Parameter and State Estimation with Singular DSGE Models*

Michal Andrle
International Monetary Fund

Preliminary and incomplete

June 20, 2016

*First version: January 24, 2014 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. I would like to thank Jan Bruha, James Costain, Martin Elisson, Omar Rachedi, Antti Ripatti, participants at the Bank of Spain research seminar in June 2016, and participants at the Computing in Economics and Finance, Oslo June 2014, for valuable comments.

Contents Page

Tables

Figures

I. Introduction

Economic models can be stochastically singular when the number of structural economic shocks is smaller than the number of the variables the model uses for identification of these shocks.

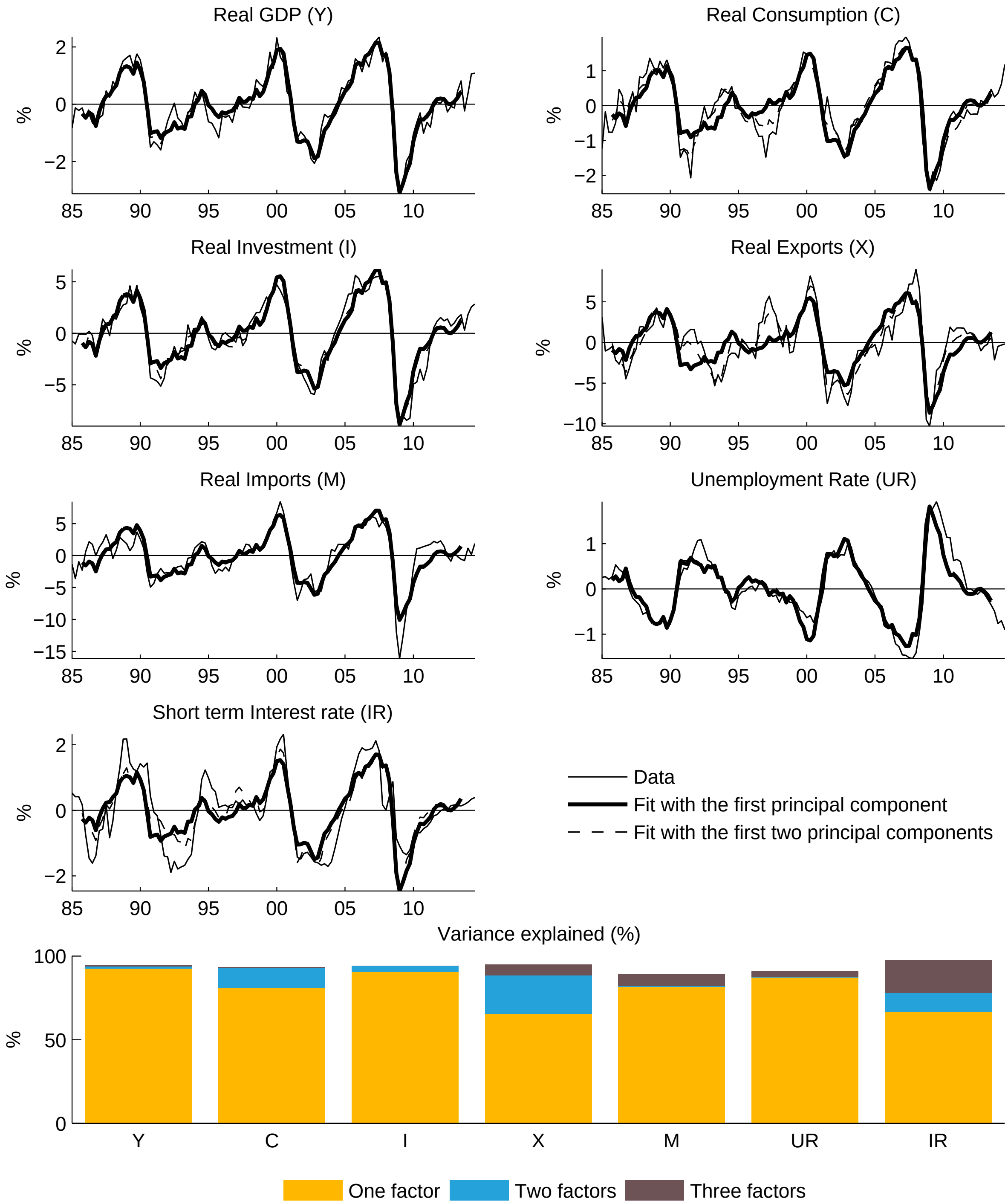
It may be realistic to assume that a small number of 'structural' shocks is causing most variation in economic data, while a myriad of other shocks contribute to relatively small proportion of data variation in the course of the business cycle. Andrle and Brůha (2013), for instance, point out striking commonality among real and nominal macroeconomic variables in OECD countries. At business-cycle frequencies, one dynamic factor explains up to 80% of economic fluctuations in consumption, investment, interest rates, inflation, and other macroeconomic variables.

