Estimating Dynamic Equilibrium Models using Mixed Frequency Macro and Financial Data

Web Appendix

Bent Jesper Christensen[(a,b)], Olaf Posch[(c,b)], and Michel van der Wel[(b,d)]
[(a)]Aarhus University, [(b)]CREATES, [(c)] Universität Hamburg,
[(d)]Erasmus University Rotterdam

April 18, 2016

Contents

A Alternative estimation approaches

A.1 The regression-based approaches

In this section we propose regression-based procedures to obtain benchmark 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.

A.1.1 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

$$y_{j,t} = x_{j,t}\beta_j + \varepsilon_{j,t}, \qquad\qquad j = C, Y, r, \tag{A.1}$$

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 $y_{r,t} = r_t^f$.[1] 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 $x_{\hat{r},t} = (1, \hat{r}_{t-\Delta})$. The reduced-form or linear parameters, β_C, $\beta_Y = (\beta_{Y,1}, \beta_{Y,2}, \beta_{Y,3})^\top$, and $\beta_r = (\beta_{r,1}, \beta_{r,2})^\top$, are given in terms of the structural parameters $\phi = (\kappa, \gamma, \eta, \rho, \delta, \sigma)^\top$ as

$$\beta_C = -\left(\rho - \tfrac{1}{2}\sigma^2\right)\Delta, \tag{A.2a}$$

$$\begin{aligned} \beta_{Y,1} &= -\left(\kappa + \rho - \tfrac{1}{2}\sigma^2\right)\Delta, \\ \beta_{Y,2} &= \kappa\gamma, \\ \beta_{Y,3} &= -\tfrac{1}{2}\eta^2, \end{aligned} \tag{A.2b}$$

$$\begin{aligned} \beta_{r,1} &= (1 - e^{-\kappa\Delta})(\gamma - \delta - \sigma^2), \\ \beta_{r,2} &= e^{-\kappa\Delta}. \end{aligned} \tag{A.2c}$$

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

$$\varepsilon_{C,t} = \sigma(Z_t - Z_{t-\Delta}), \tag{A.3a}$$

$$\varepsilon_{Y,t} = \int_{t-\Delta}^{t} \eta/\hat{r}_v dB_v + \sigma(Z_t - Z_{t-\Delta}), \tag{A.3b}$$

$$\varepsilon_{r,t} = \eta e^{-\kappa\Delta} \int_{t-\Delta}^{t} e^{\kappa(v-(t-\Delta))} dB_v. \tag{A.3c}$$

[1] In cases where δ and σ in (17) 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 $i = 1, 2, \ldots$ Alternatively, a nonlinear one-step regression-based approach could be implemented.

Using iterated expectations and the properties of stochastic integrals, if the parameters are at their true values, including δ_0 and σ_0 in (17), then 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,

$$\hat{\beta}_j = (x_j^\top x_j)^{-1} x_j^\top y_j, \qquad j = C, Y, r, \tag{A.4}$$

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 (A.4) using the link (A.2a)-(A.2c) (or by an asymptotically equivalent restricted nonlinear 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.

A.1.2 Cross-equation correlation

The estimators (A.4) allow for different variances of the error terms $\varepsilon_{j,t}$, say, σ_j^2, $j = C, Y, r$, as they are implemented separately by equation. However, they do not exploit all other properties 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 (A.3a)-(A.3c), the term $\sigma(Z_t - Z_{t-\Delta})$ is common to both $\varepsilon_{C,t}$ and $\varepsilon_{Y,t}$, whereas both $\varepsilon_{Y,t}$ and $\varepsilon_{r,t}$ involve stochastic integrals with respect to B_v. Classical seemingly unrelated regressions (SUR) analysis is intended to exploit such cross-equation correlation in the errors to improve efficiency in estimation 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

$$\hat{\beta}_{SUR} = (x^\top \hat{V}^{-1} x)^{-1} x^\top \hat{V}^{-1} y, \tag{A.5}$$

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 A.2.

A.1.3 Endogeneity

The regression approaches (OLS and SUR) do not control for possible endogeneity of right-hand side variables in (15), and hence (A.1), 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 (A.4) of $\widehat{\beta}_Y$, fitted values from the first stage regressions replace $x_{Y,t,2}$ and $x_{Y,t,3}$. Third, fitted residuals are calculated using the new second stage estimate $\widehat{\beta}_Y$ but the original $x_{Y,t,2}$ and $x_{Y,t,3}$ (not their fitted values from the first stage), and these residuals form the basis of the 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 (A.5) 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.

Note that 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 (A.3b), 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 therefore present the MEF approach in the main text, exploiting the martingale structure of the model.

A.1.4 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 (A.2a)-(A.2c), 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 (A.1). 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^2 to identify all other parameters. Instead, without the need for such additional restrictions, it is possible to separate ρ, δ, and σ^2, and thus identify all six structural parameters, by exploiting the functional form of the error variances (the variances of (A.3a)-(A.3c)). Indeed, including the variance of the residual (A.3a) from the consumption equation as a separate moment along with the relations (A.2a)-(A.2c) clearly identifies σ^2 and thereby the full parameter vector ϕ, i.e., all six structural parameters.

Why should we rely on the first moment conditions and thus the regression coefficients, only, if they do not identify all structural parameters? In models with, say, stochastic volatility or more elaborate preference specification, the error term of the consumption equation becomes intractable (like the residual of the output equation). In such a case, the econometrician may exploit the martingale property only, without considering second moment conditions - namely, the form of the error variances and covariances. Because we want to keep our analysis applicable to such specifications, we focus on 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 this 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 either of three different unrestricted parameter sets ω_i, $i = 1, 2, 3$, from the reduced-form regressions: (1) the estimates of β in (A.2), i.e., the theoretical and empirical moments to match with respect to choice of ϕ are $\omega_1(\phi) = \beta(\phi)$ and $\hat{\omega}_1 = \hat{\beta}$ (this is the first moments or regression coefficients only case); (2) β along with the variance $\sigma^2\Delta$ of the consumption equation residual in (A.3a), 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 (A.3a) and (A.3c), $\omega_3(\phi) = (\omega_2(\phi)^\top, \frac{1}{2}\eta^2(1 - e^{-2\kappa\Delta})/\kappa)^\top$ and $\hat{\omega}_3 = (\hat{\omega}_2^\top, \hat{\Sigma}_{rr})^\top$. In each of the three cases, we solve the problem

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

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

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

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

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 (A.4), then the correct

$\hat{\Omega}_1 = \hat{V}_{OLS}$ is given in Appendix A.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.

In case (1), using first moment conditions and thus β, only, to set up the minimum distance problem, estimators that are asymptotically equivalent to the resulting minimum distance estimators are alternatively obtained by restricted (nonlinear) regression, minimizing the OLS respectively the SUR objective function with respect to ϕ under the relevant restrictions (A.2a)-(A.2c) 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 $\varepsilon_{j,t}$. In cases (2) and (3), when estimated residual variances are used along with the relations (A.2a)-(A.2c) 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 modified SUR objective function, say, $SUR^*(\phi) = \sum_{t=1}^T \varepsilon(\phi)_t^\top \Sigma(\phi)^{-1} \varepsilon(\phi)_t$, or, even better, $T \log |\Sigma(\phi)| + SUR^*(\phi)$. This use of (minus twice) the Gaussian log-likelihood function amounts to quasi maximum likelihood (QML) since clearly $\varepsilon_{Y,t}$ in (A.3b) is non-Gaussian.

A.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 (A.4) 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.

A.3 Transition probability matrix

This section derives the transition probability matrix for the continuous-time Markov chain of the regime-switching model (cf. Section 3.3.2).

Consider the following question that we use in our estimation approach: If the volatility is high at time $s \leq t$, then what is the probability that volatility is high at time t?

Following Ross (2014, p.371), let $P_{ij}(t) \equiv P(\eta_t = \eta_j \mid \eta_s = \eta_i)$ for $s \leq t$ denote the probability that a process presently in state i will be in state j at time t, and ϕ_{ij} the instantaneous transition rates, when in state i, at which the process makes a transition into state j. We shall derive the desired probability, namely $P_{hh}(t)$ by solving

$$\begin{aligned} \dot{P}_{hh}(t) &= \phi_{hl}\left[P_{lh}(t) - P_{hh}(t)\right], \\ \dot{P}_{lh}(t) &= \phi_{lh}\left[P_{hh}(t) - P_{lh}(t)\right], \end{aligned}$$

$\dot{P}_{ij}(t) \equiv \lim_{h\to 0}[P_{ij}(t+h) - P_{ij}(t)]/h$ for all $i, j \in \Theta$ with initial conditions $P_{hh}(s) = 1$ and $P_{lh}(s) = 0$. The solution to this system of ODEs is given by:

$$P_{hh}(t) = \frac{\phi_{lh}}{\phi_{hl} + \phi_{lh}} + \frac{\phi_{hl}}{\phi_{hl} + \phi_{lh}} e^{-(\phi_{hl}+\phi_{lh})(t-s)}, \tag{A.6}$$

$$P_{lh}(t) = \frac{\phi_{lh}}{\phi_{hl} + \phi_{lh}} - \frac{\phi_{lh}}{\phi_{hl} + \phi_{lh}} e^{-(\phi_{hl}+\phi_{lh})(t-s)}. \tag{A.7}$$

Hence, the transition probability matrix of the continuous-time Markov chain for $s \leq t$ is

$$P(t) = \begin{bmatrix} P_{ll}(t) & P_{lh}(t) \\ P_{hl}(t) & P_{hh}(t) \end{bmatrix}, \tag{A.8}$$

in which $P_{ll}(t) = 1 - P_{lh}(t)$ and $P_{hl}(t) = 1 - P_{hh}(t)$.

If we let $P_h(s)$ denote the unconditional probability of being in state θ_h at time s, the unconditional probability of being in the same state at time $t > s$ is then

$$P_h(t) = P_h(s)P_{hh}(t) + (1 - P_h(s))\, P_{lh}(t).$$

In the limit as $t \to \infty$ the unconditional probability of being in the high regime is

$$\lim_{t\to\infty} P_h(t) = \frac{\phi_{lh}}{\phi_{hl} + \phi_{lh}}.$$

A similar procedure yields the unconditional probability of being in the low regime as

$$\lim_{t\to\infty} P_l(t) = \frac{\phi_{hl}}{\phi_{hl} + \phi_{lh}}.$$

Conversely, from (A.6) and (A.7) we can back out the instantaneous transition rates of the Poisson processes, ϕ_{hl} and ϕ_{lh}, from any given transition probability matrix.

B Comparison to the discrete-time model

We now develop the model in discrete-time formulation in order to compare both approaches. Before we start it is important to note that all state variables are directly comparable, whereas the flow variables are expressed as periodic rates (instead of instantaneous rates).

B.1 The model

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

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

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

$$A_{t+1} - A_t = \tilde{\mu}(A_t) + \tilde{\eta}(A_t)\epsilon_{A,t+1}, \quad \epsilon_A \sim \mathcal{N}(0,1), \tag{B.2}$$

with $\mu(A_t)$ and $\eta(A_t)$ generic drift and volatility functions.[2] The capital stock increases if gross investment I_t exceeds capital depreciation,

$$K_{t+1} - K_t = I_t - \tilde{\delta}K_t + \tilde{\sigma}K_t\epsilon_{K,t+1}, \quad \epsilon_K \sim \mathcal{N}(0,1), \tag{B.3}$$

where $\tilde{\delta}$ is a deterministic rate of depreciation and $\tilde{\sigma}$ determines the variance of the stochastic depreciation.[3] Similar to the continuous-time version, the stochastic depreciation does depend on the level of the predetermined capital stock.

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

$$Y_{t+1} = (A_t + \tilde{\mu}(A_t) + \tilde{\eta}(A_t)\epsilon_{A,t+1}) \, F(K_t + I_t - \tilde{\delta}K_t + \tilde{\sigma}K_t\epsilon_{K,t+1}, L). \tag{B.4}$$

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

$$U_0 \equiv E_0 \sum_{t=0}^{\infty} \tilde{\beta}^t u(C_t, A_t)dt, \quad u_C > 0, \; u_{CC} < 0, \tag{B.5}$$

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

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

subject to

$$K_{t+1} - K_t = (\tilde{r}_t - \tilde{\delta})K_t + \tilde{w}_t L - C_t + \tilde{\sigma} K_t \epsilon_{K,t+1}, \tag{B.6}$$

where $\tilde{\beta}$ is the subjective discount factor, $\tilde{r}_t$ is the rental rate of capital, and $\tilde{w}_t$ is the labor wage rate. The paths of factor rewards are taken as given by the representative consumer.

B.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

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

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

$$V(K_s, A_s) = \max_{C_s} \left\{ u(C_s, A_s) + \tilde{\beta} E_s \left[V(K_{s+1}, A_{s+1}) \right] \right\}. \tag{B.8}$$

Hence, the first-order condition for the problem is

$$u_C(C_t, A_t) = \tilde{\beta} E_t \left[V_K(K_{t+1}, A_{t+1}) \right], \tag{B.9}$$

for any $t \in [0, \infty)$, and this allows us to write consumption as a function of the state variables, $C_t = C(K_t, A_t)$. Hence, 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 (B.8) requires to solve a stochastic difference equation in contrast to a deterministic differential equation.

Using the concentrated Bellman equation,

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

we obtain

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

Note that the second term is zero in equilibrium, because from the first-order condition

$$\begin{aligned} u_C(C_t, A_t)\tilde{\sigma}\epsilon_{K,t+1} &= \tilde{\beta} E_t \left[V_K(K_{t+1}, A_{t+1}) \right] \tilde{\sigma}\epsilon_{K,t+1} \\ \Leftrightarrow \; E_t \left[u_C(C_t, A_t)\tilde{\sigma}\epsilon_{K,t+1} \right] &= \tilde{\beta} E_t \left[E_t \left[V_K(K_{t+1}, A_{t+1}) \right] \tilde{\sigma}\epsilon_{K,t+1} \right] \\ \Leftrightarrow \; u_C(C_t, A_t)\tilde{\sigma} E_t \left[\epsilon_{K,t+1} \right] &= \tilde{\beta} E_t \left[V_K(K_{t+1}, A_{t+1})\tilde{\sigma}\epsilon_{K,t+1} \right] \\ \Leftrightarrow \; 0 &= \tilde{\beta} E_t \left[V_K(K_{t+1}, A_{t+1})\tilde{\sigma}\epsilon_{K,t+1} \right] \end{aligned}$$

Hence,

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

Leading the expression one period ahead and applying expectations yields

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

Inserting back into the first-order condition (B.9) we arrive at the Euler equation

$$u_C(C_t, A_t) = \tilde{\beta} E_t \left[(1 - \tilde{\delta} + \tilde{r}_{t+1}) u_C(C_{t+1}, A_{t+1}) \right], \tag{B.10}$$

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

B.3 Equilibrium dynamics

Our equilibrium dynamics of the economy can be summarized as

$$\begin{aligned} u'(C_t) &= \tilde{\beta} E_t \left[(1 - \tilde{\delta} + \tilde{r}_{t+1}) u'(C_{t+1}) \right] && \text{(B.11a)} \\ Y_{t+1} &= (A_t + \tilde{\mu}(A_t) + \tilde{\eta}(A_t)\epsilon_{A,t+1})\, F(K_t + I_t - \tilde{\delta} K_t + \tilde{\sigma} K_t \epsilon_{K,t+1}, L) && \text{(B.11b)} \\ K_{t+1} &= (1 + \tilde{r}_t - \tilde{\delta}) K_t + \tilde{w}_t L - C_t + \tilde{\sigma} K_t \epsilon_{K,t+1} && \text{(B.11c)} \\ A_{t+1} &= A_t + \tilde{\mu}(A_t) + \tilde{\eta}(A_t)\epsilon_{A,t+1} && \text{(B.11d)} \end{aligned}$$

Provided that variables C_t, Y_t, K_t and also A_t are observed, the econometrician needs to consider the system (B.11) for statistical inference on the deep parameters.

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

$$\begin{aligned} dC_t &= \frac{u'(C_t)}{u''(C_t)}(\rho - (r_t - \delta))dt - \sigma^2 C_K K_t dt - \tfrac{1}{2}(C_A^2 \eta(A_t)^2 + C_K^2 \sigma^2 K_t^2)\frac{u'''(C_t)}{u''(C_t)}dt \\ &\quad + C_A \eta(A_t) dB_t + C_K \sigma K_t dZ_t && \text{(B.12a)} \\ dY_t &= Y_A dA_t + Y_K dK_t + \tfrac{1}{2} Y_{KK} \sigma^2 K_t^2 dt \\ &= (\mu(A_t) Y_A + (I_t - \delta K_t) Y_K + \tfrac{1}{2} Y_{KK} \sigma^2 K_t^2) dt + Y_A \eta(A_t) dB_t + \sigma Y_K K_t dZ_t && \text{(B.12b)} \\ dK_t &= (I_t - \delta K_t) dt + \sigma K_t dZ_t && \text{(B.12c)} \\ dA_t &= \mu(A_t) dt + \eta(A_t) dB_t && \text{(B.12d)} \end{aligned}$$

Provided that C_t, Y_t, K_t and also A_t are observed, the econometrician needs to consider the system (B.12) for statistical inference on the deep parameters.

In what follows, we assume that the capital stock K_t and A_t are latent variables, but we can obtain them from financial market data.

