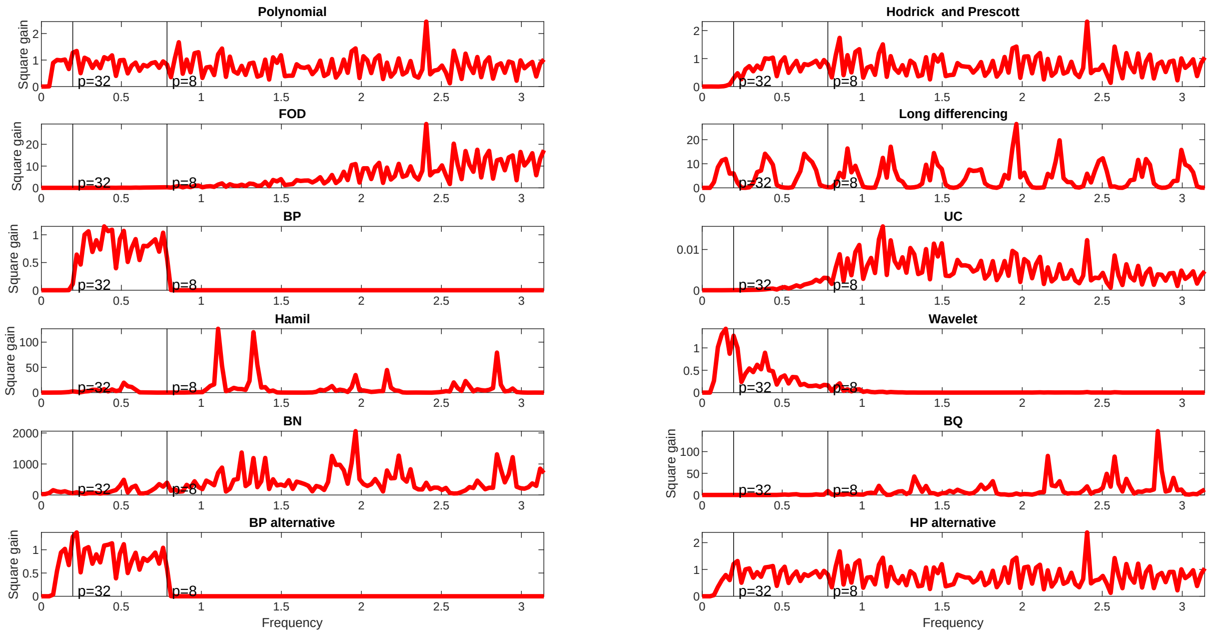

Figure 15: Estimated squared gain functions for output; unit root in TFP

Table A.1: Average MSE, T=750

Variable	HPa	LDa	BPa	Hama	BNa	BQa	FODa	Trigo	BK	UCbiv
	Gap									
Y	24.31	42.79	26.38	27.09	29.07	29.07	32.62	27.16	27.1	315.18
C	30.24	41.98	32.75	31.8	12.34	12.34	36.48	33.88	33.85	NaN
I	95.33	211.45	99.14	116.97	75.29	75.29	120.79	90.67	90.35	NaN
H	5.29	9.11	6.14	3.98	NaN	NaN	8.43	6.82	6.81	NaN
RK	**2.25**	4.88	2.49	3.3	NaN	NaN	3.2	2.58	2.57	NaN
W	22.04	29.56	23.82	22.85	NaN	NaN	26.17	24.53	24.51	NaN
CapU	**7.14**	15.47	7.88	10.47	NaN	NaN	10.13	8.16	8.15	170.23
π	0.38	0.65	0.45	0.22	NaN	NaN	0.65	0.53	0.53	NaN
R	0.38	0.76	0.53	0.21	NaN	NaN	0.76	0.61	0.61	NaN
Factor	4.07	4.66	4.51	3.66	NaN	NaN	5.06	4.84	4.83	NaN
	Transitory									
Y	41.68	65.77	39.94	56.77	33.96	33.96	41.32	35.83	35.81	119.31
C	45.84	67.11	44.14	58.69	34.14	34.14	44.75	41.05	41.03	NaN
I	230.26	357.78	222.85	293.58	198.98	198.98	229.67	198.61	198.38	NaN
H	11.55	17.01	11.09	14.61	NaN	NaN	11.99	10.13	10.12	NaN
RK	8.19	11.13	7.94	9.73	NaN	NaN	8.04	7.48	7.48	NaN
W	27.7	38.84	26.78	33.99	NaN	NaN	27.06	25.34	25.34	NaN
CapU	25.94	35.26	25.17	30.83	NaN	NaN	25.46	23.7	23.69	113.85
π	0.85	1.26	0.81	1.04	NaN	NaN	0.93	0.74	0.74	NaN
R	1.08	1.62	0.98	1.3	NaN	NaN	1.27	**0.9**	**0.9**	NaN
Factor	4.89	6.02	4.81	5.55	NaN	NaN	4.88	4.63	4.63	NaN

Notes: HPa is the modified Hodrick and Prescott filtering, FODa is the forth order differencing, LDa is 16 quarter differencing, BPa is the modified band pass filtering, Hama is modified local projection detrending, BNa and BQa are trivariate Beveridge and Nelson and Blanchard and Quah decompositions, Trigo and BK are the trigonometric and the Baxter and King versions of a BP filter, UCbiv is a bivariate UC filter. Y is output, C consumption, I investment, H hours, RK the real rate, W the real wage, CapU capacity utilization, R the nominal rate, and π the inflation rate. Factor is the first principal component of the nine series. Statistics are computed averaging over 100 data replications. For C and I BNa and BQa refer to the comparison of filtered C/Y and I/Y with the data. In bold are cases where the reported statistic improves the best result presented in table 4.

Table A.2: Average real time MSE, T=750

Variable	HPa	LDa	BPa	Hama	BNa	BQa	FODa	Trigo	BK	UCbiv
	Gap									
Y	25.26	44.46	22.44	27.71	27.93	27.93	32.59	26.02	24.83	**8.61**
C	34.6	47.47	32.24	35.45	13.45	13.45	37.88	33.83	33.4	NaN
I	92.97	259.48	86.36	128.54	72.86	72.86	126.45	87.39	81.7	NaN
H	6.15	9.11	5.77	4.06	NaN	NaN	8.61	6.67	6.51	NaN
RK	3.3	5.17	3.12	3.71	NaN	NaN	4.09	3.32	3.27	NaN
W	27.01	34.22	24.85	25.74	NaN	NaN	30.06	27	26.33	NaN
capU	10.47	16.39	9.89	11.76	NaN	NaN	12.96	10.53	10.38	**3.76**
π	0.44	0.77	0.46	**0.23**	NaN	NaN	0.67	0.53	0.49	NaN
R	0.44	0.81	0.52	**0.21**	NaN	NaN	0.79	0.65	0.61	NaN
Factor	4.96	5.08	4.72	**4**	NaN	NaN	5.37	5.15	5.09	NaN
	Transitory									
Y	40.83	62.42	39.27	55.83	37.49	37.49	42.85	38.65	38.86	**7.52**
C	47.14	68.28	**44.55**	58.09	49.59	49.59	48.62	45.08	44.92	NaN
I	236.45	397.07	237.89	345.75	230.58	230.58	243.9	218.85	219.46	NaN
H	11.92	17.11	11.58	16.05	NaN	NaN	13.17	11.63	11.65	NaN
RK	9.79	12.64	9.88	12.49	NaN	NaN	9.81	9.55	9.59	NaN
W	22.49	34.8	22.6	30.08	NaN	NaN	22.43	21.93	21.89	NaN
capU	31.02	40.04	31.29	39.58	NaN	NaN	31.1	30.25	30.37	**2.06**
π	0.84	1.3	0.79	1.13	NaN	NaN	1.03	0.78	0.75	NaN
R	1.13	1.7	1.01	1.52	NaN	NaN	1.42	1.01	0.97	NaN
Factor	6.04	7.18	5.85	6.78	NaN	NaN	6.18	5.8	5.83	NaN

