Online Appendixes: Not for Publication

A Solution Methods

This section describes the equilibrium of the model with no aggregate uncertainty as well as the details of the KS, PARAM, XPA, REITER and WINBERRY solution methods. To maintain generality the bulk of the section focuses on describing the algorithms or equilibrium concepts themselves. Therefore most numerical details (grid sizes, optimization algorithms, etc) are deferred to a listing in Appendix B.

A.1 No Aggregate Uncertainty Model

The equilibrium definition of the steady-state model or model with no aggregate uncertainty is identical to the equilibrium with aggregate uncertainty discussed in the main text with constant aggregate productivity A. Unless otherwise specified, the steady-state model will be solved with $A = 1$. With a constant aggregate state space, individual firm states are given by the far smaller state space (z, k), and solution of the no aggregate uncertainty model simply involves repeatedly guessing values of the market-clearing price or consumption, computing an ergodic cross-sectional distribution $\mu(z, k)$ based on the price-implied policies and adjustment thresholds, and then checking consistency with the guessed price level. In the code available for this paper, the price clearing is performed using bisection, and calculation of an ergodic cross-sectional distribution given a price level follows the nonstochastic or histogram-based approach of Young (2010). In turn, this approach requires a projection grid for value functions, a denser simulation grid for idiosyncratic capital, and a discretized productivity process at the idiosyncratic level. For completeness, the equilibrium equations are listed below:

$$V^A(z,k) = \max_{k',n} \left\{ p\left(Azk^\alpha n^\nu - k' + (1-\delta)k - \frac{\phi}{p}n \right) + \beta \mathbb{E}_{z'} V(z', k') \right\}$$

$$V^{NA}(z,k) = \max_{n} \left\{ p\left(Azk^\alpha n^\nu - \frac{\phi}{p}n \right) + \beta \mathbb{E}_{z'} V(z', (1-\delta)k) \right\}$$

$$V(z,k) = -\phi \int_0^{\xi^*(z,k)} \xi dG(\xi) + G(\xi^*(z,k)) V^A(z,k) + \left[1 - G(\xi^*(z,k))\right] V^{NA}(z,k)$$

$$\xi^*(z,k) = \frac{V^A(z,k) - V^{NA}(z,k)}{\phi}$$

$$\frac{1}{p} = \int\int Azk^\alpha n(z,k,\xi)^\nu d\mu(z,k) dG(\xi) - \int\int (k'(z,k,\xi) - (1-\delta)k) d\mu(z,k) dG(\xi),$$

where I have that policies for labor have a closed-form and capital policies follow a threshold rule, given optimal policies $k^{'*}(z, k)$ from $V^A(z, k)$:

$$k'(z,k,\xi) = \begin{cases} k^{'*}(z,k), & \xi \le \xi^*(z,k) \\ (1-\delta)k, & \xi > \xi^*(z,k) \end{cases},$$

and that $\mu(z, k)$ is the ergodic distribution implied by a discretization as in Young (2010) together with application of the capital policies, adjustment cost thresholds, and idiosyncratic productivity transitions. Once the equilibrium is obtained, it is trivial to compute aggregate capital as $K = \int k d\mu(z, k)$.

A.2 Krusell Smith (KS)

The KS solution method used in the paper, following Khan and Thomas (2008) assumes that the set of approximating moments m in the aggregate state space of the model is equal to the mean or aggregate capital level K. Further, the solution method discretizes idiosyncratic and aggregate productivity processes following Tauchen (1986). Conditional upon a discretized level of aggregate productivity A, the forecast rules for price and next period's aggregate capital level take a loglinear form:

$$\log(\hat{p}) = \alpha_p(A) + \beta_p(A)\log(K)$$

$$\log(\hat{K}') = \alpha_K(A) + \beta_K(A)\log(K).$$

Recall that from the household problem this yields a forecast wage level $\hat{w} = \frac{\phi}{\hat{p}}$. Given these choices, the solution algorithm works as follows. First, guess an initial forecast rule system $\hat{\Gamma}_{(1)} = (\alpha_p^{(1)}, \beta_p^{(1)}, \alpha_K^{(1)}, \beta_K^{(1)})$. Then, before the solution iteration takes place, draw a large number T of exogenous aggregate productivity values based on the discretized Markov chain for A. Finally, in iteration $s = 1, 2, ...$ do the following

1. Given forecast rule system $\hat{\Gamma}_{(s)}$, solve the following system of equations via projection on some grid of states $(z, k; A, K)$

$$V^A_{(s)}(z,k;A,K) = \max_{k',n} \left\{ \begin{array}{c} \hat{p}_{(s)}(A,K)\left(zAk^\alpha n^\nu - k' + (1-\delta)k - \hat{w}_{(s)}(A,K)n\right) \\ +\beta \mathbb{E}_{z',A'} V_{(s)}(z',k';A',\hat{K}'_{(s)}) \end{array} \right\}.$$

$$V^{NA}_{(s)}(z,k;A,K) = \max_n \left\{ \hat{p}_{(s)}(A,K)\left(zAk^\alpha n^\nu - \hat{w}_{(s)}(A,K)n\right) + \beta \mathbb{E}_{z',A'} V_{(s)}(z',(1-\delta)k;A',\hat{K}'_{(s)}) \right\}$$

$$\xi^*_{(s)}(z,k;A,K) = \frac{V^A_{(s)}(z,k;A,K) - V^{NA}_{(s)}(z,k;A,K)}{\phi},$$

$$V_{(s)}(z,k;A,K) = \begin{array}{c} -\phi \int_0^{\xi^*_{(s)}(z,k;A,K)} \xi dG(\xi) + G\left(\xi^*_{(s)}(z,k;A,K)\right) V^A_{(s)}(z,k;A,K) \\ + \left(1 - G\left(\xi^*_{(s)}(z,k;A,K)\right)\right) V^{NA}_{(s)}(z,k;A,K) \end{array}$$

   which determines value function $V_{(s)}(z,k;A,K)$. Note that expectations are computed via summation over the appropriate portion of the discretized transition matrices Π^z and Π^A for idiosyncratic and aggregate productivity, respectively, and that the solution to the Bellman equations can be achieved with policy iteration, Howard acceleration of the Bellman equations, and continuous univariate optimization techniques in next period's capital level k', such as Brent optimization or golden section search.

2. Given the value function solution from step 1, simulate the model. Simulation follows the nonstochastic approach of Young (2010), and requires initialization of the cross-sectional distribution of capital and productivity $\mu(z,k)$ over a dense, discrete grid of (z_i, k_j) with n_d points for discretized capital. In each period t, market-clearing prices p_t must be determined without reference to the forecast price level $\hat{p}_{(s)}$ (which appears only through embedded expectations in the continuation value). The price clearing algorithm, for a given guessed price value $\tilde{p}$ requires calculation of the excess demand or error function, which is done as follows:

   - Compute new capital and adjustment cutoff thresholds at each point on the dense simulation grid (z_i, k_j) through

$$\tilde{V}^A(z_i,k_j) = \max_{k',n} \left\{ \begin{array}{c} \tilde{p}\left(z_iAk_j^\alpha n^\nu - k' + (1-\delta)k_j - \frac{\phi}{\tilde{p}}n\right) \\ +\beta \mathbb{E}_{z',A'} V_{(s)}(z',k';A',\hat{K}'_{(s)}) \end{array} \right\}$$

$$\tilde{V}^{NA}(z_i,k_j) = \max_n \left\{ \tilde{p}\left(z_iAk_j^\alpha n^\nu - \frac{\phi}{\tilde{p}}n\right) + \beta \mathbb{E}_{z',A'} V_{(s)}(z',(1-\delta)k_j;A',\hat{K}'_{(s)}) \right\}$$

$$\tilde{\xi}^*(z_i,k_j) = \frac{\tilde{V}^A(z_i,k_j) - \tilde{V}^{NA}(z_i,k_j)}{\phi}$$

     where the optimization over labor is a static problem which yields an analytical reduced-form for the right hand side expressions which can be optimized in capital k' alone. Let optimal labor and capital policies conditional upon adjustment be given by $n^*(z_i,k_j)$ and $k'^*(z_i,k_j)$. I have suppressed dependence on the aggregate states (A,K) for the moment, which is constant through the market-clearing process.

- Compute implied output, investment, consumption, and labor from

$$\tilde{Y} = \sum_{z_i, k_j} \mu_t(z_i, k_j) y(z_i, k_j), \quad y(z_i, k_j) = z_i A k_j^{\alpha} n^*(z_i, k_j)^{\nu}$$

$$\tilde{I} = \sum_{z_i, k_j} \mu_t(z_i, k_j) \left[G(\tilde{\xi}^*(z_i, k_j))(k^{'*}(z_i, k_j) - (1-\delta)k_j) \right]$$

$$\tilde{C} = \tilde{Y} - \tilde{I}$$

$$\tilde{N} = \sum_{z_i, k_j} \mu_t(z_i, k_j) \left(n^*(z_i, k_j) + \int_0^{\xi^*(z_i, k_j)} \xi dG(\xi) \right)$$

- Define the excess demand function or clearing error as $\frac{1}{\tilde{p}} - \tilde{C}$
- If excess demand is suitably small, the clearing algorithm stops. If not, repeat with an updated value of $\tilde{p}$, using any preferred method such as bisection.
- With market clearing price p_t in hand (as well as Y_t, I_t, C_t, N_t), update the discretized distribution $\mu_{t+1}(z_i, k_j)$ for the next period. Based on Young (2010), distributional transitions are calculated through linear interpolation of capital policies and use of the idiosyncratic productivity transition matrix

$$\mu_{t+1}(z_{i'}, k_{j'}) = \sum_{z_i} \sum_{k_j} \Pi^z_{ii'} \mu_t(z_i, k_j) \begin{bmatrix} \omega^a(i, j, j') \mathbb{I}(k^{'*}(z_i, k_j) \in [k_{j'-1}, k_{j'+1}]) \\ +\omega^{na}(i, j, j') \mathbb{I}(k_j \in [k_{j'-1}, k_{j'+1}]) \end{bmatrix}$$

$$\omega^a(i, j, j') = \begin{cases} G\left(\xi^*(z_i, k_j)\right) \left(\frac{k^{'*}(z_i, k_j) - k_{j'}}{k_{j'+1} - k_{j'}} \right) & k^{'*}(z_i, k_j) \in [k_{j'}, k_{j'+1}], 1 < j' < n_d \\ G\left(\xi^*(z_i, k_j)\right) & k^{'*}(z_i, k_j) \geq k_{n_d}, j' = n_d \\ G\left(\xi^*(z_i, k_j)\right) & k^{'*}(z_i, k_j) \leq k_1, j' = 1 \end{cases}$$

$$\omega^{na}(i, j, j') = \begin{cases} \left[1 - G\left(\xi^*(z_i, k_j)\right)\right] \left(\frac{\tilde{k_j} - k_{j'-1}}{k_{j'} - k_{j'-1}} \right) & \tilde{k_j} \in [k_{j'-1}, k_{j'}], 1 < j' < n_d \\ \left[1 - G\left(\xi^*(z_i, k_j)\right)\right] & \tilde{k_j} \geq k_{n_d}, j' = n_d \\ \left[1 - G\left(\xi^*(z_i, k_j)\right)\right] & \tilde{k_j} \leq k_1, j' = 1 \end{cases}$$

where above $\tilde{k}_j = \max((1-\delta)k_j, k_1)$ is the non-adjustment next-period capital stock. Also, with μ_{t+1} in hand, next period's simulated capital stock is computable as

$$K_{t+1} = \sum_{z_i, k_j} \mu_{t+1}(z_i, k_j) k_j.$$

- Continue to simulate period $t+1$.