B.4 The AK-Vasicek model with logarithmic preferences

Consider an AK economy, $Y_t = A_t K_t$, which implies $\tilde{r}_t = A_t$ and $K_t = Y_t/\tilde{r}_t$, and assume that the consumer has logarithmic preferences, system (B.11) reduces to,

$$\begin{aligned} C_t^{-1} &= \tilde{\beta} E_t \left[(1 - \tilde{\delta} + \tilde{r}_{t+1}) C_{t+1}^{-1} \right] && \text{(B.13a)} \\ Y_{t+1} &= (\tilde{r}_t + \tilde{\mu}(\tilde{r}_t) + \tilde{\eta}(\tilde{r}_t)\epsilon_{A,t+1}) \left((1 + \tilde{r}_t - \tilde{\delta})(Y_t/\tilde{r}_t) - C_t + \tilde{\sigma} Y_t/\tilde{r}_t \epsilon_{K,t+1} \right) && \text{(B.13b)} \\ \tilde{r}_{t+1} &= \tilde{r}_t + \tilde{\mu}(\tilde{r}_t) + \tilde{\eta}(\tilde{r}_t)\epsilon_{A,t+1} && \text{(B.13c)} \end{aligned}$$

whereas system (B.12) reduces to

$$\begin{aligned} dC_t &= (r_t - \delta - \rho) C_t dt - \sigma^2 C_K K_t dt - (C_A^2 \eta(A_t)^2 + C_K^2 \sigma^2 K_t^2)/C_t dt \\ &\quad + C_A \eta(A_t) dB_t + C_K \sigma Y_t/r_t dZ_t && \text{(B.14a)} \\ dY_t &= (\mu(r_t) Y_t/r_t + (r_t - \delta) Y_t - r_t C_t) dt + \eta(r_t) Y_t/r_t dB_t + \sigma Y_t dZ_t && \text{(B.14b)} \\ dr_t &= \mu(r_t) dt + \eta(r_t) dB_t && \text{(B.14c)} \end{aligned}$$

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

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. The corresponding Vasicek mean-reversion model at quarterly frequency reads $\tilde{\mu}(\tilde{r}_t) = \tilde{\kappa}(\tilde{\gamma} - \tilde{r}_t)$ and $\tilde{\eta}(\tilde{r}_t) = \tilde{\eta}$ where we define

$$\tilde{\gamma} \equiv \Delta\gamma, \quad \tilde{\kappa} \equiv 1 - e^{-\Delta\kappa}, \quad \tilde{\eta} \equiv \Delta\eta \sqrt{(1 - e^{-2\kappa\Delta})/(2\kappa)} \qquad \text{(B.15)}$$

In this case, the equilibrium dynamics are

$$\begin{aligned} C_t^{-1} &= \tilde{\beta} E_t \left[(1 - \tilde{\delta} + \tilde{r}_{t+1}) C_{t+1}^{-1} \right] && \text{(B.16a)} \\ Y_{t+1} &= Y_t + (\tilde{r}_t - \tilde{\delta}) Y_t - \tilde{r}_t C_t + \tilde{\kappa}(\tilde{\gamma} - \tilde{r}_t) Y_t/\tilde{r}_t + \tilde{\eta} Y_t/\tilde{r}_t \epsilon_{A,t+1} + \tilde{\sigma} Y_t \epsilon_{K,t+1} \\ &\quad + ((\tilde{r}_t - \tilde{\delta}) Y_t/\tilde{r}_t - C_t + \tilde{\sigma} Y_t/\tilde{r}_t \epsilon_{K,t+1}) \left(\tilde{\kappa}(\tilde{\gamma} - \tilde{r}_t) + \tilde{\eta}\epsilon_{A,t+1} \right) && \text{(B.16b)} \\ \tilde{r}_{t+1} &= \tilde{r}_t + \tilde{\kappa}(\tilde{\gamma} - \tilde{r}_t) + \tilde{\eta}\epsilon_{A,t+1} && \text{(B.16c)} \end{aligned}$$

whereas system (B.14) reads

$$\begin{aligned} dC_t &= (r_t - \delta - \rho) C_t dt - \sigma^2 C_K K_t dt - (C_A^2 \eta^2 + C_K^2 \sigma^2 K_t^2)/C_t dt \\ &\quad + C_A \eta dB_t + C_K \sigma Y_t/r_t dZ_t && \text{(B.17a)} \\ dY_t &= ((\kappa\gamma/r_t - \kappa + r_t - \delta) Y_t - r_t C_t) dt + \eta Y_t/r_t dB_t + \sigma Y_t dZ_t && \text{(B.17b)} \\ dr_t &= \kappa(\gamma - r_t) dt + \eta dB_t && \text{(B.17c)} \end{aligned}$$

Before we estimate the structural parameters, we need to solve the two models. This is complicated by the fact that both models are highly nonlinear. Note that the AK-Vasicek model with logarithmic preferences in continuous time has an explicit analytical solution of the nonlinear system, whereas the discrete-time analogue can only be solved numerically. We follow a log-linear approximation of the discrete-time first-order conditions below and use log-linear representation of the equilibrium dynamics for estimation.

One simple way of proceeding with the continuous-time system in order to match it to the discrete-time nature of the data is to use an Euler scheme (as in Wang, Phillips, and Yu, 2011) to discretize the system (B.14) 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 and thus seems quite attractive. 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. This allows us to easily handle different frequencies for the estimation of structural parameters.

B.5 Log-linear approximation, discrete-time AK-Vasicek model

There are many ways to solve the discrete-time model numerically. The best practice is to solve the model through a log-linear approximation to the set of first-order conditions. For this we define auxiliary variables (which turn out to be stationary),

$$\hat{C}_t \equiv \frac{C_t}{K_t}, \quad 1 + \nu_{t+1} \equiv \frac{K_{t+1}}{K_t}$$

and which can be used to transform the Euler equation (B.16a) into

$$\begin{aligned} 1 &= \tilde{\beta} E_t \left[(1 - \tilde{\delta} + \tilde{r}_{t+1}) \frac{C_t}{C_{t+1}} \frac{K_{t+1}}{K_{t+1}} \frac{K_t}{K_t} \right] \\ &= \tilde{\beta} E_t \left[\frac{\hat{C}_t}{\hat{C}_{t+1}} \frac{1 - \tilde{\delta} + \tilde{r}_{t+1}}{1 + \nu_{t+1}} \right] \end{aligned} \tag{B.18}$$

We may write the aggregate resource constraint (B.16b) as

$$\begin{aligned} Y_{t+1} &= \tilde{r}_{t+1}(K_t + Y_t - C_t - \tilde{\delta} K_t + \tilde{\sigma} K_t \epsilon_{K,t+1}) \\ \Leftrightarrow \frac{K_{t+1}}{K_t} &= 1 + \frac{Y_t}{K_t} - \frac{C_t}{K_t} - \tilde{\delta} + \tilde{\sigma} \epsilon_{K,t+1} \end{aligned}$$

or

$$\nu_{t+1} = \tilde{r}_t - \hat{C}_t - \tilde{\delta} + \tilde{\sigma} \epsilon_{K,t+1} \tag{B.19}$$

As a reference level, with $\tilde{r}_t \equiv \tilde{\gamma}$ for all t, the non-stochastic steady-state value is given from the Euler equation, which implies steady-state value for the consumption-capital ratio

$$1 + \nu = \tilde{\beta}(1 - \tilde{\delta} + \tilde{\gamma}) \quad \Rightarrow \quad \hat{C} = (1 - \tilde{\beta})(1 - \tilde{\delta} + \tilde{\gamma}) \tag{B.20}$$

First, we rewrite (B.18) as

$$\frac{1}{\tilde{\beta}} = \frac{\hat{C}_t}{\hat{C}_{t+1}} \frac{1-\tilde{\delta}+\tilde{r}_{t+1}}{1-\tilde{\delta}+\tilde{r}_t-\hat{C}_t+\tilde{\sigma}\epsilon_{K,t+1}} - v_{t+1} \;\equiv\; G(\ln \hat{C}_{t+1}, \ln \hat{C}_t, \tilde{r}_{t+1}, \tilde{r}_t, \epsilon_{K,t+1}, v_{t+1})$$

in which we defined the expectations error

$$v_{t+1} \equiv \frac{\hat{C}_t}{\hat{C}_{t+1}} \frac{1-\tilde{\delta}+\tilde{r}_{t+1}}{1+\nu_{t+1}} - E_t\left[\frac{\hat{C}_t}{\hat{C}_{t+1}} \frac{1-\tilde{\delta}+\tilde{r}_{t+1}}{1+\nu_{t+1}}\right], \quad E_t(v_{t+1}=0)$$

Second, we log-linearize the equation about the non-stochastic steady-state values

$$\begin{aligned} 0 \;\simeq\; & -\frac{1}{\tilde{\beta}}(\ln \hat{C}_{t+1} - \ln \hat{C}) + \frac{1}{\tilde{\beta}^2}(\ln \hat{C}_t - \ln \hat{C}) + \frac{1-\tilde{\beta}}{\tilde{\beta}\hat{C}}(\tilde{r}_{t+1}-\tilde{\gamma}) - \frac{1-\tilde{\beta}}{\tilde{\beta}^2\hat{C}}(\tilde{r}_t-\tilde{\gamma}) \\ & -\frac{\tilde{\sigma}(1-\tilde{\beta})}{\tilde{\beta}^2\hat{C}}\epsilon_{K,t+1} - v_{t+1} \end{aligned}$$

or

$$\begin{aligned} \ln \hat{C}_{t+1} - \ln \hat{C} - \frac{1-\tilde{\beta}}{\hat{C}}(\tilde{r}_{t+1}-\tilde{\gamma}) \;\simeq\; & \frac{1}{\tilde{\beta}}(\ln \hat{C}_t - \ln \hat{C}) - \frac{1-\tilde{\beta}}{\tilde{\beta}\hat{C}}(\tilde{r}_t-\tilde{\gamma}) \\ & -\frac{\tilde{\sigma}(1-\tilde{\beta})}{\tilde{\beta}\hat{C}}\epsilon_{K,t+1} - v_{t+1} \end{aligned}$$

where we used

$$\begin{aligned} \left.\frac{\partial G}{\partial \ln \hat{C}_{t+1}}\right|_{ss} &= -\frac{\hat{C}_t}{\hat{C}_{t+1}} \left.\frac{1-\tilde{\delta}+\tilde{r}_{t+1}}{1-\tilde{\delta}+\tilde{r}_t-\hat{C}_t+\tilde{\sigma}\epsilon_{K,t+1}}\right|_{ss} = -\frac{1-\tilde{\delta}+\tilde{\gamma}}{1-\tilde{\delta}+\tilde{\gamma}-\hat{C}} = -\frac{1}{\tilde{\beta}}, \\ \left.\frac{\partial G}{\partial \ln \hat{C}_t}\right|_{ss} &= \frac{\hat{C}_t}{\hat{C}_{t+1}} \frac{1-\tilde{\delta}+\tilde{r}_{t+1}}{1-\tilde{\delta}+\tilde{r}_t-\hat{C}_t+\tilde{\sigma}\epsilon_{K,t+1}} - \frac{\hat{C}_t}{\hat{C}_{t+1}} \left.\frac{1-\tilde{\delta}+\tilde{r}_{t+1}}{(1-\tilde{\delta}+\tilde{r}_t-\hat{C}_t+\tilde{\sigma}\epsilon_{K,t+1})^2}(-\hat{C}_t)\right|_{ss} \\ &= \frac{1-\tilde{\delta}+\tilde{\gamma}}{1-\tilde{\delta}+\tilde{\gamma}-\hat{C}} + \hat{C}\frac{1-\tilde{\delta}+\tilde{\gamma}}{(1-\tilde{\delta}+\tilde{\gamma}-\hat{C})^2} = \frac{1}{\tilde{\beta}^2} \\ \left.\frac{\partial G}{\partial r_{t+1}}\right|_{ss} &= \frac{\hat{C}_t}{\hat{C}_{t+1}} \left.\frac{1}{1-\tilde{\delta}+\tilde{r}_t-\hat{C}_t+\tilde{\sigma}\epsilon_{K,t+1}}\right|_{ss} = \frac{1}{1-\tilde{\delta}+\tilde{\gamma}-\hat{C}} = \frac{1-\tilde{\beta}}{\tilde{\beta}\hat{C}} \\ \left.\frac{\partial G}{\partial r_t}\right|_{ss} &= -\frac{\hat{C}_t}{\hat{C}_{t+1}} \left.\frac{1-\tilde{\delta}+\tilde{r}_{t+1}}{(1-\tilde{\delta}+\tilde{r}_t-\hat{C}_t+\tilde{\sigma}\epsilon_{K,t+1})^2}\right|_{ss} = -\frac{1-\tilde{\delta}+\tilde{\gamma}}{(1-\tilde{\delta}+\tilde{\gamma}-\hat{C})^2} = -\frac{1-\tilde{\beta}}{\tilde{\beta}^2\hat{C}} \\ \left.\frac{\partial G}{\partial \epsilon_{K,t+1}}\right|_{ss} &= -\frac{\hat{C}_t}{\hat{C}_{t+1}} \left.\frac{1-\tilde{\delta}+\tilde{r}_{t+1}}{(1-\tilde{\delta}+\tilde{r}_t-\hat{C}_t+\tilde{\sigma}\epsilon_{K,t+1})^2}\tilde{\sigma}\right|_{ss} = -\frac{\tilde{\sigma}(1-\tilde{\beta})}{\tilde{\beta}^2\hat{C}} \end{aligned}$$

so that we get the matrix system

$$\begin{aligned} \begin{pmatrix} 1 & -(1-\tilde{\beta})/\hat{C} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \ln \hat{C}_{t+1} - \ln \hat{C} \\ \tilde{r}_{t+1}-\tilde{\gamma} \end{pmatrix} \;=\; & \begin{pmatrix} 1/\tilde{\beta} & -(1-\tilde{\beta})/(\tilde{\beta}\hat{C}) \\ 0 & 1-\tilde{\kappa} \end{pmatrix} \begin{pmatrix} \ln \hat{C}_t - \ln \hat{C} \\ \tilde{r}_t - \tilde{\gamma} \end{pmatrix} \\ & + \begin{pmatrix} -\tilde{\epsilon}_{K,t+1} - v_{t+1} \\ \tilde{\eta}\epsilon_{A,t+1} \end{pmatrix} \end{aligned}$$

where $\tilde{\epsilon}_{K,t+1} \equiv (\tilde{\sigma}(1-\tilde{\beta})/(\tilde{\beta}\hat{C}))\epsilon_{K,t+1}$. The matrix equation is of the form

$$\Phi_0 z_{t+1} = \Phi_1 z_t + \xi_{t+1} \quad \Rightarrow \quad z_{t+1} = (\Phi_0^{-1}\Phi_1) z_t + \Phi_0^{-1}\xi_{t+1}$$

where z_t is the vector of variables $z_t = (\ln \hat{C}_t - \ln \hat{C}, \tilde{r}_t - \tilde{\gamma})^\top$, and Φ_0 and Φ_1 are the coefficient matrices containing the structural parameters. Note that $\Phi_0^{-1}\Phi_1$ can be found as

$$\Phi_0^{-1}\Phi_1 = \begin{pmatrix} 1 & (1-\tilde{\beta})/\hat{C} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1/\tilde{\beta} & -(1-\tilde{\beta})/(\tilde{\beta}\hat{C}) \\ 0 & 1-\tilde{\kappa} \end{pmatrix} = \begin{pmatrix} 1/\tilde{\beta} & (1-\tilde{\beta})(1-\tilde{\kappa}-1/\tilde{\beta})/\hat{C} \\ 0 & 1-\tilde{\kappa} \end{pmatrix}$$

and the eigenvalues are obtained from the characteristic equation $|\Phi_0^{-1}\Phi_1 - \lambda I_2| = 0$ or

$$\begin{vmatrix} 1/\tilde{\beta} - \lambda & (1-\tilde{\beta})(1-\tilde{\kappa}-1/\tilde{\beta})/\hat{C} \\ 0 & 1-\tilde{\kappa}-\lambda \end{vmatrix} = (1/\tilde{\beta} - \lambda)(1-\tilde{\kappa}-\lambda) = 0$$

which yields $\lambda_1 = 1/\tilde{\beta}$ and $\lambda_2 = 1-\tilde{\kappa}$. While the latter is positive and less than 1, the first eigenvalue is greater than 1, that is, the economy will have a saddle path property, with a single trajectory leading to the unique steady state of the system.