Notes: HPa is the modified Hodrick and Prescott filtering, FODa is the forth order differencing, LDa is 16 quarter differencing, BPa is the modified band pass filtering, Hama is modified local projection detrending, BNa and BQa are trivariate Beveridge and Nelson and Blanchard and Quah decompositions, Trigo and BK are the trigonometric and the Baxter and King versions of a BP filter, UCbiv is a bivariate UC filter. Y is output, C consumption, I investment, H hours, RK the real rate, W the real wage, CapU capacity utilization, R the nominal rate, and π the inflation rate. Factor is the first principal component of the nine series. Statistics are computed averaging over 100 data replications. For C and I BNa and BQa refer to the comparison of filtered C/Y and I/Y with the data. In bold are cases where the reported statistic improves the best result presented in table 5.

Table A.3: Correlations, T=750

Variable	HPa	LDa	BPa	Hama	BNa	BQa	FODa	Trigo	BK	UCbiv
	Gap									
Y	0.46	0.35	0.36	0.53	0.03	0.03	0.15	0.24	0.25	-0.01
C	0.44	0.32	0.32	0.46	0.81	0.81	0.16	0.22	0.23	NaN
I	0.38	0.31	0.29	0.51	0.73	0.73	0.14	0.19	0.19	NaN
H	0.58	0.45	0.46	0.76	NaN	NaN	0.22	0.31	0.31	NaN
RK	0.5	0.35	0.39	0.44	NaN	NaN	0.16	0.29	0.29	NaN
W	0.39	0.25	0.26	0.4	NaN	NaN	0.07	0.15	0.16	NaN
capU	0.5	0.35	0.39	0.44	NaN	NaN	0.16	0.29	0.29	0.03
π	0.71	0.55	0.6	0.88	NaN	NaN	0.38	0.47	0.48	NaN
R	0.76	0.57	0.59	0.9	NaN	NaN	0.44	0.48	0.48	NaN
Factor	0.46	0.38	0.32	0.52	NaN	NaN	0.15	0.2	0.2	NaN
	Transitory									
Y	0	0	0	0	0	0	0	0	0	0.07
C	0	-0.02	0	-0.02	-0.05	-0.05	-0.01	0	0	NaN
I	0	0	0	-0.01	0.01	0.01	-0.01	0	0	NaN
H	0	0.01	0	0	NaN	NaN	0	0	0	NaN
RK	0	-0.01	0	0	NaN	NaN	-0.01	0	0	NaN
W	0	-0.01	0.01	-0.03	NaN	NaN	-0.01	0	0	NaN
CapU	0	-0.01	0	0	NaN	NaN	-0.01	0	0	-0.01
π	0.01	0.01	0.01	0.01	NaN	NaN	0	0	0	NaN
R	0	-0.01	0	-0.01	NaN	NaN	0	0	0	NaN
Factor	0	0	0	0	NaN	NaN	0	0	0	NaN

Notes: HPa is the modified Hodrick and Prescott filtering, FODa is the forth order differencing, LDa is 16 quarter differencing, BPa is the modified band pass filtering, Hama is modified local projection detrending, BNa and BQa are trivariate Beveridge and Nelson and Blanchard and Quah decompositions, Trigo and BK are the trigonometric and the Baxter and King versions of a BP filter, UCbiv is a bivariate UC filter. Y is output, C consumption, I investment, H hours, RK the real rate, W the real wage, CapU capacity utilization, R the nominal rate, and π the inflation rate. Factor is the first principal component of the nine series. Statistics are computed averaging over 100 data replications. For C and I BNa and BQa refer to the comparison of filtered C/Y and I/Y with the data. In bold are cases where the reported statistic improves the best result presented in table 6.

Table A.4: AR1 coefficient, T=750

Variable	True	HPa	LDa	BPa	Hama	BNa	BQa	FODa	Trigo	BK	UCbiv
	Gap										
Y	0.98	0.92	0.95	0.96	0.94	0.21	0.21	0.83	0.93	0.93	0.84
C	0.98	0.93	0.96	0.97	0.93	0.98	0.98	0.84	0.93	0.93	NaN
I	0.99	0.94	0.96	0.96	0.95	0.98	0.98	0.88	0.93	0.93	NaN
H	0.98	0.88	0.92	0.95	0.92	NaN	NaN	0.78	0.93	0.93	NaN
RK	0.97	0.93	0.96	**0.97**	0.93	NaN	NaN	0.84	0.94	0.94	NaN
W	0.98	0.94	0.97	0.97	0.92	NaN	NaN	0.84	0.94	0.94	NaN
CapU	0.97	0.93	0.96	**0.97**	0.93	NaN	NaN	0.84	0.94	0.94	0.78
π	0.94	0.86	0.9	**0.95**	0.91	NaN	NaN	0.77	0.92	0.92	NaN
R	0.83	0.67	0.75	0.94	0.79	NaN	NaN	0.57	0.91	0.91	NaN
Factor	0.98	0.92	0.96	0.97	0.93	NaN	NaN	0.82	0.93	0.93	NaN
	Transitory										
Y	0.97	0.92	0.95	0.96	0.93	0.2	0.2	0.82	0.93	0.93	0.96
C	0.99	0.94	0.97	0.97	0.93	0.98	0.98	0.85	0.93	0.93	NaN
I	0.98	0.94	0.96	0.96	0.95	0.98	0.98	0.87	0.93	0.93	NaN
H	0.96	0.87	0.92	0.96	0.92	NaN	NaN	0.77	0.93	0.93	NaN
RK	0.99	0.93	0.96	0.97	0.93	NaN	NaN	0.83	0.94	0.94	NaN
W	0.99	0.94	0.97	0.97	0.92	NaN	NaN	0.84	0.94	0.94	NaN
CapU	0.99	0.93	0.96	0.97	0.93	NaN	NaN	0.83	0.94	0.94	0.94
π	0.94	0.86	0.9	**0.95**	0.91	NaN	NaN	0.77	0.92	0.92	NaN
R	0.84	0.67	0.74	0.94	0.79	NaN	NaN	0.57	0.91	0.91	NaN
Factor	0.99	0.91	0.95	0.96	0.93	NaN	NaN	0.8	0.93	0.93	NaN

Notes: HPa is the modified Hodrick and Prescott filtering, FODa is the forth order differencing, LDa is 16 quarter differencing, BPa is the modified band pass filtering, Hama is modified local projection detrending, BNa and BQa are trivariate Beveridge and Nelson and Blanchard and Quah decompositions, Trigo and BK are the trigonometric and the Baxter and King versions of a BP filter, UCbiv is a bivariate UC filter. Y is output, C consumption, I investment, H hours, RK the real rate, W the real wage, CapU capacity utilization, R the nominal rate, and π the inflation rate. Factor is the first principal component of the nine series. Statistics are computed averaging over 100 data replications. For C and I BNa and BQa refer to the comparison of filtered C/Y and I/Y with the data. In bold are cases where the reported statistic improves the best result presented in table 7.

