

IMF Working Paper

Understanding DSGE Filters in Forecasting and Policy Analysis

Michal Andrle

IMF Working Paper

Research department

Understanding DSGE Filters in Forecasting and Policy Analysis

Prepared by Michal Andrle¹

Authorized for distribution by Benjamin Hunt

May 2013

This Working Paper should not be reported as representing the views of the IMF.

The views expressed in this Working Paper are those of the author(s) and do not necessarily represent those of the IMF or IMF policy. Working Papers describe research in progress by the author(s) and are published to elicit comments and to further debate.

Abstract

This paper introduces methods that allow analysts to (i) decompose the estimates of unobserved quantities into observed data, (ii) to better understand revision properties of the model, and (iii) to impose subjective prior constraints on path estimates of unobserved shocks in structural economic models. For instance, a decomposition of the flexible-price output gap, or a technology shock, into contributions of output, inflation, interest rates, and other observed variables' contribution is feasible. The intuitive nature and analytical clarity of the suggested procedures are appealing for policy-related and forecasting models.

JEL Classification Numbers: C10, E50

Keywords: filters; DSGE models; observable decomposition

Author's E-Mail Address: mandrle@imf.org

¹ I would like to thank Jan Bruha and participants of 8th Dynare conference, Zurich, for comments. I have benefited from discussions with Jaromir Benes, Jan Bruha, Ondra Kamenik, Antti Ripatti, Jan Vlcek and Anders Warne. The core of the 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

Contents

Page

I.	Introduction	3
II.	Theoretical background	4
	A. Filter representation of a state-space model	5
	B. Weights and Practical Implementation of the Method	6
	C. Effects of Data Revisions	9
	D. Effects of New Data Releases	10
III.	Missing Observations and Imposing Judgement on the Filter	11
	A. Missing observations and time-varying parameters	12
	B. Imposing Judgement	13
IV.	Examples and Applications	14
	A. Flexible price output gap and real marginal costs in Smets and Wouters (2007)	15
V.	Conclusion	19
	References	20
VI.	APPENDIX	22

Tables

Figures

1.	Flexible-price output gap in SW07	16
2.	Real marginal costs in SW07	16
3.	Filter weights and gain SW07	18

I. INTRODUCTION

This paper puts forth a decomposition of unobserved quantities into the 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 tool is that it can reveal how observed data— such as output, inflation, interest rates, etc.— contribute to the estimated total factor productivity shock, estimate of the output gap, estimated preference shocks, or other unobserved variables. 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.¹ An example of an important variable is the output gap, where one can clearly trace contributions of input data using methods presented below.

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

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 the theory of linear multivariate filters, both in time and frequency domains. The key is to recall that the Kalman filter is just that – a linear multivariate filter.

From a policy analysis point of view, there are two companion papers making use of procedures proposed in the text. [Andrle and others \(2009b\)](#) demonstrate the use of filtering techniques and observables decompositions within the forecasting and policy system of a central bank. Further, [Andrle \(2012\)](#) uses the approach described below to analyze several output gap estimation methods, their revision properties, and news effects.

The first part of the paper briefly introduces a linear state-space form and establishes its filter interpretation, its weights, and observables decomposition of the filter and forecasts. The second part deals with missing observations and illustrates a procedure to impose expert judgment on the filter exercise, together with adjustments required to carry out the observable

¹In principle, the terminology might be reversed, but the term shock decomposition is taken as an established and often used term.

decompositions. The third section provides examples of the analysis using the medium-scale DSGE model by [Smets and Wouters \(2007\)](#). The final section concludes.

II. THEORETICAL BACKGROUND

Structural models can be expressed as linear state-space models, see e.g. [Christiano, Eichenbaum, and Evans \(2005\)](#) or [Smets and Wouters \(2007\)](#). Many semi-structural models –e.g. [Benes and N’Diaye \(2004\)](#) or [Andrle and others \(2009a\)](#)– also are of this form. The following state-space form is assumed:

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

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

where \mathbf{Y}_t is an $(n_y \times 1)$ vector of observed variables, \mathbf{X}_t denotes an $(n_x \times 1)$ vector of transition variables, and ε_t stands for an $(n_e \times 1)$ vector of stochastic shocks, normally distributed with covariance $\mathbf{\Sigma}$. Parameters of the model $\{\mathbf{Z}, \mathbf{D}, \mathbf{T}, \mathbf{R}, \mathbf{O}\}$ are model dependent, and are taken as given. For convenience and with little loss in generality, a zero-mean process is assumed, with all 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. It is assumed that $\mathbf{R}\mathbf{D}' = \mathbf{0}$, which makes structural and measurement shocks uncorrelated. Lag operator L , defined as $x_{t-1} \equiv Lx_t$, is used to specify linear filters $\Phi(L) = \sum_{i=-\infty}^{\infty} \phi_i L^i$. The polynomials associated with the filter are denoted as $\Phi(z) = \sum_{i=-\infty}^{\infty} \phi_i z^i$.