Real-world data, however, are never stochastically singular, i.e. they do not have singular spectral density. Far from singularity, real-world data may be contaminated by measurement errors and omissions. If the model is stochastically singular, while the data at hand are not, inevitably the model cannot explain all variation in the data, there will be a residual, an unexplained part. Still, it is a perfectly valid question to ask how 'close' the model gets to the actual data with the given structure and set of shocks.

In this note, 'close' is being defined in a mean-square sense and the path of shocks is estimated that minimizes discrepancy of the model with respect to data, while having minimal energy (variance). For simplicity, the estimation is translated into an explicit under-determined least-squares problem, easily solved using the singular value decomposition (SVD). The non-recursive algorithm is trivial, but not suitable for larger models.

A related approach to estimating singular DSGE models and shock identification is presented in Andrle (2012), where the model is rotated into a (dynamic) principal component space of the data. Effectively, the model's measurements are as many principal components as there are stochastic shocks in the model. This paper's approach can be interpreted as mapping both the model and data observables into a dynamic principal component subspace of the model itself. An alternative approach to handle singular DSGE models proposed by Canova, Ferroni, and Mathes (2014), is to select a subset of variables that gives the strongest degree of parameter identification.

Figure 1. Business Cycle Comovement, USA (1985-2015)



Source: Andrle, Bruha, and Solmaz (2016)

II. State Estimation with Singular Models

A. Model

It is assumed that the model can be expressed as a linear state-space model and thus the whole discussion continues only in terms of state-space model. Economics comes in later. The model can be written as:

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

$$\mathbf{Y}_t = \mathbf{Z}\mathbf{X}_t + \mathbf{H}\mathbf{e}_t \qquad \text{with} \qquad \mathbf{e}_t \sim \mathbf{N}(\mathbf{0}, \boldsymbol{\Sigma}_\mathbf{e}),\ \boldsymbol{\Sigma}_\mathbf{e} = \mathbf{I} \tag{2}$$

The model can be stochastically singular if the number of shocks, $n_e = \dim(\mathbf{e}_t)$ is smaller than the number of observed variables, $n_y = \dim(\mathbf{Y}_t)$. The assumption $\boldsymbol{\Sigma}_\mathbf{e} = \mathbf{I}$ simplifies the exposition of the modeling approach with no loss of generality.

The estimation of the state variables and shocks would normally proceed with the Kalman smoother. In the case when $n_e < n_y$, this standard solution is not feasible, unless the observed data have exactly the dynamic rank equal to n_e. Such situation is rare and we will assume that the process $\{\mathbf{Y}_t^\mathbf{O}\}$ has a stochastic rank equal to n_y.

The stochastic process $\mathbf{Y}_t$ is in general an infinite-order moving-average representation

$$\mathbf{Y}_t = \left[\mathbf{Z}(\mathbf{I} - \mathbf{T}L)^{-1}\mathbf{R} + \mathbf{H}\right]\mathbf{e}_t = \mathbf{D}(L)\mathbf{e}_t, \tag{3}$$

with spectral density given by the expression $\mathbf{S}_\mathbf{Y}(\omega) = \frac{1}{2\pi}\mathbf{D}(e^{-i\omega})\boldsymbol{\Sigma}_\mathbf{e}\mathbf{D}(e^{-i\omega})^\dagger$. In the case when $n_y > n_e$, the spectrum of the model is singular. For simplicity, we will assume that unless $n_y > n_e$, there are no pathologies in $\mathbf{D}(L)$ that would result in stochastic singularity of the model.

Our analysis is carried out in two related spaces—in time domain and in frequency domain. There are three equivalent representations of the model: (i) recursive representation in time domain, (ii) stacked representation in time domain, and (iii) frequency-domain representation. Each representation may be more feasible for some analysis than the others but they are equivalent. It is common that most contributions to the literature use the first, time-domain recursive representation to estimate unobserved structural shocks and compute the likelihood function using the Kalman filter.

We will focus on the time-domain stacked representation of the model for state estimation and structural parameter estimation and link it to the frequency-domain representation later.

