

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

Bent Jesper Christensen^(a,b), Olaf Posch^{(c,b)†}, and Michel van der Wel^(b,d)
^(a)Aarhus University, ^(b)CREATES, ^(c) Hamburg University,
^(d)Erasmus University Rotterdam

October 6, 2014

Abstract

We provide a framework for inference in dynamic equilibrium models including financial market data at daily frequency, along with macro series at standard lower frequency. Our formulation of the macro-finance model in continuous-time conveniently accounts for the difference in observation frequency. We suggest the use of martingale estimating functions (MEF) to infer the structural parameters of the model directly through a nonlinear optimization scheme. This method is compared to regression-based methods and the general method of moments (GMM). We illustrate our approaches by estimating the AK-Vasicek model with mean-reverting interest rates. We provide Monte Carlo evidence on the small sample behavior of the estimators and report empirical estimates using 30 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, Eric Ghysels, Andrew Harvey, Roderick McCrorie, Hashem Pesaran, Juan Rubio-Ramírez, Frank Schorfheide, Frank Smets, and seminar participants in Mannheim, Groningen, Vienna, Atlanta (Methods and Applications for DSGE Models), and at conferences in Chicago (ASSA, SoFiE), Santa Fe (WFA), Oslo (ESEM), Florence (Econometric Modelling of Macro-Financial Linkages), Göttingen (VfS), Vancouver (CEF), Nicosia (EC2), Aarhus (Recent Developments on the Links between Financial Markets and the Real Economy), Berlin (Humboldt-Copenhagen), Tilburg (NESG), and London (IAAE) for comments. The authors appreciate financial support from the Netherlands Organisation for Scientific Research (NWO) for a Veni grant, and 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 (Email address: olaf.posch@uni-hamburg.de, Address: Hamburg University, Department of Economics, Von-Melle-Park 5, 20146 Hamburg, Germany).

1 Introduction

Dynamic stochastic general equilibrium (DSGE) models have become the workhorse in macroeconomics, capturing aggregate dynamics over the business cycle. They are frequently used in academic research and to assess various policy interventions. There is a tradition in macroeconomics to model financial frictions, but until before the recent financial turmoils, with large effects on the real economy, these models were silent about shocks that originate in the financial sector (see Jermann and Quadrini, 2012).¹ Given the importance and relevance of these models in both theoretical and applied research, surprisingly little work has been published on reconciling business cycle facts with asset pricing implications (see the results in Jermann, 1998; Tallarini, 2000; Rudebusch and Swanson, 2008). The recent economic and credit crises intensified the desire to link macro and finance.² Conceptually, the financial models should benefit from specifying the stochastic discount factor consistent with macro-dynamics, whereas the macroeconomic models could benefit from rich financial data.³ Essential features of this strand of literature are the macro-finance interaction, the mixed frequency of macroeconomic and financial data, and latent variables. Yet, there is no clear consensus on how macro and financial data should be linked consistently, and how data are used efficiently in the estimation of macro-finance models.

Our aim is to develop a paradigm in which macroeconomics, finance and econometrics are coherently linked. This paper provides a framework for estimation of dynamic equilibrium models with both macro and financial variables, taking account of mixed-frequencies and latent variables. We believe that a structural estimation approach can shed light on the channels through which financial markets and the real economy interact. This is important for the design of monetary and fiscal policies and to evaluate certain policy measures. Specifically, our contribution is threefold. First, we propose using DSGE models in continuous time to facilitate incorporation of financial market variables in a structural manner. Second, we propose and develop using martingale estimating functions (MEF) for the estimation of such macro-finance models. Our continuous-time formulation and the MEF approach naturally

¹The literature on the financial sector propagating shocks emerged from the seminal contributions by Bernanke and Gertler (1989); Kiyotaki and Moore (1997); Bernanke, Gertler, and Gilchrist (1999).

²Recent developments illustrate that such unified framework is promising and intriguing: Gertler and Karadi (2011) present a DSGE model with financial frictions in which intermediaries face balance sheet constraints. Hence, unconventional monetary policy, by expanding central bank credit intermediation, may serve as a complement to financial intermediation. Brunnermeier and Sannikov (2014) develop a macroeconomic model with a financial sector and endogenous leverage, which leads to crisis episodes showing a mechanism how small shocks can have potentially large effects on the real economy.

³There are developments incorporating macro factors in financial models of the term structure (Ang and Piazzesi, 2003; Dewachter and Lyrio, 2006; Diebold, Rudebusch, and Aruoba, 2006; Hördahl, Tristani, and Vestin, 2006; Rudebusch and Wu, 2008) and incorporating financial factors in the estimation of macro models (Ang, Piazzesi, and Wei, 2006; van Binsbergen, Fernández-Villaverde, Koijen, and Rubio-Ramírez, 2012).

incorporate variables arriving at different frequencies by a model-based time-aggregation and thus eases statistical inference in the case of mixed-frequency data. Third, we extend the baseline MEF approach to further consider the case of mixed-frequency data estimation and additionally develop techniques for estimation of dynamic equilibrium models in case of latent variables by a simulation-based MEF.

We make the link between macro and financial markets explicit by showing how financial market data (say, interest rates or return data) can be used to identify the structural parameters, which characterize preferences and technology. To make the least stringent timing assumption we cast our DSGE model in continuous time, then solve for the general equilibrium of the real economy and asset prices, and develop three alternative estimation procedures. We consider off-the-shelf regression-based methods (combined with minimum-distance to identify structural parameters), the general method of moments (GMM), and the martingale estimating functions (MEF). Our continuous-time formulation is useful in three dimensions: (i) to place structure on the residuals in the regression-based methods, (ii) to obtain the general equilibrium dynamics in terms of data and parameters for the GMM and MEF approaches, and (iii) to account for the dependence among economic variables during the observation interval.

Our analysis is motivated by the fact that financial market data typically are available at higher frequency and of better quality than aggregate macro data (e.g., not subject to revision). In a unified framework, financial variables provide an additional source of evidence on the state of the economy, beyond macro series. So far only a few researchers have made use of this property in DSGE models. One apparent challenge is that discrete-time models are not time invariant (see Marcellino, 1999; Forni and Marcellino, 2014). Put differently, the parameter estimates can only properly be interpreted in the way we solve our models and in the context of the particular rate at which we sample the underlying process.⁴ But any fixed period length is arbitrary. Moreover, the frequency at which economic agents make their decisions may not necessarily coincide with the observation frequency of either macro or financial variables. Formulating structural models in continuous time offers a way forward. Our approach yields explicit functional forms of the relations among observables without taking a stand of the frequency at which economic agents make their decisions. Having at hand these functional forms, the availability of financial data at higher frequency (say, daily) than consumption and production (monthly or quarterly) then allows precise approximation of integrals by summation over days. Because the structural parameters

⁴The temporal aggregation literature started with Amemiya and Wu (1972); Geweke (1978). Others noted how the time-invariance problem relates to the behavior of agents (Christiano and Eichenbaum, 1987). Recent contributions include Kim (2010) and Giannone, Monti, and Reichlin (2014)

enter into the coefficients on these terms, financial data help us for identification. We also show how the MEF approach can be extended to cope with (more general) mixed-frequency data and latent variables.

We depart from the traditional discrete-time formulation of DSGE models and their estimation for three related technical reasons.⁵ First, there is no need to perform numerical integration to compute expectations, since the Hamilton-Jacobi-Bellman (HJB) equation is non-stochastic, thus simplifying computation of the first-order principles. Second, some solutions allow for an analytical likelihood function, simplifying the inference on structural parameters even in the presence of nonlinearities and non-normalities (see Posch, 2009). Third, many seminal models in finance are stated in continuous time (such as the equilibrium models of Vasicek, 1977; Cox, Ingersoll, and Ross, 1985b), which is particularly useful in the development of a unified framework in macroeconomics and finance.

Our work builds on a tradition in macroeconomics estimating continuous-time models, formulated as systems of (linear) stochastic differential equations.⁶ The traditional approach is to solve the system which implies a coefficient matrix that is a function of the exponential of a matrix, which in turn depends on the structural model parameters (cf. Phillips, 1972). As illustrated in McCrorie (2009), this complicates the identification due to the aliasing phenomenon: The distinct stochastic processes may look identical when sampled at discrete intervals (see Hansen and Scheinkman, 1995, p. 769).⁷ In this paper, we adopt an alternative approach of integrating the logarithmic (nonlinear) system to get an ‘exact’ discrete-time analog. The resulting system is in logarithmic growth rates rather than in levels. It involves a coefficient matrix linear in a set of known functions of the structural parameters, and does not involve any matrix exponential. An analysis whether our approach alleviates the aliasing problem in the linear model is interesting, but beyond the scope of this paper.

Our martingale estimating functions (MEF) approach benefits from the continuous time structure of the dynamic equilibrium model. The model provides martingale increments. The martingale estimating functions are defined as a weighted sum of these martingale increments. The optimal weights in MEF are time-varying and depend on the conditional

⁵An non-exhaustive list of references on the estimation of discrete-time DSGE models are Ruge-Murcia (2007); Fernández-Villaverde and Rubio-Ramírez (2007) and An and Schorfheide (2007). While the first two references show how to use standard econometric methods and the particle filter (instead of the linear Kalman filter) for estimation, the latter reviews Bayesian methods. In an accompanying web appendix we show in Appendix A that an Euler approximation could be used to apply this toolbox to continuous-time DSGE models. We do *not* follow this route because the continuous-time formulation naturally accounts for the different observation frequencies of macro and financial market data, which in this case would be lost.

⁶Seminal papers are contributions by Bergstrom (1966); Sims (1971); Phillips (1972, 1991), whereas the rational expectation models are due to Hansen and Sargent (1991); Hansen and Scheinkman (1995).

⁷Hence, one may argue that the use of continuous-time models is not a panacea. What we can gain in time invariance can be at the cost of needing to solve a more severe econometric identification problem.

variance of the martingale increment and the conditional mean of the parameter derivatives, see Christensen and Sørensen (2008). In the models we consider, these optimal weights can be analytically derived and depend on the structural model parameters. MEF exactly identifies all structural parameters of the system, through its set-up and by having both the weights and martingale increments depend on the structural parameters. We contrast this estimation approach with GMM, which can be shown to be less efficient than the MEF estimator. In addition we discuss a variety of regression-based estimators that through a two-step approach first provide estimates of a reduced form of our models and then are mapped to structural parameters through minimum-distance estimation.

We extend the MEF approach in two directions. First, we further develop the mixed-frequency nature of our set-up. The baseline MEF approach utilizes financial market data at a daily level to develop monthly and quarterly proxies that are used for estimation at monthly or quarterly frequency. In our Mixed-Frequency MEF (MF-MEF) approach we develop an approach that allows for variables to be observed at different frequencies. We stipulate the model at a high frequency, monthly in our applications, and use the other variables in the frequency at which they occur. For example, with monthly consumption series we consider the quarterly GDP series. We implement the approach by considering the actual quarterly output in the month that it becomes available and a model based prediction for the months in the quarter where output is not available. The monthly consumption and quarterly DGP are combined with the monthly proxy based on the daily financial variable. Using mixed-frequency data for estimation recently has received a lot of attention in the literature, see, e.g., Ghysels, Sinko, and Valkanov (2007), Andreou, Ghysels, and Kourtellos (2010), Andreou, Ghysels, and Kourtellos (2013) and Schorfheide and Song (2014). Our MF-MEF approach provides a structural approach for mixed-frequency estimation.

Second, we extend the MEF set-up to allow for latent variables. In the Simulated MEF (SMEF) approach we simulate the latent variable using the process implied by the model. The simulated paths are used to obtain the conditional expectations that are used in the MEF approach. Model-consistent proxies are used to keep the simulated paths from becoming too dispersed over time. In our implementation we consider an unobserved spot rate, but the approach can be generalized to include other latent variables such as expected inflation, stochastic volatility, (stochastic) discount rates, etc.

We illustrate our approach for a simple DSGE model where the equilibrium dynamics are available in terms of observable quantities, namely, consumption, output, and interest rates. It is important to study the properties of our estimation methods in simple continuous-time models before addressing more elaborate models at the vanguard of the DSGE literature, including financial intermediation and various frictions. We assume logarithmic preferences

together with a linear technology as an important benchmark case because it allows for an analytical solution. Moreover, we have some choice in modeling the interest rate dynamics. Specifications of this kind date back to Cox, Ingersoll, and Ross (1985a), and are frequently used in macro-finance models. Since the MEF approach is not limited to analytical solutions, but applicable to all cases where the researcher is provided with a solution in the form of policy functions, our illustrating example can be used as a point of reference for exploring broader classes of dynamic general equilibrium models.

We apply our model to both simulated and empirical data on production, consumption, and interest rates. Our Monte Carlo study then examines the properties of our estimation approaches in 1,000 simulated data sets of 25 years each for both monthly and quarterly macro data, along with daily financial market data, roughly in line with the availability of empirical figures. The results show that the GMM and MEF approaches generally are able to accurately estimate the parameters of the (correctly specified) model, and that the interest rate data helps identifying the structural parameters.

Our empirical application to 30 years of U.S. data shows that the system can be applied to a combination of macro and financial series. The results indicate a long run mean of the interest rate around 10% with a 1.5% volatility annually and only weak mean reversion. Comparing our simulated results to empirical estimates and/or the estimated interest rate, when treating it as a latent variable, indicates potential misspecification of the chosen model. Nonetheless, due to its simplicity and tractability, the AK-Vasicek model specification should provide a benchmark for future research.

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),

$$Y_t = A_t F(K_t, L), \quad (1)$$

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 ,

$$dA_t = \mu(A_t)dt + \eta(A_t)dB_t, \quad (2)$$

with $\mu(A_t)$ and $\eta(A_t)$ generic drift and volatility functions satisfying regularity conditions.⁸ The capital stock increases if gross investment I_t exceeds capital depreciation,

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

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 with respect to K_t and L , 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

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

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, determined by the output elasticity of capital.⁹

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

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

subject to

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

⁸We 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.

⁹Unless 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.

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, which can be thought of being in zero net supply. 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 CIR.

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

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

i.e., the present value of expected utility along the optimal program. It is shown in Appendix 7.1 that the first-order condition for the problem is (subscripts denote derivatives)

$$u_C(C_t, A_t) = V_K(K_t, A_t), \quad (8)$$

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

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

Economically, this equation gives the pricing kernel in the economy. Hence, we may use the Euler equation (9) to shed some light how the rate of return of the physical asset is linked to any risk-free security (cf. Posch, 2011). For this purpose we apply the conditional expectation and rewrite terms to arrive at

$$\underbrace{\rho - \frac{1}{dt} E_t \left[\frac{du_C}{u_C} \right]}_{\text{cost of forgone consumption}} = \underbrace{r_t - \delta + \frac{u_{CC}(C_t, A_t)}{u_C(C_t, A_t)} C_K \sigma^2 K_t}_{\text{certainty equivalent rate of return}} \equiv r_t^f. \quad (10)$$

Optimal behavior implies that the cost of forgone consumption on the left-hand side must equal the certainty equivalent rate of return r_t^f , corresponding to the rate of the instantaneously risk-free asset on the right-hand side of the equation.¹⁰

¹⁰Note that $-u_{CC}(C_t, A_t)C_t/u_C(C_t, A_t)$ measures the degree of relative risk aversion.

Moreover, the Euler equation 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, we obtain the path for consumption,

$$dC_t = \frac{u'(C_t)}{u''(C_t)}(\rho - (r_t - \delta))dt - \sigma^2 C_K K_t dt - \frac{1}{2}(C_A^2 \eta(A_t)^2 + C_K^2 \sigma^2 K_t^2) \frac{u'''(C_t)}{u''(C_t)} dt + C_A \eta(A_t) dB_t + C_K \sigma K_t dZ_t, \quad (11)$$

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

2.3 Equilibrium dynamics of the economy

Applying the logarithm on the variables of the stochastic differentials (6), (4), and (11) the equilibrium dynamics of the economy may be summarized as instantaneous growth rates

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

If all variables C_t , Y_t , and K_t along with TFP A_t were observed, estimation could be based directly on this system and the equation (2). While 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 in a unified macro-finance framework, instead. The idea is to use model-based equilibrium conditions to identify latent state variables using financial data. Thus, suppose that an interest rate r_t is identified, either directly in the data, or in the form of an equilibrium no-arbitrage condition such as r_t^f in (10), along with C_t and Y_t .¹¹ We consider systems of stochastic differential equations that can be used for estimation in this case, based on time series for (C_t, Y_t, r_t) , by recasting the equilibrium dynamics in terms of this triple.

2.4 AK-Vasicek model with logarithmic preferences

In this section we consider an economy with technology given by $Y_t = A_t K_t$, also known as AK model (this includes the technology in Brunnermeier and Sannikov, 2014), and assume

¹¹One caveat is that some variables are observed as an integral over an interval (a flow) rather than at a point in time (a stock; Harvey and Stock, 1989). We approximate a flow variable, e.g., $Y_t \Delta$ at time t by the integral $\int_{t-\Delta}^t Y_s ds$. Observed growth rates of flow variables therefore correspond to $\ln Y_t - \ln Y_{t-\Delta}$.

preferences of the type $u(C_t) = \ln C_t$. Though labor dynamics or more general classes of preferences have been documented to be important determinants of business cycle dynamics and asset pricing implications, our specification is interpreted as a parsimonious description which allows us to study the macro-finance links within production economies and optimizing agents. It may be interpreted and used as a benchmark specification to illustrate the issues and advantages of our approach by making use of an analytical solution. With these assumptions, $A_t = Y_K = r_t$ and $K_t = Y_t/A_t = Y_t/r_t$, so the two relevant state variables (A_t, K_t) are expressed as known functions of the observable variables (Y_t, r_t) . Because of $w_t = Y_L = 0$ the equilibrium dynamics can be summarized as

$$d \ln C_t = (r_t - \rho - \delta - \frac{1}{2}\sigma^2) dt + \sigma dZ_t, \quad (12a)$$

$$d \ln Y_t = (\mu(r_t)/r_t + r_t - \rho - \delta - \frac{1}{2}\eta(r_t)^2/r_t^2 - \frac{1}{2}\sigma^2) dt + \eta(r_t)/r_t dB_t + \sigma dZ_t, \quad (12b)$$

$$dr_t = \mu(r_t)dt + \eta(r_t)dB_t. \quad (12c)$$

In general, the consumption function is non-homogeneous with respect to TFP A_t and capital K_t (or wealth, here the output-TFP ratio).¹² The functions $\mu(\cdot)$ and $\eta(\cdot)$ are chosen such that suitable boundedness conditions are met (cf. Posch, 2009). Recall that in the AK framework, the interest rate (rental rate of capital) dynamics purely reflect TFP dynamics. In the tradition of the finance literature, we illustrate the estimation of the model with the interest rate governed by a Vasicek specification (henceforth the AK-Vasicek model).