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

A. 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:

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

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

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

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

The formula above represents an *observables decomposition*. The linear filter representation (3) enables the analyst to express the estimate of an unobserved variable as a linear combination of an observed input data series. For practical implementation purposes, 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.

For the moment, let's assume the availability of a doubly-infinite sample. In such a case, the unobserved and observed variables can be expressed as a function of past structural shocks – a *shock decomposition*:

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

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

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

Application of an infinite, convergent, time-invariant filter $\mathbf{O}(z)$ on a finite sample by means of the Kalman filter/smoothen implies application of time-variant filter $\mathbf{O}(z)$, such that

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

where $Y_{\tau,j}$ denotes the j -th element of \mathbf{Y}_{τ} and the term $\mathbf{O}\mathbf{X}_0$ captures the effect of initialization of the Kalman filter/smoothen. In the case of stationary, zero-mean models $\mathbf{X}_0 = 0$, and thus the term can be omitted from the analysis. The effect of initial conditions is, however, important in the case of non-stationary models, or in the case of other value of the initial state vector.

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

B. Weights and Practical Implementation of the Method

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

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

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

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

Notably, the time-varying weights for both the Kalman filter and smoothen are derived in an important paper by [Koopman and Harvey \(2003\)](#), making direct use of the Kalman filter and smoothen 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 demonstrated 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 $\mathbf{O}(z)$. The problem is a projection of the space of a finite filter onto 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.

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

$$= \left\| \sum_{j=-\infty}^{-\infty} \mathbf{O}_j \mathbf{Y}_j - \sum_{j=-t_0}^{-T} \widehat{\mathbf{O}}_{j|t} \mathbf{Y}_j \right\|^2 \quad (10)$$

$$= \int_{-\pi}^{\pi} \|\mathbf{O}(e^{-i\omega}) - \widehat{\mathbf{O}}_t(e^{-i\omega})\|^2 \mathbf{S}_Y(\omega) d\omega, \quad (11)$$

see [Koopmans \(1974\)](#), [Christiano and Fitzgerald \(2003\)](#), or [Schleicher \(2003\)](#). The solution re-weights the time invariant weights using the model's auto-covariance generating function.

An equivalence between the projection above and padding of the sample with backcasts and forecasts can be established as long as the sample is extended enough to let the weights converge. Interestingly, a heuristic method used by many analysts to deal with end-point problems of linear two-sided filters is the optimal one, as long as the forecast is produced 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 auto-covariance generating function 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.

(a) Practical implementation As time and software implementation constraints might be binding for the analyst, having a shortcut 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 for stationary time-invariant models with the initial value for the filter given by the unconditional mean of the state vector. Most calculations above are thus easy to carry out, as easy as computing a shocks decomposition, for example. In the stationary case, the time-varying finite sample weights are identical for any input data and one can think of the problem as a solution to an associated least-squares problem, given by

$$\mathbb{X} = \mathbb{K}_O \mathbb{Y}, \quad \text{which implies} \quad (\mathbb{X}^A - \mathbb{X}^B) = \mathbb{K}_O (\mathbb{Y}^A - \mathbb{Y}^B). \quad (12)$$

Here, $\mathbb{Y} = [Y'_1 \dots Y'_T]'$, \mathbb{X} stands for stacked state vectors across time, and \mathbb{K}_O is a stacked version of weights implied by analysis of the revision problem (14), below. The simplicity of (12) greatly facilitates the implementation and the analysis of historical data revisions, news effects, and decomposition of forecast revisions into observed data.²

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

In the 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 algorithm 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 contributions of individual observed data. Any individual piece of data, i.e. a 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 one 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 contrary to researcher expectations,

²Contrast the expression (12) with the matrix algebra of the Hodrick-Prescott/Leser filter, which has both Kalman filter and Wiener-Kolmogorov implementation, along with with penalized least squares.

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.

