DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

CREATES
Center for Research in Econometric Analysis of Time Series

Estimation of continuous-time linear DSGE models from discrete-time measurements

Bent Jesper Christensen, Luca Neri and
Juan Carlos Parra-Alvarez

CREATES Research Paper 2022-12

Department of Economics and Business Economics
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

Email: oekonomi@au.dk
Tel: +45 9352 1936

Estimation of continuous-time linear DSGE models from discrete-time measurements*

Bent Jesper Christensen† Luca Neri‡ Juan Carlos Parra-Alvarez§

December 16, 2022

Abstract

We provide a general state space framework for estimation of the parameters of continuous-time linear DSGE models from data that are only available at discrete points in time. Our approach relies on the exact discrete-time representation of the equilibrium dynamics, which allows avoiding discretization errors. Using the Kalman filter, we construct the exact likelihood for data sampled either as stocks or flows, and estimate frequency-invariant parameters by maximum likelihood. We address the aliasing problem arising in multivariate settings and provide conditions for precluding it,which is required for local identification of the parameters in the continuous-time economic model. We recover the unobserved structural shocks at measurement times from the reduced-form residuals in the state space representation by exploiting the underlying causal links imposed by the economic theory and the information content of the discrete-time observations. We illustrate our approach using an off-the-shelf real business cycle model. We conduct extensive Monte Carlo experiments to study the finite sample properties of the estimator based on the exact discrete-time representation, and show they are superior to those based on a naive Euler-Maruyama discretization of the economic model. Finally, we estimate the model using postwar U.S. macroeconomic data, and offer examples of applications of our approach, including historical shock decomposition at different frequencies, and estimation based on mixed-frequency data.

Keywords: DSGE models, continuous time, exact discrete-time representation, stock and flow variables, Kalman filter, maximum likelihood, aliasing, structural shocks.
JEL classification: C13, C32, C68, E13, E32, J22

*We are grateful to Martin Møller Andreasen, Tim Bollerslev, Federico Carlini, Leopoldo Catania, Manfred Deistler, Francis X. Diebold, Dennis Kristensen, Alessia Paccagnini, Giovanni Pellegrino, Olaf Posch, Enrique Sentana, and participants at the 2022 NBER-NSF Times Series Conference (Boston, USA), the 11th Nordic Econometric Meeting (Sandbjerg Manor, Denmark), the XIIt Workshop in Time Series Econometrics (Zaragoza, Spain), the Arne Ryde Workshop: Heterogeneous Agent Models in Macroeconomics - Advances in Continuous Time (Lund University), the Nordic Junior Macro Seminar, and the Quantitative Economics seminar (Hamburg University) for valuable comments and discussions, and to the Danish Social Science Research Council (grant number 2033-00137B) for research support. Juan Carlos Parra-Alvarez acknowledges financial support from the Otto Mønsted Foundation.

†*Corresponding author*. Department of Economics and Business Economics, Aarhus University; CREATES; Dale T. Mortensen Center; Danish Finance Institute; Fuglesangs Allé 4, 8210 Aarhus V, Denmark; Email address: bjchristensen@econ.au.dk

‡Department of Economic Sciences, University of Bologna; CREATES; Dale T. Mortensen Center; Ca' Foscari University of Venice; P.zza Scaravilli 2, Italy; Email address: luca.neri19@unibo.it

§Department of Economics and Business Economics, Aarhus University; CREATES; Danish Finance Institute; Fuglesangs Allé 4, 8210 Aarhus V, Denmark; Email address: jparra@econ.au.dk

1. Introduction

Dynamic stochastic general equilibrium (DSGE) models have become a fundamental tool for macroeconomic analysis. Furthermore, the use of continuous-time methods has gained renewed interest, with a large number of models being developed and used to study the transmission mechanisms of monetary and fiscal policies, and to provide a better understanding of the interactions between the real and financial sides of the economy (see, e.g., Posch, 2011, Brunnermeier and Sannikov, 2014, Kaplan et al., 2018, Itskhoki and Moll, 2019, Posch, 2020, Fernández-Villaverde et al., 2020, Liemen and Posch, 2022). This increase in popularity has led to the development and improvement of numerical methods for approximation of the solution of both representative and heterogeneous agent continuous-time models (see, e.g., Posch, 2009, Posch and Trimborn, 2013, Parra-Alvarez, 2018, Ahn et al., 2018, Parra-Alvarez et al., 2021, Achdou et al., 2022). However, considerably less work has been devoted to study how to take these models to the data. This step requires addressing the issue that the economic theory is derived under the assumption that time evolves continuously, whereas the data used to test its validity is measured at discrete points in time. This is a particularly pressing problem in macroeconomics, where data usually are available only at low frequencies, e.g., monthly, quarterly, or annually.

In this paper, we propose and analyze a method for estimation and inference from multivariate linear continuous-time structural economic models based on discrete-time observations, with the aim to make comprehensive and systematic use of the macroeconomic data available. The approach accounts for the discrete nature of the sampling scheme, while keeping the underlying probabilistic structure unaltered. The discrete-time data generating process results from a combination of the discrete sampling scheme and the underlying continuous-time model. As the decision intervals of economic agents are not tied to the observation intervals of sampled data, the frequency with which data are measured is not relevant for the parameters of interest, namely, those of the continuous-time model. Upon estimation, these frequency-invariant structural parameters can, if desired, be mapped to parameters associated with any frequency, not only the observed data frequency.

We consider a likelihood-based framework, along the lines of Jones (1981), Harvey and Stock (1985) and Zadrozny (1988). The linearity assumption allows deriving the exact solution at the discrete observation times of the underlying continuous-time model, and the associated state space representation. Even so, the multivariate nature of the model leads to identification problems due to aliasing, i.e., that different continuous-time models can share the same discrete-time representation. We provide sufficient conditions for ruling out aliasing, and demonstrate that they are satisfied in the specific models we work out. Further, as

the residuals from the discrete-time empirical model are composites of underlying structural shocks from the continuous-time model, reflecting time-aggregation, as well as contemporaneous and dynamic relations among variables between measurements, we propose a method to approximately recover the structural shocks at measurement times. In addition, we use the continuous-time feature of the underlying model to focus on frequency-invariant structural parameters, and to accommodate either stock or flow variables in the data, as well as variables of different observation frequencies. We illustrate our approach using an off-the-shelf real business cycle model. We conduct extensive Monte Carlo experiments to study the finite sample properties of the estimator based on the exact discrete-time representation, and show they are superior to those based on a naive Euler-Maruyama discretization of the economic model. Finally, we estimate the model using postwar U.S. macroeconomic data, and offer examples of applications of our approach, including historical shock decomposition at different frequencies, and estimation based on mixed-frequency data.

The starting point of our framework is the system of equations that characterizes the solution of linear(-ized) continuous-time DSGE models. This type of model is usually represented by an autonomous system of linear stochastic differential equations (SDEs) describing the dynamics of the optimal state variables in the economy, together with a set of algebraic equations representing equilibrium conditions, optimal policy functions, or (static) no-arbitrage conditions, at any given instant. We express the solution to the system of SDEs as a discrete-time vector autoregressive (VAR) model of order one (see Bartlett and Rajalakshman, 1953, Phillips, 1959, Bergstrom, 1966, Phillips, 1973, and Bergstrom, 1984). This VAR(1) representation, usually known as the *exact discrete model* (EDM), delivers the probability distribution of any given sequence of observations sampled from the underlying continuous-time model, without introducing discretization errors that might otherwise contaminate the estimation of the model parameters, and it maintains all cross-equation restrictions implied by the economic model.[1] When combined with the algebraic equations evaluated at the same discrete points in time, we obtain a standard discrete-time and time-invariant state space representation that can be used to evaluate the likelihood function for the model using the Kalman filter, along the lines of Fernández-Villaverde and Rubio-Ramírez (2007) and Fernández-Villaverde et al. (2016). Statistical inference on the unknown value of the parameters of the continuous-time model is performed via the maximum likelihood estimator (MLE) on the basis of a sample of discrete-time measurements.

The use of state space models for conducting inference in continuous-time frameworks

[1] See McCrorie (2009) and Chambers et al. (2018) for a comprehensive review on the exact discrete-time representation of continuous-time models. Exact discrete-time representations of higher-order systems of stationary and non-stationary differential equations can be found in Chambers (1999).

from discrete observations was already acknowledged by Bergstrom (1983). The state space framework allows for estimation of models with unobserved components, and accommodates errors in measurements of the variables used for estimation, mixed frequencies of the variables, missing observations, and variables measured at unequally spaced intervals. Alternative approaches that do not rely on state space representations have also been proposed, including Hansen and Scheinkman (1995) for moment-based estimation methods, Bergstrom (1983), Lo (1988), Aït-Sahalia (2002, 2008), and Aït-Sahalia and Mykland (2003, 2004) for likelihood-based methods, Bibby and Sørensen (1995) and Sørensen (1997) for estimating function methods, and Florens-Zmirou (1993), Jiang and Knight (1997), and Fan (2005) for nonparametric techniques. Phillips and Yu (2009) provide an overview of different approaches for the estimation of continuous-time models from discrete observations. Recent applications of these methods to the estimation of DSGE models include Posch (2009), Christensen et al. (2016), Fernández-Villaverde et al. (2020), and Chambers et al. (2022).

The econometric framework proposed here is simple and convenient for estimation purposes, and the reliance on the EDM ensures that all relevant information in the observed data regarding the underlying model is retained. Nevertheless, it leads to the two challenges related to the aliasing problem and the backing out of structural shocks. Both must be addressed in order for the estimated model to be useful for economic analysis, e.g., to understand the sources of business cycle fluctuations, the propagation of exogenous shocks through the economy, or the effects of changes in macroeconomic policies. The alternative route of bypassing the two challenges by avoiding the EDM implies a loss of information. Here, we address them both.

Regarding the first challenge, the coefficient matrix of the VAR(1) representation of the EDM involves a non-linear transformation of the parameters, beyond that implied by the cross-equation restrictions associated with the rational expectations solution of the model. This additional transformation, given in terms of a matrix exponential, adds an observational equivalence problem to the list of potential identification issues that can affect DSGE model, as discussed in the literature (see, e.g., Canova and Sala, 2009, Iskrev, 2010, Komunjer and Ng, 2011, and Qu and Tkachenko, 2012, 2017). This is the *aliasing* problem, and it is particular to continuous-time models. It arises because the mapping between the coefficient matrix of the system of SDEs and the coefficient matrix of the VAR(1) is not injective. Therefore, different values of the parameters in the continuous-time model could generate the same model for the discrete observations at sampling frequency (the same EDM). Precluding aliases is necessary for local identification of the parameters in the continuous-time economic model. To address the aliasing problem, we build on the arguments in Phillips (1973), and more recently in McCrorie (2003), Kessler and Rahbek (2004), and Blevins (2017), to estab-

lish a set of sufficient conditions that rule out aliasing in multivariate settings relying on the EDM and observations in discrete time.

The second challenge is related to the fact that the transition equation in the resulting state space model implied by the EDM resembles a reduced-form VAR(1) model, with disturbances that are composites of the primitive structural shocks in the continuous-time model. Therefore, the use of the EDM for estimation purposes hinders the structural interpretation of the shocks that drive the variables in the model. The estimated reduced-form residuals reflect not only the effect of potential contemporaneous and dynamic relations among the variables, but also the effects of time-aggregation over a given interval of fixed length, i.e., the confounding of shocks that occur between measurement times. To address this issue, we propose a method to approximately recover the unobserved structural shocks of the continuous-time DSGE model at measurement times from the estimated reduced-form residuals. Our approach exploits the underlying causal links of the continuous-time model through the EDM mapping and the information contained in the sample of discrete-time observations. This strategy resembles the use of short- and long-run identifying restrictions on the variance-covariance matrix that is commonly used in the structural VAR literature, pioneered by Sims (1986), Bernanke (1986), Shapiro and Watson (1988), and Blanchard and Quah (1989).

We demonstrate how our approach can accommodate the sampling nature of the data. One of the attractive features of working with continuous-time models is that they provide a logically consistent basis for jointly accommodating variables that are sampled according to different schemes, such as stocks and flows. A stock is a variable that is measured at a given point in time, e.g., the capital stock, or bond holdings measured at the end of the period. A flow, on the contrary, is a variable whose value measures the accumulated amount over a given time interval, e.g., consumption or GDP measured from the beginning to the end of a period.

To investigate the properties of our approach, we consider a continuous-time version of the RBC model with indivisible labor of Hansen (1985), with shocks to total factor productivity and capital, as a benchmark. We run extensive Monte Carlo simulations to study the finite sample properties of the MLE, as well as the ability to recover the structural shocks, when the data used in the estimation process are sampled either as stocks or flows. Our Monte Carlo experiments shed light on the effects of model misspecification that results from using a state space model for stock variables when the observed data are in fact sampled as flows. We also show the consequences of using a state space that is derived instead from a naive discretization based on the Euler-Maruyama (EM) approximation. Our results suggest that using the EDM for the specification of the discrete-time state space model is preferred, and that the biases introduced by the alternative method are substantial. The EM approximation does not suffer from the aliasing problem, but instead from discretization

error, implying in effect that the estimates are of discrete-time parameters that are only associated with the continuous-time parameters of interest up to the degree of approximation. Similarly, discretization error in the variance-covariance structure hampers the identification and interpretation of structural shocks. While the EM approximation is popular in finance, in the presence of high-frequency data (e.g., daily, minute, second), the relatively low frequency of macroeconomic data (monthly, quarterly, annual) is particularly damaging to this method. In contrast, our approach works seamlessly.

We provide an empirical illustration of our proposed framework by estimating the benchmark model using quarterly data on macroeconomic aggregates for the U.S economy over the period from 1959:Q1 through 2019:Q4. We consider the aggregate consumption and hours worked series as a benchmark case, and compare results to those based on other data configurations, including aggregate output, and other data frequencies. The results confirm that our approach is feasible, and estimates are consistent with those in the business cycle literature. A historical shock decomposition indicates that the U.S. business cycle, as reflected in the consumption growth series, has mainly been driven by aggregate supply shocks over the period studied, whereas deviations in hours worked from the steady state have mainly been driven by aggregate demand shocks. Finally, we provide an application using data series of mixed frequencies that illustrates the generality of our approach, exploiting the frequency-invariance of the parameters of the underlying continuous-time model.

The rest of the paper is organized as follows. Section 2 presents the econometric framework, and the mapping between the continuous-time DSGE model and the analogous discrete-time state space representation. It also discusses the conditions under which the parameters of the continuous-time model can be identified from discrete-time measurements. Section 3 introduces a method to recover the sequence of the structural shocks at measurement times from the reduced-form residuals of the discrete-time state space representation. Section 4 provides Monte Carlo evidence on the finite sample properties of the MLE, and studies the accuracy of the proposed method to recover structural shocks. Section 5 presents the empirical application to U.S. data, and Section 6 concludes. All proofs of propositions, derivations, and some additional results, are provided in the Appendix.

2. The econometric framework

2.1 The economic model

Let $\mathbf{y}(t) \in \mathbb{R}^{n_y}$ denote a vector of control or jump variables at time t, and $\mathbf{x}(t) \in \mathbb{R}^{n_x}$ a vector of possibly unobserved state or predetermined variables, with $t \in \mathbb{R}$. Further, let

$\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^{\dim \boldsymbol{\theta}}$ be the vector of structural parameters characterizing preferences, technology, and/or endowments, with $\boldsymbol{\Theta}$ the parameter space of all theoretically admissible values of $\boldsymbol{\theta}$. In the following, we consider a class of continuous-time linear(ized) DGSE models whose rational expectations solution can be represented in the form

$$\begin{aligned} \mathrm{d}\mathbf{x}(t) &= \mathbf{A}(\boldsymbol{\theta})\mathbf{x}(t)\mathrm{d}t + \mathbf{B}(\boldsymbol{\theta})\mathrm{d}\mathbf{w}(t)\,, \quad \text{given } \mathbf{x}(t_0) = \mathbf{x}_0\,, && (2.1) \\ \mathbf{y}(t) &= \mathbf{C}(\boldsymbol{\theta})\mathbf{x}(t)\,, && (2.2) \end{aligned}$$

where $\mathbf{w}(t) \in \mathbb{R}^{n_w}$ is a vector of independent standard Brownian motions, $\mathbf{w}(t) \sim \mathcal{N}(\mathbf{0}, t\mathbf{I})$. By equation (2.1), $\mathbf{x}(t)$ is governed by a multivariate Ornstein-Uhlenbeck (OU) diffusion process that describes the *optimal* dynamics of the state variables, with local drift $\mathbf{A}(\boldsymbol{\theta}) \in \mathbb{R}^{n_x \times n_x}$, diffusion matrix $\mathbf{B}(\boldsymbol{\theta}) \in \mathbb{R}^{n_x \times n_w}$, $n_x \geq n_w$, and positive semi-definite instantaneous variance-covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\theta}) := \mathbf{B}(\boldsymbol{\theta})\mathbf{B}(\boldsymbol{\theta})^\top \in \mathbb{R}^{n_x \times n_x}$. The interpretation of $\mathrm{d}\mathbf{w}(t)$ is that of primitive (structural) shocks driving the economy. Equation (2.2) determines the *optimal* value of the control variables at instant t, as a function of the state variables at that time. The entries of the time-invariant matrices $\mathbf{A}(\boldsymbol{\theta})$, $\mathbf{B}(\boldsymbol{\theta})$, and $\mathbf{C}(\boldsymbol{\theta}) \in \mathbb{R}^{n_y \times n_x}$ correspond to reduced-form parameters that are nonlinear functions of the vector of structural parameters $\boldsymbol{\theta}$. The map $\boldsymbol{\theta} \mapsto (\mathbf{A}(\boldsymbol{\theta}), \mathbf{B}(\boldsymbol{\theta}), \mathbf{C}(\boldsymbol{\theta}))$ embodies the cross-equation restrictions imposed by the rational expectations solution to the DSGE model. We restrict attention to stationary models, and thus make the following assumption.

Assumption 1. For all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, the system matrix $\mathbf{A}(\boldsymbol{\theta})$ is stable, i.e., every eigenvalue of $\mathbf{A}(\boldsymbol{\theta})$ has strictly negative real part.

Remark 2.1. Given that the system matrix $\mathbf{A}(\boldsymbol{\theta})$ is stable, by Assumption 1, equations (2.1)-(2.2) define a stable linear system.

In practice, the econometrician faces two problems when using the continuous-time model (2.1)-(2.2) for statistical inference. First, the decisions made by economic agents are only measured at discrete intervals, e.g., annually, quarterly, monthly, etc. Here, it is important to differentiate between stock variables, which are sampled at a particular point in time (e.g., wages, prices, interest rates, etc.), and flow variables, which are sampled as aggregates over an interval of time (e.g., income, consumption, hours worked, etc.). Secondly, not all the variables in the model are readily available, i.e., some of the variables may be latent or unobservable. In the following, we illustrate our estimation approach by assuming that all the variables in $\mathbf{x}(t)$ are unobserved stock variables, while $\mathbf{y}(t)$ contains observable variables, all of which are sampled either as stocks or as flows. A setup that simultaneously accommodates stock and flow variables in the measurement and/or state equations can be readily

derived, e.g., following McCrorie (2000).

2.2 State space representation

To match the continuous-time model with the discrete nature of the data available, we introduce an exact discrete-time representation of the rational expectations solution that is consistent with the observations generated by the system in (2.1) and (2.2). This is achieved in two steps. First, following Bartlett and Rajalakshman (1953), Phillips (1959), Bergstrom (1966), Phillips (1973), and Bergstrom (1984), we derive a discrete-time VAR(1) process for the state vector by computing the exact solution to the differential equation in (2.1) at measurement times t_τ, $\tau \in \mathbb{Z}^+$ (with $\mathbb{Z}^+ := \{x \in \mathbb{Z} : x > 0\}$ the set of positive integers). Secondly, we evaluate the continuous-time algebraic equation for the control variables in (2.2) at measurement times. The first step is summarized in the following Proposition. Throughout, we write $\mathbf{I}$ and $\mathbf{0}$, respectively, for conformable identity and null matrices.

Proposition 1 (Exact discrete model)**.** Let $\mathbf{x}_\tau := \mathbf{x}(t_\tau)$ denote the τth realization occurring at time t_τ of the continuous-time variables $\mathbf{x}(t)$. At measurement times, equation (2.1) satisfies the VAR(1) model

$$\mathbf{x}_\tau = \mathbf{A}_h(\boldsymbol{\theta})\mathbf{x}_{\tau-1} + \boldsymbol{\eta}^s_\tau\,, \tag{2.3}$$

where $\mathbf{A}_h(\boldsymbol{\theta}) \in \mathbb{R}^{n_x \times n_x}$ is given by

$$\mathbf{A}_h(\boldsymbol{\theta}) = \exp\left(\mathbf{A}(\boldsymbol{\theta})h\right) = \sum_{i=0}^{\infty} \tfrac{(\mathbf{A}(\boldsymbol{\theta})h)^i}{i!} = \mathbf{I} + \mathbf{A}(\boldsymbol{\theta})h + \tfrac{1}{2}\mathbf{A}(\boldsymbol{\theta})^2h^2 + \dots\,, \tag{2.4}$$

with $h \geq 0$ denoting the fixed sampling interval between measurements, $h := (t_\tau - t_{\tau-1})$, and $\boldsymbol{\eta}^s_\tau$ is the $n_{\eta^s} = n_x \geq n_w$ vector of Gaussian innovations

$$\boldsymbol{\eta}^s_\tau = \int_{t_{\tau-1}}^{t_\tau} \exp\left(\mathbf{A}(\boldsymbol{\theta})(t_\tau - s)\right)\mathbf{B}(\boldsymbol{\theta})\mathrm{d}\mathbf{w}(s)\,, \tag{2.5}$$

with mean $\mathbb{E}\left[\boldsymbol{\eta}^s_\tau\right] = \mathbf{0}$ and autocovariances $\mathbb{E}[\boldsymbol{\eta}^s_\tau {\boldsymbol{\eta}^s_{\tau-\ell}}^\top] = \boldsymbol{\Sigma}_{\boldsymbol{\eta}^s,h}(\boldsymbol{\theta}) \cdot \mathbb{I}\{\ell = 0\}$, for all τ, $\ell \in \mathbb{Z}$ ($\mathbb{I}\{\cdot\}$ denotes the indicator function), where

$$\boldsymbol{\Sigma}_{\boldsymbol{\eta}^s,h}(\boldsymbol{\theta}) = \mathbb{E}[\boldsymbol{\eta}^s_\tau {\boldsymbol{\eta}^s_\tau}^\top] = \int_0^h \exp\left(\mathbf{A}(\boldsymbol{\theta})(h-s)\right)\boldsymbol{\Sigma}(\boldsymbol{\theta})\exp\left(\mathbf{A}(\boldsymbol{\theta})^\top(h-s)\right)\mathrm{d}s\,. \tag{2.6}$$

Proof. See Appendix D.1. ■

The discrete-time equation (2.3) is usually referred to as the *Exact Discrete Model* (EDM), with associated sampling frequency determined by the spacing between observations, h. Let the basic unit of time in the continuous-time model be a year. Thus, $h = 1$ refers to annual observations, $h = 1/4$ to quarterly observations, $h = 1/12$ to monthly observations, etc. Assumption 1 guarantees stationarity of the EDM. In other words, for every $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, the characteristic roots of $\mathbf{A}_h(\boldsymbol{\theta})$ have modulus less than one. Although the EDM does not account for the behavior of the state variables between measurements, the representation is exact in the sense that there are no discretization errors, and the distributions of $\mathbf{x}(t)$ and $\mathbf{x}_\tau$ coincide at measurement times t_τ. Any set of equispaced data generated from (2.3) satisfies the model in (2.1) with probability one, regardless the sampling frequency.[2]

The EDM differs from the standard discrete-time VAR(1) model in that the reduced-form parameters in the transition matrix $\mathbf{A}_h(\boldsymbol{\theta})$ are restricted in a nonlinear manner involving the matrix exponential function, and the variance-covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\eta}^s,h}(\boldsymbol{\theta})$ depends both on $\mathbf{A}(\boldsymbol{\theta})$ and $\mathbf{B}(\boldsymbol{\theta})$. We show in Appendix A how to compute these matrices using, respectively, an eigenvalue-eigenvector decomposition of the fundamental matrix $\mathbf{A}(\boldsymbol{\theta})$, and the matrix factorization approach of Van Loan (1978).

A direct consequence of the EDM is that the disturbances in (2.3) no longer represent structural shocks, but instead reduced-form innovations. By (2.5), the innovation at time t_τ can be regarded as a moving average of structural shocks $\mathrm{d}\mathbf{w}(s)$ over the time interval of length h, with time-variation in the weighting scheme, as determined by the autoregressive coefficient matrices. Geweke (1978) calls the confounding effects of the latter "contamination".

2.3 Stock variables

When the variables used as measurements are sampled as stocks, the second step amounts to evaluating (2.2) at time t_τ, i.e.,

$$\mathbf{y}_\tau = \mathbf{C}(\boldsymbol{\theta})\mathbf{x}_\tau \,, \tag{2.7}$$

where $\mathbf{y}_\tau := \mathbf{y}(t_\tau)$ defines the measurement of the control variables at time t_τ.[3]

Let $\mathbf{y}^T = \{\mathbf{y}_1, \ldots, \mathbf{y}_T\}$ be the sample of T equidistant discrete measurements of stock variables which is available to the econometrician, and potentially is subject to iid sampling or measurement errors (see, e.g., Sargent, 1989). Our empirical model is summarized by the

[2]For the more general case where the diffusion matrix in (2.1) depends on the state vector, e.g., a square root process, Nowman (1997) derives a discretization under the assumption that the volatility changes only at measurement times, then remains constant, $\mathbf{B}(\mathbf{x}(t), \boldsymbol{\theta}) \approx \mathbf{B}(\mathbf{x}(t_{\tau-1}), \boldsymbol{\theta})$. However, this approach induces some approximation error (see Yu and Phillips (2001) for a discussion).

[3]In the presence of deterministic or stochastic trends, Equation (2.7) can alternatively be written as $\Delta\mathbf{y}_\tau = \mathbf{y}_\tau - \mathbf{y}_{\tau-1} = \widetilde{\mathbf{C}}(\boldsymbol{\theta})\widetilde{\mathbf{x}}_\tau$, where $\widetilde{\mathbf{x}}_\tau = \left[\mathbf{x}_\tau^\top, \mathbf{y}_{\tau-1}^\top\right]^\top$ is an extended state vector, and $\widetilde{\mathbf{C}}(\boldsymbol{\theta}) = [\mathbf{C}(\boldsymbol{\theta}), -\mathbf{I}]$ (see Pfeifer, 2020).

discrete-time linear time-invariant state space system

$$\mathbf{x}_\tau = \mathbf{A}_h(\boldsymbol{\theta})\mathbf{x}_{\tau-1} + \boldsymbol{\eta}^s_\tau \,, \tag{2.8}$$

$$\mathbf{y}_\tau = \mathbf{C}(\boldsymbol{\theta})\mathbf{x}_\tau + \boldsymbol{\varepsilon}_\tau \,, \tag{2.9}$$

where (2.8) defines the transition equation for the latent state variables, with $\boldsymbol{\eta}^s_\tau$ from (2.5), and (2.9) the measurement equation, with $\boldsymbol{\varepsilon}_\tau$ an $n_y \times 1$ vector of serially uncorrelated Gaussian measurement errors satisfying, for all τ and ℓ, $\mathbb{E}[\boldsymbol{\varepsilon}_\tau] = \mathbf{0}$, $\mathbb{E}\left[\boldsymbol{\varepsilon}_\tau \boldsymbol{\varepsilon}^\top_{\tau-\ell}\right] = \mathbf{R} \cdot \mathbb{I}\{\ell = 0\}$, and $\mathbb{E}\left[\boldsymbol{\eta}^s_\tau \boldsymbol{\varepsilon}^\top_{\tau-\ell}\right] = \mathbf{0}$.[4] In the following, we refer to the Gaussian state space representation in (2.8)-(2.9) as the S-SSR model.[5]

2.4 Flow variables

When the variables used as measurements are sampled instead as flows, the second step of our procedure is more involved. While stocks are measured as the value of the variables at a particular point in time, flow variables such as consumption or income are defined over a particular interval of time. To accommodate the flow nature of the measurements in the state space representation, recall that, according to (2.2), the optimal values of the control variables, $\mathbf{y}(t)$, are determined in equilibrium by the state of the economy, $\mathbf{x}(t)$. Following Harvey (1990), we exploit this relation and introduce the time t_τ cumulator variable $\mathbf{y}^f(t_\tau) = \int_{t_{\tau-1}}^{t_\tau} \mathbf{y}(s)\mathrm{d}s$, measuring the cumulated values of the control variables over the time interval $[t_{\tau-1}, t_\tau]$. Here and in the following, we use the superscript f to denote the *flow* nature of the variables under consideration. Substituting in the optimal policy function from the model yields

$$\mathbf{y}^f(t_\tau) = \int_{t_{\tau-1}}^{t_\tau} \mathbf{C}(\boldsymbol{\theta})\mathbf{x}(s)\mathrm{d}s = \mathbf{C}(\boldsymbol{\theta}) \int_0^h \mathbf{x}(t_{\tau-1}+s)\mathrm{d}s \,,$$

where the dynamics of the state vector are given in (2.1). Let $\mathbf{y}^f_\tau := \mathbf{y}^f(t_\tau)$, for all τ. Under the maintained assumption that the state variables are measured as stocks, the state vector has the EDM representation in Proposition 1, and the cumulator variable at measurement times can be written as

$$\mathbf{y}^f_\tau = \mathbf{C}(\boldsymbol{\theta}) \left[\int_0^h \exp(\mathbf{A}(\boldsymbol{\theta})s)\mathrm{d}s \right] \mathbf{x}_{\tau-1} + \boldsymbol{\eta}^f_\tau \,, \tag{2.10}$$

[4] The empirical model can be augmented to accommodate serially correlated measurement errors with VAR(1) dynamics, following Ireland (2004).

[5] In our empirical work, we test the Gaussianity assumption. If necessary, the approach can be adapted as QML.

where $\boldsymbol{\eta}_\tau^f$ is an $n_y \times 1$ vector of normally distributed reduced-form disturbances, with mean $\mathbb{E}\left[\boldsymbol{\eta}_\tau^f\right] = \mathbf{0}$, and variance-covariance matrix $\mathbb{E}\left[\boldsymbol{\eta}_\tau^f \boldsymbol{\eta}_\tau^{f,\top}\right] = \boldsymbol{\Sigma}_{\boldsymbol{\eta}^f,h}(\boldsymbol{\theta})$. Using (2.5) and the fact that, by Assumption 1, the matrix $\mathbf{A}(\boldsymbol{\theta})$ is nonsingular, so that $\int_0^h \exp(\mathbf{A}(\boldsymbol{\theta})s)\mathrm{d}s = \mathbf{A}(\boldsymbol{\theta})^{-1}(\mathbf{A}_h(\boldsymbol{\theta}) - \mathbf{I})$, it follows that

$$\boldsymbol{\eta}_\tau^f = \mathbf{C}(\boldsymbol{\theta}) \int_{t_{\tau-1}}^{t_\tau} \mathbf{A}(\boldsymbol{\theta})^{-1} \left(\exp\left(\mathbf{A}(\boldsymbol{\theta})(t_\tau - s)\right) - \mathbf{I}\right) \mathbf{B}(\boldsymbol{\theta}) \mathrm{d}\mathbf{w}(s)\,. \tag{2.11}$$

The complete derivation of (2.11) is reported in Appendix D.2.

For a sample of T equidistant discrete measurements of flow variables, $\mathbf{y}^T = \{\mathbf{y}_1, \ldots, \mathbf{y}_T\}$, the continuous-time model (2.1)-(2.2) has a linear discrete-time time-invariant Gaussian state space representation given by

$$\begin{bmatrix} \mathbf{x}_\tau \\ \mathbf{y}_\tau^f \end{bmatrix} = \begin{bmatrix} \mathbf{A}_h(\boldsymbol{\theta}) & \mathbf{0} \\ \mathbf{C}(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})^{-1}(\mathbf{A}_h(\boldsymbol{\theta}) - \mathbf{I}) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{\tau-1} \\ \mathbf{y}_{\tau-1}^f \end{bmatrix} + \boldsymbol{\eta}_\tau\,, \tag{2.12}$$

$$\mathbf{y}_\tau = \begin{bmatrix} \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{x}_\tau \\ \mathbf{y}_\tau^f \end{bmatrix} + \boldsymbol{\varepsilon}_\tau\,. \tag{2.13}$$

Equation (2.12) defines the augmented transition equation for the latent state variables $\left[\mathbf{x}_\tau^\top, \mathbf{y}_\tau^{f\top}\right]^\top$. Moreover, $\boldsymbol{\eta}_\tau = \left[\boldsymbol{\eta}_\tau^{s\top}, \boldsymbol{\eta}_\tau^{f\top}\right]^\top$ is an $n_\eta = n_x + n_y$ dimensional vector of reduced-form disturbances, with mean $\mathbb{E}\left[\boldsymbol{\eta}_\tau\right] = \mathbf{0}$, variance-covariance matrix

$$\boldsymbol{\Sigma}_{\boldsymbol{\eta},h}(\boldsymbol{\theta}) = \mathbb{E}\left[\boldsymbol{\eta}_\tau \boldsymbol{\eta}_\tau^\top\right] = \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{\eta}^s,h}(\boldsymbol{\theta}) & \boldsymbol{\Sigma}_{\boldsymbol{\eta}^s\boldsymbol{\eta}^f,h}(\boldsymbol{\theta}) \\ \boldsymbol{\Sigma}_{\boldsymbol{\eta}^s\boldsymbol{\eta}^f,h}(\boldsymbol{\theta})^\top & \boldsymbol{\Sigma}_{\boldsymbol{\eta}^f,h}(\boldsymbol{\theta}) \end{bmatrix},$$

with $\boldsymbol{\Sigma}_{\boldsymbol{\eta}^s,h}(\boldsymbol{\theta})$ given in (2.6),

$$\boldsymbol{\Sigma}_{\boldsymbol{\eta}^s\boldsymbol{\eta}^f,h}(\boldsymbol{\theta}) = \left[\int_0^h \int_0^s \exp(\mathbf{A}(\boldsymbol{\theta})(s-r))\boldsymbol{\Sigma}(\boldsymbol{\theta}) \exp(\mathbf{A}(\boldsymbol{\theta})^\top r)\mathrm{d}r\mathrm{d}s\right] \mathbf{C}(\boldsymbol{\theta})^\top\,, \tag{2.14}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{\eta}^f,h}(\boldsymbol{\theta}) = \mathbf{C}(\boldsymbol{\theta}) \left[\int_0^h \int_0^s \exp(\mathbf{A}(\boldsymbol{\theta})r)\boldsymbol{\Sigma}(\boldsymbol{\theta}) \exp(\mathbf{A}(\boldsymbol{\theta})^\top r)\mathrm{d}r\mathrm{d}s\right] \mathbf{C}(\boldsymbol{\theta})^\top\,, \tag{2.15}$$

and autocovariance matrix $\mathbb{E}\left[\boldsymbol{\eta}_\tau \boldsymbol{\eta}_{\tau-\ell}^\top\right] = \mathbf{0}$, for all τ and $\ell \neq 0$. Equation (2.13) defines the measurement equation augmented with normally distributed and serially uncorrelated measurement errors, $\boldsymbol{\varepsilon}_\tau$, satisfying additionally $\mathbb{E}\left[\boldsymbol{\eta}_\tau^f \boldsymbol{\varepsilon}_{\tau-\ell}^\top\right] = \mathbf{0}$, for all τ, ℓ. It states that the observed flows are given by the latent cumulator variables plus measurement error. Similarly

to the case of stock variables, the computation of (2.14) and (2.15) is carried out via the matrix factorization approach of Van Loan (1978), described in Appendix A. Below, we refer to the state space representation (2.12)-(2.13) as the F-SSR model. Relative to the S-SSR, (2.8)-(2.9), the presence of flow variables increases the number of unobserved states in the transition equation from n_x to $n_x + n_y$.[6]

Because of the different state space representations for stock and flow variables, we now introduce an alternative representation that encompasses both the S-SSR and F-SSR models.

Definition 1 (ABCD representation)**.** The state space systems (2.8)-(2.9) and (2.12)-(2.13) admit the representation (see Fernández-Villaverde et al., 2007)

$$\mathbf{x}_{\tau+1} = \mathsf{A}(\boldsymbol{\theta})\mathbf{x}_\tau + \mathsf{B}(\boldsymbol{\theta})\boldsymbol{\epsilon}_{\tau+1}\,, \tag{2.16}$$

$$\mathbf{y}_{\tau+1} = \mathsf{C}(\boldsymbol{\theta})\mathbf{x}_\tau + \mathsf{D}(\boldsymbol{\theta})\boldsymbol{\epsilon}_{\tau+1}\,. \tag{2.17}$$

for conformable matrices ($\mathsf{A}(\boldsymbol{\theta})$, $\mathsf{B}(\boldsymbol{\theta})$, $\mathsf{C}(\boldsymbol{\theta})$, $\mathsf{D}(\boldsymbol{\theta})$), where $\boldsymbol{\epsilon}_\tau$ is the vector of errors of the systems, which include stacked reduced-form innovations, $\boldsymbol{\eta}^s_\tau$ and $\boldsymbol{\eta}^f_\tau$, and iid measurement errors, $\boldsymbol{\varepsilon}_\tau$. From the definition of the $\boldsymbol{\eta}$'s and $\boldsymbol{\varepsilon}$'s, it follows that $\boldsymbol{\epsilon}_\tau$ is white noise, i.e., for all τ, ℓ, $\mathbb{E}[\boldsymbol{\epsilon}_\tau] = \mathbf{0}$, and $\mathbb{E}[\boldsymbol{\epsilon}_\tau \boldsymbol{\epsilon}_{\tau-\ell}^\top] = \boldsymbol{\Sigma}_\epsilon(\boldsymbol{\theta}) \cdot \mathbb{I}\{\ell = 0\}$. The dimension of $\boldsymbol{\epsilon}_\tau$ depends on the features of the system. The entries of the matrices of the system (2.16)-(2.17) can be found in Appendix B.1.

