Online Appendix:
Full-Information Estimation of Heterogeneous Agent Models Using Macro and Micro Data

Laura Liu Mikkel Plagborg-Møller

June 6, 2022

Appendix B Variance-covariance matrix for moment-based methods

Here we describe how we estimate the variance-covariance matrix of the cross-sectional moments when implementing the moment-based inference approaches in Section 4.4. We take the "3rd Moment" inference approach as an example. The measurement error variance-covariance matrices of other moment-based approaches can be derived in a similar fashion.

Let $\hat{m}_{\epsilon j,t}$, for $\epsilon = 0, 1$ and $j = 1, 2, 3$, be the cross-sectional sample moments of household after-tax income in period t, and $m_{\epsilon j,t}$ be the corresponding population moments, where ϵ indicates the employment status of the group and j represents the order of the moment, such as the sample mean, variance, and third central moment. For instance,

$$\hat{m}_{11,t} = \frac{\sum_{i=1}^N \iota_{i,t}\epsilon_{i,t}}{\sum_{i=1}^N \epsilon_{i,t}}, \quad m_{11,t} = \mathbb{E}[\iota_{i,t} \mid \epsilon_{i,t} = 1,\ z_t],$$
$$\hat{m}_{1j,t} = \frac{\sum_{i=1}^N (\iota_{i,t} - \hat{m}_{11,t})^j \epsilon_{i,t}}{\sum_{i=1}^N \epsilon_{i,t}}, \quad m_{1j,t} = \mathbb{E}[(\iota_{i,t} - m_{11,t})^j \mid \epsilon_{i,t} = 1,\ z_t], \quad \text{for } j > 1.$$

Define $\hat{m}_t \equiv (\hat{m}_{01,t}, \hat{m}_{02,t}, \hat{m}_{03,t}, \hat{m}_{11,t}, \hat{m}_{12,t}, \hat{m}_{13,t})'$. To construct the measurement error variance-covariance matrix $\mathbb{V}[\hat{m}_t \mid z_t]$, we need to compute the variances and covariances across $\hat{m}_{\epsilon j,t}$'s. It is easy to see that $\hat{m}_{0j,t}$ and $\hat{m}_{1k,t}$ are asymptotically independent as $N \to \infty$ for any moment orders j, k, so we can focus on deriving the diagonal blocks where the moments share the same employment status ϵ.

Let us first consider the variance of $\hat{m}_{11,t}$. As $(\iota_{i,t}, \epsilon_{i,t})$ is cross-sectionally i.i.d. given z_t, we resort to the Central Limit Theorem and Slutsky's theorem and obtain[1]

$$\mathbb{V}[\hat{m}_{11,t} \mid z_t] = \frac{\mathbb{V}[\iota_{i,t} \mid \epsilon_{i,t} = 1,\ z_t]}{NL} + o_p\left(N^{-1}\right). \tag{B.1}$$

The sample analog of $\mathbb{V}[\iota_{i,t} \mid \epsilon_{i,t} = 1,\ z_t]$ is $\hat{m}_{12,t}$, the sample variance of the employed group in period t. As explained in the main text, we assume that the variance-covariance matrix of the moments is constant across time and estimate it using full-sample sample moments (i.e., averaging across time). Thus, the numerator in (B.1) is approximated by $\hat{m}_{12} \equiv \frac{1}{|\mathcal{T}|}\sum_{t\in\mathcal{T}} \hat{m}_{12,t}$, where $\mathcal{T}$ is the subset of time points we observe the micro data, and $|\mathcal{T}|$ gives the number of elements in set $\mathcal{T}$.[2] Similarly, the denominator in (B.1) can be approximated by $\hat{N}_1 \equiv \frac{1}{|\mathcal{T}|}\sum_{t\in\mathcal{T}}\sum_{i=1}^{N} \epsilon_{i,t}$.