The Vasicek (1977) mean-reverting specification for the rental rate of physical capital is $\mu(r_t) = \kappa(\gamma - r_t)$ and $\eta(r_t) = \eta$, where $\kappa > 0$ is the speed and γ the target rate of mean reversion, and η the constant volatility. In this case, the equilibrium dynamics are

$$d \ln C_t = (r_t - \rho - \delta - \frac{1}{2}\sigma^2) dt + \sigma dZ_t, \quad (13a)$$

$$d \ln Y_t = (\kappa\gamma/r_t - \frac{1}{2}\eta^2/r_t^2 + r_t - \kappa - \rho - \delta - \frac{1}{2}\sigma^2) dt + \eta/r_t dB_t + \sigma dZ_t, \quad (13b)$$

$$dr_t = \kappa(\gamma - r_t)dt + \eta dB_t. \quad (13c)$$

Alternative Markov diffusion specifications of the interest rate process as in Ait-Sahalia (1996, p. 528) can be implemented and the system estimated along the lines developed below, including the Cox, Ingersoll, and Ross (1985a) specification. The analytical solution does not depend on this particular choice.

In this AK-Vasicek model, the relation between the risk-free rate and the rental rate of

¹²In our benchmark case optimal consumption is linear in the capital stock, $C_t = \rho K_t$ (cf. Appendix 7.1).

capital in (10) is given by¹³

$$r_t^f = r_t - \delta - \sigma^2. \quad (14)$$

This result is quite intuitive if we recall that consumption is proportional to the capital stock, such that $C_K K_t = C_t$. The term $u_{CC}(C_t, A_t)C_K K_t/u_C(C_t, A_t)$ in (10) is the constant parameter of relative risk aversion, for logarithmic preferences equal to unity.

Economically, equation (14) is the equilibrium asset pricing relationship, which demands that the rate of return to any riskless financial asset r_t^f equals capital rewards r_t net of the rate of depreciation δ and the risk premium associated with holding the physical asset. This equation sheds light on the use of empirical data for estimating the macro-finance model: It is typically easier to relate the rental rate of capital to some observable riskless rate, such as the short rate (or 3-month interest rate) on some government bond, than to observe it directly (or to observe capital or TFP).

3 Estimation

In this section, we describe how to estimate the equilibrium system (13) using macro and financial data. First, we integrate the system to obtain an exact discrete-time analog in terms of (observable) variables. Section 3.1 presents the resulting formulation of the model for estimation purposes. Section 3.2 illustrates (i) how the reduced-form parameters can be estimated by means of standard regression-based methods, and (ii) how the structural parameters are obtained using minimum distance. Section 3.3 shows how structural parameters may alternatively be estimated directly using the martingale estimating function (MEF) approach and how this approach relates to generalized method of moments (GMM).

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. In what follows, we treat the triple of variables (C_t, Y_t, r_t^f) as being observable, though later on in an extension we consider a case with latent interest rate dynamics. Specifically, we use the risk-free rate rather than a direct measure of the rental rate of capital r_t as financial data.¹⁴ This (more realistic) assumption implies that r_t is linked to the data, but may also depend on parameters. Using the system of differential

¹³Note that δ would also capture a constant level of inflation. This assumption, however, neglects inflation dynamics. There is a number of ways to overcome this, which is part of our research agenda. Since the focus of the present paper is methodological, we leave a thorough examination for future research.

¹⁴We use daily data on the 3-month interest rate as a proxy for the risk-free rate (cf. Chapman, Long, and Pearson, 1999), along with aggregate consumption and output at lower frequencies.

equations (13) and the equilibrium asset-pricing condition (14), we obtain

$$\ln(C_s/C_t) - \int_t^s r_v^f dv = -(\rho - \frac{1}{2}\sigma^2)(s-t) + \sigma(Z_s - Z_t), \quad (15a)$$

$$\begin{aligned} \ln(Y_s/Y_t) - \int_t^s r_v^f dv &= \kappa\gamma \int_t^s 1/(r_v^f + \delta + \sigma^2)dv - \frac{1}{2}\eta^2 \int_t^s 1/(r_v^f + \delta + \sigma^2)^2 dv \\ &\quad - (\kappa + \rho - \frac{1}{2}\sigma^2)(s-t) \\ &\quad + \int_t^s \eta/(r_v^f + \delta + \sigma^2)dB_v + \sigma(Z_s - Z_t), \end{aligned} \quad (15b)$$

$$\begin{aligned} r_s^f &= e^{-\kappa(s-t)}r_t^f + (1 - e^{-\kappa(s-t)})(\gamma - \delta - \sigma^2) \\ &\quad + \eta e^{-\kappa(s-t)} \int_t^s e^{\kappa(v-t)} dB_v. \end{aligned} \quad (15c)$$

Hence, with $s - t$ fixed at Δ , the system (15) has martingale increments given by

$$\varepsilon_{C,t} = \sigma(Z_t - Z_{t-\Delta}), \quad (16a)$$

$$\varepsilon_{Y,t} = \int_{t-\Delta}^t \eta/(r_v^f + \delta + \sigma^2)dB_v + \sigma(Z_t - Z_{t-\Delta}), \quad (16b)$$

$$\varepsilon_{r,t} = \eta e^{-\kappa\Delta} \int_{t-\Delta}^t e^{\kappa(v-(t-\Delta))} dB_v. \quad (16c)$$

This system of three equations forms the basis of our empirical specifications. At the same time, it illustrates the main ideas underlying our approach. First, our 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 Riemann sums of the type $\int_t^s g(r_v^f)dv \approx (s-t) \sum_{i=1}^P g(r_{t+i(s-t)/P}^f)/P$, where $g(\cdot)$ is a smooth function of $r_{t+i(s-t)/P}^f$, the prevailing interest rate on a risk-free security on day i in the period between t and s , and P is the number of days in the period.¹⁵ Third, the structural parameters enter into the coefficients on the terms involving interest rates, thus showing that financial data indeed may serve to identify parameters of interest.

If we directly observed $r_t = r_t^f + \delta + \sigma^2$, the system would in fact be linear in a set of reduced-form parameters, which in turn are known functions of the structural parameters. This intriguing result may be exploited also when r_t^f is observed, instead, such that r_t depends on parameters: We may replace the observed time series r_t with any reasonable proxy \hat{r}_t from (14) using particular values $\delta = \delta_0$ and $\sigma = \sigma_0$. Technically, given the values for δ and σ , the rental rate of capital is uniquely identified from the risk-free rate,

$$\hat{r}_t = r_t^f + \delta_0 + \sigma_0^2. \quad (17)$$

¹⁵For notational convenience, we write P as a constant, but in our empirical approach we use the actual number of days in the period (month or quarter).

This step allows us to use off-the-shelf linear regression-based estimation methods in Section 3.2, but is not required for the GMM and MEF methods presented in Section 3.3. Because σ_0 and δ_0 are unknown, this causes problems for regression-based analysis.¹⁶

We have some choice in turning system (15) into an empirical specification. Initially, we specify a system of three regression equations for equidistant macro data, i.e., we define $\Delta \equiv s - t$ (we use $\Delta = 1/12$ for monthly data, $\Delta = 1/4$ for quarterly data). Given the higher (say, daily) frequency of financial 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 in the interest rate equation). In any case, the high-frequency property of the interest rate data are exploited in the approximation of the integrals as Riemann sums, rather than in using a separate frequency in the estimation of the third equation.

3.2 The regression-based approaches

In this section we propose 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 (IV), and estimation of structural parameters by minimum distance.

3.2.1 Reduced-form model

With $s - t$ fixed at Δ , and using the proxy series \hat{r}_t in (17), the system (15) is linear in a set of reduced-form parameters and may be recast as

$$y_{j,t} = x_{j,t}\beta_j + \varepsilon_{j,t}, \quad j = C, Y, r, \quad (18)$$

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

¹⁶In cases where δ and σ are identified by the remaining system of equations, we may interpret fixed values δ_0 and σ_0 in the construction of the auxiliary variable \hat{r}_t as starting values, then estimate the full set of parameters of the model and update the values for $\delta_i = \hat{\delta}_{i-1}$ and $\sigma_i = \hat{\sigma}_{i-1}$ recursively for $i = 1, 2, \dots$. Alternatively, a nonlinear one-step regression-based approach could be implemented.

structural parameters $\phi = (\kappa, \gamma, \eta, \rho, \delta, \sigma)^\top$ as

$$\beta_C = -\left(\rho - \frac{1}{2}\sigma^2\right)\Delta, \quad (19a)$$

$$\beta_{Y,1} = -\left(\kappa + \rho - \frac{1}{2}\sigma^2\right)\Delta, \quad (19b)$$

$$\beta_{Y,2} = \kappa\gamma, \quad (19b)$$

$$\beta_{Y,3} = -\frac{1}{2}\eta^2, \quad (19b)$$

$$\beta_{r,1} = (1 - e^{-\kappa\Delta})(\gamma - \delta - \sigma^2), \quad (19c)$$

$$\beta_{r,2} = e^{-\kappa\Delta}. \quad (19c)$$

Hence, the system (15) can be summarized in the form of simple regression equations, with error terms given by

$$\varepsilon_{C,t} = \sigma(Z_t - Z_{t-\Delta}), \quad (20a)$$

$$\varepsilon_{Y,t} = \int_{t-\Delta}^t (\eta/\hat{r}_v)dB_v + \sigma(Z_t - Z_{t-\Delta}), \quad (20b)$$

$$\varepsilon_{r,t} = \eta e^{-\kappa\Delta} \int_{t-\Delta}^t e^{\kappa(v-(t-\Delta))}dB_v. \quad (20c)$$

Using iterated expectations and the properties of stochastic integrals, the error terms are clearly serially uncorrelated, i.e., $E(\varepsilon_{j,t}\varepsilon_{j,t-\Delta}) = 0$, $j = C, Y, r$. For a simple reduced-form estimator, linearity in β suggests unrestricted equation-by-equation OLS,

$$\hat{\beta}_j = (x_j^\top x_j)^{-1}x_j^\top y_j, \quad j = C, Y, r, \quad (21)$$

where x_j is the matrix with typical row $x_{j,t}$ and y_j the vector with typical entry $y_{j,t}$. The structural parameter estimates, obtained by minimum distance applied to the reduced-form estimates (21) using the link (19a)-(19c) (or by an asymptotically equivalent restricted non-linear least squares regression), serve as useful benchmarks for assessing more elaborate structural approaches. We next discuss enhancing the basic OLS-based estimators by correction for contemporaneous cross-equation correlation of errors and endogeneity of right-hand side variables, then present the minimum distance approach yielding the structural parameter estimates.

3.2.2 Cross-equation correlation

The estimators (21) allow for different variances of the error terms $\varepsilon_{j,t}$, $j = C, Y, r$, as they are implemented separately by equation. However, it does not exploit any other property of the errors. The present model structure implies both different right-hand side variables (indeed, of different dimensions) across the equations, and cross-equation correlation of the errors. In particular, from (15), the errors in (18) take the form (20a)-(20c), where, e.g., the term $\sigma(Z_t - Z_{t-\Delta})$ is common to both (20a) and (20b). Classical seemingly unrelated

regressions (SUR) analysis is intended to exploit cross-equation correlation exactly in cases where the right-hand side variables are not common across equations. This suggests that a standard SUR correction of the reduced-form estimates should be more efficient than the OLS estimates, and, hence, that structural parameter estimates backed out from the SUR estimates (using minimum distance) should dominate those based on OLS.

Let $\hat{\varepsilon}$ be the $T \times 3$ matrix of OLS residuals, with typical row $(\hat{\varepsilon}_{C,t}, \hat{\varepsilon}_{Y,t}, \hat{\varepsilon}_{r,t})$, 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 $\beta = (\beta_C, \beta_Y^\top, \beta_r^\top)^\top$ is

$$\hat{\beta}_{SUR} = (x^\top \hat{V}^{-1} x)^{-1} x^\top \hat{V}^{-1} y, \quad (22)$$

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 variance-covariance matrices of the SUR (and OLS) estimators are given in Appendix 7.2.

3.2.3 Endogeneity

The regression approaches (OLS and SUR) do not control for possible endogeneity of right-hand side variables in (15), and hence (18), which may be an issue in the DSGE model. In particular, $x_{Y,t}$ includes two integrals involving the evolution of the auxiliary variable in (17) 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). Here, we consider first-stage regressions of each of $x_{Y,t,2} = \int_{t-\Delta}^t 1/\hat{r}_v dv$ and $x_{Y,t,3} = \int_{t-\Delta}^t 1/\hat{r}_v^2 dv$ on their respective lags $x_{Y,t-\Delta,2}$ and $x_{Y,t-\Delta,3}$ and an intercept. Next, in the computation (21) of $\hat{\beta}_Y$, 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 $\hat{\beta}_Y$ 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 IV assessment of $\hat{\Sigma}$. Finally, an FGLS-SUR-IV step is carried out using this new $\hat{\Sigma}$ in calculating $\hat{\beta}_{SUR}$ in (22) and again using the fitted values for $x_{Y,t,2}$ and $x_{Y,t,3}$. This combination of FGLS, SUR, and IV (labeled FGLS-SUR-IV) appears to be novel.

3.2.4 Minimum distance

The structural parameters are $\phi = (\kappa, \gamma, \eta, \rho, \delta, \sigma)^\top$, a total of six. They are identified by exploiting the way in which they enter into the reduced-form parameters $\beta = \beta(\phi)$. From (19a)-(19c), we may this way identify $\rho - \frac{1}{2}\sigma^2$, κ , γ , η , and $\delta + \sigma^2$, i.e., three structural

parameters, and two independent combinations of the remaining three. Note that this identification is conditional on the chosen value $\delta_0 + \sigma_0^2$ in the auxiliary variable \hat{r}_t that enters the regressors in (18). When iterating, this value is updated, as exactly the parameter combination $\delta + \sigma^2$ is one of the five that are conditionally identified. Ultimately, this identifies these five parameter functions. Instead of obtaining the five parameter functions, one can impose restrictions on ρ , δ or σ to identify all other parameters. It is also possible to separate ρ , σ , and δ , and thus identify all six structural parameters, by exploiting the functional form of the error variances (the variances of (20a)-(20c)). Indeed, including the variance of the residual (20a) from the consumption equation as a separate moment along with the relations (19a)-(19c) clearly identifies σ^2 and thereby the full parameter vector ϕ , i.e., all six structural parameters.

Why should we rely on first moments only if not all parameters are identified? In models with, say, stochastic volatility or more elaborated preference specification, the error term of the consumption equation becomes intractable (like the residual of the output equation). In such case, the econometrician may exploit the martingale property only, without considering second moments. Because we want to keep our analysis applicable to such specifications, we focus how to estimate the (identified) parameters from first moments in the main text, without going to higher moments. For comparison we show the results if we used the residual variance of the consumption and the interest rate equation in a web appendix.

In the given setup, with either five or six structural parameters thus identified, we extract estimates of them from the OLS, SUR, or FGLS-SUR-IV reduced-form parameter estimates using a minimum distance approach. We carry out minimum distance estimation based on three different unrestricted parameter sets ω_i , $i = 1, 2, 3$, from the reduced-form regressions: (1) the estimates of β in (19), i.e., the theoretical and empirical moments to match are $\omega_1(\phi) = \beta(\phi)$ and $\hat{\omega}_1 = \hat{\beta}$; (2) β along with the variance $\sigma^2\Delta$ of the consumption equation residual in (20a), so that $\omega_2(\phi) = (\omega_1(\phi)^\top, \sigma^2\Delta)^\top$ and $\hat{\omega}_2 = (\hat{\omega}_1^\top, \hat{\Sigma}_{CC})^\top$, with $\hat{\Sigma}_{CC}$ the upper left entry in the residual covariance matrix $\hat{\Sigma}$; (3) β along with the variances of the consumption and interest rate residuals (20a) and (20c), $\omega_3(\phi) = (\omega_2(\phi)^\top, \frac{1}{2}\eta^2(1-e^{-2\kappa\Delta})/\kappa)^\top$ and $\hat{\omega}_3 = (\hat{\omega}_2^\top, \hat{\Sigma}_{rr})^\top$. In each of the three cases, we solve the problem

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

Here, the relevant metrics are given by the precisions of the reduced form estimates,

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

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

with $\hat{\Sigma}^{ij}$ the (i, j) 'th entry in $\hat{\Sigma}^{-1}$.

The indicated matrix $\hat{\Omega}_1^{-1}$ is for the case where the reduced form estimates $\hat{\beta}$ are obtained using SUR, i.e., $\hat{\Omega}_1 = \hat{V}_{SUR}$. If $\hat{\beta}$ is instead obtained by OLS as in (21), then the correct $\hat{\Omega}_1 = \hat{V}_{OLS}$ is given in Appendix 7.2. A naive OLS assessment of $\hat{\Omega}_1^{-1}$ would have zero off-diagonal blocks, and diagonal blocks $\hat{\Sigma}_{jj}^{-1} x_j^\top x_j$ in the minimum distance approach. With endogeneity correction, i.e., the reduced form estimates are obtained by FGLS-SUR-IV, again the minimum distance approach requires a variance-covariance matrix, and this has the same form as in the SUR case, but with the new $\hat{\Sigma}$ from the FGLS-SUR-IV approach and with fitted values for the relevant portions of x .