Lemma 1. Under Assumption 1, for every $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, and $z \in \mathbb{C}$, $\det(\mathbf{I} - \mathsf{A}(\boldsymbol{\theta})z) = 0$ implies $|z| > 1$.

Proof. See Appendix D.3. ■

The lemma implies that if the stability condition for the continuous-time system (Assumption 1) holds, then the ABCD system (2.16)-(2.17) is stable, regardless whether the measurements are sampled as stocks or flows.

2.5 Likelihood evaluation

To estimate the unknown parameters $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ of the continuous-time model in (2.1)-(2.2) based on a random sample of discrete measurements $\mathbf{y}^T$ using the discrete-time state space representation in (2.8)-(2.9) or in (2.12)-(2.13), we use the method of maximum likelihood (ML). Under the maintained assumption that the state variables are unobserved, and that

[6] Alternative state space representations that accommodate the presence of flow variables have been proposed in the literature (see, e.g., Harvey, 1990). A study of their computational differences relative to the approach suggested here is left for future research.

the econometrician only has access to a sample of discrete-time observations of the control variables, possibly contaminated by measurement error, the *exact* likelihood function for the data implied by the economic model can be constructed and evaluated using the Kalman filter algorithm (see, e.g., Harvey, 1990, Hamilton, 1994, or Durbin and Koopman, 2012). To ensure that the Kalman gain exists, and that the Kalman filter recursion converges, we make the following additional assumption.

Assumption 2. For every $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, $\mathsf{D}(\boldsymbol{\theta})\boldsymbol{\Sigma}_{\epsilon}(\boldsymbol{\theta})\mathsf{D}(\boldsymbol{\theta})^{\top}$ is nonsingular.

Assumption 2 rules out the possibility that the state space representation is stochastically singular (see Ingram et al., 1994, Ruge-Murcia, 2007, and Fernández-Villaverde et al., 2016). It is equivalent to assuming the existence of a unique solution to the discrete algebraic Riccati equation in the Kalman filter recursion. The issue of singularity emerges in DSGE models whenever the number of observable variables exceeds the number of disturbances in the discrete-time state space representation, $n_y > \text{rank}(\boldsymbol{\Sigma}_{\boldsymbol{\eta},h}(\boldsymbol{\theta})) + \text{rank}(\mathbf{R})$.

Using the Kalman filter recursion, the conditional log-likelihood function, given $\mathbf{y}_0$, can be constructed recursively via the prediction error decomposition as

$$\mathcal{L}\left(\boldsymbol{\theta}|\mathbf{y}^T\right) = \sum_{\tau=1}^{T} \ln f\left(\mathbf{y}_\tau|\mathbf{y}_{\tau-1};\boldsymbol{\theta}\right) , \tag{2.18}$$

where $f\left(\mathbf{y}_\tau|\mathbf{y}_{\tau-1};\boldsymbol{\theta}\right)$ is the density function of the discrete measurements at time t_τ conditional on the information set at time $t_{\tau-1}$. Given the linear structure of the Gaussian state space model, $f\left(\mathbf{y}_\tau|\mathbf{y}_{\tau-1};\boldsymbol{\theta}\right)$ is multivariate normal with first- and second-order moments determined by the one-step-ahead forecast errors for the measurements, and their associated variance-covariance matrix. The MLE of $\boldsymbol{\theta}$ is then given by

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\arg\max}\ \mathcal{L}\left(\boldsymbol{\theta}|\mathbf{y}^T\right) ,$$

which, under regularity conditions, delivers consistent and asymptotically normal estimates of the parameters of the model.[7] Given $\hat{\boldsymbol{\theta}}$, it is possible to use the information content of the entire sample to predict the unobserved states, $\{\mathbf{x}_\tau\}_{\tau=1}^T$, and disturbances, $\{\boldsymbol{\eta}_\tau\}_{\tau=1}^T$. A detailed derivation of the Kalman filter state and disturbance-smoothing recursions, and the likelihood function is given in Appendix B.

It is convenient to introduce the (time-invariant) innovations representation of the state space model (see Anderson and Moore, 2012). The importance of this alternative repre-

[7] The regularity conditions include i) stability of the EDM (see Assumption 1), and ii) $\boldsymbol{\theta}_0 \in \text{Int}\boldsymbol{\Theta}$ (see Newey and McFadden, 1994).

sentation of the empirical model becomes clear below, when we establish conditions for the local identification of the model parameters. The next proposition shows how to obtain the innovations representation from the ABCD representation in (2.16)-(2.17).

Proposition 2. Under Assumptions 1 and 2, and for all $\tau = 1, .., T-1$, the system (2.16)-(2.17) admits the (time-invariant) innovations representation

$$\mathbf{x}_{\tau+1|\tau+1} = \mathsf{A}(\boldsymbol{\theta})\mathbf{x}_{\tau|\tau} + \mathbf{K}(\boldsymbol{\theta})\boldsymbol{\omega}_{\tau+1|\tau}\,, \tag{2.19}$$

$$\mathbf{y}_{\tau+1} = \mathsf{C}(\boldsymbol{\theta})\mathbf{x}_{\tau|\tau} + \boldsymbol{\omega}_{\tau+1|\tau}\,, \tag{2.20}$$

where $\mathbf{K}(\boldsymbol{\theta})$ is the (time-invariant) Kalman gain, $\mathbf{x}_{\tau|\tau}$ is the contemporaneous prediction of the state vector, and $\boldsymbol{\omega}_{\tau+1|\tau}$ the one-step-ahead forecast error for the control variables, each conditional on the information available at time t_τ, and with (time-invariant) forecast error covariance matrix $\boldsymbol{\Omega}(\boldsymbol{\theta})$. All these quantities are obtained through the Kalman filter recursion, as detailed in Appendix B.

2.6 Identification of structural parameters

The consistency of the MLE relies on the ability to identify the true parameters, $\boldsymbol{\theta}_0$, given the distribution of the data, $\mathbf{y}^T$. In general, (globally) identification requires that if $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$, then $\mathcal{L}\left(\boldsymbol{\theta}|\mathbf{y}^T\right) \neq \mathcal{L}\left(\boldsymbol{\theta}_0|\mathbf{y}^T\right)$, for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. In this case, the objective function $\mathcal{L}\left(\boldsymbol{\theta}|\mathbf{y}^T\right)$ reaches a unique maximum at $\boldsymbol{\theta}_0$ in large samples (see Rothenberg, 1971 and Newey and McFadden, 1994). Lack of identification is therefore directly linked to the shape and curvature of the probability distribution implied by the model. This is illustrated in Canova and Sala (2009), where different types of identification problems are put forward in the case of linearized DSGE models.

Since the probability distribution in linear Gaussian models is completely characterized by its first two moments, the problem of identification can be analyzed through the mapping between the parameters of the economic model and the reduced-form parameters of the corresponding empirical state space representation, i.e.,

$$\boldsymbol{\theta} \mapsto \left(\mathbf{A}_h(\boldsymbol{\theta}), \boldsymbol{\Sigma}_{\boldsymbol{\eta},h}(\boldsymbol{\theta}), \mathbf{C}(\boldsymbol{\theta})\right).$$

Using this relation, Komunjer and Ng (2011) provide necessary and sufficient conditions for local (i.e., in an open neighborhood of $\boldsymbol{\theta}_0$) identification of parameters in linearized DSGE models from the first and second moments of the data. Alternative local identification conditions have been studied by Iskrev (2010) and Qu and Tkachenko (2012), whereas Qu and Tkachenko (2017) derive necessary and sufficient conditions for global identification.

However, the validity of these conditions for the identification analysis in continuous-time DSGE models estimated from discrete-time measurements becomes blurred, because of the *aliasing* problem. This represents an additional identification failure, different in nature from those discussed in Canova and Sala (2009). The aliasing problem is specific to the analysis of multivariate continuous-time models based on discrete-time data, and applies regardless the choice of estimator. It results from the nonlinear mapping between the reduced-form parameters in the solution to the model and the reduced-form parameters in the EDM. That is, for a given $\boldsymbol{\theta}$ and sampling frequency h, the focus is on the mapping

$$(\mathbf{A}(\boldsymbol{\theta}), \boldsymbol{\Sigma}(\boldsymbol{\theta})) \mapsto (\mathbf{A}_h(\boldsymbol{\theta}), \boldsymbol{\Sigma}_{\boldsymbol{\eta},h}(\boldsymbol{\theta})) \ ,$$

defined by equations (2.4) and (2.6). Hence, a precondition to achieve (local) identification of $\boldsymbol{\theta}$ based on a sample of discrete measurements is to rule out observational equivalence due to aliasing.

2.6.1 The aliasing problem

Let us first consider the identification of the coefficient matrix $\mathbf{A}(\boldsymbol{\theta})$ in isolation. By Proposition 1, the matrix of reduced-form parameters in the transition equation (2.1) and the matrix of coefficients in the EDM in (2.8) are linked through the exponential mapping $\mathbf{A}_h(\boldsymbol{\theta}) = \exp(\mathbf{A}(\boldsymbol{\theta})h)$ (see equation 2.4), which is non-injective in multivariate settings. Hence, the inverse mapping is not uniquely defined. In general, a countably infinite number of continuous-time matrices $\mathbf{A}(\boldsymbol{\theta})$ lead to the same discrete-time matrix $\mathbf{A}_h(\boldsymbol{\theta})$. Therefore, given $\mathbf{B}(\boldsymbol{\theta})$ and $\mathbf{C}(\boldsymbol{\theta})$, aliases of $\mathbf{A}(\boldsymbol{\theta}_0)$ may lead to the same probability distribution for the data, $f(\mathbf{y}^T|\boldsymbol{\theta})$, and hence lack of identification. Thus, it is the non-uniqueness of the logarithm of a matrix that leads to the aliasing phenomenon. The problem is absent in the univariate case (see Remark 2.3 below).

To provide some intuition, Figure 1 illustrates the aliasing problem for the case of a bivariate deterministic continuous-time model, sampled at discrete intervals of length $h = 1$. The plot shows three time series of the two variables, $x_1(t)$ and $x_2(t)$, for three different models that are characterized by different matrices $\mathbf{A}_i$, $i = 0, 1, 2$, but are observationally equivalent at measurement times $t_\tau = 0, 1, \ldots, 4$. Thus, in the presence of aliasing identification problems, the entries of the matrix $\mathbf{A}(\boldsymbol{\theta})$, and hence the parameters of the economy, will not be uniquely identified from a sample of discrete-time observations.

The identification problem associated with aliasing was formally introduced in economics by Phillips (1973), and we proceed in the same way, by making the following assumption.

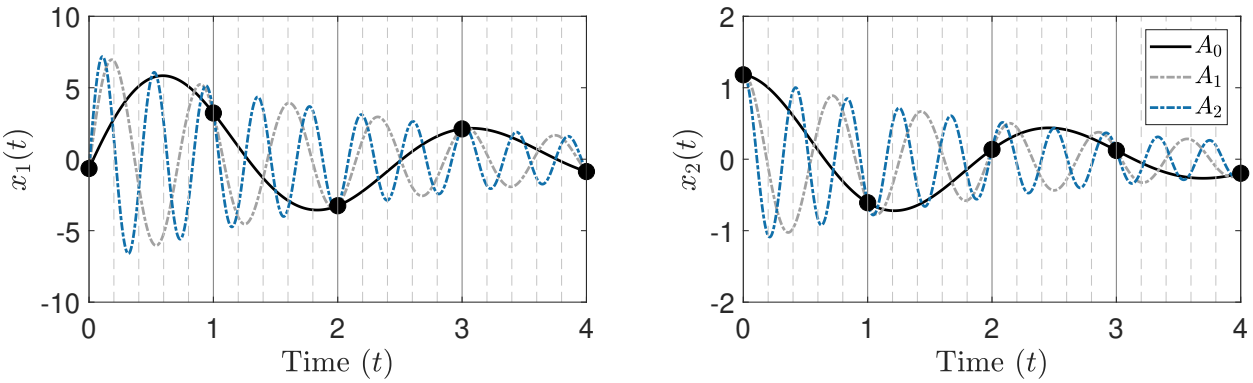


Figure 1. Aliasing phenomenon. The figure plots simulated paths for two variables, $\mathbf{X} = [x_1, x_2]^\top$, with dynamics $\mathrm{d}\mathbf{X} = \mathbf{A}_i\mathbf{X}\mathrm{d}t$, along $t \in [0, 4]$, with fixed initial values $\mathbf{X}(0) = \mathbf{X}_0$, for different coefficient matrices $\mathbf{A}_i$, $i = 0, 1, 2$. The black dots indicate discrete-time measurements generated at integer times $t_\tau = 0, 1, \ldots, 4$ from the EDM $\mathbf{X}_\tau = \exp(\mathbf{A}_i h)\mathbf{X}_{\tau-1}$, with $h = 1$. The details behind this example can be found in Appendix A.3.

Assumption 3. (i) The matrix $\mathbf{A}(\boldsymbol{\theta}_0)$ is real, with real eigenvalues, and (ii) no Jordan block belonging to any eigenvalue appears more than once.

Proposition 3. Under Assumptions 1 and 3, the mapping $\mathbf{A}_h(\boldsymbol{\theta}_0) = \exp(\mathbf{A}(\boldsymbol{\theta}_0)h)$ is injective, and hence the entries of $\mathbf{A}(\boldsymbol{\theta}_0)$ are identified from $\mathbf{A}_h(\boldsymbol{\theta}_0)$.

Proof. See Culver (1966, Theorem 2). ■

Remark 2.2. Let λ_i, $i = 1, \ldots, n_x$ be the eigenvalues of $\mathbf{A}(\boldsymbol{\theta}_0)$, i.e., $\lambda_i \in \lambda(\mathbf{A}(\boldsymbol{\theta}_0)) = \{\lambda : \det(\lambda\mathbf{I} - \mathbf{A}(\boldsymbol{\theta}_0)) = 0\}$. Some of the eigenvalues may not be distinct, i.e., an eigenvalue may occur with multiplicity greater than one. Let the Jordan block associated with the eigenvalue λ_i be denoted as $\mathbf{J}_i$. It is related to the Jordan canonical form of $\mathbf{A}(\boldsymbol{\theta}_0) = \mathbf{S}\mathbf{J}\mathbf{S}^{-1}$, with $\mathbf{J} = \mathrm{diag}(\mathbf{J}_1, ..., \mathbf{J}_k)$. Here, $k \leq n_x$ is the number of Jordan blocks of $\mathbf{A}(\boldsymbol{\theta}_0)$, which may differ from the number of distinct eigenvalues. The ith Jordan block exhibits λ_i on the main diagonal, has 1's along the first superdiagonal, and 0's everywhere else.[8] Further, $\mathbf{S}$ is an n_x-dimensional nonsingular matrix that transforms $\mathbf{A}(\boldsymbol{\theta}_0)$ into $\mathbf{J}$ (see, e.g., Gantmacher (1959) for details). The identity matrix is an example of a matrix with repeated Jordan blocks, and is therefore not identifiable from its exponential. For example, the two-dimensional identity matrix has eigenvalues $\lambda_1 = \lambda_2 = 1$, and two repeated Jordan blocks, $\mathbf{J}_1^I$ and $\mathbf{J}_2^I$, of order 1.

Let us now consider the identification of $\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)$. Without any further assumptions on the system matrices, we have the following result.

Proposition 4. Under Assumptions 1 and 3, the entries of $\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)$ are identified from $\boldsymbol{\Sigma}_{\boldsymbol{\eta},h}(\boldsymbol{\theta}_0)$.

[8] The first superdiagonal refers to the entries just above the main diagonal of a matrix.

Proof. See Appendix D.4. ■

Assumptions 1 and 3 define a set of sufficient conditions for the *sequential identification* of $(\mathbf{A}(\boldsymbol{\theta}_0), \boldsymbol{\Sigma}(\boldsymbol{\theta}_0))$ from discrete-time measurements. However, as pointed out by McCrorie (2009), there may be circumstances under which the first part of Assumption 3 is too restrictive, e.g., in models with periodic or cyclical behavior of macroeconomic variables that imply complex eigenvalues. In such cases, additional restrictions, beyond those implied by Assumptions 1 and 3, must be imposed *a priori*, to rule out aliasing identification problems. Two kinds of restrictions have been discussed in the literature. The first is based on the use of homogeneous linear restrictions, similar to the *exclusion restrictions* or *linear within-equation restrictions* used in the simultaneous equations models of the Cowles Comission for Economic Research. Phillips (1973) introduced these additional restrictions, and argued that at least $\lfloor n_x/2 \rfloor$ restrictions are necessary to rule out the aliasing identification problem, where $\lfloor \cdot \rfloor$ indicates the largest integer smaller than the argument. This condition was later extended by Blevins (2017) to include the case of inhomogeneous linear restrictions.[9] The second type relies on the *non-linear cross-equation restrictions* imposed by the assumption of rational expectations on the mapping $\boldsymbol{\theta} \mapsto (\mathbf{A}(\boldsymbol{\theta}), \boldsymbol{\Sigma}(\boldsymbol{\theta}), \mathbf{C}(\boldsymbol{\theta}))$, which Hansen and Sargent (1991) argue may provide a sufficient number of conditions to limit the number of admissible perturbations that generate the aliases of $(\mathbf{A}(\boldsymbol{\theta}_0), \boldsymbol{\Sigma}(\boldsymbol{\theta}_0))$.[10] These assumptions have also been challenged, in Hansen and Sargent (1983, Theorem 3), as they may not lead to $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ being positive semi-definite, an assumption we maintain throughout (cf. Section 2.1). To ensure the latter in the presence of complex eigenvalues, Hansen and Sargent (1983) propose exploiting the information content in the discrete-time variance-covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\eta},h}(\boldsymbol{\theta})$, to further restrict the number of potential aliases.

Remark 2.3. In the absence of aliasing, it follows from (2.4) that, given the basic unit of time in the economic model, the entries of the matrix $\mathbf{A}(\boldsymbol{\theta})$ in the continuous-time model are invariant to the sampling frequency h, as they should be. This implies that one can study the properties of a time series at any frequency $\tilde{h}$, using the matrix $\mathbf{A}(\boldsymbol{\theta})$ estimated from data at any other frequency $h \neq \tilde{h}$ (e.g., through the computation of conditional and unconditional moments, impulse-response functions, shock decompositions,

[9] A supplementary computational toolbox that adapts the identification framework of Blevins (2017) is available from the authors' webpage. The code allows testing whether the restrictions from the macroeconomic model are sufficient to rule out aliases and jointly identify the reduced-form entries of $\mathbf{A}(\boldsymbol{\theta})$ and $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ from the discrete-time measurements.

[10] An alternative source of identification, advocated in Hamerle et al. (1991), relies on the ability to sample the continuous-time model more frequently. In fact, Hansen and Sargent (1983, Theorem 5) show that, under certain conditions, there exists an $h^\star$, such that for $h \leq h^\star$, the matrices of coefficients $(\mathbf{A}(\boldsymbol{\theta}_0), \boldsymbol{\Sigma}(\boldsymbol{\theta}_0))$ are identified from $(\mathbf{A}_h(\boldsymbol{\theta}_0), \boldsymbol{\Sigma}_{\boldsymbol{\eta},h}(\boldsymbol{\theta}_0))$. However, this approach is of limited use here, due to the low-frequency sampling that characterizes macroeconomic time series.

etc.). For example, consider the univariate Ornstein-Uhlenbeck process $\{Z(t)\}_{t\geq 0}$, with drift and diffusion parameters $\boldsymbol{\theta} = (\rho_z, \sigma_z)^\top$. In this case, $a(\boldsymbol{\theta}) := \mathbf{A}(\boldsymbol{\theta}) = -\rho_z \in \mathbb{R}$, and $b(\boldsymbol{\theta}) := \mathbf{B}(\boldsymbol{\theta}) = \sigma_z \in \mathbb{R}$. By Proposition 1, the associated EDM is the AR(1) process $Z_\tau = a_h(\boldsymbol{\theta})Z_{\tau-1} + \eta_\tau$, with $h = (t_\tau - t_{\tau-1})\ \forall \tau$, $(\star)\ a_h(\boldsymbol{\theta}) = \exp(-\rho_z h)$, $\eta_\tau \sim \mathcal{N}(0, \sigma^2_{\eta,h}(\boldsymbol{\theta}))$, and $(\star\star)\ \sigma_{\eta,h}(\boldsymbol{\theta}) = b(\boldsymbol{\theta})\sqrt{(1-\exp(-2a(\boldsymbol{\theta})h))/(2a(\boldsymbol{\theta}))}$. For $(a_h(\boldsymbol{\theta}), \sigma_{\eta,h}(\boldsymbol{\theta}))^\top$ estimated from a sample of discrete measurements $\{Z_\tau\}_{\tau=0}^T$ with uniform time step h, we recover from the intermediate steps $(\star)$ and $(\star\star)$ the continuous-time parameters as $a(\boldsymbol{\theta}) = -\rho_z = \log(a_h(\boldsymbol{\theta}))/h$, and $b(\boldsymbol{\theta}) = \sigma_z = \sigma_{\eta,h}(\boldsymbol{\theta})\left((1-\exp(-2\rho_z h))/(2\rho_z)\right)^{-1/2}$. The transformations $a_{\tilde{h}}(\hat{\boldsymbol{\theta}})$ and $\sigma_{\eta,\tilde{h}}(\hat{\boldsymbol{\theta}})$ are then valid for any new system at arbitrary frequency $\tilde{h} \neq h$.

2.6.2 (Local) Identification of $\boldsymbol{\theta}_0$

In the absence of aliasing problems, it is possible to assess the local identification of $\boldsymbol{\theta}_0$ through the identification tools for DSGE models developed by Komunjer and Ng (2011). In particular, we want to investigate whether changes in some of the model parameters lead to indistinguishable outcomes, i.e., if the model is subject to observational equivalence. To proceed, we make the following assumption.

Assumption 4. The innovations representation (2.19)-(2.20) is minimal. That is, the matrices

$$[\mathbf{K}(\boldsymbol{\theta}), \mathsf{A}(\boldsymbol{\theta})\mathbf{K}(\boldsymbol{\theta}), ..., \mathsf{A}^{n_x-1}(\boldsymbol{\theta})\mathbf{K}(\boldsymbol{\theta})] \quad \text{and} \quad [\mathsf{C}(\boldsymbol{\theta})^\top, \mathsf{A}(\boldsymbol{\theta})^\top\mathsf{C}(\boldsymbol{\theta})^\top, ..., \mathsf{A}^{n_x-1}(\boldsymbol{\theta})^\top\mathsf{C}(\boldsymbol{\theta})^\top]$$

have full column rank.

Assumption 4 implies that the dimension of the state equation is the smallest possible, which is necessary for formalizing the notion of observational equivalence of $\boldsymbol{\theta}_0$ and $\boldsymbol{\theta}_1$ in the discrete-time state space system. The two matrices defined by this additional assumption are usually referred to as the reachability and observability matrices, respectively (see Hannan and Deistler, 2012).

Following Komunjer and Ng (2011), we investigate the (local) identification of parameters in continuous-time linear DSGE models with data sampled at discrete points in time by first defining the mapping

$$\delta(\boldsymbol{\theta}, \mathbf{T}) := \left[\text{vec}(\mathbf{T}\mathsf{A}(\boldsymbol{\theta})\mathbf{T}^{-1})^\top, \text{vec}(\mathbf{T}\mathbf{K}(\boldsymbol{\theta}))^\top, \text{vec}(\mathsf{C}(\boldsymbol{\theta})\mathbf{T}^{-1})^\top, \text{vech}(\boldsymbol{\Omega}(\boldsymbol{\theta}))^\top\right]^\top, \tag{2.21}$$

where $\mathbf{T}$ is any conformable square matrix of full rank, and vec and vech are the vectorization and half-vectorization operators, respectively. Under Assumptions 1, 2, 3, and 4, the parameter vector $\boldsymbol{\theta}$ is locally identified from the autocovariances of $\mathbf{y}_\tau$ at a point $\boldsymbol{\theta}_0 \in \boldsymbol{\Theta}$ if the system of equations $\delta(\boldsymbol{\theta}_0, \mathbf{I}) = \delta(\boldsymbol{\theta}_1, \mathbf{T})$ has a locally unique solution given

by $(\boldsymbol{\theta}_1, \mathbf{T}) = (\boldsymbol{\theta}_0, \mathbf{I})$. Assumption 3 is imposed to rule out the presence of aliases. While $\mathbf{T}$ can be used to control for observational equivalence between $\boldsymbol{\theta}_0$ and $\boldsymbol{\theta}_1$ in the minimal ABCD representation, it is not sufficient to rule out aliases. If the mapping $\boldsymbol{\theta} \mapsto [\text{vec}(\mathsf{A}(\boldsymbol{\theta}))^\top, \text{vec}(\mathbf{K}(\boldsymbol{\theta}))^\top, \text{vec}(\mathsf{C}(\boldsymbol{\theta}))^\top, \text{vech}(\boldsymbol{\Omega}(\boldsymbol{\theta}))^\top]^\top$ is assumed to be continuously differentiable in $\boldsymbol{\Theta}$, Komunjer and Ng (2011, Proposition 2-NS) provide necessary and sufficient rank and order conditions to locally identify $\boldsymbol{\theta}_0$ from the matrix of partial derivatives of $\delta(\boldsymbol{\theta}, \mathbf{T})$ evaluated at $(\boldsymbol{\theta}_0, \mathbf{I})$. Furthermore, the results in Komunjer and Ng (2011, Proposition 3) provide a necessary and sufficient rank condition for local identification when a subset of the model parameters is estimated conditionally on the calibrated values of the remaining ones. Without Assumption 3, these identification conditions are necessary, but not sufficient in our case, i.e., identification of the parameters of the underlying continuous-time model based on discrete-time observations. The uniqueness of the solution may be compromised, due to aliasing problems, and hence our focus on this issue.

3. Recovering structural shocks

Recall that $\boldsymbol{\eta}_\tau$ in (2.8) for the S-SSR model, or in (2.12) for the F-SSR model, does not represent structural shocks hitting the economy at observation τ. Instead, it represents a vector of reduced-form innovations, where each of its elements is potentially a composite of the true underlying structural shocks $\mathbf{w}(t)$ occurring continuously within the time interval $(t_{\tau-1}, t_\tau]$. While predicting the latter is not possible from a discrete set of observations sampled at regular intervals of length h, we introduce a simple approach that uses the reduced-form disturbances, $\{\boldsymbol{\eta}_\tau\}_{\tau \in \mathbb{Z}^+}$, to recover a sequence of structural shocks at the same discrete points in time, $\{\mathbf{u}_\tau\}_{\tau \in \mathbb{Z}^+}$, where

$$\mathbf{u}_\tau := h^{-1/2}(\mathbf{w}(t_\tau) - \mathbf{w}(t_{\tau-1})) \tag{3.1}$$

is an n_w-dimensional vector of Gaussian random variables, with $\mathbb{E}[\mathbf{u}_\tau] = \mathbf{0}$, and $\mathbb{E}\left[\mathbf{u}_\tau \mathbf{u}_\tau^\top\right] = \mathbf{I}$.

Conceptually, the proposed strategy considers partitioning the time interval $[t_{\tau-1}, t_\tau]$ into $n \geq 1$ sub-intervals, $t_{\tau-1} = t_0^\tau < t_1^\tau < \cdots < t_{n-1}^\tau < t_n^\tau = t_\tau$, each of length $h_n := h/n$. This partition, which is depicted in Figure 2, has two important implications. First, for any given partition, integrated structural shocks are given by the Riemann sum, $\sum_{i=1}^n \Delta \mathbf{w}(t_i^\tau) = \mathbf{w}(t_\tau) - \mathbf{w}(t_{\tau-1}) = \int_{t_{\tau-1}}^{t_\tau} \mathrm{d}\mathbf{w}(s)$. Second, for given $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, the relation between the reduced-

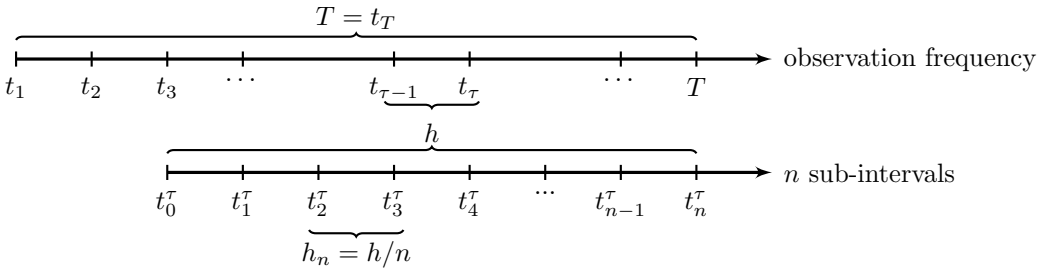


Figure 2. Partition of sampling interval. The figure illustrates the observation points in the sample, as well as the assumed subsampling scheme within each observation interval.

form innovations and the structural shocks can be written as

$$\boldsymbol{\eta}_\tau = \int_{t_{\tau-1}}^{t_\tau} \mathbf{H}(\boldsymbol{\theta}; t_\tau - s)\mathrm{d}\mathbf{w}(s) = \lim_{n\to\infty} \sum_{i=1}^{n} \mathbf{H}(\boldsymbol{\theta}; t_\tau - t_{i-1}^\tau)\Delta\mathbf{w}(t_i^\tau)\,, \tag{3.2}$$

where $\mathbf{H}(\boldsymbol{\theta}; \cdot)$ is a deterministic and square integrable function over the time interval $[t_{\tau-1}, t_\tau]$. For the S-SSR model, this function is given by the $n_x \times n_w$ matrix

$$\mathbf{H}(\boldsymbol{\theta}; t_\tau - s) = \exp\left(\mathbf{A}(\boldsymbol{\theta})(t_\tau - s)\right)\mathbf{B}(\boldsymbol{\theta})\,, \tag{3.3}$$

while for the the F-SSR model, it is given by the $(n_x + n_y) \times n_w$ matrix

$$\mathbf{H}(\boldsymbol{\theta}; t_\tau - s) = \begin{bmatrix} \exp\left(\mathbf{A}(\boldsymbol{\theta})(t_\tau - s)\right)\mathbf{B}(\boldsymbol{\theta}) \\ \mathbf{C}(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})^{-1}\left(\exp\left(\mathbf{A}(\boldsymbol{\theta})(t_\tau - s)\right) - \mathbf{I}\right)\mathbf{B}(\boldsymbol{\theta}) \end{bmatrix}. \tag{3.4}$$

The relation in (3.2) forms the basis of our framework for recovering the structural shocks at measurement times from a sample of discrete observations, building on the following proposition.

Proposition 5. Given the definition of structural shocks at measurement times in (3.1), it follows that, for any given partition $t_{\tau-1} = t_0^\tau < t_1^\tau < \cdots < t_{n-1}^\tau < t_n^\tau = t_\tau$ of the time interval $[t_{\tau-1}, t_\tau]$, the mapping in (3.2) can be written as

$$\boldsymbol{\eta}_\tau = h^{1/2}\mathbf{H}(\boldsymbol{\theta}; h)\mathbf{u}_\tau + \mathcal{R}_\tau\,, \tag{3.5}$$

where h is the fixed length of the time interval. Moreover, the remainder term, $\mathcal{R}_\tau$, which captures the effects of primitive shocks occurring between measurement times, is stochasti-

cally bounded by the sampling frequency, such that

$$\mathcal{R}_\tau = \mathcal{O}_P(h^{3/2}), \text{ as } h \to 0. \tag{3.6}$$

Proof. See Appendix D.5 ■

Remark 3.1. Using the relation between h, h_n, and n, Appendix D.5 provides alternative characterizations of the remainder in Proposition 5, in terms of the number of sub-intervals for a given partition, n, or the length of each of these sub-intervals, h_n.

The remainder $\mathcal{R}_\tau$ in (3.6) provides a measure of the *approximation error* incurred when using the approximate linear relation

$$\boldsymbol{\eta}_\tau \approx h^{1/2}\mathbf{H}(\boldsymbol{\theta}; h)\mathbf{u}_\tau \tag{3.7}$$

to back out the structural shocks at measurement times from the reduced-form innovations. This relation, which is equivalent to setting $n = 1$ in (3.2), so that $h_n = h$, indicates that the vector of structural shocks is approximately spanned by the vector of reduced-form innovations.[11] The following two examples illustrate the relation between the approximated structural shocks, $\mathbf{u}_\tau$, and the reduced-form residuals, $\boldsymbol{\eta}_\tau$, at measurement times.

Example 1 (State vector in S-SSR model)**.** Consider a bivariate state vector $[x_1(t), x_2(t)]^\top$, with dynamics governed by

$$\begin{bmatrix} \mathrm{d}x_1(t) \\ \mathrm{d}x_2(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} b_1 & 0 \\ 0 & b_2 \end{bmatrix} \begin{bmatrix} \mathrm{d}w_1(t) \\ \mathrm{d}w_2(t) \end{bmatrix},$$

where the coefficients $(a_{11}, a_{12}, a_{22}, b_1, b_2)$ may be nonlinear functions of an underlying set of parameters $\boldsymbol{\theta}$. From (2.5) and (3.7), it follows that the innovations of the corresponding EDM at measurement time τ are given approximately as

$$\begin{bmatrix} \eta^s_{1,\tau} \\ \eta^s_{2,\tau} \end{bmatrix} \approx h^{1/2} \begin{bmatrix} u_{1,\tau} b_1 \frac{(\exp(a_{11}h)-1)}{a_{11}} - u_{2,\tau} a_{12} b_2 \frac{a_{11}(\exp(a_{22}h)-1) - a_{22}(\exp(a_{11}h)-1)}{a_{11}a_{22}(a_{11}-a_{22})} \\ u_{2,\tau} b_2 \frac{(\exp(a_{22}h)-1)}{a_{22}} \end{bmatrix}.$$

Thus, the disturbances $\boldsymbol{\eta}_\tau = \left[\eta^s_{1,\tau}, \eta^s_{2,\tau}\right]^\top$ are combinations of the true underlying structural shocks $\mathbf{u}_\tau = [u_{1,\tau}, u_{2,\tau}]^\top$. The realization of $u_{2,\tau}$ leads to movements in $\eta^s_{1,\tau}$ and $\eta^s_{2,\tau}$, creating contemporaneous movements in both state variables. ■

[11] Since $\exp(\mathbf{A}(\boldsymbol{\theta})h)$ is linear in h for small h, we use instead the midpoint $\exp(\mathbf{A}(\boldsymbol{\theta})h/2)$ when computing $\mathbf{H}(\boldsymbol{\theta}; h)$ and building the approximation in (3.7) below (see Nowman, 1993).

Example 2 (State vector in F-SSR model)**.** Let the state variables be as described in Example 1. If the measurements correspond to flow variables, such that

$$y_\tau^f = \int_{t_{\tau-1}}^{t_\tau} y(s)\mathrm{d}s = \int_{t_{\tau-1}}^{t_\tau} \begin{bmatrix} c_1 & c_2 \end{bmatrix} \begin{bmatrix} x_1(s) \\ x_2(s) \end{bmatrix} \mathrm{d}s\,,$$

then the transition equation of the corresponding state space system must be extended as in (2.12), so that the vector of reduced-form residuals is now given by $\boldsymbol{\eta}_\tau = \left[\eta_{1,\tau}^s, \eta_{2,\tau}^s, \eta_{3,\tau}^f\right]^\top$, with $\eta_{1,\tau}^s$ and $\eta_{2,\tau}^s$ as in Example 1, and

$$\begin{aligned}\eta_{3,\tau}^f \approx -h^{1/2}c_1 \Bigg\{ & u_{1,\tau} b_1 \frac{a_{11}h - (\exp(a_{11}h) - 1)}{a_{11}^2} \\ & + u_{2,\tau} \frac{a_{12}b_2[a_{11}^2(\exp(a_{22}h) - 1) - a_{22}^2(\exp(a_{11}h) - 1)]}{a_{11}^2 a_{22}^2 (a_{11} - a_{22})} \\ & + u_{2,\tau} \frac{a_{12}b_2(a_{11} + a_{22})h}{a_{11}a_{22}(a_{11} - a_{22})} \Bigg\} - h^{1/2}c_2 u_{2,\tau} \frac{b_2(a_{22}h - (\exp(a_{22}h) - 1))}{a_{22}^2}\,.\end{aligned}$$

■

Remark 3.2. The dimension of $\boldsymbol{\eta}_\tau$ in Examples 1 and 2 is equal to that of the state vector, even if $b_1 = 0$, i.e., when the endogenous state variables are not subject to idiosyncratic structural shocks. Hence, the number of reduced-form residuals does not need to match the number of structural shocks, when using the EDM to estimate continuous-time models. This is particularly relevant when assessing the validity of Assumption 2, and thus whether the empirical model is subject to stochastic singularity. ■

Since the econometrician rarely has control over h_n or n, the next proposition shows that the error committed by using the approximation in (3.7) is bounded in probability. To ease notation, write $\mathbf{H}(\boldsymbol{\theta}) := h^{1/2}\mathbf{H}(\boldsymbol{\theta}; h)$. The structural shocks at measurement times can then be recovered from the reduced-form innovations as the solution to (3.7), i.e.,

$$\tilde{\mathbf{u}}_\tau = \mathbf{H}(\boldsymbol{\theta})^\dagger \boldsymbol{\eta}_\tau\,, \tag{3.8}$$

where superscript † denotes a generalized inverse. The approximation implied by (3.8) resembles the identification mechanisms commonly used in the structural VAR literature. Theoretical restrictions are imposed, here through $\mathbf{H}(\boldsymbol{\theta})$, to uncover the structural shocks that are otherwise hidden in the correlated reduced-form residuals $\boldsymbol{\eta}_\tau$.

