Dynare Working Papers Series
http://www.dynare.org/wp/

Understanding DSGE Filters in Forecasting and Policy Analysis

Michal Andrle

Working Paper no. 16

October 2012

CEPREMAP
CENTRE POUR LA RECHERCHE ECONOMIQUE ET SES APPLICATIONS
142, rue du Chevaleret — 75013 Paris — France
http://www.cepremap.ens.fr

Understanding DSGE Filters in Forecasting and Policy Analysis

Michal Andrle[1]

International Monetary Fund, Research Dept., 700 19th St. N.W., Washington, D.C., 20431, USA

Abstract

The paper introduces methods that allow analysts to (i) decompose the estimates of unobserved quantities into observed data and (ii) impose subjective prior constraints on path estimates of unobserved shocks in structural economic models. For instance, decomposition of output gap to output, inflation, interest rates and other observables contribution is feasible. The intuitive nature and the analytical clarity of procedures suggested are appealing for policy-related and forecasting models. The paper brings some of the power embodied in the theory of linear multivariate filters, namely relatinship between Kalman and Wiener-Kolmogorov filtering, into the area of structural multivariate models, expressed in linear state-space form.

Keywords: filter, DSGE, state-space, observables decomposition, judgement
JEL Codes: C10, E50

1. Introduction

This paper puts forth a decomposition of unobserved quantities into contributions of observed data as a useful tool for students of business cycles who use linearized Dynamic Stochastic General Equilibrium (DSGE) models. The reason for using such a decomposition is that it can reveal how observed data say on output, inflation, interest rates, etc., contribute to an estimated total factor productivity shock, output gap estimate, preference shocks or any other unobserved variable. This decomposition will be referred to as *observables decomposition* in the text. One can think of the procedure as a counterpart to a standard 'shock decomposition', where observed data series are expressed as contributions of estimated structural shocks.[2] An example of an important variable is the output gap, where one can clearly trace contributions of input data using methods presented below.

Techniques proposed in the paper provide also a natural approach to analysis of historical data revisions and news effects, when new data become available. A decomposition of individual data series or data points is easily mapped to model variables. This has proven to be useful in forecasting with DSGE models, as changes in the initial state of the economy and the forecast can be linked to changes in the observed data easily.

The organizing principle of the paper is the linear filter representation of the model. Acknowledging explicitly that estimated structural shocks in a structural model are identified by a two-sided linear multivariate filter, implied by the state-space form of the model is crucial for the analysis. It opens DSGE models to techniques associated with theory of linear (multivariate) filters both in time and frequency domains. Importantly, it is key to recall that a Kalman filter is just that – a linear filter.

From a policy analysis point of view there are two related papers making use of procedures proposed in the text. Andrle, Hlédik, Kameník, and Vlček (2009) demonstrate the use of filtering techniques and observables decompositions within the forecasting and policy system of the Czech National Bank. Further, Andrle (2012) uses the methods described below to analyze the output gap estimation methods, their revision properties and news effects.

☆The views expressed herein are those of the author and should not be attributed to the International Monetary Fund, its Executive Board, or its management.

[1]I would like to thank to Jan Brůha and participants of 8^{th} Dynare conference, Zurich, for comments. I have benefited from discussions with Jaromir Benes, Jan Brůha, Ondra Kameník, Antti Ripatti, Jan Vlček and Anders Warne. The core of approach introduced in the text was created for Czech National Bank core DSGE model's Forecasting and Policy Analysis System in 2007. First draft: October 2011

[2]In principle the terminology might be reversed, but the term shock decomposition is taken as an established and often used term.

The first part of the paper briefly introduces a linear state-space form, establishes its filter interpretation, its weights and observables decomposition of the filter and forecasts. The second part deals with missing observations and procedure to impose expert judgment on the filter exercise, together with adjustments to observable decompositions needed. The third section provides examples of the analysis and last section concludes.

2. Theoretical background

Structural (DSGE) models can be expressed as linear state-space models, see e.g. Christiano, Eichenbaum, and Evans (2005) or Smets and Wouters (2007). Semi-structural models e.g. Benes and N'Diaye (2004) or Andrle, Freedman, R. Garcia-Saltos, and Munandar (2009) also are of this form. The following state-space form is assumed:

$$\mathbf{Y}_t = \mathbf{Z}\mathbf{X}_t + \mathbf{D}\varepsilon_t \quad (1)$$

$$\mathbf{X}_t = \mathbf{T}\mathbf{X}_{t-1} + \mathbf{R}\varepsilon_t, \qquad \varepsilon \sim N(0, \mathbf{\Sigma}), \quad (2)$$

where $\mathbf{Y}_t$ is a $(n_y \times 1)$ vector of observed variables, $\mathbf{X}_t$ denotes a $(n_x \times 1)$ vector of transition, mostly latent, variables and ε_t stands for $(n_e \times 1)$ vector of stochastic shocks, normally distributed with covariance $\mathbf{\Sigma}$. Parameters of the model $\{\mathbf{Z}, \mathbf{D}, \mathbf{T}, \mathbf{R}, \mathbf{\Omega}\}$ are model dependent and are taken as given. For convenience and with little loss in generality a zero-mean process is assumed with eigenvalues of $\mathbf{T}$ inside the unit circle. Extensions for a non-stationary case are either obvious or discussed below when relevant. The structural nature of the shocks in ε_t is model specific and can range from theoretical structural shocks, to measurement errors. I assume that $\mathbf{RD}' = \mathbf{0}$ which makes structural and measurement shocks uncorrelated. Lag operator L defined as $x_{t-1} \equiv Lx_t$, is used to denote linear filters $\Phi(L) = \sum_{i=-\infty}^{\infty} \phi_i L^i$. Associated polynomials are denoted as $\Phi(z) = \sum_{i=-\infty}^{\infty} \phi_i z^i$.

Most often estimates of unobserved quantities $\{\mathbf{X}_t, \varepsilon_t\}$ are obtained by applying a Kalman filter (smoother) algorithm, see Durbin and Koopman (2001), inter alios. This paper focuses on a linear least squares or projection representation of the problem, while benefiting from the use of Kalman filter. For exposition purposes I denote the estimate of $\mathbf{X}_t$ using doubly-infinite sample as $\mathbf{X}_{t|\infty}$ and finite sample estimates are denoted as $\hat{\mathbf{X}}_t = \mathbf{X}_{t|T}$ for sample in periods $t \in [t_0, T]$.