Estimators that are asymptotically equivalent to these minimum distance estimators are alternatively obtained by restricted (nonlinear) regression, minimizing the OLS respectively the SUR objective function under the relevant restrictions (19a)-(19c) on β . 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 elements $\varepsilon_{j,t}$. If estimated residual variances are used along with the relations (19a)-(19c) to identify structural parameters in the minimum distance case, then an asymptotically equivalent estimator may be obtained by iterating on structural parameters as they enter both ε_t and $\Sigma = \Sigma(\phi)$, used instead of $\hat{\Sigma}$ in the SUR objective function. This use of the Gaussian log-likelihood function amounts to quasi maximum likelihood (QML) since clearly $\varepsilon_{Y,t}$ in (20b) is non-Gaussian.

3.3 GMM and the MEF approach

The issue remains whether all endogeneity issues in the structural DSGE model have been fully corrected for by FGLS-SUR-IV (or the asymptotically equivalent nonlinear regression methods). The lagged values of the relevant integrals involving the auxiliary variable \hat{r}_s , $t - 2\Delta \leq s \leq t - \Delta$, may correlate with $\hat{r}_{t-\Delta}$, and hence with $\varepsilon_{Y,t}$ from (20b), 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 implementation, and provides useful benchmark estimates and starting values for the optimal MEF estimator. A likelihood-based approach such as MEF, however, allows us to estimate the parameters without the need of constructing the auxiliary variable \hat{r}_t in (17), but directly using r_t^f in the estimation.

Let ϕ denote the parameter vector of interest, whether the structural parameters, 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, we let $m_t = \varepsilon_t = (\varepsilon_{C,t}, \varepsilon_{Y,t}, \varepsilon_{r,t})^\top$ from (20a)-(20b), so $N = 3$. Clearly, m_t is a martingale difference sequence, and from system (15) we have that in terms of data and parameters

$$m_t = \begin{pmatrix} \ln(C_t/C_{t-\Delta}) - \int_{t-\Delta}^t r_v^f dv + (\rho - \frac{1}{2}\sigma^2) \Delta \\ \ln(Y_t/Y_{t-\Delta}) - \int_{t-\Delta}^t r_v^f dv + (\kappa + \rho - \frac{1}{2}\sigma^2) \Delta - \kappa\gamma \int_{t-\Delta}^t 1/(r_v^f + \delta + \sigma^2) dv \\ \quad + \frac{1}{2}\eta^2 \int_{t-\Delta}^t 1/(r_v^f + \delta + \sigma^2)^2 dv \\ r_t^f - (1 - e^{-\kappa\Delta})(\gamma - \delta - \sigma^2) - e^{-\kappa\Delta} r_{t-\Delta}^f \end{pmatrix}, \quad (23)$$

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 From GMM to MEF

The MEF method differs slightly from the generalized method of moments (GMM) of Hansen (1982). It is at least as efficient as GMM – indeed, strictly more efficient except in the special case where the two estimators coincide. 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 right-hand side variables, $z_t = (1, \int_{t-2\Delta}^{t-\Delta} 1/(r_v^f + \delta + \sigma^2) dv, \int_{t-2\Delta}^{t-\Delta} 1/(r_v^f + \delta + \sigma^2)^2 dv, r_{t-2\Delta}^f)^\top$, since these are all in the information set at $t-\Delta$. In particular, it presents no new issue, neither for GMM nor MEF, that the instrumental variables depend not only on data, but also on parameters, $z_t = z_t(\phi)$. 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

$$H_T = \frac{1}{T} \sum_{t=1}^T h_t \quad (24)$$

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 for ϕ so as to equate 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 the squared norm $H_T(\phi)^\top W H_T(\phi)$, where W is a weight matrix. Optimal GMM is obtained by using the identity matrix $I_{\dim h}$ for W in a first step minimization, then using the resulting estimator to calculate a consistent estimate of $\text{var}(H_T)^{-1}$ that is used for W in the second step minimization. Writing $\hat{\phi}_0$ for the first step estimator using $W = I_{\dim h}$, an example could be $\text{var}(H_T)^{-1} = (\sum_t h_t(\hat{\phi}_0)h_t(\hat{\phi}_0)^\top)^{-1}$. Sometimes, a Newey and West (1987) correction is used, for robustness against serial correlation, but it is unnecessary under the null that m_t and hence h_t is a martingale difference sequence.

The important features of GMM that leave room for improvement and hence the MEF approach are now evident: First, the instruments z_t enter in the form of a vector, whereas MEF uses a matrix. Secondly, the dimension of H_T is the same as or greater than the number of parameters, whereas MEF specifically uses the same number of estimating equations and parameters. In short, the optimal estimating equations are based on matrix-valued rather than vector-valued instruments, and over-identifying restrictions ($\dim h > \dim \phi$) are unnecessary for efficiency.

To develop these ideas, note that the first order conditions for the minimization in GMM are

$$\frac{\partial H_T(\phi)^\top}{\partial \phi} W H_T(\phi) = 0, \quad (25)$$

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 (25). Thus, G could be based on the first step GMM estimator, just like W , i.e., the system is

$$\left(\sum_t \frac{\partial h_t}{\partial \phi}(\hat{\phi}_0) \right)^\top \left(\sum_t h_t(\hat{\phi}_0)h_t(\hat{\phi}_0)^\top \right)^{-1} \sum_t h_t(\phi) = 0,$$

where ϕ only appears in the last factor. The equations are solved by treating $G(\hat{\phi}_0)$ as fixed and finding ϕ that sets (25) 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 . This is the central idea of the MEF approach. Thus, there are again $\dim \phi$ equations, but they now take the more general form

$$\sum_{t=1}^T g_t(\hat{\phi}_0) h_t(\phi) = 0, \quad (26)$$

instead of $G(\hat{\phi}_0) \sum_{t=1}^T h_t(\phi) = 0$. 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

$$M_T = \sum_{t=1}^T w_t m_t, \quad (27)$$

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

$$M_T(\phi) = 0. \quad (28)$$

The optimal weights are given by

$$w_t = \psi_t^\top (\Psi_t)^{-1}, \quad (29)$$

where Ψ_t is the conditional variance of the vector martingale increment,

$$\Psi_t = \text{Var}_{t-\Delta}(m_t) = E_{t-\Delta}(m_t m_t^\top), \quad (30)$$

and ψ_t the conditional mean of its parameter derivative

$$\psi_t = E_{t-\Delta}(\partial m_t / \partial \phi^\top). \quad (31)$$

The conditioning on information available through $t - \Delta$ requires integrating out with respect to the evolution of the interest rate appearing in the integrals from $t - \Delta$ through t in (23). 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 (29) gives the optimal martingale estimating function, across choice of weights w_t . The optimal weights do depend on parameters, i.e., the martingale estimate $\hat{\phi}$ solves a system of the form $\sum_t w_t(\phi)m_t(\phi) = 0$, where the solution accounts for the parameter dependence of both w_t and m_t . Alternatively, an asymptotically equivalent estimator may be obtained by using weights evaluated at initial consistent estimates $\hat{\phi}_0$, e.g., from GMM. In this case, $\hat{\phi}$ is calculated as the solution with respect to ϕ of the system $\sum_t w_t(\hat{\phi}_0)m_t(\phi) = 0$. In both cases, the estimator is consistent (in particular, the endogeneity issue is resolved) and asymptotically normal,

$$\sqrt{T}(\hat{\phi} - \phi) \rightarrow \mathcal{N}(0, V_{MEF}), \quad (32)$$

with asymptotic variance-covariance matrix given by

$$V_{MEF} = (E(\psi_t^\top (\Psi_t)^{-1} \psi_t))^{-1}, \quad (33)$$

consistently estimated by the inverse sample average $\hat{V}_{MEF} = (T^{-1} \sum_{t=1}^T \psi_t^\top (\Psi_t)^{-1} \psi_t)^{-1}$. 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 . In the AK-Vasicek model with log utility, this Jacobian has rank six, so all six structural parameters may be identified when $N = 3$.

3.3.2 Comparison of GMM and MEF

Before applying the MEF method, let us briefly compare MEF and optimal GMM. Obviously, the GMM estimator is consistent, and the consistent estimate of the asymptotic variance takes the form $\hat{V}_{GMM} = ((T^{-1} \sum_{t=1}^T \partial h_t / \partial \phi^\top)^\top (T^{-1} \sum_{t=1}^T h_t h_t^\top)^{-1} (T^{-1} \sum_{t=1}^T \partial h_t / \partial \phi^\top))^{-1}$. 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 $V_{GMM} =$

$(E(\partial h_t/\partial \phi^\top)^\top \text{var}(h_t)^{-1} E(\partial h_t/\partial \phi^\top))^{-1}$, 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 $E(\partial h_t/\partial \phi^\top) = E(\psi_t)$, $\text{var}(h_t) = E(\Psi_t)$. It follows that the efficiency comparison is simply

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

The asymptotic variance of the MEF 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? 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 derivative ψ_t from (31) and conditional variance Ψ_t from (30) 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 special structure, i.e., it is represented in the Kronecker product form $z_t \otimes I_{\dim z}$, which is usually not the case.

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 MEF with three moment restrictions

For illustration, we report the functional form of the martingale estimating function for the AK-Vasicek model with logarithmic utility. This should be compared to the basic optimal GMM estimator, e.g., the two-step estimator based on the four-dimensional vector of instrumental variables from the previous section. 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 (23), 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 (28), we need weights w_t in (29), 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 $\Psi_{t,11} = \sigma^2\Delta$, $\Psi_{t,22} = \eta^2 E_{t-\Delta}(\int_{t-\Delta}^t 1/(r_v^f + \delta + \sigma^2)^2 dv) + \sigma^2\Delta$, 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 $\Psi_{t,23} = \eta^2 e^{-\kappa\Delta} E_{t-\Delta} \left((\int_{t-\Delta}^t 1/(r_v^f + \delta + \sigma^2) dB_v) (\int_{t-\Delta}^t e^{\kappa(v-(t-\Delta))} dB_v) \right)$. Since analytical expressions are not available, we use Euler approximations for $\Psi_{t,22}$ and $\Psi_{t,23}$,

$$\Psi_t = \begin{pmatrix} \sigma^2\Delta & \sigma^2\Delta & 0 \\ \sigma^2\Delta & \sigma^2\Delta + \eta^2\Delta/(r_{t-\Delta}^f + \delta + \sigma^2)^2 & \eta^2 e^{-\kappa\Delta}\Delta/(r_{t-\Delta}^f + \delta + \sigma^2) \\ 0 & \eta^2 e^{-\kappa\Delta}\Delta/(r_{t-\Delta}^f + \delta + \sigma^2) & \frac{1}{2}\eta^2(1 - e^{-2\kappa\Delta})/\kappa \end{pmatrix}. \quad (34)$$

Note that the consistency and the expression for the asymptotic variance are unaffected by our approximations because they enter only in the weights (29). Using martingale increments (23), we get the derivatives $(\partial m_t/\partial\phi^\top)^\top$ with respect to the parameter vector $\phi = (\kappa, \gamma, \eta, \rho, \delta, \sigma)^\top$,

$$\begin{pmatrix} 0 & \Delta - \gamma \int_{t-\Delta}^t 1/(r_v^f + \delta + \sigma^2) dv & -\Delta e^{-\kappa\Delta}\gamma + \Delta e^{-\kappa\Delta} r_{t-\Delta}^f \\ 0 & -\kappa \int_{t-\Delta}^t 1/(r_v^f + \delta + \sigma^2) dv & -(1 - e^{-\kappa\Delta}) \\ 0 & \eta \int_{t-\Delta}^t 1/(r_v^f + \delta + \sigma^2)^2 dv & 0 \\ \Delta & \Delta & 0 \\ 0 & \kappa\gamma \int_{t-\Delta}^t 1/(r_v^f + \delta + \sigma^2)^2 dv - \eta^2 \int_{t-\Delta}^t 1/(r_v^f + \delta + \sigma^2)^3 dv & (1 - e^{-\kappa\Delta}) \\ -\sigma\Delta & -\sigma\Delta + 2\sigma\kappa\gamma \int_{t-\Delta}^t 1/(r_v^f + \delta + \sigma^2)^2 dv & 2\sigma(1 - e^{-\kappa\Delta}) \\ & -2\sigma\eta^2 \int_{t-\Delta}^t 1/(r_v^f + \delta + \sigma^2)^3 dv & \end{pmatrix}. \quad (35)$$

Now apply conditional expectation to get $\psi_t = E_{t-\Delta}(\partial m_t/\partial\phi^\top)$, interchange the order of integration in (35), and use the deterministic Taylor expansion (e.g., Ait-Sahalia, 2008), which reads

$$E(g(r_s)|r_u) = \sum_{i=0}^k \frac{\Delta^i}{i!} A^i g(r_u) + O(\Delta^{k+1}), \quad s \geq u \quad (36)$$

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 example, a first-order Taylor expansion, $k = 1$, yields

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

in which $r_v = r_v^f + \delta + \sigma^2$. Thus expanding all terms involving integrals in (35), the transpose of the conditional mean of parameter derivatives ψ_t^\top reads

$$\begin{pmatrix} 0 & \Delta - \gamma \left(\Delta/r_{t-\Delta} - (\kappa(\gamma - r_{t-\Delta})/r_{t-\Delta}^2 - \eta^2/r_{t-\Delta}^3) \frac{1}{2}\Delta^2 \right) & \Delta e^{-\kappa\Delta} (r_{t-\Delta}^f - \gamma) \\ 0 & -\kappa \left(\Delta/r_{t-\Delta} - (\kappa(\gamma - r_{t-\Delta})/r_{t-\Delta}^2 - \eta^2/r_{t-\Delta}^3) \frac{1}{2}\Delta^2 \right) & -(1 - e^{-\kappa\Delta}) \\ 0 & \eta \left(\Delta/r_{t-\Delta}^2 - (2\kappa(\gamma - r_{t-\Delta})/r_{t-\Delta}^3 - 3\eta^2/r_{t-\Delta}^4) \frac{1}{2}\Delta^2 \right) & 0 \\ \Delta & \Delta & 0 \\ 0 & \kappa\gamma \left(\Delta/r_{t-\Delta}^2 - (2\kappa(\gamma - r_{t-\Delta})/r_{t-\Delta}^3 - 3\eta^2/r_{t-\Delta}^4) \frac{1}{2}\Delta^2 \right) & (1 - e^{-\kappa\Delta}) \\ -\sigma\Delta & -\eta^2 \left(\Delta/r_{t-\Delta}^3 - (3\kappa(\gamma - r_{t-\Delta})/r_{t-\Delta}^4 - 6\eta^2/r_{t-\Delta}^5) \frac{1}{2}\Delta^2 \right) & 2\sigma(1 - e^{-\kappa\Delta}) \\ -\sigma\Delta & -\sigma\Delta + 2\sigma\kappa\gamma \left(\Delta/r_{t-\Delta}^2 - (2\kappa(\gamma - r_{t-\Delta})/r_{t-\Delta}^3 - 3\eta^2/r_{t-\Delta}^4) \frac{1}{2}\Delta^2 \right) & 2\sigma(1 - e^{-\kappa\Delta}) \\ & -2\sigma\eta^2 \left(\Delta/r_{t-\Delta}^3 - (3\kappa(\gamma - r_{t-\Delta})/r_{t-\Delta}^4 - 6\eta^2/r_{t-\Delta}^5) \frac{1}{2}\Delta^2 \right) & \end{pmatrix}, \quad (37)$$

in which, again, $r_{t-\Delta} = r_{t-\Delta}^f + \delta + \sigma^2$. 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 $\hat{\phi}$. The asymptotic distribution is given by (32)-(33).

3.3.4 MEF extensions: Latent variables and mixed frequency

So far, we have considered the case where all variables in the system are observable, albeit using some mixed-frequency properties of the data. The MEF approach can be generalized to the empirically relevant cases of latent variables (e.g., unobserved real interest rates, stochastic volatility, stochastic discount rates, etc.) and mixed-frequency estimation with a different frequency for each series. To illustrate, we consider two representative cases, set in the context of the AK-Vasicek model: (i) The daily interest rate r_t is latent and cannot be backed out from data – there is no observed series proxying for this variable; (ii) Output Y_t is observed at a lower (say, quarterly) frequency than consumption C_t (say, monthly).

Case (i) serves to illustrate our approach to missing data series. For example, expected inflation and hence the real rate of interest may be treated as missing. An interesting feature of the approach is that we may infer the latent series and conduct a model specification check. Case (ii) reinforces our use of data sampled at mixed frequencies. For example, output may be proxied by industrial production at the monthly frequency (see below), 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.

For the latent variable generalization, Case (i), note that 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, 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 $E(\sum_t w_t E(m_t | \mathcal{F}_{t-\Delta})) = 0$, for weights w_t depending only on

information through $t - \Delta$. Thus, in the estimation, we may replace selected additive terms in the moments m_t by their conditional expectations given $\mathcal{F}_{t-\Delta}$. This allows replacing the integrals involving the daily interest rate by conditional expectations given monthly interest rate proxies, based on the information set.

For our application to the stochastic AK-Vasicek model, the procedure allows deriving moments for estimation, say, $m_t^* = E(m_t | \mathcal{F}_{t-\Delta})$, given by

$$\left(\begin{array}{c} \ln(C_t/C_{t-\Delta}) - E\left(\int_{t-\Delta}^t r_v dv | r_{t-\Delta}^*\right) + (\rho + \delta + \frac{1}{2}\sigma^2) \Delta \\ \ln(Y_t/Y_{t-\Delta}) + (\kappa + \rho + \frac{1}{2}\sigma^2 + \delta) \Delta - E\left(\int_{t-\Delta}^t r_v dv + \kappa\gamma \int_{t-\Delta}^t 1/r_v dv - \frac{1}{2}\eta^2 \int_{t-\Delta}^t 1/r_v^2 dv | r_{t-\Delta}^*\right) \\ r_t^* - (1 - e^{-\kappa\Delta})\gamma - e^{-\kappa\Delta}r_{t-\Delta}^* \end{array} \right), \quad (38)$$

where $r_{t-\Delta}^*$ is an interest rate proxy based on consumption and income data through $t - \Delta$. Here, $\Delta = 1/4$ 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 model-consistent proxy at the macro frequency is $r_t^* = \rho Y_t/C_t$ in which Y_t/C_t is approximated by the ratio of quarterly observed income and consumption data.

