EUROPEAN CENTRAL BANK
EUROSYSTEM

Working Paper Series

Rodolfo Dinis Rigato

A least-squares filter for sequence-space models

No 3191

Abstract

Sequence-space models are becoming increasingly popular in macroeconomics, especially in the heterogeneous-agent literature. However, the econometric toolkit for users of these models remains less developed than that available for traditional state-space methods. This note introduces an algorithm for efficiently filtering unobserved shocks in linear sequence-space models. The proposed filter solves a least-squares optimization problem in closed form and returns the expectation of unobserved shocks conditional on observed data. It handles heteroskedasticity, missing observations, measurement error, and non-Gaussian shock distributions. To illustrate its properties, I apply it to data simulated from a medium-scale heterogeneous-agent New Keynesian model and show that it accurately recovers the underlying structural shocks.

JEL classification codes: C32, E27, E32, E37

Keywords: Sequence space, Filtering, Least squares

Non-technical summary

Many economic models rely on variables that cannot be directly observed. A prominent example is the output gap, typically defined as the difference between actual output (or real GDP) and potential output. While real GDP can be measured, potential output is a theoretical concept: it is the level of economic activity consistent with stable inflation, given the economy's productive capacity, such as its labour force, capital stock, and technology. Because potential output is not observable, economists estimate it—along with many other hidden variables—using economic models and data, a process known as filtering.

This note proposes a new filtering method designed for a growing class of macroeconomic models known as sequence-space models. These models are increasingly used to study economic phenomena involving heterogeneity, such as how households with different levels of wealth respond differently to monetary policy. The proposed approach makes it possible to recover unobservable economic variables in these settings in a systematic and tractable way.

1 Introduction

Sequence-space methods are becoming increasingly popular in macroeconomics, in particular for solving heterogeneous-agent (HA) models.[1] In HA models, the entire distribution of agents, which is typically infinite-dimensional, is a state variable. Therefore, working with these models in a state-space representation can be challenging.[2] The sequence-space approach circumvents this issue by computing impulse response functions directly, relying on the equivalence between first-order perfect-foresight shocks and IRFs of linearized models. However, while there is a rich set of econometric tools available for practitioners of the state-space approach,[3] they are not immediately applicable in the sequence space.

This note proposes a simple algorithm for filtering unobserved shocks in the sequence space, accompanied by a set of public available codes.[4] The method finds the likelihood-maximizing history of shocks ε_t^j, for $j \in \{1, \ldots, J\}$, consistent with a set of $I \leq J$ observable variables y_t^i, for $i \in \{1, \ldots, I\}$. In a linearized sequence-space model, observables can be expressed as a set of MA(∞) processes of the form

$$y_t^i = \sum_{j=1}^{J} \sum_{k=0}^{\infty} \frac{\partial y_t^i}{\partial \varepsilon_{t-k}^j} \varepsilon_{t-k}^j. \tag{1}$$

When shocks are normally distributed, likelihood maximization becomes a least-squares optimization problem subject to the linear constant (1). This problem can solved in closed form, delivering filtered shocks as a linear transformation of observables. Moreover, its output corresponds to the expectation of the shocks conditional on the data.

This method has a number of desirable properties. First, it is efficient: The most time-consuming step of the algorithm consists of solving a positive definite linear system of size $(I \cdot T) \times (I \cdot T)$, where T is the sample size. For a medium-scale model with the typical sample size in macroeconomic time series, as in the application described below, the running time can be measured in milliseconds. Second, it is flexible: The method can be easily adapted to handle missing observations, heteroskedasticity, measurement error, and non-Gaussian shocks. Under non-Gaussian shocks, however, filtered shocks are no longer the expectation of shocks

[1] See Boppart et al. (2018) and Auclert et al. (2021).

[2] See Ahn et al. (2018), Winberry (2018), and Bayer et al. (2024).

[3] Fernández-Villaverde et al. (2016)

[4] See https://github.com/rdrigato/least_squares_filter/ for a Python implementation of the algorithm described here.

conditional on the data, but rather their linear projection (as long as second moments are well-defined). Last, as opposed to the Kalman filter for state-space models, it requires no distributional assumptions on initial conditions (see Fernández-Villaverde et al. 2016). One desirable property it shares with the Kalman filter, however, is that once shocks are filtered (or alternatively, the unobserved vector of states in the state space), forecasts can be immediately constructed using equation (1), as well as filtered values of endogenous unobservable variables such as the output gap. To illustrate the method, I apply it to data simulated from the medium-scale HANK model in Kase and Rigato (2025).