Hence, we obtain the linear solution to the homogeneous matrix equation

$$z_t = \begin{pmatrix} \ln \hat{C}_t - \ln \hat{C} \\ \tilde{r}_t - \tilde{\gamma} \end{pmatrix} = \mathbb{C}_1 (1/\tilde{\beta})^t \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mathbb{C}_2 (1-\tilde{\kappa})^t \begin{pmatrix} (1-\tilde{\beta})/\hat{C} \\ 1 \end{pmatrix}$$

Because we need to focus on the stable path, the stability condition requires $\mathbb{C}_1 = 0$ and from the solution of the Vasicek specification we get $\mathbb{C}_2 = \tilde{r}_0 - \tilde{\gamma}$. Hence, we find that

$$\begin{aligned} \ln \hat{C}_t - \ln \hat{C} &= \frac{1-\tilde{\beta}}{\hat{C}}(\tilde{r}_t - \tilde{\gamma}) \\ \Leftrightarrow \ \ln(C_t/K_t) &= \ln \hat{C} + \frac{1}{1-\tilde{\delta}+\tilde{\gamma}}(\tilde{r}_t - \tilde{\gamma}) \end{aligned} \tag{B.21}$$

Given any value of r_0 and initial value K_0, we obtain the optimal level of consumption C_0, the next periods capital stock K_{t+1} is obtained from (B.19).[4] Using this solution, we get

$$\begin{aligned} \ln(C_{t+1}/C_t) - \ln(K_{t+1}/K_t) &= \frac{1}{1-\tilde{\delta}+\tilde{\gamma}}(\tilde{r}_{t+1} - \tilde{r}_t) \\ &= \frac{-\tilde{\kappa}}{1-\tilde{\delta}+\tilde{\gamma}}(\tilde{r}_t - \tilde{\gamma}) + \frac{1}{1-\tilde{\delta}+\tilde{\gamma}}\tilde{\eta}\epsilon_{A,t+1} \end{aligned}$$

Using a log-linear approximation of (B.19) and insert the solution such that

$$\begin{aligned} (1+\nu)(\ln(K_{t+1}/K_t) - \ln(1+\nu)) &\simeq \tilde{r}_t - \tilde{\gamma} - \hat{C}(\ln \hat{C}_t - \ln \hat{C}) + \tilde{\sigma}\epsilon_{K,t+1} \\ &= \tilde{r}_t - \tilde{\gamma} - (1-\tilde{\beta})(\tilde{r}_t - \tilde{\gamma}) + \tilde{\sigma}\epsilon_{K,t+1} \\ &= \tilde{\beta}(\tilde{r}_t - \tilde{\gamma}) + \tilde{\sigma}\epsilon_{K,t+1} \end{aligned}$$

[4] Note that solving a linear (instead a log-linear) approximation would imply $C_t/K_t = (1-\tilde{\beta})(1-\tilde{\delta}+\tilde{r}_t)$, which could be then log-linearized to arrive at the same result $\ln(C_t/K_t) = \ln \hat{C} + (\tilde{r}_t - \tilde{\gamma})/(1-\tilde{\delta}+\tilde{\gamma})$.

or

$$\ln(K_{t+1}/K_t) = \ln(1+\nu) + \frac{1}{1-\tilde{\delta}+\tilde{\gamma}}(\tilde{r}_t - \tilde{\gamma}) + \frac{\tilde{\sigma}}{1+\nu}\epsilon_{K,t+1} \tag{B.22}$$

yields

$$\ln(C_{t+1}/C_t) = \ln(1+\nu) + \frac{1-\tilde{\kappa}}{1-\tilde{\delta}+\tilde{\gamma}}(\tilde{r}_t - \tilde{\gamma}) + \frac{\tilde{\sigma}}{1+\nu}\epsilon_{K,t+1} + \frac{\tilde{\eta}}{1-\tilde{\delta}+\tilde{\gamma}}\epsilon_{A,t+1} \tag{B.23}$$

such that the expectations error implied by our solution is (as from the Euler equation)

$$v_{t+1} \equiv -\frac{\tilde{\sigma}}{1+\nu}\epsilon_{K,t+1} \tag{B.24}$$

which implies that $E_t(v_{t+1}) = 0$ and $Var_t(v_{t+1}) = \tilde{\sigma}^2/(1+\nu)^2$.

We may use (B.22) together with a log-linear approximation of $Y_t = A_t K_t$ with $A_t = \tilde{r}_t$ or

$$\ln Y_t - \ln K_t \simeq \tilde{\gamma} + \frac{\tilde{r}_t - \tilde{\gamma}}{\tilde{\gamma}}$$

(such that the interest rate does not appear as logarithmic function) to obtain

$$\begin{aligned}
\ln(Y_{t+1}/Y_t) &= \ln(1+\nu) + \frac{\tilde{r}_{t+1} - \tilde{r}_t}{\tilde{\gamma}} + \frac{1}{1-\tilde{\delta}+\tilde{\gamma}}(\tilde{r}_t - \tilde{\gamma}) + \frac{\tilde{\sigma}}{1+\nu}\epsilon_{K,t+1} \\
&= \ln(1+\nu) + -\frac{\tilde{\kappa}}{\tilde{\gamma}}(\tilde{r}_t - \tilde{\gamma}) + \frac{1}{1-\tilde{\delta}+\tilde{\gamma}}(\tilde{r}_t - \tilde{\gamma}) + \frac{\tilde{\sigma}}{1+\nu}\epsilon_{K,t+1} + \frac{\tilde{\eta}}{\tilde{\gamma}}\epsilon_{A,t+1} \\
&= \ln(1+\nu) + \frac{\tilde{\gamma} - \tilde{\kappa}(1-\tilde{\delta}+\tilde{\gamma})}{1-\tilde{\delta}+\tilde{\gamma}}\frac{\tilde{r}_t - \tilde{\gamma}}{\tilde{\gamma}} + \frac{\tilde{\sigma}}{1+\nu}\epsilon_{K,t+1} + \frac{\tilde{\eta}}{\tilde{\gamma}}\epsilon_{A,t+1}
\end{aligned}$$

Summarizing, and using $\ln(1+\nu) = \ln\tilde{\beta} + \ln(1-\tilde{\delta}+\tilde{\gamma}) \approx \ln\tilde{\beta} - \tilde{\delta} + \tilde{\gamma}$ yields

$$\ln(C_{t+1}/C_t) = \ln\tilde{\beta} - \tilde{\delta} + \tilde{\gamma} + \frac{1-\tilde{\kappa}}{1-\tilde{\delta}+\tilde{\gamma}}(\tilde{r}_t - \tilde{\gamma}) + \frac{\tilde{\sigma}\epsilon_{K,t+1}}{\tilde{\beta}(1-\tilde{\delta}+\tilde{\gamma})} + \frac{\tilde{\eta}\epsilon_{A,t+1}}{1-\tilde{\delta}+\tilde{\gamma}} \tag{B.25a}$$

$$\ln(Y_{t+1}/Y_t) = \ln\tilde{\beta} - \tilde{\delta} + \tilde{\gamma} + \frac{\tilde{\gamma} - \tilde{\kappa}(1-\tilde{\delta}+\tilde{\gamma})}{1-\tilde{\delta}+\tilde{\gamma}}\frac{\tilde{r}_t - \tilde{\gamma}}{\tilde{\gamma}} + \frac{\tilde{\sigma}\epsilon_{K,t+1}}{\tilde{\beta}(1-\tilde{\delta}+\tilde{\gamma})} + \frac{\tilde{\eta}\epsilon_{A,t+1}}{\tilde{\gamma}} \tag{B.25b}$$

$$\tilde{r}_{t+1} = \tilde{r}_t + \tilde{\kappa}(\tilde{\gamma} - \tilde{r}_t) + \tilde{\eta}\epsilon_{A,t+1} \tag{B.25c}$$

In this AK-Vasicek model, the relation between the one-period risk-free rate and the rental rate of capital is given by (see Section B.10)

$$(\tilde{r}_t^f - \tilde{\gamma} + \tilde{\delta})(1-\tilde{\delta}+\tilde{\gamma}) + \tfrac{1}{2}\frac{(\tilde{\sigma}/\tilde{\beta})^2 + \tilde{\eta}^2}{(1-\tilde{\delta}+\tilde{\gamma})} \approx (1-\tilde{\kappa})(\tilde{r}_t - \tilde{\gamma})$$

so that we may write (B.25c) as

$$\tilde{r}_{t+1}^f = \tilde{r}_t^f + \tilde{\kappa}(\tilde{\gamma} - \tilde{\delta} - \tilde{r}_t^f) - \tfrac{1}{2}\tilde{\kappa}\frac{(\tilde{\sigma}/\tilde{\beta})^2 + \tilde{\eta}^2}{(1-\tilde{\delta}+\tilde{\gamma})^2} + \frac{1-\tilde{\kappa}}{1-\tilde{\delta}+\tilde{\gamma}}\tilde{\eta}\epsilon_{A,t+1}$$

B.6 Summary of the two empirical specifications

In what follows, we use the empirical specifications where we may use financial market data for identification and estimation of structural parameters. For the continuous-time version we use the closed-form solution $C_t = \rho K_t$ together with the nonlinear equilibrium dynamics. For the discrete-time version we have approximately $\ln(C_t/K_t) = \ln \hat{C} + (\tilde{r}_t - \tilde{\gamma})/(1 - \tilde{\delta} + \tilde{\gamma})$ together with the log-linear equilibrium dynamics.

First, based on quarterly data, the consumption (Euler) equations are

$$\ln(C_t/C_{t-1}) = \ln \tilde{\beta} + \tilde{r}^f_{t-1} + \tfrac{1}{2}\frac{(\tilde{\sigma}/\tilde{\beta})^2 + \tilde{\eta}^2}{(1-\tilde{\delta}+\tilde{\gamma})^2} + \frac{\tilde{\sigma}}{\tilde{\beta}(1-\tilde{\delta}+\tilde{\gamma})}\epsilon_{K,t} + \frac{\tilde{\eta}}{1-\tilde{\delta}+\tilde{\gamma}}\epsilon_{A,t}$$

vs.

$$\ln(C_t/C_{t-\Delta}) = \int_{t-\Delta}^{t} r^f_v dv - \left(\rho - \tfrac{1}{2}\sigma^2\right)\Delta + \sigma(Z_t - Z_{t-\Delta})$$

Second, based on quarterly data, the output equations (resource constraints) yield

$$\begin{aligned}\ln(Y_t/Y_{t-1}) &= \ln \tilde{\beta} + \tilde{r}^f_{t-1} + \tfrac{1}{2}\frac{(\tilde{\sigma}/\tilde{\beta})^2 + \tilde{\eta}^2}{(1-\tilde{\delta}+\tilde{\gamma})^2} - \frac{(1-\tilde{\delta})\tilde{\kappa}}{(1-\tilde{\kappa})\tilde{\gamma}}\left(\tilde{r}^f_{t-1} - \tilde{\gamma} + \tilde{\delta} + \tfrac{1}{2}\frac{(\tilde{\sigma}/\tilde{\beta})^2 + \tilde{\eta}^2}{(1-\tilde{\delta}+\tilde{\gamma})^2}\right) \\ &\quad + \frac{\tilde{\sigma}}{\tilde{\beta}(1-\tilde{\delta}+\tilde{\gamma})}\epsilon_{K,t} + \frac{\tilde{\eta}}{\tilde{\gamma}}\epsilon_{A,t}\end{aligned}$$

vs.

$$\begin{aligned}\ln(Y_t/Y_{t-\Delta}) &= \int_{t-\Delta}^{t} r^f_v dv + \kappa\gamma \int_{t-\Delta}^{t} 1/(r^f_v + \delta + \sigma^2)dv - \tfrac{1}{2}\eta^2 \int_{t-\Delta}^{t} 1/(r^f_v + \delta + \sigma^2)^2 dv \\ &\quad - \left(\kappa + \rho - \tfrac{1}{2}\sigma^2\right)\Delta + \int_{t-\Delta}^{t} \eta/(r^f_v + \delta + \sigma^2)dB_v + \sigma(Z_t - Z_{t-\Delta})\end{aligned}$$

Third, based on quarterly data, the interest rate equations (Vasicek specifications) read

$$\tilde{r}^f_t = (1-\tilde{\kappa})\tilde{r}^f_{t-1} + \tilde{\kappa}\left(\tilde{\gamma} - \tilde{\delta} - \tfrac{1}{2}\frac{(\tilde{\sigma}/\tilde{\beta})^2 + \tilde{\eta}^2}{(1-\tilde{\delta}+\tilde{\gamma})^2}\right) + \frac{1-\tilde{\kappa}}{1-\tilde{\delta}+\tilde{\gamma}}\tilde{\eta}\epsilon_{A,t}$$

vs.

$$r^f_t = e^{-\kappa\Delta} r^f_{t-\Delta} + (1 - e^{-\kappa\Delta})(\gamma - \delta - \sigma^2) + \eta e^{-\kappa\Delta}\int_{t-\Delta}^{t} e^{\kappa(v-(t-\Delta))} dB_v$$

Note that r_t denotes the annual interest rate whereas $\tilde{r}_t$ denotes the periodic interest rate.

To summarize, the empirical specification comprises

$$\ln(C_t/C_{t-1}) = \ln\tilde{\beta} + \tilde{r}^f_{t-1} + \tfrac{1}{2}\frac{(\tilde{\sigma}/\tilde{\beta})^2 + \tilde{\eta}^2}{(1-\tilde{\delta}+\tilde{\gamma})^2} + \frac{\tilde{\sigma}}{\tilde{\beta}(1-\tilde{\delta}+\tilde{\gamma})}\epsilon_{K,t} + \frac{\tilde{\eta}}{1-\tilde{\delta}+\tilde{\gamma}}\epsilon_{A,t} \tag{B.26a}$$

$$\begin{aligned}\ln(Y_t/Y_{t-1}) &= \ln\tilde{\beta} + \tilde{r}^f_{t-1} + \tfrac{1}{2}\frac{(\tilde{\sigma}/\tilde{\beta})^2 + \tilde{\eta}^2}{(1-\tilde{\delta}+\tilde{\gamma})^2} - \frac{(1-\tilde{\delta})\tilde{\kappa}}{(1-\tilde{\kappa})\tilde{\gamma}}\left(\tilde{r}^f_{t-1} - \tilde{\gamma} + \tilde{\delta} + \tfrac{1}{2}\frac{(\tilde{\sigma}/\tilde{\beta})^2 + \tilde{\eta}^2}{(1-\tilde{\delta}+\tilde{\gamma})^2}\right) \\ &\quad + \frac{\tilde{\sigma}}{\tilde{\beta}(1-\tilde{\delta}+\tilde{\gamma})}\epsilon_{K,t} + \frac{\tilde{\eta}}{\tilde{\gamma}}\epsilon_{A,t}\end{aligned} \tag{B.26b}$$

$$\tilde{r}^f_t = (1-\tilde{\kappa})\tilde{r}^f_{t-1} + \tilde{\kappa}\left(\tilde{\gamma} - \tilde{\delta} - \tfrac{1}{2}\frac{(\tilde{\sigma}/\tilde{\beta})^2 + \tilde{\eta}^2}{(1-\tilde{\delta}+\tilde{\gamma})^2}\right) + \frac{1-\tilde{\kappa}}{1-\tilde{\delta}+\tilde{\gamma}}\tilde{\eta}\epsilon_{A,t} \tag{B.26c}$$

or

$$\begin{aligned}\ln(C_t/C_{t-1}) &= \ln\tilde{\beta} + \tilde{r}^f_{t-1} + \mathbb{C}_0 + \varepsilon_{C,t} \\ \ln(Y_t/Y_{t-1}) &= \ln\tilde{\beta} + \tilde{r}^f_{t-1} + \mathbb{C}_0 - \mathbb{C}_2\left(\tilde{r}^f_{t-1} - \mathbb{C}_1\right) + \varepsilon_{Y,t} \\ \tilde{r}^f_t &= (1-\tilde{\kappa})\tilde{r}^f_{t-1} + \tilde{\kappa}\mathbb{C}_1 + \varepsilon_{r,t}\end{aligned}$$

where

$$\mathbb{C}_2 \equiv \frac{(1-\tilde{\delta})\tilde{\kappa}}{(1-\tilde{\kappa})\tilde{\gamma}}, \quad \mathbb{C}_1 \equiv \tilde{\gamma} - \tilde{\delta} - \mathbb{C}_0, \quad \mathbb{C}_0 \equiv \tfrac{1}{2}\frac{(\tilde{\sigma}/\tilde{\beta})^2 + \tilde{\eta}^2}{(1-\tilde{\delta}+\tilde{\gamma})^2}$$

and

$$\varepsilon_{C,t} = \frac{\tilde{\sigma}}{\tilde{\beta}(1-\tilde{\delta}+\tilde{\gamma})}\epsilon_{K,t} + \frac{\tilde{\eta}}{1-\tilde{\delta}+\tilde{\gamma}}\epsilon_{A,t}, \tag{B.27a}$$

$$\varepsilon_{Y,t} = \frac{\tilde{\sigma}}{\tilde{\beta}(1-\tilde{\delta}+\tilde{\gamma})}\epsilon_{K,t} + \frac{\tilde{\eta}}{\tilde{\gamma}}\epsilon_{A,t} \tag{B.27b}$$

$$\varepsilon_{r,t} = \frac{1-\tilde{\kappa}}{1-\tilde{\delta}+\tilde{\gamma}}\tilde{\eta}\epsilon_{A,t} \tag{B.27c}$$

We employ the following mapping of structural parameters $\phi = (\kappa, \gamma, \eta, \rho, \delta, \sigma)^\top$

$$\begin{aligned}\kappa &\simeq \tilde{\kappa} = 1 - e^{-\Delta\kappa} \\ \gamma &\simeq \tilde{\gamma} = \Delta\gamma \\ \eta &\simeq \tilde{\eta} = \Delta\eta\sqrt{(1-e^{-2\kappa\Delta})/(2\kappa)} \\ \rho &\simeq \tilde{\beta} = e^{-\Delta\rho} \\ \delta &\simeq \tilde{\delta} = 1 - e^{-\Delta\delta} \\ \sigma &\simeq \tilde{\sigma} = \Delta^{1/2}\tilde{\beta}(1-\tilde{\delta}+\tilde{\gamma})\sigma\end{aligned}$$

in which $\Delta = 1/12$ for monthly data, $\Delta = 1/4$ for quarterly data.