One possibility for implementation of the conditional expectations of the integrals in (38), i.e., integrating out the latent interest rate process r_v , is simulation. We refer to the resulting procedure as Simulated MEF, or SMEF. 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. The interest rate (or 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 (37) and (34) for ψ_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 model, through ρ) of the implied state variables is accounted for.

The SMEF approach applies generally to models involving latent variables. For example, in the stochastic volatility case, 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 model-consistent simulation.

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

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

in the intra-quarter periods when output is missing. Here, $r_v = r_v^f + \delta + \sigma^2$, using observed daily r_v^f . This is model consistent prediction, not simulation. In particular, the resulting

proxy series Y_t^* depends on the parameters. The proxy series is now substituted for Y_t in the original estimating equation $\sum_t w_t m_t = 0$. We refer to this procedure as mixed-frequency MEF, or MF-MEF. In particular, when solving for the parameter estimates, the dependence of the constructed output proxy series on trial parameter values is again accounted for.

Both generalized approaches, SMEF and MF-MEF, are akin to filtering. Thus, in the presence of latent variables, Case (i), SMEF recasts m_t in the estimating equation in terms of a set of conditional expectations or filtered predictions, given the information actually available. For mixed-frequency estimation, Case (ii), MF-MEF replaces missing data by conditional predictions given the actual observations. In both cases, standard errors may be calculated 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 report results on the sensitivity to DGP values and further results in an accompanying web appendix (cf. Appendix B).

4.1 Set-up

We simulate 25 years of both monthly and quarterly data from the model. We use simple Euler approximations to the differential equations in (13). 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 model: 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 in Section 3.1, by taking the average of the functions $g(r_v)$ over the 25 simulated days per month. For example, $\int_{t-\Delta}^t 1/(r_v^f + \delta + \sigma^2) dv$ for the monthly simulations is approximated by $(\sum_{i=1}^{25} 1/(r_{t-\Delta+i\Delta/25} + \delta + \sigma^2))\Delta/25$. 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 normalize initial output to unity and initialize the other variables consistently with theory: $\ln(Y_0) = 0$, $r_0 = \gamma$, and $\ln(C_0) = \ln(\rho \times Y_0/r_0)$.

We generate 1,000 data sets and estimate the parameters according to the approaches of Section 3. In particular, we report the parameter estimates for the OLS, FGLS-SUR-IV, GMM, and MEF methods. In the first two cases, we use the minimum distance approach to get the structural parameters from the reduced form estimates. We choose the data

generating process (DGP) parameter values in a way roughly corresponding to empirical estimates obtained in Section 5.2 below. In particular, we use $\kappa = 0.2$, $\gamma = 0.1$, $\eta = 0.01$, $\rho = 0.03$, $\delta = 0.05$, and $\sigma = 0.02$ (see column DGP in Table 1). In the web appendix, we report the sensitivity of our estimation methods to different DGP values (cf. Table A1).

4.2 Simulation results: Monthly and quarterly data

[insert Table 1]

Table 1 provides the results of the simulation study. 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 from the simulated monthly data, and in columns 6-9 the estimates from the quarterly data. For our four estimation methods we provide the median estimate of each parameter, and below the interquartile range of the 1,000 estimates. Not all six structural parameters are identified in the regression-based estimation methods, when not exploiting second moments (in particular, the residual variances of $\varepsilon_{C,t}$ and $\varepsilon_{r,t}$). In fact, five parameter combinations are identified, so one possibility is to fix one parameter at the outset and estimate the remaining five. There is some choice regarding which parameter to set, since the two combinations $\rho - \frac{1}{2}\sigma^2$ and $\delta + \sigma^2$ are identified, i.e., either ρ , δ , or σ^2 could be restricted. We choose to set δ to the DGP value of 0.05, which economically is interpreted as depreciation of physical capital of 5 percent per year, in the regression-based approaches. The same identification issue occurs for the GMM when the standard lagged right-hand side variables are used as instruments and no higher order moments are used. Also here we set δ to the DGP value of 0.05.

Overall, while simple OLS, ignoring the estimation problems described above, has some trouble identifying the structural parameters, the FGLS-SUR-IV, GMM, and MEF approaches produce estimates of γ , η and ρ that are remarkably close to the values in the DGP. Without exploiting further moments, σ seems to be only weakly identified using GMM, as reflected by the large inter-quartile range, or generates largely biased point estimates in the regression-based approaches. In the MEF approach, however, σ is already identified from considering only three moment conditions. Here, the identification works through the (optimally chosen) weight matrix and the conditional mean of parameter derivatives of the martingale increments. Similar results hold for both monthly and quarterly data.

The mean-reversion parameter κ of the Vasicek specification is more difficult to estimate: Here, a value 0.2 is used in the DGP, while the median estimates are in the range from 0.30 to 0.35 using monthly data, and only slightly better in quarterly data. This upward bias in the mean-reversion parameter estimate is well established (see Tang and Chen, 2009;

Wang, Phillips, and Yu, 2011). In particular, for values of κ close to zero, i.e., a near unit root situation typical of many financial time series, a bias correction may be preferable (Yu, 2012). We find a similar magnitude of the bias across the regression-based, GMM, and MEF approaches, and for different sets of DGP values (cf. Table A1). This finite-sample bias, however, does not seem to translate to other parameters (cf. Figure 1). Moreover, given the relatively wide inter-quartile range, the κ estimates are still within reasonable distance.

[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 data (Panel A) and quarterly data (Panel B). The figure confirms the findings from Table 1: The parameters $\gamma, \eta, \rho, \delta$ and σ are centered close to the DGP values. In addition, it becomes clear that the mode of the histograms for κ in fact is 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.

We also implement minimum distance (cf. Section 3.2.4), using the residual variance from the consumption (not reported), or both the consumption and interest rate equation as additional moments (cf. Table A2 in the web appendix), along with β in the regression-based approaches, allowing better identification of σ for monthly data (columns 2 and 3) respectively quarterly data (columns 6 and 7). A similar idea for the GMM and MEF methods is to include further moment restrictions in the martingale estimating equation (28). The results based on five moment restrictions (instead of three moment restrictions), with the conditional moments derived in Appendix 7.3, are reported in Table A2, columns 4 and 5 for monthly data respectively 8 and 9 for quarterly data. Including more moments indeed yields (better) identification of the six parameters. The relatively poor performance of the regression-based approaches clearly relates to the restriction on first moments. In cases where the econometrician is able to use second moments, it is advisable to so.

Taken together, the simulation study indicates that the GMM and MEF approaches are successful in recovering parameter estimates from the data. The regression-based methods exhibit reasonable performance, but only after accounting for potential estimation problems due to cross-equation correlation and endogeneity. In empirical work, the regression-based approaches would also require iteration over the proxy \hat{r}_t in (17). In our simulation study, the values for δ_0 and σ_0 are set to their corresponding DGP values δ and σ .¹⁷

4.3 Simulation results: Latent short rate and mixed frequency

[insert Table 2]

¹⁷We examine the sensitivity of OLS and FGLS-SUR-IV parameter estimates to δ_0 and σ_0 in Table A3.

In this section we present the simulation results for two possible MEF extensions, the missing data points in the mixed-frequency approach (MF-MEF) and in the case of latent-variables using simulation-based approach (SMEF). Table 2 provides the results for the simulation study of the AK-Vasicek model. In the first column we list the parameter values as they are used in the data generating process (DGP), in column 3 the SMEF estimates obtained on simulated monthly data, in column 5 the SMEF estimates for the simulated quarterly data, and in column 6 the MF-MEF estimates for the mixed-frequency data. For comparison, we also replicate the MEF results from Table 1 in columns 2 and 4, respectively, using monthly and quarterly data. As before, we provide the median estimate of each parameter, and below the interquartile range of the 1,000 estimates.

For the latent variable extension, Case (i), we compare SMEF (columns 3 and 5) with MEF estimates (columns 2 and 4). We find that the latent variable case is as good as the observed short rate process. At both the monthly and the quarterly observation frequency, the point estimates and interquartile ranges are estimated remarkably close to DGP values and are comparable with the MEF figures with slightly smaller interquartile ranges in the SMEF approach for ρ and σ . Of course, these findings hold true only if the data were simulated from the correct model. This fact allows us to run model-specification checks on the empirical data at hand. The simulated short rate process can actually be compared with some observed proxies (see also the discussion in Section 5.3).

For the extension to mixed-frequency data, Case (ii), we compare MF-MEF (column 6) with MEF estimates (columns 2 and 4). As one would expect, given the correct specification, for the case when output is replaced by model consistent predictions at intra-quarter periods the point estimates are remarkably close to the monthly estimates. Comparing the MF-MEF results to MEF, where consumption and output is observed at the quarterly frequency, we find that we gain better identification in σ , reflected by the smaller interquartile range.

[insert Figure 2]

In Figure 2 we provide the histograms of the 1,000 estimates that we obtain for the parameters using both the SMEF for monthly data and the MF-MEF approaches. Comparing the histograms of SMEF (Panel A) to monthly MEF in Figure 1 (also Panel A) illustrates that ρ and σ are better identified in SMEF, which also is reflected by smaller interquartile ranges above, while the histogram is slightly more narrow for δ in the MEF approach. Similarly, comparing the histograms of MF-MEF (Panel B) to monthly and quarterly MEF, respectively, in Figure 1 (Panels A and B) shows that there is a small efficiency loss with respect to monthly data, but better identification of parameters is obtained relative to the results when estimates are obtained solely from quarterly data.

Overall, both extensions work and are potentially more important for the case when we apply the method to empirical data by the same reasons that motivated our extensions.

4.4 Robustness: Time invariance and high-frequency data

In this section we present two robustness simulation results which are particularly relevant for the estimation of continuous-time models. We want to provide answers to the following two questions: (i) Are the estimates time invariant? In theory, the continuous-time model is time invariant. However, different continuous-time processes may look identical if sampled at discrete points, which sometimes is referred to as the aliasing problem. This phenomenon may prevent unique identification of the parameters of the continuous-time stochastic process from equidistant discrete-time observations. Moreover, any temporal aggregation of the data may distort our parameter estimates. For these reasons, it seems important to examine to which extent our parameter estimates change with the observation frequency. (ii) Does the high-frequency data matter? So far, we only exploit the high-frequency property of the interest rate in the approximation of the integrals as Riemann sums. Hence, we want to examine to which extent the use of daily observation respectively only considering the end-of-period figure helps to identify the parameters in our analysis.

[insert Table 3]

In order to examine to which extent the parameter estimates change with the sampling frequency, we simulate monthly and quarterly data respectively, with the same number of observations for comparison. In Table 3 we compare the usual 25 years of monthly data to 75 years of simulated quarterly data (Panel A). As before, we provide the median estimate of each parameter, and below the interquartile range of the 1,000 estimates. The results show that the bias in the κ estimate is much smaller with quarterly data than if the data were sampled at monthly frequency. Moreover, the interquartile ranges are substantially smaller with quarterly data for all four estimation methods. It reveals that the time invariance property translates to all parameters of interest except the mean-reversion parameter κ . This upward bias, however, seems to diminish if quarterly data were used provided the number of observation is sufficiently large (compare also to the results in Table 1).

We also examine what our results would like if we would not use the daily availability of interest rates but only used the end-of-period number. To this end we simulate the data as usual, but only use the end-of-month and end-of-quarter short rate in our estimation rather than the integrals. This simply neglects all within-period dynamics. In Table 3 we show the results for monthly data and quarterly data (Panel B). Comparing to the results in Table 1 shows that neglecting within-period dynamics is not innocuous. The general pattern is that

it comes at the cost of increasing inter-quartile ranges and changes in parameter estimates. In particular, the estimate for the mean-reversion parameter κ changes substantially for OLS, FGLS-SUR-IV, and GMM. Moreover, we get into more severe identification problems for σ in the regression-based approaches and now also for GMM. In contrast, we observe only minor efficiency losses for the MEF approach. Here, we still provide some information about the dynamics of the stochastic process by the deterministic Taylor expansion (36). This pattern suggests that both high-frequency data and/or more information about the within-period dynamics help to identify the parameters of interest.

5 Data and Results

In this section we estimate the AK-Vasicek model with logarithmic preferences based on empirical data. We provide results for the estimation approaches developed in Section 3, using US mixed frequency macro and financial data.

5.1 Data

[insert Figure 3]

To estimate the system (15) 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. We use real Personal Consumption Expenditures (PCE) at the monthly and the quarterly level to proxy consumption. In Figure 3 we show the plots of the monthly and quarterly growth rates of the variables (Panels A and B). Our data set spans the period from January 1982 to December 2012.

We combine the data on these aggregate macro series with financial data typically at higher frequency, in particular, the short rate. This rate is a theoretical concept and corresponds to an infinitesimal time to maturity. In applied work, the short rate is sometimes treated as a latent variable that is filtered from observed yield data (e.g., De Jong, 2000). As a starting point, we follow Chapman, Long, and Pearson (1999), and use the 3-month interest rate as a proxy for the short rate r_t^f of the risk-free financial asset, here taken as the US treasury bonds. This interest rate is available from the FRED data set at daily frequency. We use this series to obtain our monthly and quarterly figure by taking the last observation in the relevant period. Panel (C) of of Figure 3 shows the daily interest rate series. In the series, a general downward trend of the interest rate is evident.

Finally, we use the interest rate series to compute approximations to the integrals that appear in our empirical specification. We approximate the monthly and quarterly series of integrals using the daily spot rate observations. Given system (15), we approximate three integrals: $\int_{t-\Delta}^t g(r_v^f) dv \approx \Delta/P \sum_{i=1}^P g(r_{t-\Delta+i\Delta/P}^f)$, where $r_{t-\Delta+i\Delta/P}^f$ is the 3-month rate on day i of period t , and P the number of days in the period between $t - \Delta$ and t .

5.2 Estimates: Monthly and quarterly data

When taking our model to the empirical data, we experienced numerical optimization difficulties with some of the parameters, which may be due to possible model misspecification. From Section 4.2, five parameter combinations in theory are identified without exploiting second moments (specifically, the consumption growth residual variance). Hence, setting one parameter should allow estimating the remaining five, but the regression-based methods and GMM had difficulty doing so. The iterative optimization routines either diverged or produced economically unreasonable estimates that furthermore depended heavily on starting values. Consequently, in the reported results, two of the six parameters are set at pre-fixed values for these methods, instead of just one. Again, since $\rho - \frac{1}{2}\sigma^2$ and $\delta + \sigma^2$ are identified, either two of ρ , δ , or σ^2 could be restricted. From the simulation study, σ is weakly identified when not exploiting second moments, so we set this at 0.02, and δ again at 0.05 (cf. also Table A4 in the accompanying web appendix, where we use variance terms for the regression-based methods and 5 moments for both GMM and MEF).

In the MEF approach, all six parameters are identified even without using second moments. Unlike in the simulations, we implemented a slightly simplified version of MEF in order to avoid similar problems as those encountered with the other methods. Essentially, the optimal MEF weights $\psi_t^\top (\Psi_t)^{-1}$ from (29) were replaced by weights $\psi_t^\top (\hat{\Psi}_t)^{-1}$, i.e., still with time-varying conditional mean parameter derivatives of martingale increments (31), but the conditional variance (30) replaced by an estimate. Similarly to optimal two-step GMM, we first estimated with $\Psi = I_3$, then with $\hat{\Psi}_t$ computed as the outer product of the fitted residuals at time t from the first step. We label this approach two-step GMM. All six parameters were successfully estimated in both steps, and the values for δ and σ broadly in line with the pre-specified values used in the other methods.¹⁸

These identification problems can be interpreted as a first indication that the AK-Vasicek specification with logarithmic preferences probably will not match the data well. This leads to some degree of problems for all estimation procedures, in particular GMM, although less so for the MEF. The likely model misspecification is discussed further below in section 5.3.

¹⁸In the web appendix we show that the two-step MEF approach yields similar results compared to MEF with optimal weights (cf. Table A5).

[insert Table 4]

Table 4 provides structural parameter estimates based on both monthly and quarterly data, using industrial production respectively GDP for output, obtained using the OLS, FGLS-SUR-IV, GMM, and MEF approaches. The regression-based estimation methods provide fairly similar estimates of the four parameters for monthly and quarterly data. By the point estimates, the short rate is mean-reverting, but not very strongly, with speed parameter κ around 0.08 (0.047 to 0.109 for quarterly data). A speed of zero implies a unit root. The implied first order autocorrelation is $e^{-0.08 \times 1/12} = 0.99$ for monthly data. The long-term target rate γ is about 10%, and the volatility η of the short rate innovation is between 1.3% and 2.6%. As is well-known, the interest rate has been declining during the period (see Figure 3, Panel C), so the model will not yield a good fit and is likely misspecified, but it is worth noting that the 10% level makes sense. Thus, by the asset pricing equation (14), it comprises the average risk-free rate from the data, the risk premium σ^2 consistent with logarithmic preferences, and the rate δ of physical capital depreciation. For the given δ and σ , the time preference parameter ρ is estimated at around 1% in monthly and 2% in quarterly data. Of course, it is important to note that the two data sets differ not only by sampling frequency, but also by relying on industrial production respectively GDP.

The GMM estimates of κ and γ are similar to those from the regression-based methods, whereas the point estimates of ρ are about 0.5 percentage points smaller. The main difference is that GMM does not pick up any of the innovation variance in the interest rate process. The MEF results are slightly different in some respects, and it should be kept in mind that they are obtained without restricting δ and σ . In particular, in quarterly data, the depreciation rate δ is estimated at 6.2%, i.e., 25% higher than the pre-set value used for the other methods. Consistently with this, the long-run mean interest rate γ is higher, too, at 13%. In monthly data, these are lower, at 2.5% respectively 5.1%, and the κ and η estimates at both frequencies are similar to those from other methods (except that GMM had trouble estimating η). Further, MEF produces precise estimates of ρ at the quarterly frequency and σ at the monthly, even significant at conventional levels (many of the received estimates are statistically insignificant, across all parameters and methods). Indeed, both the monthly and the quarterly MEF estimate of σ confirms the value 0.02 imposed in the other methods.

