Estimating Dynamic Equilibrium Models using Mixed Frequency Macro and Financial Data*

Bent Jesper Christensen[a,b], Olaf Posch[a,b]†, and Michel van der Wel[b,c]
[a]Aarhus University, [b]CREATES,
[c]Erasmus University Rotterdam

February 15, 2013

Abstract

We show that including financial market data at daily frequency, along with macro series at standard lower frequency, facilitates statistical inference on structural parameters in dynamic equilibrium models. Our continuous-time formulation conveniently accounts for the difference in observation frequency. We suggest two approaches for the estimation of structural parameters. The first is a simple regression-based procedure for estimation of the reduced-form parameters of the model, combined with a minimum-distance method for identifying the structural parameters. The second approach uses martingale estimating functions to estimate the structural parameters directly through a non-linear optimization scheme. We illustrate both approaches by estimating the stochastic AK model with mean-reverting spot interest rates. We also provide Monte Carlo evidence on the small sample behavior of the estimators and estimate the model using 20 years of U.S. macro and financial data.

JEL classification: C13; E32; O40

Keywords: Structural estimation; AK-Vasicek model; Martingale estimating function

*We are grateful to Anton Braun, Fabio Canova, Andrew Harvey, Hashem Pesaran, Juan Rubio-Ramírez, Frank Smets, and seminar participants in Mannheim, Groningen, and at the 'Workshop on Methods and Applications for DSGE Models' in Atlanta for useful comments. The authors appreciate financial support from the Center for Research in Econometric Analysis of TimE Series (CREATES), which is funded by the Danish National Research Foundation, and the Danish Social Science Research Council.

†Corresponding author: Olaf Posch (Work: +45 8942 1938, Email address: oposch@econ.au.dk, Address: Aarhus University, School of Economics and Management, Building 1322, Bartholins Allé 10, 8000 Århus C).

1 Introduction

Dynamic stochastic general equilibrium (DSGE) models have become the workhorse in modern macroeconomics, successfully capturing aggregate dynamics over the business cycle. Compared to the importance and relevance of this research area, surprisingly little work has been published reconciling business cycle facts with asset market implications (previous results can be found in Grinols and Turnovsky, 1993; Jermann, 1998; Tallarini, 2000; Lettau and Uhlig, 2000; Boldrin, Christiano, and Fisher, 2001; Lettau, 2003, and more recent work includes Rudebusch and Swanson, 2008; Campanale, Castro, and Clementi, 2010). These papers mainly use calibration methods rather than structural estimation for their results. Moreover, the lack of financial variables in DSGE models became one of the most obvious shortcomings of macroeconomic theory (and the theory-based estimation of those systems) during the recent financial crisis, and led to fundamental critique.[1]

No clear answer has been given so far to the questions of how macro and financial data should be linked consistently within dynamic stochastic general equilibrium models, and how they can be used efficiently to shed light on macro-finance links. Financial market data are typically available at higher frequency and better quality than aggregate macroeconomic data. Hence, financial markets provide an additional source of evidence on the state of the economy, beyond macro series. Nevertheless, researchers so far have made little use of this in DSGE models. A related question is how to estimate macro-finance models without computationally costly state-space representations. In recent work, van Binsbergen, Fernández-Villaverde, Koijen, and Rubio-Ramírez (2012) do in fact solve and estimate a DSGE model using both macro data and bond yields. The authors illustrate the usefulness of incorporating financial market data into the estimation, but their analysis is purely in discrete time. As this usually requires numerical integration to compute expectations, it makes their methods computationally heavy.

In this paper, we make the link between macro and financial markets explicit by showing how financial market data facilitate the estimation of structural parameters characterizing preferences and technology. We cast our DSGE model in continuous time, solve for the general equilibrium of the real economy and financial markets, and subsequently develop the relevant estimation procedures. We consider both regression-based methods combined with a minimum-distance approach, and an alternative and asymptotically efficient martingale estimating function (MEF) technique. Our continuous-time formulation serves to (i) put structure on the residuals encountered in the regression-based estimation methods, (ii)

[1] See, for example, the debate in *The Economist* 2009, July 16th, "What went wrong with economics." In a survey article, Cochrane (2006, chap. I.1) argues: "The general equilibrium approach is a vast and largely unexplored new land. The papers [in this area] are like Columbus' report that the land is there."

obtain an explicit solution for the equilibrium dynamics of the economy in terms of data and parameters for the MEF approach, and (iii) account for the dependence among economic variables during the unit observation interval. It is important to study the estimation problems encountered in simple continuous-time models before addressing more elaborate models at the vanguard of the DSGE literature. We illustrate our approach for a class of DSGE models where the equilibrium dynamics are available in terms of observable quantities, namely, consumption, output, and spot interest rates. We consider both logarithmic and constant relative risk aversion (CRRA) preferences. Although log preferences are included in the more general CRRA class, they constitute an important benchmark case that allows for closed-form solution. Specifications of this kind date back at least to Cox, Ingersoll, and Ross (1985a). Our examples can be used as points of reference for exploring broader classes of dynamic general equilibrium models.

We introduce the linkage to financial markets by basing our approach on continuous-time equilibrium term structure models along the lines of Vasicek (1977) and Cox, Ingersoll, and Ross (1985b). We use daily data on the 3-month interest rate as a proxy for the spot rate (cf. Chapman, Long, and Pearson, 1999), along with aggregate consumption and output at lower frequency, to facilitate estimation of the structural parameters of the system in a parsimonious specification. The choice of the 3-month rate is a starting point, as it is necessary to build up empirical experience with combinations of macro and financial data available at different frequencies.

We depart from the traditional discrete-time formulation of DSGE models and their estimation for three related reasons.[2] First, the continuous-time approach has proved useful in formulating and solving dynamic models in macroeconomics and finance. There is no need to perform numerical integration to compute expectations, since the Bellman equation is non-stochastic, thus simplifying computation of the first-order conditions. Closed-form solutions are obtained in many cases and may serve as benchmarks for numerical solutions. Second, the presence of closed-form solutions can simplify inference on structural parameters even in the presence of non-linearities and non-normality (cf. Posch, 2009). Third, many financial models (e.g., equilibrium term structure models) are stated in continuous time. When linking the macro economy and financial markets, any discrete period length is arbitrary, and it seems natural to make the least stringent timing assumption, formulating both the macro and financial sides of the model in continuous time, thus avoiding specific discrete-time approximations for either of the data generating processes.

[2] A non-exhaustive list of references on structural estimation of discrete-time DSGE models is Fernández-Villaverde and Rubio-Ramírez (2007) and An and Schorfheide (2007). While the first authors show how to use the output of the particle filter to estimate the structural parameters of the model, the latter review Bayesian methods for estimating discrete-time DSGE models.

There is a tradition estimating continuous-time models in macroeconomics formulated by systems of stochastic differential equations (e.g. Bergstrom, 1966; Phillips, 1972, 1991), and rational expectations models (Hansen and Sargent, 1991; Hansen and Scheinkman, 1995). Since data are sampled only at fixed points in time, we follow Bergstrom's idea and integrate our equilibrium system of stochastic differential equations to obtain the 'exact discrete-time model'. The traditional formulations (as in Bergstrom, 1966; Phillips, 1972) typically imply a coefficient matrix that is a function of the exponential of a matrix depending on the structural parameters. As shown by McCrorie (2009), this property is problematic, as it complicates the identification of continuous-time models from discrete-time data due to the aliasing phenomenon: Distinct continuous-time processes may look identical when sampled at discrete time intervals (cf. Hansen and Scheinkman, 1995, p.769). In this paper we adopt the alternative approach of integrating the logarithmic system. In specific models we consider, the resulting system for logarithmic growth rates rather than levels involves a coefficient matrix that is linear in a set of known functions of the structural parameters. The system does not involve any matrix exponential, thus circumventing the aliasing problem. The relevant untransformed system of stochastic differential equations in the DSGE model is nonlinear and generally does not have a closed-form solution, so working with the log-transformed system involves no loss in this sense.

We apply our model to both simulated and empirical data on production, consumption, and interest rates. A Monte Carlo study examines the properties of our estimation approaches for 1,000 simulated data sets of 25 years each for both monthly and quarterly macro data, roughly in line with the availability of empirical data. The results show that both the regression-based and MEF approaches are able to accurately estimate the parameters of the model with logarithmic preferences. While our model with CRRA preferences turns out to be more challenging for the simpler regression-based approaches, the MEF approach produces reliable estimation results.

Our empirical application of more than 20 years of U.S. data shows that the system can indeed be applied to a combination of macro and financial series. The results indicate a long run mean of the short rate of interest around 5% with a 3% volatility annually and weak mean reversion, as well as higher relative risk aversion in the representative agent than under logarithmic preferences. The elasticity of consumption with respect to wealth is strongly significant but below unity, the value corresponding to log preferences, whereas the interest rate elasticity of consumption differs insignificantly from the zero value implied by log preferences.

The paper proceeds as follows. Section 2 summarizes the macroeconomic theory and solution techniques. Section 3 presents the estimation strategies. Sections 4 and 5 provide

Monte Carlo evidence on small sample properties of our estimation strategies and report empirical estimates. Section 6 concludes.

2 The macro-finance framework

We consider dynamic stochastic general equilibrium models cast in continuous time (Eaton, 1981; Cox, Ingersoll, and Ross, 1985a). This allows the application of Itô's calculus, and in some cases we can solve the model analytically to obtain closed-form expressions facilitating statistical inference.

2.1 The model

Production possibilities. At each point in time, certain amounts of capital, labor, and factor productivity are available in the economy, and these are combined to produce output. The production function is a constant returns to scale technology subject to regularity conditions (see Chang, 1988),

$$Y_t = A_t F(K_t, L), \tag{1}$$

where K_t is the aggregate capital stock, L is the constant population size, and A_t is total factor productivity (TFP), in turn driven by a standard Brownian motion B_t,[3]

$$dA_t = \mu(A_t)dt + \eta(A_t)dB_t, \tag{2}$$

with $\mu(A_t)$ and $\eta(A_t)$ generic drift and volatility functions satisfying regularity conditions.[4] The capital stock increases if gross investment I_t exceeds capital depreciation,

$$dK_t = (I_t - \delta K_t)dt + \sigma K_t dZ_t, \tag{3}$$

where δ denotes the mean and σ the volatility of the stochastic depreciation rate, driven by another standard Brownian motion Z_t.

Equilibrium properties. In equilibrium, factors of production are rewarded with marginal products $r_t = Y_K$ and $w_t = Y_L$, subscripts K and L indicating derivatives, and the goods market clears, $Y_t = C_t + I_t$. By an application of Itô's formula (e.g., Protter, 2004; Sennewald, 2007), the technology in (2), capital accumulation in (3), and market clearing condition together imply that output evolves according to

$$\begin{aligned} dY_t &= Y_A dA_t + Y_K dK_t + \tfrac{1}{2} Y_{KK}\sigma^2 K_t^2 dt \\ &= (\mu(A_t)Y_A + (I_t - \delta K_t)Y_K + \tfrac{1}{2}Y_{KK}\sigma^2 K_t^2)dt + Y_A \eta(A_t)dB_t + \sigma Y_K K_t dZ_t. \end{aligned} \tag{4}$$

[3]Since B_t is a standard Brownian motion, $B_0 = 0$, $B_{t+\Delta} - B_t \sim \mathcal{N}(0, \Delta)$, $t \in [0, \infty)$.

[4]We assume that $E(A_t) = A \in \mathbb{R}_+$ exists, and that the integral describing life-time utility in (5) below is bounded, so that the value function is well-defined.

This corresponds to equation (1) in Cox, Ingersoll, and Ross (1985a) (henceforth CIR), where $I_t - \delta K_t$ is the amount of the output good allocated to the production process. In general, Y_t can be a nonlinear activity with respect to capital, determined by its output elasticity.[5]

Preferences. We consider an economy with a single consumer, interpreted as a representative "stand in" for a large number of identical consumers. The consumer maximizes expected additively separable discounted life-time utility given by

$$U_0 \equiv E_0 \int_0^\infty e^{-\rho t} u(C_t, A_t) dt, \quad u_C > 0, \; u_{CC} < 0, \tag{5}$$

subject to

$$dK_t = ((r_t - \delta)K_t + w_t L - C_t)dt + \sigma K_t dZ_t, \tag{6}$$

where ρ is the subjective rate of time preference, r_t is the rental rate of capital, and w_t is the labor wage rate. We do not consider financial claims, although they could easily be added. The paths of factor rewards are taken as given by the representative consumer. The generic utility flow function specification $u(C_t, A_t)$ allows the possibility that technology enters as an argument. This may represent a quest for technology and is included for comparability with Cox, Ingersoll, and Ross (1985a).

2.2 The Euler equation

The relevant state variables are capital and technology, (K_t, A_t). For given initial states, the value of the optimal program is

$$V(K_0, A_0) = \max_{\{C_t\}_{t=0}^\infty} U_0 \quad s.t. \quad (6) \quad \text{and} \quad (2), \tag{7}$$

i.e., the present value of expected utility along the optimal program. It is shown in the appendix that the first-order condition for the problem is

$$u_C(C_t, A_t) = V_K(K_t, A_t), \tag{8}$$

for any $t \in [0, \infty)$, and this allows writing consumption as a function of the state variables, $C_t = C(K_t, A_t)$. The *Euler equation* is

$$\begin{aligned} \frac{du_C}{u_C} &= (\rho - (r_t - \delta))dt - \frac{u_{CC}(C_t, A_t)}{u_C(C_t, A_t)} C_K \sigma^2 K_t dt + \frac{u_{CC}(C_t, A_t)}{u_C(C_t, A_t)} C_A \eta(A_t) dB_t \\ &\quad + \frac{u_{CA}(C_t, A_t)}{u_C(C_t, A_t)} \eta(A_t) dB_t + \frac{u_{CC}(C_t, A_t)}{u_C(C_t, A_t)} C_K \sigma K_t dZ_t, \end{aligned} \tag{9}$$

[5]Unless we consider a nonlinear production process, our model is formally included in the CIR economy. We are not aware of any paper estimating the model's structural parameters using macro and financial data.

also derived in the appendix. This implicitly determines the optimal consumption path.

In the following, we restrict attention to the case $u(C_t, A_t) = u(C_t)$. Using the inverse marginal utility function, $c = g(u'(c))$, we obtain the path for consumption,

$$\begin{aligned} dC_t &= \frac{u'(C_t)}{u''(C_t)}(\rho - (r_t - \delta))dt - \sigma^2 C_K K_t dt - \tfrac{1}{2}(C_A^2 \eta(A_t)^2 + C_K^2 \sigma^2 K_t^2)\frac{u'''(C_t)}{u''(C_t)}dt \\ &\quad + C_A \eta(A_t) dB_t + C_K \sigma K_t dZ_t, \end{aligned}$$

where $u' > 0$ and $u'' < 0$ (strict concavity of preferences).

2.3 Equilibrium dynamics of the economy

The equilibrium dynamics of the economy may be represented as

$$\begin{aligned} d\ln C_t &= \left(\frac{u'(C_t)}{u''(C_t)C_t}(\rho - r_t + \delta) - \sigma^2 \frac{C_K K_t}{C_t} - \tfrac{1}{2}\frac{C_A^2 \eta(A_t)^2 + C_K^2 \sigma^2 K_t^2}{C_t^2}\frac{u'''(C_t)C_t + u''(C_t)}{u''(C_t)}\right) dt \\ &\quad + C_A \eta(A_t)/C_t dB_t + C_K \sigma K_t / C_t dZ_t, \\ d\ln Y_t &= \left(\frac{\mu(A_t)}{A_t} + \left(\frac{Y_t - C_t}{K_t} - \delta\right)\frac{K_t Y_K}{Y_t} + \tfrac{1}{2}\sigma^2 \frac{K_t^2 Y_{KK}}{Y_t}\right) dt - \tfrac{1}{2}\frac{Y_A^2 \eta(A_t)^2 + \sigma^2 Y_K^2 K_t^2}{Y_t^2} dt \\ &\quad + Y_A \eta(A_t)/Y_t dB_t + \sigma Y_K K_t / Y_t dZ_t, \\ d\ln K_t &= (r_t - \delta + w_t/K_t - C_t/K_t - \tfrac{1}{2}\sigma^2)dt + \sigma dZ_t. \end{aligned}$$

If all left-hand side variables, the logarithms of C_t, Y_t, and K_t, were observed, along with A_t, estimation could be based directly on this system and the TFP equation (2). Consumption and income are standard variables in most macro studies. Capital and technology are notoriously problematic, due to the risk of mismeasurement. This is where we propose using financial variables, instead. Thus, suppose that interest rates r_t are observed, along with C_t and Y_t.[6] We consider systems of stochastic differential equations that can be used for estimation based on data on C_t, Y_t, and r_t, by recasting the equilibrium dynamics in terms of these observables.

2.4 An illustration: The stochastic AK model

Consider an economy where labor is not an input of production, $Y_t = A_t K_t$, known as an AK model,[7] and assume that the representative consumer has constant relative risk aversion

[6] One caveat is that some variables are observed as an integral over an interval (flows) rather than at a point in time (stocks). It thus appears more difficult to estimate continuous-time models compared to their discrete-time counterparts (Harvey and Stock, 1989). In this paper we adopt a pragmatic approach: we approximate a flow variable, e.g., Y_t at time t, by the integral $\int_{t-\Delta}^{t} Y_s ds$. Observed growth rates of flow variables thus correspond to $\ln Y_t - \ln Y_{t-\Delta}$. In our simulation study we find a negligible 'time-aggregation bias'.

[7] The AK framework is also used in other macro-finance models (cf. Brunnermeier and Sannikov, 2011).