3. With the simulation from periods 1 to T completed given forecast rules $\hat{\Gamma}_{(s)}$, discard some number T_{erg} of periods as initialization, and update the forecast rules based on OLS regressions. Run the loglinear OLS regressions of *realized* prices p_t on aggregate capital stocks K_t and realized next-period capital stocks K_{t+1} on current stocks K_t, where new coefficients $(\hat{\alpha}_p, \hat{\beta}_p, \hat{\alpha}_K, \hat{\beta}_K)$ are obtained separately for each level of aggregate productivity, and compute forecast rule errors as the maximum absolute difference between assumed coefficients $\hat{\Gamma}_{(s)}$ and the new estimated values. If the coefficients have converged to an acceptable tolerance, the model is solved. If not, then update $\hat{\Gamma}_{(s+1)}$ using dampened fixed-point iteration, i.e. set $\hat{\Gamma}_{(s+1)}$ equal to a weighted average of $\hat{\Gamma}_{(s)}$ and the newly estimated coefficients. Then continue to solution iteration $s+1$.

A.3 Parameterized Distributions (PARAM)

This is a discussion of the computational algorithm due to Algan et al. (2008, 2010a) (henceforth AADH) which relies upon higher-order reference moments, as well as assumed functional forms for the cross-sectional distribution of idiosyncratic capital and productivity $\mu(z,k)$ in the solution of the model.

Just as in the KS algorithm, I first must discretize the aggregate and idiosyncratic productivity processes following Tauchen (1986). Let the number of idiosyncratic (aggregate) productivity points be given by $n_z(n_A)$. Then, prior to the solution of the model, I first determine a set of aggregate moments m to be included in the aggregate state space. Here, this will be the singleton of aggregate capital, the cross-sectional mean K. Together with aggregate productivity, (A,K) therefore forms the aggregate state. Then, determine a set of reference moments m^{ref} used to help pin down the shape of the cross-sectional distribution of idiosyncratic capital and productivity. Here, this will be the first $n_M n_z$ centered moments of the capital distribution conditional upon each value of idiosyncratic productivity. Also, compute the exogenous ergodic distribution of idiosyncratic productivity $\tilde{\pi}_z$ for future use.

The reference moments are needed in this algorithm because, together with the aggregate capital state, they jointly determine the coefficients of the flexible exponential function form for the approximation to $\mu(z_i,k)$ conditional on discretized levels of z_i:

$$P(k,\rho^{z_i}) = \rho_0^{z_i} \exp\left\{ \begin{array}{c} \rho_1^{z_i}(k-m_1^{z_i}) + \rho_2^{z_i}\left[(k-m_1^{z_i})^2 - m_2^{z_i}\right] \\ +...+\rho_{n_M}^{z_i}\left[(k-m_1^{z_i})^{n_M} - m_{n_M}^{z_i}\right] \end{array} \right\}.$$

As discussed in AADH, a well-behaved convex minimization problem which is laid out below yields a mapping between the $n_z n_M$ moments of the cross-sectional distribution and coefficient vectors $\rho^{z_i}, i = 1,...,n_z$, because the first-order conditions of this problem are identical to the moment conditions setting distributional and reference moments equal to each other. I assume when manipulating or integrating this distribution that idiosyncratic capital lies in the bounds $[\underline{k},\bar{k}]$.

To solve the model, start with an initial constant guess for these reference moments, from the distribution in the steady-state model with no aggregate uncertainty. If desired, simulation can later provide a set of reference moments which vary with the level of aggregate productivity, or these steady-state reference moments themselves can be held fixed. Although I will outline both the repeated simulation and simulation-free approaches below, the results reported in the paper hold the reference moments at their steady-state levels. Finally, another option is also to use a regression forecast system from $(A,K)\to m^{ref}$, although following AADH I forego that approach here.

1. Iterate over the reference moments or the reference moment forecasting system, which yields a constant or variable mapping $(A,K)\to m^{ref}$.

    (a) Loop over the aggregate states (A,K), on the discretized grid for A and some projection grid for K.

    - For each (A,K) and implied $m^{ref}(A,K)$, choose distributional coefficients $\rho_{A,K,1}^{z_i},...,\rho_{A,K,n_M}^{z_i}, i = 1,...,n_z$ to solve

    $$\min_{\rho_{A,K}^{z_i}, i=1,...,n_z} \sum_{i=1}^{n_z} \int_{\underline{k}}^{\bar{k}} P(k,\rho_{A,K}^{z_i})dk,$$

      noting that the constants $\rho_{a,K,0}^{z_i}$ are irrelevant for the minimization and are simply chosen to ensure integration to 1. This is a well-behaved convex minimization problem and the integral can be computed via any standard quadrature rule. Here and throughout, integrals are computed via Simpson's rule. It is important to note that any rule with fixed nodes and weights is preferable to adaptive methods because fixed rules contribute to stability in the minimization problem above. The minimization is implemented in this paper using a robust and quick quasi-Newton routine with symmetric rank-one (SR1) updating of the approximation to the inverse Hessian.

    (b) Iterate over $V_{(s)}(z,k;a,K)$ to convergence.

- Loop over the aggregate states (A, K). For each (A, K), use any nonlinear equation system solver in p, K' to obtain $p(A, K)$ and $K'(A, K)$. The method used in this paper is dampened fixed-point iteration in the pair p, K'.
  - For each value of p, K', evaluate on the discretized grid for z and some spline projection grid for k the following equations:

$$V^A_{(s+1)}(z,k;A,K) = \max_{k',n} \left\{ \begin{array}{c} p\left(zAk^\alpha n^\nu - k' + (1-\delta)k - \frac{\phi}{p}n\right) \\ +\beta \mathbb{E}_{z',A'} V_{(s)}(z',k';A',K') \end{array} \right\}$$

$$V^{NA}_{(s+1)}(z,k;A,K) = \max_{n} \left\{ p\left(zAk^\alpha n^\nu - \frac{\phi}{p}n\right) + \beta \mathbb{E}_{z',A'} V_{(s)}(z',(1-\delta)k;A',K') \right\}$$

$$\xi^*(z,k;A,K) = \frac{V^A_{(s+1)}(z,k;A,K) - V^{NA}_{(s+1)}(z,k;A,K)}{\phi}$$

$$V_{(s+1)}(z,k;A,K) = \begin{array}{c} -\phi \int_0^{\xi^*(z,k;A,K)} \xi dG(\xi) + G(\xi^*(z,k;A,K))V^A_{(s+1)}(z,k;A,K) \\ + (1 - G(\xi^*(z,k;A,K)))\, V^{NA}_{(s+1)}(z,k;A,K) \end{array}$$

  - Then, compute the errors to the system of equations in p and K' given by

$$\frac{1}{p} = \sum_{i=1}^{n_z} \tilde{\pi}^z_i \int_{\underline{k}}^{\bar{k}} \left[\begin{array}{c} zAk^\alpha n(z_i,k;A,K)^\nu \\ -G(\xi^*(z_i,k;A,K))(k'(z_i,k;A,K) - (1-\delta)k) \end{array} \right] P(k, \rho^{z_i}_{A,K}) dk$$

$$K' = \sum_{i=1}^{n_z} \tilde{\pi}^z_i \int_{\underline{k}}^{\bar{k}} \left[\begin{array}{c} G(\xi^*(z_i,k;A,K))k'(z_i,k;A,K) \\ +(1 - G(\xi^*(z_i,k;A,K)))(1-\delta)k \end{array} \right] P(k, \rho^{z_i}_{A,K}) dk,$$

    where above the calculation requires the ability to compute $n(z_i,k;A,K)$, $k'(z_i,k;A,K)$, and $\xi^*(z_i,k;A,K)$ at a set of Simpson integration nodes in k. These can be computed by recalculation of the right hand sides of the Bellman equations above at the quadrature nodes in k, and this task is simplified by noting that capital policies depend only on idiosyncratic productivity and that the labor demand optimization is static and can be obtained in closed-form. Also, the densities $P(k, \rho^{z_i}_{A,K})$ must be evaluated using moments which reflect both the reference moments for higher-order terms but also the current mean level of aggregate capital. Therefore, the first moments of capital used in the construction of $P(k, \rho^{z_i}_{A,K})$ are linearly shifted to deliver consistency with K as the current aggregate state.

- Error on the value function iteration is given by $||V_{(s+1)} - V_{(s)}||$, which can be defined as desired. The solution used in this paper is based on the max absolute percentage difference. If the value function has converged, exit the value function iteration process. If not, go iteration $s + 1$.

(c) At this point, a decision must be made. If the reference moments used in the solution are to be held fixed at their steady-state values, the model is now solved. However, if the reference moments are to be updated with simulation, then simulation must be performed as part of the solution itself. In either case, simulation proceeds as follows, noting that a set of exogenous productivity draws A_t for T periods have been made outside of the solution loop. Given converged values V, simulate the economy for a large number of periods, in each period imposing market clearing to obtain p, K'. For each period, do the following, taking advantage of the functional forms assumed for the cross-sectional density.

- Start period t with a set of coefficients $\rho^{z_i}_t$ and moments m_t (the first $n_z n_M$ moments of the idiosyncratic capital density conditional upon idiosyncratic productivity) which jointly pin down the cross-sectional density in (z, K) for the period. Then using the value function from above and the same fixed point iteration approach, compute the equilibrium p, K' consistent with both the value function and the cross-sectional distributions for the current period.