For other elements in the "employed" block, we have

$$\begin{aligned}
\mathbb{V}[\hat{m}_{12,t} \mid z_t] &= \frac{\mathbb{V}[(\iota_{i,t} - m_{11})^2 \mid \epsilon_{i,t} = 1,\ z_t]}{NL} + o_p\left(N^{-1}\right) \approx \frac{\hat{m}_{14} - \hat{m}_{12}^2}{\hat{N}_1}, \\
\mathbb{V}[\hat{m}_{13,t} \mid z_t] &= \frac{\mathbb{V}[(\iota_{i,t} - m_{11})^3 \mid \epsilon_{i,t} = 1,\ z_t]}{NL} + o_p\left(N^{-1}\right) \approx \frac{\hat{m}_{16} - 6\hat{m}_{14}\hat{m}_{12} - \hat{m}_{13}^2 + 9\hat{m}_{12}^3}{\hat{N}_1}, \\
\mathbb{C}\text{ov}[\hat{m}_{11,t}, \hat{m}_{12,t} \mid z_t] &= \frac{\text{cov}[\iota_{i,t} - m_{11},\ (\iota_{i,t} - m_{11})^2 \mid \epsilon_{i,t} = 1,\ z_t]}{NL} + o_p\left(N^{-1}\right) \approx \frac{\hat{m}_{13}}{\hat{N}_1}, \\
\mathbb{C}\text{ov}[\hat{m}_{11,t}, \hat{m}_{13,t} \mid z_t] &= \frac{\text{cov}[\iota_{i,t} - m_{11},\ (\iota_{i,t} - m_{11})^3 \mid \epsilon_{i,t} = 1,\ z_t]}{NL} + o_p\left(N^{-1}\right) \approx \frac{\hat{m}_{14} - 3\hat{m}_{12}^2}{\hat{N}_1}, \\
\mathbb{C}\text{ov}[\hat{m}_{12,t}, \hat{m}_{13,t} \mid z_t] &= \frac{\text{cov}[(\iota_{i,t} - m_{11})^2,\ (\iota_{i,t} - m_{11})^3 \mid \epsilon_{i,t} = 1,\ z_t]}{NL} + o_p\left(N^{-1}\right) \\
&\approx \frac{\hat{m}_{15} - 4\hat{m}_{13}\hat{m}_{12}}{\hat{N}_1}.
\end{aligned}$$

In each equation, the first equality is given by a similar cross-sectionally i.i.d. argument. The second approximation rewrites the variance/covariance in the numerator in terms of population moments (as in Fisher, 1930, but omitting inconsequential degree-of-freedom adjustments), and then substitutes these population moments with their sample analogs averaged over time. Note that the last terms in the equations above call for even higher-order sample moments. Specifically, to approximate the variance/covariance involving the

[1] In the denominator, $\frac{1}{N}\sum_{i=1}^{N} \epsilon_{i,t} \xrightarrow{p} \mathbb{E}[\epsilon_{i,t}]$ as $N \to \infty$. Recall that $\mathbb{E}[\epsilon_{i,t}] = L$ for all t.

[2] In our numerical experiment, the micro sample size $N_t = N$ is constant over time. If this were not the case, $\hat{m}_{12}$ should be constructed using sample size weights.

m-th and n-th order sample moments, we need sample moments up to the $(m+n)$-th order.

For the "unemployed" block, we can replace $\hat{m}_{1j,t},\ \epsilon_{i,t} = 1,\ L,\ \hat{m}_{1j}$, and $\hat{N}_1$ with $\hat{m}_{0j,t},\ \epsilon_{i,t} = 0,\ 1-L,\ \hat{m}_{0j}$, and $\hat{N}_0 \equiv \frac{1}{|\mathcal{T}|}\sum_{t\in\mathcal{T}}\sum_{i=1}^{N}(1-\epsilon_{i,t})$, respectively.

Combining all steps above, we can approximate the measurement error variance-covariance matrix using sample moments of micro data:

$$\mathbb{V}\left[\hat{m}_t \mid z_t\right] \approx \begin{pmatrix} V_{00} & 0_{3\times 3} \\ 0_{3\times 3} & V_{11} \end{pmatrix},$$

$$V_{00} \equiv \frac{1}{\hat{N}_0}\begin{pmatrix} \hat{m}_{02} & \hat{m}_{03} & \hat{m}_{04} - 3\hat{m}_{02}^2 \\ \hat{m}_{03} & \hat{m}_{04} - \hat{m}_{02}^2 & \hat{m}_{05} - 4\hat{m}_{03}\hat{m}_{02} \\ \hat{m}_{04} - 3\hat{m}_{02}^2 & \hat{m}_{05} - 4\hat{m}_{03}\hat{m}_{02} & \substack{\hat{m}_{06}-6\hat{m}_{04}\hat{m}_{02} \\ -\hat{m}_{03}^2+9\hat{m}_{02}^3} \end{pmatrix},$$

