IMF Working Paper

What Is in Your Output Gap?

Unified Framework & Decomposition into Observables

Michal Andrle

INTERNATIONAL MONETARY FUND

IMF Working Paper

Research department

What Is in Your Output Gap? Unified Framework & Decomposition into Observables

Prepared by Michal Andrle[1]

Authorized for distribution by Benjamin Hunt

Oc{ 2013

Abstract

This paper discusses several popular methods to estimate the 'output gap'. It provides a unified, natural concept for the analysis, and demonstrates how to decompose the output gap into contributions of observed data on output, inflation, unemployment, and other variables. A simple bar-chart of contributing factors, in the case of multi-variable methods, sharpens the intuition behind the estimates and ultimately shows 'what is in your output gap.' The paper demonstrates how to interpret effects of data revisions and new data releases for output gap estimates (news effects) and how to obtain more insight into real-time properties of estimators.

Contents   Page

Appendices

Tables

Figures

I. INTRODUCTION

This paper discusses several popular methods to estimate the 'output gap' using available macroeconomic data, provides a unified, natural concept for the analysis, and demonstrates how to decompose the output gap into contributions of observed data on output, inflation, unemployment, and other variables. A simple bar-chart of contributing factors, in the case of multi-variable methods, sharpens the intuition behind the estimates and ultimately shows *'what is in your output gap.'* The paper demonstrates how to interpret effects of data revisions and new data releases for output gap estimates (news effects) and how to obtain more insight into real-time properties of estimators.

The unifying approach is the *theory of linear filters.* I demonstrate that most methods for output gap estimation can be represented as a moving average of observed data – as a linear filter. Such a representation provides insight into which variables contribute to the estimate of the unobserved variables, and at what frequencies. Knowing this provides a better understanding of the estimate and of its revision properties.

Which output gap estimation approaches can be analyzed as linear filters? As demonstrated below, these range from (i) univariate or multivariate statistical filters to (ii) simple multivariate-filters with some economic theory (Phillips curve/IS curve), including also a (ii) production function approach, and even (iv) state-of-the-art DSGE (dynamic stochastic general equilibrium) models with tight theory restrictions.[1]

One thing that this paper does not intend to discuss is what method for output gap estimation is the most sensible, or the optimal one. It needs to be understood that the concept of the output gap is meaningful only when properly defined, before being embedded into an empirical model. Nevertheless, the importance of the output gap as a concept in economic policy requires a thorough understanding of model-based estimates, when used for monetary or fiscal policy.

Example To frame the discussion below, consider a very simplified, stylised example that illustrates a decomposition into observables. The multivariate model of the 'extended Hodrick-Prescott' filter, as in principle suggested in the paper by Laxton and Tetlow (1992), features a simple aggregate demand determination of the output gap, x_t, and a backward-looking Phillips

---

[1]Calculations and analysis analogous to this paper have been used since 2007 together with the Czech National Bank DSGE core projection model, see Andrle and others (2009b).

curve to determine the deviation of inflation from the target, $\hat{\pi}_t$. Aggregate output, y_t, is composed of the output gap and the potential output, τ. For simplicity, I assume that potential output growth follows a driftless random walk.

The *state-space form* of the simple model is as follows:

$$y_t \;=\; x_t + \tau_t \tag{1}$$

$$\tau_t - \tau_{t-1} \;=\; \tau_{t-1} - \tau_{t-2} + \varepsilon_t^\tau \tag{2}$$

$$x_t \;=\; \rho x_{t-1} - \kappa \hat{\pi}_t + \varepsilon_t^x \tag{3}$$

$$\hat{\pi}_t \;=\; \lambda \hat{\pi}_{t-1} + \theta x_t + \varepsilon_t^\pi. \tag{4}$$

Given observed data for $\{y_t, \hat{\pi}_t\}$ the problem is to estimate the decomposition of output into its unobserved components, $\{x_t, \tau_t\}$. For a state-space form the estimates are easily available using the well-known algorithms for the Kalman filter and smoother.

Equivalently, the implied *penalized least squares* formulation is

$$\min_{\{\tau\}_0^T} \sum_{t=0}^{T} \frac{1}{\sigma_x^2}[\varepsilon_t^x]^2 + \frac{1}{\sigma_\tau^2}[\varepsilon_t^\tau]^2 + \frac{1}{\sigma_\tau^2}[\varepsilon_t^\pi]^2. \tag{5}$$

The goal of the analysis is, however, the implied moving average representation of the model, i.e. linear filter representation, given as follows:

$$x_{t|\infty} = A(L)y_t + B(L)\hat{\pi} = \sum_{i=-\infty}^{\infty} A_i y_{t+i} + \sum_{i=-\infty}^{\infty} B_i \hat{\pi}_{t+i}, \tag{6}$$

where $A(L), B(L)$ are two-sided linear filters and L is a linear operator, such that $L^j x_t := x_{t-j}$. The weights of these filters are completely determined by the structure of the economic model and its parameters.[2]

Under rather general conditions, estimates using the state-space form (1)–(4), penalized least-squares (5), and filter specification (6) are equivalent. Simply put, the unifying approach makes use of equivalence between the methods of (i) *Penalized Least-Squares*, (ii) *Wiener-Kolmogorov filtering* and (iii) *Kalman filter* associated with these model representations. See e.g. Gomez (1999) for a lucid discussion. Each of the three approaches (least-squares, Wiener-Kolmogorov or Kalman filters) has its benefits and limitations.

---

[2] see Appendix C for details

The key ingredient for *obtaining flexible decompositions of unobservables in terms of observed data,* and for the analysis of revision properties, is the linear filter representation (6). For the model above, expression (6) clearly indicates what portion of the output gap is identified due to observations on output, y_t, and due to deviation of inflation from the inflation target, $\hat{\pi}$. Expression (6) assumes a doubly-infinite sample and is a starting point for the analysis of finite sample implementation. Extending the sample size will lead to revised estimates, to a well-known 'end-point' problems or 'news effect'. The example provides the general result that is analysed in greater detail in the rest of the paper.

Despite the extensive literature on output gap/potential output estimation using multivariate filters, beginning with the aforementioned contribution of Laxton and Tetlow (1992), to my knowledge, analysis in terms of decomposition into observables and filter analysis of output gaps has not been presented before. Also, the idea of putting most of the estimation method into a common framework of linear filters has not been explored explicitly before and is a novel approach for comparing estimates obtained by different methods.

The roadmap of the paper is as follows. The Introduction motivated a filter representation for output gap estimation and its decomposition into observables, showing 'what is in your output gap'. The Methods section demonstrates how to formulate output gap estimation methods as linear filters, decompose the output gap into observables, and demonstrates the benefits of the filter representation for understanding the revision properties of real-time estimates. The subsequent Application section gives a simple extension of the Hodrick-Prescott filter, which proves useful for multivariate models, and illustrates the main ideas of the paper using a semi-structural and fully structural DSGE model for output gap estimation.

II. Methods

This section focuses on formulating output gap estimation methods as linear filters and decomposing it into observables. State-space models, univariate statistical filters, structural vector auto-regressions, and a production function approach are considered. Subsequently, the revision properties of real-time estimates of output gap are discussed exploiting a filter representation.

Before delving into details of various estimation methods and calculations, it is crucial to understand that the goal of obtaining a linear filter representation serves practical purposes. It allows analysts to understand the weighting scheme behind the estimate, chart the output

gap as a function of underlying data and quantify the sources of revised estimates when the sample is extended or revised. A formalized analysis lowers the burden of excessive experimenting.

The benefits of the filter representation are numerous and this section focuses on a small subset only. In particular, an explicit frequency domain analysis is omitted. It is important though that the knowledge of the filter transfer function allows to design structural economic models as optimal filters.

Output gap estimates using methods discussed below can be expressed as a (multivariate) linear filter representation,

$$x_{t|T} = \sum_{i=t_0}^{T} w_{1,t} y_{1,t-i} + \cdots + \sum_{i=t_0}^{T} w_{n,t} y_{n,t-i} = \sum_{k=1}^{n} \xi_k, \tag{7}$$

where a particular unobservable variable $x_{t|T}$ —the output gap in this case–is expressed as a weighted average of an observed sample, y_t, finite or infinite, when $t_0 \to -\infty, T \to \infty$. y_i. Here $i = 1 \ldots n$ and thus $x_{t|T}$ is decomposed into contribution of n factors, ξ_j.

A multivariate version of the moving average, in case of multiple unobservables, with a doubly-infinite sample, takes the form

$$\mathbf{X}_t = \sum_{i=-\infty}^{\infty} \mathbf{W}_i \mathbf{Y}_{t-i} = \mathbf{B}(L) \mathbf{Y}_t, \tag{8}$$

which is a starting point for a theoretical analysis. Practical calculations, however, are not restricted to an infinite amount of data, nor to time-invariant weights.

The model-implied multivariate moving average, afilter $\mathbf{B}(L) = \sum_i \mathbf{B}_i z^i$, can be analyzed in time- or frequency-domain, as is the case for univariate filters, in terms of their gain, coherence or phase-shifts between variables and, the overall frequency-response function characteristics.

The subsequent subsections provide a detailed treatment of output gap estimation methods and their conversion to a filter representation analogous to (7) or (8), which answers the question, 'what is in the output gap'. Although the estimation methods are different, the principle is always the same, which allows for a direct comparison of the results.

A. Formulating Potential Output Estimates as Linear Filters

1. State-Space Forms – Semi-structural and DSGE Models

Formulating potential output estimators in a state-space form as linear filters is surprisingly simple. This is the case since the celebrated Kalman filter, see e.g. Kalman (1960) or Whittle (1983), originates from the Wiener filtering theory and deals with an important special class of stochastic processes.

The state-space formulation of the potential output estimation became very popular, partly due to its flexibility, see e.g. Kuttner (1994), Laubach and Williams (2003), inter alios. The state-space model is easy to formulate and modify, easily handles missing data or non-stationary dynamics, and is a natural representation for linearized recursive dynamic economic models.

The missing piece in the literature on the multivariate model analysis is the explicit acknowledgement and use of the fact that the Kalman filter and smoother[3] are actually just that – *filters.* As demonstrated above, an explicit linear filter formulation is useful for obtaining a decomposition into observables. The formulation of the state-space model as a linear filter, the filter weights, and a very practical implementation of decomposition into observables for state-space models, follow.

Filter representation   For the purpose of analysis, it is assumed that the model takes the following state-space form:

$$\mathbf{X}_t = \mathbf{T}\mathbf{X}_{t-1} + \mathbf{G}\varepsilon_t \tag{9}$$

$$\mathbf{Y}_t = \mathbf{Z}\mathbf{X}_t + \mathbf{H}\varepsilon_t,. \tag{10}$$

Here $\varepsilon \sim N(\mathbf{0}, \Sigma_\varepsilon)$, $\Sigma_\varepsilon = \mathbf{I}$ and thus structural shocks are not correlated with measurement error shocks, with no loss of generality. The vector of transition, or state, variables is denoted by $\mathbf{X}_t$, whereas observed variables are denoted by $\mathbf{Y}_t$. By imposing a restriction that $\mathbf{G}\Sigma_\varepsilon\mathbf{H}' = \mathbf{0}$, it is guaranteed that measurement errors and structural shocks are uncorrelated.

The state-space model (9)–(10) can be used to estimate the unobserved states and shocks $\{\mathbf{X}_t, \varepsilon_t\}$ from the available observables $\{\mathbf{Y}_t\}$. The output gap is one of the elements in $\mathbf{X}_t$. I

---

[3] The Kalman filter is a one-sided, causal estimate of the state x_t based on information up to the period $[t_0, \ldots, t]$. The Kalman smoother is a two-sided, non-causal filter that uses all available information to estimate the state $x_{t|T}$ based on $[t_0, \ldots, T]$.

shall focus mainly on the 'smoothing' case, i.e. when estimates of $\mathbf{X}_t$ are based on observations available for $t = 0, \ldots T$, using the notation $\mathbb{E}[\mathbf{X}_t | \mathbf{Y}_T, \ldots, \mathbf{Y}_0] = \mathbf{X}_{t|T}$.