2.1. *Filter representation of a state-space model*

The key to the analysis of the filter estimates and properties is the fact that the estimate $\mathbf{X}_{t|\infty}$ can be expressed literally as a linear, time-invariant two-sided (non-causal) filter:

$$\mathbf{X}_{t|\infty} = \mathbf{\Omega}(L)\mathbf{Y}_t = \sum_{i=-\infty}^{\infty} \Omega_i \mathbf{Y}_{t+i}. \quad (3)$$

The two-sided time-invariant filter $\mathbf{\Omega}$ is determined using the Wiener-Kolmogorov formula, see e.g. Whittle (1983),

$$\mathbf{\Omega}(z) = \mathbf{\Gamma}_{XY}(z)\mathbf{\Gamma}_{YY}(z)^{-1}, \quad (4)$$

where $\mathbf{\Gamma}_{XY}(z)$ and $\mathbf{\Gamma}_{YY}(z)$ are cross- and auto-covariance generating function of the model (1). The transfer function $\mathbf{\Omega}(z)$ is a function of model parameters and completely describes the properties of the filter both in time and frequency domain. Note that Kalman filter and Wiener-Kolmogorov filter can be applied to non-stationary processes, as discussed in more detail by Burridge and Wallis (1988) or Bell (1984).

The formula above represents a *observables decomposition.* The linear filter representation (3) enables the analyst to express the estimate of an unobserved variable as a linear combination of observed input data series. For practical implementation, the finite-sample version of the analysis needs to be developed and, in some cases, the finite-sample weights need to be computed.

Frequently, the model-consistent estimates of $\hat{\varepsilon}_t$ are used to provide a *historical shock decomposition*, i.e. a decomposition of observed data $\mathbf{Y}_t$ and transition variables $\hat{\mathbf{X}}_t$ into contributions of estimated structural shocks. The shock decomposition and observables decomposition are intimately connected and should be used as complements, not substitutes.

Assume for the moment availability of a doubly-infinite sample, then

$$\mathbf{X}_{t|\infty} = \mathbf{B}(L)\varepsilon_{t|\infty} \qquad \mathbf{Y}_{t|\infty} = \mathbf{C}(L)\varepsilon_{t|\infty} \tag{5}$$

In practice when only finite samples are available, the problem changes, but the main principle stays the same. A shock decomposition is given by

$$\mathbf{X}_{t|T} = \sum_{j=0}^{t-1} \mathbf{T}^j \mathbf{R}\varepsilon_{t-j|T} + \mathbf{T}^{(t-t_0)}\mathbf{X}_{t_0|T} \qquad \mathbf{Y}_{t|T} = \mathbf{Z}\mathbf{X}_{t|T} + \mathbf{D}\varepsilon_{t|T}. \tag{6}$$

Application of infinite, but convergent, time-invariant filter $\mathbf{\Omega}(z)$ on finite sample by means of Kalman filter/smoother implies application of time-variant filter $\mathbf{\Omega}(z)$ such that

$$\mathbf{X}_{t|T} = \sum_{\tau=t_0}^{T} \mathbf{\Omega}_{\tau|(t|T)}\mathbf{Y}_\tau + \mathbf{O}\mathbf{X}_0 \qquad \text{or, equivalently} \qquad \mathbf{X}_{t|T} = \sum_{j=1}^{n_y}\sum_{\tau=t_0}^{T} \mathbf{\Omega}_{\tau,j|(t|T)} Y_{\tau,j} + \mathbf{O}\mathbf{X}_0, \tag{7}$$

where $Y_{\tau,j}$ denotes the j-th element of $\mathbf{Y}_{\tau,j}$ and the term $\mathbf{O}\mathbf{X}_0$ captures the effect of initialisation of the Kalman filter/smoother. In case of stationary, zero-mean models it is common to set $\mathbf{X}_0$ and thus the term could be omitted from the analysis. The effect of initial conditions is, however, important in case of non-stationary models that not possessing unconditional mean of $\mathbf{Y}_t$ and $\mathbf{X}_t$.

The weights are time-varying in the case of a finite sample, different at each τ. The sequence of weights is conditional both on period of state estimation t and sample size T. The weights are therefore denoted using all this qualifying information, $\mathbf{\Omega}_{\tau|(t,T)}$.

2.2. Weights and Practical Implementation of the Method

It is quite clear that that a knowledge of weights implied by $\mathbf{\Omega}(z)$ or $\mathbf{\Omega}_\tau$ is a key input for the filter analysis. However, in most cases researches can dispense without actually knowing time-varying weights $\mathbf{\Omega}_\tau$ to study data revisions and news effects.

Frequency domain analysis is completely described by $\mathbf{\Omega}(z)$. The most straightforward way to obtain the weights associated with infinite, convergent filter $\mathbf{\Omega}(z)$ is evaluating the Fourier integral

$$\mathbf{\Omega}_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathbf{\Omega}(e^{-i\lambda})e^{+i\lambda k}\mathrm{d}\lambda. \tag{8}$$

A numerically more efficient procedure for calculating the filter weights associated with $\mathbf{\Omega}(z)$ is provided in a lucid article by Gomez (2006) based on Wiener-Kolmogorov theory, without explicit reliance on Kalman filter iterations. Time invariant weights, however, need to be converted to time-varying weights using projection arguments.

Importantly, time-varying weights for both Kalman filter and smoother are derived in an important paper by Koopman and Harvey (2003), making direct use of Kalman filter and smoother iterations. The time-varying weights obviously converge to time-invariant weights as the Kalman-gain and associated quantities stabilize. Although the derivation of the actual time-varying weights is provided in Koopman and Harvey (2003) and univariate examples (local linear trend) are provided, the great power of using the weights for observables decomposition comes in a multivariate context as proposed in this paper.