(CRRA) preferences, $u(C_t) = C_t^{1-\theta}/(1-\theta)$. With these assumptions, $A_t = Y_K = r_t$ and $K_t = Y_t/A_t = Y_t/r_t$, so the state variables (A_t, K_t) are expressed as known functions of the readily observable (Y_t, r_t). Since $w_t = Y_L = 0$, equilibrium dynamics are

$$
\begin{aligned}
d\ln C_t &= \left(r_t - \rho - \delta - \theta\sigma^2\pi(Y_t, r_t) + \tfrac{1}{2}(\theta\eta(r_t)\tau(Y_t, r_t))^2 + \tfrac{1}{2}(\theta\sigma\pi(Y_t, r_t))^2\right)/\theta dt \\
&\quad + \eta(r_t)\tau(Y_t, r_t)dB_t + \sigma\pi(Y_t, r_t)dZ_t, \qquad (10a) \\
d\ln Y_t &= \left(\mu(r_t)/r_t + (1 - C_t/Y_t)r_t - \delta - \tfrac{1}{2}\eta(r_t)^2/r_t^2 - \tfrac{1}{2}\sigma^2\right)dt + \eta(r_t)/r_t dB_t + \sigma dZ_t, \qquad (10b) \\
dr_t &= \mu(r_t)dt + \eta(r_t)dB_t, \qquad (10c)
\end{aligned}
$$

where $\pi(Y_t, r_t) \equiv C_K K_t/C_t$ and $\tau(Y_t, r_t) \equiv C_A/C_t$ are the relevant sensitivities of the consumption function with respect to the state variables. In particular, $\pi(Y_t, r_t)$ is the capital elasticity of consumption. Similarly, the term $\tau(Y_t, r_t)$ is an infinitesimal version of the percentage change in consumption associated with a percentage point change in the interest rate (or TFP). In general, the consumption function is non-homogeneous with respect to the interest rate r_t and capital K_t (or wealth, here the output-TFP ratio), thus implying that the time-varying portion of these sensitivities is small. The functions $\mu(\cdot)$ and $\eta(\cdot)$ are chosen such that suitable boundedness conditions are met (cf. Posch, 2009). We illustrate the estimation of the stochastic AK model with the interest rate governed by a Vasicek specification (henceforth the AK-Vasicek model).

2.4.1 AK-Vasicek model: Logarithmic preferences

With logarithmic utility, $u(C) = \ln C$, corresponding to the case of relative risk aversion $\theta = 1$, it can be shown that optimal consumption is linear in the capital stock, $C_t = \rho K_t$ (cf. appendix). This implies that the consumption function is linear-homogeneous in capital, yielding unit elasticity, $\pi(Y_t, r_t) = 1$. Moreover, consumption does not respond to changes in the interest rate or technology, $C_A = 0$, so $\tau(Y_t, r_t) = 0$. The Vasicek (1977) mean-reverting interest rate specification is $\mu(r_t) = \kappa(\gamma - r_t)$ and $\eta(r_t) = \eta$, where $\kappa > 0$ is the speed and γ the target of mean reversion, and η the constant volatility. In this case, the equilibrium dynamics are

$$
\begin{aligned}
d\ln C_t &= \left(r_t - \rho - \delta - \tfrac{1}{2}\sigma^2\right)dt + \sigma dZ_t, \qquad (11a) \\
d\ln Y_t &= \left(\kappa\gamma/r_t - \tfrac{1}{2}\eta^2/r_t^2 + r_t - \kappa - \rho - \delta - \tfrac{1}{2}\sigma^2\right)dt + \eta/r_t dB_t + \sigma dZ_t, \qquad (11b) \\
dr_t &= \kappa(\gamma - r_t)dt + \eta dB_t. \qquad (11c)
\end{aligned}
$$

Alternative specifications of the interest rate process as in Aït-Sahalia (1996, p.528) can be implemented and the system estimated along the lines developed below. The closed-form solution, however, does not depend on the particular choice.

2.4.2 AK-Vasicek model: CRRA preferences

A slightly more general specification is that of CRRA utility. As no closed-form solution is known for the case $\theta \neq 1$, our parametric estimation approach requires an approximation for the consumption function to determine the unknown functions $\tau(Y_t, r_t)$ and $\pi(Y_t, r_t)$. In particular, we assume that the consumption elasticities with respect to both state variables are roughly constant, $C_K K_t/C_t \approx \bar{\pi}$, $C_r r_t/C_t \approx \bar{\tau}$,[8] so that

$$\begin{aligned} d\ln C_t &= \left((r_t-\rho-\delta)/\theta + \tfrac{1}{2}\theta(\bar{\tau}\eta)^2/r_t^2 + \tfrac{1}{2}(\theta\bar{\pi}-2)\bar{\pi}\sigma^2\right)dt + \eta\bar{\tau}/r_t dB_t + \sigma\bar{\pi}dZ_t, && \text{(12a)} \\ d\ln Y_t &= \left(\kappa\gamma/r_t - \tfrac{1}{2}\eta^2/r_t^2 + (1-C_t/Y_t)r_t - (\kappa+\delta+\tfrac{1}{2}\sigma^2)\right)dt + \eta/r_t dB_t + \sigma dZ_t, && \text{(12b)} \\ dr_t &= \kappa(\gamma - r_t)dt + \eta dB_t. && \text{(12c)} \end{aligned}$$

For reasonable parametric restrictions, our assumptions are economically meaningful. We provide a particular numerical solution of the system (12) and the consumption elasticities with respect to the interest rate and capital stock in Appendix A.2.1. These numerical results are based on the collocation method.[9] In a nutshell, the idea is to approximate the unknown value function by a linear combination of known basis functions evaluated at the collocation nodes.[10] For the case $\theta = 2$, and with reasonable calibrations for the other parameters (cf. Appendix A.2.1), we find that the dependence of the elasticities $\pi(Y_t, r_t)$ and $\tau(Y_t, r_t)r_t$ on the state variables (A_t, K_t) is negligible, with $1.0057 < \pi(Y_t, r_t) < 1.0211$ and $0.1623 < \tau(Y_t, r_t)r_t < 0.2625$, thus confirming the relevance of the approximate system (12). In economic terms, our results are that consumption increases by about 2% per percentage point increase in the interest rate, whereas the percentage changes in consumption and wealth are about equal.

3 Estimation

In this section we describe how to estimate the equilibrium system (11) using macro and financial data. First, we integrate the system to obtain the exact discrete-time model. Section 3.1 presents the resulting formulation of the model for estimation purposes. Section 3.2 illustrates (i) how reduced-form parameters can be estimated by means of standard regression-based methods, and (ii) how structural parameters are obtained using minimum distance. Section 3.3 shows how structural parameters may alternatively be estimated directly using the martingale estimating function approach. Our illustrations are based on the

[8] We tried alternatives (e.g., C_r/C_t constant), and results were qualitatively similar.

[9] Our non-linear solver is based on the CompEcon toolbox for Matlab (cf. Miranda and Fackler, 2002).

[10] Since wealth is not bounded, we consider $K_t = 1$ as a benchmark and explore the region $0.5 < K_t < 1.25$, while for the interest rate we study a strip around its long run mean, $0.8\gamma < r_t < 1.2\gamma$. We employ 10 and 7 Chebychev polynomial bases for K_t and $A_t = r_t$, at standard Chebychev nodes.

AK-Vasicek model with logarithmic utility. In the appendix we show the generalization of our estimation approach to the case of CRRA preferences, as in system (12).

3.1 Discrete-time formulation

In order to accommodate the discrete-time nature of the data, we integrate over $s \geq t$, employing exact solutions whenever possible. Using the system of differential equations (11), we obtain

$$\ln(C_s/C_t) - \int_t^s r_v dv = -\left(\rho + \delta + \tfrac{1}{2}\sigma^2\right)(s-t) + \sigma(Z_s - Z_t), \tag{13a}$$

$$\begin{aligned}\ln(Y_s/Y_t) - \int_t^s r_v dv &= \kappa\gamma \int_t^s 1/r_v dv - \tfrac{1}{2}\eta^2 \int_t^s 1/r_v^2 dv - \left(\kappa + \rho + \delta + \tfrac{1}{2}\sigma^2\right)(s-t) \\ &\quad + \int_t^s \eta/r_v dB_v + \sigma(Z_s - Z_t),\end{aligned} \tag{13b}$$

$$r_s = e^{-\kappa(s-t)} r_t + (1 - e^{-\kappa(s-t)})\gamma + \eta e^{-\kappa(s-t)} \int_t^s e^{\kappa(v-t)} dB_v. \tag{13c}$$

This system of three equations forms the basis of our first empirical specifications. At the same time it illustrates the main ideas underlying our approach. First, the continuous-time analysis delivers the explicit functional forms of the relations among observables. Second, the availability of interest rate data at higher frequency (say, daily) than consumption and production (monthly or quarterly) allows precise approximation of the ordinary (although not the stochastic) integrals involving the interest rate by summation over days. In our applications we approximate the integrals by the Riemann sum $\int_t^s g(r_v)dv \approx (s-t)\sum_{i=1}^P g(r_{t+i(s-t)/P})/P$, where $g(\cdot)$ is a smooth function of $r_{t+i(s-t)/P}$, the prevailing interest rate on day i in the period between t and s, and P is the number of days in the period.[11] Third, the structural parameters enter into the coefficients on the terms involving interest rates, highlighting that financial data help identify the parameters of interest. The system is in fact linear in a set of reduced-form parameters that in turn are known functions of the structural parameters.

We have some choice in turning system (13) into an empirical specification. Initially, we specify a system of three regression equations for equidistant macro data, i.e., we define $\Delta \equiv s - t$ ($\Delta = 1/12$ for monthly macro data, $\Delta = 1/4$ for quarterly). Given the higher (say, daily) frequency of the interest rate data, an alternative would be to start out with separate estimation of the third equation, but the full system is likely closer to that required for more complicated models (e.g., if consumption or income enters endogenously in the interest rate

[11]For notational convenience, we write P as a constant, but in our empirical approach we use the actual number of days in the period.

equation), and the high-frequency property of the interest rate data is in any case exploited in the approximation of the integrals as Riemann sums.

3.2 A regression-based approach

In this section we propose simple regression-based procedures to obtain parameter estimates. To start with, we employ unrestricted ordinary least squares (OLS) to get reduced-form parameters, although this does not identify the structural parameters of interest. Next, we consider cross-equation correlation, controlling for endogeneity through instrumental variables, and estimation of structural parameters by minimum distance.

3.2.1 Reduced-form model

With $s - t$ fixed at Δ, system (13) is linear in a set of reduced-form parameters and may be recast as

$$y_{j,t} = x_{j,t}\beta_j + \varepsilon_{j,t}, \qquad j = C, Y, r, \tag{14}$$

where the left-hand side variables are $y_{C,t} = \ln(C_t/C_{t-\Delta}) - \int_{t-\Delta}^t r_v dv$, $y_{Y,t} = \ln(Y_t/Y_{t-\Delta}) - \int_{t-\Delta}^t r_v dv$, and $y_{r,t} = r_t$. Simlarly, the right-hand side variables $x_t = (x_{C,t}, x_{Y,t}, x_{r,t})$, with $x_{C,t} = 1$, $x_{Y,t} = (1, \int_{t-\Delta}^t 1/r_v dv, \int_{t-\Delta}^t 1/r_v^2 dv)$, and $x_{r,t} = (1, r_{t-\Delta})$. The reduced-form or linear parameters β_C, $\beta_Y = (\beta_{Y,1}, \beta_{Y,2}, \beta_{Y,3})^\top$, and $\beta_r = (\beta_{r,1}, \beta_{r,2})^\top$ are given in terms of the structural parameters as

$$\beta_C = -\left(\rho + \delta + \tfrac{1}{2}\sigma^2\right)\Delta, \tag{15a}$$
$$\beta_{Y,1} = -\left(\kappa + \rho + \delta + \tfrac{1}{2}\sigma^2\right)\Delta, \tag{15b}$$
$$\beta_{Y,2} = \kappa\gamma, \tag{15c}$$
$$\beta_{Y,3} = -\tfrac{1}{2}\eta^2, \tag{15d}$$
$$\beta_{r,1} = (1 - e^{-\kappa\Delta})\gamma, \tag{15e}$$
$$\beta_{r,2} = e^{-\kappa\Delta}. \tag{15f}$$

Hence, the system (13) can be summarized in form of simple regression equations, with error terms given by

$$\varepsilon_{C,t} = \sigma(Z_t - Z_{t-\Delta}), \tag{16a}$$
$$\varepsilon_{Y,t} = \int_{t-\Delta}^t \eta/r_v dB_v + \sigma(Z_t - Z_{t-\Delta}), \tag{16b}$$
$$\varepsilon_{r,t} = \eta e^{-\kappa\Delta}\int_{t-\Delta}^t e^{\kappa(v-(t-\Delta))} dB_v. \tag{16c}$$

For a simple reduced-form estimator, linearity in β suggests unrestricted OLS,

$$\hat{\beta}_j = (x_j^\top x_j)^{-1} x_j^\top y_j, \qquad j = C, Y, r, \tag{17}$$

where x_j is the matrix with typical row $x_{j,t}$ and y_j the vector with typical entry $y_{j,t}$. Structural parameter estimates obtained by minimum distance appliued to the reduced-form estimates (17) using the link (15a)-(15f) serve as useful benchmarks for assessing more elaborate structural approaches.

3.2.2 Cross-equation correlation

Unrestricted OLS estimation (17) allows for different variances of the error terms $\varepsilon_{j,t}$, $j = C, Y, r$, in the sense that it is carried out separately by equation, but it does not exploit any other property of the errors. Classical seemingly unrelated regressions (SUR) analysis is intended to exploit cross-equation correlation in cases where the right-hand side variables are not common across equations. The present model structure implies both different right-hand side variables (indeed, of different dimensions) across the equations, and cross-equation correlation. In particular, from (13), the errors in (14) take the form (16a)-(16b), where, e.g., the term $\sigma(Z_t - Z_{t-\Delta})$ is common to both (16a) and (16b). This suggests that a standard SUR correction of the reduced-form estimates should lead to better structural parameter estimates (using minimum distance). In particular, the reduced-form SUR estimates are more efficient than OLS (see the appendix for details on the SUR estimator and constrained nonlinear regression estimates that are asymptotically equivalent to the resulting minimum distance estimates).

Let $\hat{\varepsilon}$ be the $T \times 3$ matrix of OLS residuals, with typical element $\{\hat{\varepsilon}_{j,t}\}$, where T is the number of time periods in the data set. The SUR estimate of the 3×3 contemporaneous system variance-covariance matrix is $\hat{\Sigma} = \hat{\varepsilon}^\top \hat{\varepsilon}/T$ (in particular, the residual variance estimates along the diagonal coincide with the standard OLS assessments), and the FGLS-SUR estimate of β is

$$\hat{\beta}_{SUR} = (x^\top \hat{V}^{-1} x)^{-1} x^\top \hat{V}^{-1} y, \tag{18}$$

where y is the $3T$-vector stacking the y_j, x is the conformable matrix with the x_j along the block-diagonal, and $\hat{V}^{-1} = \hat{\Sigma}^{-1} \otimes I_T$, with I_T the identity matrix and $\otimes$ the Kronecker product. The standard SUR assessment of the asymptotic variance-covariance matrix of $\hat{\beta}_{SUR}$ is $\hat{V}_{SUR} = (x^\top \hat{V}^{-1} x)^{-1}$. Note that the (i, j)'th block of the matrix being inverted is $\hat{\Sigma}^{ij} x_i^\top x_j$, with $\hat{\Sigma}^{ij}$ the (i, j)'th entry in $\hat{\Sigma}^{-1}$, and if estimated covariances $\hat{\Sigma}_{ij}$ $(i \neq j)$ are zero then the estimated asymptotic variance of $\hat{\beta}_j$ coincides with the OLS assessment $\hat{\Sigma}_{jj}(x_j^\top x_j)^{-1}$. More generally, the SUR approach suggests that the variance-covariance matrix $\hat{V}_{OLS}$ of the unrestricted OLS estimator from (17) has blocks estimated as $\hat{\Sigma}_{ij}(x_i^\top x_i)^{-1}x_i^\top x_j(x_j^\top x_j)^{-1}$, and $\hat{V}_{OLS} \geq \hat{V}_{SUR}$ in the partial order of positive semi-definite matrices.

3.2.3 Structural parameters

The structural parameters are estimated by exploiting the manner in which they enter into the reduced-form (linear) parameters β from (15a)-(15f). The estimated variance-covariance matrix of either $\hat{\beta}_{OLS}$ or $\hat{\beta}_{SUR}$ may form the basis of a minimum distance approach (see Section 3.2.5 below). The minimum distance estimator based on SUR should be more efficient than that based on OLS. Estimators that are asymptotically equivalent to the two minimum distance estimators are alternatively obtained by minimizing the OLS respectively the SUR objective function under the relevant structural restrictions (15a)-(15f) on β. This amounts to weighted nonlinear regression and is carried out by iterating over the structural parameters. In particular, the OLS objective is $\sum_{j=C,Y,r} \varepsilon_j^\top \varepsilon_j / \hat{\Sigma}_{jj}$ and the SUR objective $\sum_{t=1}^{T} \varepsilon_t^\top \hat{\Sigma}^{-1} \varepsilon_t$, where ε_j and ε_t are residual vectors of dimension T and 3, respectively, with typical elements $\varepsilon_{j,t}$.

If estimated residual variances or covariances are used to identify structural parameters, then these may be included in the minimum distance approach, using asymptotic independence between estimated $\widehat{\beta}$ and $\widehat{\Sigma}$. Again, an asymptotically equivalent estimator may be based on the SUR objective function, with Σ as it depends on structural parameters instead of as $\hat{\Sigma}$, and the nonlinear minimization is over structural parameters as they enter Σ as well as ε. This use of the Gaussian log-likelihood function amounts to quasi maximum likelihood (QML) since clearly $\varepsilon_{Y,t}$ in (16b) is non-Gaussian.

3.2.4 Endogeneity

The regression approaches (OLS and SUR) do not control for possible endogeneity of right-hand side variables in (13)-(16), which may be an issue in the DSGE model. In particular, $x_{Y,t}$ includes two integrals involving the evolution of the interest rate from $t - \Delta$ through t and so is correlated with both $\varepsilon_{r,t}$ and $\varepsilon_{Y,t}$. The standard regression-based tool for handling endogeneity is instrumental variables (IV) or two-stages least squares (2SLS). Here, we consider first-stage regressions of each of $x_{Y,t,2} = \int_{t-\Delta}^{t} 1/r_v dv$ and $x_{Y,t,3} = \int_{t-\Delta}^{t} 1/r_v^2 dv$ on their respective lags $x_{Y,t-\Delta,2}$ and $x_{Y,t-\Delta,3}$ and an intercept. Next, in the computation (17) of $\widehat{\beta}_Y$, fitted values from the first stage regressions replace $x_{Y,t,2}$ and $x_{Y,t,3}$. Third, fitted residuals are calculated using the new second stage estimate $\widehat{\beta}_Y$ but the original $x_{Y,t,2}$ and $x_{Y,t,3}$ (not their fitted values from the first stage), and these residuals form the basis of the 2SLS assessment of $\hat{\Sigma}$. Finally, the FGLS-SUR-IV step is carried out using this new $\hat{\Sigma}$ in calculating $\hat{\beta}_{SUR}$ in (18) and again using the fitted values for $x_{Y,t,2}$ and $x_{Y,t,3}$. Once again, minimum distance or restricted (nonlinear) regression is used to estimate structural parameters. The minimum distance approach requires a variance-covariance matrix, and this has the same form as before, but with the new $\hat{\Sigma}$ and with fitted values for the relevant portions of x, and similarly for the OLS and SUR objective functions for the restricted nonlinear regressions.