It is useful and instructive to rewrite the model into a stacked form, as a function of the initial state $\mathbf{X}_0$ and stochastic shocks, $\mathbf{e}_t$, only. Denoting $\mathbf{Y} = [\mathbf{Y}'_1\ \mathbf{Y}'_2\ \ldots\ \mathbf{Y}'_N]'$ and $\mathbf{E} = [\mathbf{X}_0\ \mathbf{e}'_1\ \ldots\ \mathbf{e}'_N]'$, we c is stated as follows:

$$\mathbf{Y} = \mathbf{A} \times \mathbf{E}, \tag{4}$$

where the 'multiplier' matrix $\mathbf{A}$ is clearly given by the structure of the model and the values of T, Z, R and H. The dimension of $\mathbf{A}$ is $(n_y N) \times (n_e N + n_x)$. Further, it is trivial to see that the structure of $\mathbf{A}$ is

$$\mathbf{A} = \begin{bmatrix} \mathbf{ZT} & | & \mathbf{ZR+H} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{ZT}^2 & | & \mathbf{ZTR} & \mathbf{ZR+H} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{ZT}^3 & | & \mathbf{ZT}^2\mathbf{R} & \mathbf{ZTR} & \mathbf{ZR+H} & \mathbf{0} & \ldots & \mathbf{0} \\ \ldots & | & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \mathbf{ZT}^N & | & \mathbf{ZT}^{N-1}\mathbf{R} & \ldots & \ldots & \ldots & \ldots & \mathbf{ZR+H} \end{bmatrix} = [\mathcal{O}\ \mathcal{H}] \tag{5}$$

For cases where the model is stochastically singular, some additional care is needed when handling the estimate of the initial state. First, it is important to re-scale the estimated initial state, so the estimated vector's covariance matrix is a unitary matrix, as is the case with the shocks. Second, the unconditional variance of the state vector, $\mathbf{X}_0$, need not to be full rank and thus only the identified component of the state vector will be estimated.

More specifically, let $\mathbf{\Sigma_X}$ be an $n_x \times n_x$ unconditional covariance matrix of the stationary process $\mathbf{X}$. Let the rank of the matrix be denoted r_x and note that this static rank of the model may differ from the dynamic rank of the model, $r_e = n_e$. Using a singular-value decomposition (SVD) of the covariance matrix, $\mathbf{\Sigma}_X = \mathbf{U}_{\mathbf{X},r}\mathbf{S}_{\mathbf{X},r}\mathbf{V}'_{\mathbf{X},r}$, we can define a new, lower-rank and rescaled initial state $(r_x \times 1)$ vector $\mathbf{W}_0$ with a unitary covariance matrix and which satisfies the mapping

$$\mathbf{X}_0 = \mathbf{U}_{\mathbf{X},r}\mathbf{S}_{\mathbf{X},\mathbf{r}}^{-1/2} = \mathbf{M}\mathbf{W}_0. \tag{6}$$

The $(n_x \times r_x)$ mapping $\mathbf{M}$ transforms the estimated low-rank initial state into the original model coordinates.

B. State Estimation

State estimation makes use of equivalence between least-squares, Wiener-Kolmogorov, and Kalman filtering. After all, the very nature of the Kalman smoother is that it is a recursive and efficient version of the least squares.

To avoid explicit frequency-domain calculations, the solution stacks time into a static least-squares problem, which is solved by singular value decomposition. However, frequency-domain issues can still be explored.

The least-squares problem is formulated using the stacked form of the model (4) and employing the transformation (6):

$$\min_{\mathbf{X}_0,\{\boldsymbol{\varepsilon}\}} \quad \Lambda = \mathbf{W}_0\mathbf{W}_0' \quad + \quad \sum_{t=1}^{N} [\mathbf{Y}_t - \mathbf{Z}\mathbf{X}_t]\,(\mathbf{H}\mathbf{H}')^{-1}\,[\mathbf{Y}_t - \mathbf{Z}\mathbf{X}_t]' \tag{7}$$

$$+ \quad \sum_{t=1}^{N} [\mathbf{X}_t - \mathbf{T}\mathbf{X}_{t-1}]\,(\mathbf{R}\mathbf{R}')^{-1}\,[\mathbf{X}_t - \mathbf{T}\mathbf{X}_{t-1}]'. \tag{8}$$

It is useful to rewrite the least-squares problem in a stacked form and as a function of the initial state $\mathbf{X}_0$ and stochastic shocks, $\mathbf{e}_t$, only. Denoting $\mathbf{Y} = [\mathbf{Y}_1' \; \mathbf{Y}_2' \; \ldots \; \mathbf{Y}_N']'$ and $\boldsymbol{\mathcal{E}} = [\mathbf{W}_0 \; \mathbf{e}_1' \; \ldots \; \mathbf{e}_N']'$, the least-squares problem is stated as follows:

$$\boldsymbol{\mathcal{E}} = \text{argmin}\; ||\mathbf{Y} - \boldsymbol{\mathcal{A}} \times \boldsymbol{\mathcal{E}}||, \tag{9}$$

where the ‘multiplier’ matrix **A** is clearly given by the structure of the model and the values of $\mathbf{T}, \mathbf{Z}, \mathbf{R}$ and $\mathbf{H}$. The dimension of $\boldsymbol{\mathcal{A}}$ is $(n_y N) \times (n_e N + r_x)$. It should not be surprising to realize that the that the structure of $\boldsymbol{\mathcal{A}}$ is