C. Effects of Data Revisions

The effects of data revisions on the estimates of structural shocks or unobserved variables, the output gap for instance, is of great importance. Its importance is due to a need for historical interpretation but also because of the great role of initial conditions on forecast dynamics. One can express h -step ahead forecast as a function of all observables:

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

which provides insight into changes in a judgement-free model forecast not just using a change in initial conditions vector $\mathbf{X}_{t|t}$ –which is highly model-dependent and complex to decode– 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 an associated vector of revisions $\mathbb{R}_t = \mathbf{X}_{t|T}^A - \mathbf{X}_{t|T}^B$. It is clear that

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

where it is assumed that initialization of the filter stays unchanged, and thus does not contribute to a change in the estimate. Note that in the case of non-stationary models –where the initial conditions are identified as a function of the data e.g. as in [Rosenberg \(1973\)](#) some 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). It is demonstrated below how to decompose the difference in forecast due to new observations, e.g. $v_{t+h|t/k} = \mathbf{Y}_{t+h|t+k} - \mathbf{Y}_{t+h|t}$, where h is the horizon of the forecast and k determines the upgrade of the information sets in periods, $k \geq h$.

D. Effects of New Data Releases

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

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

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

There are several facts worth noting. First, by expanding the sample, the weights applied to a particular observation are changed. The weights are not affected by the values of the observed data; the change is only due to finite-sample implementation of an infinite two-sided filter, implied by the auto-covariance generating function of the model. Second, the problem of analyzing news can be translated into the problem of data revision analysis by ‘aligning’ the increased sample with the old (shorter) sample padded with a model prediction for observables, $\mathbf{Y}_{T+1|T} = \mathbf{Z}\mathbf{T}\mathbf{X}_{T|T}$. It is easy to demonstrate that applying the Kalman smoother to such an augmented dataset yields identical results as smoothing the original data.³

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

³This is true when the new observations are not used to update initial conditions of the model. The proof follows simply from inspection of the Kalman smoother recursion, see [Durbin and Koopman \(2001\)](#), and the fact that the prediction error in this case is trivially zero.

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

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

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

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

III. MISSING OBSERVATIONS AND IMPOSING JUDGEMENT ON THE FILTER

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

This section discusses missing observations and simple techniques of how to impose judgment on the 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 suggested above and fit naturally within the framework.

Missing observations are not just a fact of life, but a useful tool for imposing judgment on the filter. The need for a transparent and coherent implementation of expert judgement follows from the need for expert judgment itself due to stylised nature of DSGE models. The literature focuses usually on expert judgement and conditioning within the forecast horizon, see [Benes, Binning, and Lees \(2008\)](#) for instance. Often, the analyst knows 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.

A. Missing observations and time-varying parameters

Handling missing observation is not trivial using classical Wiener-Kolmogorov methods, but the Kalman filter/smoothen 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 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 analog is obtained of the form

$$\mathbf{Y}_t^w = \mathbf{Z}_t^w \mathbf{X}_t + \mathbf{D}_t^w \varepsilon_t, \quad (17)$$

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 previous state.⁴

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 sets of weights with respect to \mathbf{Y}_t^w , namely

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

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 $\mathbf{W}_t^A = \mathbf{W}_t^B \forall t$.

⁴The 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 an approach is, however, prone to numerical instability.

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

$$\mathbf{Y}_t = \mathbf{Z}_t \mathbf{X}_t + \mathbf{D}_t \varepsilon_t \quad (19)$$

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

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

B. Imposing Judgement

By ‘judgement’ it is understood imposing additional information on the estimation of $\{\mathbf{X}_t\}_{t=1}$ beyond information contained in the model (1) and available sample data $\{\mathbf{Y}\}_{t=1}^T$.

The method of imposing judgment suggested below builds on the approach of [Doran \(1992\)](#) and on the insight of [Theil and Goldberger \(1961\)](#), resulting in augmenting the model (1) with a set of stochastic linear restrictions. It can be easily understood as dummy-observations priors for state variables. The method is very simple and amenable to observables decompositions.

Assume there is some extra information available, expressed as a set of n_k stochastic restrictions

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

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

$$\begin{bmatrix} \mathbf{Y}_t \\ \mathbf{k}_t \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_t \\ \mathbf{K}_t \end{bmatrix} \mathbf{X}_t + \begin{bmatrix} \mathbf{D}_t & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ \boldsymbol{\xi}_t \end{bmatrix} \quad (22)$$

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

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

One possible practical implementation simply defines new state variables satisfying the constraint (21), augments the state vector, and defines new measurement variables via an identity matrix. With no loss of generality having $\mathbf{K}_t = \mathbf{K}$ allows for the use of this approach of impos-

ing judgment with an implementation of the Kalman filter/smoothen 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. the output gap, cost-push shocks) with an arbitrary degree of precision, given by the 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 and others \(2009b\)](#) in the case of a structural, DSGE model, as well as in [Benes and others \(2010\)](#) in the case of a semi-structural output gap estimation. Making filter ‘tunes’ explicitly stochastic provides more leeway for imposing prior judgement.