In the case of multivariate models with multiple observables, the possibilities for analysis are richer than in the case of univariate models. If meaningful, the exploration of impulse-response and transfer functions provides insights regarding the model properties, together with the popular *structural shock decomposition.* In other words, if the shocks have some structural interpretations, one can express the observed data (and unobserved states) as cumulative effects of past structural shocks,

$$\mathbf{Y}_t = [\mathbf{Z}\mathbf{A}(L) + \mathbf{H}]\hat{\varepsilon}_t, \qquad \mathbf{X}_{t|\infty} = \mathbf{A}(L)\hat{\varepsilon}_t = \sum_{j=0}^{\infty} \mathbf{T}^j \mathbf{G}\hat{\varepsilon}_{t-j}, \tag{11}$$

where $\mathbf{A}(L) = (\mathbf{I} - \mathbf{T}L)^{-1}\mathbf{G}$ and $\hat{\varepsilon}_t$, $\mathbf{X}_{t|\infty}$ denote a mean squared (Kalman smoother) estimate of the structural shocks, ε, and state variables $\mathbf{X}$.[4] Expression (11) is frequently used in a DSGE analysis for storytelling and interpretation of macroeconomic data.

Now is the time to reverse the logic and ask the question, "What observed data drive each particular unobserved structural shock and state variables?" That is the purpose of this paper – to draw a closer attention to a presentation of the unobserved state estimates as a function of the observed data. In the case of a doubly-infinite sample the model can be expressed as

$$\mathbf{X}_{t|\infty} = \mathbf{\Omega}(L)\mathbf{Y}_t = \sum_{i=-\infty}^{\infty} \mathbf{\Omega}_i \mathbf{Y}_{t+i}. \tag{12}$$

In real world applications, where the sample is always finite, the optimal finite sample implementation of (12) leads –or at least should lead– to a multivariate linear filter with *time varying weights,*

$$\mathbf{X}_{t|T} = \sum_{\tau=t_0}^{T} \mathbf{\Omega}_{\tau,t}\mathbf{Y}_\tau + \mathbf{\Omega}_{0,t}\mathbf{X}_0. \tag{13}$$

Here the weight sequence varies with every time period t. That is because the Kalman smoother carries out an optimal mean-square approximation of the infinite filter $\mathbf{\Omega}(z) = \sum_{i=-\infty}^{\infty} \mathbf{\Omega}_j z^j$ with a finite length filter $\mathbf{\Omega}_t(z) = \sum_{i=t_0}^{T} \mathbf{\Omega}_{j,t} z^j$, so as to minimize the distance $\|\mathbf{X}_{t|\infty} - \mathbf{X}_{t|T}\|^2$. It operates under the assumption that the model (9)–(10) is the data generating process for the data. More on this in a discussion of real-time properties of output gap estimates below.

---

[4]Semi-infinite sample size is assumed for simplicity only, finite sample analysis is trivial.

To decompose the output gap into observables and to analyze data revisions using (13), it remains to either calculate the time varying weights of the filter or to reformulate the problem such that one can avoid the computation of weights. Luckily, both options are readily available and their description follows.

Weights of the Filter   Weights of the filter $\boldsymbol{\Omega}(L)$ are not data-dependent and are a function of the model specification only.[5] In the case of the doubly-infinite sample, the Wiener-Kolmogorov formula, see Whittle (1983), implies that $\boldsymbol{\Omega}(z) = \boldsymbol{\Gamma}_{XY}(z)\boldsymbol{\Gamma}_Y(z)^{-1}$. Here $\boldsymbol{\Gamma}_Y(z)$ and $\boldsymbol{\Gamma}_{XY}(z)$ stand for auto-covariance and cross-covariance generating functions of the model:

$$\boldsymbol{\Gamma_{XY}}(z) = (\mathbf{I} - \mathbf{T}z)^{-1}\mathbf{G}\boldsymbol{\Sigma}_\varepsilon\mathbf{G}'(\mathbf{I} - \mathbf{T}'z^{-1})^{-1}\mathbf{Z}' \tag{14}$$

$$\boldsymbol{\Gamma_{YY}}(z) = \mathbf{Z}(\mathbf{I} - \mathbf{T}z)^{-1}\mathbf{G}\boldsymbol{\Sigma}_\varepsilon\mathbf{G}'(\mathbf{I} - \mathbf{T}'z^{-1})^{-1}\mathbf{Z}' + \mathbf{H}\boldsymbol{\Sigma}_\varepsilon\mathbf{H}'. \tag{15}$$

The transfer function of the model, $\boldsymbol{\Omega}(z)$, is the key ingredient for a frequency-domain analysis of the filter. In the Applications section below, I explore transfer function gains for a semi-structural model of the output gap, which indicates the most relevant frequencies of observed time series for the output gap estimate. The core of the analysis is in time domain and thus the weights are needed.

For all but very simple and small models, it is difficult to get an analytical description of the weights in (12) using the transfer function of the model. The weights can, however, always be obtained numerically. An inefficient, but operational way would also be to compute the inverse Fourier transform of (12). Koopman and Harvey (2003) provide a recursive way to calculate time-varying weights in (13) for general state-space models and a lucid paper by Gomez (2006) provides time-domain formulas to calculate weights in (12).

In particular, Gomez (2006) shows that for the model (9)–(10) the weights $\boldsymbol{\Omega}_j$, adjusted to model notation above, follow as

$$\boldsymbol{\Omega}_0 = \mathbf{P}(\mathbf{Z}'\boldsymbol{\Sigma}^{-1} - \mathbf{L}\mathbf{R}_{|\infty}\mathbf{K}) \tag{16}$$

$$\boldsymbol{\Omega}_j = (\mathbf{I} - \mathbf{P}\mathbf{R}_{|\infty})\mathbf{L}^{-j-1}\mathbf{K} \qquad j < 0 \tag{17}$$

$$\boldsymbol{\Omega}_j = \mathbf{P}\mathbf{L}'^j(\mathbf{Z}'\boldsymbol{\Sigma}^{-1} - \mathbf{L}'\mathbf{R}_{|\infty}\mathbf{K}) \qquad j > 0, \tag{18}$$

where $\mathbf{L} \equiv \mathbf{T} - \mathbf{KZ}$, $\mathbf{K}$ denotes the steady-state Kalman gain and $\mathbf{P}$ is the steady-state solution for the state error covariance given by a standard discrete time algebraic Ricatti equa-

---

[5]When the parameters of the model are estimated using the data, the weights become, indirectly, a function of a particular dataset.

tion (DARE) associated with a steady-state solution of the Kalman filter. $\mathbf{R}_{|\infty}$ is a solution to the Lyapounov equation, $\mathbf{R}_{|\infty} = \mathbf{L}'\mathbf{R}_{|\infty}\mathbf{L} + \mathbf{Z}(\mathbf{ZPZ}' + \mathbf{H}\Sigma_\varepsilon\mathbf{H}')^{-1}\mathbf{Z}'$, associated with the steady-state Kalman smoother solution. $\mathbf{R}_{|\infty}$ is the steady-state variance of the process, $\mathbf{r}_{t|\infty}$, in the backward recursion, $\mathbf{X}_{t|\infty} = \mathbf{X}_{t|t-1} + \mathbf{Pr}_{t|\infty}$, where in finite-data smoothing $\mathbf{r}_{t-1}$ is a weighted sum of those innovations (prediction errors) coming after period $t-1$. Finally, $\Sigma = \mathbf{Z}(\mathbf{ZPZ}' + \mathbf{H}\Sigma_\varepsilon\mathbf{H}')$.[6]

The relationship between the time-invariant weights of the filter and the time-varying weights, used in the case of the finite sample implementation by the Kalman smoother, is unique and is discussed below in the section devoted to real-time properties and news effects. The intuition behind the re-weighting is simple, though. The Kalman smoother implementation implicitly provides optimal linear forecasts and backcasts of the sample and applies the convergent time-invariant weights.

Practical Implementation of the Decomposition into Observables   To compute the observable decomposition one can always calculate the weights using Koopman and Harvey (2003) recursions and implement the moving average calculations. That requires calculating and storing large objects, pre-programmed tools, or a little bit of advanced knowledge of state-space modeling. Sometimes, time constraints might prevent analysts from using these tools. Having a shortcut is thus beneficial.

A particularly simple and accurate way is to view the Kalman smoother as a linear function of multiple inputs, denoted by $\mathbf{X} = \mathcal{F}(\mathbf{Y})$, where $\mathbf{X}$ and $\mathbf{Y}$ are $(T \times n)$ and $(T \times m)$ matrices. The great thing about that function is that it is linear. For stationary processes, the Kalman smoother provides the least squares estimates of the form $\mathbf{X} = \widehat{\Omega}\mathbf{Y}$, where $\widehat{\Omega}$ is the matrix of time-varying filter coefficients. Trivially, for two different sets of observables, $\{\mathbf{Y}^A, \mathbf{Y}^B\}$, one obtains $\mathbf{X}^A - \mathbf{X}^B = \widehat{\Omega}(\mathbf{Y}^A - \mathbf{Y}^B)$. By appropriate non-overlapping grouping of differences in inputs, one can easily obtain the effects of the change of measurements on all estimated unobservables and carry out the decomposition analysis. This method works for any model with two different sets of observables and a common initial state, unless the change in the initial state is treated as well. There is no need to know the values in $\widehat{\Omega}$; the whole decomposition of the deviations in the two estimates can be obtained by successive runs of the Kalman smoother with different inputs.

---

[6]The question of non-stationary models is more difficult, but for detectable and stabilizable models the Kalman filter/smoother converges to steady-state since, despite the infinite variance of states, the distance of the state to the estimate is stationary with finite variance. In the case of Wiener-Kolmogorov filter, the formulas apply if interpreted as a limit of minimimum mean square estimator, see Gomez (2006) or Bell (1984).

The two data input structures can bear many forms. The only requirement is that the *structure of observations* in both datasets must be identical. This setup is very feasible in the case of data revision analysis and exploration of the effects of new observations, as shown below. The counter-factual dataset can also take form of a steady-state (balanced growth path), unconditional forecasts, etc. – depending on the goals of the analysis. Decomposition into observables in this paper is consistent with a dataset featuring missing observations and direct observations implementation of linear restrictions, often used for imposing the 'expert judgement'.

2. Univariate filters – Band-Pass, Hodrick-Prescott, etc.

True, decomposition into observables in a univariate setting is not an issue, as the only variable that enters the output gap estimation is the output itself. The analysis of their real-time properties and news effects, however, follows the same principles as in the case of multivariate filters. A proper understanding of frequently used univariate filters, such as band-pass filters or the Hodrick-Prescott filter, is important, as these often form parts of multivariate models.

Univariate filters are specified either directly in terms of their weights in time domain, directly in terms of their transfer function in frequency domain, as a state-space model, or as a penalized least squares problem – e.g. the Hodrick-Prescott filter or exponential smoothing filter. Being specified in any of these ways, they have a time domain filter representation as

$$x_t = F(L)y_t = \sum_{i=-\infty}^{\infty} w_i y_{t+i}, \tag{19}$$

where w_i are the weights of the filter. This fact is well know and is restated just for clarity and completeness. Univariate filters are usually discussed in terms of their spectral properties, implied by $F(z)$, but the weights of the filter are sometimes discussed as well, see Harvey and Trimbur (2008), among others.

Most contributions to the literature focused on designing or testing the univariate filters. They are concerned with (i) approximation of the ideal band-pass filter or (ii) revision properties of the filters for increasing sample size. The ideal band-pass filter with a perfectly rectangular gain function is often considered as a natural benchmark to judge univariate statistical filters against in terms of their 'sharpness' – i.e. leakage, strength of the Gibbs effect, or ease of finite-sample implementation.

The class of linear filters is large, including a variety of bandpass and highpass filters. Apart from the Hodrick-Prescot/Leser filter, Butterworth filters analyzed in Gomez (2001), rational square wave filters suggested by Pollock (2000), or even multiresolution Haar scaling filters fit the representation (19).

Given a possibly infinite filter $F(z)$, the revision properties depend on the quality of it's finite sample approximation, which is crucially dependent on the data generating process of the data, see Christiano and Fitzgerald (2003) or Schleicher (2003). Discussion of the revision properties and optimal finite sample implementation is discussed below, since univariate filters are often part of semi-structural multivariate models or production function approach estimates of the output gap.

3. Structural VARs

Structural VARs have a natural moving average, linear filter representation. It seems therefore very desirable to express the potential output and the output gap as a linear combination of data inputs. A thorough analysis of the contributions of observed series and the SVAR estimates frequency transfer function is crucial as these often are outliers in comparison with other methods, see McNellis and Bagsic (2007), Cayen and van Norden (2005) or Scott (2000), among others. Although SVAR models may seem to be used less frequently for output gap estimation, they certainly belong in the toolbox of many central banks and applied economists.