- Compute the value of all reference moments for the next period $t+1$. These are higher-order centered moments of the cross-sectional distribution of capital next period, and can be computed directly via quadrature, given the policies and cross-sectional distribution coefficients of period t.
- Now that the reference moments are set for period $t+1$, along with the aggregate capital state, compute the coefficients of the cross-sectional distribution associated with period $t+1$ using the exact same minimization step as above.

(d) After simulation is completed for all T periods, and a certain number T_{erg} of initial periods are discarded, you have two options. If the reference moments are held constant at their steady-state values, you simply have an unconditional simulation of the model. If a fixed-point on the reference moments is desired, then update the reference moments in the outer loop now. The appropriate method depends on your assumptions for the reference moments. If you have assumed one unconditional constant set of reference moments not varying with aggregates, compute the unconditional average of each reference moment over the simulation. If you have assumed reference moments which vary with aggregate productivity, then compute the conditional average of each reference moment given the value of aggregate productivity. If you have assumed a reference moment forecasting system, update this system with OLS. Return to step (a) if the reference moments have not converged based on some criterion, say max absolute difference.

A.4 Explicit Aggregation (XPA)

First discretize the aggregate productivity process A and solve the steady-state model for each aggregate productivity value A_k, $k = 1, ..., n_A$. For each level, save values of equilibrium price and capital stocks $p^{SS}(A_k), K^{SS}(A_k)$. Also, discretize the idiosyncratic productivity process and store the exogenous ergodic distribution $\tilde{\pi}_z$ of discretized idiosyncratic productivity for future use.

Set up the aggregate state space to include (A, K) and posit forecast rules for market-clearing prices and next-period capital identical to the KS case. Then, solve the model *exactly* as in the case of the KS algorithm, with two modifications:

(2') Replace KS simulation step (2) with an "explicit aggregation" step. In particular, loop over aggregate states (A, K), where A varies over its discretized grid and K varies over the same projection grid used to compute the value functions.

- At each point (A, K), compute market-clearing prices via a solution routine like bisection over price p, with implied output, investment, consumption, and next-period aggregate capital given a guess $\tilde{p}$ defined by

$$\tilde{Y} = \sum_{z_i} \tilde{\pi}_i y(z_i, K), \quad y(z_i, K) = z_i A K^{\alpha} n^*(z_i, K)^{\nu}$$

$$\tilde{I} = \sum_{z_i} \tilde{\pi}_i \left[G(\tilde{\xi}^*(z_i, K))(k^{'*}(z_i, K) - (1-\delta)K) \right]$$

$$\tilde{C} = \tilde{Y} - \tilde{I}$$

$$\tilde{K}' = \sum_{z_i} \tilde{\pi}_i \left[G(\tilde{\xi}^*(z_i, K)) k^{'*}(z_i, K) + (1 - [G(\tilde{\xi}^*(z_i, K)))(1-\delta)K) \right]$$

with clearing error given as before by $\frac{1}{\tilde{p}} - \tilde{C}$.

- Note that the end of this iteration generates a dataset $(A, K) \to (p, K')$ as (A, K) varies over the discretization and projection grids.

(3') Replace KS update step (3) with a forecast update rule step using the dataset defined by explicit aggregation step (2').

- First, update the current vector of forecast rule coefficients $\hat{\Gamma}_{(s)}$ by estimating $(\hat{\alpha}_p(A), \hat{\beta}_p(A), \hat{\alpha}_K(A), \hat{\beta}_K(A))$ with OLS on the explicit aggregation dataset, segmented by discretized value of A.
- Then, define a vector of n_A bias correction terms

$$x_p^{Bias}(A) = \hat{\alpha}_p(A) + \hat{\beta}_p(A) \log(K^{SS}(A)) - \log(p^{SS}(A))$$

$$x_K^{Bias}(A) = \hat{\alpha}_K(A) + \hat{\beta}_K(A) \log(K^{SS}(A)) - \log(K^{SS}(A))$$

  and adjust the new forecast rule coefficients' constant terms with

$$\hat{\alpha}_p(A) \leftarrow \hat{\alpha}_p(A) - x_p^{Bias}(A)$$

$$\hat{\alpha}_K(A) \leftarrow \hat{\alpha}_K(A) - x_K^{Bias}(A).$$

- Then, check the estimated coefficients against the old coefficients $\hat{\Gamma}_{(s)}$. If they are within some tolerance according to max absolute deviations, the model is solved and exit the routine. If the forecast rules have not converged, use dampened fixed-point iteration to update the forecast rule system $\hat{\Gamma}_{(s+1)}$ based on rule (s) and the newly estimated system.

Note that the differencing off of $x^{Bias}(A)$ is an attempt to correct for the Jensen's inequality bias induced by substitution of aggregate states into idiosyncratic policies. The bias results from lack of variation in the cross-section of idiosyncratic capital when recovering market-clearing prices and next-period capital stocks. However, the steady-state model prices and capital stocks do incorporate cross-sectional integration over a distribution of idiosyncratic capital presumably similar to the distributions within the model with aggregate uncertainty. The modification by the term $x^{Bias}(A)$ requires that the estimated forecast system be able to exactly reproduce as a fixed point the steady-state prices and aggregate capital stocks $p^{SS}(A)$ and $K^{SS}(A)$, conditional upon aggregate productivity.

Note also that after the model is solved, simulation is completed exactly as in the KS algorithm, using the Young (2010) nonstochastic or histogram-based approach, and requiring market-clearing in each period with integration over the full cross-sectional distribution of idiosyncratic capital.

A.5 Projection Plus Perturbation (REITER)

The REITER solution method is based on three steps, and provides a perturbation approximation to the full rational expectations equilibrium. The first step is to solve the steady-state version of the model, with no aggregate uncertainty and aggregate productivity held fixed at a value of $A = 1$. The steady-state solution is identical to the one used, for example, as an input into the PARAM solution. The second step is to set up a system of nonlinear equations defining the model's equilibrium, which is covered in the first subsection below. The final step is to linearize and solve the system using standard numerical differentiation and solution techniques, covered in the second subsection below.

A.5.1 Nonlinear System of Equations in the Discretized Model

First establish a grid of n_z idiosyncratic productivity points and a Markov transition matrix $\Pi^z_{ij} = \mathbb{P}(z_{t+1} = z_j | z_t = z_i)$ following Tauchen (1986). Then, establish a grid of n_k idiosyncratic capital stock nodes k_i, which will function as knot points for cubic spline interpolation of the value functions. Finally, a denser simulation grid of n_d idiosyncratic points will be used to store the distribution. The following system of nonlinear equations can be linearized around the steady-state of the model.

$$V^A_{t-1}(z_i, k_j) = p_{t-1}\left(f(A_{t-1}, z_i, k_j, w_{t-1}) - k'^*_{t-1}(z_i, k_j) + (1-\delta)k_j\right) + \beta \sum_{z_{i'}} \Pi^z_{ii'} V_t(z_{i'}, k'^*_{t-1}(z_i, k_j)) + \eta^A_{t,i,j}$$

$$f(A_{t-1}, z_i, k_j, w_{t-1}) = (1-\nu)\left(\frac{\nu}{w_{t-1}}\right)^{\frac{\nu}{1-\nu}} (A_{t-1} z_i)^{\frac{1}{1-\nu}} k_j^{\frac{\alpha}{1-\nu}}$$

$$V_{t-1}^{NA}(z_i,k_j) = p_{t-1} f(A_{t-1}, z_i, k_j, w_{t-1}) + \beta \sum_{z_{i'}} \Pi^z_{ii'} V_t(z_{i'}, \tilde{k_j}) + \eta^{NA}_{t,i,j}$$

$$\tilde{k_j} = \max((1-\delta)k_j, k_1)$$

$$p_{t-1} = \beta \sum_{z_{i'}} \Pi^z_{ii'} \frac{\partial}{\partial k} V_t(z_{i'}, k'^*_{t-1}(z_i,k_j)) + \eta^{k*'}_{t,i}$$

$$V_t(z_i,k) = \begin{array}{c} -\phi \int_0^{\xi^*_t(z_i,k)} \xi dG(\xi) + G\left(\xi^*_t(z_i,k)\right) V_t^A(z_i,k) \\ + \left[1 - G\left(\xi^*_t(z_i,k)\right)\right] V_t^{NA}(z_i,k) \end{array}$$

$$\xi^*_t(z_i,k) = \frac{V_t^A(z_i,k) - V_t^{NA}(z_i,k)}{\phi}$$

$$\mu_t(z_{i'}, k_{j'}) = \sum_{z_i} \sum_{k_j} \Pi^z_{ii'} \mu_{t-1}(z_i,k_j) \left[\begin{array}{c} \omega^a(i,j,j') \mathbb{I}(k'^*_{t-1}(z_i,k_j) \in [k_{j'-1}, k_{j'+1}]) \\ + \omega^{na}(i,j,j') \mathbb{I}(\tilde{k_j} \in [k_{j'-1}, k_{j'+1}]) \end{array} \right]$$

$$\omega^a(i,j,j') = \begin{cases} G\left(\xi^*_{t-1}(z_i,k_j)\right) \left(\frac{k'^*_{t-1}(z_i,k_j) - k_{j'}}{k_{j'+1} - k_{j'}} \right) & k'^*_{t-1}(z_i,k_j) \in [k_{j'}, k_{j'+1}], 1 < j' < n_d \\ G\left(\xi^*_{t-1}(z_i,k_j)\right) & k'^*_{t-1}(z_i,k_j) \geq k_{n_d}, j' = n_d \\ G\left(\xi^*_{t-1}(z_i,k_j)\right) & k'^*_{t-1}(z_i,k_j) \leq k_1, j' = 1 \end{cases}$$

$$\omega^{na}(i,j,j') = \begin{cases} \left[1 - G\left(\xi^*_{t-1}(z_i,k_j)\right)\right] \left(\frac{\tilde{k_j} - k_{j'-1}}{k_{j'} - k_{j'-1}} \right) & \tilde{k_j} \in [k_{j'-1}, k_{j'}], 1 < j' < n_d \\ \left[1 - G\left(\xi^*_{t-1}(z_i,k_j)\right)\right] & \tilde{k_j} \geq k_{n_d}, j' = n_d \\ \left[1 - G\left(\xi^*_{t-1}(z_i,k_j)\right)\right] & \tilde{k_j} \leq k_1, j' = 1 \end{cases}$$