$$\boldsymbol{\mathcal{A}} = \begin{bmatrix} \mathbf{ZTM} & | & \mathbf{ZR}+\mathbf{H} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{ZT}^2\mathbf{M} & | & \mathbf{ZTR} & \mathbf{ZR}+\mathbf{H} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{ZT}^3\mathbf{M} & | & \mathbf{ZT}^2\mathbf{R} & \mathbf{ZTR} & \mathbf{ZR}+\mathbf{H} & \mathbf{0} & \ldots & \mathbf{0} \\ \ldots & | & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \mathbf{ZT}^N\mathbf{M} & | & \mathbf{ZT}^{N-1}\mathbf{R} & \ldots & \ldots & \ldots & \ldots & \mathbf{ZR}+\mathbf{H} \end{bmatrix}. \tag{10}$$

The importance of rewriting the model into a simple, albeit possibly large, least-squares problem is that all the modern methods for solving large-scale least-square problems are available, including various forms of *regularization* and *penalized estimation*, or natures of the

$||.||$ norm. Most notably, an extension of the form

$$\boldsymbol{\mathcal{E}} = \text{argmin}\ ||\mathbf{Y} - \boldsymbol{\mathcal{A}} \times \boldsymbol{\mathcal{E}}|| + \lambda \mathcal{P}(\mathcal{E}), \tag{11}$$

where $\mathcal{P}(\mathcal{E})$ is a penalty term with a flexible functional form.

1. Under-determined System

In the case when $n_y > n_e$, the least-square problem is under-determined and standard solutions do not apply. Fortunately, solution of such a system is a standard and well-understood problem in linear algebra. A common way of solving the problem is an application of the 'queen' of matrix transformations, the singular value decomposition (SVD).

The solution to the least-squares problem is not unique. It is chosen is such that the energy of the shocks is the smallest among those solution that provide identical mean-square error to the system. Since the problem is under-determined, the 'fitted' values of observables do not equal to observed data, unless in the very rare case of stochastically singular input data. Hence, there will be 'errors', or residuals.

The solution can be written in terms of the SVD transformation of the multiplier matrix $\boldsymbol{\mathcal{A}}$ as follows, following standard results in linear algebra.[1].

$$\boldsymbol{\mathcal{A}} = \mathbf{U}\,\mathbf{S}\,\mathbf{V}' = \begin{bmatrix}\mathbf{U}_r\ \mathbf{U}_0\end{bmatrix}\begin{bmatrix}\mathbf{S}_r & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{V}_r'\\ \mathbf{V}_{0'}\end{bmatrix} = \mathbf{U}_r\ \mathbf{S}_r\ \mathbf{V}_r' = \sum_{i=1}^{r}\mathbf{u}_i\sigma_i\mathbf{v}_i' \tag{12}$$

$$\hat{\boldsymbol{\mathcal{E}}} = \mathbf{V}_r\mathbf{S}_r^{-1}\mathbf{U}'_r\mathbf{Y} = \Psi\mathbf{Y}, \tag{13}$$

where r is the 'effective rank' of the multiplier matrix used for the estimation of shocks. In principle, the effective rank is chosen exactly to match the estimated rank of the multiplier matrix. This way, all the equations of the model are respected in their original form.

After the estimation of $\hat{\boldsymbol{\mathcal{E}}}$, it is needed to unstack the vector and feed the shocks through the original model to obtain the estimates of unobserved states $\{\mathbf{X}\}$ and the fit of the observed data. For a stochastically singular model, the fit will not be perfect, of course.

(a) Relationship to Other Methods The solution method uses efficiently the model structure and the rank of the system to compute a minimum norm solution but can be related to

[1] See e.g. Strang (200x) for introduction and Golub and van Loan (1997) for a more advanced treatment

other methods. Andrle (2012) suggest to use principal components (PCA) or dynamic principal component (DPCA) analysis to estimate shocks and parameters of stochastically singular models. In its simplest form, let's assume that the observed data covariance matrix is an $(n_y \times n_y)$ matrix $\mathbf{S_Y}$, while the model-implied covariance matrix of the data is $\mathbf{\Sigma_Y}$, consistent with (1)–(2). Given the measurement equation $\mathbf{Y}_t = \mathbf{Z}\mathbf{X}_t + \mathbf{H}\mathbf{e}_t$, one can devise a new, lower-dimensional vector of variables $\mathbf{F}_t$, which consists of 'factors' created from the observed data as $\mathbf{F}_t = \mathbf{P} \times \mathbf{Y}_t$. The whole model is then mapped into a new space of coordinates as $\mathbf{F}_t = \mathbf{PZX}_t + \mathbf{PHe}_t = \tilde{\mathbf{Z}}\mathbf{X}_t + \tilde{\mathbf{H}}\mathbf{e}_t$.[2] Andrle (2012) originally proposed to create the projection matrix $\mathbf{P}$ using the principal component analysis of the *empirical covariance* matrix of the data, $\mathbf{S_Y}$, in order to subject the model to empirical stylized facts and to make it perform in the data space. However, it is possible to form the projection matrix also using the model-implied covariance matrix, $\mathbf{\Sigma_Y}$, for given set of structural parameters.