All in all, the application to the empirical data shows that it is possible to estimate our simple benchmark macro-finance model using macro and high-frequency financial data in combination. The model is so simple as to be likely misspecified, but the exercise nevertheless suggests that the MEF is the most promising approach for our purposes.

5.3 Estimates: Latent short rate and mixed frequency

[insert Table 5]

Table 5 shows the results for our two extensions, the latent short rate and mixed-frequency estimation (columns 3 and 4). For comparison, we replicate in columns 1 and 2 the MEF results from Table 4. The latent variable extension, Case (i), reveals more evidence on the sources of misspecification: the counter-factual model-implied short rate when it is compared to the observed interest rate proxy. The results of SMEF in column 3 show that if parts of the data were model-generated, the interest rate consistent with macro dynamics would show much higher mean reversion κ about 13% and a much lower long-term target rate γ about 5.8% with very small innovation variance η close to zero. The unusual high t -values reflect the fact the standard errors are likely downward bias (so the t -values upward biased), because the additional uncertainty of drawing the short rate is not taken into account, which in principle could be accounted for by bootstrapping. For illustration, in Figure 4 we plot one simulated short rate path r_t^* . It is worth noting that in contrast to the observed risk-free rate r_t^f , any model-implied short rate r_t^* , for the given consumption and income data, will be upward sloping. Because the proxy r_t^* is set to the ratio $\rho Y_t/C_t$ at quarterly frequency (indicated by dots), the model-implied short rate follows the same pattern. For this reason, SMEF is not applicable to monthly empirical data since IP is measured as an index.

The mixed-frequency data, Case (ii), indicates that the MF-MEF long-term value for the interest rate γ is even smaller, about 3%, but quite persistent. The point estimate for the speed of mean-reversion parameter κ is 2.3%, which suggests a near unit root behavior. The innovation variance η is close to zero. Our MF-MEF approach yields a more precise estimate for ρ of about 1.3% compared to the MEF estimate of 2.1% and 0.3% for quarterly respectively monthly data. The point estimate for the depreciation rate δ using quarterly GDP and monthly consumption is smaller than the MEF values, about 1% only.

6 Conclusion

The literature has been relatively 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. In this paper we provide an econometric framework in which macroeconomics, finance and econometrics are coherently linked. The framework is developed in a continuous-time setting, that conveniently allows for thinking about variables observed at different frequencies.

This paper describes regression-based procedures, OLS and FGLS-SUR-IV, GMM and

the asymptotically efficient MEF approach in order to estimate the structural parameters of continuous-time DSGE models using mixed-frequency macro 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 and macro data can indeed be used jointly to facilitate the estimation of structural parameters in continuous-time versions of the general equilibrium models. Overall, on the methodological side, our work suggests that MEF is preferred over GMM and the regression-based approaches as long as the econometrician is restricted to first moments. It allows identifying all structural parameters already from first moments, and estimates are more precise, numerically stable, and economically meaningful. We provide two extensions of MEF akin to filtering in order to estimate models with latent variables based on simulations, SMEF, and model-consistent prediction for mixed-frequency data, MF-MEF. Development of further general equilibrium models in the Cox, Ingersoll, and Ross (1985a) framework to more elaborate specifications, and formal testing of these is part of our research agenda.

7 Appendix

7.1 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

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

Using the properties of stochastic integrals, we may write

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

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

$$\begin{aligned} \rho V(K_t, A_t) &= u(C(K_t, A_t), A_t) + ((r_t - \delta)K_t + w_t - C(K_t, A_t)) V_K \\ &\quad + \frac{1}{2} (V_{AA} \eta(A_t)^2 + V_{KK} \sigma^2 K_t^2) + V_A \mu(A_t), \end{aligned} \tag{39}$$

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

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

Collecting terms we obtain

$$\begin{aligned} (\rho - (r_t - \delta))V_K &= ((r_t - \delta)K_t + w_t - C_t)V_{KK} \\ &\quad + \frac{1}{2} (V_{AAK}\eta(A_t)^2 + V_{KKK}\sigma^2 K_t^2) + V_{KK}\sigma^2 K_t + V_{AK}\mu(A_t). \end{aligned} \quad (40)$$

Using Itô's formula, the costate obeys

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

where inserting (40) 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 (39) and the first-order condition (8) are fulfilled. Our guess is

$$V(K_t, A_t) = \mathbb{C}_1 \ln K_t + f(A_t). \quad (41)$$

From (8), optimal consumption is a constant fraction of wealth, $C_t = \mathbb{C}_1^{-1}K_t$. Now use the maximized Bellman equation (39) 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 $g(A_t) \equiv \rho f(A_t) - \frac{1}{2} (f_{AA}\eta(A_t)^2 - \sigma^2) - f_{A\mu}(A_t)$. 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 $\mathbb{C}_3 = (\kappa\gamma\mathbb{C}_2 - \ln \mathbb{C}_1 - 1 - (\delta + \frac{1}{2}\sigma^2)\mathbb{C}_1)/\rho$.

7.2 The SUR estimator

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 $\hat{\Sigma}^{-1}$. Thus,

$$\hat{V}_{SUR} = \begin{pmatrix} \hat{\Sigma}^{CC} x_C^\top x_C & \hat{\Sigma}^{CY} x_C^\top x_Y & \hat{\Sigma}^{Cr} x_C^\top x_r \\ \hat{\Sigma}^{YC} x_Y^\top x_C & \hat{\Sigma}^{YY} x_Y^\top x_Y & \hat{\Sigma}^{Yr} x_Y^\top x_r \\ \hat{\Sigma}^{rC} x_r^\top x_C & \hat{\Sigma}^{rY} x_r^\top x_Y & \hat{\Sigma}^{rr} x_r^\top x_r \end{pmatrix}^{-1}.$$

If the 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 (21) has blocks estimated as $\hat{\Sigma}_{ij} (x_i^\top x_i)^{-1} x_i^\top x_j (x_j^\top x_j)^{-1}$, i.e., \hat{V}_{OLS} equals

$$\begin{pmatrix} \hat{\Sigma}_{CC} (x_C^\top x_C)^{-1} (x_C^\top x_C) (x_C^\top x_C)^{-1} & \hat{\Sigma}_{CY} (x_C^\top x_C)^{-1} (x_C^\top x_Y) (x_Y^\top x_Y)^{-1} & \hat{\Sigma}_{Cr} (x_C^\top x_C)^{-1} (x_C^\top x_r) (x_r^\top x_r)^{-1} \\ \hat{\Sigma}_{YC} (x_Y^\top x_Y)^{-1} (x_Y^\top x_C) (x_C^\top x_C)^{-1} & \hat{\Sigma}_{YY} (x_Y^\top x_Y)^{-1} (x_Y^\top x_Y) (x_Y^\top x_Y)^{-1} & \hat{\Sigma}_{Yr} (x_Y^\top x_Y)^{-1} (x_Y^\top x_r) (x_r^\top x_r)^{-1} \\ \hat{\Sigma}_{rC} (x_r^\top x_r)^{-1} (x_r^\top x_C) (x_C^\top x_C)^{-1} & \hat{\Sigma}_{rY} (x_r^\top x_r)^{-1} (x_r^\top x_Y) (x_Y^\top x_Y)^{-1} & \hat{\Sigma}_{rr} (x_r^\top x_r)^{-1} (x_r^\top x_r) (x_r^\top x_r)^{-1} \end{pmatrix}^{-1}$$

and $\hat{V}_{OLS} \geq \hat{V}_{SUR}$ in the partial order of positive semi-definite matrices.

7.3 MEF with five moment restrictions

Let $m_t = (\varepsilon_{C,t}, \varepsilon_{Y,t}, \varepsilon_{r,t}, \varepsilon_{C,t}^2, \varepsilon_{r,t}^2)$ be the 5-vector of error terms (16), which again clearly is a martingale differences in terms of data and parameters,

$$m_t^{(5)} = \begin{pmatrix} \ln(C_t/C_{t-\Delta}) - \int_{t-\Delta}^t r_v^f dv + (\rho - \frac{1}{2}\sigma^2) \Delta \\ \ln(Y_t/Y_{t-\Delta}) - \int_{t-\Delta}^t r_v^f dv + (\kappa + \rho - \frac{1}{2}\sigma^2) \Delta - \kappa\gamma \int_{t-\Delta}^t 1/(r_v^f + \delta + \sigma^2) dv \\ \quad + \frac{1}{2}\eta^2 \int_{t-\Delta}^t 1/(r_v^f + \delta + \sigma^2)^2 dv \\ r_t^f - (1 - e^{-\kappa\Delta})(\gamma - \delta - \sigma^2) - e^{-\kappa\Delta} r_{t-\Delta}^f \\ \left(\ln(C_t/C_{t-\Delta}) - \int_{t-\Delta}^t r_v^f dv + (\rho - \frac{1}{2}\sigma^2) \Delta \right)^2 - \sigma^2 \Delta \\ \left(r_t^f - (1 - e^{-\kappa\Delta})(\gamma - \delta - \sigma^2) - e^{-\kappa\Delta} r_{t-\Delta}^f \right)^2 - \eta^2 (1 - e^{-2\kappa\Delta}) / (2\kappa) \end{pmatrix}$$

or by using the definition of three moment increments

$$m_t^{(5)} = \begin{pmatrix} m_t^{(3)} \\ \left(\ln(C_t/C_{t-\Delta}) - \int_{t-\Delta}^t r_v^f dv + (\rho - \frac{1}{2}\sigma^2) \Delta \right)^2 - \sigma^2 \Delta \\ \left(r_t^f - (1 - e^{-\kappa\Delta})(\gamma - \delta - \sigma^2) - e^{-\kappa\Delta} r_{t-\Delta}^f \right)^2 - \eta^2 (1 - e^{-2\kappa\Delta}) / (2\kappa) \end{pmatrix} \quad (42)$$

which is equivalent to consider

$$m_t^{(5)} = \begin{pmatrix} \sigma(Z_t - Z_{t-\Delta}) \\ \int_{t-\Delta}^t \eta/(r_v^f + \delta + \sigma^2) dB_v + \sigma(Z_t - Z_{t-\Delta}) \\ \eta e^{-\kappa\Delta} \int_{t-\Delta}^t e^{\kappa(v-(t-\Delta))} dB_v \\ \sigma^2(Z_t - Z_{t-\Delta})^2 - \sigma^2\Delta \\ \eta^2 e^{-2\kappa\Delta} \left(\int_{t-\Delta}^t e^{\kappa(v-(t-\Delta))} dB_v \right)^2 - \eta^2(1 - e^{-2\kappa\Delta})/(2\kappa) \end{pmatrix}$$

or

$$m_t^{(5)} = \begin{pmatrix} m_t^{(3)} \\ \sigma^2(Z_t - Z_{t-\Delta})^2 - \sigma^2\Delta \\ \eta^2 e^{-2\kappa\Delta} \left(\int_{t-\Delta}^t e^{\kappa(v-(t-\Delta))} dB_v \right)^2 - \eta^2(1 - e^{-2\kappa\Delta})/(2\kappa) \end{pmatrix}. \quad (43)$$

To construct the MEF (28), we need the weights w_t in (29), which depend on the conditional mean of the parameter derivatives, ψ_t , and the conditional variance, Ψ_t , of m_t . We have the conditional variances $\Psi_{t,11}^{(5)} = \sigma^2\Delta$, $\Psi_{t,22}^{(5)} = \eta^2 E_{t-\Delta}(\int_{t-\Delta}^t 1/(r_v^f + \delta + \sigma^2)^2 dv) + \sigma^2\Delta$, and $\Psi_{t,33}^{(5)} = \eta^2(1 - e^{-2\kappa\Delta})/(2\kappa)$, $\Psi_{t,44}^{(5)} = 2\sigma^4\Delta^2$, and $\Psi_{t,55}^{(5)} = \eta^4 e^{-4\kappa\Delta} E_{t-\Delta} \left(\left(\int_{t-\Delta}^t e^{\kappa(v-(t-\Delta))} dB_v \right)^4 \right) - \frac{1}{4}\eta^4(1 - e^{-2\kappa\Delta})^2/\kappa^2$. Similarly, the conditional covariances are $\Psi_{t,12}^{(5)} = \sigma^2\Delta$, $\Psi_{t,13}^{(5)} = 0$, $\Psi_{t,14}^{(5)} = 0$, $\Psi_{t,15}^{(5)} = 0$, $\Psi_{t,23}^{(5)} = \eta^2 e^{-\kappa\Delta} E_{t-\Delta} \left(\left(\int_{t-\Delta}^t 1/(r_v^f + \delta + \sigma^2) dB_v \right) \left(\int_{t-\Delta}^t e^{\kappa(v-(t-\Delta))} dB_v \right) \right)$, $\Psi_{t,24}^{(5)} = 0$, $\Psi_{t,25}^{(5)} = \eta^3 e^{-2\kappa\Delta} E_{t-\Delta} \left(\left(\int_{t-\Delta}^t 1/(r_v^f + \delta + \sigma^2) dB_v \right) \left(\int_{t-\Delta}^t e^{\kappa(v-(t-\Delta))} dB_v \right)^2 \right)$, $\Psi_{t,35}^{(5)} = \eta^3 e^{-3\kappa\Delta} E_{t-\Delta} \left(\left(\int_{t-\Delta}^t e^{\kappa(v-(t-\Delta))} dB_v \right)^3 \right)$, $\Psi_{t,34}^{(5)} = \Psi_{t,45}^{(5)} = 0$. We use Euler approximations for $\Psi_{t,22}^{(5)}$, $\Psi_{t,55}^{(5)}$, $\Psi_{t,23}^{(5)}$, $\Psi_{t,25}^{(5)}$ and $\Psi_{t,35}^{(5)}$,

$$\Psi_t^{(5)} = \begin{pmatrix} \sigma^2\Delta & \sigma^2\Delta & 0 & 0 & 0 \\ \sigma^2\Delta & \sigma^2\Delta + \eta^2\Delta/(r_{t-\Delta}^f + \delta + \sigma^2)^2 & \eta^2 e^{-\kappa\Delta}\Delta/(r_{t-\Delta}^f + \delta + \sigma^2) & 0 & 0 \\ 0 & \eta^2 e^{-\kappa\Delta}\Delta/(r_{t-\Delta}^f + \delta + \sigma^2) & \frac{1}{2}\eta^2(1 - e^{-2\kappa\Delta})/\kappa & 0 & 0 \\ 0 & 0 & 0 & 2\sigma^4\Delta^2 & 0 \\ 0 & 0 & 0 & 0 & \Psi_{t,55}^{(5)'} \end{pmatrix}$$

where $\Psi_{t,55}^{(5)'}$ = $3\eta^4 e^{-4\kappa\Delta}\Delta^2 - \frac{1}{4}\eta^4(1 - e^{-2\kappa\Delta})^2/\kappa^2$ or by using the definition of $\Psi_t^{(3)}$

$$\Psi_t^{(5)} = \begin{pmatrix} \Psi_t^{(3)} & & 0_{3 \times 2} \\ 0_{2 \times 3} & 2\sigma^4\Delta^2 & 0 \\ & 0 & 3\eta^4 e^{-4\kappa\Delta}\Delta^2 - \frac{1}{4}\eta^4(1 - e^{-2\kappa\Delta})^2/\kappa^2 \end{pmatrix}. \quad (44)$$

Again the consistency and the expression for the asymptotic variance are unaffected by our approximations because they enter only in the weights (29). Using martingale increments (42), we get the derivatives $(\partial m_t / \partial \phi^\top)^\top$ with respect to the parameter vector

$\phi = (\kappa, \gamma, \eta, \rho, \delta, \sigma)^\top$, such that $(\psi_t^{(5)})^\top$ reads

$$\begin{pmatrix} 0 & \frac{1}{2}\eta^2(1 - e^{-2\kappa\Delta})/\kappa^2 - \eta^2\Delta e^{-2\kappa\Delta}/\kappa \\ 0 & 0 \\ (\psi_t^{(3)})^\top & -\eta(1 - e^{-2\kappa\Delta})/\kappa \\ 0 & 0 \\ 0 & 0 \\ -2\sigma\Delta & 0 \end{pmatrix}. \quad (45)$$

where we used the fact that

$$\psi_{t,44}^{(5)} = 2E_{t-\Delta} \left(\ln(C_t/C_{t-\Delta}) - \int_{t-\Delta}^t r_v^f dv + (\rho - \frac{1}{2}\sigma^2)\Delta \right) \Delta = 0$$

which completes the construction of the estimating equations for 5 moment conditions.

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.
- AMEMIYA, T., AND R. Y. WU (1972): “The Effect of Aggregation on Prediction in the Autoregressive Model,” *J. Amer. Statistical Assoc.*, 67(339), 628–632.
- AN, S., AND F. SCHORFHEIDE (2007): “Bayesian Analysis of DSGE models,” *Econometric Rev.*, 26(2-4), 113–172.
- ANDREOU, E., E. GHYSELS, AND A. KOURTELLOS (2010): “Regression models with mixed sampling frequencies,” *J. Econometrics*, 158, 246–261.
- (2013): “Should Macroeconomic Forecasters Use Daily Financial Data and How?,” *J. Bus. Econ. Statist.*, 31, 1–12.
- ANG, A., AND M. PIAZZESI (2003): “A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables,” *J. Monet. Econ.*, 50, 745–787.
- ANG, A., M. PIAZZESI, AND M. WEI (2006): “What does the yield curve tell us about GDP growth?,” *J. Econometrics*, 131, 359–403.
- BERGSTROM, A. R. (1966): “Nonrecursive Models as Discrete Approximations to Systems of Stochastic Differential Equations,” *Econometrica*, 34(1), 173–182.