$$\frac{1}{p_{t-1}} = \sum_{z_i} \sum_{k_j} \mu_{t-1}(z_i,k_j) \left[y(A_{t-1}, z_i, k_j, w_{t-1}) - G\left(\xi^*_{t-1}(z_i,k_j)\right) \left(k'^*_{t-1}(z_i,k_j) - (1-\delta)k_j \right) \right]$$

$$w_{t-1} = \frac{\phi}{p_{t-1}}$$

$$Y_{t-1} = \sum_{z_i} \sum_{k_j} \mu_{t-1}(z_i,k_j) y(A_{t-1}, z_i, k_j, w_{t-1})$$

$$y(A_{t-1}, z_i, k_j, w_{t-1}) = \left(\frac{\nu}{w_{t-1}} \right)^{\frac{\nu}{1-\nu}} (A_{t-1} z_i)^{\frac{1}{1-\nu}} k_j^{\frac{\alpha}{1-\nu}}$$

$$I_{t-1} = \sum_{z_i} \sum_{k_j} \mu_{t-1}(z_i,k_j) G\left(\xi^*_{t-1}(z_i,k_j)\right) \left(k'^*_{t-1}(z_i,k_j) - (1-\delta)k_j \right)$$

$$N_{t-1} = \sum_{z_i} \sum_{k_j} \mu_{t-1}(z_i,k_j) \left[n(A_{t-1}, z_i, k_j, w_{t-1}) + \int_0^{\xi^*_{t-1}(z_i,k_j)} \xi dG(\xi) \right]$$

$$n(A_{t-1}, z_i, k_j, w_{t-1}) = \left(\frac{\nu}{w_{t-1}} \right)^{\frac{1}{1-\nu}} (A_{t-1} z_i)^{\frac{1}{1-\nu}} k_j^{\frac{\alpha}{1-\nu}}$$

$$\log(A_t) = \rho_A \log(A_{t-1}) + \varepsilon_t, \quad \varepsilon_t \sim_{iid} N(0, \sigma_A^2)$$

Together, the above system represents a total of $n_s = 2n_z n_k + n_z n_d + n_z + 5$ equations in the $n_s \times 1$ vector of endogenous variables $X_t = \left((V_t^A)', (V_t^{NA})', (k'^*_t)', \mu_t', \log(p_{t-1}), \log(Y_{t-1}), \log(I_{t-1}), \log(N_{t-1}), \log(A_t) \right)'$, together with the exogenous shock process ε_t. I write this nonlinear rational expectations system corresponding to the discretized model as $F(X_t, X_{t-1}, \eta_t, \varepsilon_t) = 0$, where $F(\cdot)$ is the left hand side minus the right hand side of each of the n_s equations above. Note that I know this equation is satisfied exactly at the steady-state value of $X_t = X_{t-1} = X^{SS}$ and $\eta_t = 0$, $\varepsilon_t = 0$ corresponding to no aggregate uncertainty and $A_t = 1$. The vectors X_t are $n_s \times 1$ while the vector η_t is $n_\eta \times 1$, where $n_\eta = 2n_z n_k + n_z$.

A few practical comments are in order. First, in general the approximation nodes used for interpolation of the value functions in k will be different (and less dense) than the discrete values of k used to store the cross-sectional distribution μ_t above. The value ε_t is the exogenous shock to aggregate productivity. The vector η_t is the stacked set of expectational errors using Sims (2002) notation which must be applied to the expectations in the Bellman and Euler equations above, and these expectational errors depend upon aggregates only as the idiosyncratic uncertainty reflected in the discretization of idiosyncratic productivity is already taken into account through the summation with respect to the transition matrix Π^z.

Also, note that when the system above is evaluated, the values of adjustment V_t^A and non-adjusment V_t^{NA} are used to construct the total value function V_t. Afterwards, this function is approximated using cubic splines, and the derivatives of the value function required for the Euler equation pinning down the optimal value of capital tomorrow are computed by analytically differentiating the spline approximation. The Euler equation for the optimal capital choice $k_{t-1}^{'*}(z_i, k_j)$ conditional upon adjustment only ranges over the productivity index i for z_i because the fixed nature of adjustment costs implies that tomorrow's optimal capital choice doesn't depend upon today's capital k_j.

The functions f, y, and n above represent the reduced-form expressions for revenue net of labor costs, output, and labor input at firms, conditional upon the analytic solution to the static labor optimization problem.

A.5.2 Linearizing the System

I then numerically differentiate F with respect to each of its arguments, at the steady-state, to obtain the system of n_s equations below

$$F_1(X_t - X^{SS}) + F_2(X_{t-1} - X^{SS}) + F_4\eta_t + F_5\varepsilon_t = 0,$$

where $F_1 = \frac{\partial F}{\partial X_t}(n_s \times n_s)$, $F_2 = \frac{\partial F}{\partial X_{t-1}}(n_s \times n_s)$, $F_3 = \frac{\partial F}{\partial \eta_t}(n_s \times n_\eta)$, and $F_4 = \frac{\partial F}{\partial \varepsilon_t}(n_s \times 1)$ are the approximated derivative matrices evaluated at the steady-state of the model. This system of equations can be solved using a large variety of solution methods for linear rational expectations models. In the calculations performed in this paper, the Sims (2002) algorithm as applied by the `gensys` software in MATLAB available on Chris Sims' website is used to solve the linear model, and the numerical differentiation is carried out by forward differentiation from the steady-state with relative step size 1E−6.

For concreteness, note that the linear solution to the model is simply a set of coefficient matrices A $(n_s \times n_s)$ and B $(n_s \times 1)$ such that locally around the steady-state the model satisfies

$$(X_t - X^{SS}) = A(X_{t-1} - X^{SS}) + B\varepsilon_t.$$

Immediately, the traditional local impulse responses $IRF_t, t = 1, ..., T_{IRF}$ ($n_s \times 1$ vectors) to a shock to aggregate productivity can be computed as

$$IRF_t = A^{t-1}B,$$

and the model can be simulated by drawing $N(0, \sigma_A^2)$ shocks and substituting directly into the solution equation above.

At this point, it's important to note that by numerically differentiating the system F I am assuming that although nonlinearity and threshold decision rules exist and are preserved at the microeconomic level, the dependence of these micro-level decisions, as embedded in the value functions and distributional transition weights above, on aggregate shocks to productivity is smooth. In the context of the Khan and Thomas (2008) model with stochastic adjustment costs, the resultantly smooth value functions and policies make such an assumption sensible. However, models with discrete choices and fixed, nonstochastic adjustment costs, such as those in Bloom et al. (2016), which can not be expected to see policies vary smoothly with aggregate shocks would not allow for a reasonable application of the REITER approach. Finally, for the levels of discreteness chosen in this paper's solution, the linear system is solved directly, but for a denser and infeasible levels of discretization model reduction techniques can be applied which still allow for the model's solution with standard linear rational expectations system solvers. These reduction techniques are laid out in Reiter (2010a).

A.6 Parameterization Plus Perturbation (WINBERRY)

The WINBERRY solution developed and first implemented in Winberry (2015) represents a combination of the REITER and PARAM algorithms. The core approach to the model solution, perturbation around the steady-state solution with respect to aggregate productivity, remains from the REITER method. However, by contrast with the REITER approach, the WINBERRY algorithm summarizes the cross-sectional distribution μ using only a set of reference moments together with the flexibly parameterized densities from the PARAM algorithm. Because of this parsimoniously parameterized storage convention, the dimensionality of the endogenous vector characterizing the economy in the WINBERRY method is in general significantly smaller than in REITER. Given its conceptual similarity to REITER, I simply reproduce the system of equations characterizing the WINBERRY solution here. Instructions for linearizing this system carry over directly over from the REITER subsection above.

First establish a grid of n_z idiosyncratic productivity points and a Markov transition matrix $\Pi^z_{ij} = \mathbb{P}(z_{t+1} = z_j | z_t = z_i)$ following Tauchen (1986). Let $\tilde{\pi}^z_i$, $i = 1, ..., n_z$ be the ergodic distribution of z. Then, record a set of $n_z n_M$ reference moments of the steady-state solution with no aggregate uncertainty, defined as

$$m_1^{z_i SS} = \frac{1}{\tilde{\pi}^z_i} \int_k k d\mu^{SS}(z_i, k), \quad i = 1, ..., n_z$$

$$m_j^{z_i SS} = \frac{1}{\tilde{\pi}^z_i} \int_k (k - m_1^{z_i})^j d\mu^{SS}(z_i, k), \quad i = 1, ..., n_z, \quad j = 2, ..., n_M.$$

These moments $m_j^{z_i}$ are simply the first n_M centered moments of the cross-sectional distribution μ^{SS} conditional upon each value of micro productivity z_i. Together with the parameterized set of densities

$$P(k, \rho^{z_i}) = \rho_0^{z_i} \exp \left\{ \begin{array}{c} \rho_1^{z_i}(k - m_1^{z_i}) + \rho_2^{z_i}\left[(k - m_1^{z_i})^2 - m_2^{z_i}\right] \\ + ... + \rho_{n_M}^{z_i}\left[(k - m_1^{z_i})^{n_M} - m_{n_M}^{z_i}\right] \end{array} \right\},$$

these moments define the cross-sectional distribution of capital conditional upon a given discretized level of productivity z_i. Then, record a set of grid of n_k idiosyncratic capital stock nodes k_i, which will function as knot points for cubic spline interpolation of the value functions. The following system of nonlinear equations can be linearized around the steady-state of the model.