The 'SVD filter' solution proposed in this text is related to the principal component method hinted above and can easily be used jointly with it to pre-whiten the data, focusing on the explanation of only key variance in the data. The SVD filter is, in principle, another instance of principal components itself. The vector of stacked observed data $\mathbf{Y}$ is mapped into a lower-dimensional space using the projection matrix associated with principal components, $\tilde{\mathbf{Y}} = \mathbf{U}_r'\mathbf{Y}$ and a new model is formulated for the solution, $\tilde{\mathbf{Y}} = \mathbf{U}_r'\mathcal{A}\boldsymbol{\mathcal{E}} = \mathbf{S}_r\mathbf{V}_r'\boldsymbol{\mathcal{E}}$. The intuition behind the transformation of the observed data into a new coordinate space will prove useful later when the likelihood function of the stochastically singular model is introduced.

2. Conditioning of the State Estimation

Conditioning of the least squares problem is tighly related to sensitivity of model to rounding errors, errors in the measured inputs and to weakly identified subspaces of the model.

Assume, for instance, that the measured data $\mathbf{Y}$ is now defined as true data and some 'noise', $\mathbf{Y} = \mathbf{Y}_{true} + \varepsilon\boldsymbol{\mathcal{W}}$. Using the singular value decomposition, we can rewrite the solution to the least squares problem as

$$\hat{\boldsymbol{\mathcal{E}}} = \sum_{i=1}^{r} \frac{\mathbf{u}_i'\mathbf{Y}_{true}}{\sigma_i}\mathbf{v}_i + \varepsilon \sum_{i=1}^{r} \frac{\mathbf{u}_i'\boldsymbol{\mathcal{W}}}{\sigma_i}\mathbf{v}_i, \tag{14}$$

where the solution can be seen as a function of the 'true', correctly measured data and measurement errors, or noise. Now, if the vector $\boldsymbol{\mathcal{W}}$ corresponds to (roughly) uncorrelated white

[2] In case of the dynamic principal components, the mapping is not static but a two-sided polynomial and the approach is executed in the frequency domain but the intuition is similar.

noise, the the parts of the vector in the direction of the left-singular vectors of $\mathcal{A}$ will stay roughly constant and thus the terms $\mathbf{u}_i'\mathcal{W}$ will not vary too much with i. As such, the second term can blow up with increasing i as the singular values get sufficiently small.

In order to lessen the effect of the potential noisy data, one can use a *truncated problem* and use smaller number of singular values than is the (approximate) rank of the problem, i.e. $k < r$. The truncated solution is a specific form of regularization approach out of many that are known for the least-squares problems, together with Tikhonov regularization, for instance, see **?**) and the Appendix of this paper. While not following the problem in detail and focusing on the intuition, it is interesting to note that one way of deciding on the size of problem is to inspect so called discrete Piccard condition, which requires the coefficients $|\mathbf{u}_i'\mathbf{Y}|$ to decay faster on average than the corresponding singular values.

Truncated solution is also one of the form of the *spectral filtering* and removing components of the solution associated with very low singular values corresponds to making the solution less sensitive to high-frequency variations in the data.

C. Likelihood Function and Parameter Estimation

Up to this point it has been assumed that the coefficients of the underlying model are given. However, it is rather straightforward to develop a likelihood of the state-space form of the model even with the stochastic singularity.

Let's assume the model is already cast into a stacked linear form as follows:

$$\mathbf{Y} = \mathbf{A}\,\mathbf{E} = \mathbf{A}_\theta\,\mathbf{E}, \tag{15}$$

where the multiplier matrix $\mathbf{A}_\theta$ is now associated with the subscript θ to indicate the dependence of the reduced-form coefficients on the structural coefficients. The goal now is to carry out statistical inference about the structural parameters θ. Given the Gaussian structure of components in $\mathbf{E}$, $\mathbf{Y} \sim \mathbf{N}(\mu, \mathbf{S_Y})$, where $\mathbf{S_Y} = \mathbf{A}\,\mathbf{S_E}\,\mathbf{A}'$.[3] In the case where $\text{rank}(\mathbf{A}) = \max\{r, p\}$ and thus full-rank covariance matrix $\mathbf{S_Y}$ the likelihood function is

$$L = -N \times T/2 + \log 2\pi + \log|\mathbf{S_Y}^{-1}| - \frac{1}{2}\mathbf{Y}'\mathbf{S_Y}^{-1}\mathbf{Y}, \tag{16}$$

which is easy to evaluate for small and medium-size models.