- BERNANKE, B., AND M. GERTLER (1989): “Agency Costs, Net Worth, and Business Fluctuations,” *Amer. Econ. Rev.*, 79(1), 14–31.
- BERNANKE, B., M. GERTLER, AND S. GILCHRIST (1999): *The Financial Accelerator in a Quantitative Business Cycle Framework*. in J.B. Taylor and M. Woodford, Eds., Handbook of Macroeconomics, Volume 1, Chapter 21, 1341-1393.
- BRUNNERMEIER, M. K., AND Y. SANNIKOV (2014): “A Macroeconomic Model with a Financial Sector,” *Amer. Econ. Rev.*, 104(2), 379–421.
- 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.
- CHRISTIANO, L. J., AND M. EICHENBAUM (1987): “Temporal Aggregation and Structural Inference in Macroeconomics,” *Carnegie-Rochester Conference Series on Public Policy*, 26, 63–130.
- 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.
- DEWACHTER, H., AND M. LYRIO (2006): “Macro Factors and the Term Structure of Interest Rates,” *J. Money, Credit, Banking*, 38(1), 119–140.
- DIEBOLD, F. X., G. R. RUDEBUSCH, AND S. B. ARUOBA (2006): “The macroeconomy and the yield curve: a dynamic latent factor approach,” *J. Econometrics*, 131, 309–338.
- EATON, J. (1981): “Fiscal Policy, Inflation and the Accumulation of Risky Capital,” *Rev. Econ. Stud.*, 48, 435–445.
- FERNÁNDEZ-VILLAVARDE, J., AND J. F. RUBIO-RAMÍREZ (2007): “Estimating Macroeconomic Models: A Likelihood Approach,” *Rev. Econ. Stud.*, 74, 1059–1087.

- FERNÁNDEZ-VILLAYERDE, 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.
- FORONI, C., AND M. MARCELLINO (2014): “Mixed-Frequency Structural Models: Identification, Estimation, and Policy Analysis,” *J. Appl. Econometrics*.
- GERTLER, M., AND P. KARADI (2011): “A model of unconventional monetary policy,” *J. Monet. Econ.*, 58, 17–34.
- GEWEKE, J. (1978): “Temporal Aggregation in the Multiple Regression Model,” *Econometrica*, 46(3), 643–661.
- GHYSELS, E., A. SINKO, AND R. VALKANOV (2007): “MIDAS Regressions: Further Results and New Directions,” *Econometric Rev.*, 26, 53–90.
- GIANNONE, D., F. MONTI, AND L. REICHLIN (2014): “Exploiting the monthly data-flow in structural forecasting,” *mimeo*.
- GODAMBE, V. P., AND C. C. HEYDE (1987): “Quasi Likelihood and Optimal Estimation,” *International Statistical Review*, 55, 231–244.
- 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.
- HÖRDAHL, P., O. TRISTANI, AND D. VESTIN (2006): “A joint econometric model of macroeconomic and term-structure dynamics,” *J. Econometrics*, 131, 405–444.
- JERMANN, U. J. (1998): “Asset Pricing in Production Economies,” *J. Monet. Econ.*, 41, 257–275.
- JERMANN, U. J., AND V. QUADRINI (2012): “Macroeconomic Effects of Financial Shocks,” *Amer. Econ. Rev.*, 102(1), 238–271.

- KIM, T. B. (2010): “Temporal Aggregation Bias and Mixed Frequency Estimation of a New Keynesian model,” *mimeo*.
- KIYOTAKI, N., AND J. MOORE (1997): “Credit Cycles,” *J. Polit. Economy*, 105(2), 211–248.
- KLOEDEN, P. E., AND E. PLATEN (1999): *Numerical Solution of Stochastic Differential Equations*. Springer, 3rd ed.
- MARCELLINO, M. (1999): “Some Consequences of Temporal Aggregation in Empirical Analysis,” *J. Bus. Econ. Statist.*, 17(1), 129–136.
- MCCRORIE, J. R. (2009): “Estimating Continuous-Time Models on the Basis of Discrete Data via an Exact Discrete Analog,” *Econometric Theory*, 25, 1120–1137.
- 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.
- (2011): “Risk premia in general equilibrium,” *J. Econ. Dynam. Control*, 35(9), 1557–1576.
- 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.
- RUDEBUSCH, G. D., AND T. WU (2008): “A Macro-Finance Model of the Term Structure, Monetary Policy, and the Economy,” *Econ. J.*, 118, 906–926.
- RUGE-MURCIA, F. J. (2007): “Methods to estimate dynamic stochastic general equilibrium models,” *J. Econ. Dynam. Control*, 31, 2599–2636.
- SCHORFHEIDE, F., AND D. SONG (2014): “Real-Time Forecasting with a Mixed-Frequency VAR,” *J. Bus. Econ. Statist.*, forthcoming.
- SENNEWALD, K. (2007): “Controlled Stochastic Differential Equations under Poisson Uncertainty and with Unbounded Utility,” *J. Econ. Dynam. Control*, 31(4), 1106–1131.

- SIMS, C. A. (1971): “Discrete Approximations to Continuous Time Distributed Lags in Econometrics,” *Econometrica*, 39(3), 545–563.
- TALLARINI, T. (2000): “Risk-Sensitive Real Business Cycles,” *J. Monet. Econ.*, 45, 507–532.
- TANG, C. Y., AND S. X. CHEN (2009): “Parameter estimation and bias correction for diffusion processes,” *J. Econometrics*, 149, 65–81.
- VAN BINSBERGEN, J. H., J. FERNÁNDEZ-VILLAYERDE, R. S. J. KOIJEN, AND J. F. RUBIO-RAMÍREZ (2012): “The Term Structure of Interest Rates in a DSGE Model with Recursive Preferences,” *J. Monet. Econ.*, 59, 634–648.
- VASICEK, O. (1977): “An Equilibrium Characterization of the Term Structure,” *J. Finan. Econ.*, 5, 177–188.
- WANG, X., P. C. B. PHILLIPS, AND J. YU (2011): “Bias in estimating multivariate and univariate diffusions,” *J. Econometrics*, 161, 228–245.
- YU, J. (2012): “Bias in the estimation of the mean reversion parameter in continuous time models,” *J. Econometrics*, 169, 114–122.

Table 1: Simulation Study – Monthly and Quarterly Data

The table reports output of a simulation study of the accuracy of the structural model parameters estimated using the OLS, FGLS-SUR-IV, GMM and MEF approaches for the AK-Vasicek model. 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 & Quarterly Data									
		Monthly Data				Quarterly Data			
	DGP	OLS	FGLS-SUR-IV	GMM	MEF	OLS	FGLS-SUR-IV	GMM	MEF
κ	0.2	0.349 0.286	0.299 0.134	0.345 0.345	0.354 0.284	0.354 0.290	0.225 0.119	0.287 0.319	0.353 0.305
γ	0.1	0.201 0.036	0.101 0.013	0.100 0.014	0.099 0.013	0.198 0.036	0.100 0.014	0.101 0.015	0.099 0.013
η	0.01	0.083 0.036	0.008 0.004	0.010 0.001	0.010 0.001	0.083 0.035	0.007 0.003	0.010 0.002	0.010 0.001
ρ	0.03	0.080 0.015	0.030 0.006	0.030 0.007	0.030 0.006	0.079 0.015	0.030 0.006	0.031 0.007	0.030 0.006
δ	0.05	0.05	0.05	0.05	0.050 0.002	0.05	0.05	0.05	0.050 0.003
σ	0.02	0.317 0.040	0.000 <0.001	0.027 0.047	0.023 0.005	0.312 0.044	0.000 <0.001	0.040 0.064	0.025 0.010

Table 2: Simulation Study – Latent Short Rate and Mixed Frequency

The table reports output of a simulation study of the accuracy of the structural model parameters estimated using the latent short rate and mixed-frequency MEF approaches for the AK-Vasicek model, SMEF (Latent Short Rate) and MF-MEF, respectively. 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. For completeness we include the MEF estimates from Table 1.

Parameter Estimates from Simulation Study – SMEF (Latent Short Rate) and MF-MEF						
DGP		Monthly Data		Quarterly Data		Mixed Frequency
		MEF	SMEF	MEF	SMEF	MF-MEF
κ	0.2	0.354 0.284	0.355 0.280	0.353 0.305	0.363 0.290	0.360 0.290
γ	0.1	0.099 0.013	0.099 0.012	0.099 0.013	0.107 0.020	0.099 0.013
η	0.01	0.010 0.001	0.010 0.001	0.010 0.001	0.010 0.002	0.010 0.001
ρ	0.03	0.030 0.006	0.030 0.002	0.030 0.006	0.032 0.005	0.030 0.006
δ	0.05	0.050 0.002	0.051 0.005	0.050 0.003	0.055 0.013	0.050 0.002
σ	0.02	0.023 0.005	0.021 0.003	0.025 0.010	0.021 0.006	0.022 0.005

Table 3: Robustness Simulations – Time invariance and high-frequency data

The table reports output of two simulation studies to the robustness of our estimation methods. In a first simulation (Panel A) we examine to which extent the parameter estimates change with the sampling frequency. To this end we simulate the quarterly data set-up with the same number of observations as the monthly set-up. Specifically, we compare the usual 25 years of monthly data to 75 years of quarterly data. In a second robustness simulation (Panel B) we examine what our results would like if we would not use the daily availability of interest rates but only used the end-of-period number. To this end we simulate the data as usual, but only use the end-of-month and end-of-quarter short rate in our estimation rather than the integrals. We show the accuracy of the structural model parameters estimated using OLS, FGLS-SUR-IV, GMM and MEF for the AK-Vasicek model with three moment restrictions. For 1,000 replications, we generate the 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.

Panel A: Robustness Simulations – Time invariance									
		Monthly Data				Quarterly Data (75 years)			
	DGP	OLS	FGLS-SUR-IV	GMM	MEF	OLS	FGLS-SUR-IV	GMM	MEF
κ	0.2	0.349 0.286	0.299 0.134	0.345 0.345	0.354 0.284	0.236 0.116	0.179 0.058	0.246 0.145	0.246 0.120
γ	0.1	0.201 0.036	0.101 0.013	0.100 0.014	0.100 0.013	0.190 0.023	0.101 0.008	0.101 0.009	0.100 0.008
η	0.01	0.083 0.036	0.008 0.004	0.010 0.001	0.010 0.001	0.065 0.019	0.007 0.002	0.010 0.001	0.010 0.001
ρ	0.03	0.080 0.015	0.030 0.006	0.030 0.007	0.030 0.006	0.075 0.009	0.030 0.003	0.030 0.004	0.030 0.003
δ	0.05	0.05	0.05	0.05	0.050 0.002	0.05	0.05	0.05	0.050 0.002
σ	0.02	0.317 0.040	0.000 <0.001	0.027 0.047	0.023 0.005	0.299 0.031	0.000 <0.001	0.018 0.052	0.022 0.006

Panel B: Robustness Simulations – Daily vs. Monthly and Quarterly Short Rate									
		Monthly Data				Quarterly Data			
	DGP	OLS	FGLS-SUR-IV	GMM	MEF	OLS	FGLS-SUR-IV	GMM	MEF
κ	0.2	0.188 0.444	0.395 0.260	0.164 0.164	0.356 0.286	0.184 0.444	0.387 0.287	0.144 0.137	0.353 0.305
γ	0.1	0.241 0.205	0.100 0.013	0.099 0.017	0.098 0.012	0.236 0.195	0.100 0.013	0.099 0.018	0.094 0.013
η	0.01	0.077 0.105	0.009 0.004	0.010 0.001	0.010 0.001	0.075 0.102	0.009 0.003	0.010 0.002	0.010 0.001
ρ	0.03	0.104 0.095	0.030 0.006	0.030 0.006	0.030 0.006	0.101 0.091	0.030 0.006	0.031 0.007	0.030 0.006
δ	0.05	0.05	0.05	0.05	0.048 0.002	0.05	0.05	0.05	0.045 0.004
σ	0.02	0.389 0.435	0.000 0.011	0.000 0.038	0.023 0.005	0.383 0.427	0.000 <0.001	0.032 0.057	0.024 0.010

Table 4: Estimates – Monthly and Quarterly Data

The table reports estimates for the structural model parameters estimated using OLS, FGLS-SUR-IV, GMM, and MEF approaches for the AK-Vasicek model with three moment restrictions. We run the estimation for monthly data (where production is measured by IP) and quarterly data (production measured by GDP). The sample runs from January, 1982 until December, 2012. Asymptotic t -statistics are given below the estimates.

Parameter Estimates from Empirical Data								
	Monthly Data				Quarterly Data			
	OLS	FGLS-SUR-IV	GMM	MEF	OLS	FGLS-SUR-IV	GMM	MEF
κ	0.094 0.582	0.073 2.46	0.054 0.314	0.060 0.930	0.109 1.94	0.047 0.221	0.062 0.971	0.068 0.515
γ	0.094 2.08	0.089 <0.001	0.064 0.549	0.051 0.479	0.129 <0.001	0.119 7.97	0.095 2.25	0.131 0.483
η	0.016 0.097	0.013 <0.001	0.000 <0.001	0.006 0.639	0.026 0.081	0.013 1.97	0.000 <0.001	0.014 0.221
ρ	0.014 0.976	0.014 0.430	0.005 0.129	0.003 0.802	0.021 0.974	0.020 0.871	0.015 0.753	0.021 5.95
δ	0.05	0.05	0.05	0.025 0.162	0.05	0.05	0.05	0.062 0.144
σ	0.02	0.02	0.02	0.020 3.3	0.02	0.02	0.02	0.020 1.15

Table 5: Estimates – Latent Short Rate and Mixed Frequency

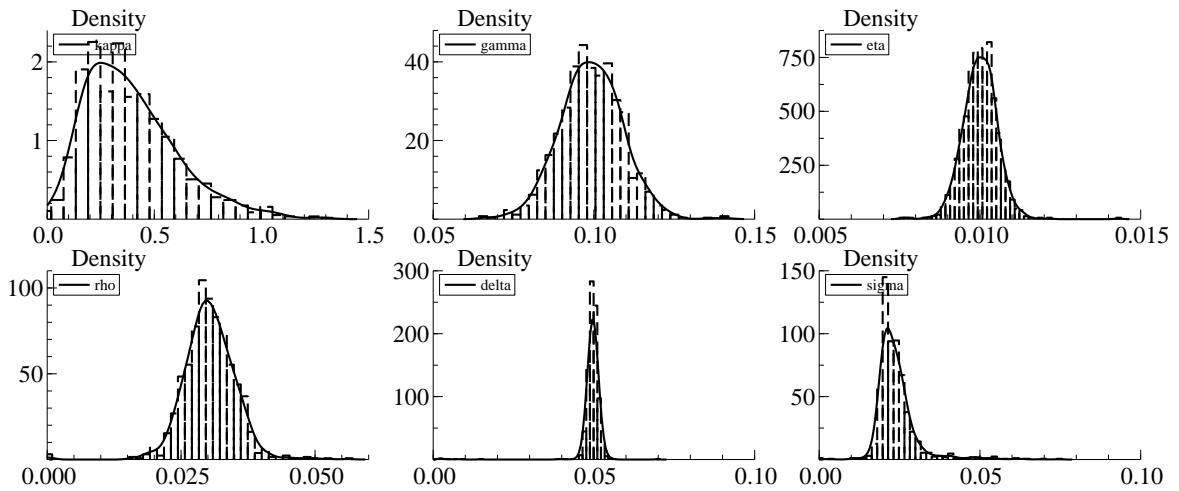
The table reports estimates for the structural model parameters estimated using the latent short rate and mixed-frequency MEF approaches for the AK-Vasicek model, SMEF (Latent Short Rate) and MF-MEF, respectively. We run the estimation for monthly data (where production is measured by IP), quarterly data (production measured by GDP) and the mixed-frequency data (where production is quarterly data using GDP, and consumption monthly data). The sample runs from January, 1982 until December, 2012. Asymptotic t -statistics are given below the estimates.