2 Filtering Shocks in the Sequence Space

Time is indexed by $t \in \{0, \dots, T-1\}$, where T is the sample size. Observables y_t^i are without loss of generality assumed to have zero mean. For now, assume there are no missing observations and that shocks are i.i.d. and normally distributed: $\varepsilon_t^j \sim N(0, (\sigma^j)^2)$. The only model objects necessary for the algorithm are impulse response functions, denoted by

$$x_k^{ij} = \frac{\partial y_{t+k}^i}{\partial \varepsilon_t^j},$$

one for each shock-observable pair. Above, $k \in \{0, \dots, T^{IRF}-1\}$, where T^{IRF} is the truncation horizon of the IRFs.

The idea is to recover the most likely history of shocks ε_t^j, going all the way from $t = -(T^{IRF}-1)$ until the end of the sample $t = T-1$. Given the normality assumption, the log likelihood of shock ε_t^j is:

$$\ell_t^j = -\frac{1}{2}\left(\frac{\varepsilon_t^j}{\sigma^j}\right)^2 - \frac{1}{2}\log\left(2\pi(\sigma^j)^2\right).$$

This gives rise to the following least-squares optimization problem:

$$\min_{\{\varepsilon_t^j\}_{t,j}} \frac{1}{2}\sum_{j=1}^{J}\sum_{t=-(T^{IRF}-1)}^{T-1}\left(\frac{\varepsilon_t^j}{\sigma^j}\right)^2 \tag{2}$$

$$\text{subject to } y_t^i = \sum_{j=1}^{J}\sum_{k=0}^{T^{IRF}-1} x_k^{ij}\varepsilon_{t-k}^j. \tag{3}$$

The objective function (2) is the log likelihood of the entire history of shocks, while the constraint (3) is the truncated form on (1).

To solve the problem, it is convenient to first write it in matrix form. Let $\mathbf{e}^j$ be the history of shock j, i.e.,

$$\mathbf{e}^j = (\varepsilon^j_{-(T^{IRF}-1)}, \ldots, \varepsilon^j_{T-1})',$$

and $\mathbf{e}$ be the stacked history of all shocks:

$$\mathbf{e} = \left(\left(\mathbf{e}^1\right)', \ldots, \left(\mathbf{e}^J\right)'\right)'.$$

Define $\mathbf{y}^i$ and $\mathbf{y}$ in a similar way:

$$\mathbf{y}^j = (y^i_0, \ldots, y^i_{T-1})'$$

$$\mathbf{y} = \left(\left(\mathbf{y}^1\right)', \ldots, \left(\mathbf{y}^I\right)'\right)'.$$

Note that shocks start from $t = -(T^{IRF} - 1)$, while observables start from $t = 0$. Now define the diagonal matrix $\mathbf{\Sigma}$ as:

$$\mathbf{\Sigma} = \text{diag}(\underbrace{(\sigma^1)^2, \ldots, (\sigma^1)^2}_{(T+T^{IRF}-1)\text{ times}}, \ldots, (\sigma^J)^2, \ldots, (\sigma^J)^2)$$

The objective function (2) can be written as $\frac{1}{2}\mathbf{e}'\mathbf{\Sigma}^{-1}\mathbf{e}$.

To write the constraint (3) in matrix form, note that

$$\mathbf{y}^i = \sum_j \underbrace{\begin{bmatrix} x^{ij}_{T-1} & \cdots & x^{ij}_0 & 0 & \cdots & 0 \\ 0 & x^{ij}_{T-1} & \cdots & x^{ij}_0 & 0 & \cdots \\ 0 & 0 & x^{ij}_{T-1} & \cdots & x^{ij}_0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & x^{ij}_{T-1} & \cdots & x^{ij}_0 \end{bmatrix}}_{\boldsymbol{X}^{ij}} \mathbf{e}^j,$$

where each matrix $\boldsymbol{X}^{ij}$ has size $T \times (T + T^{IRF} - 1)$ and contains IRFs repeated along the rows, starting from the main diagonal. The constraint (3) can be written as

$$\mathbf{y} = \boldsymbol{X}\mathbf{e},$$

where $\boldsymbol{X} = \left[\boldsymbol{X}^{ij}\right]_{ij}$ is a block matrix consisting of the $\boldsymbol{X}^{ij}$. The problem can be solved in closed form, returning filtered shocks $\hat{\mathbf{e}}$ as a linear transformation of the data:

$$\hat{\mathbf{e}} = \boldsymbol{\Sigma X}' \left(\boldsymbol{X \Sigma X}'\right)^{-1} \mathbf{y}. \tag{4}$$