$$V_{11} \equiv \frac{1}{\hat{N}_1}\begin{pmatrix} \hat{m}_{12} & \hat{m}_{13} & \hat{m}_{14} - 3\hat{m}_{12}^2 \\ \hat{m}_{13} & \hat{m}_{14} - \hat{m}_{12}^2 & \hat{m}_{15} - 4\hat{m}_{13}\hat{m}_{12} \\ \hat{m}_{14} - 3\hat{m}_{12}^2 & \hat{m}_{15} - 4\hat{m}_{13}\hat{m}_{12} & \substack{\hat{m}_{16}-6\hat{m}_{14}\hat{m}_{12} \\ -\hat{m}_{13}^2+9\hat{m}_{12}^3} \end{pmatrix}.$$

Heterogeneous household model: Parameter calibration

β	Discount factor	0.96	$\pi(0 \to 1)$	U to E trans.	0.5
α	Capital share	0.36	$\pi(1 \to 0)$	E to U trans.	0.038
δ	Capital depreciation	0.10	ρ_ζ	Agg. TFP AR(1)	0.859
b	UI replacement rate	0.15	σ_ζ	Agg. TFP AR(1)	0.014
μ_λ	Idiosyncratic distr.	-0.25	σ_e	Meas. err. in output	0.02

Table C.1: Parameter calibration in the heterogeneous household model.

Appendix C Heterogeneous household model

C.1 Calibration

Table C.1 shows the parameter calibration used to simulate the data. Here $\pi(0 \to 1)$, for example, denotes the idiosyncratic Markov transition probability $P(\epsilon_{i,t+1} = 1 \mid \epsilon_{i,t} = 0)$.

C.2 Additional simulation results

Here we provide additional results for the numerical illustration of the heterogeneous household model. Figure C.1 shows the full-information and macro-only posterior distributions of the steady-state consumption policy function for unemployed households. Figure C.2 shows the full-information and macro-only posterior distributions of the impulse response function of the asset distribution for unemployed households with respect to a TFP shock. In terms of the comparison between full-information and macro-only inference, both these figures are qualitatively similar to those for employed households, cf. Figures 3 and 4 in the main paper.

Figure C.3 shows the full-information and macro-only posterior densities of the model parameters in an alternative simulation where we only observe $N = 100$ micro draws every ten periods (instead of $N = 1000$). All other settings are the same as in Section 4. Naturally, the accuracy of posterior inference is affected by the smaller sample size, but we see that the individual heterogeneity parameter μ_λ is still precisely estimated in this simulation.

Het. household model: Consumption policy function, unemployed

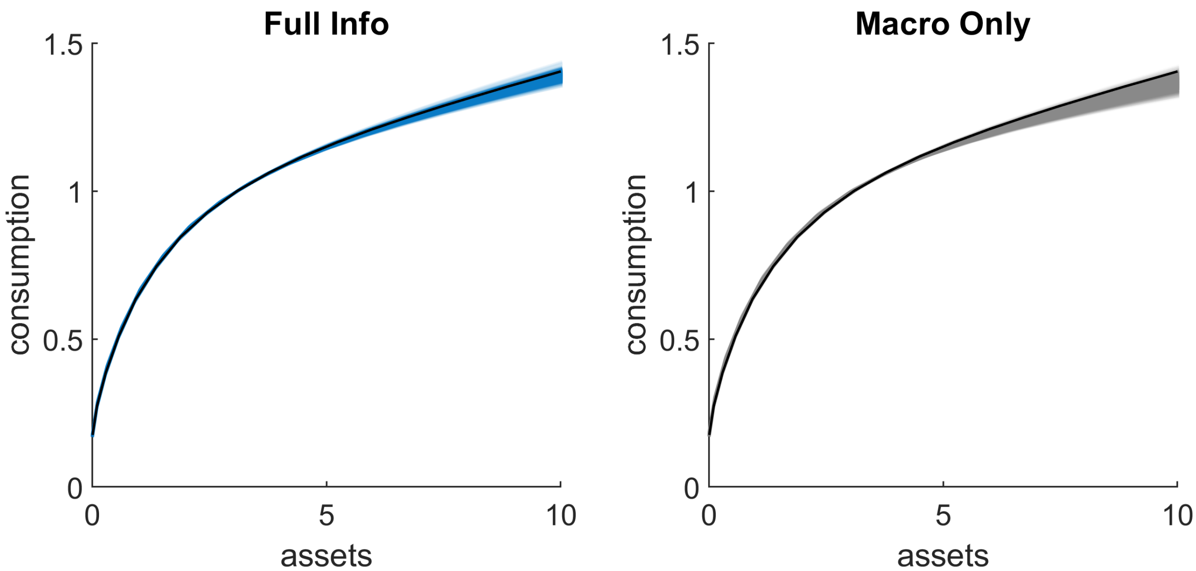

Figure C.1: Posterior draws of steady-state consumption policy function for unemployed households. See caption for Figure 3.