3.2.5 Minimum distance

The structural parameters are $\kappa, \gamma, \eta, \rho, \delta$, and σ, a total of six. From (15), we may identify four structural parameters, i.e., $\kappa, \gamma, \eta, \rho + \delta + \frac{1}{2}\sigma^2$, the latter being a combination of three of the original six, if the functional form of the error variances (the variances of (16a)-(16c)) are not exploited to help identification. By including the variance of the consumption residual (16a) as a separate moment, we may identify one additional structural parameter, i.e., $\kappa, \gamma, \eta, \rho + \delta$, and σ are identified, since σ^2 is separately identified from the variance of (16a).

We obtain the structural model parameters from the OLS, SUR, and FGLS-SUR-IV reduced form parameter estimates using a minimum distance approach. An alternative (asymptotically equivalent) method is restricted (nonlinear) regression, described in the appendix. We carry out minimum distance estimation based on three different unrestricted parameter sets from the reduced form regressions: (1) the estimates of β in (15); (2) β along with the variance $\sigma^2\Delta$ of the consumption equation residual in (16a); (3) β along with the variances of the consumption and interest rate residuals (16a) and (16c). The first of these is applied to both the OLS and SUR reduced form parameter estimators, the latter two only to OLS. In each of the cases considered, labeled with subscript i, we use a numerical optimization algorithm to solve the problem

$$\hat{\phi} = \arg\min_{\phi} \left(\omega_i(\phi) - \hat{\omega}_i\right)^\top \hat{\Omega}_i^{-1} \left(\omega_i(\phi) - \hat{\omega}_i\right),$$

where ϕ denotes the relevant vector comprising four or five structural parameters. Thus, when no variances are included, $\phi = (\kappa, \gamma, \eta, \rho+\delta+\frac{1}{2}\sigma^2)^\top$. If the variance of the consumption equation is included (or the variances of both the consumption and interest rate equations), then σ is identified and enters as a separate argument, $\phi = (\kappa, \gamma, \eta, \rho + \delta, \sigma)^\top$. The reduced form estimates are collected in

$$\hat{\omega}_1 = \hat{\beta}, \qquad \hat{\Omega}_1^{-1} = \begin{pmatrix} \hat{\Sigma}^{CC} x_C^\top x_C & \hat{\Sigma}^{CY} x_C^\top x_Y & \hat{\Sigma}^{Cr} x_C^\top x_r \\ \hat{\Sigma}^{YC} x_Y^\top x_C & \hat{\Sigma}^{YY} x_Y^\top x_Y & \hat{\Sigma}^{Yr} x_Y^\top x_r \\ \hat{\Sigma}^{rC} x_r^\top x_C & \hat{\Sigma}^{rY} x_r^\top x_Y & \hat{\Sigma}^{rr} x_r^\top x_r \end{pmatrix},$$

$$\hat{\omega}_2 = \begin{pmatrix} \hat{\omega}_1 \\ \hat{\Sigma}_{CC} \end{pmatrix}, \quad \hat{\Omega}_2^{-1} = \begin{pmatrix} \hat{\Omega}_1^{-1} & 0_{6\times 1} \\ 0_{1\times 6} & \left(2\hat{\Sigma}_{CC}^2\right)^{-1} \end{pmatrix}, \quad \hat{\omega}_3 = \begin{pmatrix} \hat{\omega}_2 \\ \hat{\Sigma}_{rr} \end{pmatrix}, \quad \hat{\Omega}_3 = \begin{pmatrix} \hat{\Omega}_2^{-1} & 0_{7\times 1} \\ 0_{1\times 7} & \left(2\hat{\Sigma}_{rr}^2\right)^{-1} \end{pmatrix},$$

as before with x_j, $j = C, Y, r$, denoting the regressors for each regression collected in a matrix for all observations. As already discussed, non-zero off-diagonal blocks is an immediate extension (the SUR case, i.e., $\hat{\Omega}_1 = \hat{V}_{SUR}$ is the inverse of the matrix with blocks $\hat{\Sigma}^{CC}(x_C^\top x_C)$, $\hat{\Sigma}^{YC}(x_Y^\top x_C)$, etc.). The vectors mapping the structural parameters to the reduced form estimates are (see (15a)-(15f))

$$\begin{aligned}
\omega_1(\phi) &= \begin{pmatrix} -(\rho+\delta+\frac{1}{2}\sigma^2)\Delta & -(\kappa+\rho+\delta+\frac{1}{2}\sigma^2)\Delta & \kappa\gamma & -\frac{1}{2}\eta^2 & (1-e^{-\kappa\Delta})\gamma & e^{-\kappa\Delta} \end{pmatrix}^\top, \\
\omega_2(\phi) &= \begin{pmatrix} \omega_1(\phi)^\top & \sigma^2\Delta \end{pmatrix}^\top, \\
\omega_3(\phi) &= \begin{pmatrix} \omega_2(\phi)^\top & \frac{1}{2}\eta^2(1-e^{-2\kappa\Delta})/\kappa \end{pmatrix}^\top.
\end{aligned}$$

Standard errors on the structural parameters are obtained using the delta method. When controlling for endogeneity using the FGLS-SUR-IV method, fitted values from the first stage are used for all regressors in the expressions for $\hat{\Omega}_i$, but not when calculating the residuals used to estimate the $\hat{\Sigma}_{ij}$ factors in the expressions, although the residuals are based on IV-corrected second-stage coefficient estimates in this case. This combination of FGLS, SUR, and IV/2SLS (labeled FGLS-SUR-IV) appears to be novel.

3.3 The martingale estimating function approach

The issue remains whether all endogeneity issues in the structural DSGE model have been fully corrected for. The lagged values of the relevant integrals involving the interest rate may be expected to correlate with $r_{t-\Delta}$, and hence with $\varepsilon_{Y,t}$ from (16b), although presumably less than without lagging (this is the idea of the instrumentation). Any such correlation between the error terms and the right-hand side variables (even when using fitted values) indicates that part of the endogeneity issue remains. For a full solution and a consistent and asymptotically efficient estimator, we turn to a computationally slightly more demanding procedure, the martingale estimating function (MEF) approach, exploiting the martingale structure of the model. This important next step builds naturally on the above regression-based approach. The latter has the advantage of permitting easy off-the-shelf implemention, and provides useful benchmark estimates and starting values for iterative solution for the optimal MEF estimator.

Let ϕ denote the parameter vector, whether the structural parameters of interest, or simply β from the reduced form. Let $m_t = m_t(\phi)$ denote the N-vector of martingale increments generated by the model, expressed in terms of data and parameters. Specifically, in the AK-Vasicek model with logarithmic utility, we let $m_t = \varepsilon_t = (\varepsilon_{C,t}, \varepsilon_{Y,t}, \varepsilon_{r,t})^\top$ from (16a)-(16b), so $N = 3$. Clearly, m_t is a martingale difference sequence, and from system (13) we have that in terms of data and parameters

$$m_t = \begin{pmatrix} \ln(C_t/C_{t-\Delta}) - \int_{t-\Delta}^{t} r_v dv + \left(\rho + \delta + \frac{1}{2}\sigma^2\right)\Delta \\ \ln(Y_t/Y_{t-\Delta}) - \int_{t-\Delta}^{t} r_v dv + \left(\kappa + \rho + \delta + \frac{1}{2}\sigma^2\right)\Delta - \kappa\gamma \int_{t-\Delta}^{t} 1/r_v dv + \frac{1}{2}\eta^2 \int_{t-\Delta}^{t} 1/r_v^2 dv \\ r_t - (1 - e^{-\kappa\Delta})\gamma - e^{-\kappa\Delta} r_{t-\Delta} \end{pmatrix}, \tag{19}$$

where the integrals are approximated by Riemann sums over days between $t-\Delta$ and t. More general versions of the model give rise to other m_t, some with higher dimension N.

3.3.1 The MEF method

The MEF method differs slightly from the generalized method of moments (GMM) of Hansen (1982). It is at least as efficient as GMM—usually strictly more efficient. It is instructive to start with the GMM, then show how to modify this appropriately, to see how the MEF method comes about. Since m_t is a martingale difference sequence, we have $E_{t-\Delta}(m_t) = 0$. The standard GMM approach is to consider instruments, say z_t, belonging to the information set and hence known at time $t - \Delta$, so that $E_{t-\Delta}(z_t \otimes m_t) = 0$, where $\otimes$ is the Kronecker product. For example, the instruments could be lagged RHS variables from the regressions, $z_t = (1, \int_{t-2\Delta}^{t-\Delta} 1/r_v dv, \int_{t-2\Delta}^{t-\Delta} 1/r_v^2 dv, r_{t-2\Delta})^\top$, since these are all in the information set at $t-\Delta$. Defining $h_t = h_t(\phi) = z_t \otimes m_t$, we have that h_t is of dimension $\dim h = \dim z \times N$, or 12 in the AK-Vasicek model with logarithmic utility. To construct the GMM estimator, let

$$H_T = \frac{1}{T} \sum_{t=1}^{T} h_t \tag{20}$$

be the sample average, evidently a martingale at the true value of the parameter ϕ. Since the unconditional expectation $E(h_t) = 0$, it would be natural to choose the estimator to match the sample analogue H_T of $E(h_t)$ to zero. Typically, $\dim h > \dim \phi$, so it is not possible to solve the equation $H_T = 0$ exactly. Instead, the GMM estimator is defined as the minimizer of $H_T(\phi)^\top W H_T(\phi)$, where W is a weight matrix. Optimal GMM is obtained by using the identity matrix I_T for W in a first step minimization, then using the resulting estimator to calculate a consistent estimate of $var(H_T)^{-1}$ that is used for W in the second step minimization.

To see how the MEF approach differs from GMM, note that the first order conditions for the minimization in GMM are

$$\frac{\partial H_T(\phi)^\top}{\partial \phi} W H_T(\phi) = 0, \tag{21}$$

that is, the same number of zero conditions as number of parameters in ϕ, as it should be. An estimator that is asymptotically equivalent to GMM may be obtained by solving the

$\dim\phi$ equation $G\sum_{t=1}^{T} h_t(\phi) = 0$, where G is an initial consistent estimate of the $\dim\phi \times \dim h$ matrix $\partial H_T(\phi)^\top/\partial\phi \cdot W$ in (21). Thus, G could be based on the first step GMM estimator, just like W. The equations are solved by treating G as fixed and finding ϕ that sets (21) exactly equal to zero, and the result is asymptotically equivalent to optimal GMM. It is now apparent that a more flexible estimation approach obtains by not just solving the equations with a fixed $\dim\phi \times \dim h$ matrix G from the first step (the approach asymptotically equivalent to optimal GMM), but instead allowing a separate $\dim\phi \times \dim h$ matrix each time period, say, g_t. Thus, there are still $\dim\phi$ equations, but they now take the more general form

$$\sum_{t=1}^{T} g_t h_t(\phi) = 0, \tag{22}$$

instead of $G\sum_{t=1}^{T} h_t(\phi) = 0$. Clearly, this is a zero-mean martingale for any choice of weight matrices g_t, which may depend on data through $t-\Delta$. They may also depend on parameters, but here we use initial consistent estimates, i.e., all g_t may be calculated after the first step estimation. The question is how to choose the g_t optimally. If they indeed vary across time, the resulting estimator differs from optimal GMM. The special case $g_t \equiv G$ returns the optimal GMM estimator. The relevant theory for optimal estimators is based on Godambe and Heyde (1987), and the dynamic case (optimal choice of time-varying g_t) is treated in Christensen and Sørensen (2008).

In fact, it is unnecessary to expand m_t to h_t by introducing the instruments z_t in $h_t = z_t \otimes m_t$, since if m_t is used instead of h_t and in fact z_t is needed in the optimal estimator, then z_t will just be part of the optimally chosen g_t. Thus, we leave the problem involving z_t and define the martingale estimating function

$$M_T = \sum_{t=1}^{T} w_t m_t, \tag{23}$$

clearly a zero-mean martingale for any choice of weight matrices w_t, which may depend on data through $t - \Delta$. A martingale estimating function (or MEF) is given by specifying w_t as a series of $d \times N$ matrices, where $d = \dim\phi$. At the true parameter value, $E(M_T) = 0$, and ϕ is estimated by solving the martingale estimating equation

$$M_T(\phi) = 0. \tag{24}$$

The optimal weights are given by

$$w_t = \psi_t^\top (\Psi_t)^{-1}, \tag{25}$$

where Ψ_t is the conditional variance of the martingale increment,

$$\Psi_t = Var_{t-\Delta}(m_t) = E_{t-\Delta}(m_t m_t^\top), \tag{26}$$

and ψ_t the conditional mean of its parameter derivative

$$\psi_t = E_{t-\Delta}(\partial m_t / \partial \phi^\top). \tag{27}$$

The conditioning on information available through $t-\Delta$ requires integrating out with respect to the evolution of the interest rate from $t-\Delta$ through t. This can be computationally more demanding than the regression-based approaches, but it does circumvent the endogeneity problem in the DSGE model. The choice of weights (25) gives the optimal martingale estimating function, across choice of weights w_t. The optimal weights do depend on parameters, but these may be replaced by initial consistent estimates, e.g., from GMM, without altering the asymptotic behavior of the martingale estimate. This is consistent (in particular, the endogeneity issue is resolved) and asymptotically normal,

$$\sqrt{T}(\hat{\phi} - \phi) \to \mathcal{N}(0, V_{MEF}), \tag{28}$$

with asymptotic variance-covariance matrix given by

$$V_{MEF} = \left(E(\psi_t^\top (\Psi_t)^{-1} \psi_t)\right)^{-1}, \tag{29}$$

consistently estimated by the inverse sample average $\hat{V}_{MEF} = \left(T^{-1} \sum_{t=1}^{T} \psi_t^\top (\Psi_t)^{-1} \psi_t\right)^{-1}$. If $\phi = \beta$, then ψ_t is block-diagonal with $x_{j,t}$ in the j'th diagonal block, $j = C, Y, r$. When ϕ consists of the structural parameters, ψ_t is this block-diagonal matrix post-multiplied by the Jacobian of the transformation $\omega_1(\phi)$ from structural parameters to β. Again, in the AK-Vasicek model with log utility, this Jacobian has rank five, so five structural parameters may be identified when $N = 3$. The martingale difference m_t may be expanded with $\varepsilon_{C,t}^2 - \sigma^2 \Delta$, and possibly $\varepsilon_{r,t}^2 - \eta^2(1 - e^{-2\kappa\Delta})/(2\kappa)$, to identify one more structural parameter, i.e., separating $\rho + \delta$ and σ^2, and ψ_t then has one more column.

3.3.2 Comparison of MEF and GMM

Before applying the MEF method to the DSGE model, let us briefly return to the comparison between MEF and optimal GMM. Obviously, the GMM estimator is consistent, and the standard consistent estimate of the asymptotic variance takes the form $\hat{V}_{GMM} = \left((T^{-1} \sum_{t=1}^{T} \partial h_t / \partial \phi^\top)^\top (T^{-1} \sum_{t=1}^{T} h_t h_t^\top)^{-1} (T^{-1} \sum_{t=1}^{T} \partial h_t / \partial \phi^\top)\right)^{-1}$. Sometimes, a Newey and West (1987) correction is used in the middle matrix, the estimate of $var(h_t)$, but it is unnecessary under the null that m_t and hence h_t is a martingale difference sequence, and in any case it makes no difference for the comparison. In particular, except in the special case where the two estimators coincide, the MEF estimator is strictly more efficient than optimal GMM,

$$\hat{V}_{MEF} < \hat{V}_{GMM},$$

in the partial order of positive semi-definite matrices. This is essentially a generalized Cauchy-Schwartz inequality, once it is recognized that h_t may be used for m_t in the MEF case (the resulting MEF estimators based on h_t and m_t coincide, as the weights if necessary incorporate z_t, following the above discussion). Specifically, we always have $V_{GMM} = \left(E(\partial h_t/\partial\phi^\top)^\top var(h_t)^{-1}E(\partial h_t/\partial\phi^\top)\right)^{-1}$, and by iterated expectations and using h_t for m_t we have $\psi_t = E_{t-\Delta}(\partial h_t/\partial\phi^\top)$, $\Psi_t = E_{t-\Delta}(h_t h_t^\top)$, and therefore $E(\partial h_t/\partial\phi^\top) = E(\psi_t)$, $var(h_t) = E(\Psi_t)$. It follows that the efficiency comparison is simply

$$V_{MEF} = \left(E(\psi_t^\top(\Psi_t)^{-1}\psi_t)\right)^{-1} < \left(E(\psi_t)^\top E(\Psi_t)^{-1}E(\psi_t)\right)^{-1} = V_{GMM}.$$

The asymptotic variance of the martingale estimator is smaller than that of GMM because the expectation is taken after multiplying the relevant matrices, instead of before, as in GMM.

When does the MEF method reduce to GMM? The GMM estimator solves $H_T(\phi) = \sum_{t=1}^T h_t(\phi) = 0$ in the exactly identified case, and (up to asymptotic equivalence) $G\sum_{t=1}^T h_t(\phi) = 0$ in the overidentified case where $\dim h > \dim\phi$, with $G = \partial H_T(\phi)^\top/\partial\phi \cdot W$. The question is when the MEF estimator solving $M_T(\phi) = \sum_{t=1}^T \psi_t^\top(\Psi_t)^{-1}m_t(\phi) = 0$ with $\psi_t = E_{t-\Delta}(\partial m_t/\partial\phi^\top)$ and $\Psi_t = E_{t-\Delta}(m_t m_t^\top)$ takes this standard GMM form. This requires that the researcher has started out with either (i) moments not given by the N-vector of martingale differences m_t, and also not by $h_t = z_t \otimes m_t$, for arbitrary z_t in the information set at $t-\Delta$, but instead given by the $\dim\phi$ vector $\psi_t^\top(\Psi_t)^{-1}m_t(\phi)$; or, (ii), moments in fact given by the N-vector m_t,in a situation with $N < \dim\phi$, and where an expansion of moment conditions from the original N-vector m_t to $h_t = z_t \otimes m_t$ happens to deliver the $\dim\phi$ vector $h_t = \psi_t^\top(\Psi_t)^{-1}m_t(\phi)$. In addition, the vector z_t that makes this happen must be in the information set at $t-\Delta$. Since the conditional mean ψ_t and conditional variance Ψ_t typically depend on parameters, this case rarely occurs for standard instrumental variables z_t in the data set. Firstly, it would require that $\dim\phi = \dim z \cdot N$. Secondly, writing $h_t = z_t \otimes m_t = (z_t \otimes I_{\dim z})\, m_t$, it also requires that $\psi_t^\top(\Psi_t)^{-1}$ has very special structure, i.e., it can be represented in the Kronecker product form $z_t \otimes I_{\dim z}$.