(b) *Observables contribution analysis with restrictions* Observables decomposition of the filter is possible, 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 the 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 the law of iterated expectations, estimates using such an augmented database $\mathbf{Y}_t^{A,\star}$ are identical to those based on \mathbf{Y}_t^A .

(c) *Restricted forecasts* Obviously, the use of time-varying restrictions could also be used for producing a forecast up to $t+h$ periods with the initial condition $\mathbf{X}_{t|t} \sim N(\mathbf{X}_{t|t}, \mathbf{0})$ and running a Kalman smoother with restrictions imposed at the forecast horizon. Clearly, 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.

IV. EXAMPLES AND APPLICATIONS

Below, an example illustrating the filter analysis suggested in the text is provided. The example decomposes identified unobservable ‘flexible-price’ output gap in the DSGE model by [Smets and Wouters \(2007\)](#) into contributions of groups of observed variables.

A. 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 parameters estimated using the Bayesian-likelihood approach. 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 filter (7) with weights implied by the structural model.⁵ Having an estimate of the unobserved flex-price output gap, it is very interesting to find out exactly how the model combines sample data when forming the estimate.

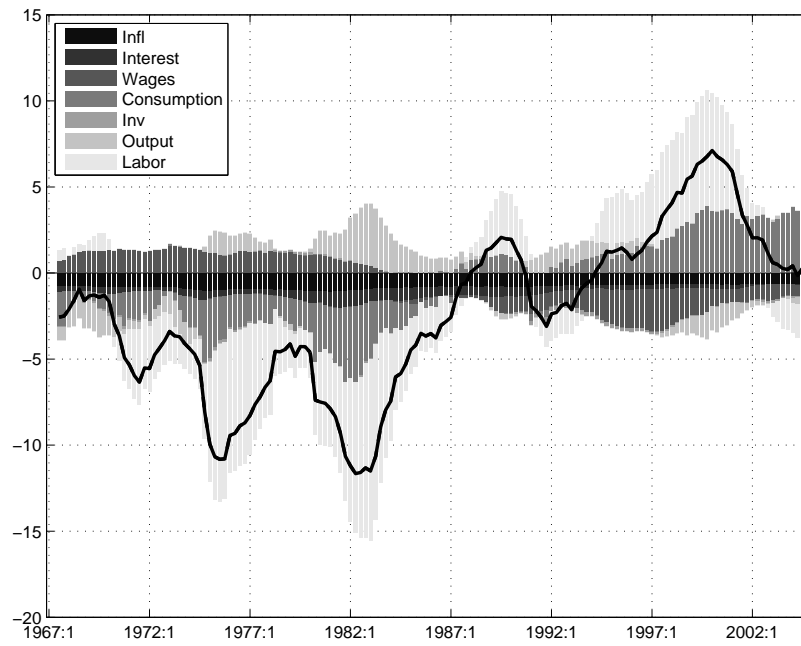
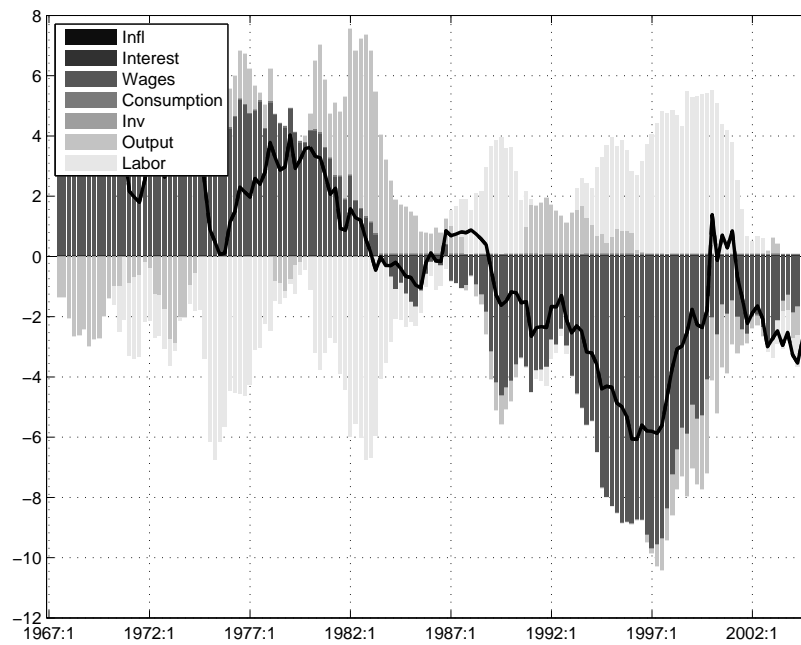
As can be seen from Fig. 1, the implicit output gap, also entering a policy rule, displays a larger degree of persistence and just a small tendency to mean reversion. The filter places optimal (in the mean-square sense) weights on available data that may corroborate similar pieces of information (e.g.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

