Dynamically Consistent Global Nonlinear Solutions in the Sequence Space: Theory and Applications[1]

Online Appendix

Contents

[1]Hanbaek Lee, University of Cambridge, Email: hanbaeklee1@gmail.com

A Implementation of the repeated transition method

A.1 Implementation

The method starts from simulating a single path of exogenous aggregate shocks for a long-enough period T, $\mathbb{S} = \{S_t\}_{t=0}^{T}$, using the aggregate transition matrix Π^S. A time partition $\mathcal{T}(S)$ is defined as group of periods that share the same aggregate exogenous state realization as follows.

$$\mathcal{T}_S := \{\tau | S_\tau = S\} \subseteq \{0, 1, 2, ..., T\} \text{ for } S \in \{B, G\}.$$

The pseudo algorithm of the RTM is as follows:

Step 1. Guess on the paths of the value functions, the endogenous states, and the prices:[2]

$$\{V_t^{(n)}, \Phi_t^{(n)}, p_t^{(n)}\}_{t=0}^{T}.$$

Step 2. Solve the model backward from the terminal period T in the following sub-steps. The explanation is based on an arbitrary period t. For an illustrative purpose, I assume $S_t = G$ and $S_{t+1} = G$:

2-a. Find $\widetilde{t}+1$ where the endogenous aggregate allocation in period is identical to the one in period $t+1$, but the shock realization is different from period $t+1$ ($S_{\widetilde{t}+1} = B$):[3]

$$\widetilde{t}+1 = \arg \inf_{\tau \in \mathcal{T}_B} ||\Phi_\tau^{(n)} - \Phi_{t+1}^{(n)}||_\infty.$$

2-b. Compute the expected future value function as follows:

$$\mathbb{E}_t \widetilde{V}_{t+1} = \pi_{G,G} V_{t+1}^{(n)} + \pi_{G,B} V_{\widetilde{t}+1}^{(n)}.$$

2-c. Using $\mathbb{E}_t \widetilde{V}_{t+1}$ and $p_t^{(n)}$, solve the individual agent's problem at the period t. Then, I obtain the solution $\{V_t^*, a_{t+1}^*\}$.

After the taking these sub-steps for $\forall t$, $\{V_t^*, a_{t+1}^*\}_{t=0}^{T}$ are available.

[2] In practice, I use the stationary equilibrium allocations for all periods as the initial guess.

[3] Such $\widetilde{t}+1$ might not be unique. However, any of such $\widetilde{t}+1$ is equally good to be used in the next step.

Step 3. Using $\{a^*_{t+1}\}_{t=0}^T$, simulate forward the time series of the endogenous states $\{\Phi^*_t\}_{t=0}^T$ starting from $\Phi^*_0 = \Phi^{(n)}_0$.[4]

Step 4. Using $\{\Phi^*_t\}_{t=0}^T$, all the aggregate allocations over the whole path such as $\{K^*_t\}_{t=0}^T$ can be obtained. Using the market-clearing condition, compute the time series of the market-clearing price. If the model features a non-trivial market clearing condition, compute the time series of the implied prices $\{p^*_t\}_{t=0}^T$.[5]

Step 5. Check the distance between the implied prices and the guessed prices.

$$\sup_{BurnIn \leq t \leq T-BurnIn} ||p^*_t - p^{(n)}_t||_\infty < tol$$

Note that the distance is measured after excluding the burn-in periods at the beginning and the end of the simulated path. This is an adjustment to handle a potential bias from the imperfect guesses on the terminal period's value function $V^{(n)}_T$ and the initial period's endogenous state $\Phi^{(n)}_0$. The convergence criterion can be augmented by including the distance in other allocations, such as value functions or endogenous states.

If the distance is smaller than the tolerance level, the algorithm is converged. Otherwise, I make the following updates on the guess:[6]

$$p^{(n+1)}_t = p^{(n)}_t \psi_1 + p^*_t (1 - \psi_1)$$

$$V^{(n+1)}_t = V^{(n)}_t \psi_2 + V^*_t (1 - \psi_2)$$

$$\Phi^{(n+1)}_t = \Phi^{(n)}_t \psi_3 + \Phi^*_t (1 - \psi_3)$$

[4] In this step, if the endogenous state is a distribution, I use the non-stochastic simulation method (Young, 2010).

[5] It is worth noting that the prices here are not the market-clearing prices that are determined from the interactions between demand and supply. Rather, they are the prices implied by the market-clearing condition given either demand or supply fixed at the n^{th} iteration:

$$p^*_t = \arg_{\widetilde{p}}\{Q^D(p^{(n)}_t, X_t, X_{t+1}) - Q^S(\widetilde{p}, X_t, X_{t+1}) = 0\} \text{ or}$$
$$p^*_t = \arg_{\widetilde{p}}\{Q^D(\widetilde{p}, X_t, X_{t+1}) - Q^S(p^{(n)}_t, X_t, X_{t+1}) = 0\}.$$

In the computation method used in Krusell and Smith (1997), a market-clearing price needs to be computed in an additional loop due to the non-trivial market-clearing condition. The implied price cannot replace the market-clearing price in this method, as the misspecified price prediction rule can lead to a divergent law of motion of the aggregate allocation. In contrast, due to the missing market clearing step, the RTM significantly saves computation time.