In all other cases, the MEF and GMM estimators differ, with $V_{MEF} < V_{GMM}$, i.e., the martingale estimator is asymptotically strictly more efficient than GMM. In our specific DSGE applications, we see below that $\psi_t^\top(\Psi_t)^{-1}$ is complicated, certainly not on Kronecker product form (case (ii)), and it is also highly unlikely that a researcher would a priori start with moment conditions $\psi_t^\top(\Psi_t)^{-1}m_t(\phi)$ rather than $m_t(\phi)$ (case (i) above), except if purposefully applying the MEF rule of always transforming from any arbitrary moment m_t (univariate or multivariate) to $\psi_t^\top(\Psi_t)^{-1}m_t(\phi)$ at the outset. In this sense, MEF could be considered GMM with optimal (typically parameter-dependent) instruments, namely, using

$\psi_t^\top(\Psi_t)^{-1}$ instead of the standard but arbitrary $z_t \otimes I_{\dim z}$.

3.3.3 A martingale estimating function with three moment restrictions

For illustration, we report the functional form of the martingale estimating function for the AK-Vasicek model with logarithmic utility. Let $m_t = \varepsilon_t = (\varepsilon_{C,t}, \varepsilon_{Y,t}, \varepsilon_{r,t})^\top$ be the 3-vector of error terms (16), clearly a martingale difference sequence. This may be expressed in terms of data and parameters as in (19), where the integrals are approximated by summation over days between $t-\Delta$ and t. This allows computing m_t at trial parameter values. To construct the MEF (24), we need the weights w_t in (25), which depend on the conditional mean of the parameter derivatives, ψ_t, and the conditional variance, Ψ_t, of m_t. Here, we have the conditional variances $\Psi_{t,11} = \sigma^2\Delta$, $\Psi_{t,22} = \eta^2 E_{t-\Delta}(\int_{t-\Delta}^t 1/r_v^2 dv) + \sigma^2\Delta$, and $\Psi_{t,33} = \eta^2(1-e^{-2\kappa\Delta})/(2\kappa)$. Similarly, the conditional covariances are $\Psi_{t,12} = \sigma^2\Delta$, $\Psi_{t,13} = 0$, and $\Psi_{t,23} = \eta^2 e^{-\kappa\Delta} E_{t-\Delta}\left((\int_{t-\Delta}^t 1/r_v dB_v)(\int_{t-\Delta}^t e^{\kappa(v-(t-\Delta))} dB_v)\right)$. Since analytical expressions are not available, we use Euler approximations for $\Psi_{t,22}$ and $\Psi_{t,23}$,

$$\Psi_t = \begin{pmatrix} \sigma^2\Delta & \sigma^2\Delta & 0 \\ \sigma^2\Delta & \sigma^2\Delta + \eta^2\Delta/r_{t-\Delta}^2 & \eta^2 e^{-\kappa\Delta}\Delta/r_{t-\Delta} \\ 0 & \eta^2 e^{-\kappa\Delta}\Delta/r_{t-\Delta} & \frac{1}{2}\eta^2(1-e^{-2\kappa\Delta})/\kappa \end{pmatrix}. \tag{30}$$

Consistency and the expression for the asymptotic variance are unaffected by this approximation. Using martingale increments (19), we get the conditional mean of the derivatives with respect to the parameter vector $\phi = (\kappa, \gamma, \eta, \rho+\delta, \sigma)^\top$ as

$$\psi_t = \begin{pmatrix} 0 & 0 & 0 & \Delta & \sigma\Delta \\ \Delta - \gamma E_{t-\Delta}\int_{t-\Delta}^t 1/r_v dv & -\kappa E_{t-\Delta}\int_{t-\Delta}^t 1/r_v dv & \eta E_{t-\Delta}\int_{t-\Delta}^t 1/r_v^2 dv & \Delta & \sigma\Delta \\ -\Delta e^{-\kappa\Delta}\gamma + \Delta e^{-\kappa\Delta} r_{t-\Delta} & -(1-e^{-\kappa\Delta}) & 0 & 0 & 0 \end{pmatrix}. \tag{31}$$

For the conditional expectations we first interchange the order of integration in (31), then use the deterministic Taylor expansion (e.g. Aït-Sahalia, 2008), which for $s \geq u$ is

$$E(g(r_s)|r_u) = \sum_{i=0}^k \frac{\Delta^i}{i!} A^i g(r_u) + O(\Delta^{k+1}), \tag{32}$$

where A is the infinitesimal generator in the Vasicek model, $Ag(x) = \kappa(\gamma - x)g'(x) + \frac{1}{2}\eta^2 g''(x)$. The function $g(\cdot)$, for example $g(x) = 1/x$ in $\psi_{t,21}$, must be sufficiently smooth. For illustration, for $\psi_{t,21}$ a first-order Taylor expansion, $k=1$, yields

$$\begin{aligned} \int_{t-\Delta}^t E_{t-\Delta}(1/r_v)dv &\approx \int_{t-\Delta}^t \left(1/r_{t-\Delta} + (v-(t-\Delta))\right)\left(-\kappa(\gamma - r_{t-\Delta})/r_{t-\Delta}^2 + \eta^2/r_{t-\Delta}^3\right)\right) dv \\ &= \Delta/r_{t-\Delta} - (t-\Delta)\left(-\kappa(\gamma - r_{t-\Delta})/r_{t-\Delta}^2 + \eta^2/r_{t-\Delta}^3\right)\Delta \\ &\quad + \tfrac{1}{2}(t^2 - (t-\Delta)^2)\left(-\kappa(\gamma - r_{t-\Delta})/r_{t-\Delta}^2 + \eta^2/r_{t-\Delta}^3\right) \\ &= \Delta/r_{t-\Delta} - \left(\kappa(\gamma - r_{t-\Delta})/r_{t-\Delta}^2 - \eta^2/r_{t-\Delta}^3\right)\tfrac{1}{2}\Delta^2. \end{aligned}$$

Similarly, for $\psi_{t,23}$ a first-order Taylor expansion yields

$$\int_{t-\Delta}^{t} E_{t-\Delta}(1/r_v^2)dv \approx \Delta/r_{t-\Delta}^2 - \left(2\kappa(\gamma - r_{t-\Delta})/r_{t-\Delta}^3 - 3\eta^2/r_{t-\Delta}^4\right)\tfrac{1}{2}\Delta^2.$$

Using a first-order Taylor expansion for $\psi_{t,21}$, $\psi_{t,22}$, and $\psi_{t,23}$ in (31), it can be easily verified that the transpose of the conditional mean of parameter derivatives $\psi_t^\top$ reads

$$\begin{pmatrix} 0 & \Delta - \gamma\left(\Delta/r_{t-\Delta} - \left(\kappa(\gamma - r_{t-\Delta})/r_{t-\Delta}^2 - \eta^2/r_{t-\Delta}^3\right)\tfrac{1}{2}\Delta^2\right) & -\Delta e^{-\kappa\Delta}\gamma + \Delta e^{-\kappa\Delta} r_{t-\Delta} \\ 0 & -\kappa\left(\Delta/r_{t-\Delta} - \left(\kappa(\gamma - r_{t-\Delta})/r_{t-\Delta}^2 - \eta^2/r_{t-\Delta}^3\right)\tfrac{1}{2}\Delta^2\right) & -(1-e^{-\kappa\Delta}) \\ 0 & \eta\left(\Delta/r_{t-\Delta}^2 - \left(2\kappa(\gamma - r_{t-\Delta})/r_{t-\Delta}^3 - 3\eta^2/r_{t-\Delta}^4\right)\tfrac{1}{2}\Delta^2\right) & 0 \\ \Delta & \Delta & 0 \\ \sigma\Delta & \sigma\Delta & 0 \end{pmatrix}. \tag{33}$$

This completes the construction of the martingale estimating function $M_T = \sum_t \psi_t^\top (\Psi_t)^{-1} m_t$. The condition $M_T(\phi) = 0$ involves the same number of equations and unknowns, and is solved exactly for the optimal estimator $\widehat{\phi}$. The asymptotic distribution is given by (28)-(29).

3.3.4 Latent Variables and Missing Data

So far, we have considered the case where all variables in the system are observable, albeit at different frequencies. Our approach can be generalized to the case of latent variables, as we now illustrate. The ability to accommodate latent variables is important for applications to a number of important DSGE models, e.g., models with stochastic volatility, stochastic discount rates, labor and capital in a production function with stochastic TFP, etc.. In this section, we consider two representative cases, set in the context of the AK-Vasicek model: (i) The interest rate r_t is unobserved, i.e., latent; and (ii) output Y_t is observed at a lower (say, quarterly) frequency than consumption C_t (observed monthly). The first case serves to illustrate the approach to the case of missing data series. For example, expected inflation and hence the real rate of interest may be treated as missing. This case also covers models involving unobserved state variables, such as stochastic volatility. Case (ii) reinforces our use of data sampled at mixed frequencies. For example, output may be proxied by industrial production at the monthly frequency, but it may be of interest to compare with results using actual output, available only quarterly. In the latter case, consumption need not be aggregated to quarterly frequency.

The basis of the MEF approach with complete data is that the condition $E(M_T) = 0$ is satisfied at the true parameter value, where $M_T = \sum_{t=1}^T w_t m_t$. In the incomplete data setting (i) where r_t is latent, define $\mathcal{F}_t$ as the information set generated by $\{C_s, Y_s\}_{s=1}^t$ (but not the missing interest rates). By $E(M_T) = 0$ and iterated expectations, we have both $E\left(\sum_t w_t E(m_t|\mathcal{F}_t)\right) = 0$ and $E\left(\sum_t w_t E(m_t|\mathcal{F}_{t-1})\right) = 0$, for weights w_t depending only on information through $t-1$. Thus, in the estimation, we may replace additive terms in the moments m_t by their conditional expectations given either $\mathcal{F}_t$ or $\mathcal{F}_{t-1}$. In particular, without daily interest rate data, this allows replacing the integrals involving the interest rate by conditional expectations given monthly interest rate proxies, based on the information set. In the stochastic AK-Vasicek model, this allows deriving moments for estimation, say, m_t^*, given by

$$E\left(m_t|\mathcal{F}_{t-\Delta}\right) = \\ \left(\begin{array}{c} \ln(C_t/C_{t-\Delta}) - E\left(\int_{t-\Delta}^t r_v dv | r_{t-\Delta}^*\right) + \left(\rho+\delta+\frac{1}{2}\sigma^2\right)\Delta \\ \ln(Y_t/Y_{t-\Delta}) + \left(\kappa+\rho+\delta+\frac{1}{2}\sigma^2\right)\Delta - E\left(\int_{t-\Delta}^t r_v dv + \kappa\gamma\int_{t-\Delta}^t 1/r_v dv - \frac{1}{2}\eta^2\int_{t-\Delta}^t 1/r_v^2 dv | r_{t-\Delta}^*\right) \\ r_t^* - (1-e^{-\kappa\Delta})\gamma - e^{-\kappa\Delta} r_{t-\Delta}^* \end{array}\right), \tag{34}$$

where $r_{t-\Delta}^*$ is an interest rate proxy based on consumption and income data through $t-\Delta$. Here, $\Delta = 1$ is used in the empirical work. From earlier, the model implies $K_t = Y_t/r_t$ and $C_t = \rho K_t$, so a valid proxy at the macro frequency is $r_t^* = \rho Y_t/C_t$. In the simulated MEF approach, the conditional expectations of the integrals in (34) are computed by integrating out the latent interest rate process r_v by simulation. Thus, each integral involves drawing a path for r_v from $dr_v = \kappa(\gamma - r_v)dv + \eta dB_v$ using an Euler scheme from $v = t-\Delta$ to t, starting at the proxy value for $r_{t-\Delta}^*$, and the expectation is formed by averaging over paths. More complicated models may be treated similarly, in each case using the model to back out latent state variables from observables. The interest rate (latent state variable) is similarly integrated out of $w_t = \psi_t^\top(\Psi_t)^{-1}$, or, in the specific case, $r_{t-\Delta}$ is simply replaced by its proxy $r_{t-\Delta}^*$ in the expressions (33) and (30) for $\psi_t^\top$ and Ψ_t. In the iterative solution of the estimating equation $\sum_t \psi_t^\top(\Psi_t)^{-1}m_t^* = 0$, the parameter dependence (in our particular model, through ρ) of the implied state variables is accounted for. The approach is general, e.g., in the stochastic volatility case, too, the relevant state variable (volatility) would be implied out of the available data for given trial parameter values and at the given sampling frequency, then integrated out between sampling periods using the model.

The mixed frequency case (ii) where output is only available quarterly is slightly different. Here, a complete (monthly) output proxy series Y_t^* is simply constructed recursively by letting $Y_t^* = Y_t$ in the (quarterly) periods where output data are available, and

$$Y_t^* = \exp\left(\ln(Y_{t-\Delta}^*) + \int_{t-\Delta}^t r_v dv - \left(\kappa+\rho+\delta+\frac{1}{2}\sigma^2\right)\Delta + \kappa\gamma\int_{t-\Delta}^t 1/r_v dv - \frac{1}{2}\eta^2\int_{t-\Delta}^t 1/r_v^2 dv\right)$$

in the intra-quarter periods where output is missing. This is not simulation, but model consistent prediction. In particular, the resulting proxy series Y_t^* depends on the parameters.

The proxy series is is now substituted for Y_t in the original estimating equation $\sum_t w_t m_t = 0$. When solving for the parameter estimates, the dependence of the constructed output proxy series on trial parameter values is again accounted for.

In both cases, (i) and (ii), the approach is akin to filtering. Thus, in the simulated MEF approach, case (i), m_t in the estimating equation is recast in terms of a set of conditional expectations or filtered predictions, given information actually available. In (ii), as well, a conditional prediction given the actual observations is used to replace missing data. In both cases, standard errors may be adjusted using the bootstrap.

4 Simulation Study

To assess the estimation methods from the previous section we run a simulation study. We first detail the set-up of our analysis. As in the previous section, our illustration is based on the AK-Vasicek model with logarithmic utility. We do point out, however, the differences for the case of the CRRA preferences. Finally, we discuss the findings for the AK-Vasicek model with logarithmic or general CRRA preferences.

4.1 Set-Up

We simulate 25 years of both monthly and quarterly data from the AK-Vasicek model as given in Section 2.4.1. We use simple Euler approximations to the differential equations in (11). The step length of the Brownian terms is taken as 1/3000. This corresponds to dividing each of the 12 months of the year into 25 days, each in turn consisting of 10 periods.

There are two further computational issues when simulating from the AK-Vasicek model with log preferences: Obtaining the integrals involving the interest rate and initialization of the simulations. Concerning the first issue, we obtain the monthly integrals over the interest rate, denoted with $\int_{t-\Delta}^{t} g(r_v)dv$ where $\Delta = 1/12$ as also in Section 3.1, by taking the average of the functions $g(r_v)$ over the 25 simulated days per month. For example, $\int_{t-\Delta}^{t} 1/r_v dv$ for the monthly simulations is approximated by $(\sum_{i=1}^{25} 1/r_{t-\Delta+i\Delta/25})\Delta/25$. For the quarterly simulated data we use a similar approximation, but now over the 75 days in the Euler approximation. Concerning the second issue, we initialize the variables as follows: $\ln(Y_0) = 0$, $r_0 = \gamma$, and $\ln(C_0) = \ln(\rho \times Y_0/r_0)$.

When simulating from the AK-Vasicek model with CRRA preferences, there are only few differences compared to the model with log preferences. Again, an Euler approximation is used, now based on (38) in the Appendix. The main difference is that an additional term is approximated using the Euler scheme, the third term $\int_{t-\Delta}^{t}(1 - C_v/Y_v)r_v dv$ in (38b) (cf. also Appendix A.2).

We generate 1,000 data sets and estimate the parameters according to the approaches of Section 3. In particular, we estimate the parameters using the OLS, FGLS-SUR, FGLS-SUR-IV, and MEF methods. In the first three cases, we use the minimum distance approach to go from the reduced form estimates to the structural model parameters; in our exposition we focus on the latter. We choose the data generating process (DGP) parameter values from the numerical solution given in Appendix A.2.1. That is, for the AK-Vasicek model with CRRA preferences we use $\kappa = 0.2$, $\gamma = 0.1$, $\eta = 0.005$, $\rho = 0.05$, $\delta = 0.05$, $\sigma = 0.05$, $\theta = 2$, $\bar{\tau} = 0.207$, and $\bar{\pi} = 1.021$. For the model with log preferences we take the same values for the common parameters (while $\theta = 1$, $\bar{\tau} = 0$, $\bar{\pi} = 1$).

4.2 Results

[insert Table 1]

Table 1 provides the results for the simulation study of the AK-Vasicek model with log preferences. In the first column we list the parameter values as they are used in the data generating process (DGP), in columns 2-5 the estimates obtained on the simulated monthly data, and in columns 6-9 the estimates for the quarterly data. For all four estimation methods we provide the median estimate of each parameter, and below this the interquartile range of the 1,000 estimates. For the three regression-based estimation methods, the estimates of γ and $\rho + \delta + \sigma^2/2$ are remarkably close to the values used in the DGP. The mean-reversion parameter κ is slightly more problematic: 0.2 is used in the DGP, and the estimates are 0.23, 0.16, and 0.22 for the OLS, FGLS-SUR, and FGLS-SUR-IV methods, respectively, using monthly data. Given the relatively wide interquartile range, this is still within reasonable distance. Thus, the estimates may be noisy, but not severely biased. The estimates of the short rate innovation variance η do deviate from the DGP value, and in case of the OLS and FGLS-SUR methods, the median estimates are relatively far from the DGP value, given the interquartile range. Especially the OLS method produces a considerable positive bias in η. Similar results hold for the quarterly data.

The fifth and ninth column of Table 1 show the median estimate obtained based on the MEF approach. With MEF we are able to identify σ separately from $\rho + \delta$, whereas the regression approaches reported only identify $\rho+\delta+\sigma^2/2$. The martingale estimating function approach is able to successfully estimate all parameters from the model. In particular, the median estimates of γ, η, $\rho + \delta$, and σ are very close to the DGP values. The κ estimate are slightly higher than the DGP value in the monthly case, but somewhat closer using quarterly data. In both cases, the median estimate and the DGP value are close relative to the interquartile range.

In unreported experiments,[12] we implemented the expanded minimum distance methods from Section 3.2.5, using the residual variance from the consumption or both the consumption and interest rate equation as additional moments along with β in the regression-based approach, thus allowing separate identification of σ and $\rho + \delta$ in this case, too. The changes in results were negligible for the reported parameters when only expanding with the consumption residual variance, but the upward bias in the η estimate was reduced by including the interest rate residual variance. The σ and $\rho + \delta$ median estimates were similar to those from the MEF approach. Overall, the preferred approach in terms of bias and interquartile range appears to be the MEF, except for the estimation of κ based on monthly data.