Assume that an estimated reduced form VAR model of order p is available, that is,

$$\mathbf{A}(L)\mathbf{Y}_t = \left(\mathbf{I} - \sum_{j=1}^{p} \mathbf{A}_j\right)\mathbf{Y}_t = \varepsilon, \qquad \mathbb{E}[\varepsilon\varepsilon^T] = \Sigma, \tag{20}$$

where residuals, ε_t (reduced-form shocks) are linked to 'structural' shocks η_t via an invertible transformation $\varepsilon_t = \mathbf{Q}\eta_t$. I assume that the dimension of $\mathbf{Y}_t$ is n. The identification often imposes long-run restrictions following Blanchard and Quah (1989) to tell apart transitory and permanent component of output, see e.g. Claus (1999). The structural VAR model is then expressed as $\mathbf{Y}_t = \mathbf{B}(L)\mathbf{Q}\mathbf{Q}^{-1}\varepsilon_t = \mathbf{S}(L)\eta_t$.

Assume that the j-th component of the data vector, $\mathbf{Y}_t$, is the GDP growth, $\mathbf{Y}_{j,t} = \Delta y_t$, then

$$\Delta y_t = \mathbf{S}_{11}(L)\eta_{1,t} + \mathbf{S}_{12}(L)\eta_{2,t} + \cdots + \mathbf{S}_{1n}(L)\eta_{n,t} \tag{21}$$

and the output gap, x_t, is the part of the GDP not affected by permanent shocks

$$x_t = \mathbf{S}_{12}(L)\eta_{2,t} + \cdots + \mathbf{S}_{1n}(L)\eta_{n,t}, \tag{22}$$

assuming $\sum_{k=0}^{\infty}\mathbf{S}_{12}(k)\eta_{2,t-k} + \cdots + \sum_{k=0}^{\infty}\mathbf{S}_{1n}(k)\eta_{n,t-k} = 0$. The structural shocks are estimated from the reduced form VAR residuals using $\eta_t = \mathbf{Q}^{-1}\varepsilon_t = \mathbf{Q}^{-1}\mathbf{A}(L)\mathbf{Y}_t$. One can thus recover the estimated output gap as a function of observations.

The identification scheme itself can be very case-specific, yet it is clear that for SVAR estimates of the output gap their concurrent estimates and final estimates coincide, unless an extended sample is used for paramter re-estimation. Investigating the spectral properties of $\mathbf{S}(z)$ is advisable, since ex-ante it is not clear which frequencies of observed series are used for estimation and SVAR estimates often stand out as outliers. The decomposition of the output gap into observables can be done using the expressions above, where the output gap is a function of structural shocks, which themselves are a function of observed data.

4. Production Function Approach

Even a production function (PF) approach can often be expressed as a filtering scheme. The real-time revision properties then crucially depend on the filter representation, as in the case of other methods. Many practitioners are, perhaps, aware of the production function approach being very much dependent on the underlying filters used in various steps of the method. This section provides an explicit formulation of the production function output gap estimate as a linear filter, along with its structure and decomposition into observables.

Assume that the value added is produced using the Cobb-Douglas production function. Denoting the logarithms of individual variables by lower-case letters, one gets

$$y_t = a_t + (1-\alpha)k_t + \alpha l_t. \tag{23}$$

Here a_t is the 'Solow residual' and k_t and l_t denote the actual levels of the capital stock and hours worked in the economy. It is common that a *trend total factor productivity*, a_t^*, is identified from the Solow residual using some smoothing procedure, often a variant of the Hodrick-Prescott or an other symmetric moving average filter. Denoting the smoothing filter as $A(L)$, it is clear that

$$a_t^* = A(L)a_t = A(L)y_t - (1-\alpha)A(L)k_t - \alpha A(L)l_t. \tag{24}$$

The next step is usually a determination of an 'equilibrium' or a trend level of hours-worked, or employment, and of the capital stock. I will consider only the estimate of the equilibrium employment, which is often cast as a filtering problem for the NAIRU, frequently in terms of inflation, capacity utilization, or other variables. Importantly, the sub-problem is most often a linear filter.

With only little loss of generality, I assume that the equilibrium employment is given by the trend component of the employment, obtained using the univariate filter as $l_t^* = E(L)l_t$. The output gap, x_t, can then be expressed then as

$$
\begin{aligned}
x_t &= y_t - y_t^* = y_t - (a_t^* + (1-\alpha)k_t^* + \alpha l_t^*) && (25) \\
&= (1 - A(L))y_t - (1-\alpha)(1 - A(L))k_t - \alpha(E(L) - A(L))l_t && (26) \\
&= \sum_{i=-\infty}^{\infty} w_{y,i} y_{t+i} + \sum_{i=-\infty}^{\infty} w_{l,i} l_{t+i} + \sum_{i=-\infty}^{\infty} w_{k,i} k_{t+i}, && (27)
\end{aligned}
$$

which is a version of a multivariate linear filter with three observables, or signals.[7] It is interesting that in this simple case, if the trend component of the Solow residual and of the employment are obtained using the same procedure, then $E(L) = A(L)$ and the contribution of observed employment data gets eliminated. Further, given the fact that the capital stock usually has little variance at business cycle frequencies–as a slow moving variable– its 'gaps' tend to be small. The smaller they become, the more the production function approach to output gap estimation approaches a simple univariate filter estimation of the output gap.[8]

In the case where $A(L)$ and $K(L)$ are transfer functions of the Hodrick-Prescott filter, which is quite usual, the production function approach results tend to be quite similar to HP filter estimates and suffer from most of the problems usually associated with the HP filter approach. See Epstein and Macciarelli (2010) as an example.

When the production function approach estimate can feasibly be expressed as a linear function of its inputs (e.g. output, labor, capital stock, etc.), providing such a decomposition is highly desirable. A finite-sample version of (25) is easy to obtain as long the process is linear. The production function estimates are often decomposed into the contributions of the

---

[7]In the case of an optimal finite sample implementation of the filter, it will be time varying, e.g. $A(z) = A(z)_t$. A decomposition into observables in a finite sample is equally simple.

[8] One can consider more involved procedures, but the principle remains the same. I can assume that the employment is determined by a working-age population, participation rate and a employment rate as $l_t = pop_t + pr_t + er_t$. If an 'equilibrium' levels of the employment rate, that is $(1 - nairu_t)$ and the participation rate are determined by a time invariant filter I can express the equilibrium employment as $l_t^* = pop_t + P(L)pr_t + E(L)er_t$, and I can proceed by substituting these expressions to the production function as in the simpler case.

total factor productivity, equilibrium employment, and capital stock, which are not all directly observable.

Incorporating any non-causal filter into the production function approach calculations impacts its real-time, revision properties. The linear filter analysis for revision properties and news effects thus relates also to the production function approach to potential output estimation.

B. Analyzing Revision Properties – Data Revisions and News Effects

Revision properties of output gap estimators are often among the crucial criteria for the evaluation of a particular method. Policymakers need accurate information for their decisions and the revisions of output gap estimates increase the uncertainty associated with policy implementation. Output gap estimates get revised due to (i) *historical data revisions* and (ii) *new information* – new data that becomes available and affects interpretation of the past estimates. Revisions of the output gap may not be a bad thing, per se, but excessively unreliable real-time estimates may lead to large policy errors or render the concept of the output gap irrelevant.

There are many contributions to the literature pointing out the real-time unreliability of many output gap estimation methods and offering various remedies to the problem, see e.g. Cayen and van Norden (2005) or Orphanides and van Norden (2002). The contribution of this section is both a conceptual and a practical one. Conceptually, it is crucial to thoroughly understand the sources of revisions and real-time unreliability. Here the linear filter framework is the most natural approach. From a practical point of view, a decomposition of the revisions into contributions of the new data is a useful analytical result, which allows researchers to asses the informativeness of available observations.

The linear filter representation allows an analysis of revisions in a very tractable way for all estimators considered. Recall that in the case of a proper finite sample implementation of filters, the method of penalized least squares, Kalman filter/smoother, or Wiener-Kolmogorov filtering yield equivalent results.

The present analysis can be used both for (i) *data revisions* and (i) *news effects* – arrival of new information. It should not come as a surprise that the key concept for the analysis is

$$\mathbf{X}_{t|T} = \sum_{j=t_0}^{T} \mathbf{\Omega}_{j|t} \mathbf{Y}_j, \tag{28}$$

a finite sample implementation of $\boldsymbol{\Omega}(z)$. Data revisions are discussed first, followed by a discussion of news effect, where the characteristics of the process $\mathbf{Y}_t$ are crucial for optimal finite sample implementation of $\boldsymbol{\Omega}(z)$.

1. Decomposing Effects of Data Revision

By data revisions, only the revision of past data releases should be understood; most often these concern GDP data. The treatment of data revisions is very simple as long as the filtering problem is applied to the same sample and number of observed data series – revised and old, say $\mathbf{Y}^A, \mathbf{Y}^B$. The revision and its decomposition into factors is then given by

$$\mathbb{R}_{t|T} = \mathbf{X}_{t|T}^A - \mathbf{X}_{t|T}^B = \sum_{\tau=t_0}^{T} \boldsymbol{\Omega}_{t|\tau}[\mathbf{Y}_\tau^A - \mathbf{Y}_\tau^B]. \tag{29}$$

Again, it is quite useful to think of (29) in a stacked form. For the Kalman filter, which is a solution to the least square problem, one has that $\mathbf{X} = \widehat{\boldsymbol{\Omega}}\mathbf{Y}$ and thus trivially $\mathbf{X}^A - \mathbf{X}^B = \widehat{\boldsymbol{\Omega}}(\mathbf{Y}^A - \mathbf{Y}^B)$. The only requirement is to keep the structure of $\widehat{\boldsymbol{\Omega}}$ fixed, for which an identical structure of observations (cross-section and time) is needed. This stacked, matrix representation of the filter is also useful for investigating news effects, treatment of missing variables and filter tunes (linear constraints) after a simple modification, see below.

The data revision decomposition is useful not just for the output gap estimates but also for understanding the revised estimates of technology, preference, and other structural shocks in a forecasting framework based on a DSGE model, see Andrle and others (2009b). For instance, the interpretation of inflationary pressures in the economy changes when the domestic absorption is revised, in the context of unchanged estimates of the CPI inflation.

2. News Effects and End-Point Bias

Implementation of a doubly-infinite, non-causal filter $\boldsymbol{\Omega}(z)$ is problematic, since all economic applications feature finite samples. Newly available data lead to some revision of past estimates of the output gap or other unobserved variables. Even in the case of an optimal finite sample approximation of the infinite filter, increasing the sample size leads to revisions.

All non-causal filters suffer the from finite-sample problem and saying that state-space models do not would not be correct. State-space models using the Kalman smoother are no excep-

tion to the rule. The statement by Proietti and Musso (2007) that their state-space signal extraction techniques 'do not suffer from what is often referred to as end-of-sample bias' is thus incorrect.

Still, the double-infinite sample size formulation of the problem is the best starting point for the analysis of news effects, restated here for convenience:

$$\mathbf{X}_{t|\infty} = \sum_{i=-\infty}^{\infty} \mathbf{\Omega}_i \mathbf{Y}_{t+i}. \tag{30}$$

The revision due to availability of new observations after period T, chronologically only, is then obviously

$$\mathbb{N}_{t|T} = \sum_{j=T}^{\infty} \mathbf{\Omega}_j \left(\mathbf{Y}_j - \mathbf{Y}_{j|T} \right), \tag{31}$$

which is just a weighted average of prediction errors, conditioned on the information set to period T, given the constant weights filter. See Pierce (1980) for a discussion of revision variance in time series models. Population revision variance can be computed given the filter and the data-generating process for the data, determining the prediction errors.

Intuitively, (i) the smaller the weight on future observations and (ii) the better the predictions of the future values, the smaller the revision variance. The discussion below conditions on a model as given and does not put forth advice on how to design a better filter with different weights.