$$V^A_{t-1}(z_i, k_j) = p_{t-1}\left(f(A_{t-1}, z_i, k_j, w_{t-1}) - k'^*_{t-1}(z_i, k_j) + (1-\delta)k_j\right) + \beta \sum_{z_{i'}} \Pi^z_{ii'} V_t(z_{i'}, k'^*_{t-1}(z_i, k_j)) + \eta^A_{t,i,j}$$

$$f(A_{t-1}, z_i, k_j, w_{t-1}) = (1-\nu)\left(\frac{\nu}{w_{t-1}}\right)^{\frac{\nu}{1-\nu}} (A_{t-1} z_i)^{\frac{1}{1-\nu}} k_j^{\frac{\alpha}{1-\nu}}$$

$$V^{NA}_{t-1}(z_i, k_j) = p_{t-1} f(A_{t-1}, z_i, k_j, w_{t-1}) + \beta \sum_{z_{i'}} \Pi^z_{ii'} V_t(z_{i'}, \tilde{k_j}) + \eta^{NA}_{t,i,j}$$

$$\tilde{k_j} = \max((1-\delta)k_j, k_1)$$

$$p_{t-1} = \beta \sum_{z_{i'}} \Pi^z_{ii'} \frac{\partial}{\partial k} V_t(z_{i'}, k'^*_{t-1}(z_i, k_j)) + \eta^{k*'}_{t,i}$$

$$V_t(z_i, k) = \begin{array}{c} -\phi \int_0^{\xi^*_t(z_i,k)} \xi dG(\xi) + G\left(\xi^*_t(z_i, k)\right) V^A_t(z_i, k) \\ + \left[1 - G\left(\xi^*_t(z_i, k)\right)\right] V^{NA}_t(z_i, k) \end{array}$$

$$\xi^*_t(z_i, k) = \frac{V^A_t(z_i, k) - V^{NA}_t(z_i, k)}{\phi}$$

$$m_{1,t}^{z_i} = \sum_{j=1}^{n_z} \tilde{\pi}_j^z \Pi_{ji}^z \int_{\underline{k}}^{\bar{k}} k_{t-1}^{'*}(z_j,k) P(k, \rho_{t-1}^{z_j}) dk, \quad i = 1,...,n_z$$

$$m_{j,t}^{z_i} = \sum_{j=1}^{n_z} \tilde{\pi}_j^z \Pi_{ji}^z \int_{\underline{k}}^{\bar{k}} (k_{t-1}^{'*}(z_j,k) - m_{1,t}^{z_i})^j P(k, \rho_{t-1}^{z_j}) dk, \quad i = 1,...,n_z, \quad j = 2,...,n_M$$

$$\frac{1}{p_{t-1}} = \sum_{i=1}^{n_z} \tilde{\pi}_i^z \int_{\underline{k}}^{\bar{k}} \left[y(A_{t-1}, z_i, k, w_{t-1}) - G\left(\xi_{t-1}^*(z_i,k)\right) \left(k_{t-1}^{'*}(z_i,k) - (1-\delta)k\right) \right] P(k, \rho_{t-1}^{z_i}) dk$$

$$w_{t-1} = \frac{\phi}{p_{t-1}}$$

$$Y_{t-1} = \sum_{i=1}^{n_z} \tilde{\pi}_i^z \int_{\underline{k}}^{\bar{k}} y(A_{t-1}, z_i, k, w_{t-1}) P(k, \rho_{t-1}^{z_i}) dk$$

$$y(A_{t-1}, z_i, k, w_{t-1}) = \left(\frac{\nu}{w_{t-1}}\right)^{\frac{\nu}{1-\nu}} (A_{t-1} z_i)^{\frac{1}{1-\nu}} k^{\frac{\alpha}{1-\nu}}$$

$$I_{t-1} = \sum_{i=1}^{n_z} \tilde{\pi}_i^z \int_{\underline{k}}^{\bar{k}} G\left(\xi_{t-1}^*(z_i,k)\right) \left(k_{t-1}^{'*}(z_i,k) - (1-\delta)k\right) P(k, \rho_{t-1}^{z_i}) dk$$

$$N_{t-1} = \sum_{i=1}^{n_z} \tilde{\pi}_i^z \int_{\underline{k}}^{\bar{k}} \left[n(A_{t-1}, z_i, k, w_{t-1}) + \int_0^{\xi_{t-1}^*(z_i,k)} \xi dG(\xi) \right] P(k, \rho_{t-1}^{z_i}) dk$$

$$n(A_{t-1}, z_i, k, w_{t-1}) = \left(\frac{\nu}{w_{t-1}}\right)^{\frac{1}{1-\nu}} (A_{t-1} z_i)^{\frac{1}{1-\nu}} k^{\frac{\alpha}{1-\nu}}$$

$$\log(A_t) = \rho_A \log(A_{t-1}) + \varepsilon_t, \quad \varepsilon_t \sim_{iid} N(0, \sigma_A^2)$$

Together, the above system represents a total of $n_s = 2n_z n_k + n_z n_M + n_z + 5$ equations in the $n_s \times 1$ vector of endogenous variables $X_t = \left((V_t^A)', (V_t^{NA})', (k_t^{'*})', (m_t^z)', \log(p_{t-1}), \log(Y_{t-1}), \log(I_{t-1}), \log(N_{t-1}), \log(A_t)\right)'$, together with the exogenous shock process ε_t. I write this nonlinear rational expectations system corresponding to the discretized model as $F(X_t, X_{t-1}, \eta_t, \varepsilon_t) = 0$, where $F(\cdot)$ is the left hand side minus the right hand side of each of the n_s equations above. Note that I know this equation is satisfied exactly at the steady-state value of $X_t = X_{t-1} = X^{SS}$ and $\eta_t = 0$, $\varepsilon_t = 0$ corresponding to no aggregate uncertainty and $A_t = 1$. The vectors X_t are $n_s \times 1$ while the vector η_t is $n_\eta \times 1$, where $n_\eta = 2n_z n_k + n_z$.

Some practical comments are in order. First, given values of $m_{t-1}^{z_i}$, the coefficients $\rho_{t-1}^{z_i}$ completing the definition of the parameterized densities $P(k, \rho_{t-1}^{z_i})$ above can computed just as in the PARAM algorithm as the solution to the convex minimization problem

$$\min_{\rho^{z_i}, i=1,...,n_z} \sum_{i=1}^{n_z} \int_{\underline{k}}^{\bar{k}} P(k, \rho_{t-1}^{z_i}) dk,$$

whose first-order conditions correspond to the n_M moment conditions for each value of z_i. With those distributional coefficients in hand, all integrals on the right hand side can be computed using standard quadrature techniques.

Table B1: Some Choices for the Numerical Solutions

Object	Value	Type or Location Used
k nodes for V spline interpolation	10	Loglinear [0.1,8.0]
k points for distribution histogram	50	Loglinear [0.1,8.0]
K points for V linear interpolation	10	Linear [1.25,2.0]
A points in discretization	5	Loglinear [0.94,1.0562]
z points in discretization	5	Loglinear [0.9176,1.0897]
V or k', ξ^* tolerance	0.0001	All methods
Forecast rule coefficient tolerance	0.001	KS, XPA methods
Howard accelerations	50	KS, XPA methods
Number of reference moments	4	PARAM & WINBERRY method densities
Quadrature nodes in k	36	PARAM method integration
Quadrature nodes in k	100	WINBERRY method integration
Nondiscarded simulation periods $T - T_{erg}$	2000	All methods
Number of IRF simulations	2000	All methods
IRF simulations length	50	All methods
IRF shock period	25	All methods

Note: The table indicates specific methods, tolerances, grid sizes and densities, and simulation lengths used in the numerical implementation of the solution techniques discussed in the paper. The code used to produce the results in this paper is available on Stephen Terry's website.

B General Numerical Choices

To complement the general discussion of each solution algorithm, it is also useful to list some practical choices made in the projection, optimization, and discretization of the model, with further information available in Table B1.

- The mean or aggregate level of capital K is used as the approximating moment m for the cross-sectional distribution $\mu(z, k)$ for the three solution methods requiring such an approximation (KS, PARAM, and XPA).
- In all solution techniques, idiosyncratic and aggregate productivity processes z and A are discretized according to Tauchen (1986) and along grids spanning two standard deviations of the unconditional process standard deviation around the process steady-state.
- Value functions are approximated as cubic splines in idiosyncratic capital k with a natural spline endpoint condition, and using linear interpolation in aggregate capital K. For the KS, XPA, and no-aggregate uncertainty models, the firm problem is solved using policy iteration with Howard acceleration of the value function, while in the PARAM method value function iteration is performed. Derivatives of value functions, where required, are computed as the derivatives of the corresponding spline approximations.
- The idiosyncratic capital policies within the Bellman equations, when optimization is required, are determined using Brent optimization.
- Price-clearing is performed during simulation using bisection (XPA, no aggregate uncertainty) or hybrid bisection/inverse quadratic interpolation (KS). Within the model solution step, joint search for clearing price and next period aggregate capital is performed using dampened fixed point iteration (PARAM).
- Minimization of the density-based objective determining distributional coefficients during simulation and solution of the PARAM and WINBERRY models is performed using a standard quasi-Newton algorithm with symmetric rank-one (SR1) updates to the inverse Hessian approximation.

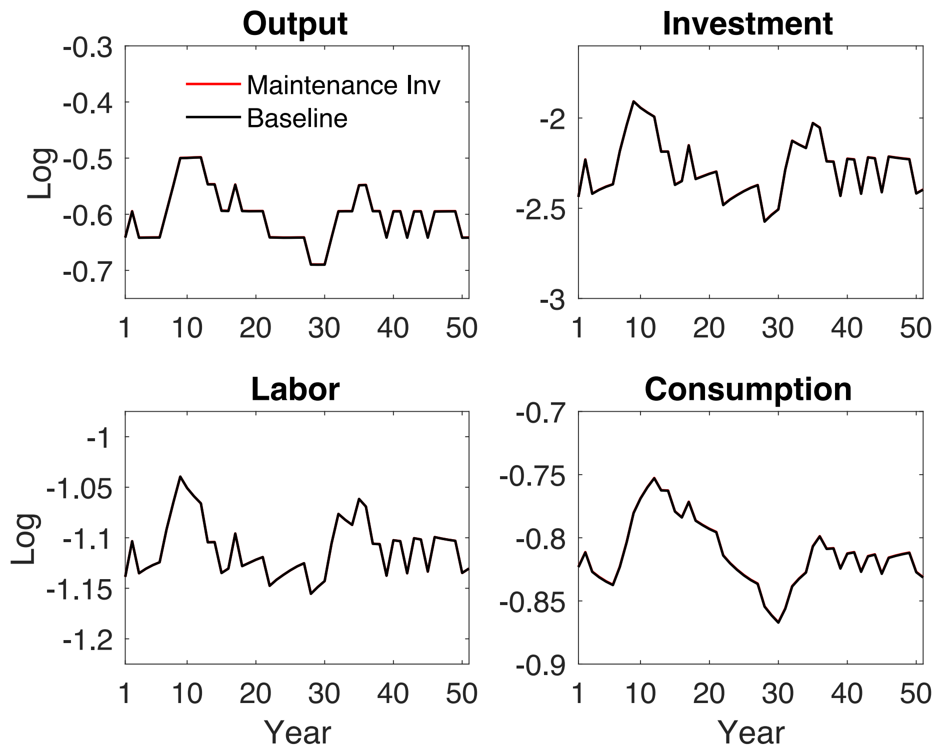

Figure B.1: Allowing for Maintenance Investment

Note: The figure plots a representative 50-year portion of selected business cycle aggregates drawn from the unconditional simulation of the model in the baseline KS solution with no maintenance investment allowed (in black) and the KS solution with maintenance investment allowed (in red). The exogenous discretized aggregate productivity process over this period is plotted in the left panel of Appendix Figure C.3 and is held constant across methods.