[insert Figure 1]

In Figure 1 we provide the histograms of the 1,000 estimates that we obtain for the parameters using the MEF approach on both monthly (Panel (A)) and quarterly (Panel (B)) data. The figure confirms the findings from the table: γ, η and $\rho + \delta$ are centered close to the DGP values. In addition, it becomes clear that the mode of the histograms for κ and σ are in fact quite close to the DGP values, but the estimates are skewed, thus causing the difference between median estimates and DGP values reported in Table 1.

[insert Table 2]

Table 2 provides the output for the simulation study of the AK-Vasicek model with CRRA preferences. Using the regression-based approach, most parameters are estimated inaccurately, γ being one exception. In particular $\bar{\tau}$, θ, and η are estimated with considerable error. The difficulty in estimating these parameters appears in all three regression-based methods. Even addressing cross-equation correlation through SUR or endogeneity of the integral terms does not yield much improvement.[13]

[insert Figure 2]

In contrast, the estimates obtained with the MEF approach as reported in Table 2 are relatively accurate. First, the approach allows identification of all nine model parameters. The regression-based methods only identify seven parameters or parameter functions, and expanding with error variances in the minimum distance approach no longer helps identifying more parameters since the errors are more complicated (see (40a)-(40c)). Second, using the MEF, all median parameter estimates are relatively close to those of the DGP. Even

[12] Results are available from the authors on request.

[13] Restricting either θ or $\bar{\tau}$ does not improve the estimates of the other parameters.

parameters that are historically found difficult to obtain from data, such as θ, are estimated accurately. In Figure 2 we provide the histograms of the 1,000 obtained series. As in the histograms for the model with logarithmic preferences, the mode of each histogram is close to the DGP value of the relevant parameter.

Taken together, the simulation study indicates that the martingale estimating function approach is successful in obtaining parameter estimates from the data. The regression-based methods exhibit reasonable performance for the AK-Vasicek model with logarithmic preferences, but encounter difficulties in case of the more complicated model with general CRRA preferences.

5 Data and Results

In this section we estimate the AK-Vasicek model with logarithmic preferences from Section 2.4.1 and with CRRA preferences from Section 2.4.2. We use the techniques from Section 3, and employ both U.S. macro and financial data in our approaches.

5.1 Data

[insert Figure 3]
[insert Figure 4]

To estimate the systems (11) and (38) we need data on production, consumption, and the short rate. We obtain these data for the US from the Federal Reserve Economic Dataset (FRED), maintained by the Federal Reserve Bank of St. Louis. To measure production, we use both real Industrial Production (IP), available at the monthly level, and real Gross Domestic Product (GDP), available at the quarterly level. In Panel (A) of Figures 3 (monthly data) and 4 (quarterly data) we show the time series plots of the growth rates of the variables, the data actually used in our analysis. Our data set spans the period from January 1982 to December 2000.

We combine the data on these aggregate macro series with financial data at higher frequency, in particular, the short rate. The short rate as a theoretical concept in principle corresponds to infinitesimal term to maturity. In applied work, it is sometimes treated as a latent variable that needs to be filtered from observed yield time series (e.g., De Jong, 2000). Chapman, Long, and Pearson (1999) show that when the short rate is proxied by available short-term interest rates, this does not lead to economically significant problems. We follow this approach, and take the 3-month interest rate as a proxy for the short rate. This rate is available from the FRED data set at daily frequency. We use this series to obtain our monthly and quarterly series by taking the last observation in the relevant period.[14] Panel (B) of Figures 3 and 4 shows the interest rate series. In both the monthly and quarterly series, the general downward trend of the series is evident. This is combined with multiple increasing and decreasing interest rate cycles in our sample period.

Finally, we use the above series to compute approximations to the integrals for both models. For the integrals that only depend on the short rate we approximate the monthly and quarterly series of integrals using the daily spot rate observations. Following the systems (11) and (38), we approximate three integrals: $\int_{t-\Delta}^{t} g(r_v)dv \approx \Delta/P \sum_{i=1}^{P} g(r_{t-\Delta+i\Delta/P})$, where $r_{t-\Delta+i\Delta/P}$ is the 3-month interest rate on day i of period t, and P the number of days in the period between $t-\Delta$ and t. For the model with CRRA preferences there is an additional integral to consider, the third term in Equation (38b). We approximate this integral using $\int_{t-\Delta}^{t}(1 - C_v/Y_v)r_v dv \approx \Delta(1 - C_{t-\Delta}/Y_{t-\Delta})r_{t-\Delta}$.[15] Panel (C) of Figures 3 and 4 show the resulting time series of approximations to the integrals.

5.2 Results

[insert Table 3]

Table 3 provides the estimates of the AK-Vasicek model with log preferences for both monthly and quarterly data (using industrial production and GDP for output, respectively) of the OLS, FGLS-SUR, FGLS-SUR-IV, and MEF approaches. The regression-based estimation methods provide fairly similar estimates for each of the data frequencies (monthly and quarterly). The short rate is mean-reverting with speed parameter κ around 0.17 (0.05 to 0.11 for quarterly data), the long-term rate it reverts to is about 9% and the volatility of the short rate innovation is about 2%. The sum of structural parameters $\rho + \delta + \sigma^2/2$ is estimated at 0.035. The MEF approach provides somewhat different results. Most notably, the mean reversion parameter estimate is much higher, at 7.9 in monthly data, and 11.6 in quarterly. In addition, σ is separately identified and estimated to about 0.9%. Based on the asymptotic t-statistics, all regression-based estimates are insignificant in monthly data, and only γ (the long run interest rate) turns significant in quarterly data. In contrast, all estimates are strongly significant at both data frequencies when using the MEF approach. In unreported results, we implemented the expanded regression-based minimum distance methods

[14] A disadvantage is that the interest rate series is in nominal terms, rather than real. There are a number of ways to overcome this, which is part of our research agenda. Dealing with this issue is not straightforward, as in periods with short-term nominal rates very low (early and late 2000's) real rates have been negative.

[15] For monthly data we only have the industrial production index, and not actual output. We obtain a monthly output estimate by for each year weighting the annual GDP (calculated as the average of the four quarterly GDP figures in a year) using as weights the monthly IP contribution to the annual total.

including in addition the error variances, but estimated precision remained less than under the MEF approach.

[insert Table 4]

Table 4 provides the estimates of the AK-Vasicek model with CRRA preferences using the four methods. The results of the simulation study generated serious worries about the performance of the regression-based methods for the general case. Though most of the parameter estimates are roughly similar across approach and data frequency, the unstable θ estimate may indicate such difficulties. Due to these reasons we focus on the estimates obtained using the MEF approach. We highlight a few findings. When allowing for general CRRA preferences, the MEF approach no longer yields a large estimate of κ. Instead, mean reversion of the interest rate is modest and insignificant, with κ estimated at about 0.4 for monthly and 0.6 for quarterly data. The mean interest rate is estimated at about 5.3%, more than one percentage point below the estimate in the log utility case, and the short rate volatility at 3.3%. The point estimate for risk aversion θ is about 6 for both data frequencies, although with a great degree of uncertainty. The MEF approach separates $\rho + \delta$ from σ under log utility and further separates ρ and δ in the general CRRA case, but the latter separation seems less successful in the data. Thus, perhaps surprisingly, the depreciation rate δ turns negative at both data frequencies, and the subjective rate of time preference ρ seems too high, at 10.9% in monthly data, and 23.6% in quarterly. The volatility of the stochastic depreciation rate σ is estimated to about 0.4 in monthly data, but is only borderline significant in quarterly data (asymptotic t-statistic of 1.75).

The parameter estimates in the AK-Vasicek model with CRRA preferences are somewhat different from the estimates obtained in the model with logarithmic preferences. This is not unexpected, as the latter model is a restricted version of the former. In order to obtain the logarithmic case out of the more general CRRA case, the parameter restrictions that must hold are $\theta = 1$, $\bar{\tau} = 0$, and $\bar{\pi} = 1$. This suggests a drop in degrees of freedom of three, going from nine free parameters to six. However, only five parameters are identified under the logarithmic null (ρ and δ cannot be separated), thus suggesting a drop of four degrees of freedom. Due to the relatively complicated model structure, the setting is one where a nuisance parameter (say, ρ or δ) is only identified under the alternative. This type of situation has been studied by Andrews and Ploberger (1994) and Hansen (1996), and the asymptotic distribution of the Wald-type test on the three parameters is non-standard. In our case, the evidence is against the restrictions, and particularly the wealth elasticity of consumption $\bar{\pi}$ at 0.25 in monthly data and 0.82 in quarterly is strongly significant and, in particular, significantly below the unit value corresponding to logarithmic preferences.

The interest sensitivity of consumption $\bar{\tau}$ is related to the marginal rate of intertemporal substitution and in particular vanishes with log utility, and this is better in line with our results.

6 Conclusion

The literature has been surprisingly quiet on the links between macroeconomics and finance, though anecdotal evidence - such as the recent financial crisis - clearly shows that financial markets and the real economy are closely linked.

This paper describes both regression-based procedures and the asymptotically efficient martingale estimating function approach in order to estimate the structural parameters of continuous-time DSGE models using macroeconomic and financial market data. We illustrate our approach by solving and estimating a stochastic AK model with mean-reverting interest rates. Our results for both simulated and empirical data are very promising and show that financial market and macro data can indeed be used jointly to facilitate the estimation of structural parameters in continuous-time versions of the DSGE models. Overall, on the methodological side, our work suggests that the martingale method is preferred over the regression-based. It allows identifying all structural parameters, and estimates are more precise, numerically stable, and economically meaningful. On the substantive side, our results indicate a long run mean of the short rate of interest around 5% with a 3% volatility annually and weak mean reversion, as well as higher relative risk aversion than logarithmic. The wealth elasticity of consumption is significantly below unity, the value corresponding to log preferences, whereas the interest rate elasticity of consumption differs insignificantly from the zero value implied by the log case. Development of further models in this class, extending the Cox, Ingersoll, and Ross (1985a) framework to more elaborate specifications, and formal testing of these is part of our research agenda.

A Appendix

A.1 The Bellman equation and the Euler equation

As a necessary condition for optimality, Bellman's principle gives at time s

$$\rho V(K_s, A_s) = \max_{C_s} \left\{ u(C_s, A_s) + \frac{1}{dt} E_s dV(K_s, A_s) \right\}.$$

Using Itô's formula yields

$$\begin{aligned} dV &= V_K dK_s + V_A dA_s + \tfrac{1}{2}\left(V_{AA}\eta(A_s)^2 + V_{KK}\sigma^2 K_s^2\right) dt \\ &= ((r_s-\delta)K_s + w_s - C_s)V_K dt + V_K \sigma K_s dZ_s + V_A \mu(A_t)dt + V_A \eta(A_s)dB_s \\ &\quad + \tfrac{1}{2}\left(V_{AA}\eta(A_s)^2 + V_{KK}\sigma^2 K_s^2\right) dt. \end{aligned}$$

Using the properties of stochastic integrals, we may write

$$\begin{aligned} \rho V(K_s, A_s) &= \max_{C_s} \{u(C_s, A_s) + ((r_s-\delta)K_s + w_s - C_s)V_K \\ &\qquad + \tfrac{1}{2}\left(V_{AA}\eta(A_s)^2 + V_{KK}\sigma^2 K_s^2\right) + V_A \mu(A_s)\} \end{aligned}$$

for any $s \in [0, \infty)$. Because it is a necessary condition for optimality, we obtain the first-order condition (8), which makes optimal consumption a function of the state variables.

For the *evolution of the costate* we use the maximized Bellman equation

$$\begin{aligned} \rho V(K_t, A_t) &= u(C(K_t, A_t), A_t) + ((r_t-\delta)K_t + w_t - C(K_t, A_t))V_K \\ &\quad + \tfrac{1}{2}\left(V_{AA}\eta(A_t)^2 + V_{KK}\sigma^2 K_t^2\right) + V_A \mu(A_t), \end{aligned} \tag{35}$$

where $r_t = r(K_t, A_t) = Y_K$ and $w_t = w(K_t, A_t) = Y_L$ to compute the costate,

$$\begin{aligned} \rho V_K &= ((r_t-\delta)K_t + w_t - C_t)V_{KK} + (r_t-\delta)V_K \\ &\quad + \tfrac{1}{2}\left(V_{AAK}\eta(A_t)^2 + V_{KKK}\sigma^2 K_t^2\right) + V_{KK}\sigma^2 K_t + V_{AK}\mu(A_t). \end{aligned}$$

Collecting terms we obtain

$$\begin{aligned} (\rho - (r_t-\delta))V_K &= ((r_t-\delta)K_t + w_t - C_t)V_{KK} \\ &\quad + \tfrac{1}{2}\left(V_{AAK}\eta(A_t)^2 + V_{KKK}\sigma^2 K_t^2\right) + V_{KK}\sigma^2 K_t + V_{AK}\mu(A_t). \end{aligned} \tag{36}$$

Using Itô's formula, the costate obeys

$$\begin{aligned} dV_K &= V_{AK}\mu(A_t)dt + V_{AK}\eta(A_t)dB_t \\ &\quad + \tfrac{1}{2}\left(V_{KAA}\eta(A_t)^2 + V_{KKK}\sigma^2 K_t^2\right) dt \\ &\quad + ((r_t-\delta)K_t + w_t - C_t)V_{KK}dt + V_{KK}\sigma K_t dZ_t, \end{aligned}$$

where inserting (36) into the last expression yields

$$dV_K = (\rho - (r_t-\delta))V_K dt - V_{KK}\sigma^2 K_t dt + V_{AK}\eta(A_t)dB_t + V_{KK}\sigma K_t dZ_t,$$

which describes the evolution of the costate variable. As a final step, we insert the first-order condition (8) to obtain the Euler equation (9).

As shown in Posch (2009), the model has a closed-form solution for $\theta = 1$, and the value function is $V(K_t, A_t) = \ln K_t/\rho + f(A_t)$, where $f(A_t)$ solves a simple ODE, which in turn depends on the functional forms of $\eta(A_t)$ and $\mu(A_t)$. The idea of this proof is as follows. We use a guess of the value function and obtain conditions under which both the maximized Bellman equation (35) and the first-order condition (8) are fulfilled. Our guess is

$$V(K_t, A_t) = \mathbb{C}_1 \ln K_t + f(A_t). \tag{37}$$

From (8), optimal consumption is a constant fraction of wealth, $C_t = \mathbb{C}_1^{-1} K_t$. Now use the maximized Bellman equation (35) and insert the candidate solution,

$$\rho \mathbb{C}_1 \ln K_t + g(A_t) = \ln K_t - \ln \mathbb{C}_1 + ((A_t - \delta)K_t - \mathbb{C}_1^{-1} K_t)\mathbb{C}_1/K_t,$$

in which $g(A_t) \equiv \rho f(A_t) - \frac{1}{2}\left(f_{AA}\eta(A_t)^2 - \sigma^2\right) - f_A \mu(A_t)$. Thus, we obtain the condition $\mathbb{C}_1 = 1/\rho$ and collect the remaining terms in $g(A_t) = \ln \rho + A_t - \delta - \rho$. In the Vasicek case, $\eta(A_t) = \eta$ and $\mu(A_t) = \kappa(\gamma - A_t)$, we get $f(A_t) = \mathbb{C}_2 A_t + \mathbb{C}_3$, in which $\mathbb{C}_2 = \mathbb{C}_1/(\rho + \kappa)$ and $\mathbb{C}_3 = (\kappa\gamma\mathbb{C}_2 - \ln \mathbb{C}_1 - 1 - (\delta + \frac{1}{2}\sigma^2)\mathbb{C}_1)/\rho$.

A.2 AK-Vasicek model (CRRA preferences)

A.2.1 Numerical solution

[insert Figure A.1]

This section illustrates one particular numerical solution to obtain reasonable values for the consumption elasticities with respect to changes in wealth and the interest rate. We parameterize our system using $\kappa = 0.2$, $\gamma = 0.1$, $\eta = 0.005$, $\rho = 0.05$, $\delta = 0.05$, $\sigma = 0.05$, and $\theta = 2$. Our solution implies elasticities of $\overline{\tau} \approx 0.207$ and $\bar{\pi} \approx 1.021$. These results are obtained using the collocation method on a 10×7 Chebychev polynomial basis at standard Chebychev nodes (for an introduction see Miranda and Fackler, 2002). Our results confirm that the time-variability of these elasticities is small (see Figure A.1).

A.2.2 Equilibrium dynamics for general CRRA preferences

Here, we show the generalization of our estimation approach to the case of general CRRA preferences. Using the equilibrium dynamics system of differential equations (12), we obtain

$$
\begin{aligned}
\ln(C_s/C_t) &= 1/\theta \int_t^s r_v dv + \tfrac{1}{2}\theta(\bar{\tau}\eta)^2 \int_t^s 1/r_v^2 dv - ((\rho+\delta)/\theta - \tfrac{1}{2}(\theta\bar{\pi}-2)\bar{\pi}\sigma^2)(s-t) \\
&\quad + \int_t^s \eta\bar{\tau}/r_v dB_v + \sigma\bar{\pi}(Z_s - Z_t), \qquad (38a) \\
\ln(Y_s/Y_t) &= \kappa\gamma \int_t^s 1/r_v dv - \tfrac{1}{2}\eta^2 \int_t^s 1/r_v^2 dv + \int_t^s (1 - C_v/Y_v) r_v dv \\
&\quad -(\kappa + \delta + \tfrac{1}{2}\sigma^2)(s-t) + \int_t^s \eta/r_v dB_v + \sigma(Z_s - Z_t), \qquad (38b) \\
r_s &= e^{-\kappa(s-t)} r_t + (1 - e^{-\kappa(s-t)})\gamma + \eta e^{-\kappa(s-t)} \int_t^s e^{\kappa(v-t)} dB_v. \qquad (38c)
\end{aligned}
$$

Comparing to the case of logarithmic utility in (13a)-(13c), only the first two equations are different. Since consumption and output data are not available at the same frequency as the spot rate, we need to make an approximation. For the integral involving output, consumption and the short rate we use an Euler approximation scheme $\int_t^s g(v)dv = g(t)(s-t)$. Alternatively, one could employ the approximation $\int_t^s g(v)dv = g(s)(s-t)$ or the trapezoidal rule (employs averages) $\int_t^s g(v)dv \approx \frac{1}{2}(s-t)\left[g(s) + g(t)\right]$.