Parameter Estimates from Empirical Data				
	Monthly Data	Quarterly Data		Mixed-Frequency
	MEF	MEF	SMEF	MF-MEF
κ	0.060 0.93	0.068 0.515	0.126 3.2	0.023 0.137
γ	0.051 0.479	0.131 0.483	0.058 55.8	0.030 0.937
η	0.006 0.639	0.014 0.221	0.000 0.3	0.001 0.005
ρ	0.003 0.802	0.021 5.95	0.009 4.39	0.013 23.2
δ	0.025 0.162	0.062 0.144	0.024 12	0.008 0.042
σ	0.020 3.3	0.020 1.15	0.022 510	0.025 1.06

Figure 1: Simulation Study – Monthly and Quarterly Data

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. 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

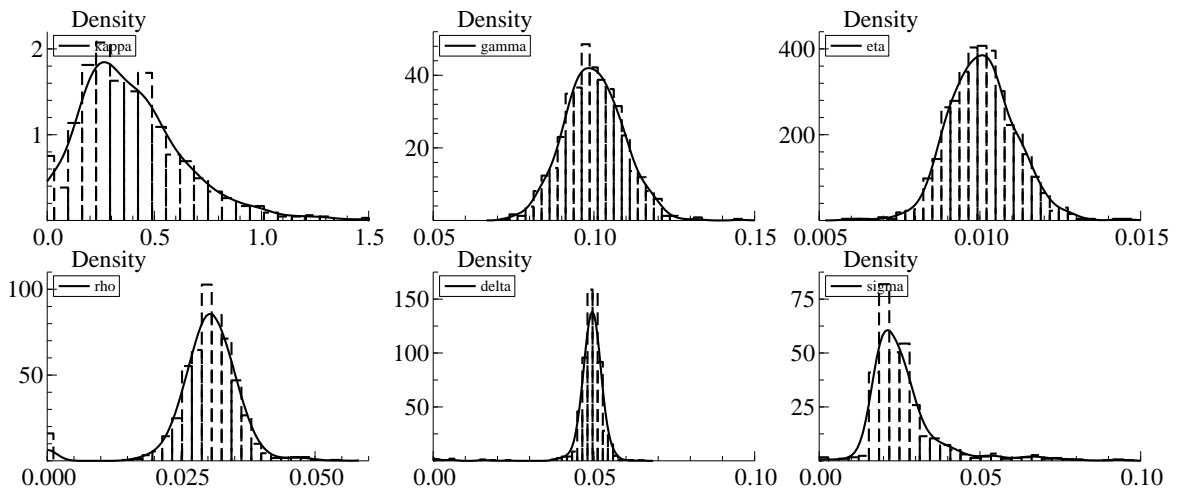
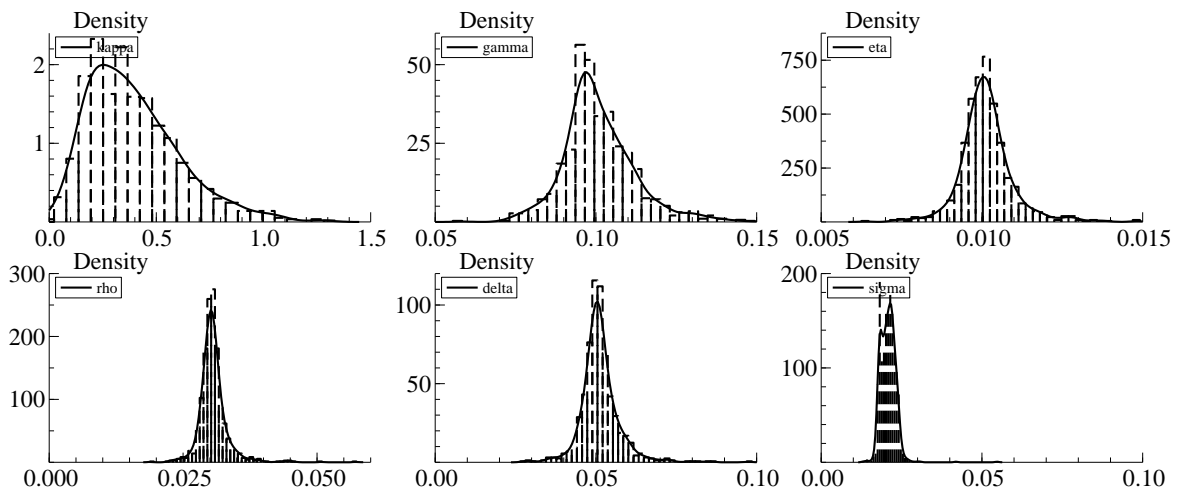


Figure 2: Simulation Study – Latent Short Rate and Mixed Frequency

The figure reports output of a simulation study of the accuracy of the structural model parameters estimated using simulated MEF and mixed-frequency approaches for the AK-Vasicek model, SMEF and MF-MEF, respectively. 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 estimates, in Panel A for the SMEF (Latent Short Rate) case based on monthly data and in Panel B for MF-MEF approach.

(A) SMEF (Latent Short Rate) for Monthly Data



(B) MF-MEF

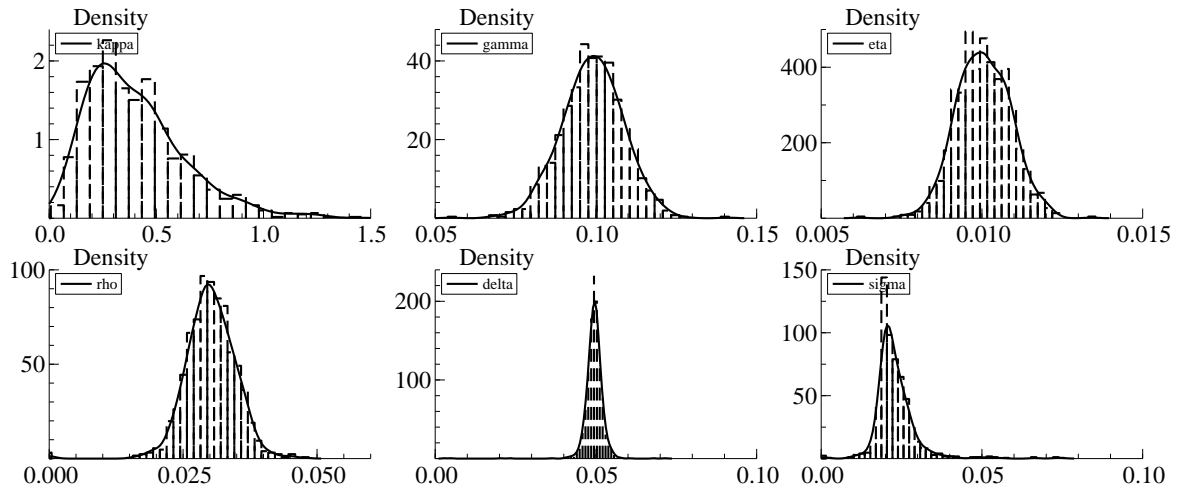
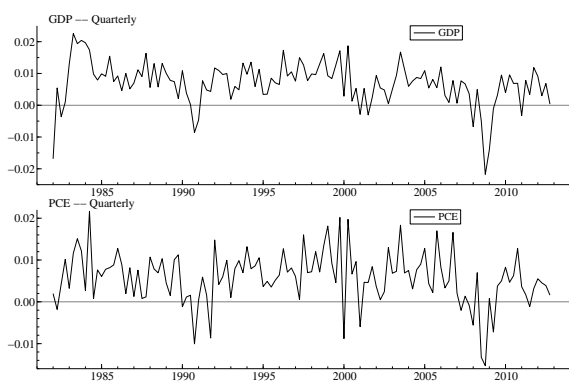


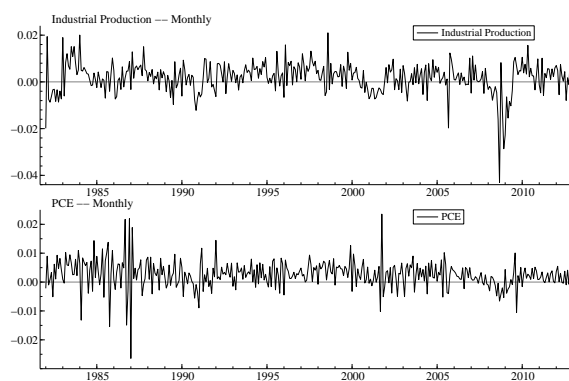
Figure 3: Overview of Quarterly, Monthly and Daily Variables

In this figure we show time series plots of the variables in our data set at the quarterly (Panel A), monthly (Panel B) and daily (Panel C) frequency. In Panel A, the top plot shows the growth rate of real Gross Domestic Product (GDP) and the bottom plot of real Personal Consumption Expenditure (PCE) both at the quarterly frequency. In Panel B, the top plot shows the growth rate of Industrial Production (IP) and the bottom plot of real PCE both at the monthly frequency. Panel C shows the nominal 3m interest rate series at the daily frequency. All series are obtained from the Federal Reserve Bank of St. Louis Economic Dataset (FRED). The sample runs from January, 1982 until December, 2012.

(A) Quarterly Macroeconomic Variables



(B) Monthly Macroeconomic Variables



(C) Daily Interest Rate

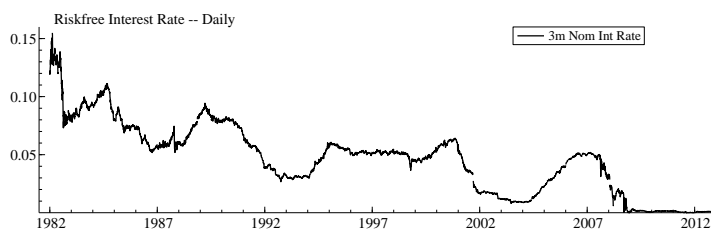
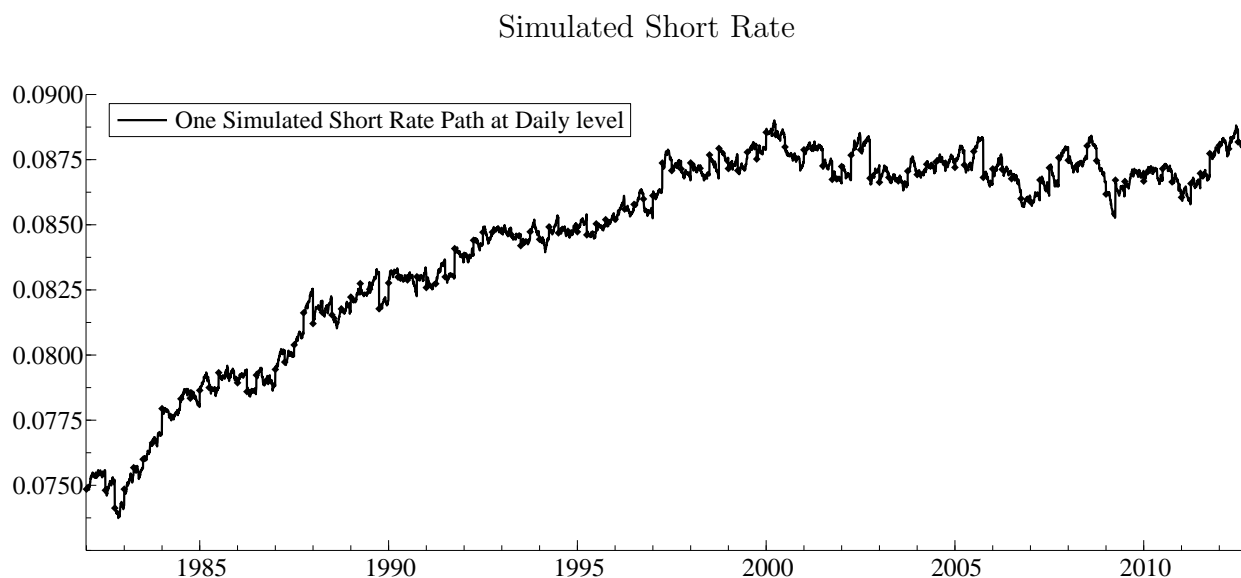


Figure 4: Latent Short Rate Estimate

In this figure we show the simulated latent short rate that is obtained from our latent short rate MEF approach for the AK-Vasicek model. We report the short rate estimate based on quarterly data. The line represents one path of the simulated short rate series, the dots the proxy $r_t^* = \rho Y_t / C_t$ for . The sample runs from January, 1982 until December, 2012.



Referee's appendix

Web appendix to “Estimating Dynamic Equilibrium Models using Mixed Frequency Macro and Financial Data” (by Bent Jesper Christensen, Olaf Posch and Michel van der Wel).

A 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.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

$$Y_t = A_t F(K_t, L), \quad (46)$$

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

$$A_{t+1} - A_t = \mu(A_t) + \eta(A_t)\epsilon_{A,t+1}, \quad \epsilon_A \sim N(0, 1), \quad (47)$$

with $\mu(A_t)$ and $\eta(A_t)$ generic drift and volatility functions.¹⁹ The capital stock increases if gross investment I_t exceeds capital depreciation,

$$K_{t+1} - K_t = I_t - \delta K_t + \sigma \epsilon_{K,t+1}, \quad \epsilon_K \sim N(0, 1), \quad (48)$$

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).²⁰

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

¹⁹We assume that $E(A_t) = A \in \mathbb{R}_+$ exists, and that the sum describing life-time utility in (51) below is bounded, so that the value function is well-defined.

²⁰It 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 $\epsilon_{K,t+1} \equiv Z_{t+1} - Z_t \sim N(0, 1)$.

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

$$Y_{t+1} = (A_t + \mu(A_t) + \eta(A_t)\epsilon_{A,t+1}) F(K_t + I_t - \delta K_t + \sigma\epsilon_{K,t+1}, L). \quad (49)$$

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

$$Y_{t+1} - Y_t = \mu(A_t)Y_A + (I_t - \delta K_t)Y_K + \frac{1}{2}Y_{KK}\sigma^2 + Y_A\eta(A_t)\epsilon_{A,t+1} + \sigma Y_K\epsilon_{K,t+1}. \quad (50)$$

Obviously, comparing both (49) and (50) 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

$$U_0 \equiv E_0 \sum_{t=0}^{\infty} \beta^t u(C_t, A_t) dt, \quad u_C > 0, \quad u_{CC} < 0, \quad (51)$$

subject to

$$K_{t+1} - K_t = (r_t - \delta)K_t + w_t L - C_t + \sigma\epsilon_{K,t+1}, \quad (52)$$

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.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

$$V(K_0, A_0) = \max_{\{C_t\}_{t=0}^{\infty}} U_0 \quad s.t. \quad (52) \quad \text{and} \quad (47), \quad (53)$$

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

$$V(K_s, A_s) = \max_{C_s} \{u(C_s, A_s) + \beta E_s [V(K_{s+1}, A_{s+1})]\}. \quad (54)$$

Hence, the first-order condition for the problem is

$$u_C(C_t, A_t) = \beta E_t [V_K(K_{t+1}, A_{t+1})], \quad (55)$$

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 (55), 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 (54) 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) gives²¹

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

Leading the expression one period ahead and applying expectations yields

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

Inserting back into the first-order condition (55) we arrive at the Euler equation

$$u_C(C_t, A_t) = \beta E_t [(1 - \delta + r_{t+1})u_C(C_{t+1}, A_{t+1})], \quad (56)$$

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

$$\begin{aligned} u_C(C_{t+1}, A_{t+1}) &= (1 + \rho - (r_t - \delta))u_C(C_t, A_t) + u_{CC}(C_t, A_t)C_K\sigma\epsilon_{K,t+1} \\ &\quad + (u_{CC}(C_t, A_t)C_A\eta(A_t) + u_{CA}(C_t, A_t)\eta(A_t))\epsilon_{A,t+1}, \end{aligned} \quad (57)$$

Again, the continuous-time formulation may help to obtain a dynamic formulation which can be used for estimation of the structural parameters. For example, (57) could be used to put structure on the residuals in a regression-based estimation approach.

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

²¹If the stochastic depreciation depends on the predetermined capital stock in (52), it is not possible to fully replace the costate variable by known functions. The corresponding expression would be

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

A.3 The equilibrium dynamics

Our equilibrium dynamics of the economy can be summarized as

$$u'(C_t) = \beta E_t [(1 - \delta + r_{t+1})u'(C_{t+1})] \quad (58a)$$

$$Y_{t+1} = (A_t + \mu(A_t) + \eta(A_t)\epsilon_{A,t+1}) F(K_t + I_t - \delta K_t + \sigma\epsilon_{K,t+1}, L) \quad (58b)$$

$$K_{t+1} = (1 + r_t - \delta)K_t + w_t L - C_t + \sigma\epsilon_{K,t+1} \quad (58c)$$

$$A_{t+1} = A_t + \mu(A_t) + \eta(A_t)\epsilon_{A,t+1} \quad (58d)$$

Provided that variables C_t , Y_t , K_t and also A_t are observed, the econometrician needs to consider the system (58) 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

$$dC_t = \frac{u'(C_t)}{u''(C_t)}(\rho - (r_t - \delta))dt - \frac{1}{2}(C_A^2\eta(A_t)^2 + C_K^2\sigma^2)\frac{u'''(C_t)}{u''(C_t)}dt + C_A\eta(A_t)dB_t + C_K\sigma dZ_t \quad (59a)$$

$$dY_t = (\mu(A_t)Y_A + (I_t - \delta K_t)Y_K + \frac{1}{2}Y_{KK}\sigma^2)dt + Y_A\eta(A_t)dB_t + \sigma Y_K dZ_t \quad (59b)$$

$$dK_t = ((r_t - \delta)K_t + w_t L - C_t)dt + \sigma dZ_t \quad (59c)$$

$$dA_t = \mu(A_t)dt + \eta(A_t)dB_t \quad (59d)$$

Provided that C_t , Y_t , K_t and also A_t are observed, the econometrician needs to consider the system (59) 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.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 (58) reduces to,

$$C_t^{-\theta} = \beta E_t [(1 - \delta + r_{t+1})C_{t+1}^{-\theta}] \quad (60a)$$

$$\begin{aligned} Y_{t+1} &= (1 + \mu(r_t)/r_t + \eta(r_t)/r_t\epsilon_{A,t+1}) (Y_t + (r_t - \delta)Y_t - r_t C_t + r_t\sigma\epsilon_{K,t+1}) \\ &= Y_t + Y_t\mu(r_t)/r_t + (r_t - \delta)Y_t - r_t C_t + Y_t\eta(r_t)/r_t\epsilon_{A,t+1} + r_t\sigma\epsilon_{K,t+1} \\ &\quad + ((r_t - \delta)Y_t - r_t C_t + r_t\sigma\epsilon_{K,t+1}) (\mu(r_t)/r_t + \eta(r_t)/r_t\epsilon_{A,t+1}) \end{aligned} \quad (60b)$$

$$r_{t+1} = r_t + \mu(r_t) + \eta(r_t)\epsilon_{A,t+1} \quad (60c)$$

whereas system (59) reduces to

$$dC_t = (r_t - \delta - \rho)C_t/\theta dt + \frac{1}{2}(1 + \theta)(C_A^2\eta(A_t)^2 + C_K^2\sigma^2)/C_t dt + C_A\eta(A_t)dB_t + C_K\sigma dZ_t \quad (61a)$$

$$dY_t = (Y_t\mu(r_t)/r_t + (r - \delta)Y_t - r_t C_t)dt + Y_t\eta(r_t)/r_t dB_t + \sigma r_t dZ_t \quad (61b)$$

$$dr_t = \mu(r_t)dt + \eta(r_t)dB_t \quad (61c)$$

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

One way of proceeding is to use an Euler scheme (as in Wang, Phillips, and Yu, 2011) to discretize the system (61) for small time intervals (no approximation error in the limit). This scheme has the nice feature that the discrete-time econometric toolbox (i.e., either linear or nonlinear estimation methods following An and Schorfheide, 2007; Fernández-Villaverde, Rubio-Ramírez, Sargent, and Watson, 2007; Fernández-Villaverde and Rubio-Ramírez, 2007) can be applied. As explained in the main text, we do not follow this route. Instead we proceed by integrating the system of equations and/or use closed-form solutions, for example for the interest rate Vasicek specification.