Proposition 6. For all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, and in the absence of additional information between measurement times, the error in the approximation of the structural shocks is stochastically

bounded, i.e.,

$$\tilde{\mathbf{u}}_\tau - \mathbf{u}_\tau = \mathcal{O}_P(1)\,.$$

Proof. See Appendix D.6. ■

Example 3 (Recovering structural shocks)**.** Consider again the model in Example 1. According to (3.7), the relation between the reduced-form innovations and the structural shocks is characterized by the approximate system of equations

$$\begin{bmatrix} \eta^s_{1,\tau} \\ \eta^s_{2,\tau} \end{bmatrix} \approx h^{1/2} \begin{bmatrix} \exp(a_{11}h) & \left(\frac{a_{12}\exp(a_{11}h) - a_{12}\exp(a_{22}h)}{a_{11}-a_{22}}\right) \\ 0 & \exp(a_{22}h) \end{bmatrix} \begin{bmatrix} b_1 & 0 \\ 0 & b_2 \end{bmatrix} \begin{bmatrix} u_{1,\tau} \\ u_{2,\tau} \end{bmatrix}.$$

Solving the system allows us to recover the structural shocks as

$$\begin{aligned} \tilde{u}_{2,\tau} &= h^{-1/2}\exp(-a_{22}h)b_2^{-1}\eta_{2,\tau}\,, \\ \tilde{u}_{1,\tau} &= h^{-1/2}\exp(-a_{11}h)b_1^{-1}\eta_{1,\tau} - h^{-1/2}\frac{a_{12}\exp(-(a_{11}+a_{22})h)(\exp(a_{11}h)-\exp(a_{22}h))}{b_1(a_{11}-a_{22})}\eta_{2,\tau}\,. \end{aligned}$$

Hence, while $\tilde{u}_{2,\tau}$ is proportional to the estimated residual $\eta_{2,\tau}$, $\tilde{u}_{1,\tau}$ is given by a linear combination of $\eta_{1,\tau}$ and $\eta_{2,\tau}$. ■

4. Monte Carlo evidence

4.1 The artificial economy

To illustrate and assess the procedures introduced in Sections 2 and 3, we use a continuous-time version of the RBC model with indivisible labor of Hansen (1985), with shocks to total factor productivity and capital accumulation. In the following, we present some of the main elements of the model. A complete derivation is relegated to Appendix E.

Preferences. Consider an economy where time evolves continuously, $t \in \mathbb{R}^+$. A representative agent maximizes her expected discounted lifetime utility from consumption $C(t)$ and is leisure $L(t)$,

$$\mathbb{E}_0\left[\int_0^\infty e^{-\rho t}\left(\ln C(t) + \psi L(t)\right)\mathrm{d}t\right], \tag{4.1}$$

where $\rho > 0$ is the subjective discount rate, ψ the weight of leisure in the instantaneous utility function, and $\mathbb{E}_0[\cdot]$ the expectation operator conditional on the information available at time $t = 0$. There is no population growth, and both the population size and the endowment of available time are normalized to one. Hence, the fraction of hours worked per unit of time,

$N(t)$, is

$$N(t) = 1 - L(t)\,. \tag{4.2}$$

The agent's income consists of wages and rents, received from selling labor and renting capital to firms, and it is allocated between consumption and investment,

$$C(t) + I(t) = W(t)N(t) + r(t)K(t)\,, \tag{4.3}$$

where $I(t)$ is investment, $W(t)$ the real wage, $r(t)$ the real interest rate, and $K(t)$ the capital stock. The latter increases whenever gross investment, $I(t)$, exceeds any capital depreciation,

$$\mathrm{d}K(t) = (I(t) - \delta K(t))\,\mathrm{d}t + \sigma_k K(t)\mathrm{d}w_k(t), \quad K\,(0) = K_0\,, \tag{4.4}$$

where $\delta \geq 0$ is the depreciation rate, and $w_k(t)$ a standard Brownian motion, representing shocks to the marginal efficiency of investment and/or the future productivity of the capital stock (cf. Furlanetto and Seneca, 2014 and Brunnermeier and Sannikov, 2014). The diffusion parameter $\sigma_k > 0$ regulates the variance of these shocks.

Technology. The one good in this economy is produced by a large number of perfectly competitive firms. The representative firm rents labor and capital from the representative agent, and combines them according to

$$Y(t) = \exp\left(Z(t)\right) K(t)^{\alpha} \left(\exp\left(\eta t\right) N(t)\right)^{1-\alpha}\,, \quad \alpha \in (0,1)\,, \tag{4.5}$$

where $Y(t)$ is aggregate output, $\eta > 1$ the constant growth rate of labor-augmenting technological progress, and $Z(t)$ a zero-mean measure of total factor productivity (TFP). The latter is assumed to evolve according to the Ornstein-Uhlenbeck process

$$\mathrm{d}Z(t) = -\rho_z Z(t)\mathrm{d}t + \sigma_z \mathrm{d}w_z(t)\,, \quad Z\,(0) = Z_0\,, \tag{4.6}$$

with mean-reversion parameter $\rho_z > 0$, diffusion parameter $\sigma_z > 0$, and $w_z(t)$ a standard Brownian motion, independent of $w_k(t)$.

Equilibrium. Both welfare theorems hold in this economy. Hence, it is possible to solve the social planner's problem directly. For $K\,(0) = K_0$ and $Z\,(0) = Z_0$ given, the planner chooses paths for consumption and the fraction of hours worked that maximize the expected lifetime utility (4.1), subject to the law of motion for the capital stock (4.4), the production function (4.5), and the evolution of TFP (4.6). In addition, the aggregate resource constraint

$$Y(t) = C(t) + I(t) \tag{4.7}$$

must hold, at all points in time. The resulting system of nonlinear equations determines the equilibrium paths of $C(t), K(t), N(t), I(t)$, and $Y(t)$.

Transformed (stationary) equilibrium. A solution to the planner's problem is not available in closed form. Therefore, we approximate its behavior around the economy's deterministic steady state. However, since the model exhibits balanced growth, we de-trend all non-stationary variables before computing the approximation. For this purpose, we define $y(t) := Y(t)/\exp(\eta t)$, $c(t) := C(t)/\exp(\eta t)$, $i(t) := I(t)/\exp(\eta t)$, and $k(t) := K(t)/\exp(\eta t)$ to be the de-trended versions of the model's variables. For notational consistency we define $n(t) := N(t)$ and $z(t) := Z(t)$, although these variables are stationary by construction.

The problem faced by the social planner can be summarized as the stochastic optimal control problem

$$J(k_0, z_0) = \max_{\{c(t), n(t)\}_{t=0}^{\infty}} \mathbb{E}_0 \left[\int_0^{\infty} e^{-\rho t} \left(\ln c(t) + \psi \left(1 - n(t) \right) \right) \mathrm{d}t \right] \quad \text{s.t.}$$

$$\begin{aligned} \mathrm{d}k(t) &= \left(\exp\left(z(t)\right) k(t)^{\alpha} n(t)^{1-\alpha} - c(t) - (\delta + \eta) k(t) \right) \mathrm{d}t + \sigma_k k(t) \mathrm{d}w_k(t) \,, \quad k(0) = k_0 \,, \\ \mathrm{d}z(t) &= -\rho_z z(t) \mathrm{d}t + \sigma_z \mathrm{d}w_z(t) \,, \quad z(0) = z_0 \,, \end{aligned}$$

in which $c(t) \in \mathbb{R}^+$ and $n(t) \in [0, 1]$ denote the control variables at instant $t > 0$, $k(t) \in \mathbb{R}^+$ and $z(t) \in \mathbb{R}$ the state variables at instant $t > 0$, and $J(k_0, z_0)$ the value of the optimal program (value function), given the initial conditions $k(0)$ and $z(0)$.

By an application of Bellman's principle of optimality, it is shown in Appendix E that the sequence of equilibrium allocations in this economy, $\{c(t), n(t), k(t)\}_{t \geq 0}$, satisfy the nonlinear system

$$0 = \psi c(t) n(t) - (1 - \alpha) \exp(z(t)) k(t)^{\alpha} n(t)^{1-\alpha} \tag{4.8}$$

$$\begin{aligned} \mathbb{E}_t \left[\mathrm{d}c(t) \right] = \Bigg[& \left(\alpha \exp\left(z(t)\right) k(t)^{\alpha - 1} n(t)^{1-\alpha} - \rho - \delta - \eta \right) - \sigma_k^2 \frac{k(t) \mathbf{c}_k \left(k(t), z(t) \right)}{c(t)} \\ & + \frac{1}{2} \left(\sigma_k^2 \left(\frac{k(t) \mathbf{c}_k \left(k(t), z(t) \right)}{c(t)} \right)^2 + \sigma_z^2 \left(\frac{\mathbf{c}_z \left(k(t), z(t) \right)}{c(t)} \right)^2 \right) \Bigg] c(t) \mathrm{d}t \end{aligned} \tag{4.9}$$

$$\mathrm{d}k(t) = \left(\exp\left(z(t)\right) k(t)^{\alpha} n(t)^{1-\alpha} - c(t) - (\delta + \eta) k(t) \right) \mathrm{d}t + \sigma_k k(t) \mathrm{d}w_k(t) \,, \tag{4.10}$$

where $z(t)$ is governed by (4.6). Equation (4.8) is the intratemporal labor supply equation, and (4.9) the Euler equation for consumption. Here, $\mathbf{c}_k \left(k(t), z(t) \right)$ and $\mathbf{c}_z \left(k(t), z(t) \right)$ denote the marginal responses of optimal consumption to changes in the capital stock and TFP.

Approximate solution. We approximate the solution of the model by first linearizing the

equilibrium conditions in (4.8)-(4.10) and the dynamics in (4.6) around the model's deterministic steady state. Next, we compute the rational expectations solution to the resulting linear system, using the approach in Sims (2002).[12]

Let circumflex on a variable denote its log-deviation from its steady state value. The log-linearized equilibrium can then be compactly written as

$$\mathbb{E}_t \begin{bmatrix} \mathrm{d}\hat{c}(t) \\ \mathrm{d}\hat{k}(t) \\ \mathrm{d}\hat{z}(t) \end{bmatrix} = \underbrace{\begin{bmatrix} \xi_{cn}\xi_{nc} & 0 & \xi_{cz}+\xi_{cn}\xi_{nz} \\ \xi_{kc}+\xi_{kn}\xi_{nc} & \xi_{kk}+\xi_{kn}\xi_{nk} & \xi_{kz}+\xi_{kn}\xi_{nz} \\ 0 & 0 & -\rho_z \end{bmatrix}}_{:=\boldsymbol{\Gamma}} \begin{bmatrix} \hat{c}(t) \\ \hat{k}(t) \\ \hat{z}(t) \end{bmatrix} \mathrm{d}t\,, \tag{4.11}$$

where we have reduced the system of equilibrium conditions from four to three variables by substituting out the linearized version of condition (4.8), $\hat{n} = \xi_{nc}\hat{c} + \xi_{nk}\hat{k} + \xi_{nz}\hat{z}$. The elements of the matrix $\boldsymbol{\Gamma}$ depend on the steady state of the model, and their values can be found in Appendix E.

Next, the rational expectations solution to the system (4.11) is built by applying the QZ (generalized Schur) decomposition to the matrix $\boldsymbol{\Gamma}$, to identify the number of stable and unstable roots of the system. If the Blanchard and Kahn (1980) conditions are satisfied, i.e., the number of stable roots equals the number of state variables, then the rational expectations solution has the continuous state space representation given in (2.1)-(2.2), with $\mathbf{x}(t) = (\hat{k}(t), \hat{z}(t))^\top$ the 2×1 vector of state variables, $\mathbf{y}(t) = (\hat{c}(t), \hat{n}(t))^\top$ the 2×1 vector of control variables, and $\mathbf{w}(t) = (w_k(t), w_z(t))^\top$ the 2×1 vector of structural shocks. The matrices $\mathbf{A}(\boldsymbol{\theta})$, $\mathbf{B}(\boldsymbol{\theta})$, and $\mathbf{C}(\boldsymbol{\theta})$ are given by

$$\mathbf{A}(\boldsymbol{\theta}) = \begin{bmatrix} \phi_{kk} & \phi_{kz} \\ 0 & -\rho_z \end{bmatrix}, \quad \mathbf{B}(\boldsymbol{\theta}) = \begin{bmatrix} \sigma_k & 0 \\ 0 & \sigma_z \end{bmatrix}, \quad \mathbf{C}(\boldsymbol{\theta}) = \begin{bmatrix} \phi_{ck} & \phi_{cz} \\ \phi_{nk} & \phi_{nz} \end{bmatrix}. \tag{4.12}$$

As shown in Appendix E, the reduced-form parameters ϕ_{ck}, ϕ_{cz}, ϕ_{nk}, ϕ_{nz}, ϕ_{yk}, ϕ_{yz}, ϕ_{kk}, and ϕ_{kz} are combinations of the eigenvalues and eigenvectors of the matrix $\boldsymbol{\Gamma}$, which depend nonlinearly on the structural parameters of the model. The latter are collected in the vector $\boldsymbol{\theta} = (\rho, \psi, \alpha, \delta, \eta, \rho_z, \sigma_z, \sigma_k)^\top$.

Lemma 2. If $(1-\alpha)(\delta+\eta+\rho) \neq \alpha\rho_z$, with $\rho_z > 0$, and $\alpha \in (0,1)$, then the matrix $\mathbf{A}(\boldsymbol{\theta})$ is not subject to aliases, and the entries of $(\mathbf{A}(\boldsymbol{\theta}), \boldsymbol{\Sigma}(\boldsymbol{\theta}))$ are identified from $(\mathbf{A}_h(\boldsymbol{\theta}), \boldsymbol{\Sigma}_{\boldsymbol{\eta},h}(\boldsymbol{\theta}))$.

Proof. See Appendix E.6. ■

[12] We use the `gensysct` routine, available at http://sims.princeton.edu/yftp/gensys/. Alternative approaches include the continuous-time version of the Blanchard and Kahn method described in Buiter (1984), the continuous-time version of the Anderson-Moore AIM algorithm in Anderson (1997), or the perturbation method in Parra-Alvarez et al. (2021).

Remark 4.1. The entries of the matrices $\mathbf{A}(\boldsymbol{\theta})$, $\mathbf{B}(\boldsymbol{\theta})$, and $\mathbf{C}(\boldsymbol{\theta})$, and thus the likelihood function, do not depend on the parameter ψ, which is therefore unidentified. It vanishes from the rational expectations solution of the model (see Canova and Sala, 2009). ■

4.2 Finite sample properties of the MLE

In this section, we investigate the finite sample properties of the MLE by running extensive Monte Carlo experiments. In particular, we study the ability of the method to estimate the parameters $\boldsymbol{\theta}$ of the underlying continuous-time model from Section 4.1, using measurements that are only available at a discrete frequency. We generate $M = 500$ samples for consumption, C, labor, N, and output, Y, from the linearized version of the RBC model, the data generating process (DGP), using the parameter values in Table 1. These values are standard in the literature, in that they match long-run values of U.S. macroeconomic aggregates observed over the postwar period.

The model is calibrated to an annual frequency, and all parameter values should be interpreted accordingly. We set $\alpha = 0.30$ to match an average labor income to GDP ratio of 70%. The values of the subjective discount rate and the depreciation rate are set to $\rho = 0.03$ and $\delta = 0.06$. These are consistent with steady-state values for the net return on capital and the investment to GDP ratio of 4% and 20%, respectively. The weight of leisure in the instantaneous utility function is fixed at $\psi = 2.686$, so that, in the steady state, agents spend 1/3 of their time working. The long-run growth of the economy is assumed to be $\eta = 2\%$. The parameters describing the dynamics of the total factor productivity are set to $\rho_z = 0.2052$ and $\sigma_z = 0.0140$, in line with standard estimates of the quarterly Solow residual for the U.S. economy (see e.g., Hansen, 1985, Hansen and Prescott, 1995, and Hansen, 1997). Finally, we fix the volatility of the capital stock at $\sigma_k = 0.0104$, based on the calibration in Ambler and Paquet (1994).

Each Monte Carlo sample contains 240 quarterly observations, corresponding to 60 years of data. The simulated observations do not include any measurement errors. While allowing for some form of measurement error may be desirable in empirical applications, we leave it out of the Monte Carlo experiments. Using the simulated samples, we estimate the model parameters using the approach described in Section 2, under the assumption that the model is correctly specified. To avoid stochastic singularity, we include only two observables in the ML estimation, namely, consumption and the fraction of hours worked, so $\mathbf{y}_\tau = [C_\tau, N_\tau]^\top$.

We compare the ML estimates from the S-SSR and F-SSR models to those obtained instead from the misspecified state space representation that results from a *naive* Euler-Maruyama (EM) discretization of the continuous-time state-transition equation (2.1). The

Table 1. Parameter values. The parameters of the RBC model are calibrated to an annual frequency, and their values should be interpreted accordingly.

Parameter	Value	Source / Target
Subjective discount rate, ρ	0.0300	Long-run (net) return on capital of 4%
Leisure weight, ψ,	2.6860	Average fraction of hours worked of 1/3
Capital share in output, α	0.3000	Average WN/Y of 0.7
Depreciation rate, δ	0.0600	Average I/Y of 0.2 and K/Y of 2.5
Labor-augmenting growth, η	0.0200	Average GDP growth
Mean-reversion of TFP, ρ_z	0.2052	Persistence of Solow residual
Volatility of TFP, σ_z	0.0140	Volatility of Solow residual
Volatility of depreciation, σ_k	0.0104	Ambler and Paquet (1994)

transition equation under this approximation is given by

$$\mathbf{x}_\tau = (\mathbf{I} + \mathbf{A}(\boldsymbol{\theta})h)\,\mathbf{x}_{\tau-1} + \sqrt{h}\mathbf{B}(\boldsymbol{\theta})\mathbf{u}_\tau\,, \tag{4.13}$$

where $\mathbf{u}_\tau$ is defined in (3.1). The state space model that results from combining (4.13) with the measurement equation in (2.9) is referred to below as the EM-SSR model. The EM-SSR model differs from the S-SSR and F-SSR models in that (i) the transition matrix is truncated to first order, and therefore ignores all the terms of order smaller than h in (2.4), (ii) the disturbances ignore the temporal aggregation of structural shocks through the system between the discrete observations, and (iii) the likelihood function is no longer exact, since the dynamics between observation points is ignored.

Similarly to Del Negro and Schorfheide (2008), we group the model parameters in two categories, $\boldsymbol{\theta}_{\text{ss}} = [\psi, \alpha, \delta, \rho, \eta]^\top$ and $\boldsymbol{\theta}_{\text{exo}} = [\rho_z, \sigma_z, \sigma_k]^\top$. The first includes parameters that can be readily identified from the model's steady state and the available measurements. The second consists of the parameters that characterize the dynamics of the exogenous processes driving the economy. For the Monte Carlo experiments, we focus on estimation of the latter, $\boldsymbol{\theta}_{\text{exo}}$, while the parameters in $\boldsymbol{\theta}_{\text{ss}}$ are held fixed at their population values from Table 1.

Remark 4.2. As indicated in Lemma 2, the model does not suffer from aliasing identification problems. Using the accompanying identification toolbox (see footnote 9), we numerically verify that the necessary and sufficient rank conditions from Komunjer and Ng (2011, Proposition 3) are satisfied, and conclude that, conditional on the calibrated values $\boldsymbol{\theta}_{\text{ss},0}$, the parameter vector $\boldsymbol{\theta}_{\text{exo}}$ is locally identified around $\boldsymbol{\theta}_{\text{exo},0}$. ■

Table 2. Finite sample properties. The table reports statistics for finite sample estimates of $\boldsymbol{\theta}_{\text{exo}}$ from $M = 500$ samples of quarterly ($h = 1/4$) consumption (C) and hours worked (N), generated over a period of 60 years ($T = 240$ observations in each sample). Simulated measurements in Panel A are sampled as stocks, and those in Panel B as flows. The parameters in $\boldsymbol{\theta}_{\text{ss}}$ are fixed at their population values in Table 1. With $\hat{\boldsymbol{\theta}}_{\text{exo},m}$ denoting the estimates from the m^{th} sample, the table displays the Bias $= M^{-1}\sum_{m=1}^{M}(\hat{\boldsymbol{\theta}}_{\text{exo},m} - \boldsymbol{\theta}_{\text{exo},0})$ and root mean squared error (RMSE $= (M^{-1}\sum_{m=1}^{M}(\hat{\boldsymbol{\theta}}_{\text{exo},m} - \boldsymbol{\theta}_{\text{exo},0})^2)^{1/2}$) across repetitions for the F-SSR, S-SSR, and EM-SSR models.

Panel A: Data is sampled as stocks

		F-SSR		S-SSR		EM-SSR	
$\boldsymbol{\theta}_{\text{exo}}$		Bias	RMSE	Bias	RMSE	Bias	RMSE
ρ_z	0.2052	-	-	0.0024	0.0188	-0.0291	0.0328
σ_z	0.014	-	-	-0.0001	0.0006	-0.0001	0.0007
σ_k	0.0104	-	-	-4.49e-5	0.0004	-0.0003	0.0006

Panel B: Data is sampled as flows

		F-SSR		S-SSR		EM-SSR	
$\boldsymbol{\theta}_{\text{exo}}$		Bias	RMSE	Bias	RMSE	Bias	RMSE
ρ_z	0.2052	0.0030	0.0194	-0.0023	0.0190	-0.0317	0.0353
σ_z	0.014	-0.0001	0.0007	-0.0026	0.0027	-0.0026	0.0027
σ_k	0.0104	-0.0001	0.0005	-0.0020	0.0020	-0.0022	0.0022

4.2.1 Stock data

Our first Monte Carlo experiment assumes that all observables are sampled at discrete points in time as stocks. Hence, we use the S-SSR model from (2.8)-(2.9) for the ML estimation of the parameters, with $\mathbf{A}(\boldsymbol{\theta})$, $\mathbf{B}(\boldsymbol{\theta})$, and $\mathbf{C}(\boldsymbol{\theta})$ given in (4.12) and Appendix E. The results are summarized in Panel A of Table 2. The table reports the bias of $\hat{\boldsymbol{\theta}}_{\text{exo}}$ across repetitions, and the root mean squared error (RMSE) of the estimates.

The Monte Carlo experiment reveals some important features that should be addressed. First, the MLE delivers small sample biases that are within a reasonable range when using the S-SSR model, and thus exact ML estimation. In particular, ρ_z is above its value in the population, by 1.17% on average. This positive bias in the estimation of the speed of mean reversion is consistent with the observations made by Merton (1980), and later verified in Tang and Chen (2009) and Yu (2012) for the univariate OU case. On the other hand, σ_z exhibits a downward bias, of around -0.7%, on average, whereas σ_k exhibits virtually no bias. Since there is no discretization error in the derivation of the state space model used to compute the likelihood function, these values reflect pure estimation bias. In contrast, the

EM-SSR model produces biases in ρ_z and σ_k that are about one order of magnitude larger. More precisely, based on the EM-SSR model, the resulting ML estimates of ρ_z and σ_k are, on average, around 14% and 3% below their true values, respectively. Interestingly, the bias in σ_z remains the same.

Previous work on the estimation of observed univariate and multivariate diffusions from discrete measurements (see, e.g., Phillips and Yu, 2005, Tang and Chen, 2009, Wang et al., 2011) has concluded that using the EM approximation inevitably induces a *discretization bias*, on top of the pure *estimation bias*. As shown by Lo (1988) for the univariate case, the discretization bias in turn implies misspecification of the likelihood function, leading to inconsistent estimates of the parameters of the continuous-time model (see also, Phillips, 1973 and Thornton and Chambers, 2016). Further, it has been shown that the discretization bias affects the estimation of the speed of mean reversion, ρ_z, more severely than the estimation of the diffusion parameter, σ_z. Naturally, the discretization bias could be reduced by sampling at a higher frequency, i.e., let $h \to 0$, if feasible. Moreover, Wang et al. (2011) show that in systems of linear SDEs, the discretization and estimation bias are of opposite sign, with the latter being larger in magnitude in financial applications, where data can be recorded at a higher frequency (see also Phillips and Yu, 2005), something that is typically not feasible in macroeconomic applications. Finally, the same authors also show that the pure estimation bias in persistence parameters is a function of the time span (Th) of the sample, i.e., the length of the time period over which the discrete observations are recorded, and, thus, cannot be reduced in large samples that result exclusively from sampling at higher frequency.

In light of the preceding evidence, we conjecture that the modest performance of the EM-SSR model, relative to the S-SSR model, can be explained by the additional misspecification bias induced by the EM approximation of the continuous-time transition. This is evident for ρ_z, where the negative discretization bias dominates the positive estimation bias when using macroeconomic data sampled at low frequency. Our results are consistent with the notion that the positive estimation bias in the rate of mean reversion (corresponding to downward bias in estimated persistence, or autocorrelation) and the opposite sign of the discretization bias carries over from the univariate and purely observed multivariate models from the literature to the state space models considered here. On the other hand, there is no evidence that the EM approximation induces any discretization bias in σ_z, so the bias stems exclusively from the ML estimation. Finally, σ_k also exhibits a negative discretization bias, which adds to the small downward estimation bias. Figure F.1 in Appendix F illustrates the effects of the discretization bias by plotting the log-likelihood profile for each of the three parameters, using one of the simulated samples. The figure reveals that using the EM-SSR model results in a likelihood function with a maximum that is located to the left of the true

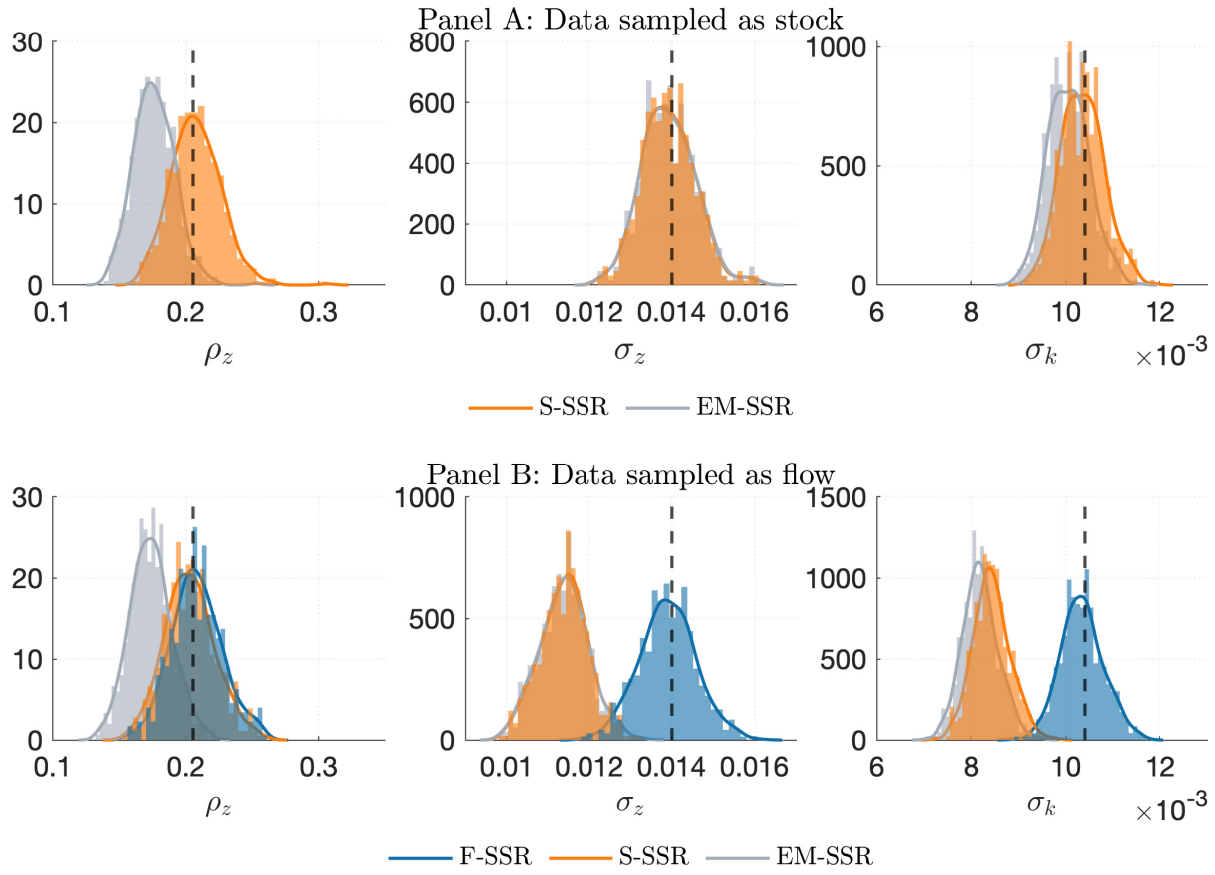


Figure 3. Finite sample distribution of parameter estimates. The graph plots the distribution of estimated parameters, $\hat{\boldsymbol{\theta}}_{\text{exo}}$, across $M = 500$ random samples of consumption (C) and hours worked (N) using different state space representations. Each sample includes $T = 240$ quarterly observations ($h = 1/4$) generated from the true data generation process. The observables in Panel A are sampled as stocks, and those in Panel B as flows. Each plot reports the density of parameter estimates obtained under the state space representation for stock variables (S-SSR), flow variables (F-SSR), or that obtained under the naive Euler-Maruyama discretization of the transition equation (EM-SSR).

parameter value. This downward bias is more pronounced for the speed of mean reversion of TFP, although it also affects the diffusion coefficients.

Further, the EM-SSR model produces RMSEs that are larger than those from the S-SSR model. This is most evident for ρ_z, where the dispersion of the estimates around the true value almost doubles, compared to the S-SSR case. However, the absolute variability of the estimates, as measured by the standard deviation of point estimates across the Monte Carlo samples, is higher in the S-SSR case (0.0187) than in the EM-SSR case (0.0153). This result is in line with those reported in Wang et al. (2011) for the case of univariate diffusions. With respect to σ_z and σ_k, the standard deviations are very low, and although always smaller in

the EM-SSR model, the difference across model specifications is negligible. These results are confirmed in Panel A of Figure 3, where we plot the finite sample distributions of the parameter estimates for both the exact and the approximate ML estimation. From the figure, the models seem to compete against a bias-variance trade-off. For ρ_z, the S-SSR model exhibits little bias, but a larger standard deviation, compared to the EM-SSR model.

Our conclusions are robust to the set of observables used in the estimation. In the absence of measurement errors, the information content in the fraction of hours worked and output, $\mathbf{y}_\tau = [N_\tau, Y_\tau]^\top$, or consumption and output, $\mathbf{y}_\tau = [C_\tau, Y_\tau]^\top$, is the same as that in consumption and the fraction of hours worked, $\mathbf{y}_\tau = [C_\tau, N_\tau]^\top$. Hence, all sets consisting of two variables are equally informative about the set of parameters under consideration. The Monte Carlo results show that the point estimates of the parameters are not affected by the set of observables used in the estimation, so the bias, standard deviation, and RMSE under both model specifications are identical to those reported in Table 2, and available upon request. This contrasts with the results in Ruge-Murcia (2007), where the accuracy with which the subjective discount rate and the persistence of TFP can be estimated depends on the observables used. Moreover, our findings are also robust to estimation of a larger group of parameters. In Appendix F, we report the results from a Monte Carlo experiment where we also estimate the subjective discount rate, ρ, and the long-run growth rate of the economy, η, besides $\boldsymbol{\theta}_{\text{exo}}$. The inclusion of these additional parameters was made after verifying the rank and order conditions for local identification (see Remark 4.2).

4.2.2 Flow data

The second Monte Carlo experiment assumes instead that all the discrete observations are sampled as flows. Therefore, we conduct ML estimation of $\boldsymbol{\theta}_{\text{exo}}$ using the exact likelihood function derived under the F-SSR model in (2.12)-(2.13). The results are summarized in Panel B of Table 2, where we also report the ML estimation results from two misspecified state space models, namely, the S-SSR and EM-SSR models, which are both intended to accommodate stock variables, and therefore do not take into account the integral nature of the observations between measurement times. This type of misspecification is referred to as *temporal aggregation* bias, since it reflects the problem of assuming that economic agents make decisions at time intervals coinciding with the sampling interval (see Christiano and Eichenbaum, 1987). For the EM-SSR model, the temporal aggregation bias adds to the discretization bias resulting from the EM approximation of the transition equation.

The distribution of the parameter estimates under the three alternative models is displayed in Panel B of Figure 3. The bias-variance trade-off appears also in the case of flow measurements, and in this case it is present for all the estimated parameters.

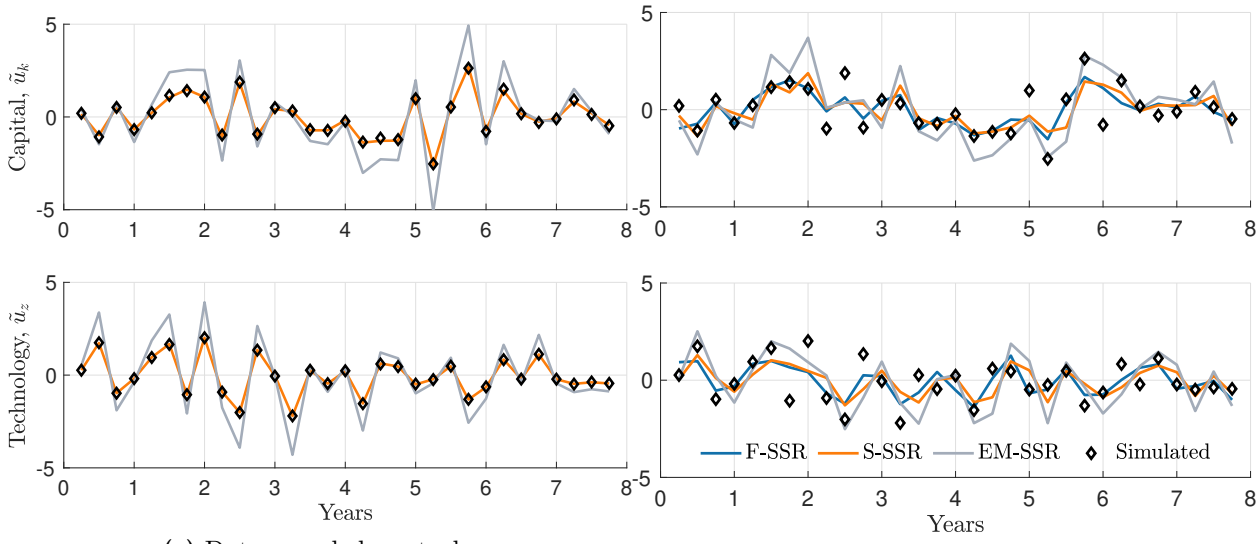
Our results suggest that using the correct specification of the state space model is critical for the MLE to deliver accurate parameter estimates in finite samples. Using the F-SSR model produces ML estimates that, on average, are close to their true value in the population. The magnitude and direction of the biases are similar to those discussed in the previous section, for stock data. Thus, the bias for ρ_z is positive, and just below 1.5% in magnitude, while the biases for σ_z and σ_k are negative and below 1%. These conclusions hold across different subsets of observables (the results are available upon request).

In contrast, using a misspecified state space model introduces substantial biases in the ML estimation of the volatility parameters, with a simultaneous increase in RMSEs. The downward bias in σ_k increases to values between 22% and 29%, depending on the model used, and the bias in σ_z increases to almost 19%. Using the S-SSR model when data are sampled as flows introduces a bias in σ_z beyond the pure estimation bias that cannot be attributed to discretization errors. In this case, the additional bias emerges due to time aggregation. On the other hand, the bias in ρ_z is negative for both the EM-SSR and S-SSR models. While the bias induced by former is considerably larger (15.5%) than that of the latter (1.1%), both values reflect an otherwise strong dominance of the misspecification bias over the pure ML estimation bias. Regarding the RMSE, we find that the variability of the ML estimates under the S-SSR model is similar to that under the F-SSR model. However, the RMSE for the EM-SSR model nearly doubles, compared to that for the F-SSR model. The downward bias induced by the model misspecification is evident from the log-likelihood profiles in Figure F.1 in Appendix F. For flow variables, the bias is noticeable, not only for the speed of mean reversion, but also for the diffusion parameters.

In summary, our Monte Carlo results suggest that ignoring the state space representation that accommodates the flow nature of the measurements comes at a cost in terms of larger estimation bias and RMSE. Although the derivation of the exact state space model is slightly more complicated for flow variables than for stock variables, or for the simple EM approximation, the benefits of pursuing it seem worthwhile in the context of estimation of DSGE models.