Importantly, equation (4) has a simple interpretation in terms of conditional expectations. Using the expressions for covariances of variables in sequence-space representation from Auclert et al. (2021), it follows that $\boldsymbol{X\Sigma X}' = \mathbb{E}\mathbf{yy}'$, i.e., the variance-covariance matrix of the vector $\mathbf{y}$. Similarly, we have $\boldsymbol{\Sigma X}' = \mathbb{E}\mathbf{ey}'$. Therefore, equation (4) becomes the familiar conditional expectation formula for jointly normally distributed vectors:

$$\begin{aligned} \hat{\mathbf{e}} &= \left(\mathbb{E}\mathbf{ey}'\right)\left(\mathbb{E}\mathbf{yy}'\right)^{-1}\mathbf{y} \\ &= \mathbb{E}(\mathbf{e}|\mathbf{y}) \end{aligned} \tag{5}$$

Moreover, it is possible to express analytically the variance of filtered shocks as well, given by:

$$\begin{aligned} Var(\mathbf{e}|\mathbf{y}) &= \mathbb{E}\mathbf{ee}' - \left(\mathbb{E}\mathbf{ey}'\right)\left(\mathbb{E}\mathbf{yy}'\right)^{-1}\left(\mathbb{E}\mathbf{ye}'\right) \\ &= \boldsymbol{\Sigma} - \boldsymbol{\Sigma X}'\left(\boldsymbol{X\Sigma X}'\right)^{-1}\boldsymbol{X\Sigma}. \end{aligned}$$

Auclert et al. (2021) solve a similar problem when computing the log likelihood of sequence space models and provide numerical methods for efficiently computing the matrices $\boldsymbol{X\Sigma X}'$ and $\boldsymbol{\Sigma X}'$ using discrete Fourier transforms. They also discuss numerical methods for inverting matrix $\boldsymbol{X\Sigma X}'$, which is the most time-consuming step in the application below.[5]

Given this setup, it is straightforward to introduce measurement error, heteroskedasticity, missing observations, and non-Gaussian shocks. Measurement error of the form $\mathbf{y} = \boldsymbol{X}\mathbf{e} + \mathbf{u}$, with $\mathbf{u} \sim \mathbf{N}(\mathbf{0}, \boldsymbol{\Omega})$ orthogonal to $\mathbf{e}$, can be introduced by noting that it only affects the variance-covariance matrix of the observables $\mathbf{y}$, but not their correlations with the shocks $\mathbf{e}$. Therefore, we have

$$\hat{\mathbf{e}} = \boldsymbol{\Sigma X}'\left(\boldsymbol{X\Sigma X}' + \boldsymbol{\Omega}\right)^{-1}\mathbf{y}.$$

[5] In the applications below, the system displays condition numbers of the order of 10^4, and therefore is not ill-conditioned. Results are also robust to changes in the truncation horizon T^{IRF}, as long as it is long enough for the effects to have converged to zero by the end of the horizon. Auclert et al. (2021) provide a more detailed discussion of the numerical issues in this problem.

Alternatively, one can solve

$$\min_{\mathbf{e},\mathbf{u}} \frac{1}{2}\mathbf{e}'\boldsymbol{\Sigma}^{-1}\mathbf{e} + \frac{1}{2}\mathbf{u}\boldsymbol{\Omega}^{-1}\mathbf{u} \quad \text{subject to} \quad \mathbf{y} = \boldsymbol{X}\mathbf{e} + \mathbf{u}.$$

Heteroskedasticity in the form of time-varying shock variances $(\sigma_t^j)^2$ can be introduced by re-defining the matrix $\boldsymbol{\Sigma}$ as:

$$\boldsymbol{\Sigma} = \text{diag}((\sigma^1_{-(T^{IRF}-1)})^2, \ldots, (\sigma^1_{T-1})^2, \ldots, (\sigma^J_{-(T^{IRF}-1)})^2, \ldots, (\sigma^J_{T-1})^2).$$

Next, missing observations can be handled by manipulating the $\boldsymbol{X}$ matrix. Notice that $\boldsymbol{X}$ has $I \cdot T$ rows, each corresponding to an observation y_t^i. If any y_t^i is missing, simply delete the corresponding row of $\boldsymbol{X}$, and equation (4) will deliver the expectation of $\mathbf{e}$ conditional on non-missing observations. Finally, if shocks are non-Gaussian, the first line in expression (5) does not necessarily correspond to their conditional expectation anymore, but is still the best linear unbiased predictor (as long as second moments are well-defined).[6]

3 Application: A Medium-Scale HANK Model

I illustrate the method using the estimated medium-scale heterogeneous-agent new Keynesian (HANK) model from Kase and Rigato (2025). The model features heterogeneous households that face idiosyncratic income risk and decide how much to consume and to save both in liquid and illiquid assets, as well as the standard ingredients of medium-scale DSGE models (Smets and Wouters, 2007), such as price and wage stickiness, investment adjustment costs, a Taylor (1993) rule, and a fiscal authority that collects taxes to finance government consumption.