B Appendix Tables

Table A1: Simulation Study – Sensitivity to DGP values

The table reports output of a simulation study into the sensitivity of the Table 1 monthly results for the OLS (Panel A), FGLS-SUR-IV (Panel B), GMM (Panel C) and MEF (Panel D) methods to the parameter settings used in the Data Generating Process (DGP). In each Panel, the top row reports the baseline DGP settings and the second row the estimates obtained for these settings (these are the estimates of Table 1). Then we vary 1 parameter at the time and consider two settings for each parameters, one value lower than the one used in the baseline settings and one value higher than those of the baseline DGP settings (while keeping all other parameters at the baseline settings). In all cases, 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.

Panel A: Parameter Estimates from Simulation Study – OLS Sensitivity to DGP values						
	κ	γ	η	ρ	δ	σ
Baseline DGP settings	0.200	0.100	0.010	0.030	0.050	0.020
OLS for Baseline DGP	0.349 0.286	0.201 0.036	0.083 0.035	0.080 0.015	0.050	0.317 0.040
DGP with $\kappa = 0.1$	0.272 0.251	0.200 0.055	0.070 0.038	0.079 0.019	0.050	0.313 0.058
DGP with $\kappa = 0.5$	0.628 0.348	0.201 0.018	0.112 0.033	0.081 0.009	0.050	0.319 0.024
DGP with $\gamma = 0.05$	0.325 0.283	0.087 0.036	0.033 0.023	0.049 0.014	0.050	0.193 0.063
DGP with $\gamma = 0.2$	0.354 0.287	0.412 0.067	0.174 0.070	0.135 0.033	0.050	0.460 0.067
DGP with $\eta = 0.005$	0.354 0.286	0.206 0.034	0.087 0.035	0.083 0.017	0.050	0.325 0.047
DGP with $\eta = 0.05$	0.175 0.310	0.054 0.221	0.002 0.040	0.032 0.009	0.050	0.000 0.094
DGP with $\rho = 0.01$	0.349 0.286	0.201 0.036	0.083 0.035	0.060 0.015	0.050	0.317 0.040
DGP with $\rho = 0.1$	0.349 0.286	0.201 0.036	0.083 0.035	0.150 0.015	0.050	0.317 0.040
DGP with $\delta = 0.01$	0.349 0.286	0.201 0.036	0.083 0.035	0.080 0.015	0.010	0.317 0.040
DGP with $\delta = 0.1$	0.349 0.286	0.201 0.036	0.083 0.035	0.080 0.015	0.100	0.317 0.040
DGP with $\sigma = 0.01$	0.348 0.284	0.201 0.034	0.083 0.035	0.080 0.012	0.050	0.317 0.037
DGP with $\sigma = 0.05$	0.351 0.289	0.200 0.045	0.083 0.038	0.080 0.026	0.050	0.319 0.061

Table A1, Panel B: Parameter Estimates from Simulation Study –
FGLS-SUR-IV Sensitivity to DGP values

	κ	γ	η	ρ	δ	σ
Baseline DGP settings	0.200	0.100	0.010	0.030	0.050	0.020
FGLS-SUR-IV for Baseline DGP	0.299 0.134	0.101 0.013	0.008 0.004	0.030 0.006	0.050	0.000 <0.001
DGP with $\kappa = 0.1$	0.263 0.146	0.101 0.022	0.009 0.004	0.030 0.006	0.050	0.000 <0.001
DGP with $\kappa = 0.5$	0.403 0.150	0.100 0.006	0.006 0.003	0.030 0.006	0.050	0.000 <0.001
DGP with $\gamma = 0.05$	0.404 0.214	0.051 0.012	0.008 0.003	0.030 0.006	0.050	0.000 <0.001
DGP with $\gamma = 0.2$	0.220 0.117	0.201 0.014	0.006 0.006	0.030 0.006	0.050	0.000 0.020
DGP with $\eta = 0.005$	0.220 0.117	0.100 0.007	0.002 0.005	0.030 0.006	0.050	0.012 0.020
DGP with $\eta = 0.05$	0.541 0.305	0.151 0.045	0.029 0.038	0.030 0.006	0.050	0.000 <0.001
DGP with $\rho = 0.01$	0.299 0.134	0.101 0.013	0.008 0.004	0.010 0.006	0.050	0.000 <0.001
DGP with $\rho = 0.1$	0.299 0.134	0.101 0.013	0.008 0.004	0.100 0.006	0.050	0.000 <0.001
DGP with $\delta = 0.01$	0.299 0.134	0.101 0.013	0.008 0.004	0.030 0.006	0.010	0.000 <0.001
DGP with $\delta = 0.1$	0.299 0.134	0.101 0.013	0.008 0.004	0.030 0.006	0.100	0.000 <0.001
DGP with $\sigma = 0.01$	0.298 0.134	0.101 0.013	0.009 0.003	0.030 0.003	0.050	0.000 <0.001
DGP with $\sigma = 0.05$	0.301 0.136	0.100 0.013	0.000 0.007	0.030 0.014	0.050	0.037 0.020

Table A1, Panel C: Parameter Estimates from Simulation Study –
GMM Sensitivity to DGP values

	κ	γ	η	ρ	δ	σ
Baseline DGP settings	0.200	0.100	0.010	0.030	0.050	0.020
GMM for Baseline DGP	0.345 0.345	0.100 0.014	0.010 0.001	0.031 0.007	0.050	0.027 0.047
DGP with $\kappa = 0.1$	0.265 0.312	0.100 0.024	0.010 0.001	0.031 0.007	0.050	0.026 0.049
DGP with $\kappa = 0.5$	0.594 0.449	0.101 0.006	0.010 0.001	0.031 0.007	0.050	0.025 0.044
DGP with $\gamma = 0.05$	0.335 0.327	0.053 0.013	0.010 0.001	0.031 0.007	0.050	0.033 0.062
DGP with $\gamma = 0.2$	0.290 0.312	0.200 0.015	0.010 0.001	0.031 0.007	0.050	0.033 0.048
DGP with $\eta = 0.005$	0.330 0.319	0.100 0.007	0.005 <0.001	0.030 0.006	0.050	0.022 0.038
DGP with $\eta = 0.05$	0.352 0.351	0.197 0.554	0.050 0.050	0.042 0.265	0.050	0.157 0.678
DGP with $\rho = 0.01$	0.344 0.348	0.100 0.014	0.010 0.001	0.011 0.007	0.050	0.027 0.047
DGP with $\rho = 0.1$	0.345 0.345	0.100 0.014	0.010 0.001	0.101 0.007	0.050	0.027 0.047
DGP with $\delta = 0.01$	0.344 0.345	0.100 0.014	0.010 0.001	0.031 0.007	0.010	0.027 0.047
DGP with $\delta = 0.1$	0.344 0.346	0.100 0.014	0.010 0.001	0.031 0.007	0.100	0.027 0.047
DGP with $\sigma = 0.01$	0.352 0.361	0.101 0.014	0.010 0.001	0.031 0.003	0.050	0.020 0.044
DGP with $\sigma = 0.05$	0.344 0.325	0.100 0.015	0.010 0.001	0.031 0.017	0.050	0.052 0.032

Table A1, Panel D: Parameter Estimates from Simulation Study –
MEF Sensitivity to DGP values

	κ	γ	η	ρ	δ	σ
Baseline DGP settings	0.200	0.100	0.010	0.030	0.050	0.020
MEF for Baseline DGP	0.354 0.284	0.099 0.013	0.010 0.001	0.030 0.006	0.050 0.002	0.023 0.005
DGP with $\kappa = 0.1$	0.212 0.239	0.099 0.026	0.010 0.001	0.030 0.006	0.050 0.002	0.020 0.003
DGP with $\kappa = 0.5$	0.624 0.350	0.100 0.005	0.010 0.001	0.030 0.006	0.050 0.001	0.021 0.003
DGP with $\gamma = 0.05$	0.393 0.324	0.050 0.014	0.010 0.001	0.030 0.006	0.050 0.002	0.021 0.006
DGP with $\gamma = 0.2$	0.356 0.282	0.199 0.013	0.010 0.001	0.030 0.006	0.050 0.002	0.021 0.004
DGP with $\eta = 0.005$	0.351 0.286	0.100 0.006	0.005 0.000	0.030 0.006	0.050 0.001	0.021 0.003
DGP with $\eta = 0.05$	0.578 0.723	0.143 0.049	0.049 0.006	0.030 0.007	0.051 0.037	0.023 0.130
DGP with $\rho = 0.01$	0.355 0.283	0.099 0.013	0.010 0.001	0.010 0.006	0.050 0.002	0.022 0.004
DGP with $\rho = 0.1$	0.356 0.284	0.099 0.013	0.010 0.001	0.100 0.006	0.050 0.002	0.019 0.001
DGP with $\delta = 0.01$	0.356 0.282	0.099 0.013	0.010 0.001	0.030 0.005	0.010 0.002	0.023 0.005
DGP with $\delta = 0.1$	0.357 0.288	0.099 0.013	0.010 0.001	0.030 0.006	0.100 0.002	0.020 0.004
DGP with $\sigma = 0.01$	0.355 0.294	0.099 0.013	0.010 0.001	0.030 0.003	0.050 0.002	0.011 0.001
DGP with $\sigma = 0.05$	0.356 0.282	0.099 0.013	0.010 0.001	0.030 0.014	0.049 0.002	0.054 0.010

Table A2: Simulation Study – Variance Terms and 5 Moment Conditions

The table reports output of a simulation study into the incorporation of additional moments for the OLS, FGLS-SUR-IV, GMM and MEF approaches for the AK-Vasicek model. 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 & Quarterly Data									
		Monthly Data				Quarterly Data			
	DGP	OLS	FGLS-SUR-IV	GMM	MEF	OLS	FGLS-SUR-IV	GMM	MEF
κ	0.200	0.168 0.141	0.299 0.134	0.311 0.348	0.285 0.425	0.171 0.155	0.227 0.119	0.177 0.265	0.244 0.422
γ	0.100	0.100 0.015	0.100 0.013	0.102 0.014	0.100 0.015	0.100 0.015	0.101 0.013	0.110 0.059	0.100 0.019
η	0.010	0.010 0.001	0.009 0.001	0.010 0.001	0.010 0.001	0.010 0.001	0.010 0.001	0.010 0.001	0.010 0.001
ρ	0.030	0.030 0.006	0.030 0.006	0.030 0.006	0.030 0.006	0.030 0.006	0.030 0.006	0.030 0.006	0.030 0.006
δ	0.050	0.050 <0.001	0.050 0.001	0.051 0.002	0.050 0.002	0.050 <0.001	0.051 0.001	0.054 0.067	0.050 0.004
σ	0.020	0.020 0.001	0.020 0.001	0.020 0.001	0.020 0.001	0.020 0.002	0.020 0.002	0.019 0.002	0.020 0.002

Table A3: Simulation Study – Sensitivity of Regression-Based Methods to δ_0 and σ_0

The table reports output of a simulation study into the sensitivity of the Table 1 monthly results for the regression-based OLS and FGLS-SUR-IV methods to the δ_0 and σ_0 settings. For δ_0 we consider 0.01, 0.05 (base) and 0.10, for σ_0 we consider 0.01, 0.02 (base) and 0.05. We set the restricted value for δ equal to δ_0 for internal consistency. Panel A reports the performance of the OLS method, Panel B of the FGLS-SUR-IV method. Each Panel consists of 9 columns, where each column represents a δ_0 and σ_0 combination. 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.

Panel A: Parameter Estimates from Simulation Study – OLS Sensitivity to δ_0 and σ_0										
	DGP	$\delta_0 = 0.01$			$\delta_0 = 0.05$			$\delta_0 = 0.10$		
		$\sigma_0 = 0.01$	$\sigma_0 = 0.02$	$\sigma_0 = 0.05$	$\sigma_0 = 0.01$	$\sigma_0 = 0.02$	$\sigma_0 = 0.05$	$\sigma_0 = 0.01$	$\sigma_0 = 0.02$	$\sigma_0 = 0.05$
κ	0.200	0.292 0.281	0.295 0.282	0.305 0.286	0.349 0.286	0.349 0.286	0.350 0.286	0.354 0.281	0.354 0.281	0.354 0.282
γ	0.100	0.108 0.034	0.109 0.034	0.114 0.033	0.201 0.036	0.201 0.036	0.206 0.036	0.312 0.053	0.313 0.054	0.317 0.054
η	0.010	0.039 0.026	0.040 0.026	0.042 0.026	0.083 0.035	0.083 0.035	0.085 0.036	0.131 0.051	0.131 0.051	0.133 0.052
ρ	0.030	0.054 0.012	0.055 0.012	0.057 0.012	0.080 0.015	0.080 0.015	0.083 0.015	0.111 0.024	0.112 0.024	0.114 0.025
δ	0.050	0.010	0.010	0.010	0.050	0.050	0.050	0.100	0.100	0.100
σ	0.020	0.222 0.051	0.224 0.050	0.234 0.047	0.316 0.041	0.317 0.040	0.324 0.040	0.401 0.057	0.402 0.057	0.408 0.058

Panel B: Parameter Estimates from Simulation Study – FGLS-SUR-IV Sensitivity to δ_0 and σ_0										
	DGP	$\delta_0 = 0.01$			$\delta_0 = 0.05$			$\delta_0 = 0.10$		
		$\sigma_0 = 0.01$	$\sigma_0 = 0.02$	$\sigma_0 = 0.05$	$\sigma_0 = 0.01$	$\sigma_0 = 0.02$	$\sigma_0 = 0.05$	$\sigma_0 = 0.01$	$\sigma_0 = 0.02$	$\sigma_0 = 0.05$
κ	0.200	0.196 0.210	0.201 0.211	0.223 0.201	0.299 0.134	0.299 0.134	0.297 0.134	0.109 0.058	0.108 0.058	0.104 0.058
γ	0.100	0.067 0.041	0.068 0.039	0.071 0.031	0.101 0.013	0.101 0.013	0.102 0.013	0.161 0.019	0.161 0.019	0.163 0.020
η	0.010	0.018 0.026	0.019 0.024	0.021 0.020	0.009 0.003	0.008 0.004	0.000 0.004	0.000 <0.001	0.000 <0.001	0.000 <0.001
ρ	0.030	0.038 0.008	0.038 0.008	0.039 0.008	0.030 0.006	0.030 0.006	0.031 0.006	0.032 0.006	0.032 0.006	0.033 0.007
δ	0.050	0.010	0.010	0.010	0.050	0.050	0.050	0.100	0.100	0.100
σ	0.020	0.132 0.033	0.133 0.032	0.140 0.028	0.000 <0.001	0.000 <0.001	0.036 0.018	0.058 0.042	0.061 0.040	0.077 0.033

Table A4: Estimates – Variance Terms and 5 Moment Conditions

The table reports estimates for the structural model parameters estimated using OLS, FGLS-SUR-IV, GMM, and MEF approaches for the AK-Vasicek model. For OLS and FGLS-SUR-IV we use the variance terms for the consumption and interest rate equation, for GMM and MEF we use 5 moment conditions. We run the estimation for monthly data (where production is measured by IP) and quarterly data (production measured by GDP). The sample runs from January, 1982 until December, 2012. Asymptotic t -statistics are given below the estimates.

Parameter Estimates from Empirical Data								
	Monthly Data				Quarterly Data			
	OLS	FGLS-SUR-IV	GMM	MEF	OLS	FGLS-SUR-IV	GMM	MEF
κ	0.096 0.436	0.083 0.270	0.030 0.185	0.069 0.679	0.114 1.064	0.065 2.083	0.045 0.769	0.048 0.697
γ	0.101 4.002	0.101 1.671	0.045 0.186	0.108 0.602	0.134 2.715	0.130 2.329	0.089 2.054	0.098 0.924
η	0.018 1.284	0.018 0.444	0.005 0.669	0.007 0.051	0.028 0.693	0.019 0.608	0.000 <0.001	0.007 0.097
ρ	0.015 0.441	0.015 1.139	0.006 0.153	0.004 0.089	0.022 0.672	0.021 0.957	0.009 0.444	0.020 0.538
δ	0.098 1.298	0.106 1.443	0.050	0.081 0.243	0.128 0.732	0.153 0.682	0.050	0.040 0.173
σ	0.018 0.065	0.018 0.082	0.014 0.994	0.018 0.008	0.018 0.078	0.017 0.003	0.000 <0.001	0.019 0.010

Table A5: Simulation Study – Iterated MEF Approach

The table reports output of a simulation study of the accuracy of the structural model parameters estimated using the iterated MEF approaches for the AK-Vasicek model. For 100 replications, we generate 25 years of data from the underlying data generating process (DGP) and apply our estimation strategy. We report estimates using the MEF approach with both 3 and 5 moment conditions for the regular MEF and iterated approach for 2 iterations. We show the median estimate, and provide the interquartile range below it.

Parameter Estimates from Simulation Study – Iterated MEF 3 and 5 Moments										
	DGP	Monthly Data				Quarterly Data				
		3 Conditions		5 Conditions		3 Conditions		5 Conditions		
		MEF	two-step MEF	MEF	two-step MEF	MEF	two-step MEF	MEF	two-step MEF	MEF
κ	0.200	0.348 0.309	0.239 0.202	0.288 0.480	0.200 0.204	0.353 0.310	0.316 0.339	0.241 0.366	0.209 0.129	
γ	0.100	0.100 0.012	0.109 0.048	0.101 0.014	0.104 0.017	0.099 0.013	0.130 0.074	0.100 0.021	0.107 0.021	
η	0.010	0.010 0.001	0.001 0.002	0.010 0.001	0.010 0.000	0.010 0.002	0.000 0.001	0.010 0.001	0.010 0.001	
ρ	0.030	0.030 0.005	0.032 0.010	0.030 0.006	0.031 0.007	0.030 0.006	0.032 0.016	0.030 0.005	0.030 0.012	
δ	0.050	0.050 0.002	0.059 0.045	0.050 0.002	0.051 0.008	0.050 0.003	0.072 0.063	0.050 0.005	0.055 0.019	
σ	0.020	0.022 0.005	0.025 0.018	0.020 0.001	0.020 0.001	0.024 0.011	0.030 0.030	0.020 0.002	0.020 0.003	