[3] This point has been noticed also by Schmitt-Grohe and Uribe (2013)

In the case when the dynamic model, and thus the coveriance matrix $\mathbf{S_Y}$ is singular, the inverse does not exists and the likelihood function is ill-defined. However, a solution with a clear and economically meaningful intuition is available. A solution that at the same time leads to a developtment of the singular multivariate Normal (SMN) distribution.

The intuition behind the simple algebra is similar to Andrle (2012), where the observations are mapped into a new sub-space, defined by the principal componets of the data. In this case, both the model and the data are mapped into a subspace defined by the regular portion of the model, as embodied in the column-rank of $\mathbf{A}$. Using the singular value decomposition $\mathbf{U\,S\,V'} = \mathbf{A}$, we can define a set of new, transformed observables $\tilde{\mathbf{Y}} = \mathbf{U_r'Y}$. Essentially, the model has been mapped into a new space of the first r principal components of the model.

The new likelihood is thus proportional to

$$L_U \propto \log|\tilde{\mathbf{S}}_{\mathbf{Y}}^{-1}| - \frac{1}{2}\tilde{\mathbf{Y}}'\tilde{\mathbf{S}}_{\mathbf{Y}}^{-1}\tilde{\mathbf{Y}}, \tag{17}$$

$$\propto \log|\mathbf{\Lambda}_r^{-1}| - \frac{1}{2}\mathrm{trace}\{\mathbf{YY'S_Y^{+}}\}. \tag{18}$$

The three equivalent forms of the likelihood function exist due to several important relationships between the SMN, SVD, and thecoordinate transformation. The last expression is the singular multivariate Normal distribution, see Rao (1973), where $\mathbf{\Lambda}$ are the ordered, non-zero eigenvalues of $\mathbf{S_Y}$ and $\mathbf{S_Y^{+}}$ is the Moore-Penrose inverse, or pseudo inverse, of the singular covariance matrix.

III. Example

A. Simple Dynamic Semi-Structural Model

The model follows a typical New-Keynesian closed economy model with price rigidities. Inflation, π_t, is driven by output in excess of its trend or equilibrium value, using a forward-looking Phillips curve. The output cycle, $\hat{y}$, is determined by an output equation derived from consumption smoothing and is interest sensitive. The monetary policy authority sets the short-term nominal interest rate, i_t, via an inflation-forecast based rule, weighting the expected deviation of year-on-year inflation from its target and the output gap.

$$\hat{y}_t = \alpha_1 y_{t+1|t} + \alpha_2 y_{t-1} + \alpha_3 (rr_t - \overline{rr}_t) + \varepsilon^y_t \tag{19}$$

$$\pi^c_t = \lambda_1 \pi^c_{t+1|t} + (1-\lambda_1)\pi^c_{t-1} + \lambda_2 \hat{y}_t + \varepsilon^\pi_t \tag{20}$$

$$i_t = \gamma_1 i_{t-1} + (1-\gamma_1) \times \left[(\overline{rr}_t + \overline{\pi}_t) + \gamma_2(\pi^{y/y}_{t+3|t} - \overline{\pi}_{t+3}) + \gamma_3 y_t\right] + \varepsilon^i_t \tag{21}$$

$$\pi_t = \pi^c_t + \varepsilon^{\pi,sr}_t \tag{22}$$

$$rr_t = i_t - \pi_{t+1|t} \tag{23}$$

$$\tag{24}$$

Despite its small size and simplicity, the model can display nontrivial dynamics in response to structural shocks. It is driven by eight parameters $\theta = \{\alpha_{1,2,3}, \lambda_{1,2}, \rho_i, \gamma_{1,2,3}\}$ and four standard deviations for structural shocks. Inflation target is assumed to be exogenous and for the purposes of the paper is kept constant. The results for the model will be presented in deviation from the steady-state values.

B. State Estimation with the Model

Given the model specification, see the appendix, the goal is to use three observed variables, the interest rate, i_t, headline inflation, π_t, and the output (gap), $\hat{y}_t$, to estimate four structural shocks, $\{\varepsilon^y_t, \varepsilon^\pi_t, \varepsilon^\pi_t \varepsilon^{\pi,sr}_t\}$. The true model is not stochastically singular, on the contrary, the number of than is the number of observables.

State estimation with the model can proceed using the Kalman smoother, as it is standard. All results will be compared with the Kalman smoother as a natural benchmark. The SVD filter for non-singular model with out use of any regularization replicates the results of the Kalman smoother exactly in the test case; an indespensible test.

To illustrate the potential use of the SVD filter, shock estimation is carried out with the stochastically singular model, misspecified model, and with observed data contaminated with a measurement noise.

(a) Estimating with the demand shocks only Let's assume that the resercher uses a theory that does not take a stand on detailed specification of more but one structural shock. In our model it is the demand shock, for proponents of the real business cycle theory it could be a productivity shock, for instance. The SVD filter, without any further setup or estimation of measurement errors will estimate the shock and leave *residuals* for series it cannot match.