All output gap estimation methods considered in the paper adopt a solution to the infinite-sample problem, either an explicit or an implicit one. Implicitly, all provide forecasts and backcasts for the actual sample, if needed. A simple *truncation of the filter weights* can be interpreted as zero-mean forecasts, which would be suitable only for an uncorrelated zero-mean stationary process. The optimal finite sample implementation of the filter is a solution to the following approximation problem:

$$\min_{\widehat{\mathbf{\Omega}}_{t,j}, j \in [t_0, \ldots, T]} \|\mathbf{X}_{t|\infty} - \mathbf{X}_{t|T}\|^2 \tag{32}$$

$$= \| \sum_{j=-\infty}^{-\infty} \mathbf{\Omega}_j \mathbf{Y}_j - \sum_{j=-t_0}^{-T} \widehat{\mathbf{\Omega}}_{j|t} \mathbf{Y}_j \|^2 \tag{33}$$

$$= \int_{-\pi}^{\pi} \|\mathbf{\Omega}(e^{-i\omega}) - \widehat{\mathbf{\Omega}}_t(e^{-i\omega})\|^2 \mathbf{S_Y}(\omega) d\omega, \tag{34}$$

see Koopmans (1974), Christiano and Fitzgerald (2003), or Schleicher (2003) for details. The solution delivers a sequence of filters $\widehat{\boldsymbol{\Omega}}_t(z)$ that, based on position in the sample, t, are the weights of the time invariant filter $\boldsymbol{\Omega}_t(z)$, adjusted by a factor derived from the auto-covariance function of the data.[9]

Importantly, the solution to a problem in (32) is equivalent to the one where the time invariant infinite filter is applied on the available sample padded with forecasts and backcasts by the j-step ahead, with j chosen such that the filter converges. The forecast is uniquely pinned down by the auto-covariance function of the data generating process. *Hence, a heuristic solution adopted by practitioners is actually the optimal one.*[10]

State-space models with the Kalman smoother and Penalized Least-Squares formulations provide implicitly the optimal finite sample approximation to $\boldsymbol{\Omega}(z)$ by re-weighting the time-invariant weights. This is equivalent to providing backcasts and forecasts and padding the sample. The crucial point is that the forecasts are based exclusively on the covariance generating function of the underlying model, which may not always represent well the covariance structure of the data generating process. The mismatch between the model and the data results in poor forecasting properties. When the weights of the filter are spread out to many periods, poor forecasting properties translate into revision variance and the so called end-point bias.

It is often the case that the filter puts a large weight on observations at the end of the sample. This is simply a consequence of the fact that for most models their forecasting formula puts larger weight on the most recent observations. When a doubly-infinite filter is applied to a padded sample, the end sample observations are implicitly counted many times due to the chain rule of projections, receiving effectively a larger weight.

Example   Consider the Hodrick-Prescott/Leser filter as an example. The filter is often mentioned for its poor revision properties, or its large 'end-point' bias. Unless one provides a factorization of the filter representation, like Kaiser and Maravall (1999), the filter takes many

---

[9]It is a projection problem, matching the auto-covariance of the $\mathbf{X}_{t|T}$ and $\mathbf{X}_{t|\infty}$ as closely as possible. For stationary processes, the Toeplitz structure of the autocovariance generating matrix allows for an efficient recursive implementation. The system of equations is not specified in full, as the paper focuses on the equivalence with the forecast and backcasts.

[10] Sometimes there are ways how to make the implementation more robust. For instance, Kaiser and Maravall (1999) factor the filter $\boldsymbol{\Omega}(z) = \mathbf{A}(z)\mathbf{A}(z^{-1})$ and use the Burnman-Wilson algorithm to implement a two-pass estimate of the HP filter, which requires only four periods of back/fore-casts to implement the infinite-order filter, using an ARIMA model for time series at hand. The same principle is the element of X12-ARIMA seasonal adjustment procedure, for instance. By lowering the number of prediction the process can be simplified and robust.

periods to converge – more than 20 data points on both sides, see Fig. 1. Viewing the HP/Leser filter as a desirable filter, a smooth transition low-pass filter, the optimal finite sample implementation then suggests re-weighting the filter weights by the auto-covariance function of the data or, equivalently, extending the sample with best linear backcasts and forecasts. Viewing the HP/Leser filter as a model, it is clear that output is assumed to be an ARIMA(0,2,2) model, where the output gap is uncorrelated white noise and the potential output growth follows a random walk. Such a model is highly implausible based on economic theory and econometric analysis to be a data generating process for any country's data. Yet, this is exactly the model that a state-space model (via the Kalman smoother) and the penalized least-squares formulation would use, yielding equivalent results.

Practical News Effect Decomposition for State-Space Models   New observations create 'news effects' only if they carry some new information, information not predictable from the past data. The representation (31) gives a simple and practical way of calculating and decomposing news effects into components of newly observed data.

The unavailable (or missing) data estimates are simply expected values conditional on the original information set. Padding (filling in) the sample data with these estimates does not change anything, since the information set is identical and there is no new information.

The problem of different sample sizes is easy to convert into a problem of identical sample sizes by padding the data with a model-based forecast and using (29) to carry out the decomposition simply by successive runs of the Kalman smoother.[11] The easiest way to see that 'padding' the data with conditional expectations or projections does not change the estimates is to realize the structure of the Kalman filter updating step: $\mathbf{X}_{t+1|t+1} = \mathbf{T}\mathbf{X}_{t|t} + \mathbf{K}(\mathbf{Y}_{t+1} - \mathbf{Y}_{t+1|t})$. In the case of data padded by forecast, the prediction error (information) is zero. The information sets of the original and padded data sets are identical.

This simple and practical approach enables the analyst to investigate a judgement-free forecast of the model and contrast it to actual data. The prediction error is then distributed into the revision of the past unobserved shocks, see e.g. Andrle and others (2009b) for examples using

---

[11] The implementation is simple, requires very little coding and allows the analyst to use a standard, existing Kalman filter routine. In comparison with computing weights explicitly, the approach is also usually faster and it is easy to code the decompositions for flexible grouping of variables, etc. In the case of non-stationary models, the situation is a little bit more involved, depending on the treatment of initial conditions, though the main principles introduced above hold. Further, using explicit time-varying weights, as in Koopman and Harvey (2003), works in every case when the Kalman smoother is applicable.

a DSGE model. The decomposition easily accounts for judgement imposed on the filter using a dummy-observations approach.

III. APPLICATIONS – DECOMPOSITION INTO OBSERVABLES & NEWS EFFECTS

This sections demonstrates applications of methods discussed in the first part of the paper. A simple extension of the Hodrick-Prescott/Leser filter is followed by an illustration of how to decompose the output gap into observables using semi-structural and structural DSGE models. Both applications thus use a state-space representation of the model.

A. Variants of Hodrick-Prescott/Leser Filter and Local Linear Trend Models

The Hodrick-Prescott/Leser filter, see Leser (1961) and Hodrick and Prescott (1997), is an undeniably popular method to estimate the output gap. Most economists either love it or hate it. Due to its important role in the applied work and in the development of many *multivariate* models, the I discuss the filter in a little bit more detail, despite its univariate nature. However, the focus will be mostly on things not dealt with in the literature and relevant for issues analyzed in the paper – most importantly an assumption of steady-state growth of output.[12]

(a) Hodrick-Prescott Filter/Leser   An often used specification of the Hodrick-Prescott filter[13] is a *penalized least squares* (PLS) form

$$\min_{\{\bar{y}_t\}_{t=-\infty}^{\infty}} \sum_{t=-\infty}^{\infty} (y_t - \bar{y}_t)^2 + \sum_{t=-\infty}^{\infty} \lambda \left[(\bar{y}_t - \bar{y}_{t-1}) - (\bar{y}_{t-1} - \bar{y}_{t-2}) \right]^2. \tag{35}$$

It is easy to see, e.g. King and Rebelo (1993), that the doubly-infinite sample model (35) implies a reduced form ARIMA(0,2,2) model for y_t.

---

[12]This paper only scratches the surface of all properties of the HP filter, see Kaiser and Maravall (2001) for many details from a frequency domain and filtering point of view.

[13]See the original paper by Leser (1961) for exactly the same idea. Ideas of variants of the filter have been around in the engineering community since 1940s.

The output gap estimate, x_t, can be then formulated as a linear time-invariant filter with a transfer function $C(L)$ and weights $w_{x,k}$, given by

$$x_t = C(L)y_t = \frac{\lambda(1-L)^2(1-L^{-1})^2}{1+\lambda(1-L)^2(1-L^{-1})}y_t = \sum_{k=-\infty}^{\infty} w_{x,k}y_{t-k} \qquad (36)$$

recall L denotes a 'lag operator', $Ly_t = y_{t-1}$. For details, see King and Rebelo (1993), inter alios.

In terms of unobserved components (UC) models, as also mentioned in Hodrick and Prescott (1997), the infinite-sample version of the Hodrick-Prescott filter can be rewritten intuitively as

$$
\begin{align}
y_t &= \bar{y}_t + x_t \tag{37}\\
x_t &= \varepsilon_t^x \qquad \varepsilon_t^x \sim N(0,\sigma_x^2) \tag{38}\\
\bar{y}_t - \bar{y}_{t-1} &= \bar{y}_{t-1} - \bar{y}_{t-2} + \varepsilon_t^g \qquad \varepsilon_t^g \sim N(0,\sigma_g^2), \tag{39}
\end{align}
$$

which clearly provides a model-based interpretation for the HP filter. The output gap is assumed to be a non-persistent random noise and the potential output growth is assumed to follow a random walk. As it is well known, $\sqrt{\lambda} = \sigma_g/\sigma_x$ is the signal-to-noise ratio.

State-space representation of the HP filter or its modifications and extensions can easily be written in a *stationary form,* where only the growth rates of the output are observed. One simply defines $\Delta y_t = \Delta x_t + g_t$, where $g_t = g_{t-1} + \varepsilon_t^g$ is coupled with a simple identity $x_t - x_{t-1} = \Delta x_t$. A stationary state-space form is easily initialized with the unconditional mean and variance of the model, avoiding the need to deal with many ways to initialize non-stationary models, using variants of a diffuse Kalman filter/smoother. The weights of the HP filter that operates on growth rates of GDP can be obtained using the integration filter transfer function to obtain $x_t = [C(L)/(1-L)]\Delta y_t$ in terms of (36).

(b) Modified Hodrick-Prescott Filter   A simple but useful modification of the HP filter is the incorporation of more realistic processes for the output gap and for the potential output. The literature is rich in these extensions, see e.g. Proietti (2009).

I will consider only a simple extension useful for better understanding the frequent treatment of the potential output in semi-structural models used for policy analysis, e.g. multivariate-filters of Benes and N'Diaye (2004) and Benes and others (2010) or semi-structural forward-

looking models of Carabenciov and others (2008) or Andrle and others (2009a). This representation is a common building block of more complex multivariate filters.[14]

The least complex model assumes that the output gap is a simple AR(1) stationary process and that the potential output is subject to (i) *level* and (ii) *growth rate* shocks. Importantly, *potential output growth is a mean-reverting,* persistent process – not a random-walk.

For many economies the assumption of mean reverting potential output growth is quite a plausible one and it is this feature of the model that stands behind its *improved revision properties.* This is an important aspect of the estimates, though most statistical or econometric literature, e.g. Proietti (2009), does not work with explicit steady-states. Steady-states are neither used in SVAR literature nor early literature on multivariate filters, e.g. Laxton and Tetlow (1992) or Conway and Hunt (1997), but are present in Benes and N'Diaye (2004), for instance.

The model, then, is as follows:

$$y_t = \bar{y}_t + x_t \tag{40}$$

$$x_t = \rho_x x_{t-1} + \varepsilon_t^x \qquad \varepsilon_t^x \sim N(0, \sigma_x) \tag{41}$$

$$\bar{y}_t = \bar{y}_{t-1} + g_t + \varepsilon_t^{\bar{y}} \qquad \varepsilon_t^x \sim N(0, \sigma_{\bar{y}}) \tag{42}$$

$$g_t = \rho_g g_{t-1} + (1 - \rho_g) g_{ss} + \varepsilon_t^g \qquad \varepsilon_t^x \sim N(0, \sigma_g). \tag{43}$$

Given this data-generating process for the GDP of a particular country, it is obviously possible –if of interest– to design a parametrization of a modified Hodrick-Prescott filter that keeps its gain function as close to HP as possible, but lowers revision variance.

(c) Example: US output gap and revision properties   As a simple example, I parametrize the modified HP as $\rho_1 = 0.70$, $\rho_g = 0.95$, $\sigma_{\bar{y}} = 0$, $\sigma_x = 1/(1 - \rho_1)$ and $\sigma_g = \sqrt{(1/\lambda)} \times [1/(1 - \rho_g)]$, $\lambda = 1600$ and apply this simple heuristic model to US output gap with sample 1967:1–2010:3. The steady-state growth of potential output is assumed to be 2 %. Fig. 1 demonstrates the difference between the output gaps estimated and the time invariant weights implied by these two filters. There cannot be any claim of optimality of this filter; far from that, it is only

---

[14] I work mainly with a state-space representation but, as explained above, penalized least squares, state-space, and linear filters methods are equivalent. The early literature on multivariate filters for output gap estimation, e.g. Laxton and Tetlow (1992), Conway and Hunt (1997) or de Brouwer (1998) somehow seems to contrast penalized least squares problems and 'unobserved component' as different methods, comparing often markedly different model specification, e.g. white noise versus AR(2) process for the output gap.

a demonstration. Revision properties are judged by the standard deviation of the difference of the final output gap estimate versus the real time estimates. For the standard HP filter, the standard deviation of the revision process is 1.489, whereas for the modified filter it is 0.858, which is essentially only 57.6% of the revision standard deviation. In this case, both the filter weights and the forecasting properties contribute to the result.