- Within the PARAM and WINBERRY solutions, integration over the cross-sectional densities is performed using fixed Simpson quadrature rules.

B.1 Maintenance Investment

The model extension allowing for maintenance investment follows Khan and Thomas (2008) in allowing for a capital adjustment within some small bounds without payment of fixed adjustment costs ξ. Firms are allowed to costlessly choose capital stocks in some range around the basic non-adjustment level: $k' \in B(k) \equiv [(-b+1-\delta)k, b+(1-\delta)k]$. When $b=0$, this nests the baseline. When b is positive but small this extension allows the model to match low observed investment inaction rates together with the observed proportion of investment spikes in the micro data. To verify that my simplified choice $b=0$ in the baseline model does not meaningfully drive overall macro dynamics, I solve the model using the KS solution for $b=0.011$, roughly following the calibration in Khan and Thomas (2008). The description of the algorithm and model environment is almost exactly identical to the description of the baseline KS solution above, with only the following five modifications required:

1. Firms in the non-adjustment case have flexible choice over capital policies $k^{NA}(z,k;A,K)$ as well as static labor choices $n^*(z,k;A,K)$ which are unchanged by future investment levels, which result from the solution to a modified Bellman equation for V^{NA}:

Table B2: Microeconomic Investment-Rate Moments with Maintenance Investment

	KS	KS-MAINT
$\frac{i}{k}$	0.0947	0.0929
$\sigma\left(\frac{i}{k}\right)$	0.2597	0.2491
$\mathbb{P}(\frac{i}{k}=0)$	0.7693	0.0239
$\mathbb{P}(\frac{i}{k}\geq 0.2)$	0.1724	0.1630
$\mathbb{P}(\frac{i}{k}\leq -0.2)$	0.0280	0.0263
$\mathbb{P}(\frac{i}{k}>0)$	0.1890	0.7521
$\mathbb{P}(\frac{i}{k}<0)$	0.0417	0.2479

Note: The rows of the table report the mean value of the indicated microeconomic moment of the cross-sectional distribution of investment rates $\frac{i}{k}$ in an unconditional simulation of the baseline KS solution with no maintenance investment (first column) and the extended version of the KS solution with maintenance investment allowed (second column). The first row reports the level of investment rates, the second row the cross-sectional standard deviation of investment rates, the third column the probability of investment inaction (defined as investment rates less than 1% in magnitude in the maintenance investment case), the fourth (fifth) columns the probability of positive (negative) investment spikes larger in magnitude than 20%, and the sixth (seventh) columns the probability of strictly positive (negative) investment rates requiring payment of fixed adjustment costs. All statistics are computed from a 2000-year unconditional simulation of the model, after first discarding an initial 500 years. The exogenous aggregate productivity series is held constant across methods.

$$V^{NA}_{(s)}(z,k;A,K) = \max_{n,k'\in B(k)} \left\{ \hat{p}_{(s)}(A,K) \left(\begin{array}{c} zAk^{\alpha}n^{\nu} - \hat{w}_{(s)}(A,K)n \\ -k' + (1-\delta)k \end{array} \right) + \beta \mathbb{E}_{z',A'} V_{(s)}(z',k';A',\hat{K}'_{(s)}) \right\}$$

2. During simulation of the model, given a guessed price level $\tilde{p}$, capital choices in the non-adjustment case $k'^{NA}(z_i,k_j)$ are now determined using the following Bellman equation:

$$\tilde{V}^{NA}(z_i,k_j) = \max_{n,k'\in B(k)} \left\{ \tilde{p} \left(\begin{array}{c} z_iAk_j^{\alpha}n^{\nu} - \frac{\phi}{\tilde{p}}n \\ -k + (1-\delta)k_j \end{array} \right) + \beta \mathbb{E}_{z',A'} V_{(s)}(z',k';A',\hat{K}'_{(s)}) \right\}$$

3. Since investment can now be non-zero when adjustment costs are not paid, the equation defining aggregate investment must be modified as

$$\tilde{I} = \sum_{z_i,k_j} \mu_t(z_i,k_j) \left[\begin{array}{c} G(\tilde{\xi}^*(z_i,k_j))(k^{'*}(z_i,k_j) - (1-\delta)k_j) \\ \left(1 - G(\tilde{\xi}^*(z_i,k_j))\right)(k^{'NA}(z_i,k_j) - (1-\delta)k_j) \end{array} \right]$$

4. Since capital next period in the non-adjustment case is not equal to $(1-\delta)k$ trivially, the evolution of the idiosyncratic capital distribution in the non-adjusting case depends on weights ω^{NA} whose definition must be modified to read:

$$\omega^{na}(i,j,j') = \left\{ \begin{array}{ll} \left[1 - G\left(\xi^*(z_i,k_j)\right)\right] \left(\frac{k^{'NA}(z_i,k_j) - k_{j'-1}}{k_{j'} - k_{j'-1}}\right) & k^{'NA}(z_i,k_j) \in [k_{j'-1},k_{j'}], 1 < j' < n_d \\ \left[1 - G\left(\xi^*(z_i,k_j)\right)\right] & k^{'NA}(z_i,k_j) \geq k_{n_d}, j' = n_d \\ \left[1 - G\left(\xi^*(z_i,k_j)\right)\right] & k^{'NA}(z_i,k_j) \leq k_1, j' = 1 \end{array} \right.$$

With the rest of the KS solution algorithm identical to the baseline case, and aggregate productivity shocks held constant, Figure B.1 plots a representative 50-year portion of the unconditional simulation of the model comparing the maintenance investment ($b > 0$) and baseline ($b = 0$) cases. The distinction makes little difference for aggregate dynamics. However, Table B2 demonstrates that, consonant with the intended purpose of this extension in Khan and Thomas (2008), the micro-level investment rate moments indicate substantially lower rates of pure investment inaction.

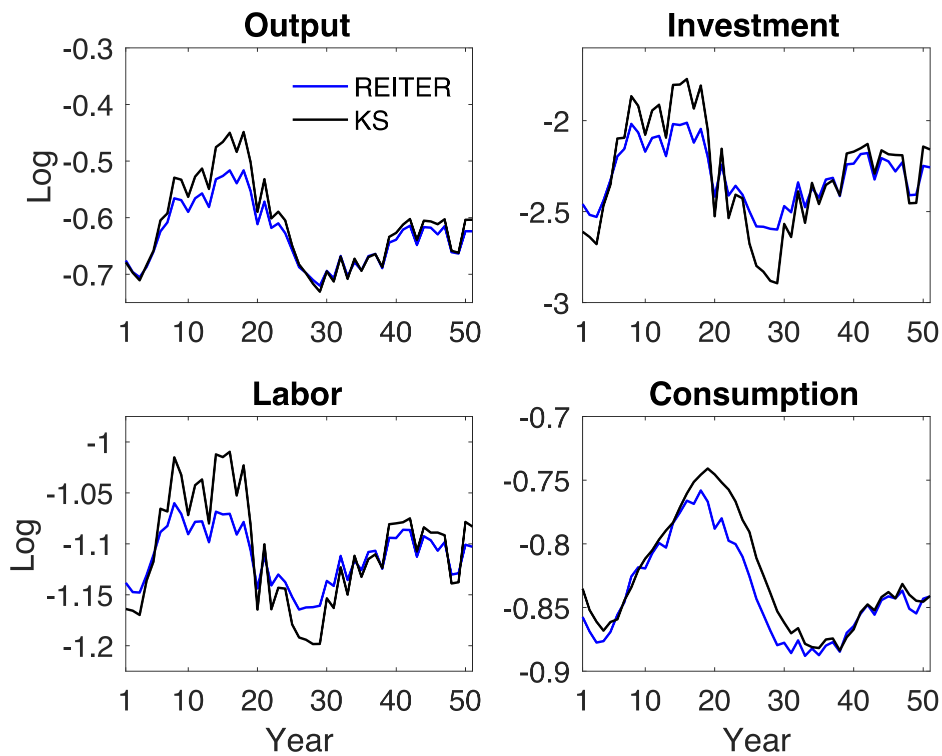

Figure B.2: Allowing for Continuous Shocks

Note: The figure plots a representative 50-year portion of selected business cycle aggregates drawn from the unconditional simulation of the model in a continuous aggregate productivity shock version of the KS solution (in black) and the REITER solution (blue). The underlying continuous exogenous aggregate productivity process is held constant across methods.

B.2 Continuous versus Discrete Aggregate Productivity

The analysis in the main text describes comparisons across solution methods in the context of discretized aggregate productivity shocks. These shocks are the fundamentals in the projection-based methods, while for the perturbation-based solutions I impose a set of continuous productivity shocks which duplicate the exogenous discretized aggregate productivity path.

This portion of the appendix lays out an extension to the KS solution of the model to allow for continuous aggregate productivity. There is no conceptual change to the algorithm required. However, I first provide a bit more detail on the calculation of expectations in the baseline discretized KS case before extending to the continuous shocks case. Model solution in the KS algorithm requires repeated calculation of expected values of the Bellman equation in future periods of the form

$$\mathbb{E}\left(V(z', k'; A', K')|z, A\right),$$

where today's exogenous states z and A are taken as given and some values for endogenous capital values k' and K' are known.

The baseline strategy is to implement discrete approximations to the stochastic processes for z and A according to Tauchen (1986), which results in grids and transition matrices $\{\tilde{z}_1, ..., \tilde{z}_{n_z}\}$, Π^z (for micro productivity), $\{\tilde{A}_1, ..., \tilde{A}_{n_A}\}$, Π^A (for aggregate productivity) such that $\mathbb{P}(z_{t+1} = \tilde{z}_j|z_t = \tilde{z}_i) = \Pi^z_{ij}$ and $\mathbb{P}(A_{t+1} = \tilde{A}_j|A_t = \tilde{A}_i) = \Pi^A_{ij}$. Given current values $z = \tilde{z}_i$ and $A = \tilde{A}_j$, in the baseline case the expectations of next period value is computed using the transition matrices for exogenous processes z and A together with cubic spline interpolation in k' and linear interpolation in K'. In particular, if a projection grid $\{\tilde{K}_1, ..., \tilde{K}_{n_K}\}$ is used in aggregate capital and today's exogenous processes satisfy $z = \tilde{z}_i$ and $A = \tilde{A}_j$ then

$$\mathbb{E}\left(V(z', k'; A', K')|z, A\right) \approx \sum_{i'=1}^{n_z} \sum_{j'=1}^{n_A} \pi^z_{ii'} \pi^A_{jj'} \left[(1 - \omega^K_{m'}(K'))V_{i'j'm'}(k') + \omega^K_{m'}(K')V_{i'j'm'+1}(k')\right]$$

where $V_{i'j'm'}$ is the cubic spline interpolant of $V(\tilde{z}_{i'}, k'; \tilde{A}_{j'}, \tilde{K'}_{m'})$ in micro capital k', $[\tilde{K}_m, \tilde{K}_{m+1}]$ is a bracketing interval to K', and $\omega_m(K') \equiv \frac{K' - \tilde{K}_m}{\tilde{K}_{m+1} - \tilde{K}_m}$ is the linear interpolation weight in K'.