(b) Estimating with the cost-push shock only This case is, surely, is a purposefuly artificial exercise just to illustrate the properties of the SVD filtering. The model is clearly misspecified as the cost-push shock contributes only a small portion of the variance in the observed data and is at odds with the strong and positive unconditional correlation between output and inflation in the model. The results demonstrate that the filter can match decently two out of three observed variables: inflation and to smaller extent the interest rate. But it does–as it should–do a rather poor job in explaing the output.

(c) Minor filter misspecification The following exercise focuses on a rather minor form of misspecification, which is also relevant in the presence of measurement noise. The baseline data-generating process is modified such that the demand shock and the short-term cost push shock affecting headline inflation are used to generate the data, with the policy shock playing negligible role just to make sure the Kalman smoother operates well. The analyst will, however, specify his model as a function of the policy shock, demand shock, and the persistent cost push shock.

(d) Noise Contamination

1. Varying the rank of the SVD-filter

The rank of the filter with one shock and the initial conditions, can be further lowered, to explore the frequency-domain implications of the SVD filter.

Figure 2. Noisy data, structural shocks, and estimates

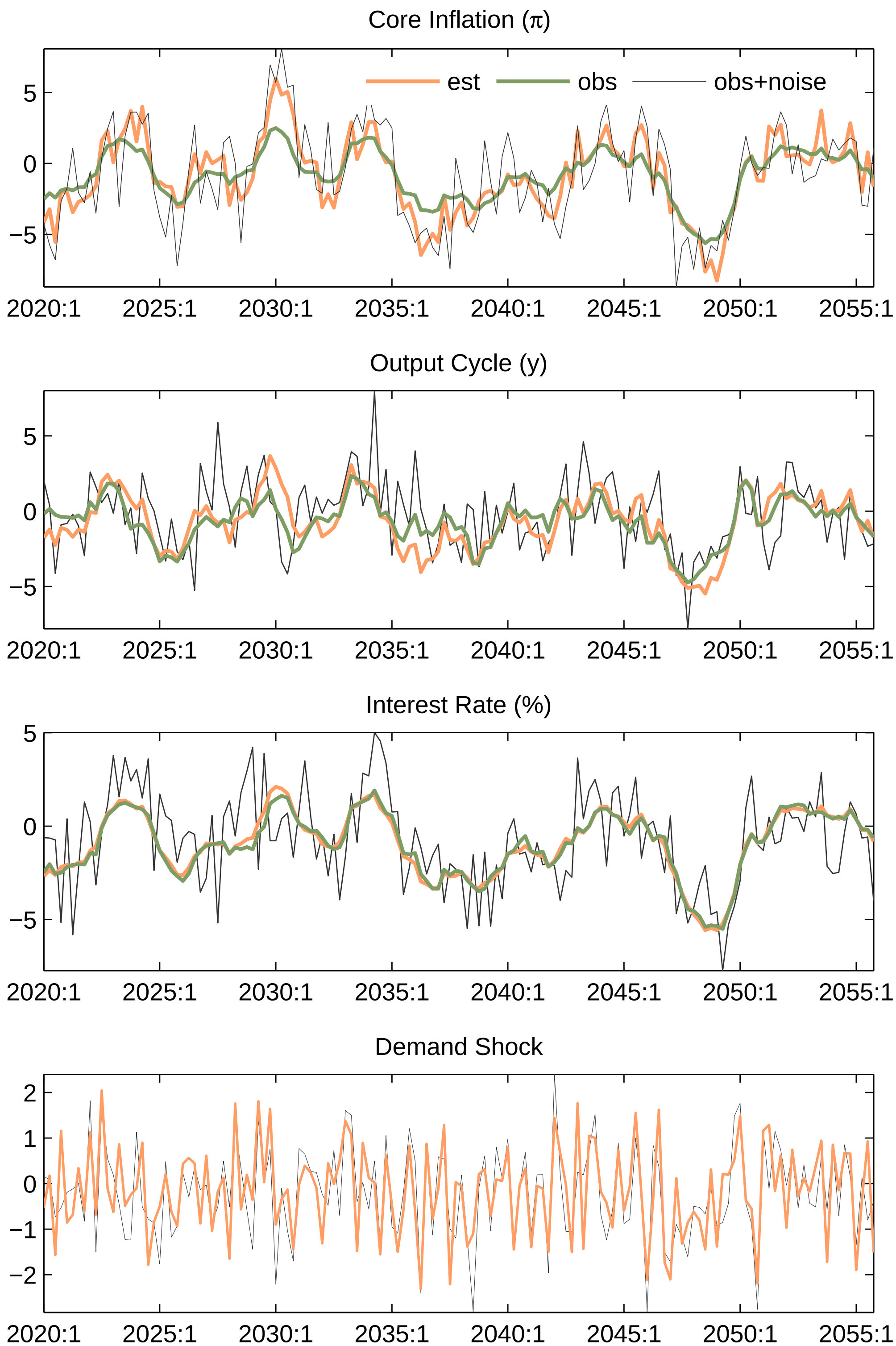
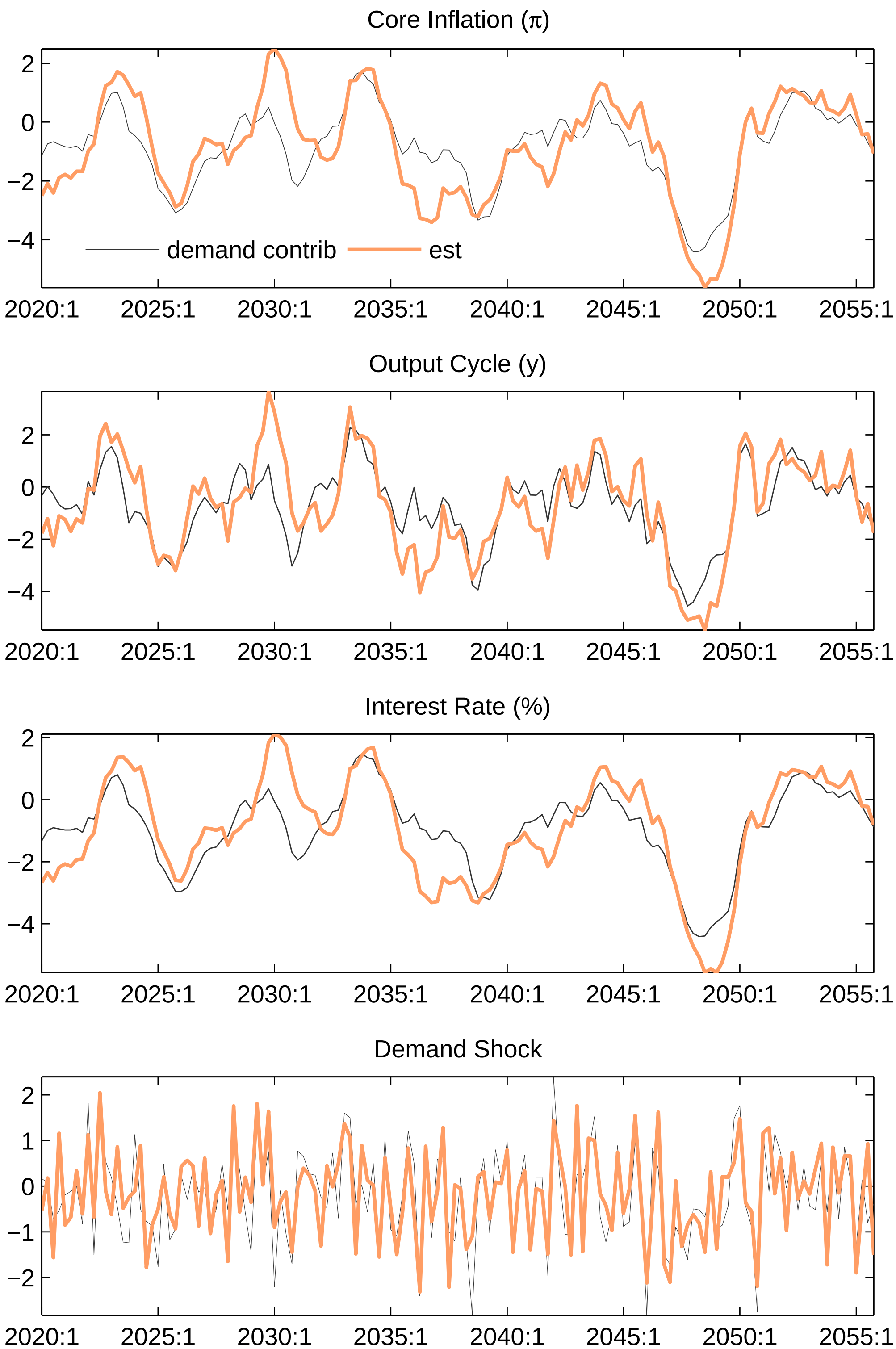


Figure 3. Effects of true vs. estimated demand shock



IV. CONCLUSION

When there is a desire to use a model with small number of shocks and test it against larger number of observables, it is feasible to estimate the shocks by least squares. This short note introduced a simple filter that can easily handle stochastically singular models, as well as regular ones.

The benefit of using the 'SVD filter' is that it can be used to enhance the robustness to measurement errors and allows to test the hypothesis of only few relevant shocks. On the other hand, in case the chosen shock is severely misspecified –or altogether missing in the data– the application of the reduced-rank filter, as well as any other filter is dangerous and misleading.