Table A.5: Variability, T=750

Variable	True	HPa	LDa	BPa	Hama	BNa	BQa	FODa	Trigo	BK	UCbiv
	Gap										
Y	23.62	9.55	33.07	7.64	**22.67**	1.55	1.55	8.87	3.5	3.45	276.56
C	28.77	5.55	21.42	4.27	14.67	13.32	13.32	4.85	1.81	1.78	NaN
I	69.37	**59.88**	195.72	50.47	127.94	94.69	94.69	55.95	24.4	24.01	NaN
H	6.43	2.63	8.1	2.13	5.36	NaN	NaN	3.02	1.17	1.15	NaN
RK	2.48	1.12	4.28	0.81	**2.84**	NaN	NaN	0.9	0.33	0.33	NaN
W	20.85	2.92	12	2.09	7.6	NaN	NaN	2.11	0.7	0.68	NaN
CapU	7.86	3.54	13.57	2.57	**8.99**	NaN	NaN	2.85	1.05	1.03	127.04
π	0.58	0.23	**0.65**	0.2	0.43	NaN	NaN	0.31	0.12	0.12	NaN
R	0.69	0.33	0.86	0.23	0.54	NaN	NaN	0.52	0.15	0.15	NaN
Factor	5.05	0.53	2	0.41	1.35	NaN	NaN	0.47	0.17	0.17	NaN
	Transitory										
Y	28.94	9.38	**32.94**	7.56	23.87	1.55	1.55	8.82	3.42	3.38	261.54
C	33.58	6.46	26.44	4.88	18.27	21.93	21.93	5.36	1.88	1.86	NaN
I	160.43	55.57	178.23	48	114.37	104.02	104.02	53.11	23.55	23.23	NaN
H	8.29	2.53	**7.89**	2.07	5.5	NaN	NaN	2.92	1.11	1.1	NaN
RK	6.32	1.03	3.87	0.78	2.54	NaN	NaN	0.85	0.31	0.31	NaN
W	20.43	3.1	13.2	2.24	8.49	NaN	NaN	2.22	0.71	0.7	NaN
CapU	20.02	3.26	12.25	2.49	8.04	NaN	NaN	2.68	0.99	0.97	89.91
π	0.57	0.22	**0.63**	0.19	0.41	NaN	NaN	0.3	0.11	0.11	NaN
R	0.69	0.32	0.84	0.23	0.53	NaN	NaN	0.51	0.15	0.15	NaN
Factor	4.52	0.41	1.5	0.32	1.03	NaN	NaN	0.38	0.14	0.14	NaN

Notes: HPa is the modified Hodrick and Prescott filtering, FODa is the forth order differencing, LDa is 16 quarter differencing, BPa is the modified band pass filtering, Hama is modified local projection detrending, BNa and BQa are trivariate Beveridge and Nelson and Blanchard and Quah decompositions, Trigo and BK are the trigonometric and the Baxter and King versions of a BP filter, UCbiv is a bivariate UC filter. Y is output, C consumption, I investment, H hours, RK the real rate, W the real wage, CapU capacity utilization, R the nominal rate, and π the inflation rate. Factor is the first principal component of the nine series. Statistics are computed averaging over 100 data replications. For C and I BNa and BQa refer to the comparison of filtered C/Y and I/Y with the data. In bold are cases where the reported statistic improves the best result presented in table 8.

Table A.6: Number of turning points, average durations, average amplitudes, T=750

Variable	True	HPa	LDa	BPa	Hama	BNa	BQa	FODa	Trigo	BK	UCbiv
	Output Gap										
Number of TP	125.87	123.44	124.31	128.28	119.83	109.11	109.11	141.49	129.37	129.74	129.2
DurE	5.67	5.77	5.75	5.55	5.99	6.68	6.68	5.03	5.52	5.5	4.93
DurR	5.71	5.83	5.77	5.58	6	6.49	6.49	5.06	5.52	5.5	6.27
AmpE	-2.49	-3.61	-5.08	-1.45	-4.95	-4.69	-4.69	-5.22	-1.43	-1.43	-21.68
AmpR	2.49	3.61	5.08	1.45	4.95	4.69	4.69	5.22	1.43	1.43	- 21.69
	Factor Gap										
Number of TP	120.07	120.69	121.06	128.56	118.84	NaN	NaN	138.45	129.98	130.44	NaN
DurE	6.01	5.93	5.9	5.57	6.05	NaN	NaN	5.16	5.49	5.49	NaN
DurR	5.92	**5.93**	5.96	5.53	6.05	NaN	NaN	5.16	5.49	5.46	NaN
AmpE	-1.08	-0.83	-1.16	-0.31	-1.25	NaN	NaN	-1.19	-0.3	-0.3	NaN
AmpR	1.08	0.83	1.16	0.31	1.25	NaN	NaN	1.19	0.3	0.3	NaN
	Transitory output										
Number of TP	125.57	123.94	124.84	128.87	119.01	109.32	109.32	142.5	130.56	130.63	127.76
DurE	5.66	5.79	5.8	5.53	6.09	6.56	6.56	5.03	5.46	5.44	5.5
DurR	5.75	**5.76**	5.68	5.55	5.97	6.55	6.55	5	5.47	5.49	5.81
AmpE	-3.4	**-3.61**	-5.1	-1.45	-5.17	-4.72	-4.72	-5.25	-1.41	-1.41	-12.25
AmpR	3.4	**3.61**	5.1	1.45	5.18	4.71	4.71	5.25	1.41	1.41	12.27
	Transitory factor										
Number of TP	121.2	120.91	120.96	129.39	118.07	NaN	NaN	140.58	131.44	131.04	NaN
DurE	5.9	5.97	5.98	5.52	6.12	NaN	NaN	5.07	5.44	5.45	NaN
DurR	5.92	5.91	5.89	5.52	6.03	NaN	NaN	5.11	5.43	5.44	NaN
AmpE	-0.98	-0.74	-1.05	-0.27	-1.07	NaN	NaN	-1.06	-0.27	-0.27	NaN
AmpR	0.98	0.74	1.05	0.27	1.08	NaN	NaN	1.06	0.27	0.27	NaN

Notes: HPa is the modified Hodrick and Prescott filtering, FODa is the forth order differencing, LDa is 16 quarter differencing, BPa is the modified band pass filtering, Hama is modified local projection detrending, BNa and BQa are trivariate Beveridge and Nelson and Blanchard and Quah decompositions, Trigo and BK are the trigonometric and the Baxter and King versions of a BP filter, UCbiv is a bivariate UC filter. Y is output, C consumption, I investment, H hours, RK the real rate, W the real wage, CapU capacity utilization, R the nominal rate, and π the inflation rate. Factor is the first principal component of the nine series. Statistics are computed averaging over 100 data replications. For C and I BNa and BQa refer to the comparison of filtered C/Y and I/Y with the data. In bold are cases where the reported statistic improves the best result presented in table 9.