A.2.3 A regression-based approach

We collect the left-hand side variables in the vector $y_t = (y_{C,t}, y_{Y,t}, y_{r,t})^\top$, where $y_{C,t} = \ln(C_t/C_{t-\Delta})$, $y_{Y,t} = \ln(Y_t/Y_{t-\Delta}) - \int_{t-\Delta}^t (1 - C_v/Y_v) r_v dv$, and $y_{r,t} = r_t$. The parameters are $\beta = \left(\beta_C^\top, \beta_Y^\top, \beta_r^\top\right)^\top$, where $\beta_C = (\beta_{C,1}, \beta_{C,2}, \beta_{C,3})^\top$, $\beta_Y = (\beta_{Y,1}, \beta_{Y,2}, \beta_{Y,3})^\top$, $\beta_r = (\beta_{r,1}, \beta_{r,2})^\top$,

$$
\begin{aligned}
\beta_{C,1} &= -\left((\rho+\delta)/\theta - \tfrac{1}{2}(\theta\bar{\pi}-2)\bar{\pi}\sigma^2\right)\Delta, \qquad (39a) \\
\beta_{C,2} &= 1/\theta, \qquad (39b) \\
\beta_{C,3} &= \tfrac{1}{2}\theta(\bar{\tau}\eta)^2, \qquad (39c) \\
\beta_{Y,1} &= -\left(\kappa + \delta + \tfrac{1}{2}\sigma^2\right)\Delta, \qquad (39d) \\
\beta_{Y,2} &= \kappa\gamma, \qquad (39e) \\
\beta_{Y,3} &= -\tfrac{1}{2}\eta^2, \qquad (39f) \\
\beta_{r,1} &= (1 - e^{-\kappa\Delta})\gamma, \qquad (39g) \\
\beta_{r,2} &= e^{-\kappa\Delta}. \qquad (39h)
\end{aligned}
$$

In particular, the system (38) is linear in the right-hand side variables in $x_t = (x_{C,t}, x_{Y,t}, x_{r,t})$, with $x_{C,t} = (1, \int_{t-\Delta}^t r_v dv, \int_{t-\Delta}^t 1/r_v^2 dv)$, $x_{Y,t} = (1, \int_{t-\Delta}^t 1/r_v dv, \int_{t-\Delta}^t 1/r_v^2 dv)$, $x_{r,t} = (1, r_{t-\Delta})$.

Hence, the system (38) can be written in the form of simple regression models (14), where

$$\varepsilon_{C,t} = \int_{t-\Delta}^{t} \eta\bar{\tau}/r_v dB_v + \sigma\bar{\pi}(Z_t - Z_{t-\Delta}), \tag{40a}$$

$$\varepsilon_{Y,t} = \int_{t-\Delta}^{t} \eta/r_v dB_v + \sigma(Z_t - Z_{t-\Delta}), \tag{40b}$$

$$\varepsilon_{r,t} = \eta e^{-\kappa\Delta} \int_{t-\Delta}^{t} e^{\kappa(v-(t-\Delta))} dB_v. \tag{40c}$$

A.2.4 A martingale estimating function with three moment restrictions

Let $m_t = \varepsilon_t = (\varepsilon_{C,t}, \varepsilon_{Y,t}, \varepsilon_{r,t})^\top$ be the 3-vector of error terms expressed in terms of data and parameters. From system (40), m_t is clearly a martingale difference series, so we use

$$m_t = \begin{pmatrix} \ln(C_t/C_{t-\Delta}) - 1/\theta \int_{t-\Delta}^{t} r_v dv - \frac{1}{2}\theta(\bar{\tau}\eta)^2 \int_{t-\Delta}^{t} 1/r_v^2 dv + ((\rho+\delta)/\theta - \frac{1}{2}(\theta\bar{\pi} - 2)\bar{\pi}\sigma^2)\Delta \\ \ln(Y_t/Y_{t-\Delta}) - \int_{t-\Delta}^{t} (1 - C_v/Y_v) r_v dv - \kappa\gamma \int_{t-\Delta}^{t} 1/r_v dv + \frac{1}{2}\eta^2 \int_{t-\Delta}^{t} 1/r_v^2 dv + (\kappa + \delta + \frac{1}{2}\sigma^2)\Delta \\ r_t - (1 - e^{-\kappa\Delta})\gamma - e^{-\kappa\Delta} r_{t-\Delta} \end{pmatrix}, \tag{41}$$

where the integrals are approximated by summation over days between $t-\Delta$ and t.

Starting with the conditional variance, Ψ_t, it is useful to recast m_t as

$$m_t = \begin{pmatrix} \int_{t-\Delta}^{t} \eta\bar{\tau}/r_v dB_v + \sigma\bar{\pi}(Z_t - Z_{t-\Delta}) \\ \int_{t-\Delta}^{t} \eta/r_v dB_v + \sigma(Z_t - Z_{t-\Delta}) \\ \eta e^{-\kappa\Delta} \int_{t-\Delta}^{t} e^{\kappa(v-(t-\Delta))} dB_v \end{pmatrix}. \tag{42}$$

Hence, we have the conditional variances $\Psi_{t,11} = (\eta\bar{\tau})^2 E_{t-\Delta}(\int_{t-\Delta}^{t} 1/r_v^2 dv) + (\sigma\bar{\pi})^2\Delta$, $\Psi_{t,22} = \eta^2 E_{t-\Delta}(\int_{t-\Delta}^{t} 1/r_v^2 dv) + \sigma^2\Delta$, and $\Psi_{t,33} = \eta^2(1 - e^{-2\kappa\Delta})/(2\kappa)$. The conditional covariances are $\Psi_{t,12} = \eta^2\bar{\tau}E_{t-\Delta}(\int_{t-\Delta}^{t} 1/r_v^2 dv) + \sigma^2\bar{\pi}\Delta$, $\Psi_{t,13} = \eta^2\bar{\tau}e^{-\kappa\Delta}E_{t-\Delta}\left((\int_{t-\Delta}^{t} 1/r_v dB_v)(\int_{t-\Delta}^{t} e^{\kappa(v-(t-\Delta))} dB_v)\right)$, and $\Psi_{t,23} = \eta^2 e^{-\kappa\Delta}E_{t-\Delta}\left((\int_{t-\Delta}^{t} 1/r_v dB_v)(\int_{t-\Delta}^{t} e^{\kappa(v-(t-\Delta))} dB_v)\right)$. Since analytical expressions are not available, we use Euler approximations for $\Psi_{t,11}$, $\Psi_{t,22}$, $\Psi_{t,12}$, $\Psi_{t,13}$ and $\Psi_{t,23}$,

$$\Psi_t = \begin{pmatrix} (\eta\bar{\tau})^2\Delta/r_{t-\Delta}^2 + (\sigma\bar{\pi})^2\Delta & \eta^2\bar{\tau}\Delta/r_{t-\Delta}^2 + \sigma^2\bar{\pi}\Delta & \eta^2\bar{\tau}e^{-\kappa\Delta}\Delta/r_{t-\Delta} \\ \eta^2\bar{\tau}\Delta/r_{t-\Delta}^2 + \sigma^2\bar{\pi}\Delta & \sigma^2\Delta + \eta^2\Delta/r_{t-\Delta}^2 & \eta^2 e^{-\kappa\Delta}\Delta/r_{t-\Delta} \\ \eta^2\bar{\tau}e^{-\kappa\Delta}\Delta/r_{t-\Delta} & \eta^2 e^{-\kappa\Delta}\Delta/r_{t-\Delta} & \frac{1}{2}\eta^2(1 - e^{-2\kappa\Delta})/\kappa \end{pmatrix}. \tag{43}$$

Using martingale increments (41), we get the transpose of the conditional mean of the derivatives $\psi_t^\top$ with respect to the parameter vector $\phi = (\kappa, \gamma, \eta, \rho, \delta, \sigma, \bar{\pi}, \bar{\tau}, \theta)^\top$ as

$$\begin{pmatrix} 0 & \Delta - \gamma E_{t-\Delta} \int_{t-\Delta}^t 1/r_v dv & -\Delta e^{-\kappa\Delta}\gamma + \Delta e^{-\kappa\Delta} r_{t-\Delta} \\ 0 & -\kappa E_{t-\Delta} \int_{t-\Delta}^t 1/r_v dv & -(1-e^{-\kappa\Delta}) \\ -\theta\bar{\tau}^2\eta E_{t-\Delta} \int_{t-\Delta}^t 1/r_v^2 dv & \eta E_{t-\Delta} \int_{t-\Delta}^t 1/r_v^2 dv & 0 \\ \Delta/\theta & 0 & 0 \\ \Delta/\theta & \Delta & 0 \\ -(\theta\bar{\pi}-2)\bar{\pi}\sigma\Delta & \sigma\Delta & 0 \\ -(\theta\bar{\pi}-1)\sigma^2\Delta & 0 & 0 \\ -\theta\bar{\tau}\eta^2 E_{t-\Delta} \int_{t-\Delta}^t 1/r_v^2 dv & 0 & 0 \\ \begin{matrix} E_{t-\Delta}\int_{t-\Delta}^t 1/\theta^2 r_v - \frac{1}{2}(\bar{\tau}\eta)^2 1/r_v^2 dv \\ -\big((\rho+\delta)/\theta^2 + \frac{1}{2}(\bar{\pi}\sigma)^2\big)\Delta \end{matrix} & 0 & 0 \end{pmatrix}.$$

For the conditional expectation we interchange the order of integration, and then use the deterministic Taylor expansion. Using a first-order Taylor expansion for $\psi_{t,21}$, $\psi_{t,22}$, $\psi_{t,23}$, $\psi_{t,13}$, $\psi_{t,18}$ and $\psi_{t,19}$ allows approximating the integrals in ψ_t by

$$\begin{aligned} E_{t-\Delta}\int_{t-\Delta}^t 1/r_v dv &\approx \Delta/r_{t-\Delta} - \big(\kappa(\gamma - r_{t-\Delta})/r_{t-\Delta}^2 - \eta^2/r_{t-\Delta}^3\big)\tfrac{1}{2}\Delta^2, \\ E_{t-\Delta}\int_{t-\Delta}^t 1/r_v^2 dv &\approx \Delta/r_{t-\Delta}^2 - \big(2\kappa(\gamma - r_{t-\Delta})/r_{t-\Delta}^3 - 3\eta^2/r_{t-\Delta}^4\big)\tfrac{1}{2}\Delta^2, \\ E_{t-\Delta}\int_{t-\Delta}^t r_v dv \quad &\approx \Delta r_{t-\Delta} + \kappa(\gamma - r_{t-\Delta})\tfrac{1}{2}\Delta^2. \end{aligned}$$

For the last approximation, we may instead use the exact solution,

$$\begin{aligned} \int_{t-\Delta}^t E_{t-\Delta}(r_v) dv &= \int_{t-\Delta}^t \big(e^{-\kappa(v-(t-\Delta))} r_{t-\Delta} + (1 - e^{-\kappa(v-(t-\Delta))})\gamma\big)\, dv \\ &= \gamma\Delta + (r_{t-\Delta} - \gamma)(1 - e^{-\kappa\Delta})/\kappa. \end{aligned}$$

This completes the construction of the estimating equation. All nine structural parameters are separately identified.

A.3 Comparison to the discrete-time model

To introduce the reader to the potential advantages of the continuous-time formulation, we shall examine the equivalent formulation (and solution) of our model in a discrete-time environment. A straight-forward way of a discrete-time formulation is to consider the Euler approximation of our model (see e.g., Kloeden and Platen, 1999).

A.3.1 The model

Production possibilities. For the ease of readability, we present the full model below. The production function is a constant returns to scale technology

$$Y_t = A_t F(K_t, L), \tag{44}$$

where K_t is the (predetermined) aggregate capital stock, L is the constant population size, and A_t is total factor productivity, which follows an autoregressive process

$$A_{t+1} - A_t = \mu(A_t) + \eta(A_t)\epsilon_{A,t+1}, \quad \epsilon_A \sim N(0,1), \tag{45}$$

with $\mu(A_t)$ and $\eta(A_t)$ generic drift and volatility functions.[16] The capital stock increases if gross investment I_t exceeds capital depreciation,

$$K_{t+1} - K_t = I_t - \delta K_t + \sigma\epsilon_{K,t+1}, \quad \epsilon_K \sim N(0,1), \tag{46}$$

where δ is a deterministic rate of depreciation and σ determines the variance of a shock to the depreciation rate. Note that the stochastic depreciation does *not* depend on the level of the predetermined capital stock. This modification is necessary to compute the discrete-time Euler equation independent from the costate variables (see below).[17]

Equilibrium properties. In equilibrium, factors of production are rewarded with marginal products $r_t = Y_K$ and $w_t = Y_L$, subscripts K and L indicating derivatives, and the goods market clears, $Y_t = C_t + I_t$. Although there is no stochastic calculus for discrete-time models, we may express the evolution of equilibrium output in this economy as

$$Y_{t+1} = (A_t + \mu(A_t) + \eta(A_t)\epsilon_{A,t+1})\, F(K_t + I_t - \delta K_t + \sigma\epsilon_{K,t+1}, L). \tag{47}$$

Alternatively, we may use an Euler scheme to approximate the next period's output for small time intervals (no approximation error in the limit) by

$$Y_{t+1} - Y_t = \mu(A_t)Y_A + (I_t - \delta K_t)Y_K + \tfrac{1}{2}Y_{KK}\sigma^2 + Y_A\eta(A_t)\varepsilon_{A,t+1} + \sigma Y_K\varepsilon_{K,t+1}. \tag{48}$$

Obviously, comparing both (47) and (48) it seems much easier to get a dynamic formulation of the model which can be used for estimation with the help of stochastic calculus.

Preferences. Consider an economy with a single consumer, interpreted as a representative "stand in" for a large number of identical consumers. The consumer maximizes expected additively separable discounted life-time utility given by

$$U_0 \equiv E_0 \sum_{t=0}^{\infty} \beta^t u(C_t, A_t)dt, \quad u_C > 0,\ u_{CC} < 0, \tag{49}$$

[16] We assume that $E(A_t) = A \in \mathbb{R}_+$ exists, and that the sum describing life-time utility in (49) below is bounded, so that the value function is well-defined.

[17] It is insightful to relate the two shocks in the system to the continuous-time counterpart by looking at the Euler approximation $\epsilon_{A,t+1} \equiv B_{t+1} - B_t \sim N(0,1)$ and $\epsilon_{K,t+1} \equiv Z_{t+1} - Z_t \sim N(0,1)$.

subject to

$$K_{t+1} - K_t = (r_t - \delta)K_t + w_t L - C_t + \sigma\epsilon_{K,t+1}, \tag{50}$$

where β is the subjective discount factor, r_t is the rental rate of capital, and w_t is the labor wage rate. The paths of factor rewards are taken as given by the representative consumer.

A.3.2 The Euler equation

The relevant state variables are capital and technology, (K_t, A_t). For given initial states, the value of the optimal program is

$$V(K_0, A_0) = \max_{\{C_t\}_{t=0}^{\infty}} U_0 \quad s.t. \quad (50) \quad \text{and} \quad (45), \tag{51}$$

i.e., the present value of expected utility along the optimal program. As a necessary condition for optimality, Bellman's principle gives at time s

$$V(K_s, A_s) = \max_{C_s} \{u(C_s, A_s) + \beta E_s [V(K_{s+1}, A_{s+1})]\}. \tag{52}$$

Hence, the first-order condition for the problem is

$$u_C(C_t, A_t) = \beta E_t [V_K(K_{t+1}, A_{t+1})], \tag{53}$$

for any $t \in [0, \infty)$, and this allows us to write consumption as a function of the state variables, $C_t = C(K_t, A_t)$. Obviously, comparing the condition (8) to (53), the discrete-time model requires evaluating an integral (integrating out expectations) to obtain the optimal consumption function. The reason is that the Hamilton-Jacobi-Bellman (HJB) equation in the discrete-time model (52) requires to solve a stochastic difference equation in contrast to a deterministic differential equation (which can be useful in finding the numerical solution).

Using the concentrated Bellman equation,

$$V(K_t, A_t) = u(C(K_t, A_t)) + \beta E_t V(K_{t+1}, A_{t+1})$$

we may replace the unknown costate variable by known functions to get the Euler equation. Differentiating with respect to capital (using the envelope theorem) gives[18]

$$\begin{aligned} V_K(K_t, A_t) &= \beta E_t [V_K(K_{t+1}, A_{t+1})(1 - \delta + r_t)] \\ &= (1 - \delta + r_t) u_C(C_t, A_t). \end{aligned}$$

[18] If the stochastic depreciation depends on the predetermined capital stock in (50), it is not possible to fully replace the costate variable by known functions. The corresponding expression would be

$$\begin{aligned} V_K(K_t, A_t) &= \beta E_t [V_K(K_{t+1}, A_{t+1})(1 - \delta + r_t)] + \beta E_t [V_K(K_{t+1}, A_{t+1})\sigma\epsilon_{K,t+1}] \\ &= (1 - \delta + r_t) u_C(C_t, A_t) + \beta E_t [V_K(K_{t+1}, A_{t+1})\sigma\epsilon_{K,t+1}]. \end{aligned}$$

Leading the expression one period ahead and applying expectations yields

$$E_t\left[V_K(K_{t+1}, A_{t+1})\right] = E_t\left[(1-\delta+r_{t+1})u_C(C_{t+1}, A_{t+1})\right].$$

Inserting back into the first-order condition (53) we arrive at the Euler equation

$$u_C(C_t, A_t) = \beta E_t\left[(1-\delta+r_{t+1})u_C(C_{t+1}, A_{t+1})\right], \tag{54}$$

Alternatively, we may use an Euler scheme to approximate the next period's marginal utility for small time intervals (no approximation error in the limit) by

$$\begin{aligned} u_C(C_{t+1}, A_{t+1}) &= (1+\rho-(r_t-\delta))u_C(C_t, A_t) + u_{CC}(C_t, A_t)C_K\sigma\epsilon_{K,t+1} \\ &\quad +(u_{CC}(C_t, A_t)C_A\eta(A_t) + u_{CA}(C_t, A_t)\eta(A_t))\epsilon_{A,t+1}, \end{aligned} \tag{55}$$

Again, the continuous-time formulation may help to obtain a dynamic formulation which can be used for estimation of the structural parameters. For example, (55) could be used to put structure on the residuals in a regression-based estimation approach.

In the following, we restrict attention to the case $u(C_t, A_t) = u(C_t)$.