The need for time-varying weights in the case of a finite sample results from the unavailability of data to apply the time-invariant weights at the beginning and at the end of the sample. Weights are thus optimally adjusted to reflect the data unavailability, based on the stochastic process (the model) underlying the filter $\mathbf{\Omega}(z)$. The problem is a projection of the space of a finite filter on the space given by the results from infinite filter and is relatively easy to solve once the covariance generating function or spectral density for the $\mathbf{Y}$ is available, i.e.

$$\min_{\widehat{\mathbf{\Omega}}_{t,j}, j\in[t_0,\dots,T]} \quad ||\mathbf{X}_{t|\infty} - \mathbf{X}_{t|T}||^2 \tag{9}$$

$$= \quad ||\sum_{j=-\infty}^{-\infty} \mathbf{\Omega}_j \mathbf{Y}_j - \sum_{j=-t_0}^{-T} \widehat{\mathbf{\Omega}}_{j|t}\mathbf{Y}_j||^2 \tag{10}$$

$$= \quad \int_{-\pi}^{\pi} ||\mathbf{\Omega}(e^{-i\omega}) - \widehat{\mathbf{\Omega}}_t(e^{-i\omega})||^2 \mathbf{S}_{\mathbf{Y}}(\omega)\mathrm{d}\omega, \tag{11}$$

see Koopmans (1974), Christiano and Fitzgerald (2003) or Schleicher (2003). The solution re-weights the time invariant weights based on auto-covariance generating function.

An equivalence between the projection above and padding of the sample with back- and forecasts can be proved, as long as the sample is extended enough to let the weights converge. Interestingly, a heuristic method used by analysts to deal with end-point problems of linear filters is an optimal one, as long as the forecast is using the data-generating process of the data.

Implicitly, the Kalman filter produces model-based backcasts and forecast of observables and applies the infinite weights to such a padded sample. The Kalman filter, however, produces the padded sample based on the moments given by the state-space model, not directly the auto-covariance generating function of the data itself. As long as the forecast is long enough for the weights of the filter to converge, this procedure is equivalent to finite sample implementation of a doubly-infinite linear filter.Poor forecasting properties of the model then lead to inaccurate finite sample implementation and potentially larger revision variance.

Practical implementation. As time and software implementation constraints might be binding for the analyst, having a short cut is always beneficial. The actual comparison of two data sets of the same dimension can be implemented without the knowledge of explicit filter weights in most cases.

It certainly is the case in case of stationary time-invariant models with initial value of the filter given by unconditional mean of the state vector. Most calculations above are thus easy to carry out as is the case of shocks decomposition, for example. In the stationary case the time-varying finite sample weights are identical for any input data of course and one can (implicitly) rewrite the problem as a solution to associated least-squares problem, given by

$$\mathbb{X} = \mathbb{K}_{\Omega}\mathbb{Y}, \qquad \text{which implies} \qquad (\mathbb{X}^A - \mathbb{X}^B) = \mathbb{K}_{\Omega}(\mathbb{Y}^A - \mathbb{Y}^B), \tag{12}$$

where $\mathbb{Y} = [Y_1' \dots Y_T']'$, $\mathbb{X}$ stands for stacked state vectors in time and $\mathbb{K}_{\Omega}$ is a stacked version of weights implied by analysis of the revision problem (14). The simplicity of (12) greatly facilitates implementation of analysis of historical data revisions, news effects and decomposition of forecast revisions into observed data.[3]

The calculation imposes very small programming costs and requires only the access to the implementation of the Kalman filter and smoother recursion. Based on number of data points (or their groups) r under inspection, the analysis requires only $r + 1$ sequential runs of the Kalman smoother. In case of structural time series models the observations are grouped by variable names or by their type, e.g. nominal or real, etc. A total differential is carried out and due to linearity of the filter, a contribution analysis can be carried out.

In case of non-stationary models extra care must be exercised with regards to initial conditions of the filter and condition on these or on part of the sample. Obviously, analysts can always rely on explicit recursive algoritm for computing the time-varying weights derived in Koopman and Harvey (2003).

Decomposition into Observables. It is clear that (7) provides a decomposition of unobserved transition variables into contribution of individual observed data. Any individual piece of data, i.e. particular observation of variable j at period τ can be singled out and inspected for its influence on identified unobservables and structural shocks. Such decomposition not only enhances intuition about identification of shocks, but also allows to point out inconsistencies in the model. For instance, when a variable contributes only to trends or low frequency variation, instead of business cycle movements as the researcher expected an explicit decomposition quickly clarifies the issue. This setup also provides an ideal framework for the analysis of *historical data revisions* effects and *effects of new data releases* on identification of structural shocks.

2.3. Effects of Data Revisions

The effects of data revisions on estimates of structural shocks or unobserved variables, output gap for instance, is of great importance. The importance is due to a need of historical interpretation but also because of a great role of on initial conditions on forecast dynamics. We have

$$\mathbf{Y}_{t+h|t} = \mathbf{Z}\mathbf{T}^h\mathbf{X}_{t|t} = \mathbf{Z}\mathbf{T}^h\left(\sum_{\tau=t_0}^{t}\mathbf{\Omega}_{\tau|(t|t)}\mathbf{Y}_\tau + \mathbf{O}\mathbf{X}_0\right), \tag{13}$$

[3]Contrast the expression (12) with the matrix algebra of the Hodrick-Prescott/Leser filter, which has both Kalman filter and Wiener-Kolmogorov implementation, alongside with penalized least squares.

which provides insight into changes in judgement-free model forecast not just using a change in initial conditions vector $\mathbf{X}_{t|t}$, which is highly model-dependent and complex, but in terms of observed macroeconomic data. Again, this has proved to be useful in actual forecasting exercises with a DSGE model.