B.7 Residual Covariance Matrix and MEF

From system (B.26) we obtain the fitted residual covariance matrix

$$\hat{\Sigma} \equiv \begin{pmatrix} \hat{\Sigma}_{CC} & \hat{\Sigma}_{CY} & \hat{\Sigma}_{Cr} \\ \hat{\Sigma}_{YC} & \hat{\Sigma}_{YY} & \hat{\Sigma}_{Yr} \\ \hat{\Sigma}_{rC} & \hat{\Sigma}_{rY} & \hat{\Sigma}_{rr} \end{pmatrix}$$

where

$$\hat{\Sigma}_{CC} = \frac{\tilde{\sigma}^2 + (\tilde{\beta}\tilde{\eta})^2}{(\tilde{\beta}(1-\tilde{\delta}+\tilde{\gamma}))^2}, \quad \hat{\Sigma}_{CY} = \frac{\tilde{\sigma}^2}{(\tilde{\beta}(1-\tilde{\delta}+\tilde{\gamma}))^2} + \frac{\tilde{\eta}^2}{\tilde{\gamma}(1-\tilde{\delta}+\tilde{\gamma})}, \quad \hat{\Sigma}_{Cr} = \frac{\tilde{\eta}^2(1-\tilde{\kappa})}{(1-\tilde{\delta}+\tilde{\gamma})^2}$$

$$\hat{\Sigma}_{YY} = \frac{\tilde{\sigma}^2}{(\tilde{\beta}(1-\tilde{\delta}+\tilde{\gamma}))^2} + \frac{\tilde{\eta}^2}{\tilde{\gamma}^2}, \quad \hat{\Sigma}_{Yr} = \frac{\tilde{\eta}^2(1-\tilde{\kappa})}{\tilde{\gamma}(1-\tilde{\delta}+\tilde{\gamma})}, \quad \hat{\Sigma}_{rr} = \frac{\tilde{\eta}^2(1-\tilde{\kappa})^2}{(1-\tilde{\delta}+\tilde{\gamma})^2}$$

which we may use for better identification of structural model parameters.

Let ϕ denote the parameter vector of interest and let $m_t = m_t(\phi)$ denote the 3-vector of martingale increments generated by the model, expressed in terms of data and parameters. Specifically, we let $m_t = \varepsilon_t = (\varepsilon_{C,t}, \varepsilon_{Y,t}, \varepsilon_{r,t})$ in (B.26) be a martingale difference sequence,

$$m_t = \begin{pmatrix} \ln(C_t/C_{t-1}) - \tilde{r}^f_{t-1} - (\ln\tilde{\beta} + \mathbb{C}_0) \\ \ln(Y_t/Y_{t-1}) - \tilde{r}^f_{t-1} - (\ln\tilde{\beta} + \mathbb{C}_0) + \mathbb{C}_2(\tilde{r}^f_{t-1} - \mathbb{C}_1) \\ \tilde{r}^f_t - (1-\tilde{\kappa})\tilde{r}^f_{t-1} - \tilde{\kappa}\mathbb{C}_1 \end{pmatrix} \tag{B.28}$$

The optimal weights are given by

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

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

$$\Psi_t = Var_{t-1}(m_t) = E_{t-1}(m_t m_t^\top)$$

and ψ_t the conditional mean of its parameter derivative

$$\psi_t = E_{t-1}(\partial m_t/\partial\phi^\top).$$

Here, the conditional variance of the vector martingale increment is constant and given by the residual covariance matrix, $\Psi_t = \hat{\Sigma}$. Using the martingale increments (B.28), we get the derivatives $(\partial m_t/\partial\phi^\top)^\top$ with respect to the parameter vector $\phi = (\tilde{\kappa}, \tilde{\gamma}, \tilde{\eta}, \tilde{\beta}, \tilde{\delta}, \tilde{\sigma})^\top$,

$$\begin{pmatrix} 0 & (\partial\mathbb{C}_2/\partial\tilde{\kappa})(r^f_{t-1} - \mathbb{C}_1) & \tilde{r}^f_{t-1} - \mathbb{C}_1 \\ -(\partial\mathbb{C}_0/\partial\tilde{\gamma}) & -(\partial\mathbb{C}_0/\partial\tilde{\gamma}) + (\partial\mathbb{C}_2/\partial\tilde{\gamma})\tilde{r}^f_{t-1} - (\partial(\mathbb{C}_1\mathbb{C}_2)/\partial\tilde{\gamma}) & -\tilde{\kappa}(1 - (\partial\mathbb{C}_0/\partial\tilde{\gamma})) \\ -(\partial\mathbb{C}_0/\partial\tilde{\eta}) & -(\partial\mathbb{C}_0/\partial\tilde{\eta}) + \mathbb{C}_2(\partial\mathbb{C}_0/\partial\tilde{\eta}) & \tilde{\kappa}(\partial\mathbb{C}_0/\partial\tilde{\eta}) \\ -1/\tilde{\beta} - (\partial\mathbb{C}_0/\partial\tilde{\beta}) & -1/\tilde{\beta} - (\partial\mathbb{C}_0/\partial\tilde{\beta}) + \mathbb{C}_2(\partial\mathbb{C}_0/\partial\tilde{\beta}) & \tilde{\kappa}(\partial\mathbb{C}_0/\partial\tilde{\beta}) \\ -(\partial\mathbb{C}_0/\partial\tilde{\delta}) & -(\partial\mathbb{C}_0/\partial\tilde{\delta}) + (\partial\mathbb{C}_2/\partial\tilde{\delta})\tilde{r}^f_{t-1} - (\partial(\mathbb{C}_1\mathbb{C}_2)/\partial\tilde{\delta}) & \tilde{\kappa}(1 + (\partial\mathbb{C}_0/\partial\tilde{\delta})) \\ -(\partial\mathbb{C}_0/\partial\tilde{\sigma}) & -(\partial\mathbb{C}_0/\partial\tilde{\sigma}) + \mathbb{C}_2(\partial\mathbb{C}_0/\partial\tilde{\sigma}) & \tilde{\kappa}(\partial\mathbb{C}_0/\partial\tilde{\sigma}) \end{pmatrix}$$

where

$$\partial\mathbb{C}_0/\partial\tilde{\gamma} = -\frac{(\tilde{\sigma}/\tilde{\beta})^2 + \tilde{\eta}^2}{(1-\tilde{\delta}+\tilde{\gamma})^3} = -2\mathbb{C}_0/(1-\tilde{\delta}+\tilde{\gamma}),$$

$$\partial\mathbb{C}_0/\partial\tilde{\eta} = \tilde{\eta}/(1-\tilde{\delta}+\tilde{\gamma})^2,$$

$$\partial\mathbb{C}_0/\partial\tilde{\beta} = -(\tilde{\sigma}/\tilde{\beta})^2/(\tilde{\beta}(1-\tilde{\delta}+\tilde{\gamma})^2),$$

$$\partial\mathbb{C}_0/\partial\tilde{\delta} = \frac{(\tilde{\sigma}/\tilde{\beta})^2 + \tilde{\eta}^2}{(1-\tilde{\delta}+\tilde{\gamma})^3} = 2\mathbb{C}_0/(1-\tilde{\delta}+\tilde{\gamma}),$$

$$\partial\mathbb{C}_0/\partial\tilde{\sigma} = \tilde{\sigma}/(\tilde{\beta}(1-\tilde{\delta}+\tilde{\gamma}))^2,$$

and

$$\partial\mathbb{C}_1/\partial\tilde{\gamma} = 1 - (\partial\mathbb{C}_0/\partial\tilde{\gamma})$$

$$\partial\mathbb{C}_1/\partial\tilde{\delta} = -1 - (\partial\mathbb{C}_0/\partial\tilde{\delta})$$

and

$$\partial\mathbb{C}_2/\partial\tilde{\kappa} = \frac{1-\tilde{\delta}}{(1-\tilde{\kappa})^2\tilde{\gamma}} = \mathbb{C}_2/(\tilde{\kappa}(1-\tilde{\kappa}))$$

$$\partial\mathbb{C}_2/\partial\tilde{\gamma} = -(1-\tilde{\delta})\tilde{\kappa}/((1-\tilde{\kappa})\tilde{\gamma}^2) = -\mathbb{C}_2/\tilde{\gamma}$$

$$\partial\mathbb{C}_2/\partial\tilde{\delta} = -\tilde{\kappa}/((1-\tilde{\kappa})\tilde{\gamma})$$

and

$$\partial(\mathbb{C}_1\mathbb{C}_2)/\partial\tilde{\gamma} = (\partial\mathbb{C}_1/\partial\tilde{\gamma})\mathbb{C}_2 + (\partial\mathbb{C}_2/\partial\tilde{\gamma})\mathbb{C}_1 = \mathbb{C}_2 - (\partial\mathbb{C}_0/\partial\tilde{\gamma})\mathbb{C}_2 - \mathbb{C}_1\mathbb{C}_2/\tilde{\gamma}$$

$$\partial(\mathbb{C}_1\mathbb{C}_2)/\partial\tilde{\delta} = (\partial\mathbb{C}_1/\partial\tilde{\delta})\mathbb{C}_2 + (\partial\mathbb{C}_2/\partial\tilde{\delta})\mathbb{C}_1$$

and with respect to the parameter vector $\phi = (\kappa, \gamma, \eta, \rho, \delta, \sigma)^\top$,

$$\begin{pmatrix} -(\partial\mathbb{C}_0/\partial\tilde{\eta})(\partial\tilde{\eta}/\partial\kappa) & \phi_{12} & \Delta e^{-\Delta\kappa}(\tilde{r}^f_{t-1} - \mathbb{C}_1) + \tilde{\kappa}(\partial\mathbb{C}_0/\partial\tilde{\eta})(\partial\tilde{\eta}/\partial\kappa) \\ -(\partial\mathbb{C}_0/\partial\tilde{\gamma})\Delta - (\partial\mathbb{C}_0/\partial\tilde{\sigma})\Delta^{3/2}\tilde{\beta}\sigma & \phi_{22} & -\Delta\tilde{\kappa}(1 - (\partial\mathbb{C}_0/\partial\tilde{\gamma}) - (\partial\mathbb{C}_0/\partial\tilde{\sigma})\Delta^{1/2}\tilde{\beta}\sigma) \\ -(\partial\mathbb{C}_0/\partial\tilde{\eta})(\tilde{\eta}/\eta) & \phi_{32} & \tilde{\kappa}(\partial\mathbb{C}_0/\partial\tilde{\eta})(\tilde{\eta}/\eta) \\ \Delta & \Delta & 0 \\ -((\partial\mathbb{C}_0/\partial\tilde{\delta}) - (\partial\mathbb{C}_0/\partial\tilde{\sigma})\Delta^{1/2}\tilde{\beta})\Delta e^{-\Delta\delta} & \phi_{52} & \tilde{\kappa}\Delta e^{-\Delta\delta}(1 + (\partial\mathbb{C}_0/\partial\tilde{\delta})) - \tilde{\kappa}(\partial\mathbb{C}_0/\partial\tilde{\sigma})\Delta^{1/2}\tilde{\beta}\sigma \\ -(\partial\mathbb{C}_0/\partial\tilde{\sigma})(\tilde{\sigma}/\sigma) & \phi_{62} & \tilde{\kappa}(\partial\mathbb{C}_0/\partial\tilde{\sigma})(\tilde{\sigma}/\sigma) \end{pmatrix}$$

where

$$\phi_{12} \equiv -(\partial\mathbb{C}_0/\partial\tilde{\eta})(\partial\tilde{\eta}/\partial\kappa) + (\partial\mathbb{C}_2/\partial\tilde{\kappa})\Delta e^{-\Delta\kappa} r^f_{t-1} + (\mathbb{C}_2(\partial\mathbb{C}_0/\partial\tilde{\eta})(\partial\tilde{\eta}/\partial\kappa) - \mathbb{C}_1(\partial\mathbb{C}_2/\partial\tilde{\kappa}))\Delta e^{-\Delta\kappa},$$

$$\phi_{22} \equiv -(\partial\mathbb{C}_0/\partial\tilde{\gamma}\Delta + \partial\mathbb{C}_0/\partial\tilde{\sigma}\Delta^{1/2}\tilde{\beta}\sigma) + (\partial\mathbb{C}_2/\partial\tilde{\gamma})\Delta\tilde{r}^f_{t-1} - (\partial(\mathbb{C}_1\mathbb{C}_2)/\partial\gamma),$$

$$\phi_{32} \equiv -(\partial\mathbb{C}_0/\partial\tilde{\eta})(\tilde{\eta}/\eta) + \mathbb{C}_2(\partial\mathbb{C}_0/\partial\tilde{\eta})(\tilde{\eta}/\eta),$$

$$\phi_{52} \equiv -(\partial\mathbb{C}_0/\partial\tilde{\delta})\Delta e^{-\Delta\delta} + (\partial\mathbb{C}_2/\partial\tilde{\delta})\Delta e^{-\Delta\delta}\tilde{r}^f_{t-1} - (\partial(\mathbb{C}_1\mathbb{C}_2)/\partial\delta),$$

$$\phi_{62} \equiv -(\partial \mathbb{C}_0/\partial \tilde{\sigma})(\tilde{\sigma}/\sigma) + \mathbb{C}_2(\partial \mathbb{C}_0/\partial \tilde{\sigma})(\tilde{\sigma}/\sigma),$$

and

$$\partial \tilde{\eta}/\partial \kappa = \tfrac{1}{2}\tilde{\eta}\left(2\Delta e^{-2\kappa\Delta}/(1-e^{-2\kappa\Delta}) - 1/\kappa\right),$$

and

$$\partial(\mathbb{C}_1\mathbb{C}_2)/\partial\gamma = (\partial\mathbb{C}_1/\partial\gamma)\mathbb{C}_2 + (\partial\mathbb{C}_2/\partial\gamma)\mathbb{C}_1 = \Delta\mathbb{C}_2 - (\partial\mathbb{C}_0/\partial\tilde{\gamma})\mathbb{C}_2 - (\partial\mathbb{C}_0/\partial\tilde{\sigma})\Delta^{3/2}\tilde{\beta}\sigma\mathbb{C}_2 - \mathbb{C}_1\mathbb{C}_2/\tilde{\gamma}\Delta,$$

$$\partial(\mathbb{C}_1\mathbb{C}_2)/\partial\delta = (-1 - (\partial\mathbb{C}_0/\partial\tilde{\delta}) + (\partial\mathbb{C}_0/\partial\tilde{\sigma})\Delta^{1/2}\tilde{\beta}\sigma)\Delta e^{-\Delta\delta}\mathbb{C}_2 + (\partial\mathbb{C}_2/\partial\tilde{\delta})\Delta e^{-\Delta\delta}\mathbb{C}_1.$$

B.8 Calibration of model parameters - periodic rates

Suppose that we want to parameterize the Ornstein-Uhlenbeck process and the first-order autoregressive process (with the discrete time process being at periodic rates)

$$dx_t = \kappa(\gamma - x_t)dt + \eta dB_t, \quad x_0 \text{ given} \quad \text{and} \quad \tilde{x}_{t+1} = \mathbb{C}_0 + \mathbb{C}_1\tilde{x}_t + \mathbb{C}_2\epsilon_{t+1}, \tag{B.29}$$

where $\tilde{x}_0 = \Delta x_0$, where $\Delta = 1/12$ for monthly and $\Delta = 1/4$ for quarterly observations, B_t a standard Brownian motion, $0 < \mathbb{C}_1 < 1$ and $\epsilon_t \sim \mathcal{N}(0,1)$.[5] The solutions are

$$x_t = e^{-\kappa t}x_0 + (1-e^{-\kappa t})\gamma + e^{-\kappa t}\eta\int_0^t e^{\kappa v}dB_v \quad \text{and} \quad \tilde{x}_t = \mathbb{C}_1^t\tilde{x}_0 + \mathbb{C}_1^t\sum_{i=1}^{t}\mathbb{C}_1^{-i}(\mathbb{C}_0 + \mathbb{C}_2\epsilon_i)$$

Let us calibrate $\mathbb{C}_i$, $i = 0,1,2$, given a parametric value for κ, γ and η at quarterly frequency, such that the expected value $\mathbb{E}_0(\Delta x_\Delta) = \mathbb{E}_0(\tilde{x}_1)$, the variance $Var_0(\Delta x_\Delta) = Var_0(\tilde{x}_1)$, and the mean of the asymptotic distribution $\mathbb{E}(\Delta x) = \mathbb{E}(\tilde{x})$ coincide. It it straightforward to show that $\mathbb{E}_0(\Delta x_\Delta) = \Delta e^{-\Delta\kappa}x_0 + \Delta(1-e^{-\Delta\kappa})\gamma$ and $\mathbb{E}_0(\tilde{x}_1) = \mathbb{C}_1\tilde{x}_0 + \mathbb{C}_0$. Moreover, $\mathbb{E}(\Delta x) = \Delta\gamma$ and $\mathbb{E}(\tilde{x}) = \mathbb{C}_0/(1-\mathbb{C}_1)$. This gives $\mathbb{C}_0 = \Delta\gamma(1-\mathbb{C}_1)$ in which $\mathbb{C}_1$ is pinned down by

$$\begin{aligned} \Delta e^{-\Delta\kappa}x_0 + \Delta\gamma - \Delta\gamma e^{-\Delta\kappa} &= \mathbb{C}_1\Delta x_0 + \Delta\gamma - \Delta\gamma\mathbb{C}_1 \\ \Leftrightarrow\ e^{-\Delta\kappa}(x_0 - \gamma) &= \mathbb{C}_1(x_0 - \gamma) \\ \Leftrightarrow\ \mathbb{C}_1 &= e^{-\Delta\kappa} \end{aligned}$$

From the Itô isometry we get

$$Var_0(\Delta x_\Delta) = \Delta^2 e^{-2\kappa\Delta}\eta^2\int_0^\Delta e^{2\kappa v}dv = \Delta^2\frac{\eta^2}{2\kappa}(1-e^{-2\kappa\Delta})$$

whereas

$$Var_0(\tilde{x}_1) = \mathbb{C}_2^2 Var_0(\epsilon_1) = \mathbb{C}_2^2$$

[5]Note that $(1+r_t)^\Delta = 1 + \tilde{r}_t$ or $\Delta\ln(1+r_t) = \ln(1+\tilde{r}_t)$ and thus $\Delta r_t \approx \tilde{r}_t$.