A possible counterpart to inflation shock-decomposition in [Smets and Wouters \(2007\)](#) could be the 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 the long-run component of the model, i.e. single technology trend and, not accounting for an implicit ‘inflation target’ of the FED, or long-term inflation expectations.

Using the [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 the same number of observed

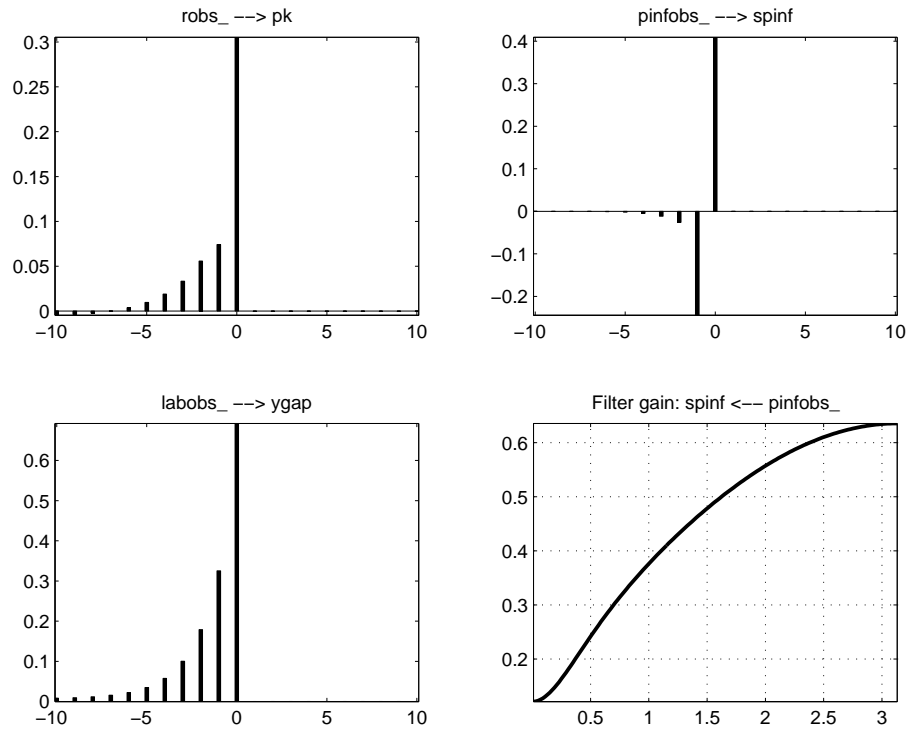
⁵Public provision of the codes, data and estimation results by F. Smets and R. Wouters is acknowledged. The calculations in the text do not rely on the Dynare version of these, but great care has been put into precise replication.

Figure 1. Flexible-price output gap in SW07**Figure 2. Real marginal costs in SW07**

variables as it has stochastic shocks. In a limiting case of known initial conditions, the shock estimation problem would cease to be a least-squares problem, but simply a linear system of equations. There are not enough degrees of freedom for efficiently imposing judgement, only small shifts of initial values.

One can also carry out a time domain and frequency domain analysis of $\mathbf{O}(z)$, as is usually the case for univariate filters. Fig. 3 depicts weights for the model and the 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 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 ($labobs$), that it is much harder to gain economic intuition than from direct filter contribution analysis exercises suggested in the text.

When the weights of the filter are inspected, 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 the 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.

Figure 3. Filter weights and gain SW07

V. CONCLUSION

This short paper puts forth methods and examples of filter analysis of structural DSGE models, allowing analysts to express estimates of unobservables in terms of observed sample data. The analysis proved useful for model development and design, along with forecasting and policy analysis. It is also easy to implement in practice.