Let's denote two vintages of data of the same sample length as $\mathbf{Y}_t^A$ and $\mathbf{Y}_t^B$ and the resulting estimates of unobserved components as $\mathbf{X}_{t|T}^A$ and $\mathbf{X}_{t|T}^B$ with associated vector of revisions $\mathbb{R}_t = \mathbf{X}_{t|T}^A - \mathbf{X}_{t|T}^B$. It follows then that

$$\mathbb{R}_{t|T} = \sum_{\tau=t_0}^{T} \mathbf{\Omega}_{\tau|(t|T)} \left(\mathbf{Y}_\tau^A - \mathbf{Y}_\tau^B\right) = \sum_{j=1}^{n_y} \sum_{\tau=t_0}^{T} \mathbf{\Omega}_{\tau,j|(t|T)} \left(Y_{\tau,j}^A - Y_{\tau,j}^B\right), \tag{14}$$

where it is assumed that initialization of the filter stays unchanged and thus does not contribute to change in the estimate. Note that in case of non-stationary models where the initial conditions are identified as function of the data, e.g. as in Rosenberg (1973), extra care must be taken to account for the change in this factor.

A decomposition of the h-step ahead forecast due to data revisions follows trivially from (13) and (14). Below it is demonstrated how to decompose the difference in forecast due to new observations, e.g. $\nu_{t+h|t/k} = \mathbf{Y}_{t+h|t+k} - \mathbf{Y}_{t+h|t}$, where h is the horizon of the forecast and k determines the upgrade of the information sets in periods, $k \geq h$.

2.4. *Effects of New Data Releases*

Understanding the effects of new data releases, e.g. moving k-period forward in time, is very useful for real-time forecasting and policy analysis. First, it is important to analyze how particular pieces of new data contribute to newly identified unobserved components, hence quantify the *news effect*. Second, due to two-sided nature of the filter the identification of structural shocks in previous periods can be changed by extension of the sample, as it is well known.

The news effect is not just the effect of new pieces of data. The news effect is the additional information conveyed by new data. The news in observation is related to a prediction error of the model. Obviously, if $\mathbf{Y}_{t+1|t} = \mathbf{Y}_{t+1}$ new data release conveys no new information beyond what has been already known. The news effect depends how the forecast prediction error translates into estimates of structural shocks and unobserved variables:

$$\mathbb{N}_{t|T+1} = \sum_{\tau=t_0}^{T+1} \widehat{\mathbf{\Omega}}_{\tau|T+1} \left(\mathbf{Y}_{\tau|T} - \mathbf{Y}_{\tau|T+1}\right) = \widehat{\mathbf{\Omega}}_{T+1|T+1} \left(\mathbb{E}\mathbf{Y}_{T+1|T} - \mathbf{Y}_{T+1}\right). \tag{15}$$

There are several facts worth noting. First, by expanding the sample the weights applied to a particular observation differ. It is not affected by the values of the observed data, the change is only due to finite-sample implementation of infinite two-sided filter, implied by ACGF of the model. Second, the problem of analyzing news can be translated to the problem of data revision analysis by 'aligning' the increased sample with the old (shorter) sample padded with a model prediction for observables, $\mathbf{Y}_{T+1|T} = \mathbf{ZTX}_{T|T}$. It is easy to demonstrate that applying Kalman smoother to such an augmented dataset yields identical results as smoothing the original data.[4]

As explained beforehand, this is how the Kalman filter/smoother augments the dataset to apply the time invariant weights stemming from $\mathbf{\Omega}(z)$ at infinite data sample. It is known that forecasts and backcasts need to be provided for implementation of two-sided filter explicitly or implicitly in Wiener-Kolmogorov approach. There is a unique mapping from $\mathbf{\Omega}(z)$ implied weights and $\mathbf{\Omega}_{\tau|T}$ depending on model's parameters and sample size.

Another intuitive form of (15) is to use the doubly-infinite sample analogue, assuming all past data are kept unchanged. This defines the revision due to new data availability as

$$\mathbb{N}_{t|\infty} = \sum_{j=0}^{\infty} \mathbf{\Omega} \left(\mathbb{E}\mathbf{Y}_{j|T} - \mathbb{E}\mathbf{Y}_{T+j}\right) = \sum_{j=T+1}^{\infty} \mathbf{\Omega}_j \nu_{T+j} = [\mathbf{\Omega}(z)]_+ \, \nu_T, \tag{16}$$

which makes clear that revisions $\mathbb{N}_{t|\infty}$ and revision variance $\operatorname{var} \mathbb{N}_{t|\infty}$ are one-sided moving averages of prediction errors ν_t of the process $\mathbf{Y}_t$ defined in (1). See Pierce (1980) for early discussion of revisions dynamics in univariate models and Maravall (1986) univariate ARMA processes analyzes closed form solutions of revision variance.

[4]This is true when the new observations are not used to update initial conditions of the model in non-stationary case. The proof follows simply from inspection of Kalman smoother recursion, see Durbin and Koopman (2001), and the fact that the prediction error in this case is trivially zero.

The procedure above is easy to extend to longer horizons to analyze difference in two forecasts due to change in the observed data, $\nu_{t+h|t/k} = \mathbf{Y}_{t+h|t+k} - \mathbf{Y}_{t+h|t}$ with $k < h$. The analysis requires running two Kalman smoothers with suitably extended observables by model-based forecast. In stationary, time-invariant models the extension by the consistent forecast does not add any information.

2.5. Value of an Observation

For interpretative purposes and model design it is key to understand the role of observables and the 'value of an observation'. More specifically, it is necessary to answer the following question: "How adding new observables (variables) changes estimates of structural shocks and state variables, and what is the impact on the forecast?" Answering this question in case of one variable added is trivial and consists of comparing unobserved states estimates using the data with and without the relevant series (or data point). In case of richer information set the procedure is simple and amounts to converting the problem into a data revision one. First information set is the actual observables considered and the other consists of conditional estimates for the exact pattern of new observables, using the smaller set of observations.

3. Missing Observations and Imposing Judgement on the Filter

The use of expert judgment is a part of practical modeling and forecasting. It is crucial to be able to impose the expert judgment in a consistent way both in the forecasting step, as well as during the estimate of the initial state of the economy – during the filtering step.

This section discusses missing observations and simple techniques how to impose judgement on estimation of structural shocks with relationship to filter decompositions into observables and analysis of news. Briefly, missing observations and judgemental adjustments impose no obstacles on procedures suggest above and fit naturally to the framework.

Missing observations are not just a fact of life, but a useful tool for imposing judgment on the filter. The need for transparent and coherent implementation of expert judgement follows from the need of the expert judgment itself due to stylised nature of DSGE models. The literature focuses usually on expert judgement and conditioning on the forecast horizon, see e.g. Benes, Binning, and Lees (2008) for instance. Often, the analyst knows also much more about the history and interpretation of particular events that the model cannot cope with without extra information in the filtering step or data transformation.