4.3 Structural shocks

We now proceed to evaluate the accuracy of the approach introduced in Section 3 to approximately recover the model's underlying structural shocks at measurement times, using (3.8). To control for the variability from the ML estimation, and to separate effects, we refrain from using the ML estimates at this stage, and instead compute our approximations conditional on the true DGP. That is, we fix all the parameter values at those in Table 1. In our experiments, we employ the same simulated data sets on consumption and the fraction



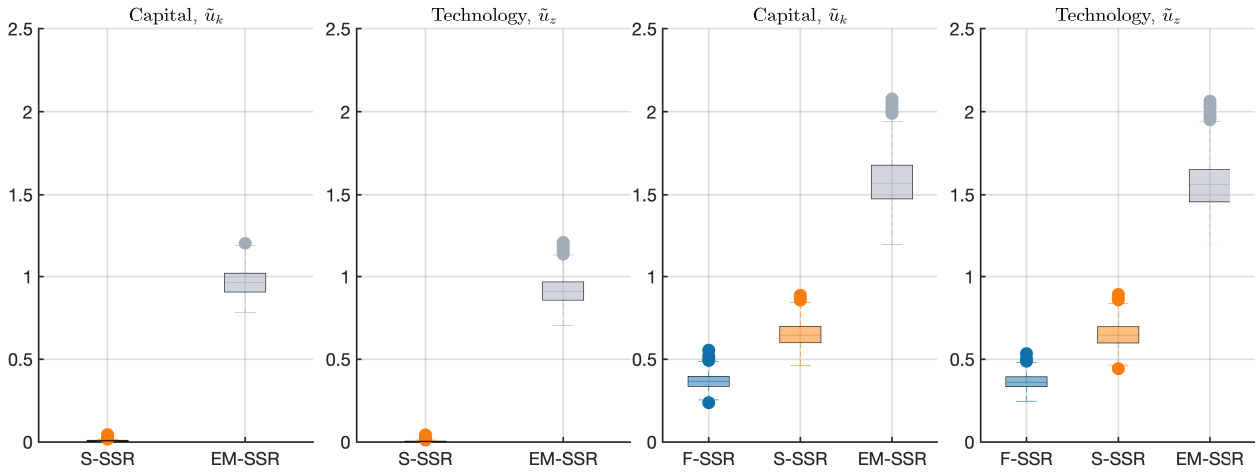
(a) Data sampled as stocks

(b) Data sampled as flow

Figure 4. Structural shocks. The graph plots the time series of structural shocks to TFP, $\tilde{u}_z$, and capital stock, $\tilde{u}_k$, recovered from one sample of simulated data on consumption, C, and hours worked, N. Simulated measurements are sampled as stocks in Exhibit (a), and as flows in Exhibit (b).

of hours worked that were used in the Monte Carlo exercises, both for the stock and flow data. For each sampling scheme, we compare the structural shocks recovered from the alternative state space representations, i.e., the S-SSR, F-SSR, and EM-SSR models. In (3.8), the matrix $\mathbf{H}(\boldsymbol{\theta})$ is defined as $h^{1/2}\mathbf{H}(\boldsymbol{\theta}, h)$, with $\mathbf{H}(\boldsymbol{\theta}, h)$ given by (3.3) for the S-SSR model, (3.4) for the F-SSR model, and simply by $\mathbf{B}(\boldsymbol{\theta})$ for the EM-SSR model. Recall that for stock data, the S-SSR model is an exact representation of the DGP at measurement times, while the EM-SSR model is an approximation that induces a misspecification error associated with temporal aggregation. For flow data, the F-SSR model is the exact discrete representation, and the other two misspecified.

To gain some intuition on the mechanics of our approach, Figure 4 displays the time series of structural shocks to the capital stock, $\tilde{u}_{k,\tau}$ (top exhibits), and to TFP, $\tilde{u}_{z,\tau}$ (bottom exhibits), for the first eight years of a given simulated sample. Left exhibits (labeled (a)) show the time series recovered using the S-SSR and EM-SSR models, together with the true underlying structural shocks, denoted with a $\Diamond$, for the case when data are simulated as stocks. Right exhibits (labeled (b)) present the time series recovered using the F-SSR, S-SSR, and EM-SSR models for the case when data are simulated as flows. For the particular sample used, we observe that the method performs remarkably well in the case of stock data, conditional on correct state space representation. Thus, the plots in exhibits (a) reveal that by using the S-SSR model we are able to recover the structural shocks at measurement times



(a) Stock variables **(b)** Flow variables

Figure 5. Approximation errors in estimated structural shocks. The graph shows the distribution of the mean squared error (MSE) between the true (simulated) structural shocks and their estimated (smoothed) counterparts. Each boxplot represents the distribution of MSE across Monte Carlo simulations. Results are shown for simulated data sampled as stocks in Exhibit (a), and as flows in Exhibit (b).

with great precision, while using the EM-SSR model misses the task in many of the periods under consideration. The poor performance of the latter model is driven by the misspecification induced by the EM approximation that disregards all the model-based information between measurements when computing the variance-covariance matrix of the shocks. For the case of flow data, the conclusions are less evident. Although it is clear that using the EM-SSR model produces the worst of the three approximations to the structural shocks, it is not obvious whether the F-SSR model produces a more accurate approximation than the misspecified S-SSR model.

Our approximation drops the remainder term, $\mathcal{R}_\tau$, in both the stock and flow cases. The results suggest that the remainder is small, particularly in the stock case, at least for the particular sample considered. To better understand the quality of the approximations, we repeat the exercise on all the $M = 500$ samples. The results are summarized in Figure 5, where we plot the mean squared errors (MSE) and their dispersion across simulations. Several conclusions emerge. First, using the exact discrete representation always outperforms the EM-SSR model. The S-SSR and F-SSR models produce lower MSEs in general. Second, for variables sampled as stocks, not only is the MSE for both capital and technology shocks obtained from the S-SSR model low on average, it also displays minimal dispersion. By contrast, the MSE obtained from the EM-SSR model exhibits both a large mean and high variability, with some notable outliers. Third, when the observables are sampled as flows, the MSEs are, on average, larger than those obtained for the stock case. However, the

resulting MSEs from the F-SSR and S-SSR models exhibit lower mean and variability than those produced by the EM-SSR model. Fourth, despite being the exact discrete representation for flow data, the F-SSR model does not produce (approximately) zero MSEs, as the S-SSR model does for stock variables. To understand why this is the case, recall that the approach requires solving the system of equations in (3.7). For the RBC model in Section 4.1, this requires solving a square system when data are sampled as stocks, and a rectangular system for flows. Thus, there are more equations than unknowns in the flow case, and some equations are inevitably solved with error.[13]

5. Empirical illustration

This section provides an illustration of the maximum likelihood estimation method of Section 2, and of the structural shock identification of Section 3, using U.S. data. In the first exercise, we estimate a subset of the parameters of the RBC model of Section 4, using quarterly data on aggregate consumption and the fraction of hours worked for the period 1959 to 2019.[14] Both variables are obtained from the Federal Reserve Economic Data database, FRED. Aggregate consumption is measured by the monthly nominal personal consumption expenditures (PCE), deflated by the corresponding monthly price index (PCEPI), both from the NIPA tables. Aggregate consumption at quarterly frequency is computed by aggregating monthly real expenditures over the quarter. Quarterly hours worked correspond to the number of hours from wage and salary workers on nonfarm payrolls (TOTLQ). All variables are transformed into per-capita values using the civilian, non-institutional population, aged 16 and over (CNP16OV) from the U.S. Bureau of Labor Statistics. With the exception of population, all variables are seasonally adjusted. We assume that the observed variables are trending exponentially at a constant growth rate of 2% per year that captures the long-run economic growth rate in the model. Figure 6 shows the time series used in the ML estimation. The gray areas represent recession periods, as dated by the NBER.

In a second exercise, we use the estimated residuals to recover the structural shocks at measurement times, conditional on the ML estimates. We then use the estimated sequence of structural shocks to build a historical shock decomposition of the observed variables. This exercise is usually employed in economic policy circles to build narratives around the sources of past economic fluctuations. Other potential applications include impulse-response analysis, and forecast error variance decompositions (see, e.g., Lütkepohl, 2005, Canova 2007,

[13] Dropping the remainder term in our approximation means that there is some error in all cases, but the results show that this is small for stock sampling, particularly under correct specification.

[14] Estimation results using alternative subsets of observables are available in Appendix F.

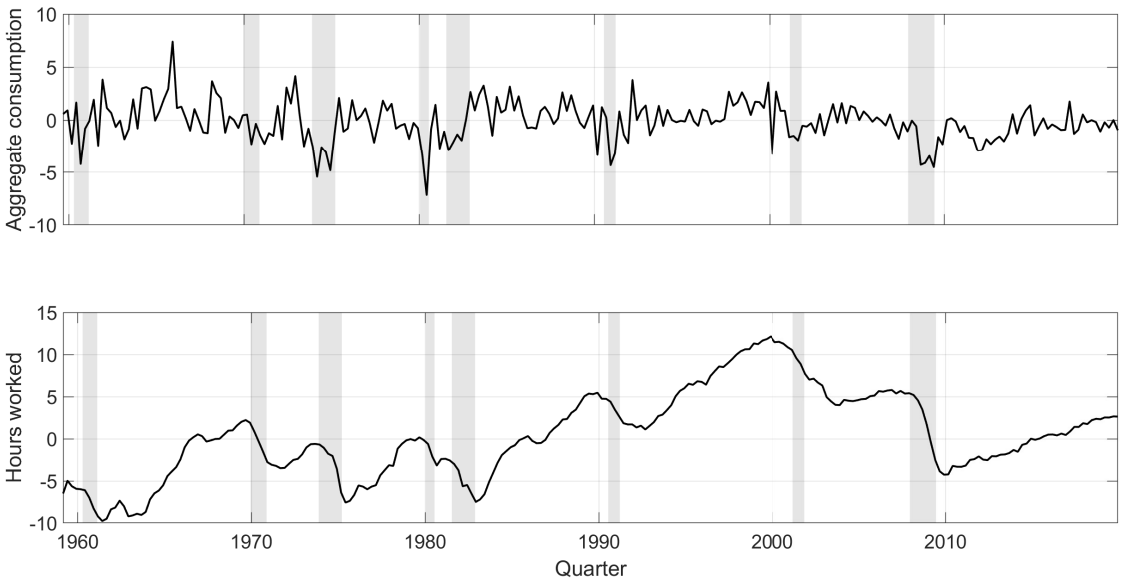


Figure 6. Measurements. The figure shows the annualized growth rates of the real aggregate consumption per capita series in the upper panel ($400 \cdot (\log(C_\tau/C_{\tau-1}) - \eta h)$), and in the lower panel the logarithm of hours worked, expressed in percentage deviation from the steady state ($100 \cdot (\log(N_\tau) - n^\star)$). The sample spans the period from 1959:Q1 to 2019:Q4.

and Fernández-Villaverde et al., 2016).

The final exercise reports ML estimation results from a state space representation that accommodates observations at mixed frequencies. In particular, we estimate the model parameters using monthly (rather than quarterly) data on aggregate real consumption, together with quarterly data on hours worked. The time series for monthly aggregate consumption is shown in Appendix F.

5.1 ML estimation

Analogously to the Monte Carlo experiments in Section 4, we report ML estimates for $\boldsymbol{\theta}_{\text{exo}}$, while fixing $\boldsymbol{\theta}_{\text{ss}}$ at the values in Table 1. We estimate the model parameters with and without allowance for (uncorrelated) measurement errors in the observables. For each case, we report parameter estimates from the three different state space representations analyzed earlier, i.e., the F-SSR, the S-SSR, and EM-SSR models. This allows us to remain agnostic about the way in which the data are sampled. The state vector in each model is initialized at the unconditional mean and the unconditional covariance matrix. Table 3 summarizes the results for the case in which the observables are aggregate consumption, C, and hours worked, N. Besides the point estimates, the table reports estimated standard errors (in parentheses), computed

Table 3. ML estimates. The table reports the ML estimates of $\boldsymbol{\theta}_{\text{exo}} = [\rho_z, \sigma_z, \sigma_k]^\top$ for the model in Section 4, using quarterly data on aggregate consumption, C, and hours worked, N, for the U.S., over the period from 1959:Q1 to 2019:Q4. The remaining parameters of the model, $\boldsymbol{\theta}_{\text{ss}}$, are fixed at the values in Table 1. Wild bootstrap standard errors computed from $B = 499$ samples are reported in parentheses.

	without measurement error			with measurement error		
$\boldsymbol{\theta}_{\text{exo}}$	F-SSR	S-SSR	EM-SSR	F-SSR	S-SSR	EM-SSR
ρ_z	0.0345	0.0421	0.0356	0.0258	0.0370	0.0303
	(0.0175)	(0.0213)	(0.0183)	(0.0144)	(0.0169)	(0.0167)
σ_z	0.0125	0.0109	0.0110	0.0127	0.0110	0.0111
	(0.0013)	(0.0010)	(0.0011)	(0.0012)	(0.0010)	(0.0011)
σ_k	0.0170	0.0149	0.0147	0.0169	0.0147	0.0145
	(0.0014)	(0.0012)	(0.0012)	(0.0014)	(0.0012)	(0.0012)

using the wild bootstrap algorithm, with $B = 499$ samples (see Angelini et al., 2021).[15] We choose random sampling methods for the computation of the standard errors, instead of the asymptotic formulas, since they better accommodate the finite-sample nature of the data. In particular, they allow for heteroscedasticity in the forecast errors, which is a desirable feature when conducting empirical analysis. Moreover, the bootstrap procedure is well suited for situations in which a parameter estimate is close to the boundary of its domain, thus leaving the computation of numerical derivatives of the maximized log-likelihood function challenging .

In general, our results show that the point estimates of the parameters representing the driving forces behind the real business cycle model are similar across the various specifications, and quite reasonable. In the absence of measurement errors, the speed of mean reversion of TFP is estimated at $\rho_z = 0.0345$ in the F-SSR model, 0.0421 in the S-SSR model, and 0.0356 in the EM-SSR model. Given the sampling frequency of the data, these estimates imply a first order autocorrelation coefficient for quarterly TFP between 0.9895 and 0.9914 ($= \exp(-\rho_z/4)$). Our results are close to the quarterly estimates reported in Ireland (2001, 2004) and Malley and Woitek (2010) for the U.S. economy, and thus consistent with the arguments in King et al. (1988a,b) and Hansen (1997), according to which the shocks to TFP should be highly persistent for the RBC model to match key features of the U.S. data.[16] Further, the point estimates of the instantaneous volatility of TFP shocks, σ_z, imply a quarterly conditional volatility between 0.54% and 0.62% (see Remark 2.3), values that are

[15] Assumptions 1–3, A4′, and 5 in Angelini et al. (2021) are verified prior to the implementation of the boostrapping procedure.

[16] The persistence of TFP in Table 1 is based on quarterly autoregressions for the Solow residual, with coefficients estimated at around 0.95, whereas we use state space methods (as in Ireland (2001, 2004) and Malley and Woitek (2010)) to estimate the speed of mean reversion of TFP directly.

again close to those reported by the same authors, and of the same order of magnitude as the those commonly used in the RBC literature. Moreover, our estimates suggest that the variability of the shocks to the capital stock is greater than that in Ambler and Paquet (1994), and greater than that to TFP.[17] In particular, the estimates of σ_k imply a volatility between 0.74% and 0.86% per quarter ($\approx \sqrt{(1/4)}\sigma_k$). When comparing across model specifications, however, the F-SSR model usually produces larger estimates of the instantaneous volatilities, σ_z and σ_k, than the S-SSR and EM-SSR models, in line with the Monte Carlo evidence provided earlier (see Panel B of Table 2, and Panel B of Figure 3). Finally, Figure 7 shows the filtered time series of unobserved aggregate capital stock and TFP from the F-SSR model at quarterly frequency, in percentage deviations from their corresponding steady-state values.[18]

In order to assess the effects of allowing for iid measurement errors in the ML estimation, we fix their standard deviations to the quarterly estimates reported in Ireland (2004, Table 6), namely, $\sigma_{\varepsilon_C} = 0.0098$ for aggregate consumption, and $\sigma_{\varepsilon_N} = 0.0020$ for hours worked. Our results (right side of Table 3) suggest that allowing for measurement error does not have sizable effects on the structural parameter estimates. Overall, the ML estimates imply persistence levels of quarterly TFP that are of the same order as those without measurement errors, across model specifications. The quarterly volatility of the the capital stock is estimated at about 1% below the no-measurement error case, and the difference is small relative to the standard errors.

5.2 Historical shock decomposition

Using the ML estimates, we employ the approach of Section 3 to recover the time series of structural shocks at measurement times. Given these, we next investigate the historical contribution over time of the structural shocks to aggregate consumption and hours worked in the U.S., through the lens of the continuous-time RBC model. The results from this exercise are displayed in Figure 8, based on the historical shock decomposition

$$HSD(\mathbf{y}_\tau, i) = \mathsf{C}(\boldsymbol{\theta})\mathsf{A}(\boldsymbol{\theta})^{\tau-1}\mathbf{x}_0 + \mathsf{D}(\boldsymbol{\theta})\mathbf{H}(\boldsymbol{\theta})\mathbf{S}_i\tilde{\mathbf{u}}_\tau + \mathsf{C}(\boldsymbol{\theta})\sum_{s=0}^{\tau-2}\mathsf{A}(\boldsymbol{\theta})^s\mathsf{B}(\boldsymbol{\theta})\mathbf{H}(\boldsymbol{\theta})\mathbf{S}_i\tilde{\mathbf{u}}_{\tau-1-s}\,, \quad (5.1)$$

for $i = \{k, z\}$ computed from the ABCD representation (2.16)-(2.17) associated with the F-SSR model. Here, $\mathbf{S}_i$ is a selection matrix picking out shocks to the capital stock or TFP,

[17] In contrast, for the parameter values in Table 1, the variability of the shocks to TFP is greater than that of the shocks to the aggregate capital stock.

[18] Table F.3 in Appendix F reports the results from a battery of tests assessing the univariate and multivariate normality of the one-step ahead prediction errors of the measurements, $\boldsymbol{\omega}_{\tau+1|\tau}$, in the F-SSR model across different subsets of the data used in the estimation. The results are generally consistent with Gaussianity.

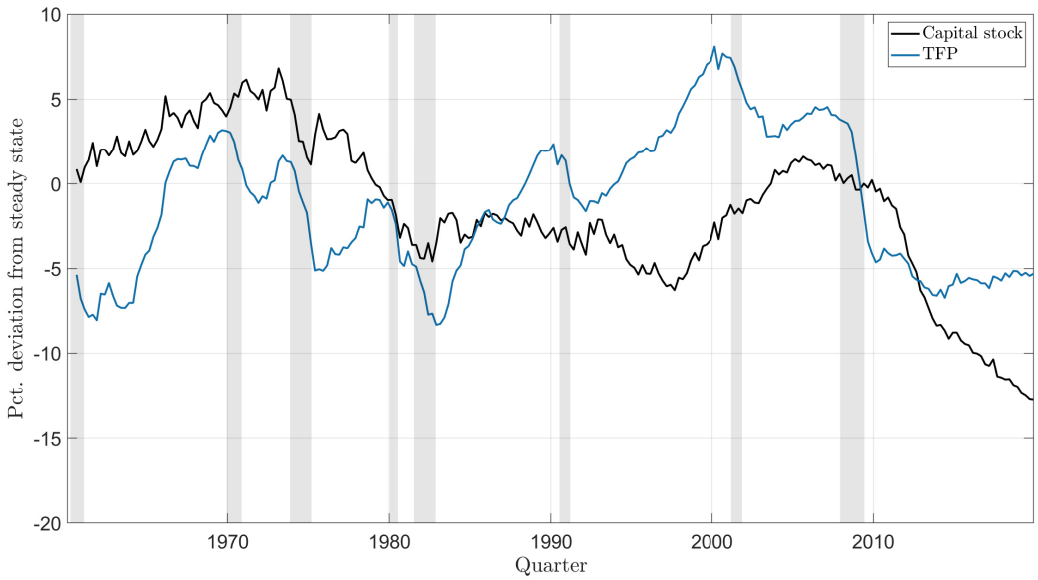


Figure 7. Latent states. Filtered quarterly series of the aggregate capital stock and TFP (% deviation from their corresponding steady-state values). Sample spans from 1960:Q1 to 2019:Q4. The series are obtained from the F-SSR model after parameter estimation and filtering using as measurements quarterly real PCE and quarterly hours worked.

and $\mathbf{x}_0$ contains the initial values. The derivation of (5.1) is given in Appendix D.7.

The top panel of Figure 8 displays the contributions from shocks to TFP, $\tilde{u}_z$, and to the capital stock, $\tilde{u}_k$, to annual consumption growth, and the bottom panel their contributions to quarterly deviations of hours worked from their steady state value. Any discrepancies between the sum of the bars and the observed series at a particular point in time can be attributed to the $\mathcal{O}_P(h^{3/2})$ term in Proposition 6. The structural shocks represent the driving forces behind the dynamics of the aggregate capital stock and TFP in Figure 7. Historical decompositions using the S-SSR and the EM-SSR models can be found in Appendix F. Note that for the case of aggregate consumption, we report the historical decomposition of the annual growth rates, although the model is estimated using quarterly data. As indicated in Remark 2.3, this shift is achieved seamlessly, due to the frequency invariance of the structural parameters in the continuous-time model. Once the model parameters have been estimated using quarterly data, it is straightforward to recalibrate the ABCD representation to the frequency of interest, in this case annual.

If shocks to TFP are interpreted as "aggregate supply shocks", and shocks to the capital stock as "aggregate demand shocks", the decomposition in Figure 8 suggests that the U.S. business cycle over the period 1962:Q1 through 2019:Q4 mainly has been driven by aggregate supply shocks. Consumption growth is explained by shocks to firm productivity, both during

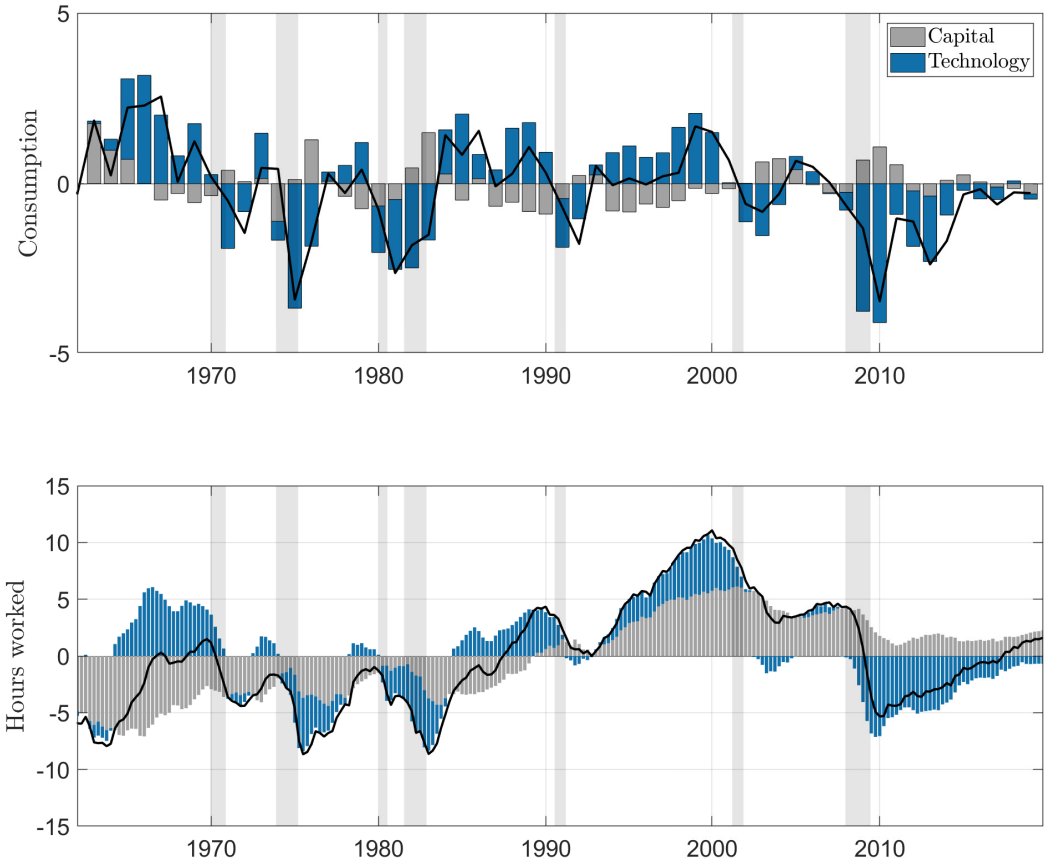


Figure 8. Historical shock decomposition (F-SSR model). The plot shows the historical contribution of each of the structural shocks recovered from the estimated F-SSR model on the observed measurements over the period 1962:Q1-2019:Q4, expressed in percent. The black solid line in the upper panel represents annual consumption growth rates. The black solid line in the lower panel represents quarterly percentage deviations of hours worked from its steady state ($n^\star = 33\%$). The light gray vertical bands indicate NBER recessions.

expansions and contractions. In contrast, with only a few exceptions, the short-run variability of hours worked, relative to their long-run value, has mostly been driven by shocks to capital accumulation, i.e., aggregate demand shocks. This is consistent with the evidence on the limited power of TFP shocks to explain the behavior of hours worked in RBC models, in particular their unconditional variance (see, e.g., Cooley and Prescott, 1995 and Ireland, 2004).

Table 4. Mixed-frequency estimates. The table reports the ML estimates of $\boldsymbol{\theta}_{\text{exo}} = [\rho_z, \sigma_z, \sigma_k]^\top$ for the model in Section 4, using monthly data on real PCE (per capita), and quarterly non-farm hours worked. The sample spans the period from March 1959 ,to December 2019. The remaining parameters of the model, $\boldsymbol{\theta}_{\text{ss}}$, are fixed at the values in Table 1. Wild bootstrap standard errors computed from $B = 499$ samples are reported in parentheses.

	Mixed Frequency without measurement error		
$\boldsymbol{\theta}_{\text{exo}}$	MXF-SSR	MXS-SSR	MXEM-SSR
ρ_z	2.97E-06	0.0044	0.0031
	(0.0496)	(0.0093)	(0.0083)
σ_z	0.0118	0.0133	0.0133
	(0.0007)	(0.0009)	(0.0009)
σ_k	0.0277	0.0179	0.0178
	(0.0021)	(0.0012)	(0.0012)

5.3 Mixed-frequency estimation

In our final exercise, we estimate the same subset of parameters, $\boldsymbol{\theta}_{\text{exo}}$, using monthly (rather than quarterly) data on aggregate real consumption per capita, together with quarterly observations on the fraction of hours worked. In this case of mixed frequencies, the state space representation must be adjusted accordingly. If observations are sampled as stocks, we simply write the S-SSR and EM-SSR models in (2.8)-(2.9) and (4.13)-(2.9), respectively, in terms of the variable(s) observed at the highest frequency, then modify the filtering algorithm to accommodate the corresponding missing observations of the variable(s) observed at any other lower frequencies (see Durbin and Koopman, 2012, Sec. 4.10).[19] A similar treatment is applied when data are sampled as flows. However, in this case, we extend the state vector of the F-SSR model in (2.12)-(2.13) with additional deterministic variables that allow keeping track of the time aggregation at high frequencies of the variable(s) that are sampled at lower frequencies. We refer to the mixed-frequency state space representations as the MXF-SSR, MXS-SSR, and MXEM-SSR models. The details of the derivations of the state space representations in the case of mixed frequencies are given in Appendix C.

Table 4 displays the mixed-frequency ML estimates of $\boldsymbol{\theta}_{\text{exo}}$ from the MXF-SSR, MXS-SSR, and MXEM-SSR models without measurement errors. Standard errors (in parentheses) are computed using the wild bootstrap algorithm with $B = 499$ samples. Although not directly comparable, the results do not differ considerably from those obtained under a

[19] The case of mixed-frequency sampling frequency has been addressed in the context of state space models by Harvey and Pierse (1984), Zadrozny (1988), Harvey (1990), Zadrozny (1990), Mariano and Murasawa (2003), Aruoba et al. (2009), Ghysels and Wright (2009), Kuzin et al. (2011), and Bai et al. (2013), among others.

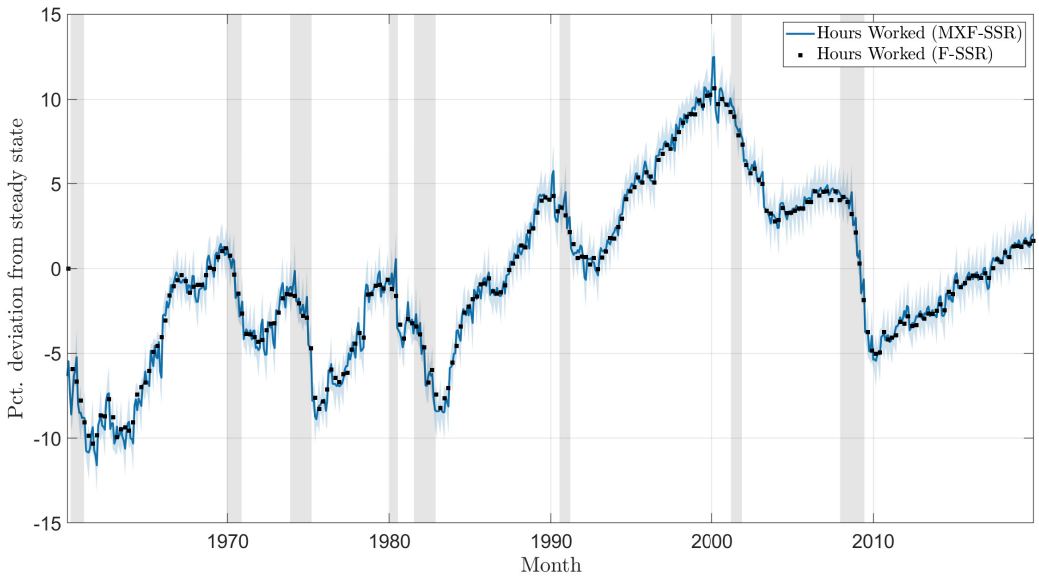


Figure 9. Filtered series of hours worked. The figure shows the time series of filtered values of the annualized fraction of hours worked, as percentage deviation from its steady-state value. Black dots represent quarterly values generated from the F-SSR model. The continuous blue line represents monthly values generated from the MXF-SSR model. Light blue bands indicate uncertainty about the monthly state estimates. The sample spans the period from January 1960 to December 2019.

common quarterly sampling frequency, as reported in Table 3. This holds particularly for the point estimates of the volatility parameters, σ_z and σ_k, and their associated standard errors. The implied estimates of the quarterly conditional volatilities of TFP and the capital stock oscillate between 0.59% and 0.67%, and between 0.89% and 1.35%, respectively.

The conclusions are less clear for the speed of mean reversion, ρ_z. The point estimates under mixed-frequency sampling are closer to zero across the different model specifications than in the case of a common frequency sampling. Our estimates imply a first order auto-correlation coefficient for quarterly TFP of 0.999. This confirms the (near) unit root, and thus highly persistent, behavior of TFP that is necessary in RBC models in order to match the time-series properties of macroeconomic data. On the other hand, the standard errors increase in the MXF-SSR model, and decrease in the MXS-SSR and MXEM-SSR models, relative those obtained under common frequency sampling.

Despite the small differences in some of the estimated parameters, the models produce similar estimates of the state variables. As an example, Figure 9 plots the time series of annualized values of the fraction of hours worked in the U.S. (as percentage deviation from its steady-state value) predicted by the model when using a state space representation for flow data. The black dots represent the quarterly filtered values recovered from the F-SSR

model using the parameter estimates in Table 3, and the continuous blue line shows the monthly filtered series from the MXF-SSR model using the parameter estimates in Table 4. From the figure, we conclude that the model produces similar predicted values of hours worked when using different state space representations, with different (estimated) parameter values. Similar conclusions are obtained for the unobserved states (capital stock and TFP) by comparing the time series of filtered states from the F-SSR model in Figure 7 to the filtered states from the MXF-SSR model in Appendix F Figure F.5. This exercise highlights the potentials of using a mixed-frequency approach in continuous-time linear models, viz., estimating time series for the variables in the model at a sampling frequencies that differ from that of the data, while simultaneously obtaining ML estimates of the model parameters that are invariant to any frequency.

6. Conclusions

We introduce a state space framework to conduct maximum likelihood (ML) inference on continuous-time linear(-ized) DSGE models from a sample of regularly spaced macroeconomic data observed at discrete points in time, i.e., not observed continuously. The approach uses the exact discrete representation of the equilibrium transition dynamics to derive a reduced-form discrete-time state space representation that does not involve discretization errors. Using the Kalman filter, it is possible to compute the exact likelihood function of the data by taking into account the sampling nature of the data, i.e., stock versus flow sampling. Using a continuous-time version of an otherwise standard RBC model, Monte Carlo experiments show that our approach delivers accurate estimates under the null of correct model specification. Indeed, using a state space representation that correctly accommodates the sampling nature of the data is important for the ML estimator to deliver accurate estimates in finite samples. The results obtained from using the exact state space model outperform those from a state space representation based on a naive Euler-Maruyama discretization of the equilibrium transition dynamics. This is particularly true in our setup, since the macroeconomic data usually employed in the estimation of DSGE models is only sampled at low frequencies (e.g., monthly, quarterly, annually). Our state space based results confirm previous evidence on the direction of the observed biases in finite samples.

We provide a detailed discussion of the aliasing identification problem that emerges when using the discrete-time representation of a continuous-time model in multivariate settings. We provide necessary and sufficient conditions for ruling out the existence of aliases, and argue why this is a precondition for achieving local identification of the model parameters along the lines of Komunjer and Ng (2011). Obtaining necessary and sufficient conditions

for local identification that simultaneously rule out aliases is left for future research.

Further, we introduce a method to approximately recover the underlying structural shocks of the continuous-time model at measurement times from the reduced-form residuals of the state space representation. The method is analogous to the use of short- and long-run identifying restrictions common in the structural VAR literature. Our Monte Carlo results suggest that the method recovers the unobserved realizations of the structural shocks with great precision when using the exact discrete-time state space representation under correct model specification. Under these circumstances, our approach is successful at incorporating sufficient instantaneous causal links from the economic theory so as to identify independent structural shocks. In contrast, the poor performance of the alternative naive discretization is explained by its inability to account for the information from model between measurements when computing the variance-covariance matrix of the structural shocks.

We apply our approach to the same RBC model, using quarterly U.S. data from 1959:Q1 through 2019:Q4. We consider the aggregate consumption and hours worked series as a benchmark case, and compare results to those based on other data configurations, including aggregate output, and other data frequencies. The results confirm that our approach is feasible, and estimates make sense. Estimates of the rate of mean reversion of TFP, as well as the volatilities of shocks to TFP and the aggregate capital stock, both treated as latent state variables, are consistent with those in related literature. Based on the structural shocks recovered from the observed series, a historical decomposition indicates that the U.S. business cycle has mainly been driven by aggregate supply shocks over the period analyzed, as indicated by the dominant contribution of TFP shocks to consumption growth. At the same time, deviations in hours worked from the steady state have mainly been driven by aggregate demand shocks, as identified with the recovered shocks to capital. Finally, an application using data series of mixed frequencies, monthly consumption and quarterly hours worked, illustrates the generality of the approach, with the empirical methods all based on the underlying continuous-time model, and thus frequency-invariant parameters, hence facilitating consistent analysis across all desired frequencies in data, predictions, and so forth.

References

Achdou, Y., J. Han, J.-M. Lasry, P.-L. Lions, and B. Moll (2022): "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach," *The Review of Economic Studies*, 89, 45–86. (Cited on page 2.)

Ahn, S., G. Kaplan, B. Moll, T. Winberry, and C. Wolf (2018): "When Inequality Matters for Macro and Macro Matters for Inequality," *NBER Macroeconomics Annual*, 32, 1–75. (Cited on pages 2 and 80.)

Aït-Sahalia, Y. (2002): "Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-form Approximation Approach," *Econometrica*, 70, 223–262. (Cited on page 4.)

——— (2008): "Closed-form Likelihood Expansions for Multivariate Diffusions," *The Annals of Statistics*, 36, 906–937. (Cited on page 4.)

Aït-Sahalia, Y. and P. A. Mykland (2003): "The Effects of Random and Discrete Sampling when Estimating Continuous–Time Diffusions," *Econometrica*, 71, 483–549. (Cited on page 4.)

——— (2004): "Estimators of Diffusions with Randomly Spaced Discrete Observations: A General Theory," *The Annals of Statistics*, 32, 2186–2222. (Cited on page 4.)

Ambler, S. and A. Paquet (1994): "Stochastic Depreciation and the Business Cycle," *International Economic Review*, 35, 101–116. (Cited on pages 27 and 39.)

Anderson, B. D. and J. B. Moore (2012): *Optimal Filtering*, Courier Corporation. (Cited on page 13.)

Anderson, G. (1997): "Continuous Time Application of the Anderson-Moore (AIM) Algorithm for Imposing the Saddle Point Property in Dynamic Models," *Unpublished*. (Cited on page 26.)

Angelini, G., G. Cavaliere, and L. Fanelli (2021): "Bootstrap Inference and Diagnostics in State Space Models: With Applications to Dynamic Macro Models," *Journal of Applied Econometrics*, 1–20. (Cited on page 38.)

Aruoba, S., F. X. Diebold, and C. Scotti (2009): "Real-Time Measurement of Business Conditions," *Journal of Business & Economic Statistics*, 27, 417–427. (Cited on page 42.)

Bai, J., E. Ghysels, and J. Wright (2013): "State Space Models and MIDAS Regressions," *Econometric Reviews*, 32, 779–813. (Cited on page 42.)