Equating terms implies:

$$\mathbb{C}_2 = \Delta\eta\sqrt{(1-e^{-2\kappa\Delta})/(2\kappa)}$$

As an example, our DGP with $\Delta = 1/4$, $\kappa = 0.2$, $\gamma = 0.1$ and $\eta = 0.01$ corresponds to the discrete-time process with parameterization $\mathbb{C}_0 \approx 0.001$, $\mathbb{C}_1 \approx 0.951$, and $\mathbb{C}_2 \approx 0.001$, or

$$dx_t = 0.2(0.1 - x_t)dt + 0.01dB_t \quad \simeq \quad \tilde{x}_{t+1} = 0.001 + 0.951\tilde{x}_t + 0.001\epsilon_{t+1}$$

Observe that $B_\Delta - B_0 \sim \mathcal{N}(0, \Delta)$ so we may use $\epsilon_t \equiv \Delta^{-1/2}(B_t - B_{t-\Delta}) \sim \mathcal{N}(0,1)$ in order to match the size of the realized shocks. Economically, $\tilde{x}_t$ now matches the Vasicek interest rate dynamics of quarterly interest rates observed at the quarterly frequency,

$$\tilde{x}_{t+1} - \tilde{x}_t = \tilde{\kappa}(\tilde{\gamma} - \tilde{x}_t) + \tilde{\eta}\epsilon_{t+1}, \quad \tilde{x}_0 = \Delta x_0$$

where

$$\tilde{\gamma} \equiv \Delta\gamma, \quad \tilde{\kappa} \equiv 1 - e^{-\Delta\kappa}, \quad \tilde{\eta} \equiv \Delta\eta\sqrt{(1-e^{-2\kappa\Delta})/(2\kappa)}.$$

B.9 Calibration of model parameters - stochastic depreciation

Next we want to relate the dynamics of the discrete-time resource constraint

$$\ln(K_{t+1}/K_t) = \ln(1+\nu) + \frac{1}{1-\tilde{\delta}+\tilde{\gamma}}(\tilde{r}_t - \tilde{\gamma}) + \frac{\tilde{\sigma}}{1+\nu}\epsilon_{K,t+1}$$

to the corresponding continuous-time formulation

$$\ln(K_{t+\Delta}/K_t) = \int_t^{t+\Delta} r_s ds - (\delta + \rho + \tfrac{1}{2}\sigma^2)\Delta + \sigma(Z_{t+\Delta} - Z_t),$$

where $\epsilon_{K,t+1} = \Delta^{-1/2}(Z_{t+\Delta} - Z_t) \sim \mathcal{N}(0,1)$ so in order to get the same conditional moments

$$\frac{\tilde{\sigma}}{1+\nu}\Delta^{-1/2} = \sigma,$$

$$-(\delta - \gamma + \rho + \tfrac{1}{2}\sigma^2)\Delta = \ln\tilde{\beta} + \ln(1 - \tilde{\delta} + \tilde{\gamma}).$$

It can be simplified to $(\gamma - \delta - \frac{1}{2}\sigma^2)\Delta = \ln(1 + \tilde{\gamma} - \tilde{\delta})$ which corresponds to our definitions.

B.10 Asset pricing

From (B.25a) we obtain $E_t(\exp(-\ln(C_{t+1}/C_t)))$ as

$$E_t\left(\exp\left(-\left[\ln(1+\nu) + \frac{1-\tilde{\kappa}}{1-\tilde{\delta}+\tilde{\gamma}}(\tilde{r}_t - \tilde{\gamma}) + \frac{\tilde{\sigma}}{1+\nu}\epsilon_{K,t+1} + \frac{\tilde{\eta}}{1-\tilde{\delta}+\tilde{\gamma}}\epsilon_{A,t+1}\right]\right)\right)$$
$$= \exp\left(-\ln(1+\nu) - \frac{1-\tilde{\kappa}}{1-\tilde{\delta}+\tilde{\gamma}}(\tilde{r}_t - \tilde{\gamma}) + \tfrac{1}{2}\frac{\tilde{\sigma}^2}{(1+\nu)^2} + \tfrac{1}{2}\frac{\tilde{\eta}^2}{(1-\tilde{\delta}+\tilde{\gamma})^2}\right)$$
$$= \tilde{\beta}^{-1}(1-\tilde{\delta}+\tilde{\gamma})^{-1}\exp\left(-\frac{1-\tilde{\kappa}}{1-\tilde{\delta}+\tilde{\gamma}}(\tilde{r}_t - \tilde{\gamma}) + \frac{1}{2}\frac{(\tilde{\sigma}/\tilde{\beta})^2 + \tilde{\eta}^2}{(1-\tilde{\delta}+\tilde{\gamma})^2}\right).$$

Observe that from (B.16a) the one-period risk-free rate of some bond, r_t^f, which is determined at the end of period t for the following period $t+1$ must satisfy

$$\begin{aligned} 1+\tilde{r}_t^f &= \left(\tilde{\beta} E_t[\exp\left(-\ln(C_{t+1}/C_t)\right)]\right)^{-1} \\ &= (1-\tilde{\delta}+\tilde{\gamma})\exp\left(\frac{1-\tilde{\kappa}}{1-\tilde{\delta}+\tilde{\gamma}}(\tilde{r}_t-\tilde{\gamma})-\tfrac{1}{2}\frac{(\tilde{\sigma}/\tilde{\beta})^2+\tilde{\eta}^2}{(1-\tilde{\delta}+\tilde{\gamma})^2}\right) \end{aligned}$$

and so

$$\tilde{r}_t^f \approx \tilde{\gamma}-\tilde{\delta}+\frac{1-\tilde{\kappa}}{1-\tilde{\delta}+\tilde{\gamma}}(\tilde{r}_t-\tilde{\gamma})-\tfrac{1}{2}\frac{(\tilde{\sigma}/\tilde{\beta})^2+\tilde{\eta}^2}{(1-\tilde{\delta}+\tilde{\gamma})^2} \tag{B.30}$$

and the expected risk premium over net capital rewards is $\frac{1}{2}((\tilde{\sigma}/\tilde{\beta})^2+\tilde{\eta}^2)/(1-\tilde{\delta}+\tilde{\gamma})^2$.

C MEF extensions, and MEF with five conditional moment restrictions

C.1 AK-Vasicek-RS model

In the case of the AK-Vasicek model with regime switching (cf. Section 3.3.2), the system of equilibrium dynamics reads[6]

$$\begin{aligned} d\ln C_t &= \left(r_t-\rho-\delta-\tfrac{1}{2}\sigma^2\right)dt+\sigma dZ_t, && \text{(C.1a)} \\ d\ln Y_t &= \left(\kappa\gamma/r_t-\tfrac{1}{2}(\eta_t/r_t)^2+r_t-\kappa-\rho-\delta-\tfrac{1}{2}\sigma^2\right)dt+\eta_t/r_t dB_t+\sigma dZ_t, && \text{(C.1b)} \\ dr_t &= \kappa(\gamma-r_t)dt+\eta_t dB_t, && \text{(C.1c)} \\ d\eta_t &= (\eta_l-\eta_h)dq_{1,t}+(\eta_h-\eta_l)dq_{2,t}. && \text{(C.1d)} \end{aligned}$$

Using system (C.1) and the equilibrium asset-pricing condition (14), we obtain

$$\begin{aligned} \ln(C_t/C_{t-\Delta})-\int_{t-\Delta}^t r_v^f dv &= -\left(\rho-\tfrac{1}{2}\sigma^2\right)\Delta+\varepsilon_{C,t}, && \text{(C.2a)} \\ \ln(Y_t/Y_{t-\Delta})-\int_{t-\Delta}^t r_v^f dv &= \kappa\gamma\int_{t-\Delta}^t 1/(r_v^f+\delta+\sigma^2)dv-\tfrac{1}{2}\int_{t-\Delta}^t \eta_v^2/(r_v^f+\delta+\sigma^2)^2 dv \\ &\quad -\left(\kappa+\rho-\tfrac{1}{2}\sigma^2\right)\Delta+\varepsilon_{Y,t}, && \text{(C.2b)} \\ r_t^f &= e^{-\kappa\Delta}r_{t-\Delta}^f+(1-e^{-\kappa\Delta})(\gamma-\delta-\sigma^2)+\varepsilon_{r,t}, && \text{(C.2c)} \\ \eta_t &= \eta_{t-\Delta}+(\eta_l-\eta_h)\int_{t-\Delta}^t(\phi_1(\eta_v)-\phi_2(\eta_v))dv+\varepsilon_{\eta,t}. && \text{(C.2d)} \end{aligned}$$

[6]It can be shown that the analytical solution $C_t=\rho K_t$ is not affected by the presence of regime switches such that the relation between the risk-free rate and the rental rate of capital is still given by (14).

with martingale increments given by

$$\varepsilon_{C,t} = \sigma(Z_t - Z_{t-\Delta}), \tag{C.3a}$$

$$\varepsilon_{Y,t} = \int_{t-\Delta}^{t} \eta_v/(r_v^f + \delta + \sigma^2)dB_v + \sigma(Z_t - Z_{t-\Delta}), \tag{C.3b}$$

$$\varepsilon_{r,t} = e^{-\kappa\Delta}\int_{t-\Delta}^{t} \eta_v e^{\kappa(v-(t-\Delta))}dB_v, \tag{C.3c}$$

$$\varepsilon_{\eta,t} = (\eta_l - \eta_h)\int_{t-\Delta}^{t}(dq_{1,v} - \phi_1(\eta_v)dv - dq_{2,v} + \phi_2(\eta_v)dv). \tag{C.3d}$$

We let $m_t = \varepsilon_t = (\varepsilon_{C,t}, \varepsilon_{Y,t}, \varepsilon_{r,t})^\top$ from (C.3a)-(C.3c), so $\dim m = 3$. Clearly, m_t is a martingale difference sequence, and from system (C.2) we have that in terms of data and parameters

$$m_t = \begin{pmatrix} \ln(C_t/C_{t-\Delta}) - \int_{t-\Delta}^{t} r_v^f dv + \left(\rho - \frac{1}{2}\sigma^2\right)\Delta \\ \ln(Y_t/Y_{t-\Delta}) - \int_{t-\Delta}^{t} r_v^f dv + \left(\kappa + \rho - \frac{1}{2}\sigma^2\right)\Delta - \kappa\gamma\int_{t-\Delta}^{t} 1/(r_v^f + \delta + \sigma^2)dv \\ +\frac{1}{2}\int_{t-\Delta}^{t}\eta_v^2/(r_v^f + \delta + \sigma^2)^2 dv \\ r_t^f - (1 - e^{-\kappa\Delta})(\gamma - \delta - \sigma^2) - e^{-\kappa\Delta}r_{t-\Delta}^f \end{pmatrix}, \tag{C.4}$$

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

Similarly to the case of latent variables, our procedure is to derive some proxy moments for estimation, say, $m_t^* = E\left(m_t|\mathcal{F}_{t-\Delta}\right)$, given by (46). We may also obtain $\Psi_{t,11} = \sigma^2\Delta$, $\Psi_{t,22} = E_{t-\Delta}(\int_{t-\Delta}^{t}\eta_v^2/(r_v^f+\delta+\sigma^2)^2 dv)+\sigma^2\Delta$, $\Psi_{t,33} = e^{-2\kappa\Delta}E_{t-\Delta}(\int_{t-\Delta}^{t}\eta_v^2 e^{2\kappa(v-(t-\Delta))}dv)$, $\Psi_{t,12} = \sigma^2\Delta$, $\Psi_{t,13} = 0$, and $\Psi_{t,23} = E_{t-\Delta}((\int_{t-\Delta}^{t}\eta_v/(r_v^f+\delta+\sigma^2)dB_v)(e^{-\kappa_1\Delta}\int_{t-\Delta}^{t}\eta_v e^{\kappa_1(v-(t-\Delta))}dB_v))$.We use Euler approximations for $\Psi_{t,22}$, $\Psi_{t,23}$, and $\Psi_{t,33}$ such that

$$\Psi_t = \begin{pmatrix} \sigma^2\Delta & \sigma^2\Delta & 0 \\ \sigma^2\Delta & \eta_{t-\Delta}^2\Delta/(r_{t-\Delta}^f + \delta + \sigma^2)^2 + \sigma^2\Delta & \eta_{t-\Delta}^2 e^{-\kappa_1\Delta}\Delta/(r_{t-\Delta}^f + \delta + \sigma^2) \\ 0 & \eta_{t-\Delta}^2 e^{-\kappa\Delta}\Delta/(r_{t-\Delta}^f + \delta + \sigma^2) & e^{-2\kappa\Delta}\Delta\eta_{t-\Delta}^2 \end{pmatrix}, \tag{C.5}$$

where in the estimation we simply replace η_t by its proxy η_t^*. Note that this is time-varying, i.e., MEF is strictly more efficient than GMM, and consistency of the parameter estimates is not affected since the approximations only enter the weights.

Using moments m_t^* given by (46), we get the derivatives $(\partial m_t^*/\partial\phi^\top)^\top$ with respect to the parameter vector $\phi = (\kappa, \gamma, \eta_l, \eta_h, \phi_{lh}, \phi_{hl}, \rho, \delta, \sigma)^\top$

$$\begin{pmatrix} 0 & \Delta - \gamma \int_{t-\Delta}^{t} 1/(r_v^f + \delta + \sigma^2) dv & \begin{matrix} -\Delta e^{-\kappa\Delta}(\gamma - \delta - \sigma^2) \\ +\Delta e^{-\kappa\Delta} r_{t-\Delta}^f \end{matrix} \\ 0 & -\kappa \int_{t-\Delta}^{t} 1/(r_v^f + \delta + \sigma^2) dv & -(1 - e^{-\kappa\Delta}) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \Delta & \Delta & 0 \\ 0 & \begin{matrix} \kappa\gamma \int_{t-\Delta}^{t} 1/(r_v^f + \delta + \sigma^2)^2 dv \\ -E\left(\int_{t-\Delta}^{t} \eta_v^2/(r_v^f + \delta + \sigma^2)^3 dv | \mathcal{F}_{t-\Delta} \right) \end{matrix} & 1 - e^{-\kappa\Delta} \\ -\sigma\Delta & E\left(\begin{matrix} -\sigma\Delta \\ +2\sigma\kappa\gamma \int_{t-\Delta}^{t} 1/(r_v^f + \delta + \sigma^2)^2 dv \\ -2\sigma \int_{t-\Delta}^{t} \eta_v^2/(r_v^f + \delta + \sigma^2)^3 dv \end{matrix} \middle| \mathcal{F}_{t-\Delta} \right) & 2\sigma(1 - e^{-\kappa\Delta}) \end{pmatrix}. \tag{C.6}$$

We also use an Euler approximation for the unknown integrals, such that $\psi_t^\top$ reads

$$\begin{pmatrix} 0 & \Delta - \gamma\Delta/(r_{t-\Delta}^f + \delta + \sigma^2) & \begin{matrix} -\Delta e^{-\kappa\Delta}(\gamma - \delta - \sigma^2) \\ +\Delta e^{-\kappa\Delta} r_{t-\Delta}^f \end{matrix} \\ 0 & -\kappa\Delta/(r_{t-\Delta}^f + \delta + \sigma^2) & -(1 - e^{-\kappa\Delta}) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \Delta & \Delta & 0 \\ 0 & \begin{matrix} \kappa\gamma\Delta/(r_{t-\Delta}^f + \delta + \sigma^2)^2 \\ -(\eta_{t-\Delta}^*)^2\Delta/(r_v^f + \delta + \sigma^2)^3 \end{matrix} & 1 - e^{-\kappa\Delta} \\ -\sigma\Delta & \begin{matrix} -\sigma\Delta + 2\sigma\kappa\gamma\Delta/(r_{t-\Delta}^f + \delta + \sigma^2)^2 \\ -2\sigma\Delta(\eta_{t-\Delta}^*)^2/(r_{t-\Delta}^f + \delta + \sigma^2)^3 \end{matrix} & 2\sigma(1 - e^{-\kappa\Delta}) \end{pmatrix}. \tag{C.7}$$

This completes the construction of the estimating function $M_T = \sum_t \psi_t^\top (\Psi_t)^{-1} m_t^*$.