A.3.3 The equilibrium dynamics

Our equilibrium dynamics of the economy can be summarized as

$$u'(C_t) = \beta E_t\left[(1-\delta+r_{t+1})u'(C_{t+1})\right] \tag{56a}$$

$$Y_{t+1} = (A_t + \mu(A_t) + \eta(A_t)\epsilon_{A,t+1})\, F(K_t + I_t - \delta K_t + \sigma\epsilon_{K,t+1}, L) \tag{56b}$$

$$K_{t+1} = (1 + r_t - \delta)K_t + w_t L - C_t + \sigma\epsilon_{K,t+1} \tag{56c}$$

$$A_{t+1} = A_t + \mu(A_t) + \eta(A_t)\epsilon_{A,t+1} \tag{56d}$$

Provided that variables C_t, Y_t, K_t and also A_t are observed, the econometrician needs to consider the system (56) for statistical inference on the deep parameters.

For comparison, the equilibrium dynamics the corresponding continuous-time economy analogous to the model used in the main text can be summarized as

$$\begin{aligned} dC_t &= \frac{u'(C_t)}{u''(C_t)}(\rho-(r_t-\delta))dt - \tfrac{1}{2}(C_A^2\eta(A_t)^2 + C_K^2\sigma^2)\frac{u'''(C_t)}{u''(C_t)}dt \\ &\quad +C_A\eta(A_t)dB_t + C_K\sigma dZ_t \end{aligned} \tag{57a}$$

$$dY_t = (\mu(A_t)Y_A + (I_t - \delta K_t)Y_K + \tfrac{1}{2}Y_{KK}\sigma^2)dt + Y_A\eta(A_t)dB_t + \sigma Y_K dZ_t \tag{57b}$$

$$dK_t = ((r_t-\delta)K_t + w_t L - C_t)dt + \sigma dZ_t \tag{57c}$$

$$dA_t = \mu(A_t)dt + \eta(A_t)dB_t \tag{57d}$$

Provided that C_t, Y_t, K_t and also A_t are observed, the econometrician needs to consider the system (57) for statistical inference on the deep parameters.

In what follows, we assume that the capital stock K_t is a latent variable, but we can obtain the real interest rate from financial market data

A.3.4 An illustration: the stochastic AK model

Consider an AK economy, $Y_t = A_t K_t$, which implies $r_t = A_t$ and $K_t = Y_t/r_t$, and assume that the consumer has CRRA preferences with risk aversion θ, system (56) reduces to,

$$\begin{aligned} C_t^{-\theta} &= \beta E_t \left[(1-\delta+r_{t+1})C_{t+1}^{-\theta}\right] && \text{(58a)} \\ Y_{t+1} &= (1+\mu(r_t)/r_t + \eta(r_t)/r_t\epsilon_{A,t+1})\,(Y_t + (r_t-\delta)Y_t - r_tC_t + r_t\sigma\epsilon_{K,t+1}) \\ &= Y_t + Y_t\mu(r_t)/r_t + (r_t-\delta)Y_t - r_tC_t + Y_t\eta(r_t)/r_t\epsilon_{A,t+1} + r_t\sigma\epsilon_{K,t+1} \\ &\quad +((r_t-\delta)Y_t - r_tC_t + r_t\sigma\epsilon_{K,t+1})\,(\mu(r_t)/r_t + \eta(r_t)/r_t\epsilon_{A,t+1}) && \text{(58b)} \\ r_{t+1} &= r_t + \mu(r_t) + \eta(r_t)\epsilon_{A,t+1} && \text{(58c)} \end{aligned}$$

whereas system (57) reduces to

$$\begin{aligned} dC_t &= (r_t-\delta-\rho)C_t/\theta dt + \tfrac{1}{2}(1+\theta)(C_A^2\eta(A_t)^2 + C_K^2\sigma^2)/C_t dt \\ &\quad +C_A\eta(A_t)dB_t + C_K\sigma dZ_t && \text{(59a)} \\ dY_t &= (Y_t\mu(r_t)/r_t + (r-\delta)Y_t - r_tC_t)dt + Y_t\eta(r_t)/r_t dB_t + \sigma r_t dZ_t && \text{(59b)} \\ dr_t &= \mu(r_t)dt + \eta(r_t)dB_t && \text{(59c)} \end{aligned}$$

Both systems give the model in terms of observables (macro and financial market data).

One way of proceeding is to use an Euler scheme to discretize the system (59) for small time intervals (no approximation error in the limit). We do *not* follow this route because the continuous-time formulation naturally accounts for the different observation frequencies of macro and financial market data. Instead we proceed by integrating the system of equations and/or use closed-form solutions, for example for the interest rate.

A.3.5 Estimation

There is now a vast literature on estimating the deep parameters in discrete-time dynamic general equilibrium models. For example, An and Schorfheide (2007) use likelihood-based methods to estimate the structural parameters of system (58). One important caveat is that the nonlinear dynamics do not readily imply a likelihood function. To get around this problem, most of the previous literature has used the approximated likelihood derived from a linearized version of the model (Uhlig, 1999). Together with a state-space formulation, the system can now be estimated using the Kalman filter (see e.g., Fernández-Villaverde, Rubio-Ramírez, Sargent, and Watson, 2007). As an alternative approach the particle filter has been suggested to evaluate the likelihood function of the nonlinear model (Fernández-Villaverde and Rubio-Ramírez, 2007). This procedure combines a nonlinear solution method with a state-space representation to estimate the deep parameters.

We show in this paper that a continuous-time formulation of the model can facilitate estimating the deep parameters of the model. We make use of the fact that system (59) provides a much clearer description of the equilibrium dynamics. This paper now provides an alternative route which fully accounts for nonlinear structure of the equilibrium dynamics without leaving the regression-based and/or simulation-based estimation framework.

A.4 On problems of measurement with flow variables

The logarithmic specification in the main text was chosen in order to get the reduced-form parameters as a linear function of the observables. One caveat is that we do *not* observe the rate at which a flow variable, say C_t, changes over time, we only observe $\int_t^{t+1} C_s ds$.

Consider the Euler equation for consumption with CRRA preferences,

$$\begin{aligned} dC_t &= (r_t - \delta - \rho)C_t/\theta dt - \sigma^2 C_K K_t dt + \tfrac{1}{2}(1+\theta)(C_A^2 \eta(A_t)^2 + C_K^2 \sigma^2 K_t^2)/C_t dt \\ &\quad + C_A \eta(A_t) dB_t + C_K \sigma K_t dZ_t, \end{aligned}$$

Integrating the equation yields

$$\begin{aligned} C_{t+1} - C_t &= \int_t^{t+1} \left[(r_s - \delta - \rho)C_s/\theta - \sigma^2 C_K K_s + \tfrac{1}{2}(1+\theta)(C_A^2 \eta(A_s)^2 + C_K^2 \sigma^2 K_s^2)/C_s \right] ds \\ &\quad + \int_t^{t+1} C_A \eta(A_s) dB_s + \int_t^{t+1} C_K \sigma K_s dZ_s, \end{aligned}$$

It seems infeasible to consider changes in *observed* consumption (even further integration). Our approach therefore is to approximate the growth rates

$$\ln C_{t+1} - \ln C_t \;\approx\; \ln \int_t^{t+1} C_s ds - \ln \int_{t-1}^{t} C_s ds$$

References

Aït-Sahalia, Y. (1996): "Nonparametric Pricing of Interest rate Derivative Securities," *Econometrica*, 64(3), 527–560.

——— (2008): "Closed-form Likelihood Expansions for Multivariate Diffusions," *Annals of Statistics*, 36(2), 906–937.

An, S., and F. Schorfheide (2007): "Bayesian Analysis of DSGE models," *Econometric Rev.*, 26(2-4), 113–172.

Andrews, D. W. K., and W. Ploberger (1994): "Optimal Tests when a Nuisance Parameter is Present Only Under the Alternative," *Econometrica*, 62(6), 1383–1414.

Bergstrom, A. R. (1966): "Nonrecursive Models as Discrete Approximations to Systems of Stochastic Differential Equations," *Econometrica*, 34(1), 173–182.

Boldrin, M., L. J. Christiano, and J. D. M. Fisher (2001): "Habit Persistence, Asset Returns, and the Business Cycle," *Amer. Econ. Rev.*, 91(1), 149–166.

Brunnermeier, M. K., and Y. Sannikov (2011): "A Macroeconomic Model with a Financial Sector," *unpublished*.

Campanale, C., R. Castro, and G. L. Clementi (2010): "Asset Pricing in a Production Economy with Chew-Dekel Preferences," *Rev. Econ. Dynam.*, 13, 379–402.

Chang, F.-R. (1988): "The Inverse Optimal Problem: A Dynamic Programming Approach," *Econometrica*, 56(1), 147–172.

Chapman, D. A., J. B. Long, and N. D. Pearson (1999): "Using Proxies for the Short rate: When are Three Months Like an Instant?," *Rev. Finan. Stud.*, 12(4), 763–806.

Christensen, B. J., and M. Sørensen (2008): "Optimal Inference in Dynamic Models with Conditional Moment Restrictions," *CREATES Research Paper*, 51.

Cochrane, J. H. (2006): *Financial Markets and the Real Economy*, vol. 18 of *International Library of Critical Writings in Financial Economics*. Edward Elgar, London.

Cox, J. C., J. E. Ingersoll, and S. A. Ross (1985a): "An Intertemporal General Equilibrium Model of Asset Prices," *Econometrica*, 53(2), 363–384.

——— (1985b): "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53(2), 385–407.

De Jong, F. (2000): "Time Series and Cross-Section Information in Affine Term-Structure Models," *J. Bus. Econ. Statist.*, 18(3), 300–314.

Eaton, J. (1981): "Fiscal Policy, Inflation and the Accumulation of Risky Capital," *Rev. Econ. Stud.*, 48, 435–445.

Fernández-Villaverde, J., and J. F. Rubio-Ramírez (2007): "Estimating Macroeconomic Models: A Likelihood Approach," *Rev. Econ. Stud.*, 74, 1059–1087.

Fernández-Villaverde, J., J. F. Rubio-Ramírez, T. J. Sargent, and M. W. Watson (2007): "ABCs (and Ds) of Understanding VARs," *Amer. Econ. Rev.*, 97(3), 1021–1026.

Godambe, V. P., and C. C. Heyde (1987): "Quasi Likelihood and Optimal Estimation," *International Statistical Review*, 55, 231–244.

Grinols, E. L., and S. J. Turnovsky (1993): "Risk, the Financial Market, and Macroeconomic Equilibrium," *J. Econ. Dynam. Control*, 17, 1–36.

Hansen, B. E. (1996): "Inference When a Nuisance Parameter Is Not Identified Under the Null Hypothesis," *Econometrica*, 64(2), 413–430.

Hansen, L. P. (1982): "Large sample properties of generalized method of moments estimators," *Econometrica*, 50(4), 1029–1054.

Hansen, L. P., and T. J. Sargent (1991): *Rational Expectations Econometrics*. Westview Press, Boulder.

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

Harvey, A., and J. H. Stock (1989): "Estimating Integrated Higher-Order Continuous Time Autoregressions with an Application to Money-Income Causality," *J. Econometrics*, 42, 319–336.

Jermann, U. J. (1998): "Asset Pricing in Production Economies," *J. Monet. Econ.*, 41, 257–275.

Kloeden, P. E., and E. Platen (1999): *Numerical Solution of Stochastic Differential Equations*. Springer, 3rd ed.

Lettau, M. (2003): "Inspecting the Mechanism: Closed-Form Solutions for Asset Prices in Real Business Cycle Models," *Econ. J.*, 113, 550–575.

Lettau, M., and H. Uhlig (2000): "Can Habit Formation Be Reconciled with Business Cycle Facts?," *Rev. Econ. Dynam.*, 3, 79–99.

McCrorie, J. R. (2009): "Estimating Continuous-Time Models on the Basis of Discrete Data via an Exact Discrete Analog," *Econometric Theory*, 25, 1120–1137.

Miranda, M. J., and P. L. Fackler (2002): *Applied Computational Economics and Finance*. MIT Press, Cambridge.

Phillips, P. (1972): "The Structural Estimation of a Stochastic Differential Equation System," *Econometrica*, 40(6), 1021–1041.

——— (1991): "Error-Correction and Long-Run Equilibrium in Continuous-Time," *Econometrica*, 59(4), 967–980.

Posch, O. (2009): "Structural Estimation of Jump-Diffusion Processes in Macroeconomics," *J. Econometrics*, 153(2), 196–210.

Protter, P. E. (2004): *Stochastic Integration and Differential Equations.* Springer, Berlin.

Rudebusch, G. D., and E. T. Swanson (2008): "Examining the Bond Premium Puzzle with a DSGE Model," *J. Monet. Econ.*, 55, S111–S126.

Sennewald, K. (2007): "Controlled Stochastic Differential Equations under Poisson Uncertainty and with Unbounded Utility," *J. Econ. Dynam. Control*, 31(4), 1106–1131.

Tallarini, T. (2000): "Risk-Sensitive Real Business Cycles," *J. Monet. Econ.*, 45, 507–532.

Uhlig, H. (1999): *A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily.* Computational Methods for the Study of Dynamic Economies, Marimon and Scott (eds.).

van Binsbergen, J. H., J. Fernández-Villaverde, R. S. J. Koijen, and J. F. Rubio-Ramírez (2012): "The Term Structure of Interest Rates in a DSGE Model with Recursive Preferences," *J. Mon. Econ*, forthcoming.

Vasicek, O. (1977): "An Equilibrium Characterization of the Term Structure," *J. Finan. Econ.*, 5, 177–188.

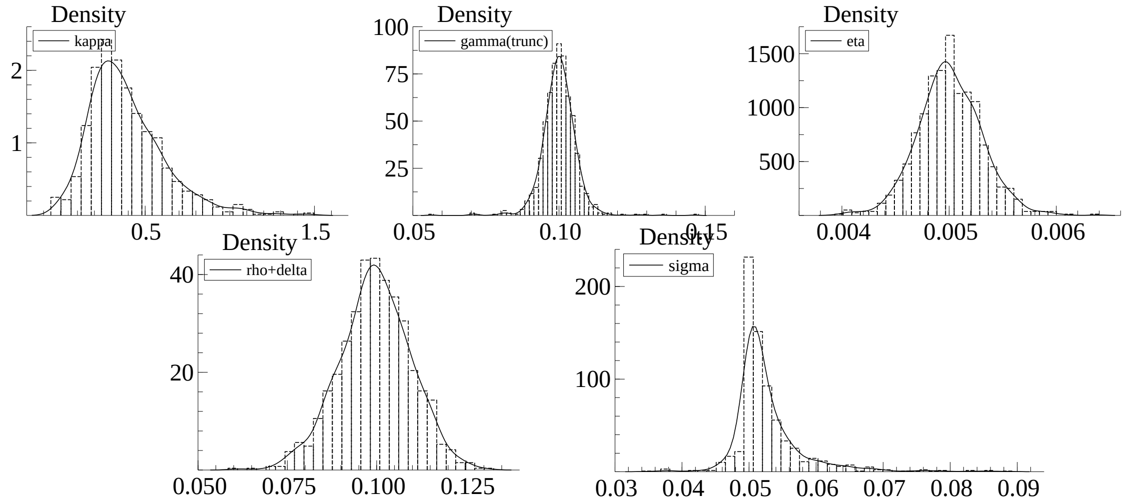
Table 1: Simulation Study – AK-Vasicek Model (Logarithmic Preferences)

The table reports output of a simulation study of the accuracy of the structural model parameters estimated using OLS, FGLS-SUR, FGLS-SUR-IV, and MEF approaches for the AK-Vasicek model with logarithmic preferences. For 1,000 replications, we generate 25 years of data from the underlying data generating process (DGP) and apply our estimation strategy. We show the median estimate, and provide the interquartile range below it.

Parameter Estimates from Simulation Study									
		Monthly Data				Quarterly Data			
	DGP	OLS	FGLS-SUR	FGLS-SUR-IV	MEF	OLS	FGLS-SUR	FGLS-SUR-IV	MEF
κ	0.2	0.227	0.162	0.220	0.351	0.227	0.173	0.180	0.235
		0.189	0.13	0.126	0.273	0.197	0.131	0.147	0.104
γ	0.1	0.100	0.100	0.100	0.100	0.100	0.0999	0.100	0.0999
		0.00703	0.00748	0.00664	0.00624	0.00722	0.00769	0.00733	0.00692
η	0.005	0.00865	0.00353	0.00408	0.00499	0.00842	0.00366	0.00366	0.00528
		0.00188	0.000521	0.000888	0.000383	0.0021	0.000772	0.00125	0.000383
$\rho+\delta+\frac{1}{2}\sigma^2$	0.101	0.100	0.101	0.101		0.100	0.101	0.101	
		0.0127	0.0128	0.013		0.0128	0.0129	0.0125	
$\rho+\delta$	0.1				0.0997				0.0995
					0.0129				0.0201
σ	0.05				0.0511				0.0573
					0.00352				0.0272

Figure 1: Simulation Study MEF Approach – AK-Vasicek Model (Logarithmic Preferences)
The figure reports output of a simulation study of the accuracy of the structural model parameters estimated using the MEF approach for the AK-Vasicek model with logarithmic preferences. For 1,000 replications, we generate 25 years of data from the underlying data generating process (DGP) and apply our estimation strategy. We plot the distribution of the obtained estimates, in Panel (A) for monthly data and in Panel (B) for quarterly data.

(A) Monthly Data



(B) Quarterly Data

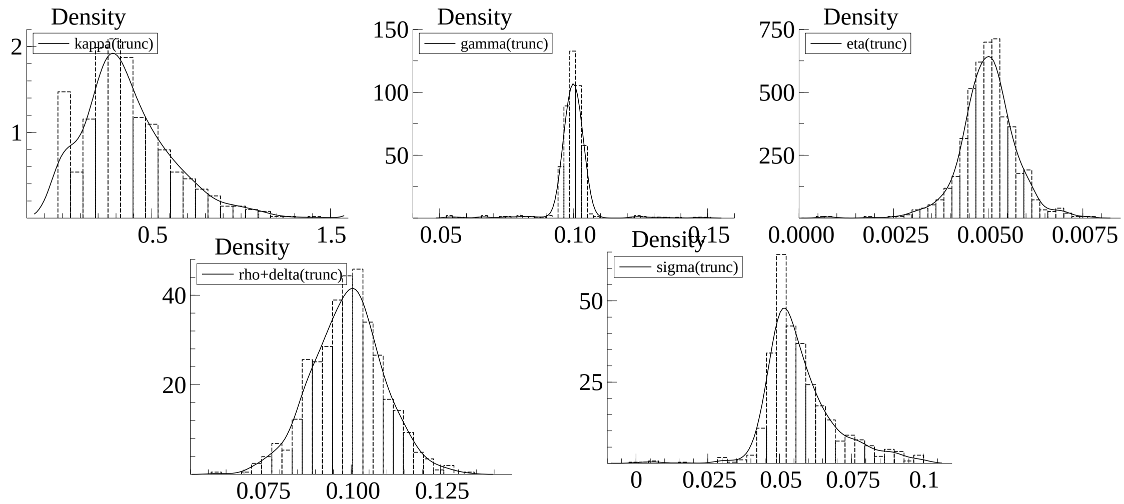


Table 2: Simulation Study – AK-Vasicek Model (CRRA Preferences)

The table reports output of a simulation study of the accuracy of the structural model parameters estimated using OLS, FGLS-SUR, FGLS-SUR-IV, and MEF approaches for the AK-Vasicek model with CRRA preferences. For 1,000 replications, we generate 25 years of data from the underlying data generating process (DGP) and apply our estimation strategy. We show the median estimate, and provide the interquartile range below it.