By contrast, the continuous shock case allows for continuous or unmodified evolution of the aggregate productivity process

$$\log(A') = \rho_A \log(A) + \sigma_A \varepsilon'_A, \quad \varepsilon'_A \sim N(0, 1)$$

by computing expectations using Gauss-Hermite quadrature over the future values of ε'_A and bilinear interpolation in (A', K'). This approach broadly follows the strategy in McGrattan (1996), with standard quadrature formulas (Judd, 1998) implying a discrete set of integration nodes $\{\tilde{\varepsilon}_{A_1}, ..., \tilde{\varepsilon}_{A_{n_A}}\}$ and weights $\tilde{\pi}^A_i$, $i = 1, ..., n_A$. In this context, since the process A itself is not discretized, I must also use a separate projection grid $\{\tilde{A}_1, ..., \tilde{A}_{n_A}\}$ in aggregate productivity itself. If today's exogenous micro productivity process satisfies $z = \tilde{z}_i$ and values for k', A, and K' are given, then

$$\mathbb{E}\left(V(z', k'; A', K')|z, A\right) \approx \sum_{i'=1}^{n_z} \sum_{j'=1}^{n_A} \pi^z_{ii'} \tilde{\pi}^A_{j'} \left[\begin{array}{c} (1 - \omega^K_{m'}(K'))(1 - \omega^A_{l'(j')}(A))V_{i'l'(j')m'}(k') + \\ \omega^K_{m'}(K')(1 - \omega^A_{l'(j')}(A))V_{i'l'(j')m'+1}(k') + \\ (1 - \omega^K_{m'}(K'))\omega^A_{l'(j')}(A)V_{i'l'(j')+1m'}(k') + \\ \omega^K_{m'}(K')\omega^A_{l'(j')}(A)V_{i'l'(j')+1m'+1}(k') \end{array}\right]$$

where $[\tilde{A}'_l, \tilde{A}_{l'+1}]$ is the bracketing or nearest interval for $\exp(\rho \log(A) + \sigma_A \varepsilon_{Aj'})$ and

$$\omega^A_{l'(j')}(A) \equiv \frac{\exp(\rho \log(A) + \sigma_A \varepsilon_{Aj'}) - \tilde{A}'_l}{\tilde{A}_{l'+1} - \tilde{A}_{l'}}$$

is the linear interpolation/extrapolation weight for aggregate productivity next period. The other notation is similar to the discretized case.

With the approximation to the expectation of the value function next period defined in practical terms, the rest of the KS method is almost identical to before, although the exogenous set of aggregate productivity values used for simulation and update of prediction rules is given by a continuous simulation of the process.[16] Also, note that the prediction rules are modified to flexibly depend on a continuous A via the two unified rules

$$\log(\hat{K}'(A, K)) = \alpha_K + \beta_K \log(K) + \gamma_K \log(A) + \delta_K \log(K) \log(A)$$
$$\log(\hat{p}(A, K)) = \alpha_p + \beta_p \log(K) + \gamma_p \log(A) + \delta_p \log(K) \log(A).$$

Figure B.2 plots a representative 50-year period of the simulation of the continuous KS solution and the baseline REITER solution with the same continuous exogenous process for A.

C Simulated IRFs with Nonlinear Discretized Models

As noted in the main text, nonlinear impulse response analysis must take into account variation in the initial conditions and size of the imposed shocks, and following Koop et al. (1996) I take the following approach to compute the average conditional response to a one-standard deviation aggregate productivity shock:

1. Fix a large number N of simulations, a per-simulation length T_{IRF}, and a shock-period T_{shock}.

[16]In practice, I hold n_A and the projection grid for A equal to the n_A and discretized grid for A used in the baseline Tauchen (1986) discretization of A. I also simulate continuous standard normal draws for $\varepsilon'_A \sim N(0, 1)$ by using standard uniform draws and the Box-Muller transform.

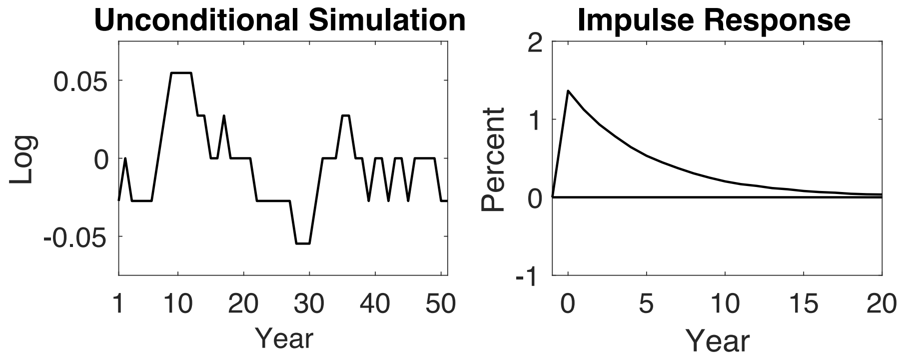

Figure C.3: Exogenous Productivity Series: Unconditional Simulation and Impulse Response

Note: The left panel of the figure plots a representative 50-year portion of the exogenous aggregate productivity series drawn from unconditional simulation of the model. This aggregate productivity series underlies portions of the plots in Figures 1, 5, B.1, D.4, and D.5. The right panel of the figure plots the exogenous positive one-standard deviation shock to aggregate productivity underlying the simulation-based generalized impulse responses following Koop et al. (1996) and plotted in Figure 3.

2. Independently draw exogenous $u_{it} \sim U(0,1)$ shocks for each period t of each simulation i, as well as one "shock occurrence" $s_i \sim U(0,1)$ draw for each simulation i.
3. For each simulation i, based on comparison of the shocks u_{it} with the entries of the cumulated transition matrix of discretized aggregate productivity, create two versions of simulated aggregate productivity series A_{it}^{shock} and $A_{it}^{noshocks}$, which are identical until the shock-period for $t < T_{shock}$. In the shock period T_{shock} for each simulation, compare the shock occurrence draw s_i for simulation i with the cutoff thresholds $\bar{s}$ defined below. If $s_i \leq \bar{s}$, set $A_{iT_{shock}}^{shock} = A_{nA}$, where n_A is the highest level of the discretized aggregate productivity process, and then allow A_{it}^{shock} to transit normally based on u_{it} for $t > T_{shock}$. If $s_i > \bar{s}$, allow A_{it}^{shock} to transit normally based on u_{it} for all periods $t \geq T_{shock}$. In either case, allow $A_{it}^{noshocks}$ to transit normally based on u_{it} for all $t \geq T_{shock}$.
4. Simulate any aggregate series of interest X twice, using both A_{it}^{shock} and $A_{it}^{noshock}$ exogenous processes. The impulse responses $\hat{x}_t$ of series X in period t to an aggregate productivity shock in period T_{shock} is given by

$$\hat{x}_t = 100\frac{1}{N}\sum_{i=1}^{N}\log\left(\frac{X_{it}^{shock}}{X_{it}^{noshock}}\right).$$

To obtain an average percentage innovation in aggregate productivity which equals σ_A exactly, I choose the shock threshold $\bar{s}$ to solve

$$\sigma_A = \bar{s}\sum_{k=1}^{n_A}\tilde{\pi}_{Ak}\left(\log(A_{n_A}) - \log(A_k)\right),$$

where $\tilde{\pi}_A$ is the ergodic distribution of the discretized aggregate productivity process A.

As noted in Appendix B, to compute the impulse responses plotted in the main text, I set $T_{IRF} = 50$, $T_{shock} = 25$, and $N = 2000$, and I hold exogenous draws u_{it}, s_i constant across simulation methods. I also set $n_A = 5$.

One final comment is in order regarding the REITER and WINBERRY solutions. Because these approaches yields linearized solutions the simulation-based analysis of Koop et al. (1996) is unnecessary. Although for completeness and comparability I perform the simulation-based impulse response with the REITER and WINBERRY methods, a much simpler alternative, invariant to shock scaling or initial conditions, is available. In particular, when writing the solution as $X_t = AX_{t-1} + B\varepsilon_{At}$, where X_t is the endogenous vector and ε_{At} is a continuous shock to aggregate productivity $\varepsilon_{At} \sim N(0, \sigma_A^2)$, the linear impulse response is given by

$$\hat{x}_t^{LIN} = A^{t-1}B,$$

Table D1: A Range of Shock Sizes
% Absolute Differences: KS vs. REITER

	Output		Investment		Labor		Consumption	
σ_A: Rel. to Baseline	Mean	Max	Mean	Max	Mean	Max	Mean	Max
5%	0.0977	0.1234	0.2648	0.4293	0.0405	0.0810	0.0576	0.0786
25%	0.0860	0.1716	0.2481	0.7624	0.0453	0.1749	0.0529	0.1272
Baseline (100%)	0.2195	0.5627	0.7400	9.2213	0.0964	0.4834	0.1658	0.4402
150%	0.8174	1.3791	1.9267	21.8248	0.3084	1.0287	0.5523	1.0712
200%	0.8254	1.6006	2.9494	53.7001	0.3130	1.4937	0.6256	1.5183

Note: The table reports the mean and maximum percentage difference between the indicated business cycle aggregate in the KS and REITER solutions over a 2000-year unconditional simulation of the model. Across rows, the standard deviation σ_A of aggregate productivity shocks varies. The underlying innovations or shocks to aggregate productivity are held constant as the scaling of these shocks varies across comparisons.

which is a vector of responses of each endogenous variable to an innovation in ε_{At}. For comparability, although this choice has little quantitative significance, I rely upon the Koop et al. (1996) concept for the perturbation-based methods in Figure 3 rather than the direct calculations based on the linear solutions.