Table A.7 Phillips curve and Okun law predictions, T=750

Variable	HPa	LDa	BPa	Hama	BNa	BQa	FODa	Trigo	BK	UCbiv
	Phillips curve prediction: Output Gap									
1 step	0.21	0.36	0.25	0.34	0.27	0.27	0.32	0.26	0.26	0.32
4 steps	1.48	1.81	1.88	1.74	1.67	1.67	1.69	1.68	1.67	1.6
	Phillips curve prediction: Transitory Output									
1 step	0.36	0.15	0.39	0.13	0.19	0.19	0.16	0.42	0.43	0.15
4 steps	1.03	0.54	3.21	0.5	0.6	0.6	0.62	2.87	2.89	0.73
	Phillips curve prediction: Output Gap									
1 step	0.55	0.21	4.67	0.21	0.24	0.24	0.23	3.75	3.79	4.91
4 steps	17.25	1.99	55.2	2.56	2.49	2.49	2.27	45.59	45.94	17.28
	Okun law prediction: Transitory Output									
1 step	0.32	0.31	0.32	0.31	0.33	0.33	0.31	0.31	0.31	0.34
4 steps	2.85	2.69	3.21	2.97	2.76	2.76	2.63	3.03	3.05	2.95

Notes: HPa is the modified Hodrick and Prescott filtering, FODa is the forth order differencing, LDa is 16 quarter differencing, BPa is the modified band pass filtering, Hama is modified local projection detrending, BNa and BQa are trivariate Beveridge and Nelson and Blanchard and Quah decompositions, Trigo and Bk are the trigonometric and the Baxter and King versions of a BP filter, UCbiv is a bivariate UC filter. The Phillips curve and the Okun law predictions are regression of the form ; $x_{t+m} = \alpha_0 + \alpha_1 x_t + \sum_{j=1}^{3} \beta_j y_{t-j}$ where y_{t-j} is the true gap (transitory) or the estimated one, $x_t = \pi_t$ or h_t and m=1,4. Reported is the difference in variance of the prediction error between each procedure and the true prediction error, averaged over 100 data replications. In bold cases where the reported statistic improves the best result presented in table 10.

Table A.8: Average MSE, T=150

Variable	POLY	HP	FOD	LD	BP	Wa	Ham	UC	BN	BQ	BW
	Gap										
Y	27.26	**26.32**	29.77	51.56	27.55	29.84	28.02	104.64	1146.1	140.71	20.06(*)
C	**31.97**	33.63	36.48	50.27	34.8	36.07	34.25	112.85	NaN	NaN	22.36(*)
I	111.95	**86**	87.82	250.93	89.22	117.93	107.82	159.06	NaN	NaN	90
H	5.51	6.58	7.93	8.9	7.02	5.69	5.27	89.77	**3.58**	**3.58**	1.08(*)
RK	2.63	**2.5**	3.02	6.64	2.69	3.54	3.19	2.71	NaN	NaN	4.76
W	19.78	21.09	22.65	30.86	21.79	22.57	21.86	**18.93**	NaN	NaN	11.04(*)
CapU	8.32	**7.92**	9.58	21.03	8.51	11.22	10.09	22.82	NaN	NaN	15.08
π	0.42	0.51	0.69	0.66	0.57	0.48	**0.37**	0.59	NaN	NaN	0.03(*)
R	**0.4**	0.5	0.79	0.76	0.63	0.54	0.35	0.6	NaN	NaN	0.04(*)
Factor	**4.01**	4.5	5.21	6.56	4.77	4.36	4.4	4.72	NaN	NaN	2.49(*)
	Transitory										
Y	47.3	37.73	**35.29**	77.94	36.87	49.42	48.31	87.87	1849.5	120.7	50.38
C	52.96	44.44	**43.08**	80.37	43.87	55.58	52.1	101.61	NaN	NaN	55.32
I	263.42	210.79	192.59	403.39	207.75	267.84	263.75	**172.93**	NaN	NaN	280.51
H	12.5	10.31	**9.61**	19.49	10.04	12.84	13.02	62.44	53.43	53.43	13.38
RK	9.09	7.52	7.31	12.45	7.44	9.07	8.54	**3.16**	NaN	NaN	9.59
W	33.57	28.47	28.05	48.54	28.21	35.16	31.96	**23.57**	NaN	NaN	36.51
CapU	28.82	23.83	23.16	39.46	23.58	28.74	27.06	**21.37**	NaN	NaN	30.38
π	0.88	0.77	**0.71**	1.3	0.75	0.86	0.91	0.79	NaN	NaN	0.93
R	1.12	0.99	0.99	1.68	0.91	1.06	1.19	**0.86**	NaN	NaN	1.2
Factor	6.25	4.74	4.39	10.55	**4.61**	6.53	6.14	7.91	NaN	NaN	6.89

Notes: POLY is polynomial detrending, HP is Hodrick and Prescott filtering, FOD is first order differencing, LD is 24 quarter differencing, UC is an unobservable component model, BP is band pass filtering, Ham is local projection detrending, BN and BQ are bivariate Beveridge and Nelson and Blanchard and Quah decompositions, BW the Butterworth filter. Y is output, C consumption, I investment, H hours, RK the real rate, W the real wage, CapU capacity utilization, R the nominal rate, and π the inflation rate. Factor is the first principal component of the nine series. The MSE is computed averaging over 100 data replications. In bold the best approach. A (*) in the last column indicates that the BW filter improves over or is comparable to the best approach.

Table A.9: Average real time MSE

Variable	POLY	HP	FOD	LD	BP	Wa	Ham	UC	BN	BQ	BW
	Gap										
Y	36.98	35.83	38.66	42.87	36.09	**26.59**	33.48	120.31	334.35	269.97	18.85(*)
C	45.82	44.24	46.91	40.11	44.31	**34.39**	44.88	129.96	NaN	NaN	21.2(*)
I	134.64	**110.61**	111.67	248.47	110.73	115.56	117.75	287.04	NaN	NaN	78.04(*)
H	9.46	8.98	10.28	9.09	9.05	6.39	5.98	82.42	**3.41**	**3.41**	1.04(*)
RK	3.83	**2.41**	2.69	8.21	2.44	3.89	3.5	7.61	NaN	NaN	5.24
W	24.24	22.62	23.18	27.01	22.66	**20.14**	23.12	29.28	NaN	NaN	10.19(*)
CapU	12.15	**7.62**	8.52	26	7.72	12.34	11.08	37.28	NaN	NaN	16.62
π	0.68	0.65	0.87	0.8	0.67	0.6	**0.49**	0.67	NaN	NaN	0.04(*)
R	0.68	0.65	0.93	0.88	0.74	0.64	**0.45**	0.78	NaN	NaN	0.05(*)
Factor	10.19	8.01	8.21	9.51	7.95	**7.43**	8.22	9.33	NaN	NaN	3.81(*)
	Transitory										
Y	45.68	39.3	**38.8**	76.63	39.31	49.81	49.64	122.66	459.47	217.52	52.81
C	53.27	44.02	**43.52**	89.39	44.09	59.56	52.25	127.17	NaN	NaN	56.19
I	231.54	187.57	**173.41**	377.49	187.57	250.4	241.36	279.46	NaN	NaN	244.74
H	11.27	10.02	9.98	16.49	**9.83**	11.2	12.54	61.70	61.36	61.36	11.8
RK	9.74	7.7	7.64	14.38	7.56	10.04	9.1	**5.41**	NaN	NaN	8.05
W	38.01	32.35	32.08	58.36	**32.32**	42.02	35.64	37.61	NaN	NaN	40.23
CapU	30.85	24.41	24.22	45.56	**23.96**	31.81	28.82	25.24	NaN	NaN	25.52
π	0.83	0.73	**0.69**	1.26	0.73	0.86	0.95	0.75	NaN	NaN	0.91
R	1.11	0.92	0.92	1.8	0.83	1.11	1.34	**0.75**	NaN	NaN	1.3
Factor	8.41	6.44	**6.38**	11.44	6.42	8.05	8.02	7.22	NaN	NaN	7.24