C.2 AK-Vasicek-SV model

In the case of the AK-Vasicek model with stochastic volatility (cf. Section 3.3.3), the system of equilibrium dynamics reads[7]

$$
\begin{aligned}
d\ln C_t &= \left(r_t - \rho - \delta - \tfrac{1}{2}\sigma^2\right) dt + \sigma dZ_t, && \text{(C.8a)}\\
d\ln Y_t &= \left(\kappa_1\gamma_1/r_t - \tfrac{1}{2}(\eta_t/r_t)^2 + r_t - \kappa_1 - \rho - \delta - \tfrac{1}{2}\sigma^2\right) dt + \eta_t/r_t dB_t + \sigma dZ_t, && \text{(C.8b)}\\
dr_t &= \kappa_1(\gamma_1 - r_t)dt + \eta_t dB_t, && \text{(C.8c)}\\
d\log(\eta_t^2) &= \kappa_2(\gamma_2 - \log(\eta_t^2))dt + \xi dW_t. && \text{(C.8d)}
\end{aligned}
$$

Using system (C.8) and the equilibrium asset-pricing condition (14), we obtain

$$
\begin{aligned}
\ln(C_t/C_{t-\Delta}) - \int_{t-\Delta}^{t} r_v^f dv &= -\left(\rho - \tfrac{1}{2}\sigma^2\right)\Delta + \varepsilon_{C,t}, && \text{(C.9a)}\\
\ln(Y_t/Y_{t-\Delta}) - \int_{t-\Delta}^{t} r_v^f dv &= \kappa_1\gamma_1 \int_{t-\Delta}^{t} 1/(r_v^f + \delta + \sigma^2)dv - \tfrac{1}{2}\int_{t-\Delta}^{t} \eta_v^2/(r_v^f + \delta + \sigma^2)^2 dv\\
&\quad - \left(\kappa_1 + \rho - \tfrac{1}{2}\sigma^2\right)\Delta + \varepsilon_{Y,t}, && \text{(C.9b)}\\
r_t^f &= e^{-\kappa_1\Delta} r_{t-\Delta}^f + (1 - e^{-\kappa_1\Delta})(\gamma_1 - \delta - \sigma^2) + \varepsilon_{r,t}, && \text{(C.9c)}\\
\log(\eta_t^2) &= e^{-\kappa_2\Delta}\log(\eta_{t-\Delta}^2) + (1 - e^{-\kappa_2\Delta})\gamma_2 + \varepsilon_{\eta,t}, && \text{(C.9d)}
\end{aligned}
$$

with martingale increments given by

$$
\begin{aligned}
\varepsilon_{C,t} &= \sigma(Z_t - Z_{t-\Delta}), && \text{(C.10a)}\\
\varepsilon_{Y,t} &= \int_{t-\Delta}^{t} \eta_v/(r_v^f + \delta + \sigma^2)dB_v + \sigma(Z_t - Z_{t-\Delta}), && \text{(C.10b)}\\
\varepsilon_{r,t} &= e^{-\kappa_1\Delta}\int_{t-\Delta}^{t} \eta_v e^{\kappa_1(v-(t-\Delta))}dB_v, && \text{(C.10c)}\\
\varepsilon_{\eta,t} &= e^{-\kappa_2\Delta}\int_{t-\Delta}^{t} \xi e^{\kappa_2(v-(t-\Delta))}dW_v. && \text{(C.10d)}
\end{aligned}
$$

We let $m_t = \varepsilon_t = (\varepsilon_{C,t}, \varepsilon_{Y,t}, \varepsilon_{r,t}, \varepsilon_{\eta,t})^\top$ from (C.10a)-(C.10d), i.e., using four moments instead of three. Clearly, m_t is a martingale difference sequence, and from system (C.9) we have that in terms of data and parameters

$$
m_t = \begin{pmatrix}
\ln(C_t/C_{t-\Delta}) - \int_{t-\Delta}^{t} r_v^f dv + \left(\rho - \tfrac{1}{2}\sigma^2\right)\Delta\\
\ln(Y_t/Y_{t-\Delta}) - \int_{t-\Delta}^{t} r_v^f dv + \left(\kappa_1 + \rho - \tfrac{1}{2}\sigma^2\right)\Delta - \kappa_1\gamma_1\int_{t-\Delta}^{t} 1/(r_v^f + \delta + \sigma^2)dv\\
+\tfrac{1}{2}\int_{t-\Delta}^{t} \eta_v^2/(r_v^f + \delta + \sigma^2)^2 dv\\
r_t^f - (1 - e^{-\kappa_1\Delta})(\gamma_1 - \delta - \sigma^2) - e^{-\kappa_1\Delta} r_{t-\Delta}^f\\
\log(\eta_t^2) - (1 - e^{-\kappa_2\Delta})\gamma_2 - e^{-\kappa_2\Delta}\log(\eta_{t-\Delta}^2)
\end{pmatrix}, \tag{C.11}
$$

[7]It can be shown that the analytical solution $C_t = \rho K_t$ is not affected by the presence of stochastic volatility such that the relation between the risk-free rate and the rental rate of capital is still given by (14).

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

Similarly to the case of latent interest rates, our procedure is to derive moments for estimation, say, $m_t^* = E\left(m_t|\mathcal{F}_{t-\Delta}\right)$, given by (48). We may also obtain $\Psi_{t,11} = \sigma^2\Delta$, $\Psi_{t,22} = E_{t-\Delta}(\int_{t-\Delta}^t \eta_v^2/(r_v^f+\delta+\sigma^2)^2 dv) + \sigma^2\Delta$, $\Psi_{t,33} = e^{-2\kappa_1\Delta}E_{t-\Delta}(\int_{t-\Delta}^t \eta_v^2 e^{2\kappa_1(v-(t-\Delta))})dv$, $\Psi_{t,44} = \xi^2(1-e^{-2\kappa_2\Delta})/(2\kappa_2)$, $\Psi_{t,12} = \sigma^2\Delta$, $\Psi_{t,13} = 0$, $\Psi_{t,14} = 0$, $\Psi_{t,23} = E_{t-\Delta}((\int_{t-\Delta}^t \eta_v/(r_v^f+\delta+\sigma^2)dB_v)(e^{-\kappa_1\Delta}\int_{t-\Delta}^t \eta_v e^{\kappa_1(v-(t-\Delta))}dB_v))$, $\Psi_{24}=0$, and $\Psi_{34}=0$. We use Euler approximations for the unknown integrals $\Psi_{t,22}$, $\Psi_{t,23}$ and $\Psi_{t,33}$, such that

$$\Psi_t = \begin{pmatrix} \sigma^2\Delta & \sigma^2\Delta & 0 & 0 \\ \sigma^2\Delta & \eta_{t-\Delta}^2\Delta/(r_{t-\Delta}^f+\delta+\sigma^2)^2+\sigma^2\Delta & \eta_{t-\Delta}^2 e^{-\kappa_1\Delta}\Delta/(r_{t-\Delta}^f+\delta+\sigma^2) & 0 \\ 0 & \eta_{t-\Delta}^2 e^{-\kappa_1\Delta}\Delta/(r_{t-\Delta}^f+\delta+\sigma^2) & e^{-2\kappa_1\Delta}\Delta\eta_{t-\Delta}^2 & 0 \\ 0 & 0 & 0 & \Psi_{t,44} \end{pmatrix}, \tag{C.12}$$

where in the estimation we simply replace η_t by its proxy η_t^*. Again, this is time-varying, and MEF strictly more efficient than GMM.

Using moments m_t given by (C.11), we get the derivatives $(\partial m_t/\partial\phi^\top)^\top$ with respect to the parameter vector $\phi = (\kappa_1,\gamma_1,\kappa_2,\gamma_2,\xi,\rho,\delta,\sigma)^\top$. We also use an Euler approximation for the unknown integrals, such that $\psi_t^\top$ reads

$$\begin{pmatrix} 0 & \Delta-\gamma_1\Delta/(r_{t-\Delta}^f+\delta+\sigma^2) & \begin{array}{c}-\Delta e^{-\kappa_1\Delta}(\gamma_1-\delta-\sigma^2)\\ +\Delta e^{-\kappa_1\Delta}r_{t-\Delta}^f\end{array} & 0 \\ 0 & -\kappa_1\Delta/(r_{t-\Delta}^f+\delta+\sigma^2)dv & -(1-e^{-\kappa_1\Delta}) & 0 \\ 0 & 0 & 0 & \begin{array}{c}-\Delta e^{-\kappa_2\Delta}\gamma_2\\ +\Delta e^{-\kappa_2\Delta}2\log(\eta_{t-\Delta}^*)\end{array} \\ 0 & 0 & 0 & -(1-e^{-\kappa_2\Delta}) \\ 0 & 0 & 0 & 0 \\ \Delta & \Delta & 0 & 0 \\ 0 & \begin{array}{c}\kappa_1\gamma_1\Delta/(r_{t-\Delta}^f+\delta+\sigma^2)^2\\ -(\eta_{t-\Delta}^*)^2\Delta/(r_{t-\Delta}^f+\delta+\sigma^2)^3\end{array} & 1-e^{-\kappa_1\Delta} & 0 \\ -\sigma\Delta & \begin{array}{c}-\sigma\Delta+2\sigma\kappa_1\gamma_1\Delta/(r_{t-\Delta}^f+\delta+\sigma^2)^2\\ -2\sigma(\eta_{t-\Delta}^*)^2\Delta/(r_{t-\Delta}^f+\delta+\sigma^2)^3\end{array} & 2\sigma(1-e^{-\kappa_1\Delta}) & 0 \end{pmatrix}. \tag{C.13}$$

This completes the construction of the estimating function $M_T = \sum_t \psi_t^\top(\Psi_t)^{-1}m_t^*$.

C.3 MEF with five conditional moment restrictions

The 5-vector $m_t = (\varepsilon_{C,t}, \varepsilon_{Y,t}, \varepsilon_{r,t}, \varepsilon^2_{C,t} - \sigma^2\Delta, \varepsilon^2_{r,t} - \eta^2(1-e^{-2\kappa\Delta})/(2\kappa))$ based on the error terms from (16) is clearly a martingale difference, given in terms of data and parameters as

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

or, by using the definition of three moment increments $m_t^{(3)}$ from (21),

$$m_t^{(5)} = \begin{pmatrix} m_t^{(3)} \\ \left(\ln(C_t/C_{t-\Delta}) - \int_{t-\Delta}^t r_v^f dv + \left(\rho - \frac{1}{2}\sigma^2\right)\Delta\right)^2 - \sigma^2\Delta \\ \left(r_t^f - (1-e^{-\kappa\Delta})(\gamma-\delta-\sigma^2) - e^{-\kappa\Delta} r_{t-\Delta}^f\right)^2 - \eta^2(1-e^{-2\kappa\Delta})/(2\kappa) \end{pmatrix} \quad \text{(C.14)}$$

which is equivalent to considering

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

or

$$m_t^{(5)} = \begin{pmatrix} m_t^{(3)} \\ \sigma^2 (Z_t - Z_{t-\Delta})^2 - \sigma^2\Delta \\ \eta^2 e^{-2\kappa\Delta} \left(\int_{t-\Delta}^t e^{\kappa(v-(t-\Delta))} dB_v\right)^2 - \eta^2(1-e^{-2\kappa\Delta})/(2\kappa) \end{pmatrix}. \quad \text{(C.15)}$$

To construct the MEF (27), 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 $\Psi^{(5)}_{t,11} = \sigma^2\Delta$, $\Psi^{(5)}_{t,22} = \eta^2 E_{t-\Delta}(\int_{t-\Delta}^t 1/(r_v^f + \delta + \sigma^2)^2 dv) + \sigma^2\Delta$, and $\Psi^{(5)}_{t,33} = \eta^2(1-e^{-2\kappa\Delta})/(2\kappa)$, $\Psi^{(5)}_{t,44} = 2\sigma^4\Delta^2$, and $\Psi^{(5)}_{t,55} = \eta^4 e^{-4\kappa\Delta} E_{t-\Delta}\left((\int_{t-\Delta}^t e^{\kappa(v-(t-\Delta))} dB_v)^4\right) - \frac{1}{4}\eta^4(1-e^{-2\kappa\Delta})^2/\kappa^2$. Similarly, the conditional covariances are $\Psi^{(5)}_{t,12} = \sigma^2\Delta$, $\Psi^{(5)}_{t,13} = 0$, $\Psi^{(5)}_{t,14} = 0$, $\Psi^{(5)}_{t,15} = 0$, $\Psi^{(5)}_{t,23} = \eta^2 e^{-\kappa\Delta} E_{t-\Delta}\left((\int_{t-\Delta}^t 1/(r_v^f + \delta + \sigma^2) dB_v)(\int_{t-\Delta}^t e^{\kappa(v-(t-\Delta))} dB_v)\right)$, $\Psi^{(5)}_{t,24} = 0$, $\Psi^{(5)}_{t,25} = \eta^3 e^{-2\kappa\Delta} E_{t-\Delta}\left((\int_{t-\Delta}^t 1/(r_v^f + \delta + \sigma^2) dB_v)(\int_{t-\Delta}^t e^{\kappa(v-(t-\Delta))} dB_v)^2\right)$, $\Psi^{(5)}_{t,35} = \eta^3 e^{-3\kappa\Delta} E_{t-\Delta}\left((\int_{t-\Delta}^t e^{\kappa(v-(t-\Delta))} dB_v)^3\right)$, $\Psi^{(5)}_{t,34} = \Psi_{t,45} = 0$. We use Euler approximations for

$\Psi^{(5)}_{t,22}$, $\Psi^{(5)}_{t,55}$, $\Psi^{(5)}_{t,23}$, $\Psi^{(5)}_{t,25}$ and $\Psi^{(5)}_{t,35}$,

$$\Psi^{(5)}_t = \begin{pmatrix} \sigma^2\Delta & \sigma^2\Delta & 0 & 0 & 0 \\ \sigma^2\Delta & \sigma^2\Delta + \eta^2\Delta/(r^f_{t-\Delta}+\delta+\sigma^2)^2 & \eta^2 e^{-\kappa\Delta}\Delta/(r^f_{t-\Delta}+\delta+\sigma^2) & 0 & 0 \\ 0 & \eta^2 e^{-\kappa\Delta}\Delta/(r^f_{t-\Delta}+\delta+\sigma^2) & \frac{1}{2}\eta^2(1-e^{-2\kappa\Delta})/\kappa & 0 & 0 \\ 0 & 0 & 0 & 2\sigma^4\Delta^2 & 0 \\ 0 & 0 & 0 & 0 & \Psi^{(5)\prime}_{t,55} \end{pmatrix}$$

where $\Psi^{(5)\prime}_{t,55} = 3\eta^4 e^{-4\kappa\Delta}\Delta^2 - \frac{1}{4}\eta^4(1-e^{-2\kappa\Delta})^2/\kappa^2$, or using $\Psi^{(3)}_t$ from (41),

$$\Psi^{(5)}_t = \begin{pmatrix} \Psi^{(3)}_t & & 0_{3\times 2} \\ & 2\sigma^4\Delta^2 & 0 \\ 0_{2\times 3} & 0 & 3\eta^4 e^{-4\kappa\Delta}\Delta^2 - \frac{1}{4}\eta^4(1-e^{-2\kappa\Delta})^2/\kappa^2 \end{pmatrix}. \tag{C.16}$$

Again, $\Psi^{(5)}_t$ is time-varying, i.e., this is a conditionally heteroskedastic case, and optimal MEF is strictly more efficient than GMM. Consistency and the expression for the asymptotic variance are unaffected by our approximations because they enter only in the weights (29). Using martingale increments (C.14), we get the derivatives $(\partial m^{(5)}_t(\phi)/\partial\phi^\top)^\top$ with respect to the parameter vector $\phi = (\kappa,\gamma,\eta,\rho,\delta,\sigma)^\top$, such that $(\psi^{(5)}_t)^\top$ reads

$$\begin{pmatrix} & 0 & \frac{1}{2}\eta^2(1-e^{-2\kappa\Delta})/\kappa^2 - \eta^2\Delta e^{-2\kappa\Delta}/\kappa \\ & 0 & 0 \\ (\psi^{(3)}_t)^\top & 0 & -\eta(1-e^{-2\kappa\Delta})/\kappa \\ & 0 & 0 \\ & 0 & 0 \\ & -2\sigma\Delta & 0 \end{pmatrix}, \tag{C.17}$$

with $(\psi^{(3)}_t)^\top$ from (44), and where we use the fact that

$$\psi^{(5)}_{t,44} = 2E_{t-\Delta}\left(\ln(C_t/C_{t-\Delta}) - \int_{t-\Delta}^{t} r^f_v dv + (\rho - \tfrac{1}{2}\sigma^2)\Delta\right)\Delta = 0.$$

This completes the construction of the martingale estimating function for five conditional moment restrictions $M^{(5)}_T = \sum_t(\psi^{(5)}_t)^\top(\Psi^{(5)}_t)^{-1}m^{(5)}_t$.

D Additional Simulation Evidence

D.1 Simulation results: MEF extensions

In this section we present simulation results for two possible MEF extensions, namely, accommodating missing data points in the mixed-frequency approach (MF-MEF), and latent variables using the simulation-based approach (SMEF). Table D1 provides the results for the simulation study of the AK-Vasicek model for the case of truly missing data. In the first

Table D1: 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						
		Monthly Data		Quarterly Data		Mixed Frequency
	DGP	MEF	SMEF	MEF	SMEF	MF-MEF
κ	0.2	0.354	0.355	0.353	0.363	0.360
		0.284	0.280	0.305	0.290	0.290
γ	0.1	0.099	0.099	0.099	0.107	0.099
		0.013	0.012	0.013	0.020	0.013
η	0.01	0.010	0.010	0.010	0.010	0.010
		0.001	0.001	0.001	0.002	0.001
ρ	0.03	0.030	0.030	0.030	0.032	0.030
		0.006	0.002	0.006	0.005	0.006
δ	0.05	0.050	0.051	0.050	0.055	0.050
		0.002	0.005	0.003	0.013	0.002
σ	0.02	0.023	0.021	0.025	0.021	0.022
		0.005	0.003	0.010	0.006	0.005

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. The interest rate data are missing in the SMEF cases, and two of every three monthly output observations are missing in the MF-MEF case. For comparison, we also replicate the complete data MEF results from Table 1 in columns 2 and 4, respectively, using monthly and quarterly data. For the case of observed data (consumption, output and interest rate), but with latent volatility, Table D2 provides the results for regime switching (Panel A) and stochastic volatility (Panel B). 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) from Section 3.3, we compare SMEF (columns 3 and 5 of Table 2) with MEF estimates (columns 2 and 4 of Table D1). 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 the DGP values and are comparable with the MEF figures, with slightly smaller interquartile ranges in the SMEF approach for ρ and σ. This suggests that the model-consistent interest rate proxy $r_t^* = \rho Y_t / C_t$ is particularly fortunate in the AK-Vasicek model. 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).