Het. household model: Impulse responses of asset distribution, unemployed

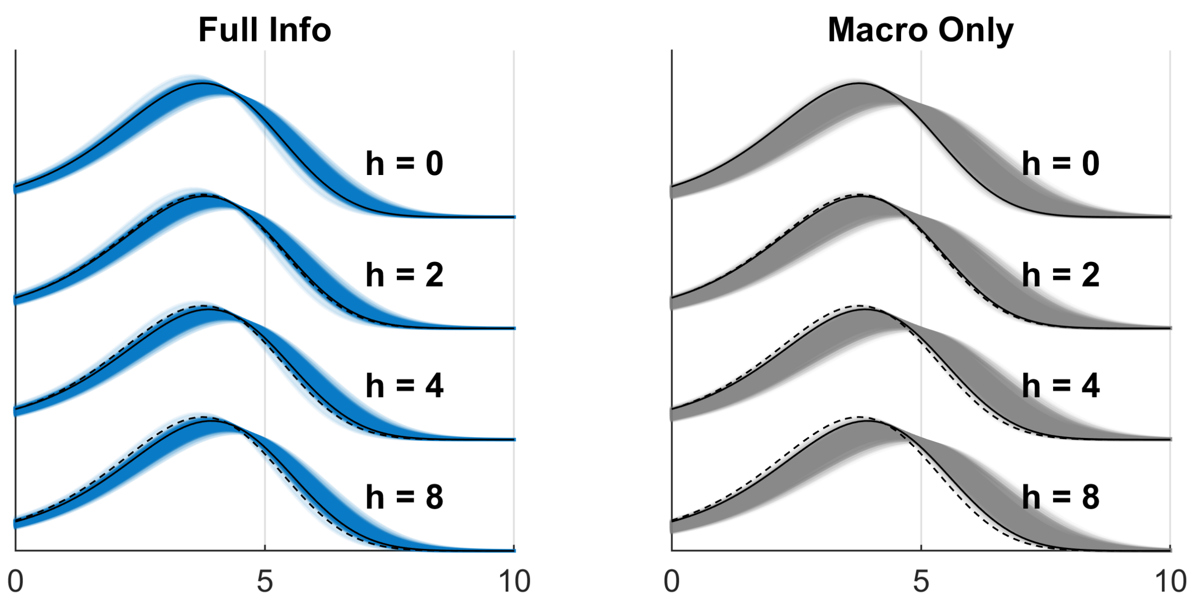

Figure C.2: Posterior of impulse response function of unemployed households' asset distribution with respect to an aggregate productivity shock. See caption for Figure 4.