Bartlett, M. S. and D. V. Rajalakshman (1953): "Goodness of Fit Tests for Simultaneous Autoregressive Series," *Journal of the Royal Statistical Society. Series B (Methodological)*, 15, 107–124. (Cited on pages 3 and 8.)

Bergstrom, A. R. (1966): "Nonrecursive Models as Discrete Approximations to Systems of Stochastic Differential Equations," *Econometrica*, 34, 173–182. (Cited on pages 3 and 8.)

——— (1983): "Gaussian Estimation of Structural Parameters in Higher Order Continuous Time Dynamic Models," *Econometrica*, 51, 117–152. (Cited on page 4.)

——— (1984): "Continuous Time Stochastic Models and Issues of Aggregation over Time," in *Handbook of Econometrics*, Elsevier, vol. 2, 1145–1212. (Cited on pages 3 and 8.)

Bernanke, B. S. (1986): "Alternative Explanations of the Money-Income Correlation," *Carnegie-Rochester Conference Series on Public Policy*, 25, 49–99. (Cited on page 5.)

Bibby, B. M. and M. Sørensen (1995): "Martingale Estimation Functions for Discretely Observed Diffusion Processes," *Bernoulli*, 1, 17–39. (Cited on page 4.)

Blanchard, O. J. and C. M. Kahn (1980): "The Solution of Linear Difference Models under Rational Expectations," *Econometrica*, 48, 1305–1311. (Cited on page 26.)

Blanchard, O. J. and D. Quah (1989): "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, 79, 655–673. (Cited on page 5.)

Blevins, J. R. (2017): "Identifying Restrictions for Finite Parameter Continuous Time Models with Discrete Time Data," *Econometric Theory*, 33, 739–754. (Cited on pages 4 and 17.)

Bowman, K. O. and L. R. Shenton (1975): "Omnibus Test Contours for Departures from Normality based on $\sqrt{b_1}$ and b_2," *Biometrika*, 62, 243–250. (Cited on page 90.)

Brunnermeier, M. and Y. Sannikov (2014): "A Macroeconomic Model with a Financial Sector," *American Economic Review*, 104, 379–421. (Cited on pages 2 and 24.)

Buiter, W. H. (1984): "Saddlepoint Problems in Continuous Time Rational Expectations Models: A General Method and Some Macroeconomic Examples," *Econometrica*, 52, 665–680. (Cited on pages 26 and 81.)

Canova, F. (2007): *Methods for Applied Macroeconomic Research*, Princeton University Press. (Cited on page 36.)

Canova, F. and L. Sala (2009): "Back to Square One: Identification Issues in DSGE models," *Journal of Monetary Economics*, 56, 431–449. (Cited on pages 4, 14, 15, and 27.)

Chambers, M. J. (1999): "Discrete Time Representation of Stationary and Non-Stationary Continuous Time Systems," *Journal of Economic Dynamics and Control*, 23, 619–639. (Cited on page 3.)

Chambers, M. J., J. R. McCrorie, and M. A. Thornton (2018): "Continuous Time Modelling Based on an Exact Discrete Time Representation," in *Continuous Time Modeling in the Behavioral and Related Sciences*, ed. by K. van Montfort, J. H. L. Oud, and M. C. Voelkle, Springer International Publishing, 317–357. (Cited on page 3.)

Chambers, M. J., T. Simos, and M. Tsionas (2022): "Locally Exact Discrete Time Representations of Non-Linear Continuous Time Models with an Application to the Estimation of a DSGE Model," Tech. rep., mimeo. (Cited on page 4.)

Chang, F.-R. (2009): *Stochastic Optimization in Continuous Time*, Cambridge University Press. (Cited on page 76.)

Christensen, B. J., O. Posch, and M. Van Der Wel (2016): "Estimating Dynamic Equilibrium Models using Mixed Frequency Macro and Financial Data," *Journal of Econometrics*, 194, 116–137. (Cited on page 4.)

Christiano, L. J. and M. Eichenbaum (1987): "Temporal Aggregation and Structural Inference in Macroeconomics," *Carnegie-Rochester Conference Series on Public Policy*, 26, 63–130. (Cited on page 32.)

Cooley, T. F. and E. C. Prescott (1995): "Economic Growth and Business Cycles," in *Frontiers of Business Cycle Research*, ed. by T. F. Cooley, Princeton, New Jersey: Princeton University Press, chap. 1, 1–38. (Cited on page 41.)

Culver, W. J. (1966): "On the Existence and Uniqueness of the Real Logarithm of a Matrix," *Proceedings of the American Mathematical Society*, 17, 1146–1151. (Cited on pages 16 and 71.)

Del Negro, M. and F. Schorfheide (2008): "Forming Priors for DSGE Models (and How it Affects the Assessment of Nominal Rigidities)," *Journal of Monetary Economics*, 55, 1191–1208. (Cited on page 28.)

Doornik, J. A. and H. Hansen (2008): "An Omnibus Test for Univariate and Multivariate Normality," *Oxford Bulletin of Economics and Statistics*, 70, 927–939. (Cited on page 90.)

Durbin, J. and S. J. Koopman (2012): *Time Series Analysis by State Space Methods*, Oxford University Press. (Cited on pages 13, 42, 62, 63, and 64.)

Fan, J. (2005): "A Selective Overview of Nonparametric Methods in Financial Econometrics," *Statistical Science*, 20, 317–337. (Cited on page 4.)

Fernández-Villaverde, J., S. Hurtado, and G. Nuño (2020): "Financial Frictions and the Wealth Distribution," CESifo Working Paper Series 8482, CESifo. (Cited on pages 2 and 4.)

Fernández-Villaverde, J., J. F. Rubio-Ramírez, T. J. Sargent, and M. W. Watson (2007): "ABCs (and Ds) of Understanding VARs," *American Economic Review*, 97, 1021–1026. (Cited on page 12.)

Fernández-Villaverde, J. and J. F. Rubio-Ramírez (2007): "Estimating Macroeconomic Models: A Likelihood Approach," *Review of Economic Studies*, 74, 1059–1087. (Cited on page 3.)

Fernández-Villaverde, J. F., J. Rubio-Ramírez, and F. Schorfheide (2016): "Solution and Estimation Methods for DSGE Models," in *Handbook of Macroeconomics*, ed. by J. B. Taylor and H. Uhlig, Elsevier, vol. 2 of *Handbook of Macroeconomics*, 527–724. (Cited on pages 3, 13, and 37.)

Florens-Zmirou, D. (1993): "On Estimating the Diffusion Coefficient from Discrete Observations," *Journal of Applied Probability*, 30, 790–804. (Cited on page 4.)

Furlanetto, F. and M. Seneca (2014): "Investment Shocks and Consumption," *European Economic Review*, 66, 111–126. (Cited on page 24.)

Gantmacher, F. R. (1959): *The Theory of Matrices*, Chelsea, New York. (Cited on page 16.)

Geweke, J. (1978): "Temporal Aggregation in the Multiple Regression Model," *Econometrica*, 46, 643–661. (Cited on page 9.)

Ghysels, E. and J. H. Wright (2009): "Forecasting Professional Forecasters," *Journal of Business & Economic Statistics*, 27, 504–516. (Cited on page 42.)

Golub, G. H. and C. F. Van Loan (2013): *Matrix Computations*, Johns Hopkins University Press, 4th ed. (Cited on page 70.)

Hamerle, A., W. Nagl, and H. Singer (1991): "Problems with the Estimation of Stochastic Differential Equations using Structural Equations Models," *Journal of Mathematical Sociology*, 16, 201–220. (Cited on page 17.)

Hamilton, J. (1994): *Time Series Analysis*, Princeton University Press. (Cited on page 13.)

Hannan, E. J. and M. Deistler (2012): *The Statistical Theory of Linear Systems*, SIAM. (Cited on page 18.)

Hansen, G. D. (1985): "Indivisible Labor and the Business Cycle," *Journal of Monetary Economics*, 16, 309–327. (Cited on pages 5, 23, and 27.)

——— (1997): "Technical Progress and Aggregate Fluctuations," *Journal of Economic Dynamics and Control*, 21, 1005–1023. (Cited on pages 27 and 38.)

Hansen, L. P. and E. C. Prescott (1995): "Recursive Methods for Computing Equilibria of Business Cycle Models," in *Frontiers of Business Cycle Research*, ed. by T. F. Cooley, Princeton, New Jersey: Princeton University Press. (Cited on page 27.)

Hansen, L. P. and T. J. Sargent (1983): "The Dimensionality of the Aliasing Aroblem in Models with Rational Spectral Densities," *Econometrica*, 51, 377–387. (Cited on page 17.)

——— (1991): "Identification of Continuous Time Rational Expectations Models from Discrete Time Data," in *Rational Expectations Econometrics*, ed. by L. P. Hansen and T. J. Sargent, Westview Press, chap. 9, 219–235. (Cited on page 17.)

Hansen, L. P. and J. A. Scheinkman (1995): "Back to the Future: Generating Moment Implications for Continuous-Time Markov Processes," *Econometrica*, 63, 767–804. (Cited on page 4.)

Harvey, A. C. (1990): *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press. (Cited on pages 10, 12, 13, 42, and 62.)

Harvey, A. C. and R. G. Pierse (1984): "Estimating Missing Observations in Economic Time Series," *Journal of the American Statistical Association*, 79, 125–131. (Cited on page 42.)

Harvey, A. C. and J. H. Stock (1985): "The Estimation of Higher-Order Continuous Time Autoregressive Models," *Econometric Theory*, 1, 97–117. (Cited on page 2.)

Ingram, B. F., N. R. Kocherlakota, and N. E. Savin (1994): "Explaining Business Cycles: A Multiple-Shock Approach," *Journal of Monetary Economics*, 34, 415–428. (Cited on page 13.)

Ireland, P. (2001): "Technology Shocks and the Business Cycle: An Empirical Investigation," *Journal of Economic Dynamics and Control*, 25, 703–719. (Cited on page 38.)

Ireland, P. N. (2004): "A Method for Taking Models to the Data," *Journal of Economic Dynamics and Control*, 28, 1205–1226. (Cited on pages 10, 38, 39, and 41.)

Iskrev, N. (2010): "Local Identification in DSGE Models," *Journal of Monetary Economics*, 57, 189–202. (Cited on pages 4 and 14.)

Itskhoki, O. and B. Moll (2019): "Optimal Development Policies With Financial Frictions," *Econometrica*, 87, 139–173. (Cited on page 2.)

Jewitt, G. and J. R. McCrorie (2005): "Computing Estimates of Continuous Time Macroeconometric Models on the Basis of Discrete Data," *Computational Statistics & Data Analysis*, 49, 397–416. (Cited on page 56.)

Jiang, G. J. and J. L. Knight (1997): "A Nonparametric Approach to the Estimation of Diffusion Processes, With an Application to a Short-Term Interest Rate Model," *Econometric Theory*, 13, 615–645. (Cited on page 4.)

Jones, R. H. (1981): "Fitting a Continuous Time Autoregression to Discrete Data," in *Applied Time Series Analysis II*, ed. by D. Findlay, New York: Academic Press, 651–682. (Cited on page 2.)

Kaplan, G., B. Moll, and G. L. Violante (2018): "Monetary Policy According to HANK," *American Economic Review*, 108, 697–743. (Cited on page 2.)

Kessler, M. and A. Rahbek (2004): "Identification and Inference for Multivariate Cointegrated and Ergodic Gaussian Diffusions," *Statistical Inference for Stochastic Processes*, 7, 137–151. (Cited on page 4.)

King, R. G., C. I. Plosser, and S. T. Rebelo (1988a): "Production, Growth and Business Cycles : I. The Basic Neoclassical Model," *Journal of Monetary Economics*, 21, 195–232. (Cited on page 38.)

——— (1988b): "Production, Growth and Business Cycles : II. New Directions," *Journal of Monetary Economics*, 21, 309–341. (Cited on page 38.)

Komunjer, I. and S. Ng (2011): "Dynamic Identification of Dynamic Stochastic General Equilibrium Models," *Econometrica*, 79, 1995–2032. (Cited on pages 4, 14, 18, 19, 28, and 44.)

Kuzin, V., M. Marcellino, and C. Schumacher (2011): "MIDAS vs. Mixed-Frequency VAR: Nowcasting GDP in the Euro Area," *International Journal of Forecasting*, 27, 529–542. (Cited on page 42.)

Liemen, M. O. and O. Posch (2022): "FTPL and the Maturity Structure of Government Debt in the New Keynesian Model," CESifo Working Paper Series 9840, CESifo. (Cited on page 2.)

Lo, A. W. (1988): "Maximum Likelihood Estimation of Generalized Itô Processes with Discretely Sampled Data," *Econometric Theory*, 4, 231–247. (Cited on pages 4 and 30.)

Lütkepohl, H. (2005): *New Introduction to Multiple Time Series Analysis*, Springer Books, Springer. (Cited on page 36.)

Malley, J. and U. Woitek (2010): "Technology Shocks and Aggregate Fluctuations in an Estimated Hybrid RBC Model," *Journal of Economic Dynamics and Control*, 34, 1214–1232. (Cited on page 38.)

Mariano, R. S. and Y. Murasawa (2003): "A New Coincident Index of Business Cycles based on Monthly and Quarterly Series," *Journal of Applied Econometrics*, 18, 427–443. (Cited on page 42.)

McCrorie, J. R. (2000): "Deriving the Exact Discrete Analog of a Continuous Time System," *Econometric Theory*, 16, 998–1015. (Cited on page 8.)

——— (2003): "The Problem of Aliasing in Identifying Finite Parameter Continuous Time Stochastic Models," *Acta Applicandae Mathematicae*, 79, 9–16. (Cited on pages 4 and 71.)

——— (2009): "Estimating Continuous-Time Models On The Basis Of Discrete Data Via An Exact Discrete Analog," *Econometric Theory*, 25, 1120–1137. (Cited on pages 3 and 17.)

Merton, R. C. (1980): "On Estimating the Expected Return on the Market: An Exploratory Investigation," *Journal of Financial Economics*, 8, 323–361. (Cited on page 29.)

Newey, W. K. and D. McFadden (1994): "Large sample estimation and hypothesis testing," *Handbook of econometrics*, 4, 2111–2245. (Cited on pages 13 and 14.)

Nowman, K. B. (1993): "Finite-Sample Properties of the Gaussian Estimation of an Open Higher-Order Continuous-Time Dynamic Model with Mixed Stock and Flow Data," in *Continuous-Time Econometrics: Theory and Applications*, ed. by G. Gandolfo, Springer, Dordrecht, chap. 5, 93–116. (Cited on page 21.)

——— (1997): "Gaussian Estimation of Single-Factor Continuous Time Models of the Term Structure of Interest Rates," *Journal of Finance*, 52, 1695–1706. (Cited on page 9.)

Parra-Alvarez, J. C. (2018): "A Comparison of Numerical Methods for the Solution of Continuous-time DSGE Models," *Macroeconomic Dynamics*, 22, 1555 – 1583. (Cited on page 2.)

Parra-Alvarez, J. C., H. Polattimur, and O. Posch (2021): "Risk Matters: Breaking Certainty Equivalence in Linear Models," *Journal of Economics Dynamics and Control*, 133. (Cited on pages 2, 26, and 80.)

Pfeifer, J. (2020): "A Guide to Specifying Observation Equations for the Estimation of DSGE Models," Tech. rep., mimeo. (Cited on page 9.)

Phillips, A. W. (1959): "The Estimation of Parameters in Systems of Stochastic Differential Equations," *Biometrika*, 46, 67–76. (Cited on pages 3 and 8.)

Phillips, P. C. B. (1973): "The Problem of Identification in Finite Parameter Continuous Time Models," *Journal of Econometrics*, 1, 351–362. (Cited on pages 3, 4, 8, 15, 17, and 30.)

Phillips, P. C. B. and J. Yu (2005): "Jackknifing Bond Option Prices," *Review of Financial Studies*, 18, 707–742. (Cited on page 30.)

——— (2009): "Maximum Likelihood and Gaussian Estimation of Continuous Time Models in Finance," in *Handbook of financial time series*, Springer, Handbook of Financial Time Series, 497–530. (Cited on page 4.)

Posch, O. (2009): "Structural Estimation of Jump-Diffusion Processes in Macroeconomics," *Journal of Econometrics*, 153, 196–210. (Cited on pages 2 and 4.)

——— (2011): "Risk Premia in General Equilibrium," *Journal of Economic Dynamics and Control*, 35, 1557–1576. (Cited on page 2.)

——— (2020): "Resurrecting the New-Keynesian Model: (Un)conventional Policy and the Taylor rule," CESifo Working Paper Series 6925, CESifo. (Cited on page 2.)

Posch, O. and T. Trimborn (2013): "Numerical Solution of Dynamic Equilibrium Models under Poisson Uncertainty," *Journal of Economic Dynamics and Control*, 37, 2602–2622. (Cited on page 2.)

Qu, Z. and D. Tkachenko (2012): "Identification and Frequency Domain Quasi-Maximum Likelihood Estimation of Linearized Dynamic Stochastic General Equilibrium Models," *Quantitative Economics*, 3, 95–132. (Cited on pages 4 and 14.)

——— (2017): "Global Identification in DSGE Models Allowing for Indeterminacy," *The Review of Economic Studies*, 84, 1306–1345. (Cited on pages 4 and 14.)

Rothenberg, T. J. (1971): "Identification in Parametric Models," *Econometrica*, 39, pp. 577–591. (Cited on page 14.)

Ruge-Murcia, F. J. (2007): "Methods to Estimate Dynamic Stochastic General Equilibrium Models," *Journal of Economic Dynamics and Control*, 31, 2599–2636. (Cited on pages 13 and 32.)

Sargent, T. J. (1989): "Two Models of Measurements and the Investment Accelerator," *Journal of Political Economy*, 97, 251–287. (Cited on page 9.)

Shapiro, M. D. and M. W. Watson (1988): "Sources of Business Cycle Fluctuations," *NBER Macroeconomics Annual*, 3, 111–148. (Cited on page 5.)

Shapiro, S. S. and M. B. Wilk (1965): "An Analysis of Variance Test for Normality (Complete Samples)," *Biometrika*, 52, 591–611. (Cited on page 90.)

Simon, H. A. (1956): "Rational Choice and the Structure of the Environment." *Psychological review*, 63, 129. (Cited on page 80.)

Sims, C. A. (1986): "Are Forecasting Models Usable for Policy Analysis?" *Quarterly Review*, 10, 2–16. (Cited on page 5.)

——— (2002): "Solving Linear Rational Expectations Models," *Computational Economics*, 20, 1–20. (Cited on pages 26 and 80.)

Sørensen, M. (1997): "Estimating Functions for Discretely Observed Diffusions: A Review," *Lecture Notes-Monograph Series*, 32, 305–325. (Cited on page 4.)

Tang, C. Y. and S. X. Chen (2009): "Parameter Estimation and Bias Correction for Diffusion Processes," *Journal of Econometrics*, 149, 65–81. (Cited on pages 29 and 30.)

Theil, H. (1957): "A Note on Certainty Equivalence in Dynamic Planning," *Econometrica: Journal of the Econometric Society*, 346–349. (Cited on page 80.)

Thornton, M. A. and M. J. Chambers (2016): "The Exact Discretisation of CARMA Models with Applications in Finance," *Journal of Empirical Finance*, 38, 739–761. (Cited on page 30.)

Van Loan, C. (1978): "Computing Integrals Involving the Matrix Exponential," *IEEE Transactions on Automatic Control*, 23, 395–404. (Cited on pages 9, 12, 56, and 71.)

Wang, X., P. C. B. Phillips, and J. Yu (2011): "Bias in Estimating Multivariate and Univariate Diffusions," *Journal of Econometrics*, 161, 228–245. (Cited on pages 30 and 31.)

Yu, J. (2012): "Bias in the Estimation of the Mean Reversion Parameter in Continuous Time Models," *Journal of Econometrics*, 169, 114–122. (Cited on page 29.)

Yu, J. and P. C. B. Phillips (2001): "A Gaussian Approach for Continuous Time Models of the Short-term Interest Rate," *The Econometrics Journal*, 4, 210–224. (Cited on page 9.)

Zadrozny, P. (1988): "Gaussian Likelihood of Continuous-Time ARMAX Models when Data are Stocks and Flows at Different Frequencies," *Econometric Theory*, 4, 108–124. (Cited on pages 2 and 42.)

Zadrozny, P. A. (1990): "Forecasting U.S. GNP at Monthly Intervals with an Estimated Bivariate Time Series Model," *Economic Review*, 2–15. (Cited on page 42.)

Appendix

A. Matrix computations in the exact discrete model

A.1 Computation of $\mathbf{A}_h$

From Proposition 1, the $n_x \times n_x$ matrix $\mathbf{A}_h(\boldsymbol{\theta})$ is defined as

$$\mathbf{A}_h(\boldsymbol{\theta}) = \exp\left(\mathbf{A}(\boldsymbol{\theta})h\right) = \mathbf{I} + \mathbf{A}(\boldsymbol{\theta})h + \tfrac{1}{2}\mathbf{A}^2(\boldsymbol{\theta})h^2 + \tfrac{1}{3!}\mathbf{A}^3(\boldsymbol{\theta})h^3 + \ldots,$$

where $\mathbf{A}^j(\boldsymbol{\theta})$ indicates right multiplication of j copies of the $n_x \times n_x$ matrix $\mathbf{A}(\boldsymbol{\theta})$.

Assume $\mathbf{A}(\boldsymbol{\theta})$ is diagonalizable.[20] Then, $\mathbf{A}(\boldsymbol{\theta})$ can be factorized as

$$\mathbf{A}(\boldsymbol{\theta}) = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{-1},$$

where $\mathbf{V}$ is a square matrix whose columns correspond to the eigenvectors of $\mathbf{A}(\boldsymbol{\theta})$ and $\boldsymbol{\Lambda}$ is a diagonal matrix whose elements are the corresponding eigenvalues,

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_{n_x} \end{bmatrix}.$$

Therefore, the exponential matrix $\exp(\mathbf{A}(\boldsymbol{\theta})h)$ can be computed as

$$\exp(\mathbf{A}(\boldsymbol{\theta})h) = \mathbf{V}\exp(\boldsymbol{\Lambda}h)\mathbf{V}^{-1} = \mathbf{V}\begin{bmatrix} e^{\lambda_1 h} & 0 & \ldots & 0 \\ 0 & e^{\lambda_2 h} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & e^{\lambda_{n_x} h} \end{bmatrix}\mathbf{V}^{-1}.$$

A.2 Computation of $\boldsymbol{\Sigma}_{\boldsymbol{\eta},h}(\boldsymbol{\theta})$

Here, we show how to implement the matrix decomposition method in Van Loan (1978, Theorem 1) to compute the variance-covariance matrix of the reduced-form innovation $\boldsymbol{\Sigma}_{\boldsymbol{\eta}^s,h}(\boldsymbol{\theta})$ in the S-SSR model (2.8)-(2.9), and $\boldsymbol{\Sigma}_{\boldsymbol{\eta},h}(\boldsymbol{\theta})$ in the F-SSR model (2.12)-(2.13). Jewitt and

[20] The matrix $\mathbf{A}(\boldsymbol{\theta})$ is diagonalizable if: i) it has n_x distinct eigenvalues; or ii) the sum of the geometric multiplicities of its eigenvalues is equal to n_x; or iii) the sum of the algebraic multiplicities of its eigenvalues is equal to n_x, and for each eigenvalue, the geometric multiplicity equals the algebraic multiplicity.

McCrorie (2005) provide a comparison of different computational methods to implement this decomposition.

A.2.1 The S-SSR model: measurements are stock variables

Consider the continuous-time model for the state vector $\mathbf{x}(t)$ in (2.1) where $\mathbf{A}(\boldsymbol{\theta})$ is the stable $n_x \times n_x$ drift matrix, and $\mathbf{B}(\boldsymbol{\theta})$ is the $n_x \times n_w$ diffusion matrix with associated $n_x \times n_x$ instantaneous variance-covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \mathbf{B}(\boldsymbol{\theta})\mathbf{B}(\boldsymbol{\theta})^\top$. Then, when the observables are sampled as stocks, we define the $2n_x \times 2n_x$ block triangular matrix

$$\boldsymbol{\Xi}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{A}(\boldsymbol{\theta}) & \boldsymbol{\Sigma}(\boldsymbol{\theta}) \\ \mathbf{0} & -\mathbf{A}(\boldsymbol{\theta})^\top \end{bmatrix},$$

with exponential

$$\exp(\boldsymbol{\Xi}(\boldsymbol{\theta})h) = \begin{bmatrix} \mathbf{A}_h(\boldsymbol{\theta}) & \mathbf{M}_h(\boldsymbol{\theta}) \\ \mathbf{0} & (\mathbf{A}_h(\boldsymbol{\theta})^{-1})^\top \end{bmatrix},$$

where $\mathbf{A}_h(\boldsymbol{\theta}) = \exp(\mathbf{A}(\boldsymbol{\theta})h)$ and $\mathbf{M}_h(\boldsymbol{\theta}) = \left(\int_0^h \mathbf{A}_s(\boldsymbol{\theta})\boldsymbol{\Sigma}(\boldsymbol{\theta})\mathbf{A}_s(\boldsymbol{\theta})^\top \mathrm{d}s\right) \exp\left(-\mathbf{A}(\boldsymbol{\theta})^\top h\right)$, and where we used the fact that $\mathbf{A}_h(\boldsymbol{\theta})^\top = \exp(\mathbf{A}(\boldsymbol{\theta})^\top h)$. Then, the $n_x \times n_x$ variance-covariance matrix of the reduced-form innovations in the transition equation (2.8) is obtained as

$$\boldsymbol{\Sigma}_{\boldsymbol{\eta}^s,h}(\boldsymbol{\theta}) = \mathbf{M}_h(\boldsymbol{\theta})\mathbf{A}_h(\boldsymbol{\theta})^\top, \tag{A.1}$$

which can be alternatively written as

$$\begin{aligned} \mathrm{vech}(\boldsymbol{\Sigma}_{\boldsymbol{\eta}^s,h}(\boldsymbol{\theta})) &= (\mathbf{A}_h(\boldsymbol{\theta}) \otimes \mathbf{I})\mathrm{vech}\big(\mathbf{M}_h(\boldsymbol{\theta})\big) \\ &= (\mathbf{A}_h(\boldsymbol{\theta}) \otimes \mathbf{I})(\mathbf{A}_h(\boldsymbol{\theta})^{-1} \otimes \mathbf{I})\mathrm{vech}\left[\left(\int_0^h \mathbf{A}_s(\boldsymbol{\theta})\boldsymbol{\Sigma}(\boldsymbol{\theta})\mathbf{A}_s(\boldsymbol{\theta})^\top \mathrm{d}s\right)\right] \\ &= \left[\left(\int_0^h \mathbf{A}_s(\boldsymbol{\theta}) \otimes \mathbf{A}_s(\boldsymbol{\theta})\mathrm{d}s\right)\right] \mathrm{vech}(\boldsymbol{\Sigma}(\boldsymbol{\theta})) \\ &= \Big(\mathbf{A}(\boldsymbol{\theta}) \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{A}(\boldsymbol{\theta})\Big)^{-1}\Big(\mathbf{A}_h(\boldsymbol{\theta}) \otimes \mathbf{A}_h(\boldsymbol{\theta}) - \mathbf{I}\Big)\mathrm{vech}\big(\boldsymbol{\Sigma}(\boldsymbol{\theta})\big), \end{aligned} \tag{A.2}$$

where vech is the half-vectorization operator, and $\otimes$ denotes the right-hand Kronecker product. The second equality uses $\exp(-\mathbf{A}(\boldsymbol{\theta})h) = \mathbf{A}_h(\boldsymbol{\theta})^{-1}$, the third equality uses the fact that $(\mathbf{A}_h(\boldsymbol{\theta})^{-1} \otimes \mathbf{I}) = (\mathbf{A}_h(\boldsymbol{\theta}) \otimes \mathbf{I})^{-1}$, whereas the last equality computes $\left(\int_0^h \mathbf{A}_s(\boldsymbol{\theta}) \otimes \mathbf{A}_s(\boldsymbol{\theta})\mathrm{d}s\right)$ in closed form as a linear function of $\mathbf{A}(\boldsymbol{\theta})$ and $\mathbf{A}_h(\boldsymbol{\theta})$. Notice that $\mathbf{A}_h(\boldsymbol{\theta}) \otimes \mathbf{A}_h(\boldsymbol{\theta}) = \exp\left((\mathbf{A}(\boldsymbol{\theta}) \oplus \mathbf{A}(\boldsymbol{\theta}))h\right)$.

A.2.2 The F-SSR model: measurements are flow variables

A similar approach can be used to compute $\mathbf{\Sigma}_{\boldsymbol{\eta}^s\boldsymbol{\eta}^f,h}(\boldsymbol{\theta})$ in (2.14) and $\mathbf{\Sigma}_{\boldsymbol{\eta}^f,h}(\boldsymbol{\theta})$ in (2.15) when the observables are measured as flows. In particular, define the augmented upper-triangular matrix

$$\mathbf{\Xi}^f(\boldsymbol{\theta}) = \begin{bmatrix} -\mathbf{A}(\boldsymbol{\theta}) & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{A}(\boldsymbol{\theta}) & \mathbf{\Sigma}(\boldsymbol{\theta}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}(\boldsymbol{\theta})^\top & \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

with exponential

$$\exp(\mathbf{\Xi}^f(\boldsymbol{\theta})h) = \begin{bmatrix} \mathbf{F}_{1,h}(\boldsymbol{\theta}) & \mathbf{G}_{1,h}(\boldsymbol{\theta}) & \mathbf{H}_{1,h}(\boldsymbol{\theta}) & \mathbf{K}_{1,,h}(\boldsymbol{\theta}) \\ \mathbf{0} & \mathbf{F}_{2,h}(\boldsymbol{\theta}) & \mathbf{G}_{2,h}(\boldsymbol{\theta}) & \mathbf{H}_{2,h}(\boldsymbol{\theta}) \\ \mathbf{0} & \mathbf{0} & \mathbf{F}_{3,h}(\boldsymbol{\theta}) & \mathbf{G}_{3,h}(\boldsymbol{\theta}) \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{F}_{4,h}(\boldsymbol{\theta}) \end{bmatrix},$$

where every block is n_x-by-n_x. Then,

$$\mathbf{\Sigma}_{\boldsymbol{\eta}^s\boldsymbol{\eta}^f,h}(\boldsymbol{\theta}) = \left\{\mathbf{F}_{3,h}(\boldsymbol{\theta})^\top \mathbf{H}_{2,h}(\boldsymbol{\theta})\right\} \mathbf{C}(\boldsymbol{\theta})^\top,$$

$$\mathbf{\Sigma}_{\boldsymbol{\eta}^f,h}(\boldsymbol{\theta}) = \mathbf{C}(\boldsymbol{\theta}) \left\{ \left[\mathbf{F}_{3,h}(\boldsymbol{\theta})^\top \mathbf{K}_{1,h}(\boldsymbol{\theta})\right] + \left[\mathbf{F}_{3,h}(\boldsymbol{\theta})^\top \mathbf{K}_{1,h}(\boldsymbol{\theta})\right]^\top \right\} \mathbf{C}(\boldsymbol{\theta})^\top,$$

and $\mathbf{A}(\boldsymbol{\theta})^{-1}(\mathbf{A}_h(\boldsymbol{\theta}) - \mathbf{I}) = \int_0^h \exp(\mathbf{A}(\boldsymbol{\theta})s)\mathrm{d}s = \mathbf{G}_{3,h}(\boldsymbol{\theta})^\top$.

A.3 An example with aliases of a matrix A

The example illustrated in Figure 1 assumes that the dynamics of the vector $\mathbf{X}(t) = [x_1(t), x_2(t)]^\top$ is given by

$$\mathrm{d}\mathbf{X}^{(0)}(t) = \mathbf{A}_0\mathbf{X}^{(0)}(t)\mathrm{d}t, \quad \text{with} \quad \mathbf{A}_0 = \begin{bmatrix} -0.4 & 16 \\ -0.4 & -0.4 \end{bmatrix}.$$

There are two Jordan blocks that are scalars and the solution to the mapping $\mathbf{A}_h(\boldsymbol{\theta}_0) = \exp(\mathbf{A}(\boldsymbol{\theta}_0)h)$ is therefore not uniquely defined. The Jordan blocks of $\mathbf{A}_0$ coincide with its eigenvalues, i.e.,

$$\mathbf{\Lambda} = \mathbf{J} = \begin{bmatrix} -0.4 + 2.5298i & 0 \\ 0 & -0.4 - 2.5298i \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix},$$

where $i = \sqrt{-1}$. The Jordan blocks $J_1 = -0.4 + 2.5298i$ and $J_2 = -0.4 - 2.5298i$ are associated with the eigenvalues $\lambda_1 = -0.4 + 2.5298i$ and $\lambda_2 = -0.4 - 2.5298i$, and have algebraic and geometric multiplicity 1. The matrix of eigenvectors is given by

$$\mathbf{V} = \begin{bmatrix} 0.9877 & 0.9877 \\ 0.0000 + 0.1562i & 0.0000 - 0.1562i \end{bmatrix},$$

and the transformation matrix of the Jordan canonical form is given by

$$\mathbf{S} = \begin{bmatrix} 0.0000 - 6.3246i & 0.0000 + 6.3246i \\ 1.0000 & 1.0000 \end{bmatrix}.$$

Matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ (aliases of $\mathbf{A}_0$) are constructed using

$$\mathbf{A}_1 = \mathbf{A}_0 + 2\pi i \cdot \mathbf{V} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \mathbf{V}^{-1},$$
$$\mathbf{A}_2 = \mathbf{A}_0 + 2\pi i \cdot \mathbf{V} \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \mathbf{V}^{-1}.$$

Then, Figure 1 plots the simulated values for x_1 and x_2 with dynamics

$$\begin{aligned} A_0: &\quad \mathrm{d}\mathbf{X}^{(0)}(t) = \mathbf{A}_0 \mathbf{X}^{(0)}(t)\mathrm{d}t, \\ A_1: &\quad \mathrm{d}\mathbf{X}^{(1)}(t) = \mathbf{A}_1 \mathbf{X}^{(1)}(t)\mathrm{d}t, \\ A_2: &\quad \mathrm{d}\mathbf{X}^{(2)}(t) = \mathbf{A}_2 \mathbf{X}^{(2)}(t)\mathrm{d}t. \end{aligned}$$

B. Kalman Recursions and the Likelihood Function

B.1 The ABCD representation

It is convenient to work with a state space which does not make any explicit assumption on the nature of the measurements. Namely, the ABCD representation of (2.16)-(2.17), restated here for completeness,

$$\begin{aligned}\mathbf{x}_{\tau+1} &= \mathsf{A}(\boldsymbol{\theta})\mathbf{x}_\tau + \mathsf{B}(\boldsymbol{\theta})\boldsymbol{\epsilon}_{\tau+1}\\ \mathbf{y}_{\tau+1} &= \mathsf{C}(\boldsymbol{\theta})\mathbf{x}_\tau + \mathsf{D}(\boldsymbol{\theta})\boldsymbol{\epsilon}_{\tau+1}.\end{aligned}$$

According to whether the measurements are either stocks or flows or else whether the EM approximation is used instead of the EDM, the entries of the matrices change. Suppose $n_\varepsilon = \{0, n_y\}$ where n_ε defines the number of measurement errors associate to the vector of measurements $\mathbf{y}_\tau$. For clarity, let us be explicit on the dimensions of matrices $\mathbf{I}$ and $\mathbf{0}$, such that $\mathbf{I}_n$ and $\mathbf{0}_{n\times m}$ indicate the n-dimensional identity matrix and the n-by-m null matrix, respectively. Let us describe the different elements of the system matrices according to the type of data (stock or flow) and/or the discretization scheme (EDM or EM):

The S-SSR case.

$$\mathsf{A}(\boldsymbol{\theta}) := \mathbf{A}_h(\boldsymbol{\theta}), \qquad \mathsf{B}(\boldsymbol{\theta}) := [\mathbf{I}_{n_x}, \mathbf{0}_{n_x\times n_\varepsilon}], \qquad \widetilde{\mathsf{C}}(\boldsymbol{\theta}) := \mathbf{C}(\boldsymbol{\theta}),$$

$$\boldsymbol{\Sigma}_\epsilon(\boldsymbol{\theta}) := \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{\eta}^s,h}(\boldsymbol{\theta}) & \mathbf{0}_{n_y\times n_x}^\top \\ \mathbf{0}_{n_y\times n_x} & \mathbf{R}\end{bmatrix}, \qquad \mathsf{C}(\boldsymbol{\theta}) := \widetilde{\mathsf{C}}(\boldsymbol{\theta})\mathsf{A}(\boldsymbol{\theta}), \qquad \mathsf{D}(\boldsymbol{\theta}) := [\widetilde{\mathsf{C}}(\boldsymbol{\theta}), \mathbf{I}_{n_\varepsilon}];$$

The F-SSR case.

$$\mathsf{A}(\boldsymbol{\theta}) := \begin{bmatrix} \mathbf{A}_h(\boldsymbol{\theta}) & \mathbf{0}_{n_x\times n_y} \\ \mathbf{C}(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})^{-1}(\mathbf{A}_h(\boldsymbol{\theta}) - \mathbf{I}_{n_x}) & \mathbf{0}_{n_y\times n_y}\end{bmatrix}, \qquad \begin{aligned}\mathsf{B}(\boldsymbol{\theta}) &:= [\mathbf{I}_{(n_x+n_y)}, \mathbf{0}_{(n_x+n_y)\times n_\varepsilon}],\\ \widetilde{\mathsf{C}}(\boldsymbol{\theta}) &:= [\mathbf{0}_{n_y\times n_x}, \mathbf{I}_{n_y}],\end{aligned}$$

$$\boldsymbol{\Sigma}_\epsilon(\boldsymbol{\theta}) := \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{\eta},h}(\boldsymbol{\theta}) & \mathbf{0}_{n_y\times(n_x+n_y)}^\top \\ \mathbf{0}_{n_y\times(n_x+n_y)} & \mathbf{R}\end{bmatrix}, \qquad \mathsf{C}(\boldsymbol{\theta}) := \widetilde{\mathsf{C}}(\boldsymbol{\theta})\mathsf{A}(\boldsymbol{\theta}), \qquad \mathsf{D}(\boldsymbol{\theta}) := [\widetilde{\mathsf{C}}(\boldsymbol{\theta}), \mathbf{I}_{n_\varepsilon}];$$