Notes: POLY is polynomial detrending, HP is Hodrick and Prescott filtering, FOD is first order differencing, LD is 24 quarter differencing, UC is an unobservable component model, BP is band pass filtering, Ham is local projection detrending, BN and BQ are bivariate Beveridge and Nelson and Blanchard and Quah decompositions, BW the Butterworth filter. Y is output, C consumption, I investment, H hours, RK the real rate, W the real wage, CapU capacity utilization, R the nominal rate, and π the inflation rate. Factor is the first principal component of the nine series. The MSE is computed averaging over 100 data replications. In bold the best approach. A (*) in the last column indicates that the BW filter improves over or is comparable to the best approach.

Table A.10: Contemporaneous correlations, T=150

Variable	POLY	HP	FOD	LD	BP	Wa	Ham	UC	BN	BQ	BW
	Gaps and Filtered variables										
Y	**0.59**	0.45	0.04	0.48	0.33	0.52	0.54	0.13	-0.58	0.15	0.74(*)
C	**0.58**	0.45	0.08	0.46	0.33	0.48	0.48	0.11	NaN	NaN	0.76(*)
I	**0.51**	0.37	0.04	0.45	0.27	0.49	0.51	0.29	NaN	NaN	0.66(*)
H	0.75	0.53	0.09	0.62	0.41	0.68	0.77	0.1	**0.93**	**0.93**	0.93(*)
RK	**0.6**	0.5	0.03	0.42	0.37	0.45	0.43	0.34	NaN	NaN	0.61(*)
W	**0.57**	0.4	0.02	0.44	0.26	0.45	0.39	0.13	NaN	NaN	0.74(*)
CapU	**0.6**	0.5	0.03	0.42	0.37	0.45	0.43	0.25	NaN	NaN	0.61(*)
π	0.87	0.74	0.22	0.65	0.6	0.72	**0.91**	0.58	NaN	NaN	1(*)
R	0.89	0.79	0.34	0.66	0.59	0.68	**0.93**	0.77	NaN	NaN	1(*)
Factor	**0.49**	0.34	0.01	0.38	0.24	0.41	0.39	0.32	NaN	NaN	0.7(*)
	Transitory and Filtered variables										
Y	-0.01	-0.01	-0.02	-0.02	-0.01	-0.03	-0.03	-0.02	-0.08	0.02	-0.03
C	0	0.01	-0.01	-0.04	0.02	-0.03	-0.02	-0.03	NaN	NaN	0
I	-0.03	0	-0.02	-0.02	-0.01	-0.01	-0.02	-0.03	NaN	NaN	0
H	-0.02	-0.01	-0.02	-0.02	-0.01	-0.02	-0.03	0	-0.02	-0.02	-0.02
RK	-0.04	-0.01	-0.01	-0.02	-0.01	-0.01	-0.01	0.02	NaN	NaN	-0.04
W	-0.02	0.01	-0.03	-0.06	0.01	-0.05	-0.04	0.02	NaN	NaN	-0.06
CapU	-0.04	-0.01	-0.01	-0.02	-0.01	-0.01	-0.01	0	NaN	NaN	-0.04
π	0.01	0.01	0	0	0.01	0.02	0.03	-0.01	NaN	NaN	-0.01
R	0.01	0.01	0	-0.01	0.01	-0.01	0.02	0.01	NaN	NaN	0
Factor	-0.04	-0.01	-0.01	-0.06	-0.01	-0.06	-0.04	-0.01	NaN	NaN	-0.04

Notes: POLY is polynomial detrending, HP is Hodrick and Prescott filtering, FOD is first order differencing, LD is 24 quarter differencing, UC is an unobservable component model, BP is band pass filtering, Ham is local projection detrending, BN and BQ are bivariate Beveridge and Nelson and Blanchard and Quah decompositions, BW the butterworth filter. Y is output, C consumption, I investment, H hours, RK the real rate, W the real wage, CapU capacity utilization, R the nominal rate, and π the inflation rate. Factor is the first principal component of the nine series. Correlations are computed averaging over 100 data replications.In bold the best approach. A (*) in the last column indicates that the BW filter improves over or is comparable to the best approach.

Table A.11: AR1 coefficient, T=150

Variable	True	POLY	HP	FOD	LD	BP	Wa	Ham	UC	BN	BQ	BW
	Gaps and filtered variables											
Y	0.96	0.94	0.81	0.2	**0.95**	0.93	0.98	0.89	1.00	0.97	0.79	0.96(*)
C	0.96	0.95	0.81	0.21	**0.96**	0.93	0.98	0.88	1.00	NaN	NaN	0.97
I	0.98	0.96	0.88	0.5	0.97	0.93	**0.98**	0.91	0.99	NaN	NaN	0.97
H	0.96	0.9	0.76	0.11	0.92	0.92	**0.97**	0.88	1.00	0.93	0.93	0.93
RK	0.95	**0.95**	0.81	0.19	0.96	0.93	0.98	0.88	0.92	NaN	NaN	0.98
W	0.95	**0.96**	0.8	0.1	0.97	0.93	0.99	0.86	0.93	NaN	NaN	0.98
CapU	0.95	**0.95**	0.81	0.19	0.96	0.93	0.98	0.88	0.92	NaN	NaN	0.98
π	0.9	0.87	0.76	0.16	**0.88**	0.92	0.96	0.86	0.77	NaN	NaN	0.9 (*)
R	0.76	0.69	0.53	-0.21	**0.73**	0.91	0.94	0.72	0.29	NaN	NaN	0.78(*)
Factor	0.97	0.95	0.81	0.19	**0.96**	0.93	**0.98**	0.88	0.91	NaN	NaN	0.98(*)
	Transitory and filtered variables											
Y	0.95	0.93	0.81	0.2	**0.95**	0.93	0.98	0.89	1.00	0.95	0.86	0.94
C	0.96	**0.95**	0.83	0.24	**0.97**	0.93	0.98	0.88	1.00	NaN	NaN	0.96(*)
I	0.96	0.95	0.88	0.48	**0.96**	0.93	0.98	0.9	0.99	NaN	NaN	0.96(*)
H	0.94	0.89	0.76	0.09	0.92	0.92	0.97	0.87	0.99	**0.93**	**0.93**	0.91
RK	0.97	0.95	0.81	0.18	**0.96**	0.93	**0.98**	0.88	0.93	NaN	NaN	0.95
W	0.97	0.95	0.81	0.11	**0.97**	0.93	0.99	0.87	0.92	NaN	NaN	0.96
CapU	0.97	0.95	0.81	0.18	**0.96**	0.93	**0.98**	0.88	0.92	NaN	NaN	0.95
π	0.9	0.86	0.76	0.13	**0.89**	0.92	0.96	0.86	0.77	NaN	NaN	0.88
R	0.75	0.69	0.52	-0.21	**0.73**	0.91	0.94	0.72	0.29	NaN	NaN	0.73(*)
Factor	0.97	0.93	0.81	0.17	**0.96**	0.93	**0.98**	0.88	0.91	NaN	NaN	0.95