Heterogeneous household model: Posterior density, $N = 100$

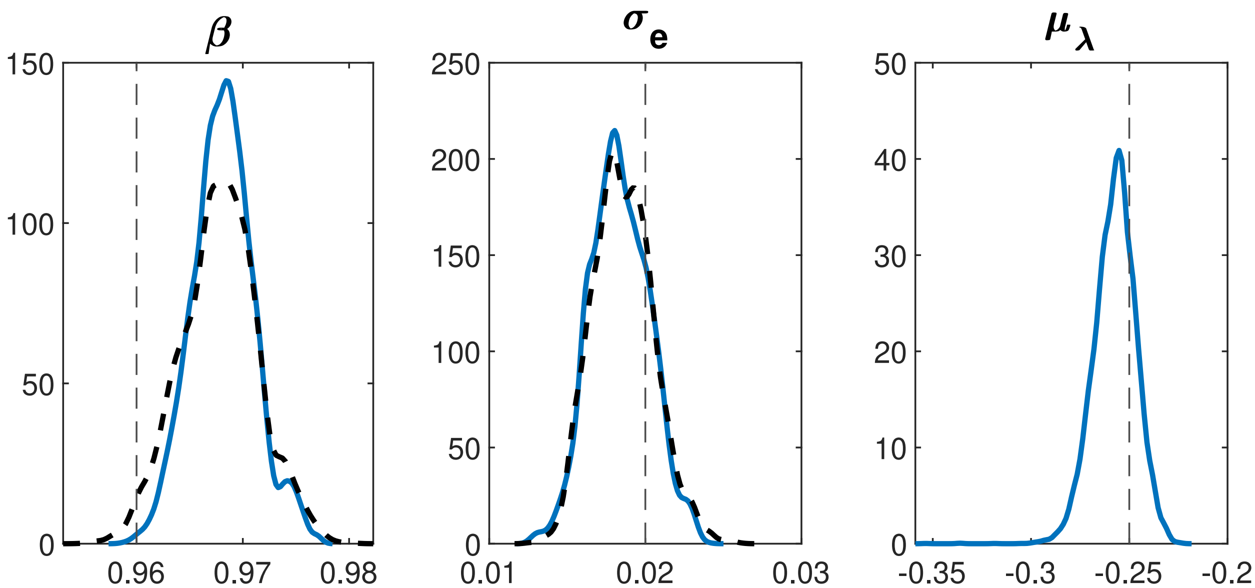

Figure C.3: Posterior densities with (blue solid curves) and without (black dashed curves) conditioning on the micro data, for cross-sectional sample size $N = 100$. See caption for Figure 2.

Appendix D Heterogeneous firm model

D.1 Model assumptions

We here briefly describe the assumptions of the heterogeneous firm model. See Khan and Thomas (2008) and Winberry (2018) for more complete discussions of the model. Note that the notation in this section recycles some of the notation used for the household model in Section 2.2.

A unit mass of firms $i \in [0,1]$ have decreasing returns to scale production functions $Y_{i,t} = e^{\zeta_t + \epsilon_{i,t}} k_{i,t}^{\alpha} n_{i,t}^{\nu}$, where $k_{i,t}$ and $n_{i,t}$ denote firm-specific capital and labor inputs ($\alpha + \nu < 1$). Labor $n_{i,t}$ is hired in a competitive labor market with aggregate wage rate w_t. Aggregate log TFP ζ_t evolves as an AR(1) process $\zeta_t = \rho_\zeta \zeta_{t-1} + \varepsilon_{\zeta,t}$, $\varepsilon_{\zeta,t} \overset{i.i.d.}{\sim} N(0, \sigma_\zeta^2)$. The firm-specific log productivity levels evolve as independent AR(1) processes $\epsilon_{i,t} = \rho_\epsilon \epsilon_{i,t-1} + \omega_{i,t}$, where the idiosyncratic shocks $\omega_{i,t} \overset{i.i.d.}{\sim} N(0, \sigma_\epsilon^2)$ are independent across i and are dynamically independent of aggregate TFP.

After production, firms can choose to turn part of their production good into investment in next-period capital. A gross investment level of $I_{i,t}$ yields next-period capital $k_{i,t+1} = (1-\delta)k_{i,t} + e^{q_t} I_{i,t}$. The aggregate investment efficiency shifter q_t follows an AR(1) process $q_t = \rho_q q_{t-1} + \varepsilon_{q,t}$, where the aggregate shock $\varepsilon_{q,t} \overset{i.i.d.}{\sim} N(0, \sigma_q^2)$ is independent of the aggregate TFP shock $\varepsilon_{\zeta,t}$. Investment activity is free if $|I_{i,t}/k_{i,t}| \leq a$, where $a \geq 0$ is a parameter. Otherwise, firms pay a fixed adjustment cost of $\xi_{i,t}$ in units of labor (i.e., the monetary cost is $\xi_{i,t} \times w_t$). $\xi_{i,t}$ is drawn at the beginning of every period from a uniform distribution on the interval $[0, \bar{\xi}]$, independently across firms and time. Here $\bar{\xi} \geq 0$ is another parameter.

A representative household chooses consumption C_t and labor supply L_t to maximize

$$\mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \left\{ \log C_t - \chi \frac{L_t^{1+\varphi}}{1+\varphi} \right\} \right],$$

where φ is the inverse Frisch elasticity of labor supply. The household owns all firms and markets are complete. Market clearing requires $C_t = \int (Y_{i,t} + I_{i,t})\, di$ and $L_t = \int n_{i,t}\, di$.

See Winberry (2018, section 2.2) for the Bellmann equations implied by the firms' and household's optimality conditions.