The model features 7 shocks, affecting: the consumers' discount factor, firms' investment technology, price and wage markups, total factor productivity, government spending, and monetary policy. I simulate $T = 100$ periods of data from the model using IRFs of size $T^{IRF} = 300$. To do so, I simulate 400 periods and discard the first 300, which, given the truncation horizon of the IRFs, ensures that variables are sampled from their joint ergodic distribution. I use 7 observables to filter the structural shocks: the growth rates of aggregate consump-

[6] Let $\boldsymbol{B} = \boldsymbol{\Sigma}\boldsymbol{X}'\left(\boldsymbol{X}\boldsymbol{\Sigma}\boldsymbol{X}'\right)^{-1}$ and $\tilde{\boldsymbol{B}}$ be any other matrix of the same size. If shocks are non-Gaussian but still have finite second moments, one can show that $\mathbb{E}\left\|\boldsymbol{e} - \tilde{\boldsymbol{B}}\boldsymbol{x}\right\|^2 = \mathbb{E}\left\|\boldsymbol{e} - \boldsymbol{B}\boldsymbol{x}\right\|^2 + \mathbb{E}\left\|(\boldsymbol{B} - \tilde{\boldsymbol{B}})\boldsymbol{x}\right\|^2$, from which the claim follows.

tion, investment, output, hours worked, and compensation per employee, as well as inflation and the short-term nominal interest rate. Parameter values correspond to the posterior mode in Kase and Rigato (2025)

Figure 1 shows simulated and filtered shocks[7] in a baseline case in which shocks are i.i.d. and normally distributed. Figure 2 introduces measurement error by assuming that each observation is subject to an i.i.d error $u_{it} \sim (0, 0.1^2)$. Figure 3 shows results under heteroskedasticity, assuming that investment shocks are 5 times more volatile starting from period $t = 50$. As can be seen, the filter performs well in these cases, recovering shocks close to the simulated ones.

With missing data, as expected, the filter's performance is hampered. Figure 4 shows results when investment data is missing in the second half of the sample. Interestingly, missing investment data reduces more severely the accuracy of the filtered government spending shocks, suggesting that investment shocks are more easily identifiable due to larger effects on other observable variables. All other shocks are still accurately filtered. Last, figure 5 shows results when shocks are simulated from a Student's t-distribution with 4 degrees of freedom (normalized to have unit variance). This distribution is severely fat-tailed; in fact, it features infinite kurtosis. Nevertheless, the filter still works well. Finally, table 1 provides accuracy metrics. It shows correlations between filtered and simulated shocks and root-mean-square errors (RMSE) under the specifications above, computed as averages of 500 Monte Carlo simulations.

4 Conclusion

This note introduces a least-squares algorithm for filtering unobserved shocks in sequence-space models. The problem can be solved in closed form and delivers the expectation of unobserved shocks conditional on the data, and can handle missing observations, heteroskedasticity, measurement error, and non-Gaussian shocks. An application to data simulated from a medium-scale HANK model shows that it works well in practice.

[7] To keep results comparable to state-space methods, I display filtered shocks only for $t \geq 0$.

References

Ahn, SeHyoun, Greg Kaplan, Benjamin Moll, Thomas Winberry, and Christian Wolf, "When inequality matters for macro and macro matters for inequality," *NBER macroeconomics annual*, 2018, *32* (1), 1–75.

Auclert, Adrien, Bence Bardóczy, Matthew Rognlie, and Ludwig Straub, "Using the sequence-space Jacobian to solve and estimate heterogeneous-agent models," *Econometrica*, 2021, *89* (5), 2375–2408.

Bayer, Christian, Benjamin Born, and Ralph Luetticke, "Shocks, Frictions, and Inequality in US Business Cycles," *American Economic Review*, 2024, *114* (5), 1211–1247.

Boppart, Timo, Per Krusell, and Kurt Mitman, "Exploiting MIT shocks in heterogeneous-agent economies: the impulse response as a numerical derivative," *Journal of Economic Dynamics and Control*, 2018, *89*, 68–92.

Fernández-Villaverde, Jesús, Juan Francisco Rubio-Ramírez, and Frank Schorfheide, "Solution and estimation methods for DSGE models," in "Handbook of macroeconomics," Vol. 2, Elsevier, 2016, pp. 527–724.

Kase, Hanno and Rodolfo Rigato, "Beyond averages: heterogeneous effects of monetary policy in a HANK model for the euro area," 2025.

Smets, Frank and Rafael Wouters, "Shocks and frictions in US business cycles: A Bayesian DSGE approach," *American economic review*, 2007, *97* (3), 586–606.

Taylor, John B, "Discretion versus policy rules in practice," in "Carnegie-Rochester conference series on public policy," Vol. 39 Elsevier 1993, pp. 195–214.