3.1. Missing observations and time-varying parameters

Handling missing observation is not trivial using classical Wiener-Kolmogorov methods, but Kalman filter/smoother is well equipped to handle the problem, see e.g. Durbin and Koopman (2001) or Harvey (1989). Standard treatment of missing observations is to set up the time-varying system of in terms of the measurement equations. Let $\mathbf{W}_t$ denote a time-varying selection matrix relating unavailable full data vector $\mathbf{Y}_t$ to actually observed linear combination of data $\mathbf{Y}^w$ such that $\mathbf{Y}^w = \mathbf{W}_t\mathbf{Y}_t$. By left-multiplying the measurement equation in (1) by $\mathbf{W}_t$ a time-varying analogue is obtained of the form

$$\mathbf{Y}^w_t = \mathbf{Z}^w_t\mathbf{X}_t + \mathbf{D}^w_t\varepsilon_t, \tag{17}$$

keeping the transition equation time-invariant and unmodified. In case of no observations for a particular time period, the prediction error is undefined and the updating step of the Kalman filter is just a propagation of the last state.[5]

In finite samples the implied filter is time-varying even for time-invariant state-space forms. Adding time-varying parameters of the model results in different set of weights with respect to $\mathbf{Y}^w_t$, i.e.

$$\mathbf{X}_{t|T} = \sum_{\tau=t_0}^{T} \mathbf{\Omega}^w_{\tau|(t|T)}\mathbf{Y}^w_\tau + \mathbf{O}\mathbf{X}_0, \tag{18}$$

[5]The time-varying feature of the model can be explicitly provided by the researcher or implicitly carried out by the computer code by inspection of pattern of missing data. An alternative approach is to make the distribution of measurement shocks time-varying, fill-in missing data with arbitrary information and impose infinite variance on the measurement error for that particular observation. Such approach is, however, prone to numerical instability.

where, conditioned on the sequence $\{\mathbf{W}_t\}_{t_0}^T$ the weights are deterministically related to weights in the problem with constant parameters.

Decomposition of $\mathbf{X}_{t|T}$ into actually observed data is thus feasible. The analysis of change in the identified state variables and structural shocks due to historical data revisions or new observations available can be carried out using a modified version of (14) and (15). The key requirement though is that $\mathbf{W}_t^A = \mathbf{W}_t^B\ \forall t$.

In case of finite sample analysis of the filter adopting a linear time-varying state-space model of the form

$$\mathbf{Y}_t = \mathbf{Z}_t\mathbf{X}_t + \mathbf{D}_t\varepsilon_t \tag{19}$$

$$\mathbf{X}_t = \mathbf{T}_t\mathbf{X}_{t-1} + \mathbf{R}_t\varepsilon_t, \qquad \varepsilon \sim N(0, \mathbf{\Sigma}), \tag{20}$$

does not affect the analysis if the conditions stated above are satisfied. Obviously the infinite-sample time-invariant analysis based upon $\mathbf{\Omega}(z)$ is no longer relevant description of the filter properties, except in case of missing observations with random pattern.

3.2. Imposing Judgement

By 'judgement' we understand imposing additional information on the estimation of $\{\mathbf{X}_t\}_{t=1}$ beyond information contained in the model (1) and available data $\{\mathbf{Y}\}_{t=1}^T$. Most often, the judgmental extra information

The method of imposing judgment suggested below builds the approach of Doran (1992) and insight of Theil and Goldberger (1961), resulting in augmenting the model (1) with a set of stochastic linear restrictions. It can be understood easily as dummy-observations priors for state variables.

Assume there is an extra information available expressed as set of n_k stochastic restrictions

$$\mathbf{k}_t = \mathbf{K}_t\mathbf{X}_t + \xi \qquad \xi \sim N(\mathbf{0}, \mathbf{\Sigma}_\xi), \tag{21}$$

which can be used to augment the time varying version of the model (19) to obtain a judgment-adjusted model

$$\begin{bmatrix}\mathbf{Y}_t\\ \mathbf{k}_t\end{bmatrix} = \begin{bmatrix}\mathbf{Z}_t\\ \mathbf{K}_t\end{bmatrix}\mathbf{X}_t + \begin{bmatrix}\mathbf{D}_t & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\begin{bmatrix}\varepsilon_t\\ \xi_t\end{bmatrix} \tag{22}$$

$$\mathbf{X}_t = \mathbf{T}_t\mathbf{X}_{t-1} + \mathbf{R}_t\varepsilon_t, \tag{23}$$

or simply defining a new measurement equation as $\mathbf{Y}_t^\star = \mathbf{Z}_t^\star\mathbf{X}_t + \mathbf{R}_t^\star\varepsilon_t^\star$ with obvious definition of new variables, parameters and covariance matrix of extended shocks $\mathbf{\Sigma}^\star$.

In practical implementation it is equivalent to define new state variables satisfying the constraint (21) and augment the state vector, defining also new measurement variables via an identity matrix. With no loss of generality having $\mathbf{K}_t = \mathbf{K}$ allows then to use this approach of imposing judgment with an implementation of Kalman filter/smoother supporting only missing observations, not time-varying parameters per se.

The interpretation of judgement as direct observations on structural shocks and state variables (e.g. output gap, cost push shocks, etc.) with arbitrary degree of precision, given by variance of the measurement error ξ_j is very intuitive and allows researches to provide extra information for the model to interpret historical data.

Exact, or 'hard', filter tunes were applied in Andrle et al. (2009) in the case of structural, DSGE model or Benes, Clinton, Garcia-Saltos, Johnson, Laxton, Manchev, and Matheson (2010) in the case of semi-structural output gap estimation. Making the filter tune explicitly stochastic provides more leeway for prior judgement.