Heterogeneous firm model: Parameter calibration

β	Discount factor	0.961	$\bar{\xi}$	Fixed cost bound	0.0083
χ	Labor disutility	$\bar{N} = \frac{1}{3}$	ρ_ζ	Agg. TFP AR(1)	0.859
φ	Inverse Frisch	10^{-5}	σ_ζ	Agg. TFP AR(1)	0.014
ν	Labor share	0.64	ρ_q	Agg. inv. eff. AR(1)	0.859
α	Capital share	0.256	σ_q	Agg. inv. eff. AR(1)	0.014
δ	Capital depreciation	0.085	ρ_ϵ	Idio. TFP AR(1)	0.53
a	No fixed cost region	0.011	σ_ϵ	Idio. TFP AR(1)	0.0364

Table D.1: Parameter calibration in the heterogeneous firm model.

D.2 Calibration

Table D.1 shows the parameter calibration used to simulate the data. The labor disutility parameter χ is chosen so that steady-state hours equal $\bar{N} = 1/3$, given all other parameters. As explained in Section 5.1, the only difference from Winberry (2018) is that the idiosyncratic productivity process uses the alternative (less persistent) parametrization from Khan and Thomas (2008). We do this because Winberry's Dynare code appears to be more numerically stable in a neighborhood of these alternative parameter values.

D.3 Estimating the parameters of the firms' productivity process

In this subsection we run the same estimation exercise as in Section 5, except that we here estimate the AR(1) parameter ρ_ϵ and innovation standard deviation σ_ϵ of the firms' idiosyncratic log productivity process. All other structural parameters (including the adjustment cost parameters) are assumed known for simplicity. The data and estimation settings are similar to those in Section 5.3 except that micro cross sections are observed at each of the five time points $t = 10, 20, \ldots, 50$.

Figure D.1 shows the full-information posterior densities of the idiosyncratic productivity parameters $(\rho_\epsilon, \sigma_\epsilon)$, across 9 simulated data sets.[3] The posterior densities are well-centered and concentrated near the true parameter values. We refrain from comparing with posterior inference that only exploits macro data, as these posteriors are extremely diffuse, consistent with Khan and Thomas (2008).

[3] We simulated 10 data sets but discarded one, as standard MCMC convergence diagnostics showed a failure of convergence of the Metropolis-Hastings sampler for that particular data set.

Heterogeneous firm model: Posterior densities of productivity parameters

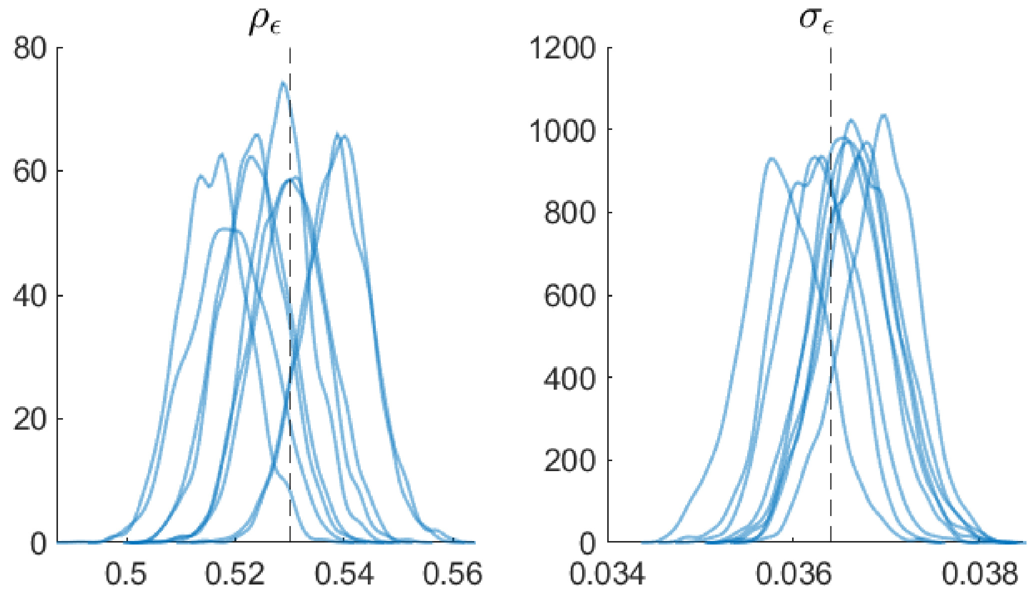

Figure D.1: Posterior densities across 9 simulated data sets. Vertical dashed lines indicate true parameter values. Posterior density estimates from the 9,000 retained MCMC draws using Matlab's `ksdensity` function with default settings.