Winberry, Thomas, "A method for solving and estimating heterogeneous agent macro models," *Quantitative Economics*, 2018, *9* (3), 1123–1151.

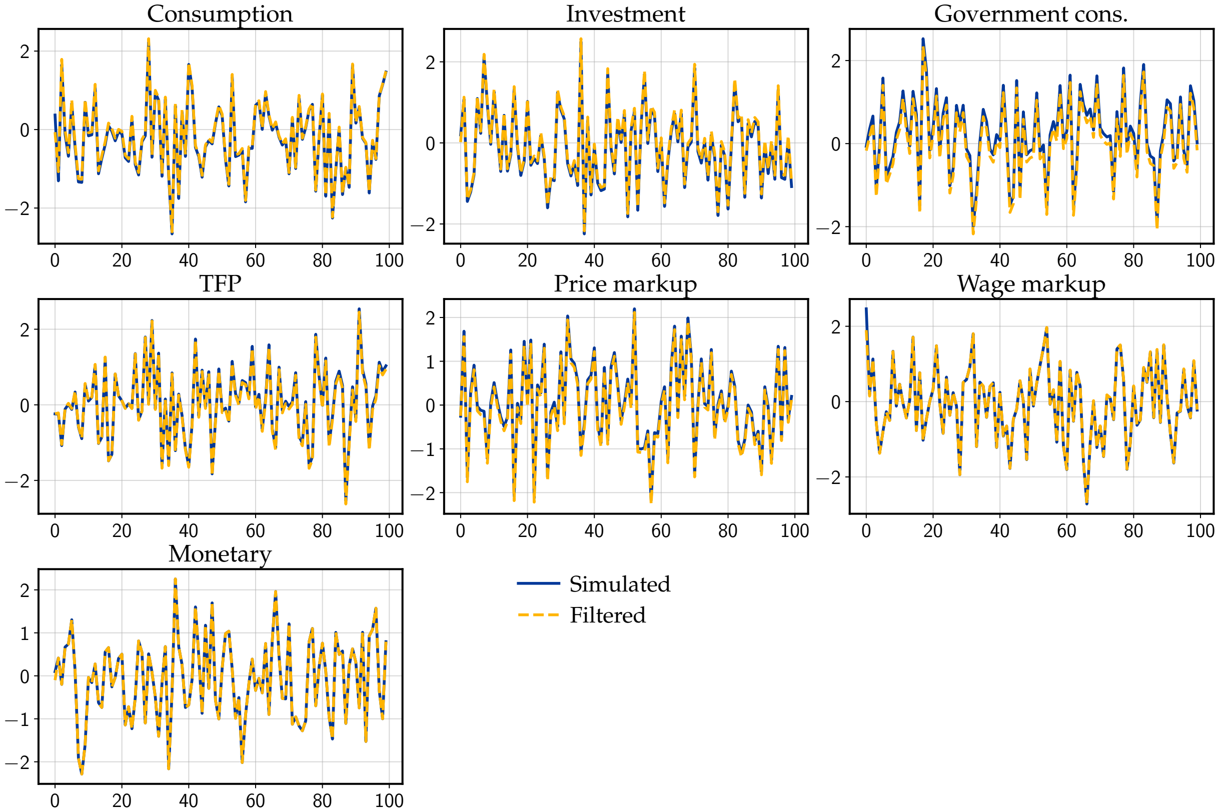

Figure 1: Simulated and filtered shocks

	Baseline		Meas. error		Missing obs.		Heteroskedastic		Fat-tailed	
	Corr.	RMSE	Corr.	RMSE	Corr.	RMSE	Corr.	RMSE	Corr.	RMSE
Consumption	1.0	0.06	0.97	0.23	1.0	0.07	1.0	0.06	1.0	0.06
Investment	1.0	0.05	1.0	0.09	0.97	0.22	1.0	0.05	1.0	0.05
Government cons.	1.0	0.01	0.94	0.34	0.74	0.67	1.0	0.01	1.0	0.01
TFP	1.0	0.05	0.99	0.11	1.0	0.06	1.0	0.05	1.0	0.05
Price markup	1.0	0.04	1.0	0.08	1.0	0.04	1.0	0.04	1.0	0.04
Wage markup	1.0	0.05	0.99	0.11	1.0	0.05	1.0	0.05	1.0	0.05
Monetary	1.0	0.06	0.95	0.31	1.0	0.06	1.0	0.06	1.0	0.06