As is well known, the standard Hodrick-Prescott/Leser filter, without any priors or modifications, is an extremely infeasible model for real-time estimation of the output gap or any cyclical features of the data, due to its very poor revision properties. This is well known, but the reason is often poorly understood.

Figure 1. HP filter vs. Modified HP filter – estimate & weights
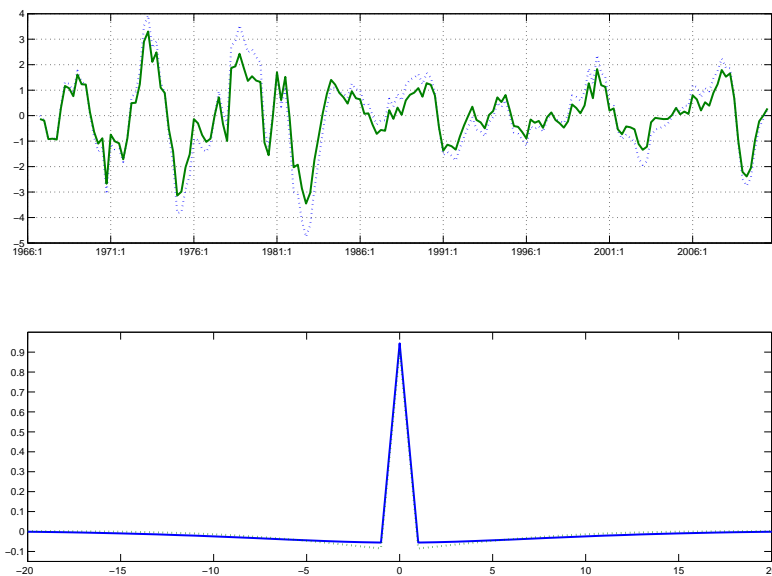
B. Output Gap Estimation using a Multivariate Semi-Structural Filter

This part of the paper discusses the results of a state-of-the-art multivariate model (filter) for the output gap estimation, for the US economy, developed by Benes and others (2010).[15] This

---

[15]The replication materials are publicly available and can be freely downloaded from `www.douglaslaxton.com`

particular model structure has been and is being used in policy analysis in many instances, see Cheng (2011), Scott and Weber (2011) or Babihuga (2011), among others. This paper suggests some additional angles of viewing the model properties that practitioners could use to gain more insight into the model as a filtering device, along its economic structure.

I analyze how the application of the model for the US economy makes use of the observed data and I provide an elementary frequency domain analysis of the implied filter. The model features an exceptionally small revision variance, together with very good forecasting properties, see Benes and others (2010), hence these will not be discussed in detail.

The model is specified by equations (44)–(55). The authors formulate a simple backward-looking output gap equation, the Phillips curve, Okun's law, and also use the capacity utilization series as an additional measurement of the cyclical signal in the economy. A deviation of the year-over-year inflation from long-term inflation expectations contributes to the output gap negatively, as a supply shock and an implicit tightening of the monetary policy stance. On the other hand, a positive output gap increases the inflation due to excess demand pressures.

The output gap, y_t, is linked to capacity utilisation gap, c_t, unemployment gap, u_t, via simple measurement relationship and the Okun's law, unlike a bit more structural output gap and Phillips curve relationship. Capacity utilisation, unemployment, and GDP feature the trend-component specification essentially identical to (40)–(43). An interesting aspect of the model is that a year-over-year inflation, π_t^4, follows a unit root process, though it is anchored by long-term inflation expectations process, π_t^{4LTE}.[16] The authors consider the model to be a simple and pragmatic way to obtain a measure of the output gap, which has outstanding revision properties and is thus suitable for real-time policy making.[17]

---

[16] The specification of inflation in year-over-year terms also has structural implications. First, a year-over-year filter $(1 - L^4)$ attenuates high and seasonal frequencies, as a high frequency noise is hardly expected to be related to the output gap. Second, the filter implies a phase delay of around 5.5 months, since it essentially is a one-sided geometric moving average. Inflation developments thus propagate only very gradually to output gap.

[17] Authors also suggest that more complex, forward-looking models are the subject of their further research.

The model is as follows:

$$y_t = Y_t - \bar{Y}_t \qquad c_t = C_t - \bar{C}_t \qquad u_t = \bar{U}_t - U_t \tag{44}$$

$$y_t = \rho y_{t-1} - \tilde{\rho}_2(\pi_{t-1}^4 - \pi_{t-1}^{4LTE}) + \varepsilon^y \tag{45}$$

$$\pi_t^4 = \pi_{t-1}^4 + \beta y_t + \Omega(y_t - y_{t-1}) + \varepsilon_t^{\pi 4} \tag{46}$$

$$u_t = \phi_1 u_{t-1} + \phi_2 y_t + \varepsilon_t^u \tag{47}$$

$$c_t = \kappa_1 u_{t-1} + \kappa_2 y_t + \varepsilon_t^c \tag{48}$$

$$\pi_t^{4LTE} = \pi_t^{4LTE} + \varepsilon_t^{\pi 4LTE} \tag{49}$$

$$\bar{Y}_t = \bar{Y}_{t-1} + G_t^{\bar{Y}}/4 + \theta(\bar{U} - \bar{U}_{t-1}) - (1-\theta)(\bar{U} - \bar{U}_{t-20})/19 + \varepsilon_t^{\bar{Y}} \tag{50}$$

$$G_t^{\bar{Y}} = \tau G_{SS}^{\bar{Y}} + (1-\tau)G_{t-1}^{\bar{Y}} + \varepsilon_t^{G^{\bar{Y}}} \tag{51}$$

$$\bar{U}_t = \bar{U}_{t-1} + G_t^{\bar{U}} - \tilde{\omega} y_{t-1} - \tilde{\lambda}(\bar{U}_{t-1} - U^{SS}) + \varepsilon_t^{\bar{U}} \tag{52}$$

$$G_t^{\bar{U}} = (1-\alpha)G_{t-1}^{\bar{U}} + \varepsilon_t^{G^{\bar{U}}} \tag{53}$$

$$\bar{C}_t = \bar{C}_{t-1} + G_t^{\bar{C}} + \varepsilon_t^{\bar{C}} \tag{54}$$

$$G_t^{\bar{C}} = (1-\delta)G_{t-1}^{\bar{C}} + \varepsilon_t^{G^{\bar{C}}}, \tag{55}$$

where all innovation processes ε_t^i are uncorrelated and follow the Gaussian distribution with a zero mean and variance specified in Benes and others (2010). Obviously, the standard deviations of all innovations are a crucial part of the model's transfer function, although the impulse-response function remains unaffected.

In terms of shock decomposition, the model's output gap can be function of its own output gap innovations and innovations to inflation or inflation expectations and thus is not very interesting, though it cannot be omitted from the analysis. A more interesting and non-standard analysis uses the filter representation and provides the decomposition into observables. Such analysis is given below.

1. Decomposition into Observables

An interesting question is how individual observables (GDP, y/y inflation, capacity utilization, or unemployment) contribute to the final estimate of the output gap. This is easily answered by carrying out *the decomposition into observables* of the model's state-space form. This complements the analysis in Benes and others (2010) and provides the example use of the methods discussed above.

Contribution of all observables to the output gap is depicted at the Fig. 2. The first thing to notice is that the contribution of the GDP growth to the 'Great Recession' estimates started in 2007 and is smaller than in the case of the largest recession within the sample (in terms of the

Figure 2. Output-Gap Observable Decomposition of Benes et al. (2010) model
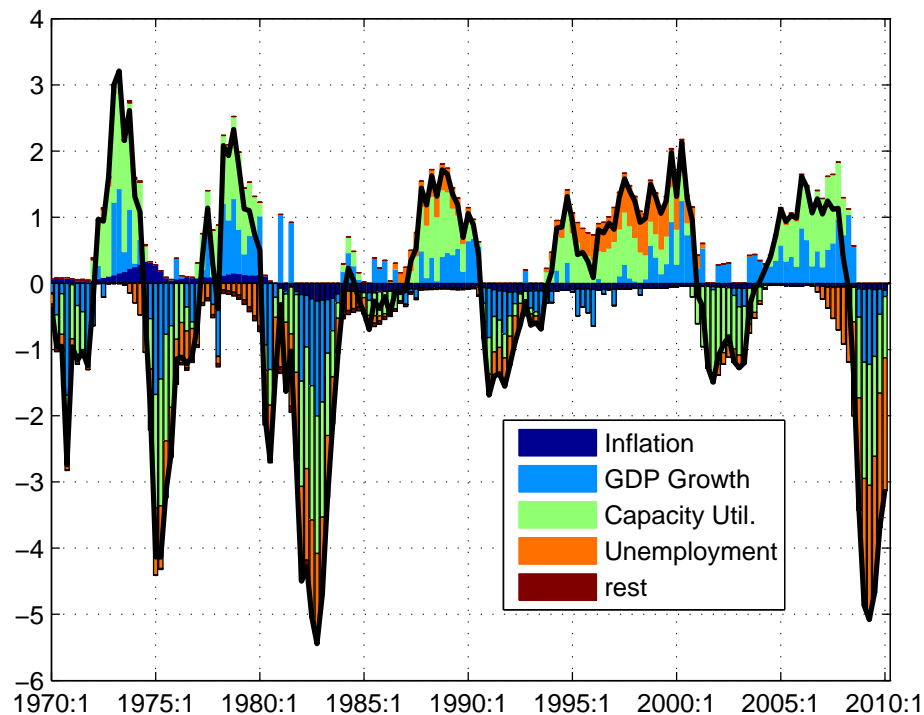

output gap) in 1980–1985, whereas the contribution of the unemployment series is the greatest ever. Second, the information extracted from observing the inflation data is rather limited. The fact that the parameterization of the model attributes a negligible role for inflation in identification of the output gap could be controversial. It may also undermine any potential interpretation of the output gap with respect to non-accelerating inflation or New Keynesian theories.

The contribution of unemployment is large and more persistent than the contribution of GDP and capacity utilization. That is quite consistent with the jobless recovery that the US economy has been experiencing since the 1990s, where the GDP and capacity utilization recover faster than the unemployment rate and the inflation expectations are well anchored. The model is parametrized to imply that NAIRU takes more time to change than potential output. Both NAIRU and potential output growth vary less than most models based on definition of cycle with frequency 6–32 periods, which contributes to realistic and interpretable magnitudes of the output gap.

Empirically, the unemployment lags output at business cycle frequencies, with lower amplitude and high coherence. This is one of the most robust stylised facts across most developed economies. Fig. 10 depicts output, capacity utilisation and unemployment (shift to lead by one quarter) after applying a low-pass filter (HP filter, $\lambda = 1600$ for simplicity[18]) and scaled to output gap volatility, with a phase-shift. The tight co-movement is important for a robust signal extraction and lower revision variance of the model, see the Appendix for details.

These results complement the analysis in Benes and others (2010). The low role for inflation signal can be easily understood by looking at parameter estimates. Importantly, the loading coefficient of the 'inflation gap' to the output gap is only $\tilde{\rho}_2 = 0.005$ and also the parameters determining output gap effect on inflation β, Ω are small.