The EM-SSR case.

$$\mathsf{A}(\boldsymbol{\theta}) := \mathbf{I}_{n_x} + \mathbf{A}(\boldsymbol{\theta})h, \qquad \mathsf{B}(\boldsymbol{\theta}) := [\mathbf{I}_{n_x}, \mathbf{0}_{n_x\times n_\varepsilon}], \qquad \widetilde{\mathsf{C}}(\boldsymbol{\theta}) := \mathbf{C}(\boldsymbol{\theta}),$$

$$\mathbf{\Sigma}_{\epsilon}(\boldsymbol{\theta}) := \begin{bmatrix} h\mathbf{B}(\boldsymbol{\theta})\mathbf{B}(\boldsymbol{\theta})^{\top} & \mathbf{0}_{n_y \times n_x}^{\top} \\ \mathbf{0}_{n_y \times n_x} & \mathbf{R} \end{bmatrix}, \qquad \mathsf{C}(\boldsymbol{\theta}) := \widetilde{\mathsf{C}}(\boldsymbol{\theta})\mathsf{A}(\boldsymbol{\theta}), \qquad \mathsf{D}(\boldsymbol{\theta}) := [\widetilde{\mathsf{C}}(\boldsymbol{\theta}), \mathbf{I}_{n_\varepsilon}].$$

B.2 Kalman filter

Let $\mathbf{y}^{\tau-1} = \{\mathbf{y}_1, \ldots, \mathbf{y}_{\tau-1}\}$ denote the history of measurements up to time $t_{\tau-1}$. Additionally, let $\mathbf{x}_{\tau|\tau-1} = \mathbb{E}\left[\mathbf{x}_\tau | \mathbf{y}^{\tau-1}\right]$ denote the forecast of the state vector conditional on the information available at the time $t_{\tau-1}$, and $\mathbf{P}_{\tau|\tau-1} = \mathbb{E}\left[\left(\mathbf{x}_\tau - \mathbf{x}_{\tau|\tau-1}\right)\left(\mathbf{x}_\tau - \mathbf{x}_{\tau|\tau-1}\right)^{\top}\right]$ the corresponding forecast error covariance matrix.[21] Similarly, let $\mathbf{y}_{\tau|\tau-1} = \mathbb{E}\left[\mathbf{y}_\tau | \mathbf{y}^{\tau-1}\right]$ denote the forecast of the control variables conditional on past information, and

$$\mathbf{\Omega}_{\tau|\tau-1} = \mathbb{E}\left[\left(\mathbf{y}_\tau - \mathbf{y}_{\tau|\tau-1}\right)\left(\mathbf{y}_\tau - \mathbf{y}_{\tau|\tau-1}\right)^{\top}\right]$$

its associated forecast error covariance matrix. By exploiting the linearity of (2.16)-(2.17), the forecast of the state variables, and their associated variance-covariance, are

$$\mathbf{x}_{\tau|\tau-1} = \mathsf{A}(\boldsymbol{\theta})\mathbf{x}_{\tau-1|\tau-1}, \tag{B.4}$$

$$\mathbf{P}_{\tau|\tau-1} = \mathsf{A}(\boldsymbol{\theta})\mathbf{P}_{\tau-1|\tau-1}\mathsf{A}(\boldsymbol{\theta})^{\top} + \mathsf{B}(\boldsymbol{\theta})\mathbf{\Sigma}_{\epsilon}(\boldsymbol{\theta})\mathsf{B}(\boldsymbol{\theta})^{\top}, \tag{B.5}$$

given some initial conditions $\mathbf{x}_0$ and $\mathbf{P}_0$. Since the state vector is stationary, we use as initial values its unconditional mean $\mathbf{x}_0 = \mathbf{x}_{1|0} = \mathbb{E}\left[\mathbf{x}_1\right] = \mathbf{0}$, and its unconditional covariance matrix

$$\text{vec}\left(\mathbf{P}_0\right) = \text{vec}\left(\mathbf{P}_{1|0}\right) = \text{vec}\left(\mathbb{E}\left[\mathbf{x}_1\mathbf{x}_1^{\top}\right]\right) = \left[\mathbf{I}_{n_x^2} - \left(\mathsf{A}(\boldsymbol{\theta}) \otimes \mathsf{A}(\boldsymbol{\theta})\right)\right]^{-1}\left[\mathsf{B}(\boldsymbol{\theta}) \otimes \mathsf{B}(\boldsymbol{\theta})\right]\text{vec}\left(\mathbf{\Sigma}_{\epsilon}(\boldsymbol{\theta})\right),$$

where vec is the vectorization operator. Given the predictions for the state variables, the Kalman filter recursively computes the one-step-ahead forecast error of the control variables and associated variance-covariance matrix

$$\boldsymbol{\omega}_{\tau|\tau-1} = \mathbf{y}_\tau - \mathbf{y}_{\tau|\tau-1} = \mathbf{y}_\tau - \mathsf{C}(\boldsymbol{\theta})\mathbf{x}_{\tau-1|\tau-1} \tag{B.6}$$

$$\mathbf{\Omega}_{\tau|\tau-1} = \tilde{\mathbf{C}}(\boldsymbol{\theta})\mathbf{P}_{\tau|\tau-1}\tilde{\mathbf{C}}(\boldsymbol{\theta})^{\top} + \mathsf{D}(\boldsymbol{\theta})\mathbf{\Sigma}_{\epsilon}(\boldsymbol{\theta})\mathsf{D}(\boldsymbol{\theta})^{\top}. \tag{B.7}$$

Using the information above, we update the state variables according to

$$\mathbf{x}_{\tau|\tau} = \mathbf{x}_{\tau|\tau-1} + \mathbf{K}_{\tau|\tau-1}\boldsymbol{\omega}_{\tau|\tau-1} \tag{B.8}$$

$$\mathbf{P}_{\tau|\tau} = \mathbf{P}_{\tau|\tau-1} - \mathbf{K}_{\tau|\tau-1}\mathbf{\Omega}_{\tau|\tau-1}\mathbf{K}_{\tau|\tau-1}^{\top}, \tag{B.9}$$

[21] Here, $\mathbf{x}_\tau$ refers to the n_x dimensional state vector in (2.8) for the S-SSR model or in (4.13) for the EM-SSR model, or the $(n_x + n_y)$ dimensional vector in (2.12) for the F-SSR model.

where

$$\mathbf{K}_{\tau|\tau-1} = \mathbf{P}_{\tau|\tau-1}\tilde{\mathbf{C}}(\boldsymbol{\theta})^{\top}\boldsymbol{\Omega}_{\tau|\tau-1}^{-1} \tag{B.10}$$

is referred to as the Kalman gain. Equations (B.4)-(B.10) together with initial conditions $\mathbf{x}_0$ and $\mathbf{P}_0$ define the *Kalman filter recursion* for $\tau = 1, 2, \ldots, T$. If convergence occurs at a given point $t_{\tau-1}$, then $\mathbf{K} = \mathbf{K}_{s+1|s} = \mathbf{K}_{s|s-1}$ and $\boldsymbol{\Omega} = \boldsymbol{\Omega}_{s+1|s} = \boldsymbol{\Omega}_{s|s-1}$ for all $s \geq \tau \in \mathbb{N}$. Namely, $\mathbf{K} = \mathbf{K}(\boldsymbol{\theta})$ and $\boldsymbol{\Omega} = \boldsymbol{\Omega}(\boldsymbol{\theta})$ are the two (time-variant) matrices, after convergence.

B.3 Likelihood function

Given the linear structure of the Gaussian state-space model (2.8)-(2.9), it follows that the (joint) probability density function of the discrete measurements can be written as

$$f(\mathbf{y}^T;\boldsymbol{\theta}) = f(\mathbf{y}_1,\ldots,\mathbf{y}_T;\boldsymbol{\theta}) = f(\mathbf{y}_1;\boldsymbol{\theta})\prod_{\tau=2}^{T} f(\mathbf{y}_\tau|\mathbf{y}_{\tau-1};\boldsymbol{\theta}),$$

where $f\left(\mathbf{y}_\tau|\mathbf{y}_{\tau-1};\boldsymbol{\theta}\right) = \mathcal{N}\left(\mathsf{C}(\boldsymbol{\theta})\mathbf{x}_{\tau-1|\tau-1}, \boldsymbol{\Omega}_{\tau|\tau-1}\right)$. Then, using the Kalman filter recursion, the log-likelihood function can be constructed recursively via the prediction error decomposition as (see Harvey, 1990)

$$\begin{aligned}
\mathcal{L}\left(\boldsymbol{\theta}|\mathbf{y}^T\right) &= \sum_{\tau=1}^{T} \ln f\left(\mathbf{y}_\tau|\mathbf{y}_{\tau-1};\boldsymbol{\theta}\right) \\
&= -\frac{n_y T}{2}\ln(2\pi) - \frac{1}{2}\sum_{\tau=1}^{T}\ln|\boldsymbol{\Omega}_{\tau|\tau-1}| - \frac{1}{2}\sum_{\tau=1}^{T}\boldsymbol{\omega}_{\tau|\tau-1}^{\top}\boldsymbol{\Omega}_{\tau|\tau-1}^{-1}\boldsymbol{\omega}_{\tau|\tau-1},
\end{aligned}$$

with $f(\mathbf{y}_1|\mathbf{y}_0) = f(\mathbf{y}_1)$, and the maximum-likelihood (ML) estimator of $\boldsymbol{\theta}$ by

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\arg\max}\ \mathcal{L}\left(\boldsymbol{\theta}|\mathbf{y}^T\right).$$

B.4 State smoothing

For $\tau = T, ..., 1$, we smooth states $\mathbf{x}_\tau$, given observations $\{\mathbf{y}_1, ..., \mathbf{y}_T\}$ using the fast state-smoothing recursion in Durbin and Koopman (2012, Chapter 4). The first part of the procedure consists of running a backward disturbance smoother algorithm. In particular, let

$\mathbf{r}_T = \mathbf{0}$. Then compute

$$\begin{aligned}
\hat{\mathbf{d}}_\tau &= \mathsf{D}(\boldsymbol{\theta})\boldsymbol{\Sigma}_\epsilon(\boldsymbol{\theta})\mathsf{D}(\boldsymbol{\theta})^\top \mathbf{v}_\tau \\
\hat{\mathbf{n}}_\tau &= \mathsf{B}(\boldsymbol{\theta})\boldsymbol{\Sigma}_\epsilon(\boldsymbol{\theta})\mathsf{B}(\boldsymbol{\theta})^\top \mathbf{r}_\tau \\
\mathbf{v}_\tau &= \boldsymbol{\Omega}_{\tau|\tau-1}^{-1}\boldsymbol{\omega}_{\tau|\tau-1} - \mathbf{K}_{\tau|\tau-1}^\top \mathbf{r}_\tau \\
\mathbf{r}_{\tau-1} &= \tilde{\mathbf{C}}(\boldsymbol{\theta})^\top \mathbf{v}_\tau + \mathsf{A}(\boldsymbol{\theta})^\top \mathbf{r}_\tau,
\end{aligned}$$

where disturbances $\mathbf{d}_\tau$ and $\mathbf{n}_\tau$ are used for diagnostics, (cf. Durbin and Koopman 2012, Chapter 7) and have variance-covariance matrices

$$\begin{aligned}
\mathrm{Var}(\mathbf{d}_\tau|\mathbf{y}^T) &= \mathsf{D}(\boldsymbol{\theta})\boldsymbol{\Sigma}_\epsilon(\boldsymbol{\theta})\mathsf{D}(\boldsymbol{\theta})^\top - \mathsf{D}(\boldsymbol{\theta})\boldsymbol{\Sigma}_\epsilon(\boldsymbol{\theta})\mathsf{D}(\boldsymbol{\theta})^\top \mathbf{D}_\tau \mathsf{D}(\boldsymbol{\theta})\boldsymbol{\Sigma}_\epsilon(\boldsymbol{\theta})\mathsf{D}(\boldsymbol{\theta})^\top, \\
\mathrm{Var}(\mathbf{n}_\tau|\mathbf{y}^T) &= \mathsf{B}(\boldsymbol{\theta})\boldsymbol{\Sigma}_\epsilon(\boldsymbol{\theta})\mathsf{B}(\boldsymbol{\theta})^\top - \mathsf{B}(\boldsymbol{\theta})\boldsymbol{\Sigma}_\epsilon(\boldsymbol{\theta})\mathsf{B}(\boldsymbol{\theta})^\top \mathbf{N}_\tau \mathsf{B}(\boldsymbol{\theta})\boldsymbol{\Sigma}_\epsilon(\boldsymbol{\theta})\mathsf{B}(\boldsymbol{\theta})^\top,
\end{aligned}$$

with

$$\mathbf{D}_\tau = \boldsymbol{\Omega}_{\tau|\tau-1} + \mathbf{K}_{\tau|\tau-1}^\top \mathbf{N}_\tau \mathbf{K}_{\tau|\tau-1},$$

and

$$\mathbf{N}_{\tau-1} = \tilde{\mathbf{C}}(\boldsymbol{\theta})^\top \mathbf{D}_\tau \tilde{\mathbf{C}}(\boldsymbol{\theta}) + \mathsf{A}(\boldsymbol{\theta})^\top \mathbf{N}_\tau \mathsf{A}(\boldsymbol{\theta}) + \tilde{\mathbf{C}}(\boldsymbol{\theta})^\top \mathbf{K}_{\tau|\tau-1}^\top \mathbf{N}_\tau \mathsf{A}(\boldsymbol{\theta}) - \mathsf{A}(\boldsymbol{\theta})^\top \mathbf{N}_\tau \mathbf{K}_{\tau|\tau-1} \tilde{\mathbf{C}}(\boldsymbol{\theta}).$$

Initialize the state vector by $\hat{\mathbf{x}}_1 = a_1 + \mathbf{P}_1 \mathbf{r}_0$. Then, for $\tau = 1, ..., T$, the second part of the procedure consists of computing the recursion

$$\begin{aligned}
\hat{\boldsymbol{\eta}}_{\tau+1} &= \mathsf{B}(\boldsymbol{\theta})\boldsymbol{\Sigma}_\epsilon(\boldsymbol{\theta})\mathsf{B}(\boldsymbol{\theta})^\top \mathbf{r}_\tau, \\
\hat{\mathbf{x}}_{\tau+1} &= \mathsf{A}(\boldsymbol{\theta})\hat{\mathbf{x}}_\tau + \hat{\boldsymbol{\eta}}_{\tau+1}.
\end{aligned}$$

C. State-space representation for mixed-frequency data

Consider the case where measurements are sampled at two different frequencies, i.e., a high frequency $\bar{h}$, and a low frequency $\underline{h}$, with $\underline{h} > \bar{h}$. The associated state-space representation in the case of mixed-frequency observations depends on the sampling nature of the data:

Stock data. Rewrite the transition equation (2.3) for the S-SSR model, or in (4.13) for the EM-SSR model, in terms of the time step associated with the highest frequency available, i.e., $\bar{h}$. On the other hand, rewrite the measurement equation in (2.7) as $\mathbf{y}_\tau = \mathbf{W}_\tau \mathbf{C}(\boldsymbol{\theta})$, where $\mathbf{W}_\tau$ is a known time-varying matrix whose rows at a given point in time are a subset of the rows of $\mathbf{I}_{n_y}$. More specifically, the number of rows at time t_τ is determined by the number of variables for which observations are available in t_τ.

For instance, consider the estimation of the RBC model discussed in the main text using as observables monthly data ($\bar{h} = 1/12$) on aggregate consumption, c_τ, and quarterly data ($\underline{h} = 1/4$) on the fraction of hours worked, n_τ. Further assume that there is no measurement error, i.e., $\boldsymbol{\varepsilon}_\tau = \mathbf{0}$ for all τ. Then, the S-SSR model for mixed-frequency data sampling is given by

$$\begin{bmatrix} k_\tau \\ z_\tau \end{bmatrix} = \mathbf{A}_{\bar{h}}(\boldsymbol{\theta}) \begin{bmatrix} k_{\tau-1} \\ z_{\tau-1} \end{bmatrix} + \boldsymbol{\eta}^s_\tau \tag{C.1}$$

$$\begin{bmatrix} c_\tau \\ n_\tau \end{bmatrix} = \mathbf{W}_\tau \mathbf{C}(\boldsymbol{\theta}) \begin{bmatrix} k_\tau \\ z_\tau \end{bmatrix} \tag{C.2}$$

with $\mathbf{W}_\tau = [\ 1 \quad 0\ ]$ when τ coincides with observations recorder at points in time within a quarter (so only c_τ is available), and $\mathbf{W}_\tau = \mathbf{I}_{n_y}$ when τ coincides with observations recorded at the end of a quarter (so both c_τ and n_τ are available). The Kalman filter recursions proceeds accordingly by accommodating the missing values of the variables sampled at low frequencies as shown in Durbin and Koopman (2012, Chp. 4.10).

Notice that the state-space representation formed by (C.1)-(C.2) is the same as that in (2.3)-(2.7) for the S-SSR model, or in (4.13)-(2.7) for the EM-SSR model, with $\mathbf{W}_\tau = \mathbf{I}_{n_y}$ for all τ in the case of a common sampling frequency among observables. Despite their similarity, we refer to this new state-space representations as the MXS-SSR and MXEM-SSR models to emphasize their application to mixed-frequency sampling.

Flow data. The state-space representation for flow variables is more involved than that for stock data because we need to keep track of the unobserved time aggregation occurring at "high" frequencies of the variables sampled at lower frequencies. This can be achieved by

introducing a number of additional deterministic states that measure the unobserved behavior of the flow variables within the time intervals for which observations are not available. As for the stock case, the state-space representation is also written in terms of the time step associated with the highest frequency, $\bar{h}$.

So if we consider again the estimation of the RBC model using as observables monthly data ($\bar{h} = 1/12$) on aggregate consumption, c_τ, and quarterly data ($\underline{h} = 1/4$) on the fraction of hours worked, n_τ. Further assume that there is no measurement error, i.e., $\boldsymbol{\varepsilon}_\tau = \mathbf{0}$ for all τ. Then, the F-SSR model for mixed-frequency sampling is given by

$$\begin{bmatrix} k_\tau \\ z_\tau \\ c^f_\tau \\ n^f_\tau \\ \hdashline n^f_{\tau-1} \\ n^f_{\tau-2} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{\bar{h}}(\boldsymbol{\theta}) & \mathbf{0}_{4\times 4} \\ \mathbf{C}(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})^{-1}(\mathbf{A}_{\bar{h}}(\boldsymbol{\theta}) - \mathbf{I}) & \\ \hdashline \mathbf{0}_{2\times 2} & \begin{matrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{matrix} \end{bmatrix} \begin{bmatrix} k_{\tau-1} \\ z_{\tau-1} \\ c^f_{\tau-1} \\ n^f_{\tau-1} \\ \hdashline n^f_{\tau-2} \\ n^f_{\tau-3} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\eta}^s_\tau \\ \boldsymbol{\eta}^f_\tau \\ \hdashline 0 \\ 0 \end{bmatrix} \tag{C.3}$$

$$\begin{bmatrix} c_\tau \\ n_\tau \end{bmatrix} = \mathbf{W}_\tau \left[\begin{array}{cc:cc} \mathbf{0}_{2\times 2} & \mathbf{I}_2 & \begin{matrix} 0 & 0 \\ 1 & 1 \end{matrix} \end{array} \right] \begin{bmatrix} k_\tau \\ z_\tau \\ c^f_\tau \\ n^f_\tau \\ \hdashline n^f_{\tau-1} \\ n^f_{\tau-2} \end{bmatrix}, \tag{C.4}$$

where $\mathbf{W}_\tau$ is defined as before. Following the arguments in Section 2.4, notice that any point in time, t_τ, the observed fraction of hours worked in (C.4), sampled at a lower frequency $\underline{h}$, n_τ, is related to the higher frequency state equations according to

$$\begin{aligned} n_\tau = n^f(t_\tau) &= \int_{t_\tau - \underline{h}}^{t_\tau} n(s)\mathrm{d}s \\ &= \int_{t_\tau - \bar{h}}^{t_\tau} n(s)\mathrm{d}s + \int_{t_\tau - 2\bar{h}}^{t_\tau - \bar{h}} n(s)\mathrm{d}s + \int_{t_\tau - \underline{h}}^{t_\tau - 2\bar{h}} n(s)\mathrm{d}s \\ &= n^f_\tau + n^f_{\tau-1} + n^f_{\tau-2}, \end{aligned}$$

where $n^f_\tau := \int_{t_\tau - \bar{h}}^{t_\tau} n(s)\mathrm{d}s$, $n^f_{\tau-1} := \int_{t_\tau - 2\bar{h}}^{t_\tau - \bar{h}} n(s)\mathrm{d}s$, and $n^f_{\tau-2} := \int_{t_\tau - \underline{h}}^{t_\tau - 2\bar{h}} n(s)\mathrm{d}s$. In the main text, we refer to this extended state-space representation for flow data in the case of mixed-frequency data as the MXF-SSR model.

Remark C.1. Notice that the transition equations for the additional deterministic state variables in (C.3), i.e., those associated with $n^f_{\tau-2}$ and $n^f_{\tau-3}$, do not depend on the vector of

unknown parameters. Then, it is possible to use $\boldsymbol{\theta}$ estimated from observations sampled a low frequency $\underline{h}$ to recover state estimates at a higher frequency $\overline{h}$ by simply using the state-space representation in (C.1)-(C.2) or (C.3)-(C.4) together with any filtering algorithm that handles missing observations. This means that the model does not need to be re-estimated if the analysis entails the same variables observed at different frequencies.

D. Proofs and derivations

D.1 Proof of Proposition 1

The exact discrete model (EDM) corresponds to the solution, in the narrow-sense, of the linear SDE in (2.1)

$$\mathrm{d}\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{x}(t)\mathrm{d}t + \mathbf{B}(\boldsymbol{\theta})\mathrm{d}\mathbf{w}(t),$$

with fixed initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$, and where $\mathbf{A}(\boldsymbol{\theta})$ and $\mathbf{B}(\boldsymbol{\theta})$ are time-invariant matrices. In particular, consider the solution to the associated vector-valued homogeneous ordinary differential equation

$$\mathrm{d}\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{x}(t)\mathrm{d}t, \quad \mathbf{x}\left(t_0\right) = \mathbf{x}_0,$$

which is given by

$$\mathbf{x}_t = \exp\left(\int_{t_0}^{t} \mathbf{A}(\boldsymbol{\theta})\mathrm{d}\tau\right)\mathbf{x}_0 = \exp(\mathbf{A}(\boldsymbol{\theta})h)\mathbf{x}_0 = \boldsymbol{\Phi}\left(t, t_0\right)\mathbf{x}_0,$$

where we have defined $h = t - t_0$, and where $\boldsymbol{\Phi}\left(t_0, t_0\right) = \mathbf{I}$. An application of Itô's formula to the transformation $\boldsymbol{\Phi}\left(t, t_0\right)^{-1}\mathbf{x}$ yields

$$\begin{aligned}\mathrm{d}\left(\boldsymbol{\Phi}\left(t, t_0\right)^{-1}\mathbf{x}_t\right) &= \left(\frac{\partial \boldsymbol{\Phi}\left(t, t_0\right)^{-1}}{\partial t}\mathbf{x}_t + \boldsymbol{\Phi}\left(t, t_0\right)^{-1}\mathbf{A}(\boldsymbol{\theta})\mathbf{x}_t\right)\mathrm{d}t + \boldsymbol{\Phi}\left(t, t_0\right)^{-1}\mathbf{B}(\boldsymbol{\theta})\mathrm{d}\mathbf{w}(t) \\ &= \boldsymbol{\Phi}\left(t, t_0\right)^{-1}\mathbf{B}(\boldsymbol{\theta})\mathrm{d}\mathbf{w}(t), \end{aligned} \tag{D.1}$$

where $\boldsymbol{\Phi}\left(t, t_0\right)^{-1}$, is called the integrating factor. Integrating both sides of (D.1) we obtain the solution to (2.1) as

$$\mathbf{x}(t) = \boldsymbol{\Phi}\left(t, t_0\right)\left(\mathbf{x}_0 + \int_{t_0}^{t} \boldsymbol{\Phi}\left(s, t_0\right)^{-1}\mathbf{B}(\boldsymbol{\theta})\mathrm{d}\mathbf{w}(s)\right).$$

By setting $t_0 = t_{\tau-1}$ and $t = t_\tau$, the solution can be written as

$$\begin{aligned}\mathbf{x}(t_\tau) &= \exp(\mathbf{A}(\boldsymbol{\theta})h)\mathbf{x}(t_{\tau-1}) + \exp(\mathbf{A}(\boldsymbol{\theta})h)\int_{t_{\tau-1}}^{t_\tau} \exp(\mathbf{A}(\boldsymbol{\theta})(t_{\tau-1} - s))\mathbf{B}(\boldsymbol{\theta})\mathrm{d}\mathbf{w}(s) \\ &= \exp(\mathbf{A}(\boldsymbol{\theta})h)\mathbf{x}(t_{\tau-1}) + \exp(\mathbf{A}(\boldsymbol{\theta})t_\tau)\int_{t_{\tau-1}}^{t_\tau} \exp(-\mathbf{A}(\boldsymbol{\theta})s)\mathbf{B}(\boldsymbol{\theta})\mathrm{d}\mathbf{w}(s) \\ &= \exp(\mathbf{A}(\boldsymbol{\theta})h)\mathbf{x}(t_{\tau-1}) + \int_{t_{\tau-1}}^{t_\tau} \exp\left(\mathbf{A}(\boldsymbol{\theta})(t_\tau - s)\right)\mathbf{B}(\boldsymbol{\theta})\mathrm{d}\mathbf{w}(s) \\ &= \mathbf{A}_h(\boldsymbol{\theta})\mathbf{x}(t_{\tau-1}) + \boldsymbol{\eta}^s(t_\tau).\end{aligned}$$

Let $\boldsymbol{\eta}(t_\tau) = \boldsymbol{\eta}_\tau$. The error term is a flow variable and, from the definition of Brownian increment, it has mean $\mathbb{E}[\boldsymbol{\eta}^s(t_\tau)] = \mathbf{0}$. The covariance matrix of the error term (2.6) is proved by first applying the Itô isometry property,

$$\begin{aligned}\mathbb{E}[\boldsymbol{\eta}_\tau^s \boldsymbol{\eta}_\tau^{s\top}] &= \mathbb{E}\left[\left(\int_{t_{\tau-1}}^{t_\tau} \exp\left(\mathbf{A}(\boldsymbol{\theta})(t_\tau - s)\right)\mathbf{B}(\boldsymbol{\theta})\mathrm{d}\mathbf{w}(s)\right)\right. \\ &\quad \left.\times\left(\int_{t_{\tau-1}}^{t_\tau} \exp\left(\mathbf{A}(\boldsymbol{\theta})(t_\tau - s)\right)\mathbf{B}(\boldsymbol{\theta})\mathrm{d}\mathbf{w}(s)\right)^\top\right] \\ &= \int_{t_{\tau-1}}^{t_\tau} \exp\left(\mathbf{A}(\boldsymbol{\theta})(t_\tau - s)\right)\mathbf{B}(\boldsymbol{\theta})\mathbf{B}(\boldsymbol{\theta})^\top \exp\left(\mathbf{A}(\boldsymbol{\theta})^\top(t_\tau - s)\right)\mathrm{d}s. \qquad \text{(D.2)}\end{aligned}$$

The equation is time-invariant. It only depends on the fixed time interval between measurements, $h = t_\tau - t_{\tau-1}$. Therefore, a change of variable leads to

$$\begin{aligned}\boldsymbol{\Sigma}_{\boldsymbol{\eta}^s,h}(\boldsymbol{\theta}) = \mathbb{E}[\boldsymbol{\eta}_\tau^s \boldsymbol{\eta}_\tau^{s\top}] &= \int_0^h \exp\left(\mathbf{A}(\boldsymbol{\theta})(h - s)\right)\mathbf{B}(\boldsymbol{\theta})\mathbf{B}(\boldsymbol{\theta})^\top \exp\left(\mathbf{A}(\boldsymbol{\theta})^\top(h - s)\right)\mathrm{d}s \\ &= \int_0^h \exp\left(\mathbf{A}(\boldsymbol{\theta})(h - s)\right)\boldsymbol{\Sigma}(\boldsymbol{\theta}) \exp\left(\mathbf{A}(\boldsymbol{\theta})^\top(h - s)\right)\mathrm{d}s.\end{aligned}$$

Moreover, notice that the errors are serially uncorrelated at all leads and lags,

$$\mathbb{E}[\boldsymbol{\eta}_\tau^s \boldsymbol{\eta}_{\tau-\ell}^{s\top}] = \mathbf{0}, \text{ for all } \ell \neq 0.$$

■

D.2 Derivation of Equation (2.11)

Equation (2.11) characterizes the form of reduced-form residuals, $\boldsymbol{\eta}_\tau^f$, in

$$\mathbf{y}_\tau^f = \mathbf{C}(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})^{-1}(\exp(\mathbf{A}(\boldsymbol{\theta})h) - \mathbf{I})\mathbf{x}_{\tau-1} + \boldsymbol{\eta}_\tau^f.$$

The cumulator operator within the interval $[t_{\tau-1}, t_\tau]$ leads to

$$\boldsymbol{\eta}_\tau^f = \mathbf{C}(\boldsymbol{\theta}) \int_0^h \int_{t_{\tau-1}}^{t_{\tau-1}+s} \exp(\mathbf{A}(\boldsymbol{\theta})(t_{\tau-1} + s - r))\mathbf{B}(\boldsymbol{\theta})\mathrm{d}\mathbf{w}(r)\,\mathrm{d}s$$

by redefining the bounds of the inner stochastic integral, and then exchanging the order of integration,

$$\begin{aligned}\boldsymbol{\eta}_\tau^f &= \mathbf{C}(\boldsymbol{\theta}) \int_0^h \int_0^s \exp(\mathbf{A}(\boldsymbol{\theta})(s-r))\mathbf{B}(\boldsymbol{\theta})\mathrm{d}\mathbf{w}(t_{\tau-1}+r)\,\mathrm{d}s \\ &= \mathbf{C}(\boldsymbol{\theta}) \int_0^h \int_r^h \exp(\mathbf{A}(\boldsymbol{\theta})(s-r))\mathbf{B}(\boldsymbol{\theta})\mathrm{d}s\,\mathrm{d}\mathbf{w}(t_{\tau-1}+r). \end{aligned} \tag{D.3}$$

At this point, we notice that the definite integral

$$\int_r^h \exp(\mathbf{A}(\boldsymbol{\theta})(s-r))\mathbf{B}(\boldsymbol{\theta})\mathrm{d}s = \mathbf{A}(\boldsymbol{\theta})^{-1}(\exp(\mathbf{A}(\boldsymbol{\theta})(h-r)) - \mathbf{I})\mathbf{B}(\boldsymbol{\theta}).$$

By substituting the solution of the definite integral into the stochastic integral, setting $s = t_{\tau-1} + r$, and using the fact that the uniform time-step $h = t_\tau - t_{\tau-1}$, we conclude

$$\boldsymbol{\eta}_\tau^f = \mathbf{C}(\boldsymbol{\theta}) \int_{t_{\tau-1}}^{t_\tau} \mathbf{A}(\boldsymbol{\theta})^{-1}(\exp(\mathbf{A}(\boldsymbol{\theta})(t_\tau - s)) - \mathbf{I})\mathbf{B}(\boldsymbol{\theta})\mathrm{d}\mathbf{w}(s).$$

The quadrants of the covariance matrix of $\boldsymbol{\eta}_\tau = \left[\boldsymbol{\eta}_\tau^{s\top}, \boldsymbol{\eta}_\tau^{f\top}\right]^\top$ in the F-SSR model apply the Itô Isometry property similarly to (D.2) to the expected values $\mathbb{E}[\boldsymbol{\eta}_\tau^s \boldsymbol{\eta}_\tau^{f\top}]$ and $\mathbb{E}[\boldsymbol{\eta}_\tau^f \boldsymbol{\eta}_\tau^{f\top}]$, where $\boldsymbol{\eta}_\tau^s$ and $\boldsymbol{\eta}_\tau^f$ come from (2.5) and (D.3), respectively.

D.3 Proof of Lemma 1

The proof of the Lemma requires that we show that the eigenvalues of $\mathsf{A}(\boldsymbol{\theta})$ lie within the unit circle for all the considered models.

The S-SSR case. From Assumption 1 we know that the eigenvalues of $\mathbf{A}(\boldsymbol{\theta})$ have negative real part. Therefore, the elements of $\boldsymbol{\Lambda}$ in section A.1 have all negative real part. Since

$$\mathsf{A}(\boldsymbol{\theta}) := \exp(\mathbf{A}(\boldsymbol{\theta})h) = \mathbf{V}\exp(\boldsymbol{\Lambda}h)\mathbf{V}^{-1},$$

with $h \geq 0$, and elements on the diagonal of $\exp(\boldsymbol{\Lambda}h) = \mathrm{diag}(\mathrm{e}^{\lambda_1 h}, ..., \mathrm{e}^{\lambda_{n_x} h})$ are the eigenvalues of $\mathsf{A}(\boldsymbol{\theta})$, with $\lambda_i < 0$ for all $i = 1, ..., n_x$. It follows that every diagonal element of $\exp(\boldsymbol{\Lambda}h)$ lies within the unit circle.

The F-SSR case. We just proved that the upper-left block of (2.12), $\exp(\mathbf{A}(\boldsymbol{\theta})h)$, has eigenvalues within the unit circle. From (2.12), using decoupling arguments, is evident that, for $\sigma \in \mathbb{C}$, the nonzero roots from the characteristic equation $\det(\sigma\mathbf{I} - \mathsf{A}(\boldsymbol{\theta}))$ are the same as those of $\det(\sigma\mathbf{I} - \exp(\mathbf{A}(\boldsymbol{\theta})h))$ (see e.g. Golub and Van Loan, 2013, Lemma 7.1.1). In fact, let $\sigma(\mathbf{A}) = \{\sigma : \det(\sigma\mathbf{I} - \mathbf{A}) = 0\}$ and partition the discrete-time drift matrix of (2.12) as follows

$$\mathsf{A}(\boldsymbol{\theta}) = \begin{bmatrix} \mathsf{A}_{11}(\boldsymbol{\theta}) & \mathbf{0} \\ \mathsf{A}_{21}(\boldsymbol{\theta}) & \mathsf{A}_{22}(\boldsymbol{\theta}) \end{bmatrix},$$

then $\sigma(\mathsf{A}(\boldsymbol{\theta})) = \sigma(\mathsf{A}_{11}(\boldsymbol{\theta})) \cup \sigma(\mathsf{A}_{22}(\boldsymbol{\theta}))$, where however $\mathsf{A}_{22}(\boldsymbol{\theta}) := \mathbf{0}$ and $\sigma(\mathbf{0}) = \text{diag}(0, ..., 0)$. It follows that $\mathsf{A}(\boldsymbol{\theta})$ has n_x eigenvalues given by $0 \leq \mathrm{e}^{\lambda_i h} < 1$ for $i = 1, ..., n_x$, and n_y eigenvalues that are exactly zero. Therefore, all the eigenvalues are within the unit circle also for the F-SSR model.

The EM-SSR case. From Assumption 1 we know that the eigenvalues of $\mathbf{A}(\boldsymbol{\theta})$ have negative real part. Additionally, $\mathsf{A}(\boldsymbol{\theta}) = \mathbf{I}_{n_x} + \mathbf{A}(\boldsymbol{\theta})h$. The eigenvalues of $\mathsf{A}(\boldsymbol{\theta})$ are of the type $1 + \lambda_i h$, $i = 1, ..., n_x$, which lie within the unit circle if and only if $\lambda_i h > -2$ for all $i = 1, ..., n_x$. ■

D.4 Proof of Proposition 4

If Assumptions 1 and 3 hold, it follows from Proposition 3 that $\mathbf{A}(\boldsymbol{\theta})$ is identified from $\mathbf{A}_h(\boldsymbol{\theta})$. Then, if $\int_0^h \mathbf{A}_s(\boldsymbol{\theta}) \otimes \mathbf{A}_s(\boldsymbol{\theta}) \mathrm{d}s$ is invertible, $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ is the solution of (A.2). Notice that $\int_0^h \mathbf{A}_s(\boldsymbol{\theta}) \otimes \mathbf{A}(\boldsymbol{\theta}) \mathrm{d}s := (\mathbf{A}(\boldsymbol{\theta}) \otimes \mathbf{I}_{n_x} + \mathbf{I}_{n_x} \otimes \mathbf{A}(\boldsymbol{\theta}))^{-1}(\mathbf{A}_h(\boldsymbol{\theta}) \otimes \mathbf{A}_h(\boldsymbol{\theta}) - \mathbf{I}_{n_x^2})$ and $(\mathbf{A}(\boldsymbol{\theta}) \otimes \mathbf{I}_{n_x} + \mathbf{I}_{n_x} \otimes \mathbf{A}(\boldsymbol{\theta}))^{-1}$ exists with negative eigenvalues of the type $1/\lambda_i + 1/\lambda_j$ for all $i, j = 1, ..., n_x$.