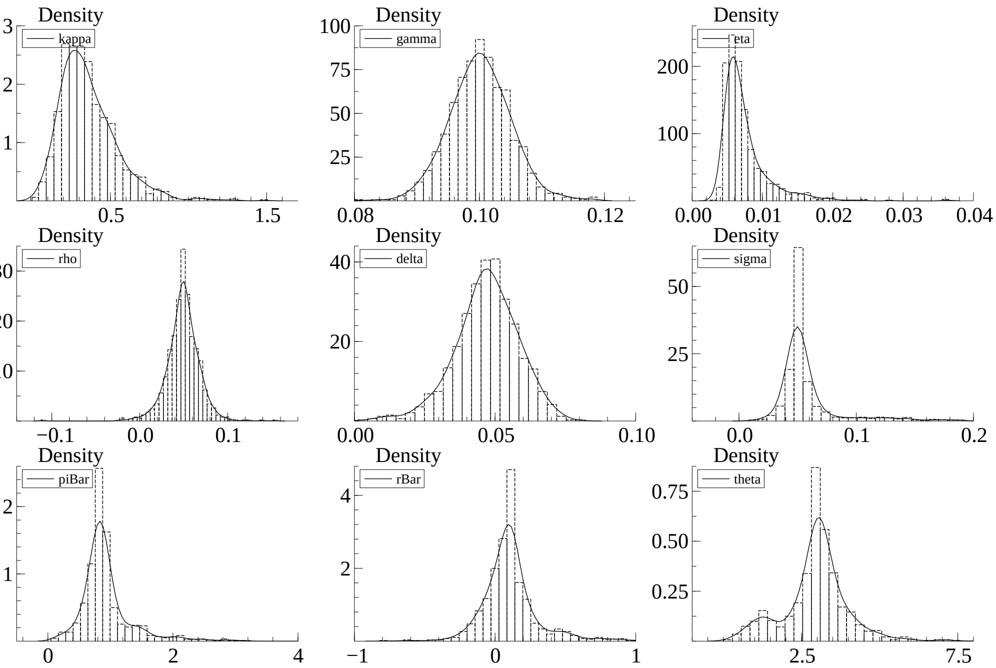
Parameter Estimates from Simulation Study									
		Monthly Data				Quarterly Data			
	DGP	OLS	FGLS-SUR	FGLS-SUR-IV	MEF	OLS	FGLS-SUR	FGLS-SUR-IV	MEF
κ	0.2	0.351	0.151	0.187	0.304	0.348	0.139	0.128	0.285
		0.27	0.108	0.0943	0.188	0.274	0.123	0.11	0.253
γ	0.1	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100
		0.00619	0.00756	0.00699	0.00631	0.00632	0.00811	0.00834	0.00631
η	0.005	0.0607	0.00421	0.00000170	0.00669	0.0597	0.00334	0.00000295	0.00695
		0.028	0.0301	0.0000163	0.0027	0.0306	0.0319	0.019	0.00669
ρ	0.05				0.0519				0.0515
					0.015				0.0226
δ	0.05				0.0475				0.0484
					0.0138				0.0186
σ	0.05				0.0518				0.0558
					0.0123				0.035
$\bar{\pi}$	1.02				1.00				1.01
					0.347				0.524
$\bar{\tau}$	0.207	0.00000229	3.96	0.745	0.161	0.00000201	2.94	0.377	0.116
		4.08	216	3.03	0.233	3.98	218	2.53	0.546
θ	2	0.301	0.630	1.73	2.06	0.308	0.659	1.33	2.29
		0.985	0.348	0.728	0.676	1.22	0.591	0.733	1.91
$\delta + \frac{1}{2}\sigma^2$	0.0513	-0.134	0.0400	0.0493		-0.125	0.0405	0.0469	
		0.18	0.047	0.017		0.181	0.0514	0.0257	
$\rho/\theta - A^{\#}$	0.0488	0.910	0.139	0.0282		0.900	0.133	0.0370	
		3.52	0.129	0.0147		3.39	0.211	0.0519	

$^{\#}$: For brevity we write $A = \frac{1}{2}((\theta\bar{\pi} - 2)\bar{\pi} + \frac{1}{\theta})\sigma^2$

Figure 2: Simulation Study MEF Approach – AK-Vasicek Model (CRRA Preferences)

The figure reports output of a simulation study of the accuracy of the structural model parameters estimated using the MEF approach for the AK-Vasicek model with CRRA preferences. For 1,000 replications, we generate 25 years of data from the underlying data generating process (DGP) and apply our estimation strategy. We plot the distribution of the obtained estimates, in Panel (A) for monthly data and in Panel (B) for quarterly data.

(A) Monthly Data



(B) Quarterly Data

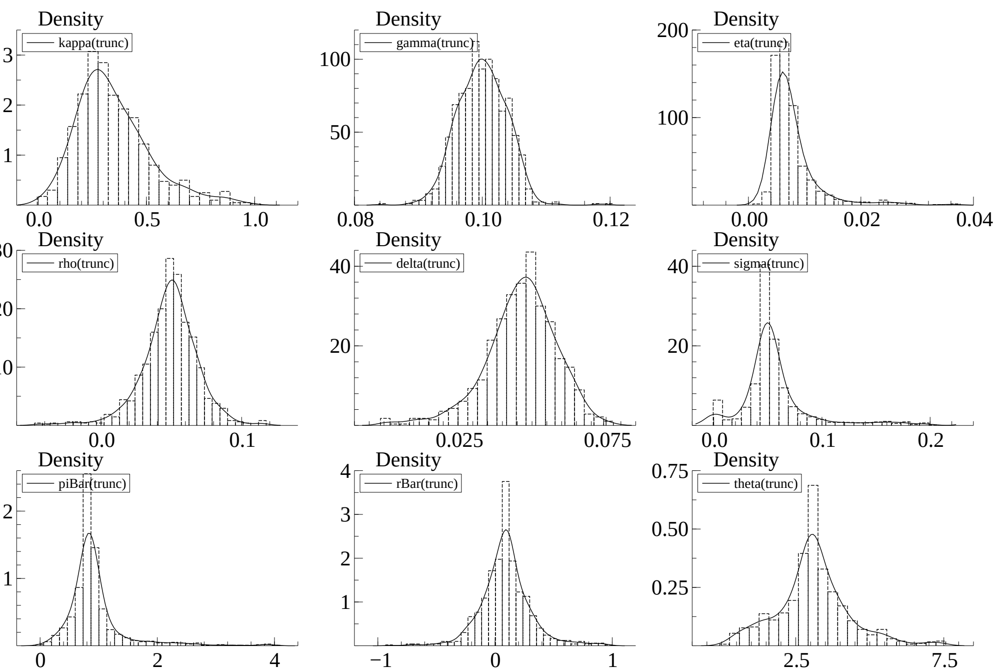
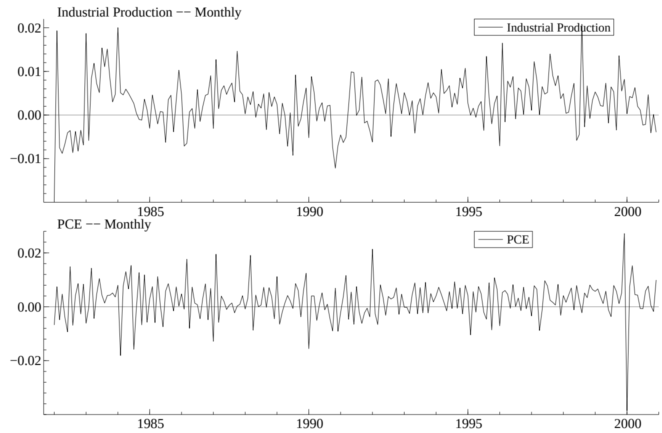


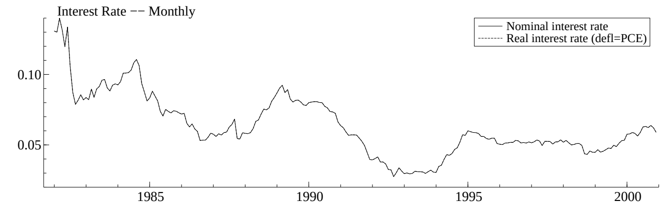
Figure 3: Overview of Monthly Variables

In this figure we show time series plots of the macroeconomic variables in our data set at the monthly frequency. In Panel (A) we show the growth rate of Industrial Production (IP) and Real Personal Consumption Expenditure (PCE), both from the Federal Reserve Bank of St. Louis Economic Dataset (FRED). In Panel (B) we show the nominal 3m interest rate series also obtained from the FRED (last day of month observation from the daily data set). Panel (C) shows the approximations to the three integrals from the structural model based on the daily nominal 3m interest rate series, and the integral based on consumption, income, and the interest rate. In all cases, the sample runs from January, 1982 until December, 2000.

(A) Monthly Macroeconomic Variables



(B) Monthly Interest Rate



(C) Monthly Integrals

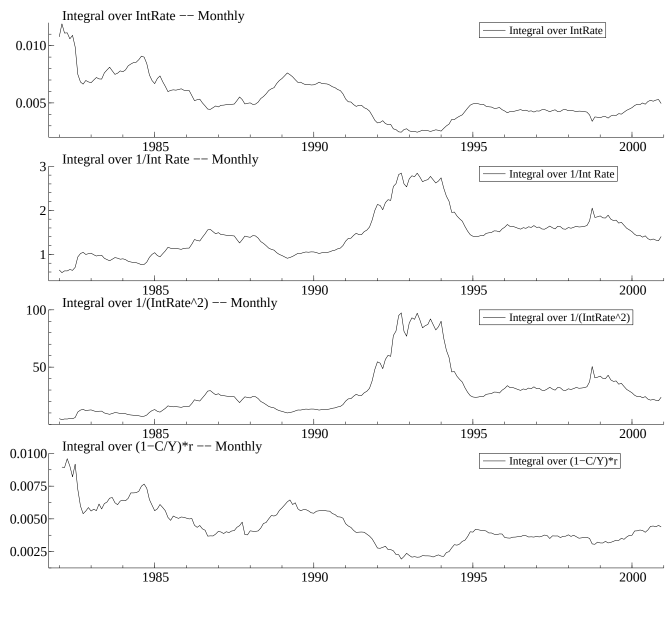
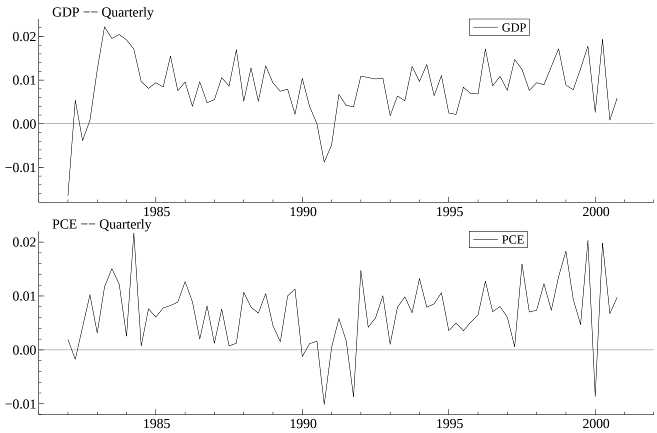


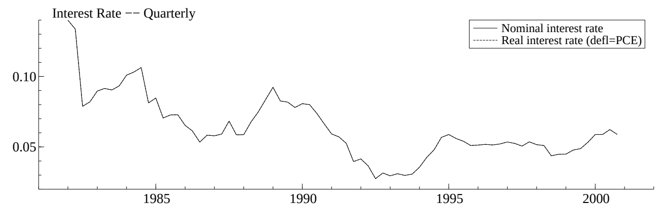
Figure 4: Overview of Quarterly Variables

In this figure we show time series plots of the macroeconomic variables in our data set at the quarterly frequency. In Panel (A) we show the growth rate of Real Gross Domestic Product (GDP) and Real Personal Consumption Expenditure (PCE), both from the Federal Reserve Bank of St. Louis Economic Dataset (FRED). In Panel (B) we show the nominal 3m interest rate series also obtained from the FRED (last day of quarter observation from the daily data set). Panel (C) shows the approximations to the three integrals from the structural model based on the daily nominal 3m interest rate series, and the integral based on consumption, income, and the interest rate. In all cases, the sample runs from 1982:Q1 until 2000:Q4.

(A) Quarterly Macroeconomic Variables



(B) Quarterly Interest Rate



(C) Quarterly Integrals

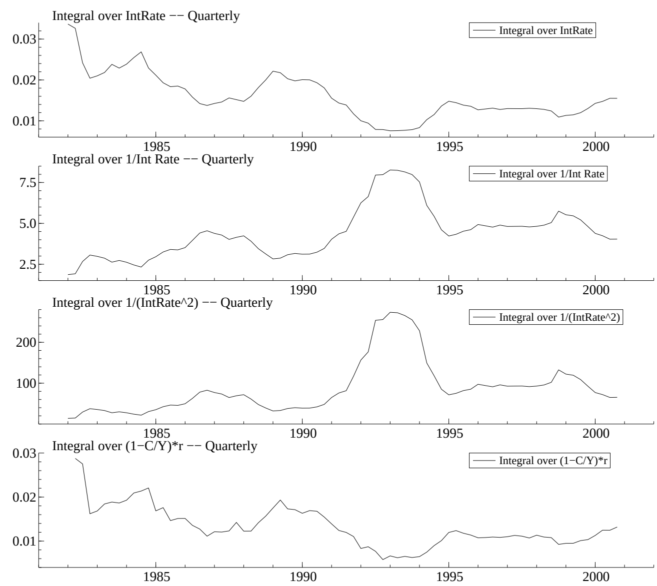


Table 3: Estimates – AK-Vasicek Model (Logarithmic Preferences)

The table reports estimates for the structural model parameters estimated using OLS, FGLS-SUR, FGLS-SUR-IV, and MEF approaches for the AK-Vasicek model with logarithmic preferences. We run the estimation for monthly data (where production is measured by IP) and quarterly data (production measured by GDP). The sample runs from 1982 until 2001. Asymptotic t-statistics are given below the estimates.

Parameter Estimates from Empirical Data								
	Monthly Data				Quarterly Data			
	OLS	FGLS-SUR	FGLS-SUR-IV	MEF	OLS	FGLS-SUR	FGLS-SUR-IV	MEF
κ	0.170	0.154	0.186	7.91	0.107	0.0591	0.0545	11.6
	0.699	0.357	0.346	19.3	0.176	0.195	0.399	4.03
γ	0.0857	0.0876	0.0888	0.0649	0.0876	0.0926	0.0930	0.0767
	1.04	0.000199	0.00000436	74.2	3.27	2.94	6.27	21.2
η	0.0221	0.0213	0.0241	0.0301	0.0173	0.0129	0.0125	0.154
	0.0224	2.59	0.000000781	2.59	0.981	0.0748	0.238	13.3
$\rho+\delta+\frac{1}{2}\sigma^2$	0.0334	0.0346	0.0345		0.0344	0.0348	0.0335	
	0.157	0.259	0.303		0.643	0.774	0.357	
$\rho+\delta$				0.0355				0.0351
				18				18.6
σ				0.00855				0.00817
				508				530

Table 4: Estimates – AK-Vasicek Model (CRRA Preferences)

The table reports estimates for the structural model parameters estimated using OLS, FGLS-SUR, FGLS-SUR-IV, and MEF approaches for the AK-Vasicek model with CRRA preferences. We run the estimation for monthly data (where production is measured by IP) and quarterly data (production measured by GDP). The sample runs from 1982 until 2001. Asymptotic t-statistics are given below the estimates.

Parameter Estimates from Empirical Data								
	Monthly Data				Quarterly Data			
	OLS	FGLS-SUR	FGLS-SUR-IV	MEF	OLS	FGLS-SUR	FGLS-SUR-IV	MEF
κ	0.336	0.335	0.388	0.398	0.269	0.246	0.247	0.562
	0.136	0.298	0.243	1.13	0.32	0.287	0.0503	1.33
γ	0.0494	0.0493	0.0516	0.0531	0.0360	0.0321	0.0309	0.0544
	1.95	0.000155	1.06	2.52	0.508	0.000361	0.000179	3.24
η	0.0248	0.0247	0.0275	0.0329	0.0184	0.0159	0.0156	0.0379
	0.695	0.145	0.0644	3.37	3.3	0.0000361	0.0897	5.76
ρ				0.109				0.236
				0.795				6.44
δ				-0.266				-0.272
				-2.06				-3.24
σ				0.372				0.0968
				8.28				1.75
$\bar{\pi}$				0.253				0.818
				3.47				102
$\bar{\tau}$	0.0224	0.00104	0.000459	0.00384	0.000576	0.000820	0.104	-0.00000285
	0.00559	0.000000287	0.000314	0.00132	0.00214	0.00000107	0.0000266	-0.00000840
θ	91.1	37669	17202	6.33	15735	10918	11.2	6.00
	0.0000157	1.69e+005	1.11e+003	0.487	6.35e+005	2.16e+005	0.00019	0.202
$\delta + \frac{1}{2}\sigma^2$	-0.132	-0.13	-0.146		-0.139	-0.131	-0.136	
	-0.0493	-0.161	-0.115		-0.164	-0.136	-0.0284	
$\rho/\theta - A^{\#}$	-0.0208	-0.0236	-0.0228		-0.0245	-0.0240	-0.00459	
	-0.00614	-0.00356	-0.00416		-0.0088	-0.00887	-0.000231	

$^{\#}$: For brevity we write $A = \frac{1}{2}((\theta\bar{\pi} - 2)\bar{\pi} + \frac{1}{\theta})\sigma^2$

Figure A.1: Numerical solution of the AK-Vasicek model (CRRA preferences)

In this figure we show (from left to right) optimal consumption as a function of wealth and the interest rate, the residuals of the Bellman equation, the consumption elasticity with respect to changes in wealth, $\bar{\pi}$, and the interest rate, $\bar{\tau}$, for a parametrization $(\kappa, \gamma, \eta, \rho, \delta, \sigma, \theta) = (0.2, 0.1, 0.005, 0.05, 0.05, 0.05, 2)$.