Table D2: Simulation Study – Regime Switching and Stochastic Volatility

The table reports output of a simulation study of the accuracy of the structural model parameters estimated using the MEF approaches for the AK-Vasicek model with regime switching for the interest rate volatility (Panel A) and with latent stochastic volatility (Panel B). We use three strategies in the latter case for identifying ξ. First, we estimate it along with the other parameters ("MEF"). Second we fix it at the known value ("Fix ξ"). Third, we estimate it separately by looking at the residuals from an autoregressive process for the proxied volatility series and plug this value into the MEF procedure ("Proxy ξ"). 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: Simulation Study – Regime Switching (η_l and η_h)
MEF (Latent Short Rate Volatility)

		Monthly	Quarterly		Monthly	Quarterly
	DGP	MEF	MEF	DGP	MEF	MEF
κ	0.2	0.045	0.032	0.5	0.542	0.587
		0.090	0.128		0.388	0.423
γ	0.1	0.081	0.094	1	1.002	1.009
		0.085	0.092		0.086	0.134
η_l	0.005	0.019	0.016	0.1	0.113	0.118
		0.011	0.013		0.037	0.097
η_h	0.02	0.022	0.023	0.25	0.240	0.225
		0.009	0.019		0.064	0.106
ϕ_{lh}	1.1	1.075	1.034	1	1.045	1.327
		0.533	1.058		0.967	9.023
ϕ_{hl}	1.5	1.403	1.176	5	4.304	3.567
		0.756	1.191		3.471	47.847
ρ	0.03	0.030	0.031	0.03	0.030	0.029
		0.006	0.007		0.006	0.008
δ	0.05	0.080	0.091	0.05	0.049	0.050
		0.104	0.176		0.002	0.033
σ	0.02	0.021	0.021	0.02	0.023	0.027
		0.009	0.029		0.013	0.062

Panel B: Simulation Study – Stochastic Volatility (Latent η_t)
MEF (Latent Short Rate Volatility)

		Monthly			Quarterly		
	DGP	MEF	Fix ξ	Proxy ξ	MEF	Fix ξ	Proxy ξ
κ_1	0.2	0.243	0.250	0.251	0.244	0.268	0.269
		0.201	0.215	0.208	0.253	0.285	0.289
γ_1	0.1	0.099	0.102	0.101	0.103	0.113	0.113
		0.079	0.080	0.080	0.095	0.136	0.134
κ_2	2	2.065	2.173	2.180	0.267	1.299	1.310
		1.084	0.706	0.712	1.588	1.631	1.647
γ_2	-10	-10.033	-10.034	-10.030	-9.891	-9.887	-9.892
		0.413	0.358	0.357	3.469	0.495	0.487
ξ	2.5	4.726	2.500	2.565	326.666	2.500	1.968
		156.150		0.165	6091.628		0.222
ρ	0.03	0.031	0.031	0.031	0.031	0.032	0.032
		0.006	0.006	0.006	0.007	0.007	0.007
δ	0.05	0.047	0.050	0.050	0.050	0.059	0.058
		0.073	0.077	0.075	0.089	0.119	0.110
σ	0.02	0.022	0.022	0.022	0.025	0.028	0.028
		0.027	0.032	0.035	0.052	0.060	0.059

For the latent variable extension, case (ii), we find that the identifiability of the structural parameters largely depends on the calibration of the DGP values (cf. Table D2, Panel A). This is intuitive because the embedded filter may have problems identifying the transition probabilities and/or the size of the two volatility regimes if the difference between them is negligible. One DGP in column 1 is taken roughly in line with the interest rate data, while another DGP in column 4 illustrates that performance improves if the difference between the two regimes is more pronounced. The regime-switching model also has strong implications for the estimate (and bias) of the mean-reversion parameter κ. The point estimates and interquartile ranges for parameters γ, ρ and σ are estimated remarkably close to DGP values. The upward bias in δ in columns 2 and 3 may be explained by a weak identification of η_l. For the case where the two regimes are well identified in columns 5 and 6, all parameter estimates are close to their DGP values. Overall, sampling data at a higher frequency works better.

For the latent variable extension, case (iii), we find that the parameter ξ, the variance of the stochastic volatility process, is weakly identified (cf. Table D2, Panel B). This explains the large interquartile range for the SMEF estimates (columns 2 and 5). Hence, in columns 3 and 7 we fix ξ at its DGP value. We then compare SMEF estimates (columns 4 and 7) with the benchmark estimates when ξ is known. We find that estimating ξ at the outset and using this value as a proxy for ξ works remarkably well. 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 benchmark figures.

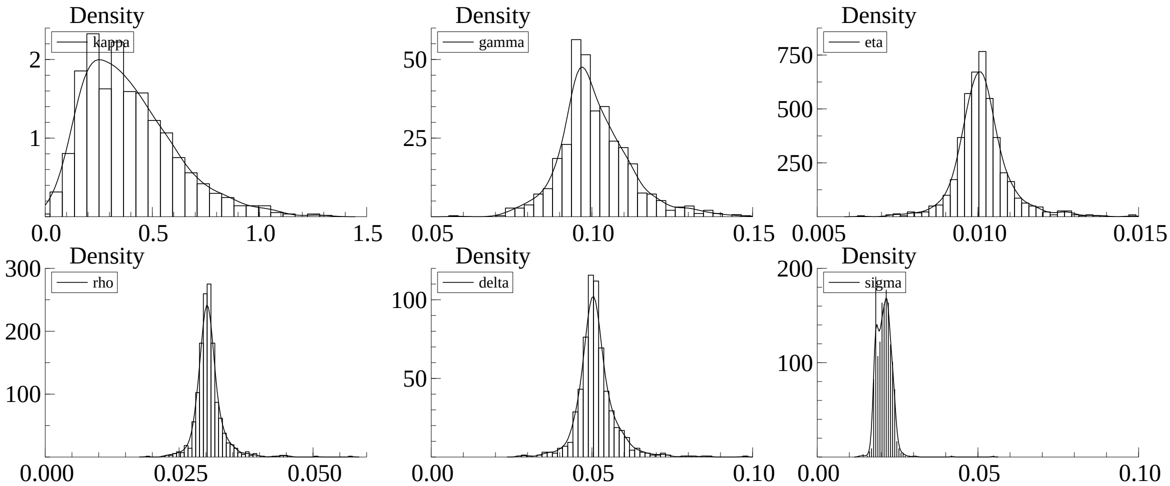
For the extension to mixed-frequency data, case (iv), we compare MF-MEF (column 6) with MEF estimates (columns 2 and 4). As one would expect, given the correct specification, for the case when output is replaced by model consistent predictions at intra-quarter periods the point estimates are remarkably 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.

In Figure D1 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 (Table D1). Comparing the histograms of SMEF to monthly MEF in Figure 1 (both 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.

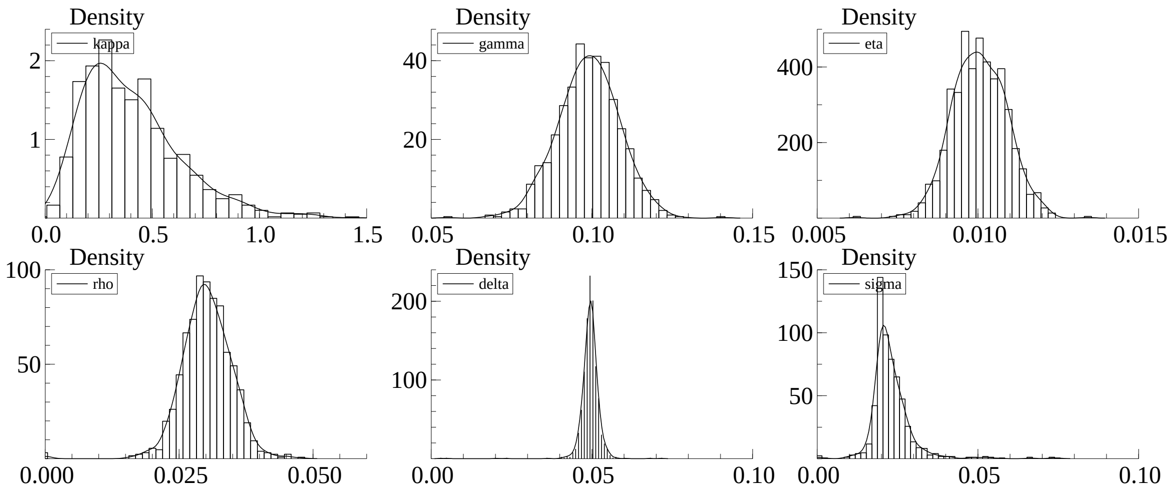
Figure D1: Simulation Study – Latent Short Rate and Mixed Frequency

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 the MF-MEF approach.

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



(B) MF-MEF



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

D.2 Robustness: Time invariance, high-frequency data, and the comparison to discrete time

In this section we present robustness simulation results which are particularly relevant for the estimation of continuous-time models. We want to provide answers to the following three 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) Do 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 observations versus only considering the end-of-period figure helps to identify the parameters in our analysis. (iii) What happens if the true DGP is the continuous-time model and the researcher specifies a discrete-time model, then estimates that system to obtain parameter estimates. Is this problematic?

In order to examine (i), 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 D3 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. This 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 observations is sufficiently large (compare also to the results in Table 1).

Table D3: Robustness Simulations – Time invariance and High-frequency data

The table reports output of two simulation studies of 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 in 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 look like without exploiting the availability of daily interest rates, using instead 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 conditional 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	0.299	0.345	0.354	0.236	0.179	0.246	0.246
		0.286	0.134	0.345	0.284	0.116	0.058	0.145	0.120
γ	0.1	0.201	0.101	0.100	0.100	0.190	0.101	0.101	0.100
		0.036	0.013	0.014	0.013	0.023	0.008	0.009	0.008
η	0.01	0.083	0.008	0.010	0.010	0.065	0.007	0.010	0.010
		0.036	0.004	0.001	0.001	0.019	0.002	0.001	0.001
ρ	0.03	0.080	0.030	0.030	0.030	0.075	0.030	0.030	0.030
		0.015	0.006	0.007	0.006	0.009	0.003	0.004	0.003
δ	0.05	0.05	0.05	0.05	0.050	0.05	0.05	0.05	0.050
					0.002				0.002
σ	0.02	0.317	0.000	0.027	0.023	0.299	0.000	0.018	0.022
		0.040	<0.001	0.047	0.005	0.031	<0.001	0.052	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	0.395	0.164	0.356	0.184	0.387	0.144	0.353
		0.444	0.260	0.164	0.286	0.444	0.287	0.137	0.305
γ	0.1	0.241	0.100	0.099	0.098	0.236	0.100	0.099	0.094
		0.205	0.013	0.017	0.012	0.195	0.013	0.018	0.013
η	0.01	0.077	0.009	0.010	0.010	0.075	0.009	0.010	0.010
		0.105	0.004	0.001	0.001	0.102	0.003	0.002	0.001
ρ	0.03	0.104	0.030	0.030	0.030	0.101	0.030	0.031	0.030
		0.095	0.006	0.006	0.006	0.091	0.006	0.007	0.006
δ	0.05	0.05	0.05	0.05	0.048	0.05	0.05	0.05	0.045
					0.002				0.004
σ	0.02	0.389	0.000	0.000	0.023	0.383	0.000	0.032	0.024
		0.435	0.011	0.038	0.005	0.427	<0.001	0.057	0.010

We also examine (ii), what our results would look like if we did not use the daily availability of interest rates, but only 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 D3 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 changing 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 using the deterministic Taylor expansion (43). Table D4 shows that the use of high-frequency observations may be more important for different models and/or data. In particular, if the speed of mean reversion κ is high, the daily approximations of integrals of financial interest rate data are important for identification of the structural parameters of the macro model, such as the depreciation rate δ. For example, using DGP values $\kappa = 1$ and $\delta = 0.05$, the end-of-quarter approximation of integrals suggests that $\delta = 0$, and at the same time the high-frequency data approximation yields $\delta = 0.049$. These patterns suggest that both high-frequency data and/or more information about the within-period dynamics help identify the parameters of interest.

To examine (iii) we simulate the DGP from the continuous-time system (15) and then estimate the discrete time system (19). It turns out that while the model is misspecified, at a first glance, it seems to be a (surprisingly) good approximation and the structural parameters can be estimated from the simulated data (cf. Table D5). We may directly compare the results to Table 1. We find that the median MEF point estimates are similar in both approaches. A second look, however, reveals that the approximation has strong consequences on the identifiability of structural parameters. As shown by Canova and Sala (2009), many (linear) DSGE models share identification problems for (a subset of) model parameters. Figure D2, which shows the elasticity of the objective function to the parameter values for one draw of the simulated data, suggests that the continuous-time approach, where we use the nonlinear model, may help the identifiability of structural parameters.[8] The objective function is much steeper around parameter estimates in the continuous-time model (Panel A), which implies elasticities at several orders of magnitude higher in the continuous-time

[8] A nonlinear analysis of the discrete-time model requires solving (18) with some nonlinear approximation scheme. Along those lines, we would arrive at an (implicit) dynamic equilibrium system which no longer allows the application of our estimation methods. This is in contrast to our continuous-time approach where an explicit dynamic equilibrium system is obtained using the stochastic calculus.

Table D4: Robustness Simulations – High-frequency Data for Different DGPs

The table reports output of a simulation study of the accuracy of the structural model parameters estimated using the MEF approach for the AK-Vasicek model, to illustrate the benefits of high-frequency data. In each panel we show two types of results. First, estimation where the Riemann sum for the integral is replaced by the end of month ("EoMth") and end of quarter value ("EoQrt"). Second, the usual estimation using the Riemann sum approximation of the integral ("Daily"). 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. In Panel A (a) we replicate the results from Tables 1 and D3. In Panel A (b) we report results for a DGP setting with a relatively high value of η, whereas in Panel B (a) and (b) we report results for relatively high values of both κ and η.

Panel A: Robustness Simulations – Daily vs. Monthly and Quarterly Short Rate (high η)

	(a)	Monthly Data		Quarterly Data		(b)	Monthly Data		Quarterly Data	
	DGP	Daily	EoMth	Daily	EoQrt	DGP	Daily	EoMth	Daily	EoQrt
κ	0.2	0.354	0.356	0.353	0.353	0.2	0.228	0.212	0.228	0.189
		0.284	0.286	0.305	0.305		0.342	0.328	0.342	0.374
γ	0.1	0.099	0.098	0.099	0.094	0.5	0.521	0.521	0.521	0.589
		0.013	0.012	0.013	0.013		0.308	0.707	0.308	42.475
η	0.01	0.010	0.010	0.010	0.010	0.1	0.099	0.099	0.099	0.099
		0.001	0.001	0.001	0.001		0.009	0.009	0.009	0.021
ρ	0.03	0.030	0.030	0.030	0.030	0.03	0.031	0.031	0.031	0.032
		0.006	0.006	0.006	0.006		0.006	0.006	0.006	0.006
δ	0.05	0.050	0.048	0.050	0.045	0.05	0.043	0.037	0.043	0.028
		0.002	0.002	0.003	0.004		0.034	0.028	0.034	0.052
σ	0.02	0.023	0.023	0.025	0.024	0.02	0.022	0.022	0.022	0.020
		0.005	0.005	0.010	0.010		0.007	0.007	0.007	0.772

Panel B: Robustness Simulations – Daily vs. Monthly and Quarterly Short Rate (high κ and η)

	(a)	Monthly Data		Quarterly Data		(b)	Monthly Data		Quarterly Data	
	DGP	Daily	EoMth	Daily	EoQrt	DGP	Daily	EoMth	Daily	EoQrt
κ	0.5	0.754	0.740	0.754	0.724	1	1.129	1.123	1.129	1.162
		0.434	0.463	0.434	0.552		0.436	0.443	0.436	0.985
γ	0.5	0.538	0.532	0.538	0.541	0.5	0.498	0.476	0.498	0.452
		0.125	0.122	0.125	0.217		0.029	0.032	0.029	0.032
η	0.2	0.196	0.197	0.196	0.197	0.1	0.100	0.100	0.100	0.105
		0.019	0.019	0.019	0.042		0.007	0.006	0.007	0.018
ρ	0.03	0.031	0.031	0.031	0.032	0.03	0.031	0.031	0.031	0.031
		0.006	0.006	0.006	0.007		0.006	0.006	0.006	0.006
δ	0.05	0.048	0.038	0.048	0.011	0.05	0.049	0.026	0.049	0.000
		0.038	0.042	0.038	0.053		0.011	0.013	0.011	0.001
σ	0.02	0.022	0.022	0.022	0.020	0.02	0.022	0.022	0.022	0.001
		0.007	0.010	0.007	0.355		0.003	0.004	0.003	0.018

approach (Panel B). Moreover, within-period dynamics or the mixed-frequency property of macro and financial data can no longer be exploited (as discussed above). Hence, in models where the within-period dynamics are economically relevant and/or the nonlinearities are economically important, this will probably also show up in forecasting performance. Ultimately, it will be an empirical question whether these features matter in more elaborate models.

Our conclusion from the robustness analysis is that if parameters are well identified, a (log-linear) approximation of a more elaborate continuous-time model where no analytical solution is available seems a promising route and may be the best-practice approach. While it easily allows using mixed-frequency data, it keeps efficiency losses when estimating the model at a minimum. We leave further analysis of this for future research.