Appendix E Alternative panel data approach

We here propose an alternative approach to handling panel data that may sometimes be more convenient than the procedure discussed in Section 6. Another view of the challenge described in that section is that, while the solution to the heterogeneous agent model supplies the marginal cross-sectional distribution of micro state variables at any point in time $p(s_{i,t} \mid z_t, \theta)$, it does not directly supply the *joint* density of micro state variables across multiple points in time $p\left(\{s_{i,t}\}_{t \in \mathcal{T}} \mid \mathbf{z}, \theta\right)$. To get around this issue, our second proposal is to artificially expand the micro state vector when solving the heterogeneous agent model, by including the previous periods' state variables in the current state vector.

Again, consider for illustration the heterogeneous household model in Section 2.2, and suppose each household is observed for two consecutive periods. Rather than treating the micro state vector as simply $s_{i,t} = (\lambda_i, \epsilon_{i,t}, a_{i,t-1})$ (permanent productivity, current-period employment, and predetermined normalized assets), we now expand it to be $\tilde{s}_{i,t} = (\lambda_i, \epsilon_{i,t}, a_{i,t-1}, \epsilon_{i,t-1}, a_{i,t-2})$ (thus including previous-period employment and normalized assets). Let $\tilde{\psi}_t$ denote the parameters governing the full four-dimensional cross-sectional distribution for the expanded micro state variables $\tilde{s}_{i,t}$. Then $\tilde{\psi}_t$ is part of the expanded aggregate state $\tilde{z}_t$.

Having expanded the micro state vector, we now modify the numerical model solver so that it outputs the joint cross-sectional density of $\tilde{s}_{i,t}$, not just of $s_{i,t}$. This is achieved by altering the model solution code and continuing to impose all model-implied restrictions on the evolution of the cross-sectional distribution (which is summarized by distribution parameters $\tilde{\psi}_t$). Note that the model-implied restrictions include those implied by the individual saving behavior $a'_t(a, \epsilon)$ and the exogenous Markov process for employment $p(\epsilon_{i,t} \mid \epsilon_{i,t-1}, \theta)$. We also continue to apply the Reiter (2009) type model solution method, leading to a state space model for macro variables characterized by equations (2) and (3).

Given a draw of the macro states $\tilde{\mathbf{z}}$ outputted from the modified model solution, we can easily compute the micro likelihood for two consecutive periods of household employment and income $(y_{i,t}, y_{i,t-1}) = (\epsilon_{i,t}, \iota_{i,t}, \epsilon_{i,t-1}, \iota_{i,t-1})$. This can be calculated as a simple distributional transformation of the four-dimensional distribution of the expanded micro state vector $\tilde{s}_{i,t}$, using the definition $\iota_{i,t'} = \lambda_i\{w_{t'}[(1-\tau)\epsilon_{i,t'} + b(1-\epsilon_{i,t'})] + (1+r_{t'})a_{i,t'-1}\}$ for $t' \in \{t-1, t\}$.

Though conceptually simple, the downside of the approach described in this section is that the expansion of the micro state vector is computationally demanding, for two reasons. First, as the dimension of the effective micro state vector $\tilde{s}_{i,t}$ increases, so does the dimension of the distribution parameters $\tilde{\psi}_t$, and then both the speed and the precision of off-the-shelf numerical methods for solving heterogeneous agent models tend to deteriorate. Second, achieving a sufficiently accurate approximation of the cross-sectional distribution of the higher-dimensional expanded micro state vector $\tilde{s}_{i,t}$ may require using a large number q of basis functions in the finite-dimensional distributional approximation described in Section 2.2, which further increases the dimension of $\tilde{\psi}_t$ in the equilibrium approximation.

References

Fisher, R. A. (1930): "Moments and Product Moments of Sampling Distributions," *Proceedings of the London Mathematical Society*, 2(1), 199–238.

Khan, A., and J. K. Thomas (2008): "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics," *Econometrica*, 76(2), 395–436.

Reiter, M. (2009): "Solving Heterogeneous-Agent Models by Projection and Perturbation," *Journal of Economic Dynamics and Control*, 33(3), 649–665.

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