Figure 3. Transfer function gains, Beneš et al. (2010) model

---
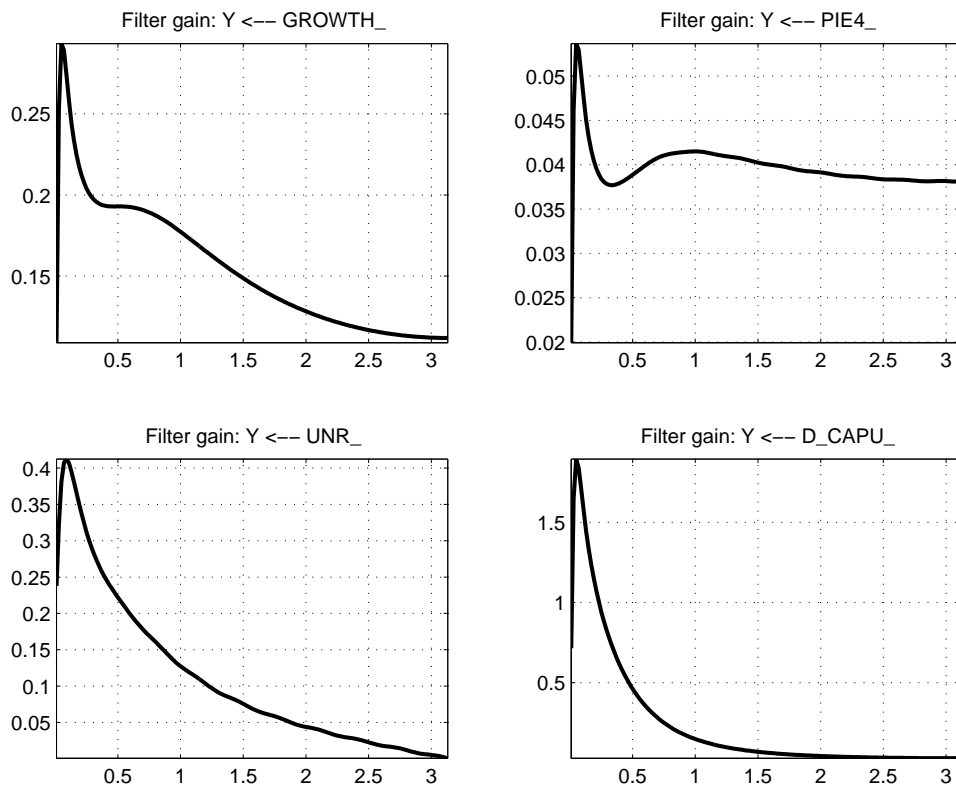
[18]There is nothing magical about the value of 1600, arguments can be found as to why such a value is inappropriate when the recovered cycle should be used as an output gap concept.

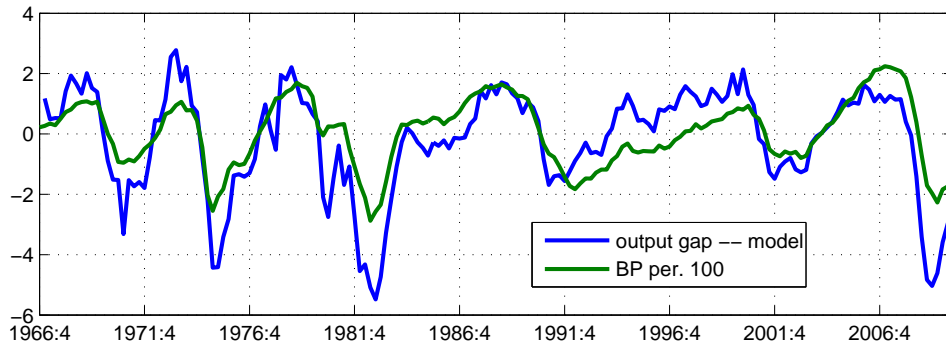**Figure 4. Univariate approximation using unemployment**
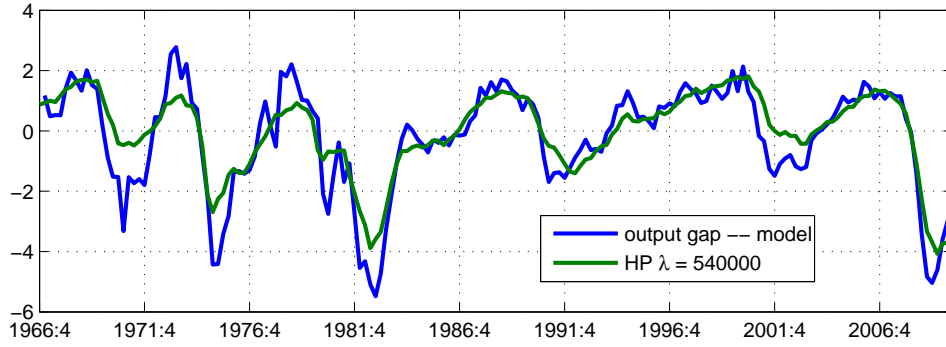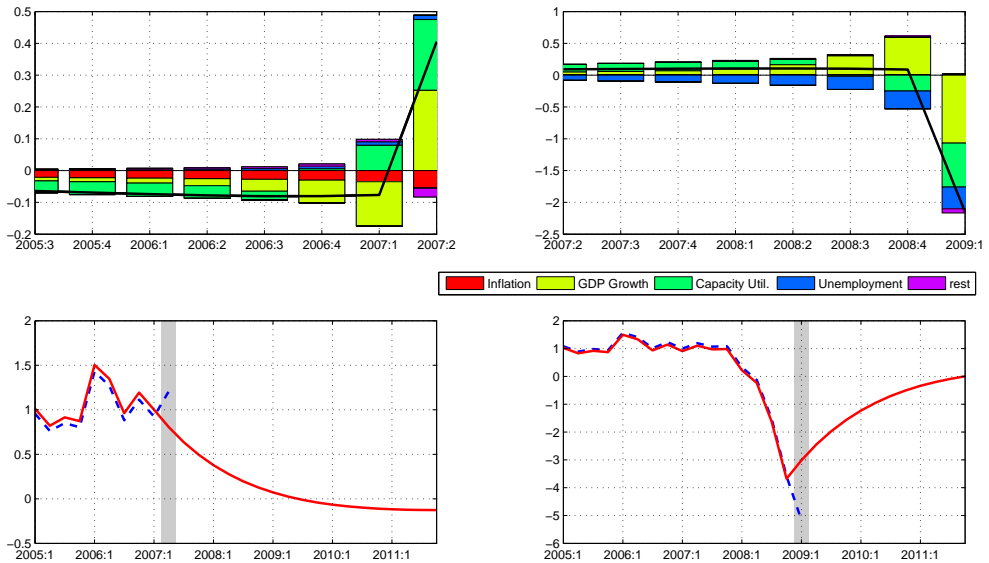

Figure 5. News Effects Decomposition

The most influential observable in the model turns out to be the unemployment. This can be easily seen at Fig. 6, where the output gap estimates using all observables and the unemployment as the only input are contrasted. Adding the capacity utilization series brings the model closer to the final estimates, but the series all alone performs worse in after 1990's. Using the GDP growth observations alongside the unemployment and capacity utilisation modifies the estimates just a tiny bit, inflation and the inflation expectations bear close to no information in the model framework.

Figure 6. Output-Gap Estimates from the Beneš et al. (2010) model
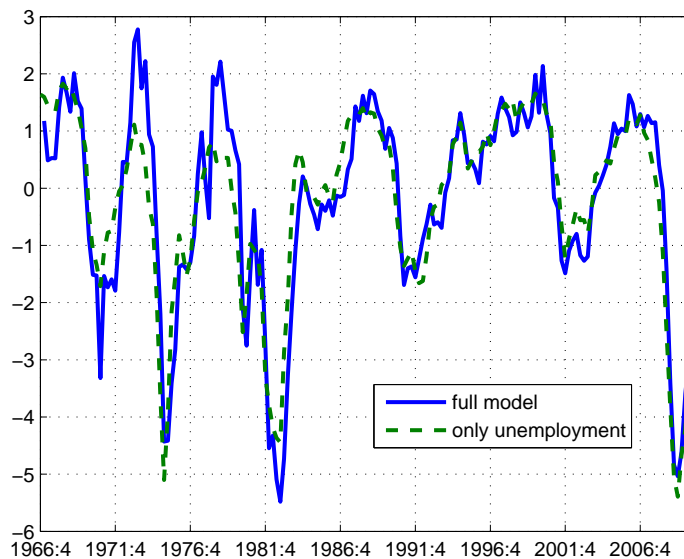
Analysis of the transfer function of the model   As in the case of univariate filters it is feasible to analyze the transfer function of the model in the frequency domain. Fig. 3 presents the gain of the model's transfer function for the output gap, Y, and four key observables in the model – GDP growth, inflation, unemployment, and first change in the capacity utilisation.

The interpretation of the gain is standard, as in the univariate filters analysis. It demonstrates which frequencies of the observed variables spill into the output gap estimate. One can see that the gain from the inflation is rather flat across the whole spectrum and does not distinguish between business cycle and high frequency dynamics. It is also very small. Looking at the gain of output with respect to the GDP growth and changes in the capacity utilization,

indicates that the filter places large weight at lower frequencies, some weight at business cycle frequencies (6-32 quarters) and little weight at high frequencies. One has to take into account, however, that the first difference filter itself boosts high frequencies. The analysis of the output gain for the level of the observed variables requires a straightforward application of the integration filter, the inverse of the first difference filter with the transfer function $1/(2 - 2\cos(\omega))$.

The gain profile going from the level of unemployment to the output gap is of great interest, following the result on the importance of unemployment observations for determination of the gap in the model. The weight on low frequencies and only a gradual decline of the gain with an increase in the periodicity suggest rather large spillover of longer cycles into the output gap. This is also clear from the spectral density of the output gap implied by the model.

To frame the discussion within a time domain, one can search for specification of the simple Hodrick-Prescott filter or a frequency cutoff of Christiano and Fitzerald's band pass filter that would minimize the distance to the output gap estimated by the model. The model draws more cyclical information from the unemployment series rather from the GDP. Unemployment is used to back out the output gap so as to match the model's estimate. Fig. 4 demonstrates that a univariate approximation using the HP filter with a large value of the smoothing parameter λ, i.e. smooth trend, is quite successful. The results are not surprising, as the importance of unemployment series and weight on lower frequencies (longer cycles) was explained.

2. News Effects Decomposition

Despite its excellent revision properties, the model can be used to illustrate the news effect and a decomposition of output gap revisions into relevant observables. The illustration focuses 2007Q2 and 2009Q1, both being interesting periods with respect to the 'Great Recession'. The results are depicted in Fig. 5. As can be seen, the revision properties of the model are quite favourable. A news effect is defined as a projection error, and thus it can be easily quantified and decomposed into components.

The data arriving in 2009Q1 resulted into a further deepening of the output gap estimate. The drop in the output gap was a complete news, as the model dynamics would imply a start of a recovery and closing of the gap. The largest contribution to the news is due to new data observation for the GDP growth, followed by capacity utilization, and unemployment numbers.

Consistently with above findings observed data on inflation or inflation expectations contribute only modestly and revise the estimate in a persistent way.

Clearly, to properly interpret the contribution graph in Fig. 5 it is important to understand not only the derivative of the filter to the new data, but also the differential, i.e. the difference between forecast of observables and the actual outcome. In the case of the example using the 2009Q1 the GDP growth, capacity utilization, and unemployment were both lower than the model would predict. The more complex is the model, the bigger is the value of a formal and automatized approach.

C.   Natural Output Gap in a DSGE Model

This section decomposes the 'flexible-price equilibrium output gap' into observables using the model of Smets and Wouters (2007). The model is an estimated model of the US economy.[19] The model is a medium-sized DSGE model that uses the output gap concept consistent with a theoretical definition of the 'natural rate of output' in the absence of nominal wage and price rigidities and with no 'mark-up' shocks to wages and prices.

The output gap in the model is defined as a deviation of the 'natural rate of output' from the actual output. Despite a very different definition and modeling framework from most empirical measures of the output gap, all tools discussed in the paper continue to apply. The decomposition of the flexible-price output gap is depicted in Fig. 7

In a closed economy model, the output gap is usually analyzed only in terms of the shock decomposition, which provides a structural interpretation of the historical developments. The decomposition into observables is essentially a reverse process, indicating which observables are bing used by the structure of the model to identify latent variables – technology, preference, and other shocks. One can observe that some variables are important at high and business cycle frequencies (hours worked, consumption, and output), whereas inflation or interest rates contribute mainly to low frequency dynamics of the output gap in this case. The model thus extracts the information about the natural rate of output without placing much weight on the observed inflation data, similarly to a semi-structural model above. The small weight on inflation would be easy to explain if the observations of real wages would contribute significantly to business cycle dynamics of the output gap, yet that is not the case.

---

[19]I would like to thank to authors for making available the codes for replicating their work. As I have moved the codes from Dynare, I retain the blame for any errors.