Observables contribution analysis with restrictions. Observables decomposition of the filter is possible, again, under conditions specified for time-varying systems or systems with missing observations. The pattern of time variation or missing observations must be identical for the two datasets $\mathbf{Y}_t^{A,\star}$ and $\mathbf{Y}_t^{B,\star}$ in all periods. To align judgement-free problem with the new information-augmented problem requires adding observations at periods when judgment was applied for particular variables, the conditional estimates of relevant priors $E[\mathbf{k}_t|\{\mathbf{Y}_t^A\}_{t=t_0}^T]$. Again, due to law of iterated expectations estimates using thus augmented database $\mathbf{Y}_t^{A,\star}$ are identical to those based on $\mathbf{Y}_t^A$.

Restricted forecasts. Obviously the use of time-varying restrictions could be used also for producing a forecast for $t + h$ using with the initial condition $\mathbf{X}_{t|t} \sim N(\mathbf{X}_{t|t}, \mathbf{0})$ and running the Kalman smoother with restrictions imposed at forecast horizon. Unless the initial state variance is set to zero, the restrictions in the forecast horizon would change the historical estimates of the state variables and shocks.

4. Examples & Applications

Bellow several examples illustrated the filter analysis suggested in the text are provided. The first example decomposes identified unobservable 'flexible-price' output gap in the DSGE model by Smets and Wouters (2007) into observed variables, grouped. All examples were coded in Matlab.

4.1. Flexible price output gap and real marginal costs in Smets and Wouters (2007)

In their well known contribution Smets and Wouters (2007) construct a DSGE model with parameter estimated using Bayesian-likelihood approach with strong econometric focus of the model. In their paper the authors present the structure of the model, parameter estimates, selected impulse-response functions and a decomposition of inflation and GDP growth into contributions of structural shocks. To illustrate the decompositions suggested above this paper decomposes 'flexible-price' output gap and real marginal costs into contributions of observed time series, using the filter (7) with weights implied by the structural model.[6]

As can be seen from Fig. 1 the implicit output gap, also entering a policy rule, displays a larger degree of persistence and just a small tendency to mean reversion. The filter places optimal (in mean-square sense) weights on available data that may corroborate similar pieces of information (e.g. real wages or consumption) and are linked by the theoretical structure of the model. Little weight is put directly on inflation and interest rates, most cyclical contributions are coming from observing employment and output. Interestingly, consumption and real-wage observables contributions seem to be trending, as well as the real consumption series 'gap' defined as real consumption deflated by level of technological progress.

A possible counterpart to inflation shock-decomposition in Smets and Wouters (2007) could be inspection of real marginal costs in terms of the importance placed by the model-based filter.

Looking at Fig. 2 one can see that given the default information set of the paper, there is a little weight put on inflation and that real marginal cost is dominated by information in hours worked, output and real wages. The real marginal costs feature a pronounced downward-sloping trend. Presumably this is a consequence of potential misspecification of long-run component of the model, i.e. single technology trend and, not accounting for an implicit inflation target of the Federal reserve.

Using Smets and Wouters (2007) model as an example for imposing judgement is possible, yet there is little leeway to do so as the model features identical number of observed variables as stochastic shocks. In a limiting case of known initial conditions, the shock estimation problem would cease to be a least squares, but a linear system of equations. There are not many degrees of freedom for efficiently imposing judgement, just small shifts of initial values.

One can also carry out a time- and frequency domain analysis of $\mathbf{\Omega}(z)$ as is usually the case for univariate filters. Fig. 3 depicts weights for the model and a gain of the filter for the cost-push (markup) shock with respect to inflation. Confirming the economic intuition both weights and the gain of the model's filter suggest that the markup shock to inflation is driven mostly by high-frequency variation of the series. On the other hand it seems that from the weight functions of price of capital (pk) with respect to observed real interest rate (robs) and of the flex-price output gap (ygap) with respect to observed hours worked (labobs) may be less insightful than direct filter contribution analysis exercises suggested in the text.

Looking at weights implied by the model one sees another peculiar nature of the model – steady-state weights (and only those) are one sided only. This reflects the stochastic rank of the model coincident with number of observables. In a steady-state version of the model, where initial conditions effects is zero in the limit, the model is simply an inverted system of equations and not a least squares signal extraction problem.

[6]Public provision of the codes, data and estimation results by F. Smets and R. Wouters is acknowledged. The calculations in the text do not rely on Dynare version of these, but great care has been put into precise replication.