Notes: POLY is polynomial detrending, HP is Hodrick and Prescott filtering, FOD is first order differencing, LD is 24 quarter differencing, UC is an unobservable component model, BP is band pass filtering, Ham is local projection detrending, BN and BQ are bivariate Beveridge and Nelson and Blanchard and Quah decompositions, BW the Butterworth filter. Y is output, C consumption, I investment, H hours, RK the real rate, W the real wage, CapU capacity utilization, R the nominal rate, and π the inflation rate. Factor is the first principal component of the nine series. The AR1 coefficient is computed averaging over 100 data replications. In bold the best approach. A (*) in the last column indicates that the BW filter improves over or is comparable to the best approach.

Table A.12: Variability, T=150

Variable	True	POLY	HP	FOD	LD	BP	Wa	Ham	UC	BN	BQ	BW
	Gaps and filtered variables											
Y	15.65	**14.85**	3.95	1.56	36.28	3.21	12.1	13.7	192.03	425.21	24.07	20.88
C	17.59	**9.23**	2.06	0.81	24.54	1.67	8.11	7.98	190.93	NaN	NaN	18.35(*)
I	44.67	87.39	**25.58**	6.85	205.24	22.25	70.02	80.12	242.9	NaN	NaN	93.09
H	4.66	3.4	1.34	0.65	7.72	1.08	2.58	3.58	186.35	**5.41**	**5.41**	3.54
RK	1.9	**1.92**	0.38	0.15	4.87	0.3	1.61	1.53	0.23	NaN	NaN	2.65
W	11.4	5.54	0.86	0.38	**14.81**	0.64	4.81	3.75	0.51	NaN	NaN	9.36
CapU	6.02	**6.07**	1.2	0.48	15.44	0.94	5.12	4.85	29.82	NaN	NaN	8.41
π	0.42	0.27	0.14	0.07	**0.54**	0.11	0.18	0.3	0.08	NaN	NaN	0.26
R	0.53	**0.38**	0.23	0.22	0.75	0.15	0.22	0.41	0.06	NaN	NaN	0.32
Factor	5.36	1.92	0.49	0.27	**4.82**	0.4	1.56	1.7	0.26	NaN	NaN	2.8
	Transitory and filtered variables											
Y	18.21	**13.18**	3.93	1.57	33.83	3.1	11.28	13.14	204.37	652.08	25.43	19.44(*)
C	17.43	**10.23**	2.28	0.87	27.95	1.75	9.1	8.86	207.63	NaN	NaN	15.46(*)
I	113.84	**70.9**	24.77	6.64	168.26	21.36	58.08	68.56	237.58	NaN	NaN	102.97(*)
H	5.68	3.39	1.31	0.65	**7.98**	1.04	2.61	3.55	169.43	6.93	6.93	4.72(*)
RK	3.95	1.62	0.36	0.14	**4.02**	0.29	1.32	1.31	0.23	NaN	NaN	2.2
W	10.97	5.44	0.89	0.38	**15.03**	0.66	4.85	3.83	0.45	NaN	NaN	8.79(*)
CapU	12.52	5.15	1.14	0.46	**12.75**	0.91	4.18	4.14	22.5	NaN	NaN	6.99
pi	0.38	0.25	0.13	0.06	**0.53**	0.1	0.17	0.27	0.08	NaN	NaN	0.32(*)
R	0.5	0.36	0.22	0.21	**0.71**	0.14	0.21	0.4	0.06	NaN	NaN	0.45(*)
Factor	4.53	1.82	0.54	0.27	**4.58**	0.43	1.51	1.72	4.38	NaN	NaN	2.64

Notes: POLY is polynomial detrending, HP is Hodrick and Prescott filtering, FOD is first order differencing, LD is 24 quarter differencing, UC is an unobservable component model, BP is band pass filtering, Ham is local projection detrending, BN and BQ are bivariate Beveridge and Nelson and Blanchard and Quah decompositions, BW the Butterworth filter. Y is output, C consumption, I investment, H hours, RK the real rate, W the real wage, CapU capacity utilization, R the nominal rate, and π the inflation rate. Factor is the first principal component of the nine series. The variability is computed averaging over 100 data replications. In bold the best approach. A (*) in the last column indicates that the BW filter improves over or is comparable to the best approach.

Table A.13: Number of turning points, average durations, average amplitudes, T=150

Variable	True	POLY	HP	FOD	LD	BP	Wa	Ham	UC	BN	BQ	BW
	Output Gap and filtered output											
Number of TP	19.97	20.44	20.56	17.9	**19.66**	20.36	21.63	19.47	20.07	22.12	17.18	21.08
DurE	5.94	**5.75**	**5.75**	6.91	6.19	5.64	5.34	6	5.64	5.33	6.81	5.61
DurR	5.75	5.61	5.57	6.35	5.66	5.51	5.24	**5.83**	5.64	5.31	6.9	5.38
AmpE	-2.47	-3.58	-3.55	-4.7	-5.12	-1.41	**-1.48**	-5.29	-1.83	-11.88	-5.9	-2.66(*)
AmpR	2.46	3.59	3.55	4.72	5.1	1.4	**1.47**	5.27	1.81	11.86	5.92	2.67(*)
	factor Gap and filtered factor											
Number of TP	19.37	19.08	**19.23**	17.73	19.71	20.23	18.5	21.48	21.68	NaN	NaN	19.89
DurE	6.06	6.01	**6.02**	6.77	6.15	5.66	6.32	5.37	5.24			5.8
DurR	5.92	6.13	**5.94**	6.55	5.79	5.54	6.26	5.25	5.69	NaN	NaN	5.84
AmpE	-1.53	-1.21	-1.19	**-1.55**	-1.68	-0.46	-1.82	-0.48	-0.46	NaN	NaN	-0.89
AmpR	1.53	1.21	1.19	**1.55**	1.68	0.46	1.83	0.48	0.46	NaN	NaN	0.89
	Output Transitory and filtered output											
Number of TP	21.05	19.99	20.11	17.67	19.6	20.47	22.04	19.05	**20.44**	22.77	17.87	19.85
DurE	5.47	5.66	5.63	6.85	5.98	**5.48**	5.08	6.34	5.82	4.7	6.73	5.72
DurR	5.57	6.01	5.84	6.38	5.85	**5.61**	5.18	6.16	**5.61**	5.63	6.8	6.06
AmpE	-3.42	-3.6	**-3.56**	-4.75	-5.18	-1.42	-1.47	-5.25	-1.84	-15.51	-5.46	-3.61
AmpR	3.42	3.61	**3.57**	4.74	5.17	1.42	1.46	5.24	1.81	15.46	5.48	3.61
	Factor transitory and filtered factor											
Number of TP	19.97	19.07	19.1	17.85	**19.59**	20.33	19.02	21.68	21.84	NaN	NaN	18.98
DurE	5.91	6.14	6.14	6.62	**6.04**	5.58	6.05	5.28	5.24	NaN	NaN	6.2
DurR	5.73	6.02	5.97	6.44	**5.71**	5.61	6.26	5.26	5.39	NaN	NaN	6.03
AmpE	-1.34	**-1.27**	-1.26	-1.66	-1.8	-0.49	-1.87	-0.5	-1.41	NaN	NaN	-1.28(*)
AmpR	1.34	**1.27**	1.26	1.67	1.81	0.5	1.88	0.5	1.41	NaN	NaN	1.28(*)