Explicit understanding of the estimation of unobserved structural shocks in terms of linear filters is beneficial and can enhance insight into the economic structure of the model. The method allows researchers to understand how, for instance, the output gap, preference or technology shocks were identified from the data and what are the influential observables at what 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 of data arrive. Explicit decompositions are useful namely in the case of medium- to large-scale DSGE forecasting models. These decompositions are essential for economic story telling, alongside the decompositions of variables in terms of structural shocks.

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 surprise in the actual 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 [Andrle \(2012\)](#), where the revisions and misunderstanding of factors identifying the estimates can lead to serious policy errors.

REFERENCES

- Andrle, M., 2012, "What is in Your Output Gap," Techn. rep., IMF Working Paper, forthcoming.
- Andrle, M., Ch. Freedman, D. Laxton R. Garcia-Saltos, D. Hermawan, and H. Munandar, 2009a, "Adding Indonesia to the Global Projection Model," Techn. rep., IMF Working Paper 09/253.
- 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.
- Bell, W.R., 1984, "Signal extraction for nonstationary time series," *Annals of Statistics*, Vol. 12, pp. 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, "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.
- Burrige, P., and K.F. Wallis, 1988, "Prediction Theory for Autoregressive-Moving Average Processes," *Econometric Reviews*, Vol. 7, No. 1, pp. 65–95.
- Christiano, L.J., M. Eichenbaum, and Ch. Evans, 2005, "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, Vol. 113, No. 1, pp. 1–45.
- Christiano, L.J., and T.J. Fitzgerald, 2003, "The Bandpass Filter," *International Economic Review*, Vol. 44, No. 2, pp. 435–465.
- Doran, H.E., 1992, "Constraining Kalman Filter and Smoothing Estimates to Satisfy Time-Varying Restrictions," *Review of Economics and Statistics*, Vol. 74, pp. 568–572.
- Durbin, J., and S.J. 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*, Vol. 28, No. 3, pp. 361–385.
- Harvey, A.C., 1989, *Forecasting, Structural Time Series Models and the Kalman Filter* (Cambridge: Cambridge University Press).
- 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).
- Maravall, A., 1986, "Revisions in ARIMA Signal Extractions," *Journal of American Statistical Association*, Vol. 81, No. 395, pp. 736–740.
- Pierce, D.A., 1980, "Data revisions in moving average seasonal adjustment procedures," *Journal of Econometrics*, Vol. 14, No. 1, pp. 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*, Vol. 2, pp. 399–428.
- Schleicher, Ch., 2003, "Kolmogorov-Wiener Filters for Finite Time Series," Techn. rep., 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*, Vol. 97, No. 3, pp. 586–606.
- Theil, H., and A.S. Goldberger, 1961, "On Pure and Mixed Statistical Estimation in Economics," *International Economic Review*, Vol. 2, No. 1, pp. 65–78.
- Whittle, P., 1983, *Prediction and Regulation by Linear Least Squares Methods* (Minneapolis: University of Minnesota Press).

VI. APPENDIX

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

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

$$\mathbf{X}_{t|\infty} = \mathbf{O}(L)\mathbf{Y}_t = \mathbf{G}_{\mathbf{X}\mathbf{Y}}(L)\mathbf{G}_{\mathbf{Y}}^{-1}(L)\mathbf{Y}_t = \sum_{j=-\infty}^{\infty} \mathbf{O}_j \mathbf{Y}_{t+j}, \quad (24)$$

where $\mathbf{G}_{\mathbf{Y}}(z)$, $\mathbf{G}_{\mathbf{X}\mathbf{Y}}(z)$ stands for the multivariate auto-covariance generating function (ACGF) of the state-space model, e.g.

$$\mathbf{G}_{\mathbf{X}\mathbf{Y}}(z) = (\mathbf{I} - \mathbf{T}z)^{-1} \mathbf{G} \mathbf{\Sigma}_\varepsilon \mathbf{G}' (\mathbf{I} - \mathbf{T}'z^{-1})^{-1} \mathbf{Z}' \quad (25)$$

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

Note the usefulness of the (24), which clearly defines the filter and its transfer function. Analysts can study links from the data to unobservables across frequencies.

Regarding a time domain expression for the time invariant weights, Gomez (2006) shows that for (1) the weights \mathbf{O}_j follow as (adjusted to our model specification)

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

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

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

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 Riccati equation (DARE) associated with the 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 also easily obtained after carrying out the Kalman filter iterations (which are data independent) until convergence.