Figure 1: Flexible-price output gap in SW07

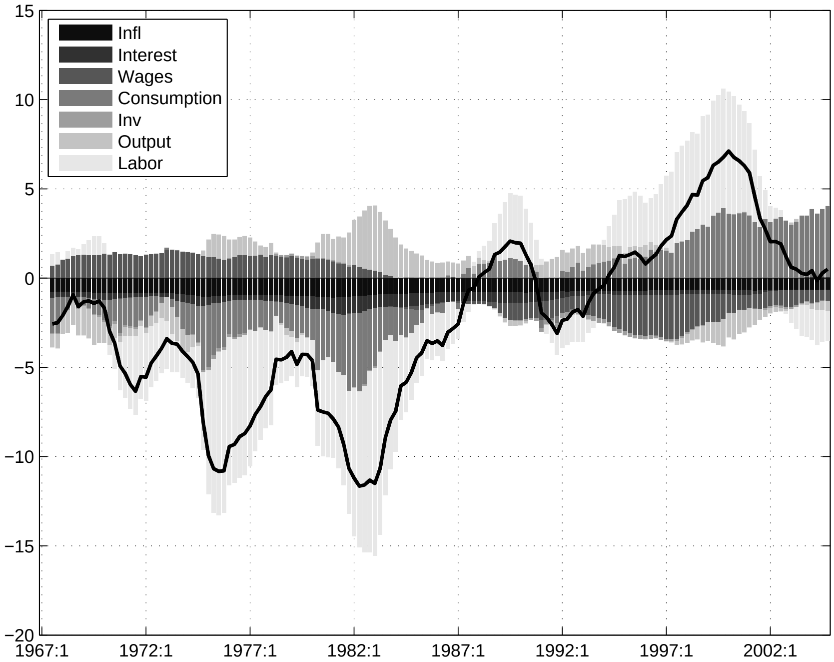


Figure 2: Real marginal costs in SW07

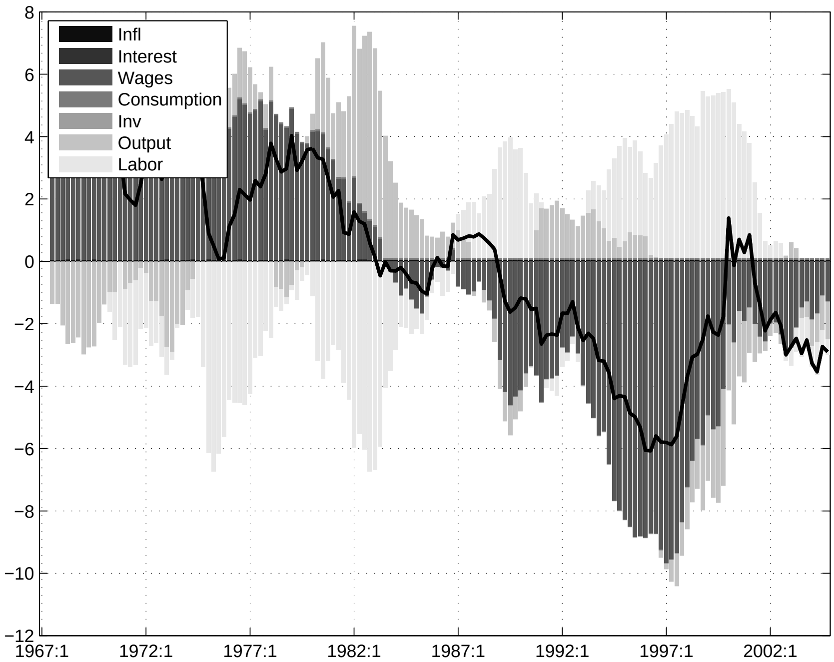
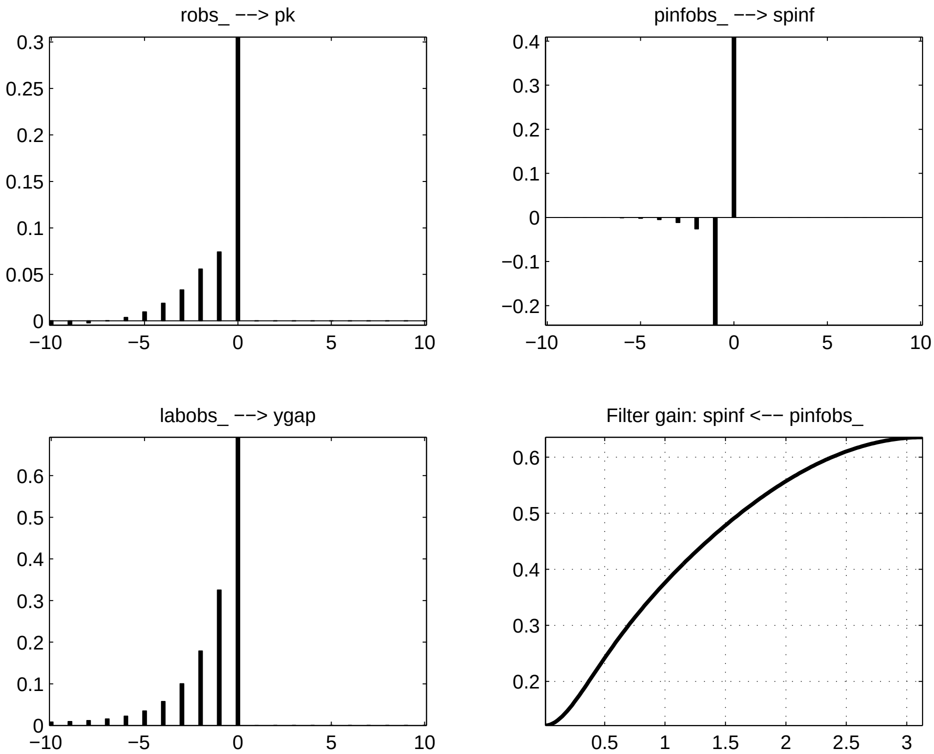


Figure 3: Filter weights & gain SW07



5. Conclusion

This short paper advances methods and examples of filter analysis of structural DSGE models. The analysis above proved useful for model development and design, forecasting and policy analysis. Is is also easy to implement in practice.

Explicit understanding of the estimation of unobserved structural shocks in terms of linear filter is beneficial and can increase insight into economic structure of the model. The method allows researches to understand how, for instance, the output gap, preference or technology shocks were identified from the data and what are the influential observables, dates. Such analysis serves as a complement, not a substitute, for historical shock analysis.

Decomposition of unobserved quantities into contributions of observed data is also useful for historical data revisions and successive re-estimates of unobservables as new pieces data arrive. Explicit decomposition are useful namely in case of medium- to large-scale DSGE forecasting models. These decomposition are essential for economic story telling.

The paper also illustrates a simple way of imposing prior insight (expert judgement) on the identified unobserved quantities. As imputation of extraneous, expert judgement is no news in the forecasting process with structural models, there is no reason to expect that identification of initial state of the (model) economy and structural shocks should be judgement free. Despite the impressive progress in model development in recent years, structural economic models are and have to be rather stylised depictions of reality.

The approach suggested in this paper has been recently applied to a prominent unobserved variable, the output gap, where the revisions and misunderstanding of factors identifying the estimates can lead to serious policy errors, see Andrle (2012).

References

Andrle, M. (2012). What is in your output gap. Technical report, IMF Working Paper, forthcoming.

Andrle, M., C. Freedman, D. L. R. Garcia-Saltos, D. Hermawan, and H. Munandar (2009, November). Adding Indonesia to the Global Projection Model. Technical report, IMF Working Paper 09/253.

Andrle, M., T. Hlédik, O. Kameník, and J. Vlček (2009, December). Implementing the new structural model of the Czech national bank. Working paper no. 2, Czech National Bank.

Bell, W. (1984). Signal extraction for nonstationary time series. *Annals of Statistics 12*, 646–664.

Benes, J., A. Binning, and K. Lees (2008). Incorporating judgement with DSGE models. Discussion paper, dp2008/10, Reserve Bank of New Zealand.