Figure 7. Output Gap Observable Decomposition from a DSGE Model
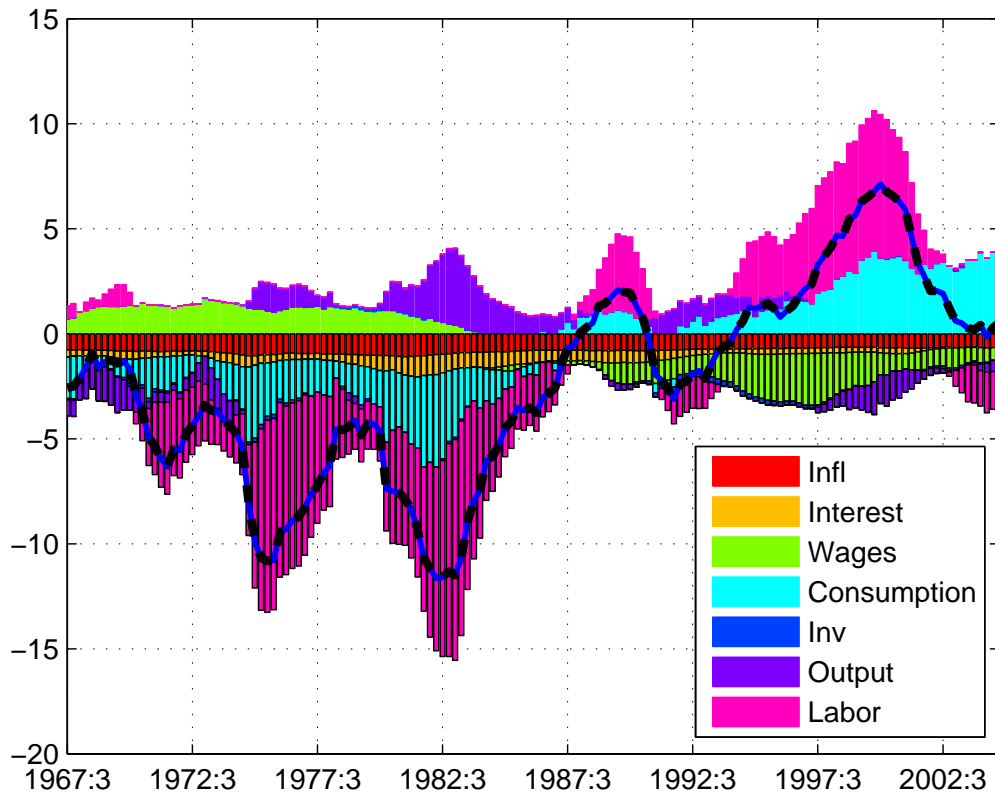
IV. Conclusions

This paper suggests a simple and useful method for exploring 'what is in the output gap.' A decomposition of the unobservable output gap in terms of observed inputs, e.g. output, inflation, or unemployment rate is provided. The procedure answers the question of what observed variables and at which periodicity contribute to the estimate of an unobservable quantity – the output gap. Importantly, the decomposition into observables also allows researches to quantify the 'news effects' caused by new available data, generating revisions of the output gap estimates. A better understanding of the role that new data play for a change in the estimate is easier to obtain using the decomposition into observables.

The paper demonstrates that the most frequently used methods for potential output estimation can be cast in terms of a linear filter theory. This enables both a frequency- and time-domain analysis and provides insights into the nature of revisions of the unobserved variables estimates. The analysis in the paper applies to simple multivariate filters, semi-structural models, production function method estimates, and fully articulated DSGE models.

The method is illustrated using a semi-structural multivariate filter and a fully articulated DSGE model. Using the multivariate filter for potential output estimation, the paper demonstrates that new insight is obtained due to the decomposition into observables. Both models considered attribute very low weight to observed data on inflation when identifying the output gap.

Revision properties and the 'end-point bias' of individual approaches can be better understood as properties of the two-sided moving average, or, the filter representation. The more spread out the weights and the worse the forecasting properties of the filter-implied model, the larger the real-time revision variance of the estimate. For a particular data-generating process a population or analytical exploration of revision error stationary stochastic process can be easily performed.

The paper also shows that, a priori, there is no reason to expect that multivariate filters, expressed in a state-space form, should feature better revision properties than univariate filters. The key is the structure of the model, providing the link between the economic theory and optimal signal extraction principles.

References

Andrle, M., 2009, "DSGE Filters for Forecasting and Policy Analysis," Techn. rep., Czech National Bank/European Central Bank.

Andrle, M., Ch. Freedman, R. Garcia-Saltos, D. Hermawan, D. Laxton, and H. Munandar, 2009a, "Adding Indonesia to the Global Projection Model," Working Paper 09/253, International Monetary Fund, Washington DC.

Andrle, M., T. Hlédik, O. Kameník, and J. Vlček, 2009b, "Implementing the New Structural Model of the Czech National Bank," Working paper no. 2, Czech National Bank.

Babihuga, R., 2011, "How Large is Sweden's Output Gap," Sweden: 2011 Article IV Consultation–Staff Report; Public Information Notice on the Executive Board Discussion; and Statement by the Executive Director for Sweden. IMF Country Report No. 11/171 11/171, International Monetary Fund, Washington DC.

Bell, W.R., 1984, "Signal extraction for nonstationary time series," *Annals of Statistics*, Vol. 12, pp. 646–664.

Benes, J., K. Clinton, R. Garcia-Saltos, M. Johnson, D. Laxton, P. Manchev, and T. Matheson, 2010, "Estimating Potential Output with a Multivariate Filter," Working Paper 10/285, International Monetary Fund, Washington DC.

Benes, J., and P. N'Diaye, 2004, "A Multivariate Filter for Measuring Potential Output and the Nairu," Working Paper 04/45, International Monetary Fund, Washington DC.

Blanchard, O., and D. Quah, 1989, "The dynamic effects of aggregate supply and demand disturbances," *American Economic Review*, Vol. 79, pp. 655–673.

Carabenciov, I., I. Ermolaev, Ch. Freedman, M. Juillard, O. Kamenik, D. Korshunov, and D. Laxton, 2008, "A Small Quarterly Projection Model of the US Economy," Working Paper 08/278, International Monetary Fund, Washington DC.

Cayen, J.-P., and S. van Norden, 2005, "The Reliability of Canadian Output-Gap Estimates," *The North American Journal of Economics and Finance*, Vol. 16, pp. 373–393.

Cheng, K. C., 2011, "France's Potential Output during the Crisis and Recovery," France: Selected Issues Paper, IMF Country Report No. 11/212 11/212, International Monetary Fund, Washington DC.

Christiano, L.J., and T.J. Fitzgerald, 2003, "The Bandpass Filter," *International Economic Review*, Vol. 44, No. 2, pp. 435–465.

Claus, I., 1999, "Estimating potential output for New Zealand: a structural VAR approach," Discussion Paper 2000/03, Reserve Bank of New Zealand.

Conway, P., and B. Hunt, 1997, "Estimating Potential Output: A Semi-Structural Approach," Discussion Paper G97/9, Reserve Bank of New Zealand.

de Brouwer, G., 1998, "Estimating Output Gaps," Research Discussion Paper 9809, Reserve Bank of Australia.

Epstein, N., and C. Macciarelli, 2010, "Estimating Poland's Potential Output: A Production Function Approach," Working Paper 10/15, International Monetary Fund, Washington DC.

Gomez, V., 1999, "Three Equivalent Methods for Filtering Finite Nonstationary Time Series," *Journal of Business & Economics Statistics*, Vol. 17, No. 1, pp. 109–116.

———, 2001, "The Use of Butterworth Filters for Trend and Cycle Estimation in Economic Time Series," *Journal of Business & Economic Statistics*, Vol. 19, pp. 365–373.

———, 2006, "Wiener-Kolmogorov Filtering and Smoothing for Multivariate Series with State-Space Structure," *Journal of Time Series Analysis*, Vol. 28, No. 3, pp. 361–385.

Harvey, A., and T. Trimbur, 2008, "Trend Estimation and the Hodrick-Prescott Filter," *Journal of Japan Statistical Society*, Vol. 38, No. 1, pp. 41–49.

Hodrick, R., and E. Prescott, 1997, "Post-War Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking*, Vol. 29, No. 1, pp. 1–16.

Kaiser, R., and A. Maravall, 1999, "Estimation of the Business Cycle: A Modified Hodrick-Prescott Filter," *Spanish Economic Review*, Vol. 1, pp. 175–206.

———, 2001, *Measuring Business Cycles in Economic Time Series* (New York: Springer-Verlag, Lecture Notes on Statistics 154).

Kalman, R.E., 1960, "A new approach to linear filtering and prediction problems," *Trans, ASME, Ser. D., Journal of Basic Engineering*, Vol. 82, pp. 35–45.

King, R.G., and S.T. Rebelo, 1993, "Low frequency filtering and real business cycles," *Journal of Economic Dynamics and Control*, Vol. 17, No. 1–2, pp. 207–231.

Koopman, S.J., and A. Harvey, 2003, "Computing observation weights for signal extraction and filtering," *Journal of Economic Dynamics and Control*, Vol. 27, pp. 1317–1333.

Koopmans, L.H., 1974, *The Spectral Analysis of Time Series* (San Diego, CA: Academic Press).

Kuttner, K., 1994, "Estimating Potential Output as a Latent Variable," *Journal of Business and Economic Statistics*, Vol. 12, No. 3, pp. 361–368.

Laubach, T., and J.C. Williams, 2003, "Measuring the Natural Rate of Interest," *The Review of Economics and Statistics*, Vol. 85, No. November, pp. 1063–1070.

Laxton, D., and R. Tetlow, 1992, "A Simple Multivariate Filter for the Measurement of the Potential Output," Technical Report 59 (June), Bank of Canada.

Leser, C.E.V., 1961, "A Simple Method of Trend Construction," *Journal of the Royal Statistical Society, Series B (Methodological)*, Vol. 23, No. 1, pp. 91–107.

McNellis, P.D., and C.B. Bagsic, 2007, "Output Gap Estimation for Inflation Forecasting: the Case of the Philippines," Techn. rep., Bangko Sentral ng Pilipina.

Orphanides, A., and S. van Norden, 2002, "The Unreliability of Output-Gap Estimates in Real Time," *Review of Economics and Statistics*, Vol. 84, pp. 569–583.

Pierce, D.A., 1980, "Data revisions in moving average seasonal adjustment procedures," *Journal of Econometrics*, Vol. 14, No. 1, pp. 95–114.

Pollock, D.S.G., 2000, "Trend Estimation and De-trending via Rational Square-wave Filters," *Journal of Econometrics*, Vol. 98, No. 1-3, pp. 317–334.

Proietti, T., 2009, "On the Model-Based Interpretation of Filters and the Reliability of Trend-Cycle Estimates," *Econometric Reviews*, Vol. 28, No. 1-3, pp. 186–208.

Proietti, T., and A. Musso, 2007, "Growth Accounting for the Euroarea – a Structural Approach," Working Paper 804, European Central Bank.

Schleicher, Ch., 2003, "Wiener-Kolmogorov Filters for Finite Time Series," Techn. rep., University of British Columbia.

Scott, A., 2000, "Stylised Facts from Output Gap Measures," Discussion Paper DP2000/07, Reserve Bank of New Zealand, Wellington.

Scott, A., and S. Weber, 2011, "Potential Output Estimates and Structural Policy," Kingdom of The Netherlands–Netherlands: Selected Issues and Analytical Notes, IMF Country Report No. 11/212 11/143, International Monetary Fund, Washington DC.

Smets, F., and R. Wouters, 2007, "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, Vol. 97, No. 3, pp. 586–606.

Whittle, P., 1983, *Prediction and Regulation by Linear Least-Square Methods, Second Ed.* (Minneapolis: University of Minnesota Press).

Appendix A. Parameter Estiates from Beneš et al. (2010)

The model by Benes and others (2010), as specified by equations (44)–(55) is econometrically estimated using United States data for period 1967:1–2010:2. The approach is a Bayesian-likelihood, more specifically a 'regularized maximum likelihood' – a method popular in engineering. The method is equivalent to a likelihood estimation with an independent joint prior on parameters coming from truncated-Normal distribution, as upper and lower bound for parameter are estimated.