The proof of proposition 4 requires to prove (i) nonsingularity of $\mathcal{A}_h = (\mathbf{A}_h(\boldsymbol{\theta}) \otimes \mathbf{A}_h(\boldsymbol{\theta}) - \mathbf{I}_{n_x^2})$ (all the eigenvalues of $\mathcal{A}_h$ are different from zero) and that (ii) $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ is a positive semi-definite matrix; that is $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ is an admissible instantaneous covariance matrix.

(i) Nonsingularity of $\mathcal{A}_h$ follows from Assumption 1. In fact, notice that eigenvalues of $(\mathbf{A}(\boldsymbol{\theta}) \otimes \mathbf{A}(\boldsymbol{\theta}))$, are given by the pairwise multiplications of the eigenvalues in $\lambda(\mathbf{A}(\boldsymbol{\theta}))$ of $\mathbf{A}(\boldsymbol{\theta})$. Namely, the n_x^2 eigenvalues of $(\mathbf{A}(\boldsymbol{\theta}) \otimes \mathbf{A}(\boldsymbol{\theta}))$, are formed as $\lambda_i \lambda_j > 0$ for all $i, j = 1, ..., n_x$. Then the eigenvalues of $\mathcal{A}_h$ are of the form $\exp((\lambda_i + \lambda_j)h - 1)$, which is negative, therefore different from zero, for all $i, j = 1, ..., n_x$ by Assumption 1.

Finally (ii) the solution of (A.2) is a positive semi-definite matrix. In fact $\boldsymbol{\Sigma}_{\boldsymbol{\eta}^s,h}(\boldsymbol{\theta})$ is positive semi-definite, the eigenvalues of $(\mathbf{A}(\boldsymbol{\theta}) \otimes \mathbf{I}_{n_x} + \mathbf{I}_{n_x} \otimes \mathbf{A}(\boldsymbol{\theta}))$ are negative and of the type $(\lambda_i + \lambda_j)$, and the eigenvalues of $\mathcal{A}_h^{-1}$ are also negative. Hence, all the eigenvalues of $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ are non-negative. ■

Remark D.1 (Joint identification)**.** While Propositions 3 and 4 achieve identification of $\mathbf{A}(\boldsymbol{\theta}_0)$

and $\boldsymbol{\Sigma}(\boldsymbol{\theta}_0)$ sequentially, it is also possible to derive conditions for their *joint identification*. In particular, a sufficient condition for the identification of $(\mathbf{A}(\boldsymbol{\theta}_0), \boldsymbol{\Sigma}(\boldsymbol{\theta}_0))$ from discrete-time measurements can be derived by constructing the augmented matrix (see Van Loan, 1978)

$$\boldsymbol{\Xi}(\boldsymbol{\theta}_0) = \begin{bmatrix} \mathbf{A}(\boldsymbol{\theta}_0) & \boldsymbol{\Sigma}(\boldsymbol{\theta}_0) \\ \mathbf{0} & -\mathbf{A}(\boldsymbol{\theta}_0)^\top \end{bmatrix}.$$

Then, we have the following result.

Proposition 7 (McCrorie, 2003)**.** If the eigenvalues of the augmented matrix $\boldsymbol{\Xi}(\boldsymbol{\theta}_0)$ are strictly real and no redundant Jordan block occurs more than once, then $(\mathbf{A}(\boldsymbol{\theta}_0), \boldsymbol{\Sigma}(\boldsymbol{\theta}_0))$ is identified from $(\mathbf{A}_h(\boldsymbol{\theta}_0), \boldsymbol{\Sigma}_{\boldsymbol{\eta},h}(\boldsymbol{\theta}_0))$.

The proof of Proposition 7 uses the results in Culver (1966, Theorem 2) for the uniqueness of the matrix logarithm on the extended matrix $\boldsymbol{\Xi}(\boldsymbol{\theta}_0)$. ■

D.5 Proof of Proposition 5

For possibly large but finite n, the proof is constructed in three steps. First, we show that the proposition holds for $\boldsymbol{\eta}^s$, which it is the case for the S-SSR model. Then, we do the same for $\boldsymbol{\eta}^f$ in isolation. Finally, the proof of the proposition for the F-SSR model follows from stacking the previous two cases.

The S-SSR case. Consider $n \in \mathbb{Z}^+$ such that the Riemann–Stieltjes sum

$$\boldsymbol{\eta}^s_\tau = \sum_{i=1}^{n} \exp(\mathbf{A}(\boldsymbol{\theta})(t_\tau - t^\tau_{i-1}))\mathbf{B}(\boldsymbol{\theta})\Delta\mathbf{w}(t^\tau_i) + o_P(1).$$

Expanding the sum on the right-hand side yields

$$\begin{aligned} \boldsymbol{\eta}^s_\tau = {} & \exp(\mathbf{A}(\boldsymbol{\theta})h)\mathbf{B}(\boldsymbol{\theta})\Delta\mathbf{w}(t^\tau_1) + \exp(\mathbf{A}(\boldsymbol{\theta})(h - h_n))\mathbf{B}(\boldsymbol{\theta})\Delta\mathbf{w}(t^\tau_2)\ldots \\ & + \exp(\mathbf{A}(\boldsymbol{\theta})(h - (n-2)h_n))\mathbf{B}(\boldsymbol{\theta})\Delta\mathbf{w}(t^\tau_{n-1})\ldots \\ & + \exp(\mathbf{A}(\boldsymbol{\theta})(h - (n-1)h_n))\mathbf{B}(\boldsymbol{\theta})\Delta\mathbf{w}(t^\tau_n) + o_P(1), \end{aligned}$$

which can be rewritten as

$$\begin{aligned}\boldsymbol{\eta}_\tau^s &= (\mathbf{I} + \mathbf{A}(\boldsymbol{\theta})h + \mathbf{A}(\boldsymbol{\theta})^2h^2/2 + \dots)\mathbf{B}(\boldsymbol{\theta})\Delta\mathbf{w}(t_1^\tau)\dots \\ &+ (\mathbf{I} + \mathbf{A}(\boldsymbol{\theta})(h-h_n) + \mathbf{A}(\boldsymbol{\theta})^2(h-h_n)^2/2 + \dots)\mathbf{B}(\boldsymbol{\theta})\Delta\mathbf{w}(t_2^\tau)\dots \\ &+ (\mathbf{I} + \mathbf{A}(\boldsymbol{\theta})(h-(n-2)h_n) + \mathbf{A}(\boldsymbol{\theta})^2(h-(n-2)h_n)^2/2 + \dots)\mathbf{B}(\boldsymbol{\theta})\Delta\mathbf{w}(t_{n-1}^\tau)\dots \\ &+ (\mathbf{I} + \mathbf{A}(\boldsymbol{\theta})(h-(n-1)h_n) + \mathbf{A}(\boldsymbol{\theta})^2(h-(n-1)h_n)^2/2 + \dots)\mathbf{B}(\boldsymbol{\theta})\Delta\mathbf{w}(t_n^\tau) + o_P(1).\end{aligned}$$

Collecting terms with same coefficients yields

$$\begin{aligned}\boldsymbol{\eta}_\tau^s &= \left(\mathbf{I} + \mathbf{A}(\boldsymbol{\theta})h + \mathbf{A}(\boldsymbol{\theta})^2(h^2/2) + \dots\right)\mathbf{B}(\boldsymbol{\theta})\sum_{i=1}^{n}\Delta\mathbf{w}(t_i^\tau)\dots \\ &+ \left[\mathbf{A}(\boldsymbol{\theta})^2h_n^2/2 - \mathbf{A}(\boldsymbol{\theta})h_n - \mathbf{A}(\boldsymbol{\theta})^2hh_n - \dots\right]\mathbf{B}(\boldsymbol{\theta})\Delta\mathbf{w}(t_2^\tau)\dots \\ &+ \left[(n-2)^2\mathbf{A}(\boldsymbol{\theta})^2h_n^2/2 - \mathbf{A}(\boldsymbol{\theta})(n-2)h_n - (n-2)\mathbf{A}(\boldsymbol{\theta})^2hh_n + \dots\right]\mathbf{B}(\boldsymbol{\theta})\Delta\mathbf{w}(t_{n-1}^\tau)\dots \\ &+ \left[(n-1)^2\mathbf{A}(\boldsymbol{\theta})^2h_n^2/2 - \mathbf{A}(\boldsymbol{\theta})(n-1)h_n - (n-1)\mathbf{A}(\boldsymbol{\theta})^2hh_n + \dots\right]\mathbf{B}(\boldsymbol{\theta})\Delta\mathbf{w}(t_n^\tau) + o_P(1) \\ &= h^{1/2}\exp(\mathbf{A}(\boldsymbol{\theta})h)\mathbf{B}(\boldsymbol{\theta})\mathbf{u}_\tau + \mathcal{R}_\tau,\end{aligned}$$

where we have used the definitions of the matrix exponential in (2.4), of $\mathbf{u}_\tau$ in (3.1), and the fact that $t_0^\tau = t_{\tau-1}$ and $t_n^\tau = t_\tau$.

The term $\mathcal{R}_\tau$ denotes the remainder of the approximation $h^{1/2}\mathbf{H}(\boldsymbol{\theta}, h)\mathbf{u}_\tau$ for $\boldsymbol{\eta}_\tau^s$ where $\mathbf{H}(\boldsymbol{\theta}, h) = \exp(\mathbf{A}(\boldsymbol{\theta})h)\mathbf{B}(\boldsymbol{\theta})$. By recalling that $h_n = h/n$, the properties of $\mathcal{R}_\tau$ can be characterized according to the following three limiting behaviors: (i) one for increasing number of sub-intervals $n \to \infty$, while keeping h constant; (ii) one for decreasing length between sub-intervals, $h_n \to 0$; (iii) and, analogously, one for increasing frequency of data $h \to 0$. Note that shrinking h_n or h describes a similar behavior, that is, increasing data availability.

(i) As $n \to \infty$, it follows that $\mathbf{A}(\boldsymbol{\theta})h_n = \mathcal{O}(n^{-1})$, and the term $\mathbf{A}(\boldsymbol{\theta})(n-1)h_n \to \mathbf{A}(\boldsymbol{\theta})h$. Similarly, the term $\Delta\mathbf{w}(t_i^\tau) = \mathcal{O}_P(n^{-1/2})$ for all $1 < i \leq n$. Then, the remainder

$$\mathcal{R}_\tau = \mathcal{O}_P(n^{-1/2}),$$

i.e., it is bounded in probability by $n^{-1/2}$, a scalar that decreases with increasing number of sub-intervals.

(ii) As $h_n \to 0$, it follows that $\mathbf{A}(\boldsymbol{\theta})h_n = \mathcal{O}(h_n)$, $\mathbf{A}(\boldsymbol{\theta})(n-1)h_n = \mathcal{O}(h_n)$, and $\Delta\mathbf{w}(t_i^\tau) = \mathcal{O}_P(h_n^{1/2})$. Then,

$$\mathcal{R}_\tau = \mathcal{O}_P(h_n^{3/2}).$$

(iii) For $h \to 0$, it follows that $\mathbf{A}(\boldsymbol{\theta})h_n = \mathcal{O}(h)$, $\mathbf{A}(\boldsymbol{\theta})(n-1)h_n = \mathcal{O}(h)$, and $\Delta\mathbf{w}(t_i^\tau) = \mathcal{O}_P(h^{1/2})$, for all $1 < i \leq n$. Then,

$$\mathcal{R}_\tau = \mathcal{O}_P(h^{3/2}).$$

The F-SSR case. Let us first consider the case of $\boldsymbol{\eta}_\tau^f$ in isolation. In particular, consider an n such that the Reimann-Stieltjes sum

$$\boldsymbol{\eta}_\tau^f = \mathbf{C}(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})^{-1}\sum_{i=1}^{n}[\exp(\mathbf{A}(\boldsymbol{\theta})(t_\tau - t_{i-1}^\tau)) - \mathbf{I}]\mathbf{B}(\boldsymbol{\theta})\Delta\mathbf{w}(t_i^\tau) + o_P(1).$$

Similar to the case of the S-SSR model, we arrive to

$$\boldsymbol{\eta}_\tau^f = h^{1/2}\mathbf{C}(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})^{-1}\left[\exp(\mathbf{A}(\boldsymbol{\theta})h) - \mathbf{I}\right]\mathbf{B}(\boldsymbol{\theta})\mathbf{u}_\tau + \mathcal{R}_\tau$$

as $h \to 0$. Finally, recall that $\boldsymbol{\eta}_\tau = \left[\boldsymbol{\eta}_\tau^{s,\top}, \boldsymbol{\eta}_\tau^{f,\top}\right]^\top$. Thus, it follows that

$$\boldsymbol{\eta}_\tau = h^{1/2}\begin{bmatrix} \exp(\mathbf{A}(\boldsymbol{\theta})h)\mathbf{B}(\boldsymbol{\theta}) \\ \mathbf{C}(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta})^{-1}\left[\exp(\mathbf{A}(\boldsymbol{\theta})h) - \mathbf{I}\right]\mathbf{B}(\boldsymbol{\theta}) \end{bmatrix}\mathbf{u}_\tau + \mathcal{R}_\tau.$$

The properties of the remainder term in the expansion of $\boldsymbol{\eta}_\tau^f$ can be shown are the same as those for the expansion $\boldsymbol{\eta}_\tau^s$ in the S-SSR model. The proof is analogous to that above. Now, since $\boldsymbol{\eta}_\tau = \left[\boldsymbol{\eta}_\tau^{s,\top}, \boldsymbol{\eta}_\tau^{f,\top}\right]^\top$ in the F-SSR model it must follow that

(i) $\mathcal{R}_\tau = \mathcal{O}_P(n^{-1/2})$ as $n \to \infty$

(ii) $\mathcal{R}_\tau = \mathcal{O}_P(h_n^{3/2})$ as $h_n \to 0$

(iii) $\mathcal{R}_\tau = \mathcal{O}_P(h^{3/2})$ as $h \to 0$. ■

D.6 Proof of Proposition 6

Using Proposition 5 and given that $\mathbf{H}(\boldsymbol{\theta})$ is $\mathcal{O}(h^{3/2})$, it follows that $\mathbf{H}(\boldsymbol{\theta})^{-1}\mathcal{R}_\tau = \mathcal{O}_P(1)$. ■

D.7 Derivation of the historical shock decomposition in Equation (5.1)

The historical shock decomposition links the structural shocks of the system to the observables. Given Assumption 1 and Proposition 6, let $\boldsymbol{\epsilon} := \boldsymbol{\eta}$ in the ABCD representation (2.16)-(2.17) such that there is no measurement error wlog. Then the transition equation

can be written as

$$\begin{aligned}\mathbf{x}_\tau &= (\mathbf{I}_{n_x} - \mathsf{A}(\boldsymbol{\theta})L)^{-1}\mathbf{H}(\boldsymbol{\theta})\mathbf{u}_\tau + \mathcal{O}_P(h^{3/2}) \\ &= \sum_{s=0}^{\infty} \mathsf{A}(\boldsymbol{\theta})^s \mathbf{H}(\boldsymbol{\theta})\mathbf{u}_{\tau-s} + \mathcal{O}_P(h^{3/2}).\end{aligned}$$

We disregard the $\mathcal{O}_P(h^{3/2})$ terms and substitute the definition of $\mathbf{x}_\tau$ in the measurement equation to obtain

$$\begin{aligned}\mathbf{y}_{\tau+1} &= \mathsf{C}(\boldsymbol{\theta})(\mathbf{I}_{n_x} - \mathsf{A}(\boldsymbol{\theta})L)^{-1}\mathbf{H}(\boldsymbol{\theta})\mathbf{u}_\tau + \mathsf{D}(\boldsymbol{\theta})\mathbf{H}(\boldsymbol{\theta})\mathbf{u}_{\tau+1} \\ &= \mathsf{C}(\boldsymbol{\theta}) \sum_{s=0}^{\infty} \mathsf{A}(\boldsymbol{\theta})^s \mathbf{H}(\boldsymbol{\theta})\mathbf{u}_{\tau-s} + \mathsf{D}(\boldsymbol{\theta})\mathbf{H}(\boldsymbol{\theta})\mathbf{u}_{\tau+1}.\end{aligned}$$

Conditioning on the beginning of the sample, the transition equation can be alternatively written as

$$\mathbf{x}_\tau = \mathsf{A}(\boldsymbol{\theta})^\tau \mathbf{x}_0 + \sum_{s=0}^{\tau-1} \mathsf{A}(\boldsymbol{\theta})^s \mathbf{H}(\boldsymbol{\theta})\mathbf{u}_{\tau-s},$$

so the structural shocks are related to the measurements through

$$\begin{aligned}\mathbf{y}_{\tau+1} &= \mathsf{A}(\boldsymbol{\theta})^\tau \mathbf{x}_0 + \mathsf{C}(\boldsymbol{\theta})(\mathbf{I}_{n_x} - \mathsf{A}(\boldsymbol{\theta})L)^{-1}\mathbf{H}(\boldsymbol{\theta})\mathbf{u}_\tau + \mathsf{D}(\boldsymbol{\theta})\mathbf{H}(\boldsymbol{\theta})\mathbf{u}_{\tau+1} \\ &= \mathsf{C}(\boldsymbol{\theta})\mathsf{A}(\boldsymbol{\theta})^\tau \mathbf{x}_0 + \mathsf{D}(\boldsymbol{\theta})\mathbf{H}(\boldsymbol{\theta})\mathbf{u}_{\tau+1} + \mathsf{C}(\boldsymbol{\theta}) \sum_{s=0}^{\tau-1} \mathsf{A}(\boldsymbol{\theta})^s \mathsf{B}(\boldsymbol{\theta})\mathbf{H}(\boldsymbol{\theta})\mathbf{u}_{\tau-s}.\end{aligned}$$

The historical contributions of the individual structural shock u_i, $i = 1, ..., n_q$ to the measurements at the τth observation are then given by selecting and propagating the individual shocks one at a time. In practice, we use ML estimates $\hat{\boldsymbol{\theta}}$ of the structural parameters. Additionally, we replace the unobserved $\mathbf{x}_0$ and $\mathbf{u}$ with their smoothed estimates using approximations $\tilde{\mathbf{u}}$ instead of $\mathbf{u}$.

E. Stochastic optimal control problem

E.1 The HJB equation and the first-order conditions

The social planner chooses paths for consumption and the fraction of hours worked in order to maximize the expected lifetime utility

$$J(K_0, Z_0) = \max_{\{C(t),N(t)\}_{t=0}^{\infty}} \mathbb{E}_0 \left[\int_0^{\infty} e^{-\rho t} \left(\ln C(t) + \psi \left(1 - N(t) \right) \right) \mathrm{d}t \right]$$

subject to

$$\begin{aligned} \mathrm{d}K(t) &= \left(\exp\left(Z(t) \right) K(t)^{\alpha} \left(\exp\left(\eta t \right) N(t) \right)^{1-\alpha} - C(t) - \delta K(t) \right) \mathrm{d}t + \sigma_k K(t) \mathrm{d}w_k(t), \\ \mathrm{d}Z(t) &= -\rho_z Z(t) \mathrm{d}t + \sigma_z \mathrm{d}w_z(t), \end{aligned}$$

in which $C(t) \in \mathbb{R}^+$ and $N(t) \in [0, 1]$ are the control variables at instant $t > 0$, $K(t) \in \mathbb{R}^+$ and $Z(t) \in \mathbb{R}$ are the state variables at instant t, and $J(K, Z)$ is the value of the optimal program (value function) given the initial conditions $K(0) = K_0$ and $Z(0) = Z_0$.

The economy exhibits balanced growth path, i.e., over the long run the variables in the economy, with the exception of hours worked and TFP, will grow at the gross rate $\eta > 1$. A stationary version of the model can be obtained by defining $y(t) := Y(t)/\exp(\eta t)$, $c(t) := C(t)/\exp(\eta t)$, $k(t) := K(t)/\exp(\eta t)$ to be the de-trended values of the macroeconomic variables. For notation consistency, we also define $n(t) := N(t)$ and $z(t) := Z(t)$. Using these definitions, the planner's optimal control problem can be rewritten as[22]

$$J(k_0, z_0) = \max_{\{c(t),n(t)\}_{t=0}^{\infty}} \mathbb{E}_0 \left[\int_0^{\infty} e^{-\rho t} \left(\ln c(t) + \psi \left(1 - n(t) \right) \right) \mathrm{d}t \right]$$

subject to

$$\begin{aligned} \mathrm{d}k(t) &= \left(\exp\left(z(t) \right) k(t)^{\alpha} n_t^{1-\alpha} - c(t) - (\delta + \eta) k(t) \right) \mathrm{d}t + \sigma_k k(t) \mathrm{d}w_k(t), \quad k(0) = k_0 \\ \mathrm{d}z(t) &= -\rho_z z(t) \mathrm{d}t + \sigma_z \mathrm{d}w_z(t), \quad z(0) = z_0. \end{aligned}$$

A recursive representation of the planner's problem is given by the Hamilton-Jacobi-

[22] We use the fact that $\int_{t=0}^{\infty} e^{-\rho t} \eta t \mathrm{d}t = \frac{\eta}{\rho^2}$ for $\rho > 0$, and hence it is just a constant that we omit without affecting the optimization problem. In discrete-time this is equivalent to omitting $\sum_{t=0}^{\infty} \beta^t \eta t = \frac{\eta\beta}{(\beta-1)^2}$ as long as $\beta \in (0, 1)$.

Bellman (HJB) equation,[23]

$$\rho J(k,z) = \max_{c,n} \Bigg\{ (\ln c + \psi(1-n)) + \left(\exp(z)k^{\alpha}n^{1-\alpha} - c - (\delta+\eta)k\right) J_k(k,z) \\ - \rho_z z J_z(k,z) + \frac{1}{2}\sigma_k^2 k^2 J_{kk}(k,z) + \frac{1}{2}\sigma_z^2 J_{zz}(k,z) \Bigg\}, \quad \text{(E.1)}$$

where subscripts denote partial derivatives. The first order conditions for an interior solution are given by

$$c = \left(J_k(k,z)\right)^{-1},$$
$$\psi = (1-\alpha)\exp(z)k^{\alpha}n^{-\alpha}J_k(k,z),$$

which implicitly define optimal consumption and hours worked as functions of the state variables of the economy, $c = \mathbf{c}(k,z)$ and $n = \mathbf{n}(k,z)$.

The maximized HJB equation reads

$$\rho J(k,z) = \ln \mathbf{c}(k,z) + \psi(1-\mathbf{n}(k,z)) \\ + \left(\exp(z)k^{\alpha}\mathbf{n}(k,z)^{1-\alpha} - \mathbf{c}(k,z) - (\delta+\eta)k\right) J_k(k,z) \\ - \rho_z z J_z(k,z) + \frac{1}{2}\sigma_k^2 k^2 J_{kk}(k,z) + \frac{1}{2}\sigma_z^2 J_{zz}(k,z),$$

from which it is follows that the co-state variable associated with the capital stock must satisfy (using the envelope condition)

$$\rho J_k(k,z) = \left(\alpha \exp(z) k^{\alpha-1}\mathbf{n}(k,z)^{1-\alpha} - (\delta+\eta)\right) J_k(k,z) \\ + \left(\exp(z)k^{\alpha}\mathbf{n}(k,z)^{1-\alpha} - \mathbf{c}(k,z) - (\delta+\eta)k\right) J_{kk}(k,z) \\ - \rho_z z J_{kz}(k,z) + \sigma_z^2 k J_{kk}(k,z) + \frac{1}{2}\sigma_k^2 k^2 J_{kkk}(k,z) + \frac{1}{2}\sigma_z^2 J_{kkz}(k,z).$$

Collecting terms yields

$$\left(\rho - \alpha \exp(z) k^{\alpha-1}\mathbf{n}(k,z)^{1-\alpha} + \delta + \eta\right) J_k(k,z) = \Big(\exp(z)k^{\alpha}\mathbf{n}(k,z)^{1-\alpha} \\ - \mathbf{c}(k,z) - (\delta+\eta)k \Big) J_{kk}(k,z) - \rho_z z J_{kz}(k,z) \\ + \sigma_k^2 k J_{kk}(k,z) + \frac{1}{2}\sigma_k^2 k^2 J_{kkk}(k,z) + \frac{1}{2}\sigma_z^2 J_{kkz}(k,z). \quad \text{(E.2)}$$

[23] See Chang (2009) for a formal derivation of the HJB equation.

Using Ito's formula, the co-state variable evolves according to

$$\begin{aligned}\mathrm{d}J_k\left(k,z\right)=&\Bigg[\left(\exp(z)k^{\alpha}n^{1-\alpha}-c-\left(\delta+\eta\right)k\right)J_{kk}\left(k,z\right)\\&-\rho_z z J_{kz}\left(k,z\right)+\frac{1}{2}\sigma_k^2k^2J_{kkk}\left(k,z\right)+\frac{1}{2}\sigma_z^2J_{kkz}\left(k,z\right)\Bigg]\mathrm{d}t\\&+\sigma_k kJ_{kk}\left(k,z\right)\mathrm{d}w_k+\sigma_z J_{kz}\left(k,z\right)\mathrm{d}w_z,\end{aligned}$$

where substituting for the optimal costate in (E.2), we obtain the equilibrium dynamics of the marginal utility of consumption

$$\begin{aligned}\mathrm{d}J_k\left(k,z\right)=&\Bigg[\left(\rho-\alpha\exp\left(z\right)k^{\alpha-1}n^{1-\alpha}+\delta+\eta\right)J_k\left(k,z\right)-\sigma_k^2kJ_{kk}\left(k,z\right)\Bigg]\mathrm{d}t\\&+\sigma_k kJ_{kk}\left(k,z\right)\mathrm{d}w_k+\sigma_z J_{kz}\left(k,z\right)\mathrm{d}w_z.\end{aligned}\tag{E.3}$$

After some algebra, one obtains the Euler equation for consumption:

$$\begin{aligned}\frac{\mathrm{d}c}{c}=&\Bigg[\left(\alpha\exp\left(z\right)k^{\alpha-1}n^{1-\alpha}-\rho-\delta-\eta\right)-\sigma_k^2\frac{k\mathbf{c}_k\left(k,z\right)}{c}\\&+\frac{1}{2}\left(\sigma_k^2\left(\frac{k\mathbf{c}_k\left(k,z\right)}{c}\right)^2+\sigma_z^2\left(\frac{\mathbf{c}_z\left(k,z\right)}{c}\right)^2\right)\Bigg]\mathrm{d}t\\&+\sigma_k\frac{k\mathbf{c}_k\left(k,z\right)}{c}\mathrm{d}w_k+\sigma_z\frac{\mathbf{c}_z\left(k,z\right)}{c}\mathrm{d}w_z.\end{aligned}\tag{E.4}$$

Given the properties of stochastic integrals for Brownian motions, the Euler equation for consumption can be alternatively written in expected terms as

$$\begin{aligned}\frac{1}{\mathrm{d}t}\mathbb{E}_t\left[\frac{\mathrm{d}c}{c}\right]=&\left(\alpha\exp\left(z\right)k^{\alpha-1}\mathbf{n}\left(k,z\right)^{1-\alpha}-\rho-\delta-\eta\right)-\sigma_k^2\frac{k\mathbf{c}_k\left(k,z\right)}{c}\\&+\frac{1}{2}\sigma_k^2\left(\frac{k\mathbf{c}_k\left(k,z\right)}{c}\right)^2+\frac{1}{2}\sigma_z^2\left(\frac{\mathbf{c}_z\left(k,z\right)}{c}\right)^2.\end{aligned}\tag{E.5}$$

E.2 Equilibrium

The general equilibrium in this economy can be characterized in the time-domain by the following system of nonlinear stochastic differential equations:

$$\mathbb{E}_t\left[\frac{\mathrm{d}c}{c}\right] = \left[\left(\alpha \exp\left(z\right) k^{\alpha-1} n^{1-\alpha} - \rho - \delta - \eta\right) - \sigma_k^2 \frac{k\mathbf{c}_k\left(k,z\right)}{c}\right. \tag{E.6}$$
$$\left. + \frac{1}{2}\left(\sigma_k^2\left(\frac{k\mathbf{c}_k\left(k,z\right)}{c}\right)^2 + \sigma_z^2\left(\frac{\mathbf{c}_z\left(k,z\right)}{c}\right)^2\right)\right]\mathrm{d}t,$$

$$\mathrm{d}k = \left(\exp\left(z\right) k^{\alpha} n^{1-\alpha} - c - \left(\delta+\eta\right)k\right)\mathrm{d}t + \sigma_k k \mathrm{d}w_k, \quad k\left(0\right) = k_0, \tag{E.7}$$

$$\mathrm{d}z = -\rho_z z \mathrm{d}t + \sigma_z \mathrm{d}w_z, \quad z\left(0\right) = z_0, \tag{E.8}$$

together with the algebraic (static) equation for the optimal fraction of hours worked,

$$\psi c n = \left(1-\alpha\right)\exp(z) k^{\alpha} n^{1-\alpha}. \tag{E.9}$$

Collecting the model variables in the vector $\check{\mathbf{x}} = \left[c, k, z, n\right]^{\top}$, and using the properties of stochastic integrals for Brownian motions, we compactly write the nonlinear equilibrium as

$$\mathrm{d}\check{\mathbf{x}}(t) = \mathbf{G}_0\left(\check{\mathbf{x}}(t)\right)\mathrm{d}t + \mathbf{G}_1\left(\check{\mathbf{x}}(t)\right)\mathrm{d}\mathbf{w}(t) + \widetilde{\mathbf{\Pi}}\mathrm{d}\varepsilon(t), \tag{E.10}$$

where $\mathbf{w}(t) = \left[w_k(t), w_z(t)\right]^{\top}$ is the vector of structural shocks, $\varepsilon(t)$ is an expectation error defined as the difference between the actual and unexpected change in consumption, i.e., $\mathrm{d}\varepsilon(t) = \mathbb{E}_t\left[\mathrm{d}c(t)\right] - \mathrm{d}c(t)$, satisfying $\mathbb{E}_t\left[\mathrm{d}\varepsilon(t)\right] = 0$, and $\widetilde{\mathbf{\Pi}}$ is a selection matrix.

E.3 Deterministic steady state

In the absence of uncertainty, the (de-trended) economy converges over time to a fixed point, or steady-state equilibrium, in which all variables are idle. We denote such point by $\check{\mathbf{x}}^{\star} = \left(c^{\star}, n^{\star}, k^{\star}, z^{\star}\right)^{\top}$. Therefore, imposing $\sigma_k = \sigma_z = 0$ together with the no-growth condition $\mathrm{d}\check{\mathbf{x}}(t)/\mathrm{d}t = \mathbf{0}$ to the system (E.10) yields

$$z^{\star} = 0, \qquad n^{\star} = \left(1-\alpha\right)\left(\psi\left(1 - \tfrac{\alpha(\delta+\eta)}{\rho+\delta+\eta}\right)\right)^{-1},$$

$$k^{\star} = \left(\frac{\alpha}{\rho+\delta+\eta}\right)^{\frac{1}{1-\alpha}} n^{\star}, \text{ and } c^{\star} = \left(k^{\star}\right)^{\alpha}\left(n^{\star}\right)^{1-\alpha} - \left(\delta+\eta\right)k^{\star}.$$

E.4 Log-linearized equilibrium

The nonlinear system formed by (E.6)-(E.9) can be linearized in order to study the dynamic behavior of the stationary variables as they fluctuate in close proximity of their deterministic steady-state values. Let $\hat{c} = \ln c - \ln c^{\star}$, $\hat{n} = \ln n - \ln n^{\star}$, $\hat{k} = \ln k - \ln k^{\star}$ and $\hat{z} = z - z^{\star}$ denote log-deviations of the variables with respect to their steady-state values. Then, a first-order Taylor expansion of (E.10) yields

$$\begin{bmatrix} \mathrm{d}\hat{c} \\ \mathrm{d}\hat{k} \\ \mathrm{d}\hat{z} \\ 0 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & \xi_{ck} & \xi_{cz} & \xi_{cn} \\ \xi_{kc} & \xi_{kk} & \xi_{kz} & \xi_{kn} \\ 0 & 0 & -\rho_z & 0 \\ \xi_{nc} & \xi_{nk} & \xi_{nz} & -1 \end{bmatrix}}_{\equiv \widetilde{\mathbf{\Gamma}}} \begin{bmatrix} \hat{c} \\ \hat{k} \\ \hat{z} \\ \hat{n} \end{bmatrix} \mathrm{d}t + \underbrace{\begin{bmatrix} 0 & 0 \\ \sigma_k & 0 \\ 0 & \sigma_z \\ 0 & 0 \end{bmatrix}}_{\equiv \widetilde{\mathbf{\Psi}}} \begin{bmatrix} \mathrm{d}w_k \\ \mathrm{d}w_z \end{bmatrix} + \underbrace{\begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{\equiv \widetilde{\mathbf{\Pi}}} \mathrm{d}\varepsilon$$

where $\widetilde{\mathbf{\Gamma}}$ is the Jacobian matrix of the log-transformed equilibrium evaluated at the deterministic steady state, and $\widetilde{\mathbf{\Psi}}$ is the corresponding diffusion matrix. The log-transformation is obtained via an application of Itô's formula to (E.10). The coefficients in $\widetilde{\mathbf{\Gamma}}$ are given by $\xi_{ck} = (\alpha-1)(\rho+\delta+\eta)$, $\xi_{cz} = (\rho+\delta+\eta)$, $\xi_{cn} = (1-\alpha)(\rho+\delta+\eta)$, $\xi_{kc} = -(\rho+(1-\alpha)(\delta+\eta))/\alpha$, $\xi_{kk} = \rho$, $\xi_{kz} = (\rho+\delta+\eta)/\alpha$, $\xi_{kn} = (1-\alpha)(\rho+\delta+\eta)/\alpha$, $\xi_{nc} = -1/\alpha$, $\xi_{nk} = 1$ and $\xi_{nz} = 1/\alpha$.

Next, we substitute out the intratemporal labor supply condition $\hat{n} = \xi_{nc}\hat{c} + \xi_{nk}\hat{k} + \xi_{nz}\hat{z}$, to obtained a linearized equilibrium consisting of the 3×3 system of linear stochastic differential equations

$$\begin{bmatrix} \mathrm{d}\hat{c} \\ \mathrm{d}\hat{k} \\ \mathrm{d}\hat{z} \end{bmatrix} = \underbrace{\begin{bmatrix} \xi_{cn}\xi_{nc} & 0 & \xi_{cz} + \xi_{cn}\xi_{nz} \\ \xi_{kc} + \xi_{kn}\xi_{nc} & \xi_{kk} + \xi_{kn}\xi_{nk} & \xi_{kz} + \xi_{kn}\xi_{nz} \\ 0 & 0 & -\rho_z \end{bmatrix}}_{\equiv \mathbf{\Gamma}} \begin{bmatrix} \hat{c} \\ \hat{k} \\ \hat{z} \end{bmatrix} \mathrm{d}t$$

$$+ \underbrace{\begin{bmatrix} 0 & 0 \\ \sigma_k & 0 \\ 0 & \sigma_z \end{bmatrix}}_{\equiv \mathbf{\Psi}} \begin{bmatrix} \mathrm{d}w_k \\ \mathrm{d}w_z \end{bmatrix} + \underbrace{\begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}}_{\equiv \mathbf{\Pi}} \mathrm{d}\varepsilon,$$

which can be compactly written as

$$\mathrm{d}\tilde{\mathbf{x}}(t) = \mathbf{\Gamma}\tilde{\mathbf{x}}(t)\mathrm{d}t + \mathbf{\Psi}\mathrm{d}\mathbf{w}(t) + \mathbf{\Pi}\mathrm{d}\varepsilon(t), \tag{E.11}$$

where $\tilde{\hat{\mathbf{x}}} = [\hat{c}, \hat{k}, \hat{z}]^{\top}$ denotes the vector of variables in deviations from their deterministic steady state, and where $\mathbf{\Gamma}$, $\mathbf{\Psi}$, and $\mathbf{\Pi}$ are the adjusted versions of $\widetilde{\mathbf{\Gamma}}$, $\widetilde{\mathbf{\Psi}}$, and $\widetilde{\mathbf{\Pi}}$. Note that

the volatility parameters σ_k and σ_z do not affect the matrix $\mathbf{\Gamma}$ that characterizes the endogenous persistence in the linearized equilibrium system. Therefore, they will not have any effects on the implied optimal decision rules, and hence our approximated solution exhibits certainty equivalent in the sense of Simon (1956) and Theil (1957) (see Ahn et al., 2018; Parra-Alvarez et al., 2021).