Notes: POLY is polynomial detrending, HP is Hodrick and Prescott filtering, FOD is first order differencing, LD is 24 quarter differencing, UC is an unobservable component model, BP is band pass filtering, Ham is local projection detrending, BN and BQ are bivariate Beveridge and Nelson and Blanchard and Quah decompositions, BW is the Butterworth filter. Y is output and Factor is the first principal component of the nine series. DurE and DurR are the durations of expansions and recessions; AmpE and AmpR the amplitude of expansions and recessions. Statistics are computed averaging over 100 data replications. In bold the best approach. A (*) in the last column indicates that the BW filter improves over or is comparable to the best approach.

Table A.14: Phillips curve and Okun law predictions, T=150

Step ahead	POLY	HP	FOD	LD	BP	Wa	Ham	UC	BN	BQ	BW
	Phillips curve prediction: Output Gap										
1 step	0.46	0.6	0.43	0.48	0.64	0.44	0.49	**0.36**	0.42	0.52	0.35(*)
4 steps	2.49	2.76	2.06	2.75	3.92	2.5	2.55	**2.08**	2.8	2.93	1.65(*)
	Phillips curve prediction: Transitory Output										
1step	0.58	0.64	0.53	**0.45**	0.53	0.5	0.55	0.61	0.49	0.52	0.57
4step	2.52	2.33	2.22	**2.16**	3.26	2.39	2.24	2.6	2.28	2.21	2.45
	Okun law prediction: Output Gap										
1step	0.79	1.21	0.75	0.71	3.26	**0.68**	0.69	0.7	0.84	0.69	0.84
4step	12.92	26.59	**7.96**	8.98	39.17	8.92	8.25	8.31	14.3	11.56	9.36
	Okun law prediction: transitory Output										
1step	1.46	1.44	1.53	1.39	1.45	**1.38**	1.53	1.45	1.51	1.45	1.48
4step	13.94	13.69	**12.89**	13.54	15.62	15.03	15.63	16.69	15.78	13.5	15.48

Notes: POLY is polynomial detrending, HP is Hodrick and Prescott filtering, FOD is first order differencing, LD is 24 quarter differencing, UC is an unobservable component model, BP is band pass filtering, Ham is local projection detrending, BN and BQ are bivariate Beveridge and Nelson and Blanchard and Quah decompositions, BW is the Butterworth filter. The Phillips curve and the Okun law predictions are regression of the form ; $x_{t+m} = \alpha_0 + \alpha_1 x_t + \sum_{j=1}^{3} \beta_j y_{t-j}$ where y_{t-j} is the true gap (transitory) or the estimated one, $x_t = \pi_t$ or h_t and m=1,4. Reported the difference in variance of the prediction error between each procedure and the true prediction error, averaged over 100 data replications. In bold the best approach. A (*) in the last column indicates that the BW filter improves over or is comparable to the best approach.