Table D5: Simulation Study – Monthly and Quarterly Data Discrete Time

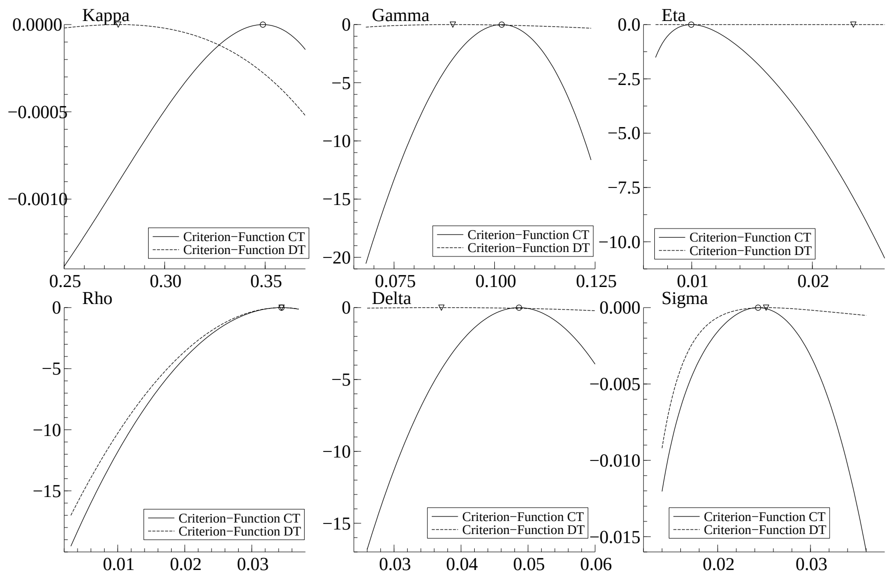
The table reports output of a simulation study of the accuracy of the structural model parameters estimated using the OLS, FGLS-SUR-IV, GMM, and MEF approaches to the AK-Vasicek model in Discrete Time. For 1,000 replications, we generate 25 years of data from the underlying continuous time data generating 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	0.361	0.135	0.365	0.282	0.366	0.114	0.385	0.335
		0.280	0.084	0.268	0.249	0.286	0.056	0.310	0.279
γ	0.100	0.103	0.101	0.099	0.096	0.107	0.103	0.108	0.105
		0.016	0.015	0.013	0.015	0.019	0.018	0.017	0.023
η	0.010	0.010	0.010	0.010	0.012	0.010	0.010	0.010	0.011
		0.010	0.010	0.000	0.014	0.011	0.010	0.000	0.008
ρ	0.030	0.034	0.033	0.031	0.030	0.039	0.034	0.039	0.036
		0.011	0.009	0.007	0.006	0.017	0.012	0.012	0.013
δ	0.050	0.050	0.050	0.050	0.047	0.050	0.050	0.050	0.050
					0.011				0.008
σ	0.020	0.072	0.034	0.020	0.020	0.134	0.067	0.133	0.109
		0.135	0.117	0.001	0.002	0.187	0.150	0.077	0.136

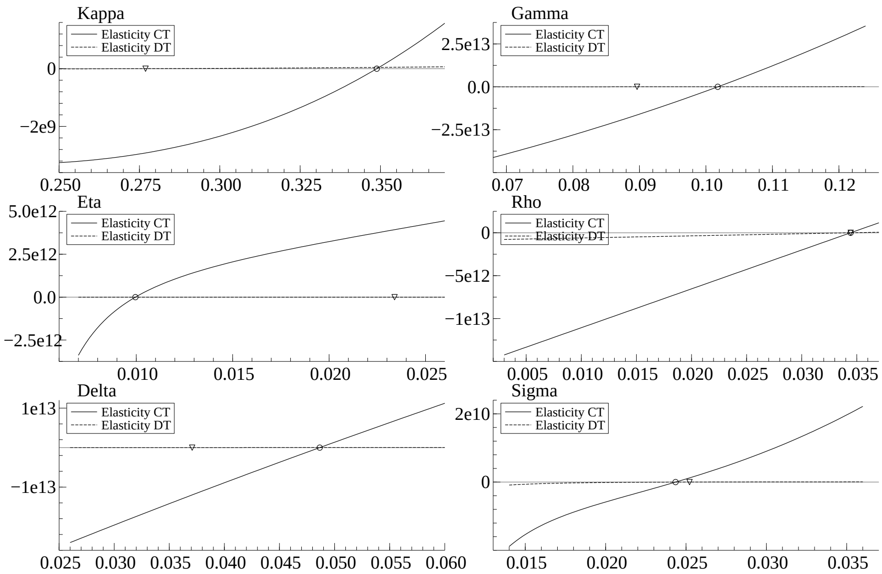
Figure D2: Simulation Study – Objective Function with Elasticity

The figure reports the elasticity of the objective function to the parameter values for both the continuous time and discrete time models. For each model, one parameter is varied over the range of the horizontal axis while the other parameters are fixed at the estimated values for each method. Both the criterion function and elasticity (percentage change of objective function divided by percentage change of parameter value) are reported, where the objective function is -1 times the inner product of the elements of $M_T(\phi)$. The figure is an illustration for 1 of the 1,000 replications in the simulation study, with generated 25 years of data from the underlying data generating process (DGP) in the case of monthly data. Panel A plots the criterion function for both the continuous (solid line) and discrete time (dashed line) models, where the dot (reverse triangle) is the estimated parameter value. Panel B plots the elasticity for both the continuous (solid line) and discrete time (dashed line) models, where the dot (reverse triangle) is the estimated parameter value.

(A) Objective Function – Continuous Time (CT) and Discrete Time (DT)



(B) Objective Function Elasticity – Continuous Time (CT) and Discrete Time (DT)



E Appendix Tables

Table E1: 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 one parameter at a time and consider two settings for each, one value lower than the one used in the baseline setting, and one value higher than that of the baseline DGP setting (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 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	0.201	0.083	0.080	0.050	0.317
	0.286	0.036	0.035	0.015		0.040
DGP with $\kappa = 0.1$	0.272	0.200	0.070	0.079	0.050	0.313
	0.251	0.055	0.038	0.019		0.058
DGP with $\kappa = 0.5$	0.628	0.201	0.112	0.081	0.050	0.319
	0.348	0.018	0.033	0.009		0.024
DGP with $\gamma = 0.05$	0.325	0.087	0.033	0.049	0.050	0.193
	0.283	0.036	0.023	0.014		0.063
DGP with $\gamma = 0.2$	0.354	0.412	0.174	0.135	0.050	0.460
	0.287	0.067	0.070	0.033		0.067
DGP with $\eta = 0.005$	0.354	0.206	0.087	0.083	0.050	0.325
	0.286	0.034	0.035	0.017		0.047
DGP with $\eta = 0.05$	0.175	0.054	0.002	0.032	0.050	0.000
	0.310	0.221	0.040	0.009		0.094
DGP with $\rho = 0.01$	0.349	0.201	0.083	0.060	0.050	0.317
	0.286	0.036	0.035	0.015		0.040
DGP with $\rho = 0.1$	0.349	0.201	0.083	0.150	0.050	0.317
	0.286	0.036	0.035	0.015		0.040
DGP with $\delta = 0.01$	0.349	0.201	0.083	0.080	0.010	0.317
	0.286	0.036	0.035	0.015		0.040
DGP with $\delta = 0.1$	0.349	0.201	0.083	0.080	0.100	0.317
	0.286	0.036	0.035	0.015		0.040
DGP with $\sigma = 0.01$	0.348	0.201	0.083	0.080	0.050	0.317
	0.284	0.034	0.035	0.012		0.037
DGP with $\sigma = 0.05$	0.351	0.200	0.083	0.080	0.050	0.319
	0.289	0.045	0.038	0.026		0.061

Table E1, 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	0.101	0.008	0.030	0.050	0.000
	0.134	0.013	0.004	0.006		<0.001
DGP with $\kappa = 0.1$	0.263	0.101	0.009	0.030	0.050	0.000
	0.146	0.022	0.004	0.006		<0.001
DGP with $\kappa = 0.5$	0.403	0.100	0.006	0.030	0.050	0.000
	0.150	0.006	0.003	0.006		<0.001
DGP with $\gamma = 0.05$	0.404	0.051	0.008	0.030	0.050	0.000
	0.214	0.012	0.003	0.006		<0.001
DGP with $\gamma = 0.2$	0.220	0.201	0.006	0.030	0.050	0.000
	0.117	0.014	0.006	0.006		0.020
DGP with $\eta = 0.005$	0.220	0.100	0.002	0.030	0.050	0.012
	0.117	0.007	0.005	0.006		0.020
DGP with $\eta = 0.05$	0.541	0.151	0.029	0.030	0.050	0.000
	0.305	0.045	0.038	0.006		<0.001
DGP with $\rho = 0.01$	0.299	0.101	0.008	0.010	0.050	0.000
	0.134	0.013	0.004	0.006		<0.001
DGP with $\rho = 0.1$	0.299	0.101	0.008	0.100	0.050	0.000
	0.134	0.013	0.004	0.006		<0.001
DGP with $\delta = 0.01$	0.299	0.101	0.008	0.030	0.010	0.000
	0.134	0.013	0.004	0.006		<0.001
DGP with $\delta = 0.1$	0.299	0.101	0.008	0.030	0.100	0.000
	0.134	0.013	0.004	0.006		<0.001
DGP with $\sigma = 0.01$	0.298	0.101	0.009	0.030	0.050	0.000
	0.134	0.013	0.003	0.003		<0.001
DGP with $\sigma = 0.05$	0.301	0.100	0.000	0.030	0.050	0.037
	0.136	0.013	0.007	0.014		0.020

Table E1, 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	0.100	0.010	0.031	0.050	0.027
	0.345	0.014	0.001	0.007		0.047
DGP with $\kappa = 0.1$	0.265	0.100	0.010	0.031	0.050	0.026
	0.312	0.024	0.001	0.007		0.049
DGP with $\kappa = 0.5$	0.594	0.101	0.010	0.031	0.050	0.025
	0.449	0.006	0.001	0.007		0.044
DGP with $\gamma = 0.05$	0.335	0.053	0.010	0.031	0.050	0.033
	0.327	0.013	0.001	0.007		0.062
DGP with $\gamma = 0.2$	0.290	0.200	0.010	0.031	0.050	0.033
	0.312	0.015	0.001	0.007		0.048
DGP with $\eta = 0.005$	0.330	0.100	0.005	0.030	0.050	0.022
	0.319	0.007	<0.001	0.006		0.038
DGP with $\eta = 0.05$	0.352	0.197	0.050	0.042	0.050	0.157
	0.351	0.554	0.050	0.265		0.678
DGP with $\rho = 0.01$	0.344	0.100	0.010	0.011	0.050	0.027
	0.348	0.014	0.001	0.007		0.047
DGP with $\rho = 0.1$	0.345	0.100	0.010	0.101	0.050	0.027
	0.345	0.014	0.001	0.007		0.047
DGP with $\delta = 0.01$	0.344	0.100	0.010	0.031	0.010	0.027
	0.345	0.014	0.001	0.007		0.047
DGP with $\delta = 0.1$	0.344	0.100	0.010	0.031	0.100	0.027
	0.346	0.014	0.001	0.007		0.047
DGP with $\sigma = 0.01$	0.352	0.101	0.010	0.031	0.050	0.020
	0.361	0.014	0.001	0.003		0.044
DGP with $\sigma = 0.05$	0.344	0.100	0.010	0.031	0.050	0.052
	0.325	0.015	0.001	0.017		0.032

Table E1, 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 E2: Simulation Study – Monthly Data with Bias Correction

The table reports output of a simulation study of the accuracy of the structural model parameters estimated using the MEF approach to the AK-Vasicek model, where bias correction methods are applied. We apply the formulas from Yu (2012, eq. (17)), Tang and Chen (2009, Theorem 3.1.1), and two bootstrap methods inspired by Tang and Chen (2009, Section 4), where we bias correct based on both the mean and median simulated bias. For 1,000 replications, we generate 25 years of data from the underlying continuous time data generating process (DGP) and apply our estimation strategy. We show the median estimate, and provide the interquartile range below it.

Parameter Estimates from Simulation Study – Monthly Data with Bias Correction Methods						
			Yu (2012)	Tang and Chen (2009)	Bootstrapped	
	DGP	MEF	(17)	Theorem 3.1.1	Mean	Median
κ	0.200	0.355	0.278	0.192	0.150	0.204
		0.285	0.280	0.283	0.306	0.313
γ	0.100	0.099			0.099	0.099
		0.013			0.013	0.013
η	0.010	0.010		0.010	0.010	0.010
		0.001		0.001	0.001	0.001
ρ	0.030	0.030			0.030	0.030
		0.006			0.006	0.005
δ	0.050	0.050			0.050	0.050
		0.002			0.002	0.002
σ	0.020	0.023			0.020	0.022
		0.005			0.006	0.005

Table E3: Simulation Study – Variance Terms and Five Conditional Moment Restrictions

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

Table E4: Simulation Study – Sensitivity of Regression-Based Methods to δ_0 and σ_0

The table reports output of a simulation study of 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 the values 0.01, 0.05 (base), and 0.10, and for σ_0, we consider the values 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, and Panel B that of the FGLS-SUR-IV method. Each panel consists of nine 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

		$\delta_0 = 0.01$			$\delta_0 = 0.05$			$\delta_0 = 0.10$		
	DGP	$\sigma_0 = 0.01$	$\sigma_0 = 0.02$	$\sigma_0 = 0.05$	$\sigma_0 = 0.01$	$\sigma_0 = 0.02$	$\sigma_0 = 0.05$	$\sigma_0 = 0.01$	$\sigma_0 = 0.02$	$\sigma_0 = 0.05$
κ	0.200	0.292	0.295	0.305	0.349	0.349	0.350	0.354	0.354	0.354
		0.281	0.282	0.286	0.286	0.286	0.286	0.281	0.281	0.282
γ	0.100	0.108	0.109	0.114	0.201	0.201	0.206	0.312	0.313	0.317
		0.034	0.034	0.033	0.036	0.036	0.036	0.053	0.054	0.054
η	0.010	0.039	0.040	0.042	0.083	0.083	0.085	0.131	0.131	0.133
		0.026	0.026	0.026	0.035	0.035	0.036	0.051	0.051	0.052
ρ	0.030	0.054	0.055	0.057	0.080	0.080	0.083	0.111	0.112	0.114
		0.012	0.012	0.012	0.015	0.015	0.015	0.024	0.024	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	0.224	0.234	0.316	0.317	0.324	0.401	0.402	0.408
		0.051	0.050	0.047	0.041	0.040	0.040	0.057	0.057	0.058

Panel B: Parameter Estimates from Simulation Study – FGLS-SUR-IV Sensitivity to δ_0 and σ_0

		$\delta_0 = 0.01$			$\delta_0 = 0.05$			$\delta_0 = 0.10$		
	DGP	$\sigma_0 = 0.01$	$\sigma_0 = 0.02$	$\sigma_0 = 0.05$	$\sigma_0 = 0.01$	$\sigma_0 = 0.02$	$\sigma_0 = 0.05$	$\sigma_0 = 0.01$	$\sigma_0 = 0.02$	$\sigma_0 = 0.05$
κ	0.200	0.196	0.201	0.223	0.299	0.299	0.297	0.109	0.108	0.104
		0.210	0.211	0.201	0.134	0.134	0.134	0.058	0.058	0.058
γ	0.100	0.067	0.068	0.071	0.101	0.101	0.102	0.161	0.161	0.163
		0.041	0.039	0.031	0.013	0.013	0.013	0.019	0.019	0.020
η	0.010	0.018	0.019	0.021	0.009	0.008	0.000	0.000	0.000	0.000
		0.026	0.024	0.020	0.003	0.004	0.004	$<$0.001	$<$0.001	$<$0.001
ρ	0.030	0.038	0.038	0.039	0.030	0.030	0.031	0.032	0.032	0.033
		0.008	0.008	0.008	0.006	0.006	0.006	0.006	0.006	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	0.133	0.140	0.000	0.000	0.036	0.058	0.061	0.077
		0.033	0.032	0.028	$<$0.001	$<$0.001	0.018	0.042	0.040	0.033

Table E5: 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, and for GMM and MEF, we use five conditional 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, through 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	0.083	0.030	0.069	0.114	0.065	0.045	0.048
	0.436	0.270	0.185	0.679	1.064	2.083	0.769	0.697
γ	0.101	0.101	0.045	0.108	0.134	0.130	0.089	0.098
	4.002	1.671	0.186	0.602	2.715	2.329	2.054	0.924
η	0.018	0.018	0.005	0.007	0.028	0.019	0.000	0.007
	1.284	0.444	0.669	0.051	0.693	0.608	<0.001	0.097
ρ	0.015	0.015	0.006	0.004	0.022	0.021	0.009	0.020
	0.441	1.139	0.153	0.089	0.672	0.957	0.444	0.538
δ	0.098	0.106	0.050	0.081	0.128	0.153	0.050	0.040
	1.298	1.443		0.243	0.732	0.682		0.173
σ	0.018	0.018	0.014	0.018	0.018	0.017	0.000	0.019
	0.065	0.082	0.994	0.008	0.078	0.003	<0.001	0.010

Table E6: 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 three and five moment conditions for the regular MEF and iterated approach, for two iterations. We show the median estimate, and provide the interquartile range below it.

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

References

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

Canova, F., and L. Sala (2009): "Back to square one: Identification issues in DSGE models," *J. Monet. Econ.*, 56, 431–449.

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

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

Ross, S. M. (2014): *Introduction to Probability Models.* Elsevier, 11th ed., Amsterdam.

Tang, C. Y., and S. X. Chen (2009): "Parameter estimation and bias correction for diffusion processes," *J. Econometrics*, 149, 65–81.

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.