Benes, J., K. Clinton, R. Garcia-Saltos, M. Johnson, D. Laxton, P. Manchev, and T. Matheson (2010, December). Estimating potential output with a multivariate filter. Working Paper 10/285, International Monetary Fund, Washington DC.

Benes, J. and P. N'Diaye (2004, March). A multivariate filter for measuring potential output and the nairu. Working Paper 04/45, International Monetary Fund, Washington DC.

Burridge, P. and K. Wallis (1988). Prediction theory for autoregressive-moving average processes. *Econometric Reviews 7*(1), 65–95.

Christiano, L., M. Eichenbaum, and C. Evans (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy 113*(1), 1–45.

Christiano, L. and T. Fitzgerald (2003). The bandpass filter. *International Economic Review 44*(2), 435–465.

Doran, H. (1992). Constraining Kalman filter and smoothing estimates to satisfy time-varying restrictions. *Review of Economics and Statistics 74*, 568–572.

Durbin, J. and S. Koopman (2001). *Time Series Analysis by State Space Methods*. New York: Oxford UP.

Gomez, V. (2006). Wiener-Kolmogorov filtering and smoothing for multivariate series with state-space structure. *Journal of Time Series Analysis 28*(3), 361–385.

Harvey, A. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

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

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

Maravall, A. (1986). Revisions in ARIMA signal extractions. *Journal of American Statistical Association 81*(395), 736–740.

Pierce, D. (1980). Data revisions in moving average seasonal adjustment procedures. *Journal of Econometrics 14*(1), 95–114.

Rosenberg, B. (1973). Random coefficient models: the analysis of a cross-section of time series by stochastically convergent parameter regression. *Annals of Economic and Social Measurement 2*, 399–428.

Schleicher, C. (2003, January). Kolmogorov-Wiener filters for finite time series. Technical report, University of British Columbia, mimeo.

Smets, F. and R. Wouters (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *American Economic Review 97*(3), 586–606.

Theil, H. and A. Goldberger (1961). On pure and mixed statistical estimation in economics. *International Economic Review 2*(1), 65–78.

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

6. APPENDIX

This appendix provides the formulas for the Wiener-Kolmogorov expression for time invariant weights and also result of Gomez (2006).

In case of doubly inifinite sample the estimate of unobserved components can be obtained using Wiener-Kolmogorov formula

$$\mathbf{X}_{t|\infty} = \mathbf{\Omega}(L)\mathbf{Y}_t = \mathbf{\Gamma}_{\mathbf{XY}}(L)\mathbf{\Gamma}_{\mathbf{Y}}^{-1}(L)\mathbf{Y}_t = \sum_{j=-\infty}^{\infty} \mathbf{\Omega}_j \mathbf{Y}_{t+j}, \tag{24}$$

in case of doubly-infinite data sample, where $\mathbf{\Gamma}_Y(z), \mathbf{\Gamma}_{XY}(z)$ stands for multivariate auto-covariance generating function (ACGF) of the state-space model, e.g.

$$\mathbf{\Gamma}_{\mathbf{XY}}(z) = (\mathbf{I} - \mathbf{T}z)^{-1}\mathbf{G}\mathbf{\Sigma}_\varepsilon\mathbf{G}'(\mathbf{I} - \mathbf{T}'z^{-1})^{-1}\mathbf{Z}' \tag{25}$$

$$\mathbf{\Gamma}_{\mathbf{Y}}(z) = \mathbf{Z}(\mathbf{I} - \mathbf{T}z)^{-1}\mathbf{G}\mathbf{\Sigma}_\varepsilon\mathbf{G}'(\mathbf{I} - \mathbf{T}'z^{-1})^{-1}\mathbf{Z}' + \mathbf{H}\mathbf{\Sigma}_\varepsilon\mathbf{H}'. \tag{26}$$

Note the usefullness of the , which clearly defines the filter and its transfer function. Analysts can carefuly study links from the data to unobservables accross frequencies.

Regarding a time domain expression for the time invariant weights, Gomez (2006) shows that for (**??**) the weights $\mathbf{\Omega}_j$ follow as (adjusted to our model specification)

$$\mathbf{\Omega}_0 = \mathbf{P}(\mathbf{Z}'\mathbf{\Sigma}^{-1} - \mathbf{L}\mathbf{R}_{|\infty}\mathbf{K}) \tag{27}$$

$$\mathbf{\Omega}_j = (\mathbf{I} - \mathbf{P}\mathbf{R}_{|\infty})\mathbf{L}^{-j-1}\mathbf{K} \qquad j < 0 \tag{28}$$

$$\mathbf{\Omega}_j = \mathbf{P}\mathbf{L}'^{j}(\mathbf{Z}'\mathbf{\Sigma}^{-1} - \mathbf{L}'\mathbf{R}_{|\infty}\mathbf{K}) \qquad j > 0, \tag{29}$$

where $\mathbf{L} \equiv \mathbf{T} - \mathbf{K}\mathbf{Z}$, $\mathbf{K}$ 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 equation (DARE) associated with steady-state solution of the Kalman filter. $\mathbf{R}_{|\infty}$ is a solution to Lyapounov equation $\mathbf{R}_{|\infty} = \mathbf{L}'\mathbf{R}_{|\infty}\mathbf{L} + \mathbf{Z}(\mathbf{Z}\mathbf{P}\mathbf{Z}' + \mathbf{H}\mathbf{\Sigma}_\varepsilon\mathbf{H}')^{-1}\mathbf{Z}'$, associated with the steady-state Kalman smoother solution. $\mathbf{R}_{|\infty}$ is the steady-state variance of the process $r_{t|\infty}$ in the backward recursion $X_{t|\infty} = X_{t|t-1} + Pr_{t|\infty}$, where in finite-data smoothing r_{t-1} is a weighted sum of those innovations (prediction errors) coming after period $t-1$. Finally $\mathbf{\Sigma} = \mathbf{Z}(\mathbf{Z}\mathbf{P}\mathbf{Z}' + \mathbf{H}\mathbf{\Sigma}_\varepsilon\mathbf{H}')$. All the quantities introduced are easily also obtained after carrying out the Kalman filter iterations (which are data independent) until convergence.