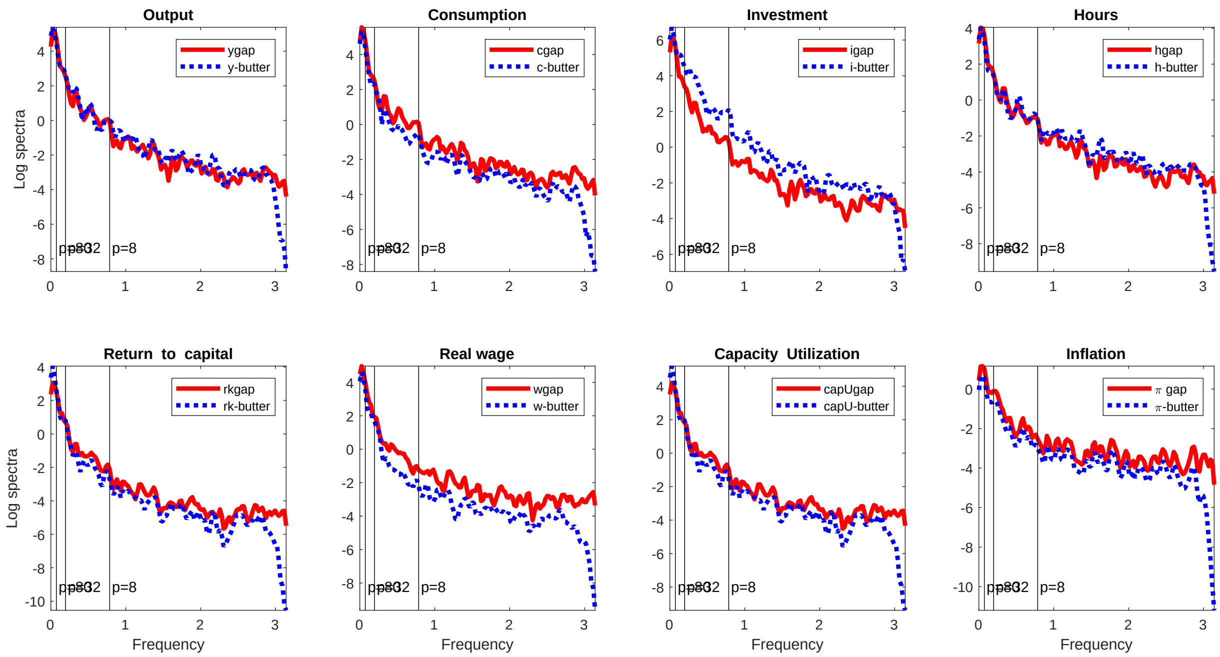

Figure A.1: Log spectra of gaps and of Butterworth filtered data

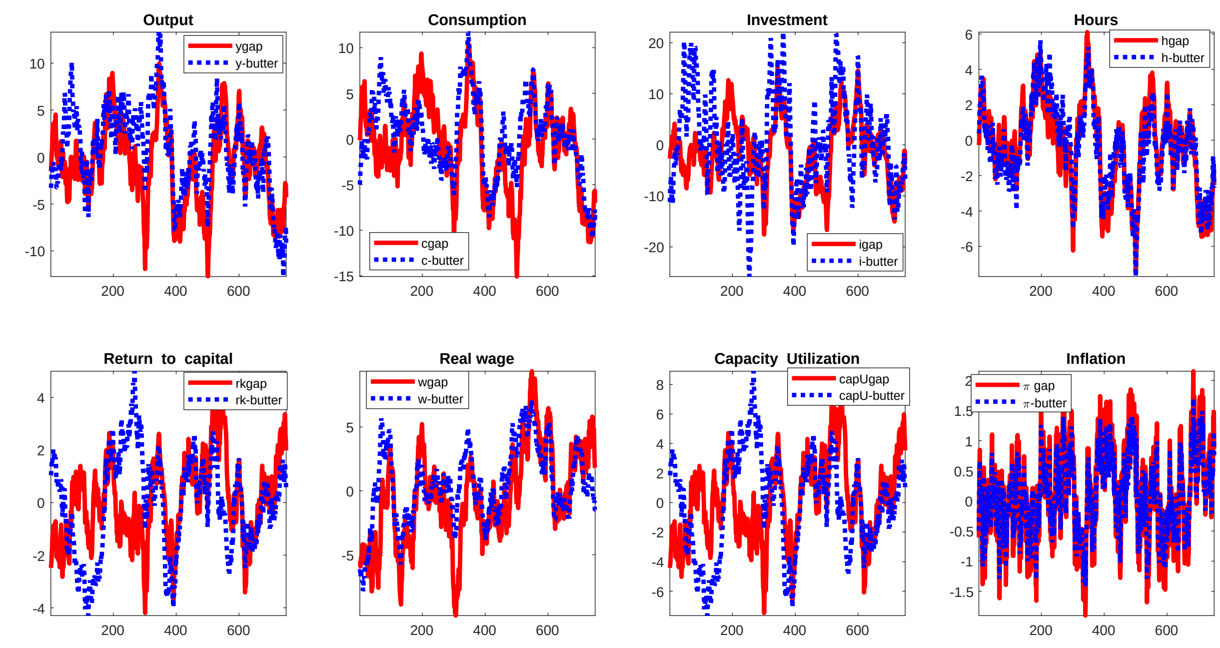

Figure A.2: Time series of gaps and of Butterworth filtered data

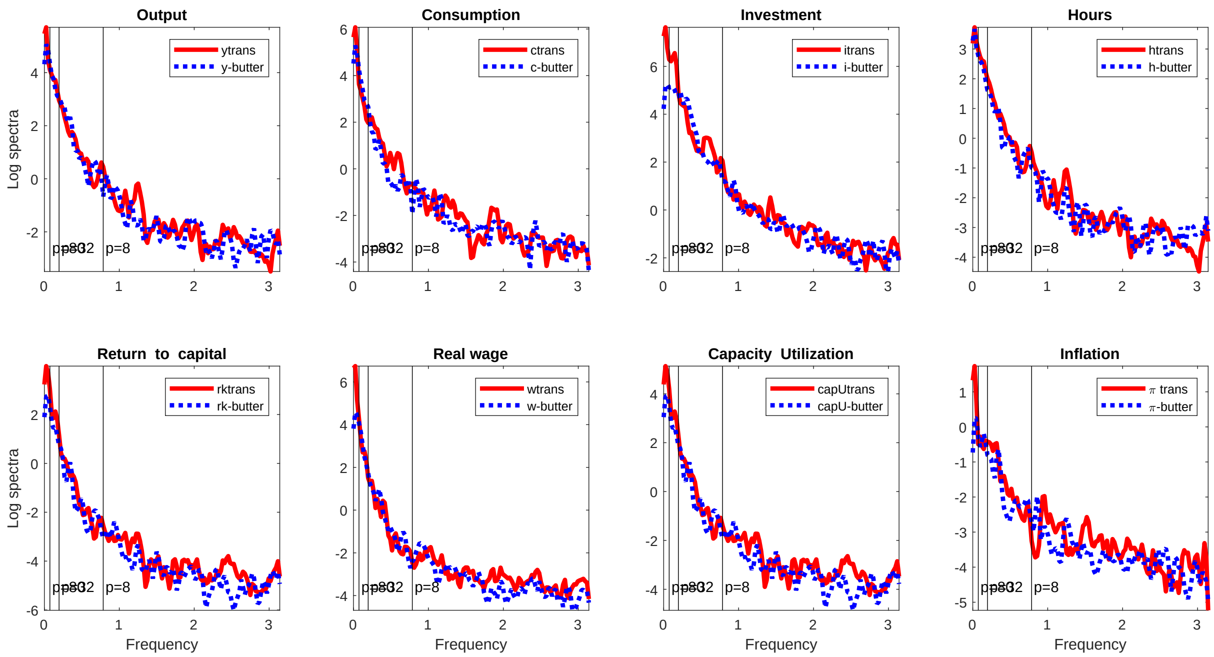

Figure A.3: Log spectra of transitory components and of Butterworth filtered data

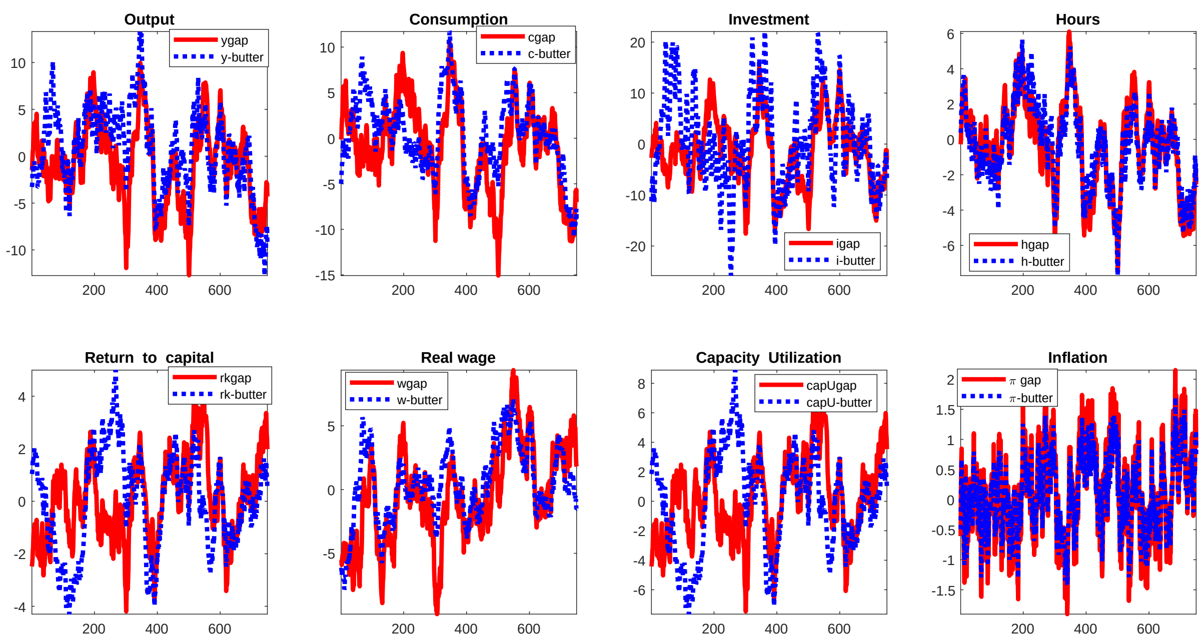

Figure A.4: Time series of transitory data and of Butterworth filtered data