Table 1: Correlations and RMSE between filtered and simulated shocks

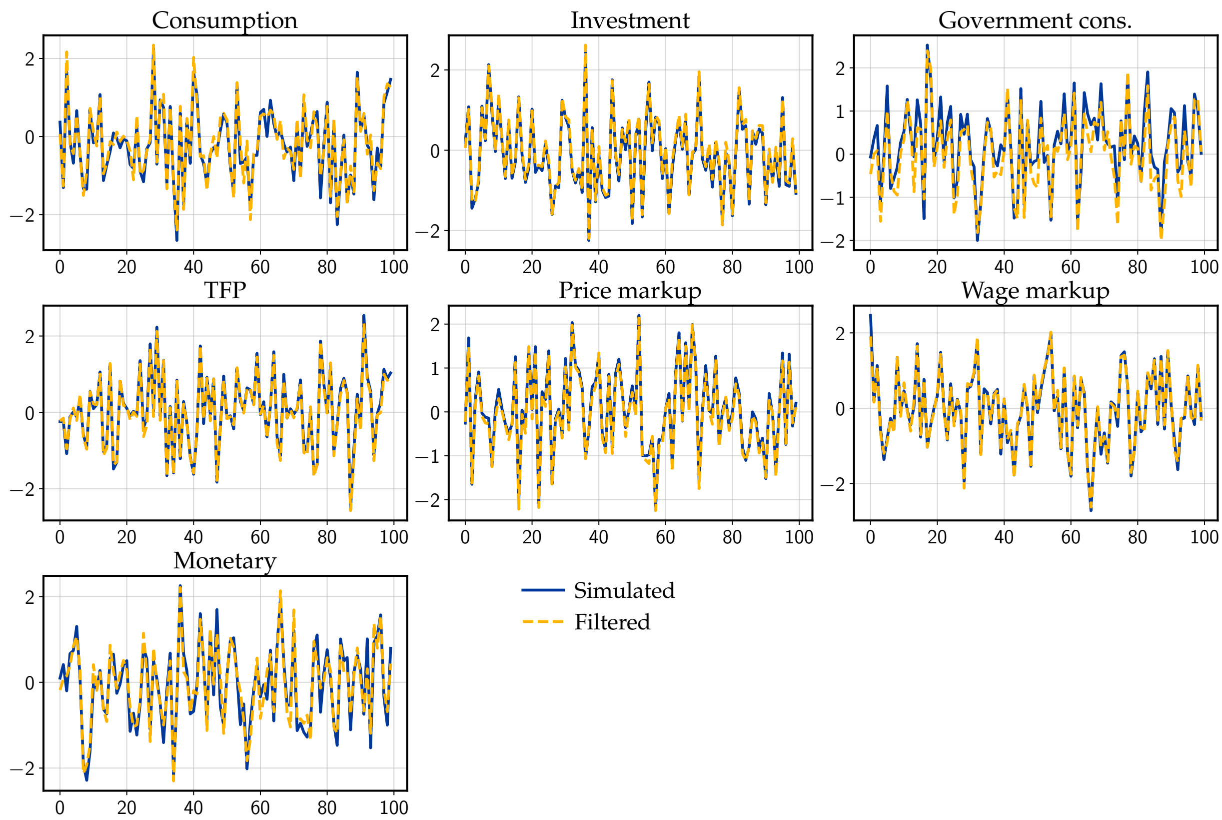

Figure 2: Simulated and filtered shocks with measurement error

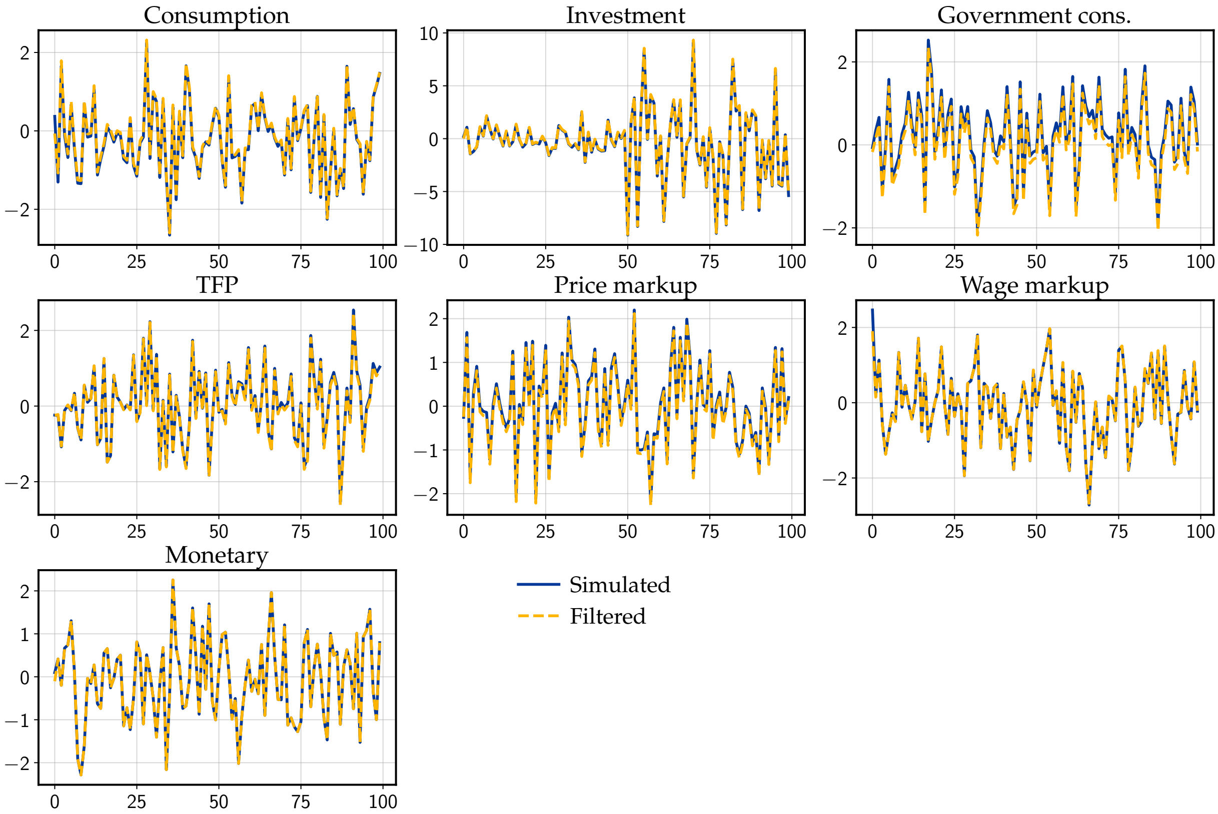

Figure 3: Simulated and filtered shocks with heteroskedastic investment shocks

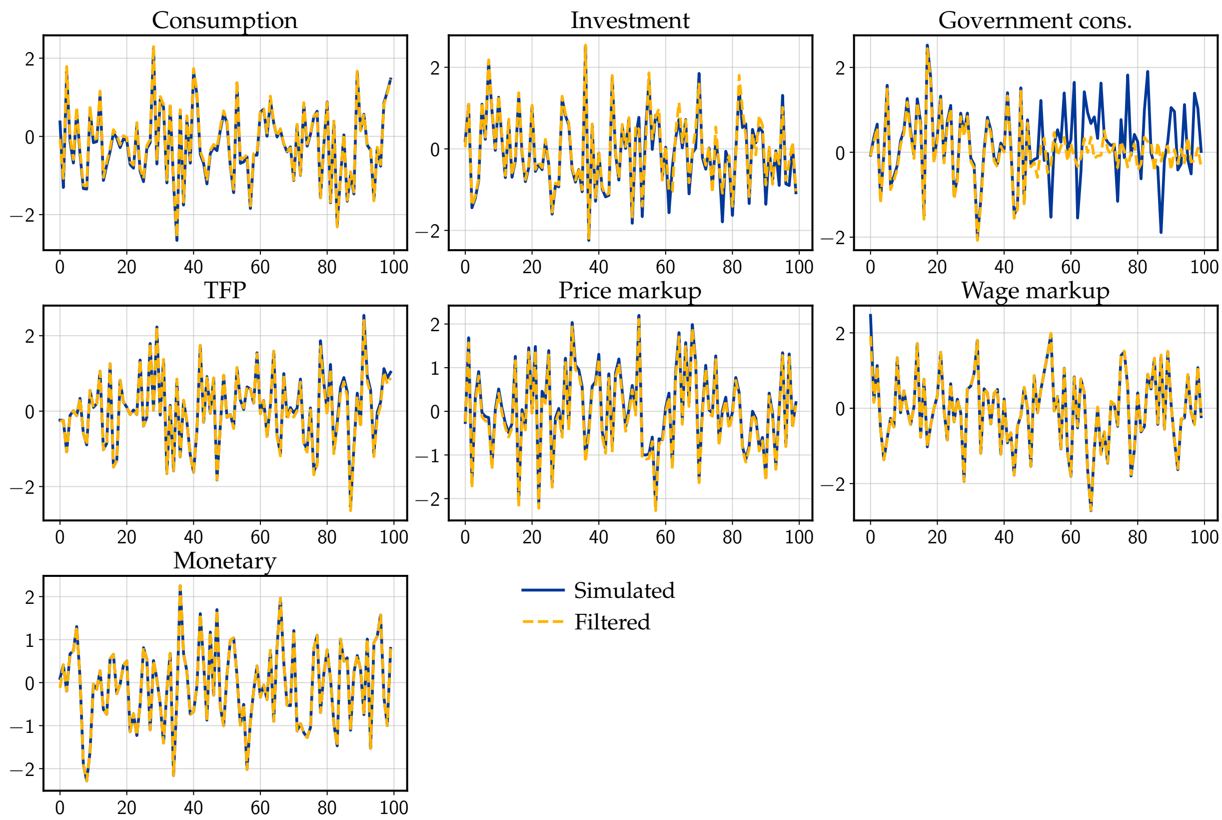

Figure 4: Simulated and filtered shocks with missing investment data

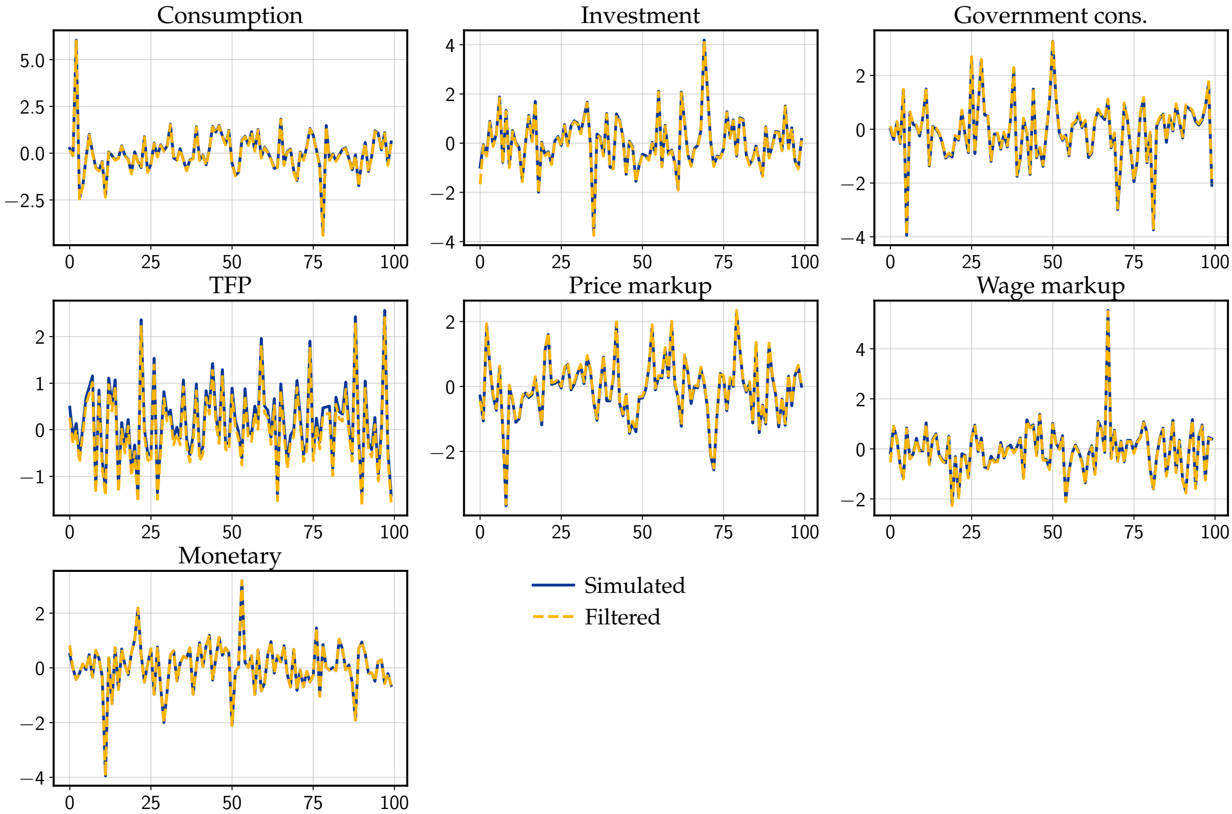

Figure 5: Simulated and filtered shocks with a fat-tailed distribution

Acknowledgements

I thank Marta Banbura, Hanno Kase, Michele Lenza, and an anonymous referee for helpful comments.
The views expressed here are solely those of the author and do not necessarily reflect the views of the European Central Bank.

Rodolfo Dinis Rigato
European Central Bank, Frankfurt am Main, Germany; email: rodolfo.dinis_rigato@ecb.europa.eu

Postal address 60640 Frankfurt am Main, Germany
Telephone +49 69 1344 0
Website www.ecb.europa.eu

This paper can be downloaded without charge from www.ecb.europa.eu, from the Social Science Research Network electronic library or from RePEc: Research Papers in Economics. Information on all of the papers published in the ECB Working Paper Series can be found on the ECB's website.

PDF ISBN 978-92-899-7714-2 ISSN 1725-2806 doi:10.2866/0708911 QB-01-26-074-EN-N