D Extended Versions of the Model

Changing only the value of σ_A in the baseline model, Table D1 reports the mean and maximum percentage differences between the various business cycle aggregates in the resulting KS and REITER solutions over a 2000-year unconditional simulation of the model. Now, I proceed to describe the two extensions discussed in Section 5, namely size-dependent taxation and micro uncertainty fluctuations.

D.1 Size-Dependent Taxation

This extension to the baseline model allows for a system of distortionary size-dependent taxes and subsidies on the labor inputs of firms which are classified as large or small relative to a fixed reference level of idiosyncratic capital k^*. This system of subsidies and taxes varies over the cycle according to the following equation

$$\log(1+\tau(A)) = \gamma_\tau \log(A).$$

When $\gamma_\tau = 0$ and at steady-state, this extension nests the baseline model. The system of taxes and transfers is funded by lump-sum transfers to representative households which balance the fiscal authority's budget in each period. Since households own all of the firms in the economy and receive both the taxes/subsidies as well as the funding lump-sum transfers within the period, no changes need to be made to the clearing conditions of the economy. Furthermore, I simply re-interpret the wage w from the description of the baseline model as equal to the household marginal rate of substitution between consumption and leisure.

Because there is no additional aggregate state variable which needs to be tracked in the KS solution, and because the baseline is nested at the steady-state input into the REITER method, implementation requires only a minor set of modifications in practice. In particular, for all locations in which the wage w was previously referenced in the firm investment and labor decisions, insert instead the following net wage after the application of the size-dependent tax or subsidy.

$$\omega(A,k) = \begin{cases} w(1-\tau(A)), & k \geq k^* \\ w(1+\tau(A)), & k < k^* \end{cases},$$

In the case described in the main text, I solve the model with the round value of $k^* = 1.5$, approximately in the middle of the cross-sectional capital distributions plotted in Figure 2, as well as values for γ_τ equal to 10%, 25%, and 33%. Figure D.4 plots a representative 50-year period of the resulting unconditional simulation of the extended model for both the KS and REITER cases.

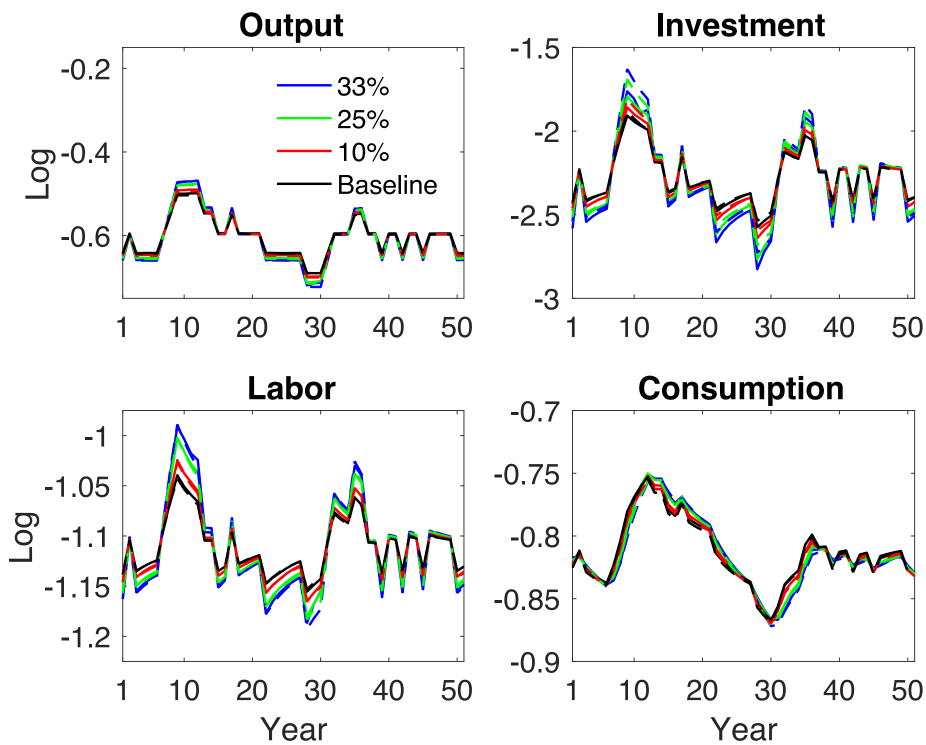

Figure D.4: Size-Dependent Taxation: Business Cycle Simulation

Note: The figure plots a representative 50-year portion of selected business cycle aggregates drawn from the unconditional simulation of the model in the KS (solid lines) and REITER (dashed lines) solutions. Each set of lines is based on a different level of the cyclical elasticity γ_τ of size-dependent taxation which is described in the main text. The baseline case of $\gamma_\tau = 0$ is in black, $\gamma_\tau = 10\%$ in red, $\gamma_\tau = 25\%$ in green, and $\gamma_\tau = 33\%$ in blue. The exogenous discretized aggregate productivity process over this period is plotted in the left panel of Appendix Figure C.3 and is held constant across all comparisons here.

D.2 Micro Uncertainty Fluctuations

The extension to the baseline model allowing for fluctuations in micro uncertainty links the volatility of micro productivity shocks to the realization of aggregate productivity in the next period via the equations

$$\log(z') = \rho_z \log(z) + s(A')\sigma_z \varepsilon_z, \quad \varepsilon_z \sim N(0,1)$$

$$\log(s(A)) = -\gamma_s \log(A).$$

When $\gamma_s = 0$ and at steady-state this extension nests the baseline model.

For the same reasons as in the size-dependent extension above, implementation requires only a minor set of modifications in practice to the baseline algorithms. In particular, only the expectations on the right hand side of Bellman equations and the distributional transitions must be modified to take into account the dependence of micro-level volatility next period on the realization of the aggregate productivity state A.

In the KS extension, there are a finite number n_A of possible realizations of micro-level volatility in the next period. The transition to each separate aggregate productivity, and hence micro-level volatility, state is governed by the transition matrix for discretized A: Π^A. This logic can be used to define any relevant joint discretized transition probability for z and A via

$$P(z_{t+1} = i', A_{t+1} = A'_j | z_t = z_i, A_t = A_j) = \Pi^A_{jj'} \Pi^{zA'_j}_{ii'},$$

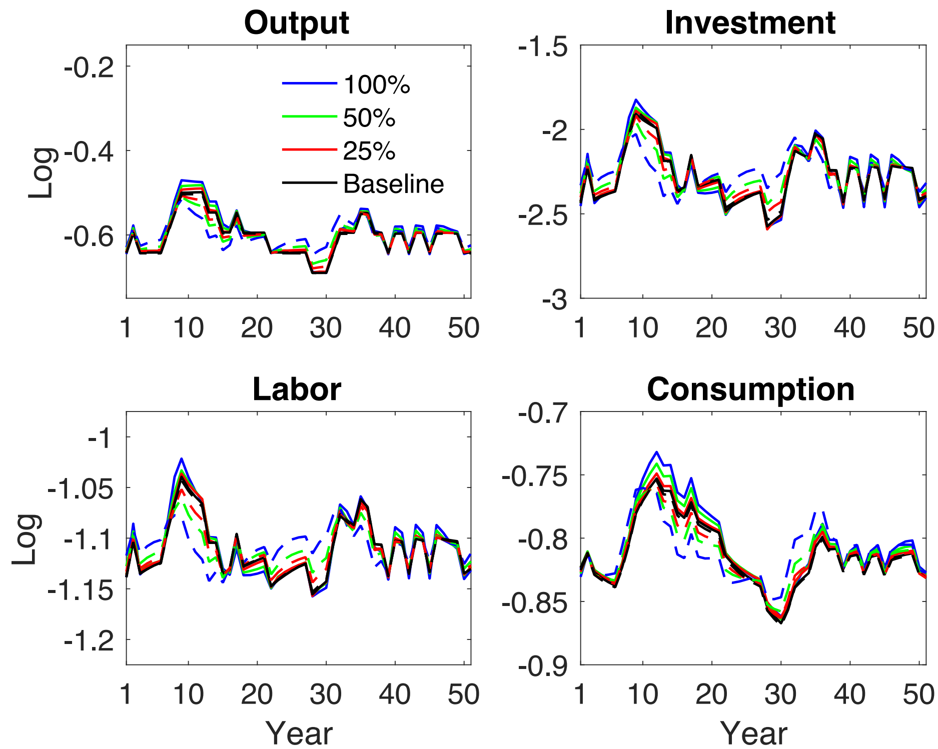

Figure D.5: Micro Uncertainty Fluctuations: Business Cycle Simulation

Note: The figure plots a representative 50-year portion of selected business cycle aggregates drawn from the unconditional simulation of the model in the KS (solid lines) and REITER (dashed lines) solutions. Each set of lines is based on a different level of the parameter γ_s described in the main text and governing the countercyclicality of micro volatility. The baseline case of $\gamma_s = 0$ is in black, $\gamma_s = 25$ in red, $\gamma_s = 50$ in green, and $\gamma_s = 100$ in blue. The exogenous discretized aggregate productivity process over this period is plotted in the left panel of Appendix Figure C.3 and is held constant across all comparisons here.

where the conditional transition matrix $\Pi^{zA'_j}$ is obtained just as in Tauchen (1986) by integration over intervals under the standard normal error term in the equation $\log(z_{t+1}) = \rho_z \log(z_i) + \sigma_z s(A_{j'})\varepsilon_z$. This formula can be used to compute all expectations on the right hand side of Bellman equations in the KS solution. The matrices $\Pi^{zA_{t+1}}$ alone are used to compute distributional transitions from t to $t+1$ in the simulation of the model, given a realization of aggregate productivity A_{t+1}. All of the relevant transition probabilities can be pre-computed and stored.

In the REITER extension, simply replace all appearances of the micro-level transition matrix Π^z with $\Pi^z(A_{t+1})$, where $\Pi^z(A_{t+1})_{ij} = P(z_{t+1} = z_j | z_t = z_i, A_{t+1})$ is obtained just as in Tauchen (1986) by integration over intervals under the standard normal error term in the equation $\log(z_{t+1}) = \rho_z \log(z_i) + \sigma_z s(A_{t+1})\varepsilon_z$. This discretization must be recomputed for each evaluation of the nonlinear system F defined in the REITER solution description above.

Figure D.5 plots a representative 50-year period of the resulting unconditional simulation of the extended model for both the KS and REITER cases with the choices γ_s equal to 25, 50, and 100.