Table 1. Estimated parameters for Beneš et al. (2010)

Parameter	Prior		Posterior	
	Mode	Dispersion	Mode	Dispersion
$G_{SS}^{\bar{Y}}$	2.140	0.300	2.143	0.041
$U_S S$	5.800	0.300	5.801	0.042
θ	0.700	0.030	0.700	0.004
α	0.900	0.150	0.895	0.022
β	0.400	0.300	0.176	0.024
Ω	0.500	0.300	0.265	0.038
ρ	0.800	0.150	0.821	0.019
$100\tilde{\rho}_2$	5.000	3.000	5.145	0.418
κ_1	0.100	0.600	0.353	0.037
κ_2	1.500	1.500	1.421	0.070
ϕ_1	0.800	0.150	0.729	0.018
ϕ_2	0.300	0.150	0.241	0.014
τ	0.100	0.150	0.113	0.021
δ	0.500	0.150	0.496	0.021
$100\tilde{\omega}$	3.000	1.500	2.796	0.207
λ	2.000	3.000	2.078	0.371
σ_Y	1.000	0.300	0.888	0.039
σ_G	1.000	0.300	0.921	0.042
σ_u	0.500	0.300	0.198	0.024
$\sigma_{\bar{U}}$	0.100	0.150	0.079	0.018
$\sigma_{G_{\bar{U}}}$	0.100	0.150	0.087	0.017
σ_u	0.400	0.300	0.482	0.033
σ_c	0.250	0.150	0.288	0.021
$\sigma_{G_{\bar{C}}}$	0.075	0.030	0.076	0.004
$\sigma_{\pi 4}$	0.500	0.300	0.554	0.030
$\sigma_{\pi 4LTE}$	0.300	0.300	0.081	0.008
$\sigma_{\bar{Y}}$	0.250	0.100	0.217	0.013

APPENDIX B. NOT FOR PUBLICATION: DIFFERENCE BETWEEN TWO REPRESENTATIONS OF THE HP FILTER

The HP filter state-space form is often represented in the following form:

$$y_t = \bar{y}_t + x_t \tag{56}$$

$$x_t = \varepsilon_t^x \qquad \varepsilon_t^x \sim N(0, \sigma_x^2) \tag{57}$$

$$\bar{y}_t = \bar{y}_{t-1} + \beta_{t-1} \tag{58}$$

$$\beta_t = \beta_{t-1} + \varepsilon_t^g \qquad \varepsilon_t^g \sim N(0, \sigma_g^2), \tag{59}$$

which is equivalent the the HP filter state-space representation included in the text and the results from both state-space implementations are identical. Moreover, these state-space representations are identical to a standard matrix formulas of HP filter implementations.

APPENDIX C. EXAMPLE: SIMPLE MULTIVARIATE FILTER – THREE REPRESENTATIONS

In this section I provide a simple example, beyond univariate HP filter, of a semi-structural multivariate filters represented as a (i) state-space model, (ii) Wiener-Kolmogorov filter and (iii) penalized least squares.

The state-space form of the model is

$$y_t = x_t + \tau_t \tag{60}$$

$$\tau_t - \tau_{t-1} = \tau_{t-1} - \tau_{t-2} + \varepsilon_t^\tau \tag{61}$$

$$x_t = \rho x_{t-1} - \kappa \pi_t^{gap} + \varepsilon_t^x \tag{62}$$

$$\pi_t^{gap} = \lambda \pi_{t-1}^{gap} + \theta x_t + \varepsilon_t^\pi \tag{63}$$

and can be casted in a penalized least squares problem

$$\min_{\{\tau\}_0^T} \sum_{t=0}^T \frac{1}{\sigma_x^2}[\varepsilon_t^x]^2 + \frac{1}{\sigma_\tau^2}[\varepsilon_t^\tau]^2 + \frac{1}{\sigma_\tau^2}[\varepsilon_t^\pi]^2, \tag{64}$$

which is in a form of 'multivariate Hodrick-Prescott filter', as suggested by Laxton and Tetlow (1992) and for $\rho = \kappa = \theta = 0$ and $\lambda = (\sigma_x/\sigma_\tau)^2$ is equivalent to the penalized least squares formulation of the HP filter. The problem (64) is easy to solve by finding first-order conditions with respect to $\{\tau_t\}_{t=0}^T$.

The implied Winer-Kolmogorov estimate of the output gap x_t is given by the multivariate filter in terms of output level and inflation (or output growth and inflation) is as follows

$$x_t = (1 - W_y(L))y_t - W_\pi(L)\pi_t^{gap} = \frac{(1 - W_y(L))}{(1 - L)}\Delta y_t - W_\pi(L)\pi_t^{gap} \qquad (65)$$

where

$$W_y(L) \equiv \frac{(\lambda_1 + \lambda_2\alpha^2 + \rho^2\lambda_1) - \lambda_1\rho(L - L^{-1})}{(\lambda_1 - 2\lambda_3 + \alpha^2\lambda_2 + \rho^2\lambda_1) - (4\lambda_3 + \lambda_1\rho)(L + L^{-1}) + \lambda_3(L^2 + L^{-2})} \qquad (66)$$

$$W_\pi(L) \equiv \frac{-(\lambda_1\kappa + \alpha\lambda_2) + \alpha\lambda_2\phi L + \rho\lambda_1\kappa L^{-1}}{(\lambda_1 - 2\lambda_3 + \alpha^2\lambda_2 + \rho^2\lambda_1) - (4\lambda_3 + \lambda_1\rho)(L + L^{-1}) + \lambda_3(L^2 + L^{-2})} \qquad (67)$$

express the z-transform of the two-sided filter weights. The time domain weights profile can be easily obtained either from a state-space representation using the theory outlined in the paper, or from (66) and (67) by computing inverse z-transform either numerically or analytically by factorizing the formulas.

Appendix D. Not for publication: Variance Reduction via Common Component and Multiple Measures

In this section I briefly review the STAT-101 intuition about revision variance reduction by adding multiple relevant measures of an underlying signals. In case there are more noisy, but *relevant* measurement available, then (i) the variance of the estimates is lowered and, equivalently, (ii) the weights of the filter are less spread-out, lowering revision variance.

The most simple example is an estimate of a deterministic signal μ from one or two noisy signals: $z_1 = \mu + u_1$ and $z_2 = \mu + u_2$, with $u_1 \sim N(0, \sigma_1^2)$ and $u_1 \sim N(0, \sigma_2^2)$. In case of just one signal, z_1, the estimate is simply $\hat{\mu} = z_1$, with variance of the estimate σ_1^2. When both measurements are available, the estimate is given by

$$\hat{\mu} = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}z_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}z_2 \qquad \text{MSE}_{\hat{\mu}} = \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)^{-1}. \qquad (68)$$

It is clear from (68) that as long as the second measurement is available, i.e. $\sigma_2^2 < \infty$, the precision of the estimate is increasing. This is a principle that carries over into more complex, dynamic models analyzed using Kalman or Wiener-Kolmogorov filtering. The larger the precision of the estimate, ceteris paribus, the less spread out weights of dynamic models are.

I can also analyze more realistic model. Still, for simplicity I will ignore the trend components of the signal. Let us assume a model of the AR(1) signal x_t and two available measure-

ments y_1 and y_2:

$$x_t = \rho x_{t-1} + \nu \tag{69}$$

$$y_{1,t} = x_t + \varepsilon_{1,t} \tag{70}$$

$$y_{2,t} = \phi x_t + \varepsilon_{2,t}, \tag{71}$$

where the parameter ϕ is indicating the degree of relevance of the signal, together with variances associated with error terms, σ_ν^2, σ_1^2 and σ_2^2. The signal extraction differs dramatically for various parameter values of ρ and variances.

This is rather simple and well understood problem with analytical solution. The weights of the two-sided filter will be symmetric and follow the scheme

$$x_{t|\infty} = c_1 \sum_{i=-\infty}^{\infty} \lambda^{|i|} X_{t-i}, \tag{72}$$

where a variable X denotes convex combination of both observables, following essentially the weighting scheme (68). The parameter λ can be recovered from a solution of quadratic equation associated with the transfer function of the filter. For the clarity of exposition a numerical examples are provided below.

(D.0.0.1) **Single measurement**   In case of single measurement only, or, equivalently, $\phi = 0$, the estimate of x_t using doubly infinite sample is given by

$$x_t = \frac{q}{q + |1 - \rho L|^2} y_{1,t}, \tag{73}$$

which implies symmetric two-sided filter with weights decaying in an exponential way. The higher is the persistence ρ the slower is the decay and the larger is the revision variance. For $\rho = 0$ only the concurrent values of $y_{1,t}$ are used – $x_t = q/(1+q) = \sigma_x^2/(\sigma_x^2 + \sigma_1^2)y_t$.

(D.0.0.2) **Two measurements of the signal**   The easiest case to analyze is when there is no dynamics, $\rho = 0$, as the estimator uses only current period values of observables $y_{1,t}, y_{2,t}$.

The estimation with dynamics and two observables yields the two-sided filter of the form

$$x_t = \frac{q_1}{(q_1 + \phi q_2) + |1 - \rho L|^2} y_{1,t} + \frac{\phi q_2}{(q_1 + \phi q_2) + |1 - \rho L|^2} y_{2,t}. \tag{74}$$

In case $\rho = 0$ the problem is trivial and solved above. Note that the weights for both observables will by symmetric and proportional, only rescaled by appropriate factor in terms of their relative informativeness, given by the variance of the measurement error and cross-correlation ϕ. Fig. 8 depicts the problem with $\rho = 0.50$ and $\rho = 1.0$, given $\phi = 1$ and $\sigma_{y_1} = \sigma_{y_2} = 0.9$.

Figure 8. Weights of AR(1) model
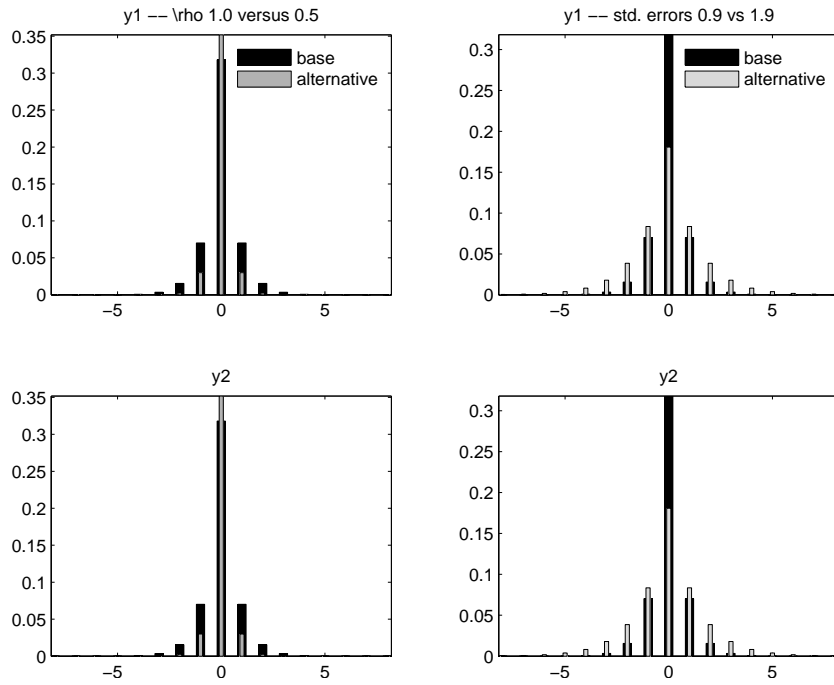
APPENDIX E.   NOT FOR PUBLICATION: ADDITIONAL GRAPHS
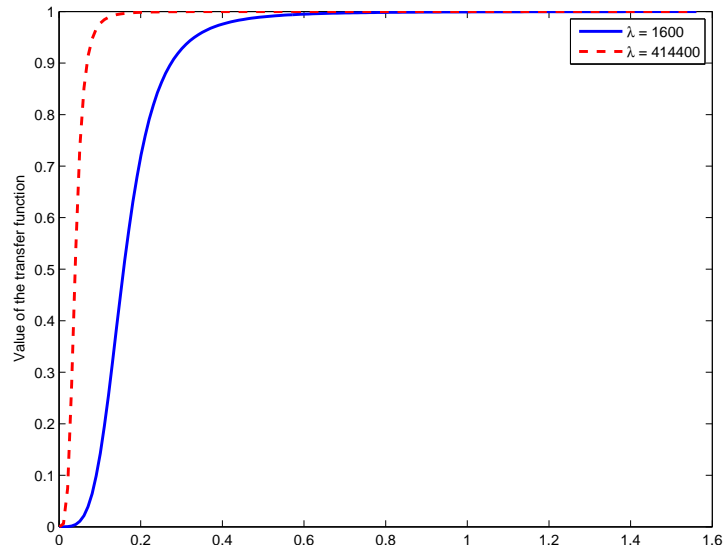
Figure 9. HP filter transfer function

Figure 10. Output, unemployment (+1Q) and cap. utilization detrended