E.5 Rational expectation solution

Following Sims (2002), let us assume that the matrix $\mathbf{\Gamma}$ can be diagonalized according to

$$\mathbf{\Gamma} = \mathbf{T}\mathbf{\Upsilon}\mathbf{T}^{-1}, \tag{E.12}$$

where $\mathbf{T}$ is a 3×3 matrix of right-eigenvectors of $\mathbf{\Gamma}$, and $\mathbf{\Upsilon}$ is a diagonal matrix whose diagonal elements are the eigenvalues of $\mathbf{\Gamma}$. Premultiplying (E.11) by $\mathbf{T}^{-1}$ and defining $\mathbf{z}(t) = \mathbf{T}^{-1}\tilde{\mathbf{x}}(t)$ yields

$$\mathrm{d}\mathbf{z}(t) = \mathbf{\Upsilon}\mathbf{z}(t)\mathrm{d}t + \mathbf{T}^{-1}\mathbf{\Psi}\mathrm{d}\mathbf{w}(t) + \mathbf{T}^{-1}\mathbf{\Pi}\mathrm{d}\varepsilon(t). \tag{E.13}$$

The eigenvalues of the matrix $\mathbf{\Gamma}$ solve the characteristic equation $|\mathbf{\Gamma} - \upsilon\mathbf{I}_3| = 0$. Thus, it follows that the eigenvalues of $\mathbf{\Gamma}$ are given by $\upsilon_1 = -\rho_z$ and the roots of the quadratic equation

$$a_0\upsilon^2 + a_1\upsilon + a_2 = 0,$$

with $a_0 = 1$, $a_1 = -\left(\xi_{cn}\xi_{nc} + \xi_{kk} + \xi_{kn}\xi_{nk}\right)$, and

$$a_2 = \left(\xi_{cn}\xi_{nc}\xi_{kk} + \xi_{cn}\xi_{nc}\xi_{kn}\xi_{nk}\right).$$

After some algebra, it is possible to show that

$$a_1 = -\rho < 0, \quad \text{and} \quad a_2 = -\frac{(1-\alpha)\left(\rho+\delta+\eta\right)}{\alpha}\left(\frac{\rho+(1-\alpha)\left(\delta+\eta\right)}{\alpha}\right) < 0.$$

Since $a_2^2 - 4a_0a_1 > 0$ (the discriminant of the quadratic equation) and $a_2 < 0$, the quadratic equation has two distinct real roots of opposite sign given by

$$\upsilon_2 = -\frac{(1-\alpha)(\delta+\eta+\rho)}{\alpha} < 0 \text{ and } \upsilon_3 = \frac{(1-\alpha)(\delta+\eta)+\rho}{\alpha} > 0.$$

Hence, the linearized system has two stable roots (non-positive eigenvalues, υ_1 and υ_2) and one unstable root (positive eigenvalue, υ_3). Since the reduced model in (E.11) has two state

variables and one control/jump variable, the Blanchard and Kahn conditions are satisfied, and the model has a unique rational expectation solution (see Buiter, 1984). The eigenvectors of $\mathbf{\Gamma}$ associated to each of its eigenvalues are given by multiples of the following vectors

$$T_1 = \frac{1}{\iota}\begin{bmatrix} -(\delta+\eta+\rho) \\ -\frac{(\delta+\eta+\rho)((1-\alpha)(\delta+\eta)+\rho+\rho_z)}{(1-\alpha)(\delta+\eta)+\rho+\alpha\rho_z} \\ \iota \end{bmatrix}, T_2 = \begin{bmatrix} \frac{\alpha(2(1-\alpha)(\delta+\eta)+(2-\alpha)\rho)}{(1-\alpha^2)(\delta+\eta)+\rho} \\ 1 \\ 0 \end{bmatrix}, \text{ and } T_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},$$

where $\iota = (\alpha-1)(\delta+\eta+\rho)+\alpha\rho_z$.

Let $\mathbf{M}_+$ be a 1×3 vector that selects the rows of $\mathbf{T}^{-1}$ corresponding to eigenvalues with positive real parts, and $\mathbf{M}_-$ a 2×3 matrix that selects the rows of $\mathbf{T}^{-1}$ corresponding to eigenvalues with non-positive real parts. It follows that $\left(\mathbf{M}_+^\top\mathbf{M}_+ + \mathbf{M}_-^\top\mathbf{M}_-\right) = \mathbf{I}_3$. Then

$$\mathbf{M}_+ \mathrm{d}\mathbf{z}(t) = \mathbf{M}_+\mathbf{\Upsilon}\mathbf{z}(t)\mathrm{d}t + \mathbf{M}_+\mathbf{T}^{-1}\mathbf{\Psi}\mathrm{d}\mathbf{w}(t) + \mathbf{M}_+\mathbf{T}^{-1}\mathbf{\Pi}\mathrm{d}\varepsilon(t), \tag{E.14}$$

defines the equation associated with the unstable eigenvalue. To rule out explosive paths, i.e., to ensure that $\lim_{s\to\infty}\mathbb{E}_t\left[\mathbf{z}(s)\right] < \infty$ for $s > t$, and thus to satisfy the model's transversality conditions, we impose

$$\mathbf{M}_+\mathbf{z}(t) = 0, \ \forall t, \tag{E.15}$$

implying that

$$\mathrm{d}\varepsilon(t) = -\left[\mathbf{M}_+\mathbf{T}^{-1}\mathbf{\Pi}\right]^{-1}\mathbf{M}_+\mathbf{T}^{-1}\mathbf{\Psi}\mathrm{d}\mathbf{w}(t). \tag{E.16}$$

In other words, the stability condition imposes an exact relationship between the vector of structural shocks and the expectation error, such that the system does not exhibit explosive paths.

Once we impose the stability conditions (E.15) and (E.16), it is possible to compute the solution associated with the stable eigenvalues by computing

$$\mathbf{M}_- \mathrm{d}\mathbf{z}(t) = \mathbf{M}_-\mathbf{\Upsilon}\mathbf{z}(t)\mathrm{d}t + \mathbf{M}_-\mathbf{T}^{-1}\mathbf{\Psi}\mathrm{d}\mathbf{w}(t) + \mathbf{M}_-\mathbf{T}^{-1}\mathbf{\Pi}\mathrm{d}\varepsilon(t),$$

which in turn implies that

$$\mathrm{d}\mathbf{z}(t) = \mathbf{\Upsilon}^\star\mathbf{z}(t)\mathrm{d}t + \mathbf{\Psi}^\star\mathrm{d}\mathbf{w}(t), \tag{E.17}$$

where $\mathbf{\Upsilon}^\star = \mathbf{M}_-^\top\mathbf{M}_-\mathbf{\Upsilon}\mathbf{M}_-^\top\mathbf{M}_-$ is the 3×3 matrix of eigenvalues with zeros in the position of the explosive paths, and

$$\mathbf{\Psi}^\star = \mathbf{M}_-^\top\mathbf{M}_-\mathbf{T}^{-1}\left[\mathbf{I}_3 - \mathbf{\Pi}\left[\mathbf{M}_+\mathbf{T}^{-1}\mathbf{\Pi}\right]^{-1}\mathbf{M}_+\mathbf{T}^{-1}\right]\mathbf{\Psi}$$

is a 3×2 matrix. Finally, we use the definition $\mathbf{z}(t) = \mathbf{T}^{-1}\tilde{\hat{\mathbf{x}}}(t)$ to recover the autoregressive representation of the rational expectation solution in the original variables

$$\mathrm{d}\tilde{\hat{\mathbf{x}}}(t) = \widetilde{\mathbf{A}}\tilde{\hat{\mathbf{x}}}(t)\mathrm{d}t + \widetilde{\mathbf{B}}\mathrm{d}\mathbf{w}(t) \tag{E.18}$$

where $\widetilde{\mathbf{A}} = \mathbf{T}\boldsymbol{\Upsilon}^{\star}\mathbf{T}^{-1}$ and $\widetilde{\mathbf{B}} = \mathbf{T}\boldsymbol{\Psi}^{\star}$.

From the stability condition (E.15) and the definition of the transformed variable $\mathbf{z}(t)$, we recover the optimal policy for consumption as

$$\hat{c}(t) = \phi_{ck}\hat{k}(t) + \phi_{cz}\hat{z}(t), \tag{E.19}$$

where $\phi_{ck} = -(T_{21}/T_{11})$ and $\phi_{cz} = -(T_{31}/T_{11})$, with T_{ij} the (i,j)-th element of the matrix of $\mathbf{T}^{-1}$. Using the linearized condition for hours worked, the optimal policy for labor is given by

$$\hat{n}(t) = \phi_{nk}\hat{k}(t) + \phi_{nz}\hat{z}(t), \tag{E.20}$$

where $\phi_{nk} = \left(\xi_{nk} - \xi_{nc}\frac{T_{21}}{T_{11}}\right)$ and $\phi_{nz} = \left(\xi_{nz} - \xi_{nc}\frac{T_{31}}{T_{11}}\right)$.

As a final step, we eliminate the dependence of the system in (E.18) on the control variables to obtain and system of SDEs that only describes the optimal dynamics of the state variables. After some algebra we obtain

$$\begin{bmatrix} \mathrm{d}\hat{k}(t) \\ \mathrm{d}\hat{z}(t) \end{bmatrix} = \begin{bmatrix} \phi_{kk} & \phi_{kz} \\ 0 & -\rho_z \end{bmatrix} \begin{bmatrix} \hat{k}(t) \\ \hat{z}(t) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} \sigma_k & 0 \\ 0 & \sigma_z \end{bmatrix} \begin{bmatrix} \mathrm{d}w_k(t) \\ \mathrm{d}w_z(t) \end{bmatrix}, \tag{E.21}$$

with

$$\begin{aligned} \phi_{kk} &= -(\tilde{a}_{21}T_{12})/T_{11} + \tilde{a}_{22} = -\frac{(1-\alpha)(\delta+\eta+\rho)}{\alpha} < 0, \\ \phi_{kz} &= -(\tilde{a}_{21}T_{13})/T_{11} + \tilde{a}_{23} = \frac{(\delta+\eta+\rho)((1-\alpha)(\delta+\eta)+\rho+\rho_z)}{\alpha((1-\alpha)(\delta+\eta)+\rho+\alpha\rho_z)} > 0, \end{aligned}$$

where $\tilde{a}_{ij}$ is the (i,j)-th element of the matrix of $\tilde{\mathbf{A}}$.

Let $\hat{\mathbf{y}}_t = [\hat{c}_t, \hat{n}_t]^\top$ and $\hat{\mathbf{x}}_t = [\hat{k}_t, \hat{z}_t]^\top$ denote, respectively, the vector of control and state variables in log-deviations from their steady state values. Then (E.21), together with (E.19) and (E.20), have the continuous-time state space representation in (2.1) and (2.2), i.e.,

$$\begin{aligned} \mathrm{d}\hat{\mathbf{x}}(t) &= \mathbf{A}\hat{\mathbf{x}}(t)\mathrm{d}t + \mathbf{B}\mathrm{d}\mathbf{w}(t), \\ \hat{\mathbf{y}}(t) &= \mathbf{C}\hat{\mathbf{x}}(t). \end{aligned}$$

E.6 Proof of Lemma 2

Let the economy follow dynamics of the state-space system (4.12), where the transition equation is given by (E.21). The state-space system results from the rational expectation solution of the equilibrium (4.8), (4.9), (4.10), and (4.6). Then, Assumption 1 implies that the eigenvalues of $\mathbf{A}(\boldsymbol{\theta})$ have strictly negative real part. For every $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, $\mathbf{A}(\boldsymbol{\theta})$ is a real matrix. Then, in order to rule out aliases of $\mathbf{A}(\boldsymbol{\theta})$, it is left to verify that no Jordan block of $\mathbf{A}(\boldsymbol{\theta})$ occur more than once and that the eigenvalues are strictly real. Notice that $\mathbf{A}(\boldsymbol{\theta})$, the drift matrix of (E.21), is a 2-by-2 matrix, and let $\{\lambda_1, \lambda_2\} := \lambda(\mathbf{A}(\boldsymbol{\theta})) = \{\lambda : \det(\lambda \mathbf{I}_{n_x} - \mathbf{A}(\boldsymbol{\theta}) = 0\}$. A sufficient and necessary condition for the eigenvalues λ_1 and λ_2 to have imaginary part different from zero is that

$$\text{trace}(\mathbf{A}(\boldsymbol{\theta}))^2 < 4 \cdot \det(\mathbf{A}(\boldsymbol{\theta})).$$

In the specific problem, $\lambda_1 = \phi_{kk}$ and $\lambda_2 = -\rho_z$, are both real and negative, $\lambda_1, \lambda_2 < 0$. In general, however, given the triangular structure of $\mathbf{A}(\boldsymbol{\theta})$ and the analytical form of (E.21), complex eigenvalues require that $(\phi_{kk} + \rho_z)^2 < 0$, which is impossible. Additionally, the assumption that $(1-\alpha)(\delta + \eta + \rho) \neq \alpha\rho_z$, ensures that the eigenvalues are distinct, and no Jordan block is therefore repeated.

Assumptions 1 and 3 are satisfied. Namely, the eigenvalues of $\mathbf{A}(\boldsymbol{\theta})$ are distinct (no Jordan blocks occur more than once), real, and negative. Finally, the matrix $\mathbf{A}(\boldsymbol{\theta})$ does not have any aliases, so the entries of $\mathbf{A}(\boldsymbol{\theta})$ are identified from $\mathbf{A}_h(\boldsymbol{\theta})$. Further, $\left(\int_0^h \mathbf{A}_s(\boldsymbol{\theta}) \otimes \mathbf{A}_s(\boldsymbol{\theta}) \text{d}s\right)$ is nonsingular, and identification of $\mathbf{A}(\boldsymbol{\theta})$ implies identification of $\boldsymbol{\Sigma}(\boldsymbol{\theta})$. ■

Remark E.1. The assumption $(1-\alpha)(\delta + \eta + \rho) \neq \alpha\rho_z$ implies Assumption 3, but it is used only for simplifying the proof. In fact, it is stronger than what is necessary for the assumption to hold. To see this, let $\lambda_1 = \lambda_2 = \phi_{kk} < 0$, such that $\phi_{kk} = -\rho_z$ and therefore $(1-\alpha)(\delta+\eta+\rho) = \alpha\rho_z$. Then, given $\mathbf{A}(\boldsymbol{\theta})$ in (4.12), $(\mathbf{A}(\boldsymbol{\theta}) - \phi_{kk}\mathbf{I}_{n_x}) \neq \mathbf{0}$, whereas $(\mathbf{A}(\boldsymbol{\theta}) - \phi_{kk}\mathbf{I}_{n_x})^2 = \mathbf{0}$, indicating that the block related to eigenvalue ϕ_{kk} is 2-by-2. Hence, there is only one individual 2-by-2 Jordan block associated to ϕ_{kk} and Assumption 3 is satisfied.

F. Additional tables and figures

Table F.1. Finite sample properties for extended vector of parameters. The table reports finite sample estimates of $\boldsymbol{\theta} = [\rho, \eta, \rho_z, \sigma_z, \sigma_k]^\top$ from $M = 500$ samples of quarterly ($h = 1/4$) observations on aggregate consumption (C) and hours worked (N), generated over a period of 60 years ($T = 240$ observations in each sample). Simulated measurements in Panel A are sampled as stocks, and those in Panel B as flows. The share of capital in output, α, and the depreciation rate, δ, are calibrated to their population values in Table 1. Let $\hat{\boldsymbol{\theta}}_m$ denote the estimates from the $m-$th sample. The table displays mean bias (Bias $= M^{-1}\sum_{m=1}^{M}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_0)$) and root mean squared errors (RMSE $= (M^{-1}\sum_{m=1}^{M}(\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_0)^2)^{1/2}$) across repetitions for the F-SSR, S-SSR, and EM-SSR models.

Panel A: Data is sampled as stock

		F-SSR		S-SSR		EM-SSR	
	$\boldsymbol{\theta}$	Bias	RMSE	Bias	RMSE	Bias	RMSE
ρ	0.03	-	-	0.0002	0.0015	0.0001	0.0015
η	0.02	-	-	0.0003	0.0021	0.0001	0.0021
ρ_z	0.2052	-	-	0.0044	0.0196	-0.0282	0.0322
σ_z	0.014	-	-	-0.0001	0.0007	-0.0001	0.0007
σ_k	0.0104	-	-	-0.0001	0.0005	-0.0004	0.0006

Panel B: Data is sampled as flows

		F-SSR		S-SSR		EM-SSR	
	$\boldsymbol{\theta}$	Bias	RMSE	Bias	RMSE	Bias	RMSE
ρ	0.03	0.0002	0.0015	1.91E-05	0.0015	-0.0001	0.0015
η	0.02	0.0003	0.0021	1.15E-05	0.0021	-0.0002	0.0021
ρ_z	0.2052	0.0051	0.0206	-0.0008	0.0195	-0.0314	0.0352
σ_z	0.014	-0.0001	0.0007	-0.0026	0.0027	-0.0026	0.0027
σ_k	0.0104	-0.0001	0.0005	-0.0020	0.0020	-0.0022	0.0023

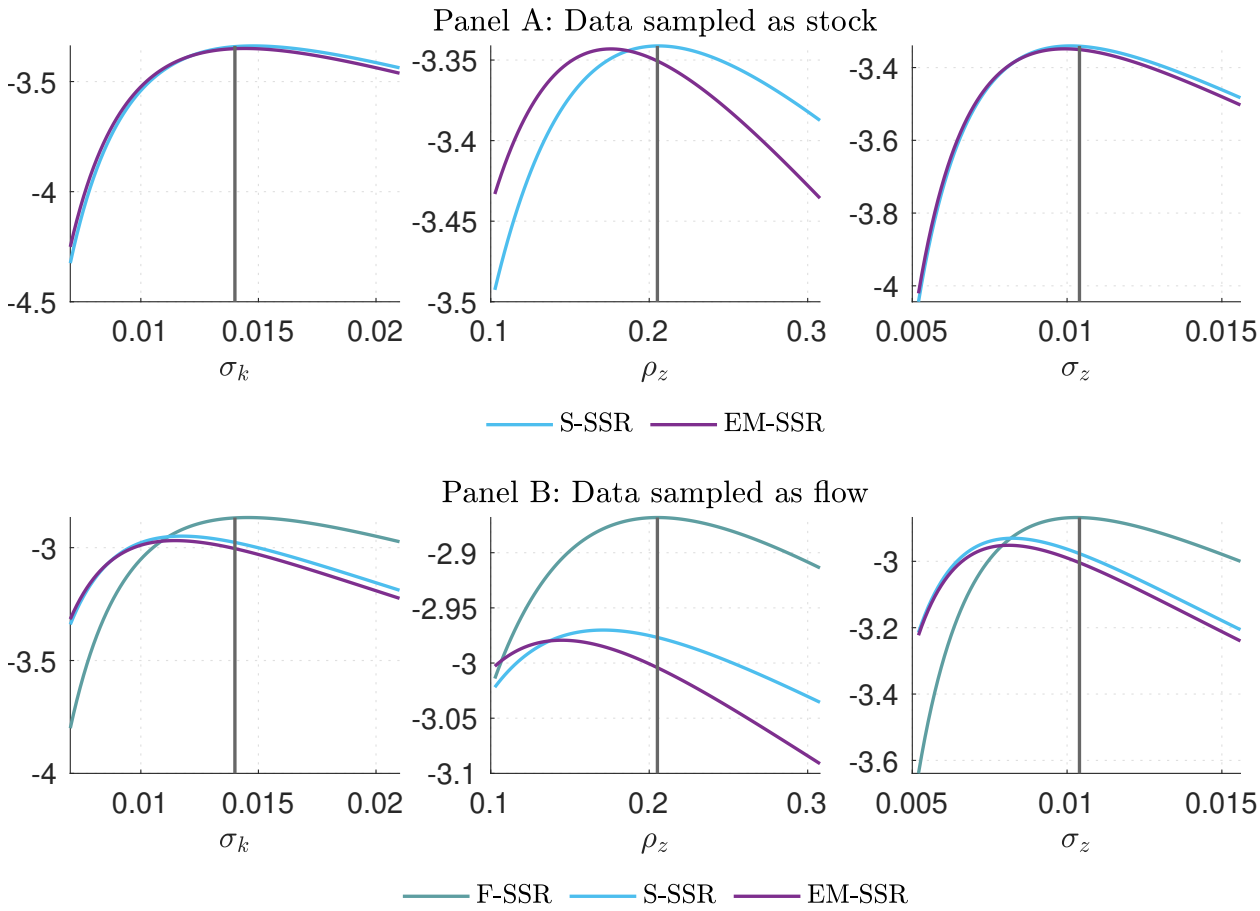


Figure F.1. Log-likelihood profile. The graph shows the log-likelihood function $\mathcal{L}(\boldsymbol{\theta}|\mathbf{y}^T)$ for selected parameters $(\sigma_k, \rho_z, \sigma_z)^\top \in \boldsymbol{\Theta}$, while keeping the remaining parameters at their population values in Table 1. The plots are generated using a single simulated sample of quarterly observations for a period of 60 years. Different samples provide the same conclusions. Panel A corresponds to simulated measurements that have been sampled as stocks, and those in Panel B as flows. The vertical black solid line denotes the true population value of the parameter.

Table F.2. ML estimates. The table reports the ML estimates of $\boldsymbol{\theta}_{\text{exo}} = [\rho_z, \sigma_z, \sigma_k]^\top$ for the model in Section 4 using two additional sets of measurements for the U.S. from 1959:Q1 to 2019:Q4: quarterly data on hours worked, N, and aggregate output, Y; and aggregate consumption, C, and aggregate output, Y. The remaining parameters of the model $\boldsymbol{\theta}_{\text{ss}}$ are fixed to the values in Table 1. Bootstrap standard errors computed from $B = 499$ samples are reported in parentheses.

	N and Y			C and Y		
$\boldsymbol{\theta}_{\text{exo}}$	F-SSR	S-SSR	EM-SSR	F-SSR	S-SSR	EM-SSR
ρ_z	0.0194	4.87E-08	8.44E-07	0.0633	0.0607	0.0565
	(0.0744)	(0.0071)	(0.0079)	(0.0508)	(0.0383)	(0.0342)
σ_z	0.0133	0.0104	0.0104	0.0107	0.0089	0.0089
	(0.0016)	(0.0011)	(0.0010)	(0.0011)	(0.0010)	(0.0009)
σ_k	0.0318	0.0221	0.0217	0.0263	0.0177	0.0174
	(0.0065)	(0.0019)	(0.0019)	(0.0020)	(0.0013)	(0.0012)

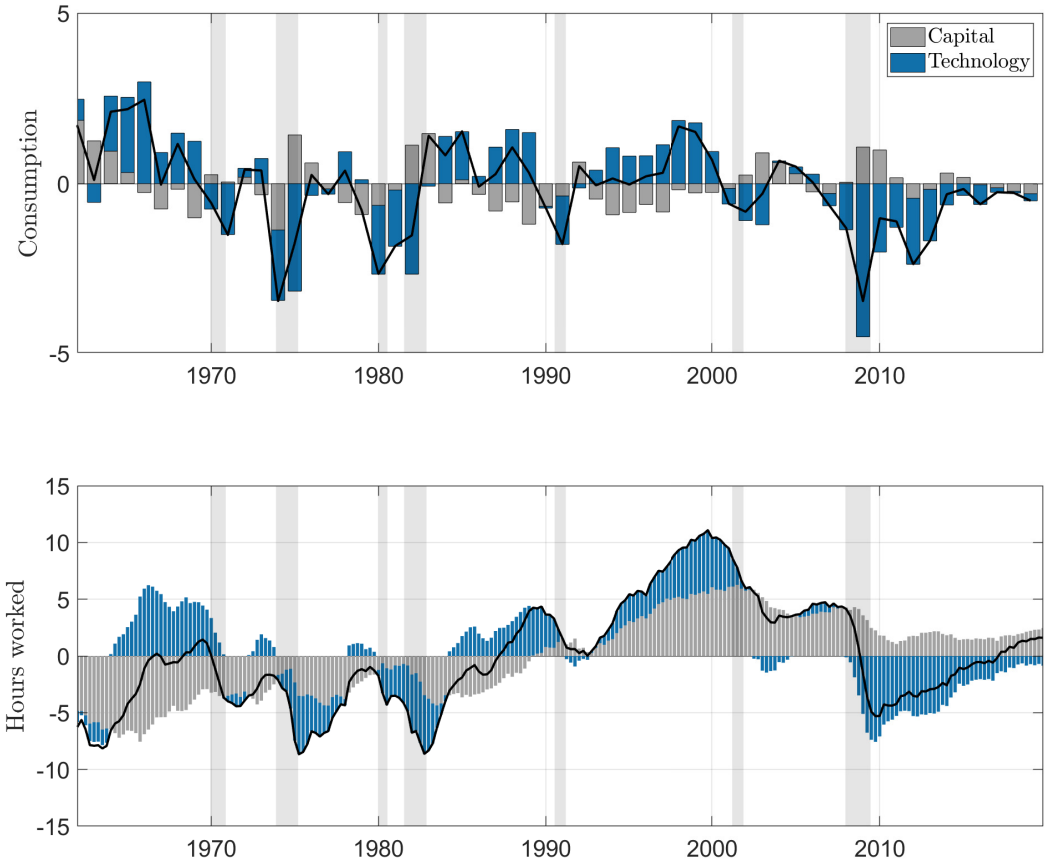


Figure F.2. Historical shock decomposition (S-SSR model). The plot shows the historical contribution of each of the structural shocks recovered from the estimated S-SSR model on the observed measurements over the period 1975:Q1-2019:Q4 (expressed in percentages). The black solid line in the upper panel represents annual consumption growth rates. The black solid line in the lower panel represents quarterly percentage deviations of hours worked from its steady state ($n^\star = 33\%$). NBER recession are reported with the light gray vertical bands.

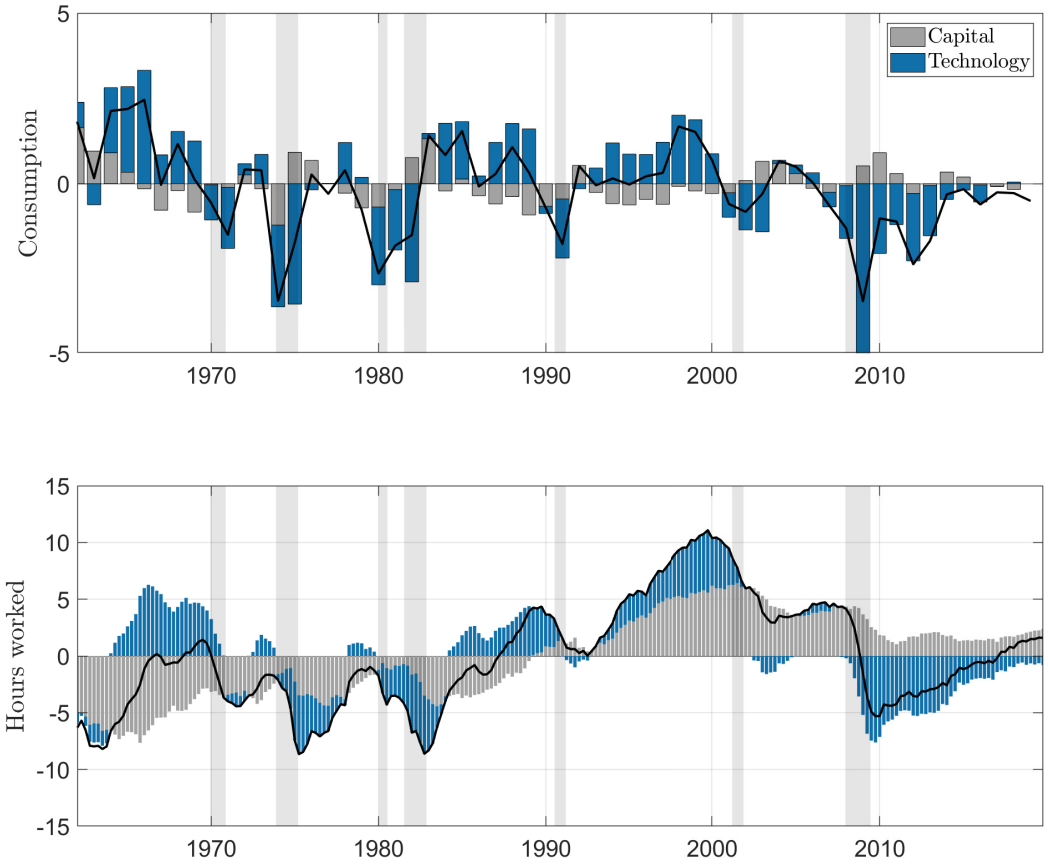


Figure F.3. Historical shock decomposition (EM-SSR model). The plot shows the historical contribution of each of the structural shocks recovered from the estimated EM-SSR model on the observed measurements over the period 1948:Q1-2019:Q4 (expressed in percentages). The black solid line in the upper panel represents annual consumption growth rates. The black solid line in the lower panel represents quarterly percentage deviations of hours worked from its steady state ($n^\star = 33\%$). NBER recession are reported with the light gray vertical bands.

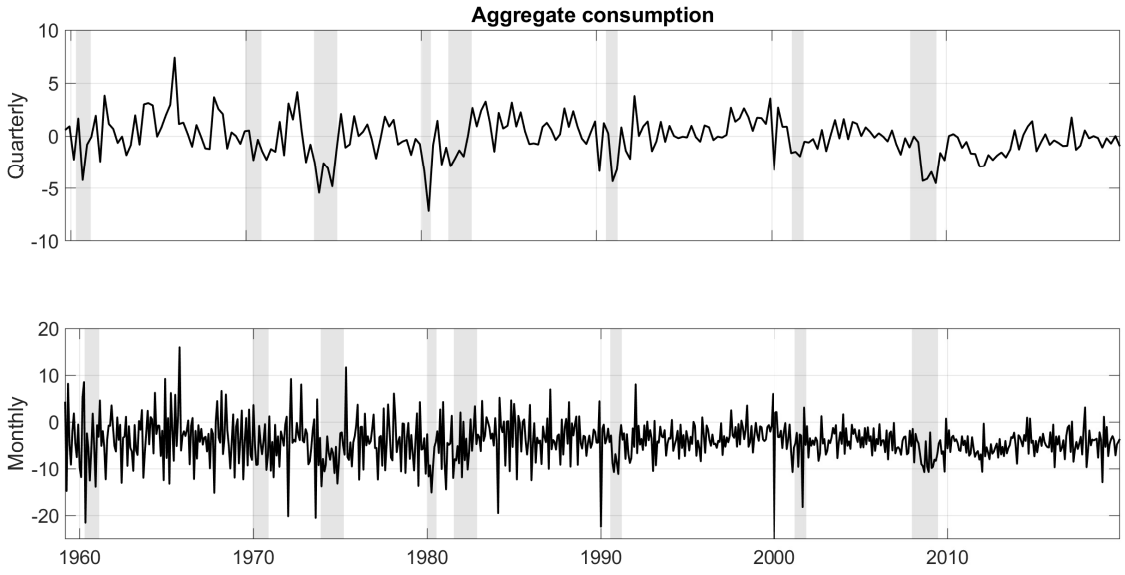


Figure F.4. Aggregate real consumption per capita. The figure plots the annualized growth rate of aggregate real consumption per capita, $100/h \cdot (\log(C_\tau/C_{\tau-1}) - \eta h)$, where h is the sampling frequency. The top panel samples the measurement at a quarterly frequency, $h = 1/4$, while the bottom panel samples at a monthly frequency, $h = 1/12$. In both cases, the sample spans from 1959:M1 to 2019:M12.

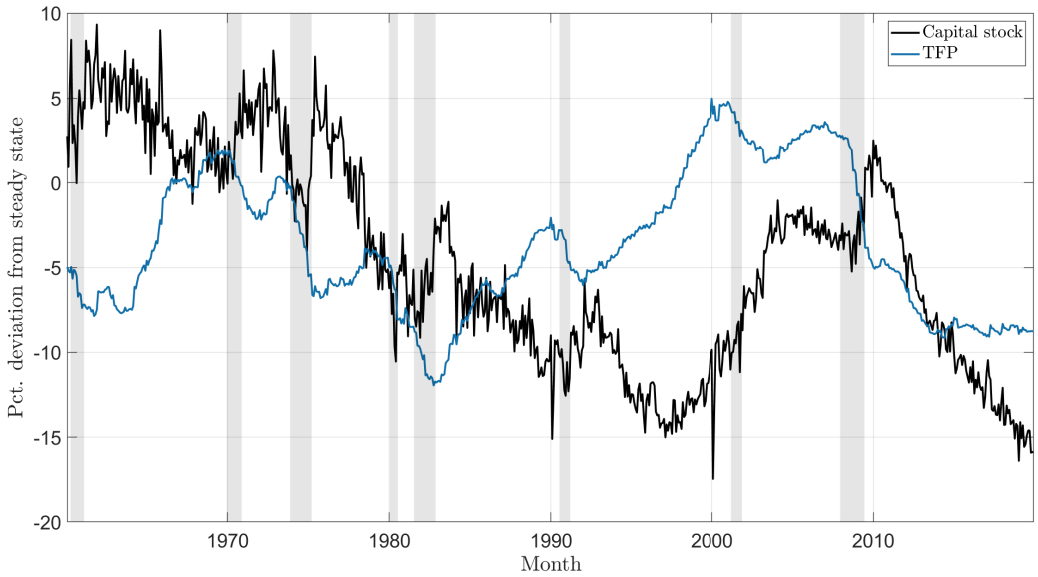


Figure F.5. Latent states, monthly. Filtered monthly series of aggregate capital stock and TFP (as % deviations from their corresponding steady-state values). Sample spans from January 1960 to December 2019. The series are obtained from the MXF-SSR model (C.3)-(C.4) after parameter estimation and filtering using as measurements monthly real PCE and quarterly hours worked.

Table F.3. Normality tests. Battery of normality tests for the estimated F-SSR model without measurement error (see Table 3, first column). The table reports the values of the test statistics and their associated p-values (in parenthesis) for three tests that assess the normality of the prediction errors, $\boldsymbol{\omega}_{\tau|\tau-1}$. The tests are performed over different subsamples of the sample used in the estimation, as indicated in the first column. The second and third columns report the results from a Shapiro-Wilk test (SW, Shapiro and Wilk, 1965) on the univariate normality of the prediction errors associated with consumption, ω_C, and hours worked, ω_N. The fourth column reports the results from the omnibus asymptotic multivariate normality test in (BS, Bowman and Shenton, 1975). The last column reports the results from the omnibus approximate normality test in (DH, Doornik and Hansen, 2008). All the tests evaluate the null hypothesis that the underlying data is normally distributed with a given mean and variance. P-values larger than 0.05 indicate that the test fails to reject the null hypothesis at the 5% level of significance.

Period	SW, ω_C	SW, ω_N	BS	DH
01-Jun-1960 - 01-Dec-1966	0.97	0.94	1.79	2.77
	(0.51)	(0.15)	(0.77)	(0.60)
01-Dec-1966 - 01-Jun-1973	0.92	0.93	5.30	9.37
	(0.04)	(0.08)	(0.26)	(0.05)
01-Jun-1973 - 01-Mar-1980	0.99	0.87	16.96	17.20
	(0.95)	(0.00)	(0.00)	(0.00)
01-Mar-1980 - 01-Sep-1986	0.94	0.94	4.44	6.69
	(0.11)	(0.16)	(0.35)	(0.15)
01-Sep-1986 - 01-Jun-1993	0.96	0.94	1.68	4.78
	(0.34)	(0.14)	(0.79)	(0.31)
01-Jun-1993 - 01-Dec-1999	0.93	0.96	2.89	4.49
	(0.07)	(0.43)	(0.58)	(0.34)
01-Dec-1999 - 01-Sep-2006	0.95	0.95	13.17	11.00
	(0.22)	(0.19)	(0.01)	(0.03)
01-Sep-2006 - 01-Mar-2013	0.97	0.93	1.95	3.88
	(0.52)	(0.07)	(0.74)	(0.42)
01-Mar-2013 - 01-Dec-2019	0.96	0.98	1.72	3.18
	(0.40)	(0.81)	(0.79)	(0.53)

Research Papers

2021-12: Mikkel Bennedsen, Asger Lunde, Neil Shephard and Almut E. D. Veraart: Inference and forecasting for continuous-time integer-valued trawl processes and their use in financial economics

2021-13: Anthony D. Hall, Annastiina Silvennoinen and Timo Teräsvirta: Four Australian Banks and the Multivariate Time-Varying Smooth Transition Correlation GARCH model

2021-14: Ulrich Hounyo and Kajal Lahiri: Estimating the Variance of a Combined Forecast: Bootstrap-Based Approach

2021-15: Salman Huseynov: Long and short memory in dynamic term structure models

2022-01: Jian Kang, Johan Stax Jakobsen, Annastiina Silvennoinen, Timo Teräsvirta and Glen Wade: A parsimonious test of constancy of a positive definite correlation matrix in a multivariate time-varying GARCH model

2022-02: Javier Haulde and Morten Ørregaard Nielsen: Fractional integration and cointegration

2022-03: Yue Xu: Spillovers of Senior Mutual Fund Managers' Capital Raising Ability

2022-04: Morten Ørregaard Nielsen, Wonk-ki Seo and Dakyung Seong: Inference on the dimension of the nonstationary subspace in functional time series

2022-05: Kristoffer Pons Bertelsen: The Prior Adaptive Group Lasso and the Factor Zoo

2022-06: Ole Linnemann Nielsen and Anders Merrild Posselt: Betting on mean reversion in the VIX? Evidence from ETP flows

2022-07: Javier Hualde and Morten Ørregaard Nielsen: Truncated sum-of-squares estimation of fractional time series models with generalized power law trend

2022-08: James MacKinnon and Morten Ørregaard Nielsen: Cluster-Robust Inference: A Guide to Empirical Practice

2022-09: Mikkel Bennedsen, Eric Hillebrand and Sebastian Jensen: A Neural Network Approach to the Environmental Kuznets Curve

2022-10: Yunus Emre Ergemen: Parametric Estimation of Long Memory in Factor Models

2022-11: Yue Xu: Reallocation of Mutual Fund Managers and Capital Raising Ability

2022-12: Bent Jesper Christensen, Luca Neri and Juan Carlos Parra-Alvarez: Estimation of continuous-time linear DSGE models from discrete-time measurements