[6] In highly nonlinear aggregate dynamics, I have found that the log-convex combination updating rule marginally dominates the standard convex combination updating rule in terms of convergence speed. The log-convex combination rule is as follows:

$$log(p^{(n+1)}_t) = log(p^{(n)}_t)\psi_1 + log(p^*_t)(1 - \psi_1).$$

for $\forall t \in \{0, 1, 2, 3, ..., T\}$. With the updated guess $\{V_t^{(n+1)}, \Phi_t^{(n+1)}, p_t^{(n+1)}\}_{t=0}^{T}$, I go back to Step 1.

(ψ_1, ψ_2, ψ_3) are the parameters of convergence speed in the algorithm. If ψ_i is high, then the algorithm conservatively updates the guess, leaving the algorithm to converge slowly. If the equilibrium dynamics are almost linear, as in Krusell and Smith (1998), uniformly setting ψ_i at around 0.9 guarantees convergence at a fairly high convergence speed. However, if a model is highly nonlinear, the convergence speed needs to be controlled to be substantially slower than the one in the linear models. This is because the nonlinearity can lead to a sudden jump in the realized allocations during the iteration if a new guess is too dramatically changed from the last guess. A heterogeneous updating rule $\psi_i \neq \psi_j$ $(i \neq j)$ is also helpful in cases where the dynamics of certain allocations are particularly more nonlinear than the others.

Depending on the models, the guessed bundle can include the paths of the optimal decision rules on top of the basic three allocations. For example, in an application for a global nonlinear solution for the canonical New Keynesian model with Rotemberg price adjustment cost, the paths of labor demand and consumption are all included in the bundle of guess and simultaneously updated with the inter-temporal policy functions, endogenous states, and prices.[7]

As can be seen from the convergence criterion in Step 5, the algorithm stops when the predicted allocation paths (n^{th} iteration) are close enough to the realized allocation paths (with asterisks). Therefore, once the convergence is achieved, the solution is guaranteed to be dynamically consistent: the predicted path coincides with the realized path. If the accuracy is measured in R^2 or in the mean-squared errors, as in Krusell and Smith (1998), the RTM features R^2 at 1, and its mean-squared error becomes negligibly different than zero.

A.2 The required length of simulation path

In this section, I discuss how long a simulation needs to be for the RTM. First, one of the most crucial determinants of the desired length is the assumed Markov process of the exogenous aggregate states. That is, the number of realizations of each exogenous state during the simulation is the key information. As in Krusell and Smith (1998), if only two aggregate states are realized based on a symmetric transition probability of a moderate level (0.875), the simulation of 500 periods is good enough to make the solution stay unaffected by further lengthening. However, for models

[7] The sample code and the model elaboration is available in the Online Appendix.

with more complex exogenous state processes, a longer simulation may be required. If an aggregate TFP process is discretized by the Tauchen method, covering three standard deviation ranges, at least 3,000 periods are needed to have enough realizations (at least 30) for both ends of the grids. The persistence of the exogenous state process also plays a role, with greater persistence requiring longer simulation periods.[8]

The nonlinearity of the model is another critical factor in determining the required simulation length. In practice, the RTM groups periods with similar endogenous aggregate states and uses piecewise interpolation to fill in missing observations. For highly nonlinear DSGE models, having additional observations in each exogenous state realization can substantially improve the accuracy of the solution by providing more nodes for interpolation. In the case of the highly nonlinear RBC model with irreversible investment, increasing the simulation length gradually changes the solutions until 4,000 periods are reached, after which the solution remains unaffected by additional periods.

While the required simulation length for the RTM is not necessarily longer than that used in the state space-based approach, it varies depending on the shock's persistence and the model's nonlinearity. The accuracy of the RTM stems from its ability to update the entire predicted path of allocations based on the entire realized path, which maximizes the information passed on to the next iteration. This contrasts with the state space-based approach, which relies on a functional relationship between the current and future periods that only summarily captures the dynamics. As a result, the RTM can achieve high accuracy without requiring a particularly longer simulation period compared to alternative methods.

A.3 Implementation of the sufficient statistic approach

In the algorithm explained in the previous section, Step 2-a is the most demanding step for heterogeneous-agent models, as it needs to find a period $\widetilde{t}+1$ that is identical to period $t+1$ in terms of distribution. Therefore, the similarity of the distributions across the periods needs to be measured, which is computationally costly.

However, if there is a sufficient statistic that can perfectly represent a period's endogenous aggregate state, the computational efficiency can be substantially improved. This enables to locate the target period $\widetilde{t}+1$ by only comparing the distance between these sufficient statistics instead of the distributions. For example, in Krusell and Smith (1998), if the aggregate capital is the sufficient

[8] It is important to note that the RTM is not unique in its sensitivity to simulation length. Other approaches, such as the state space-based approach, also require a sufficient number of observations in each exogenous state realization to ensure accuracy.

statistic, Step 2-a becomes easier as follows:

$$\widetilde{t}+1 = \arg \inf_{\tau \in \mathcal{T}_B} ||K_{\tau}^{(n)} - K_{t+1}^{(n)}||_{\infty}.$$

As the algorithm relies on the ergodicity, a sufficiently long period of simulation is needed for accurate computation. However, in practice, the simulation ends in finite periods. Therefore, the period $\widetilde{t}+1$ that shares exactly identical sufficient statistic as period t+1 might not exist. For this hurdle, the following adjusted versions of Step 2-a and Step 2-b help improve the accuracy of the solution:

2-a′. Find $\widetilde{t}^{up}+1$ where the sufficient statistic of the endogenous aggregate state is closest to the one in period $t+1$ from above, but the shock realization is different from period $t+1$:

$$\widetilde{t}^{up}+1 = \arg \inf_{\tau \in \mathcal{T}_B \text{ s.t. } e_{\tau}^{(n)} \geq e_{t+1}^{(n)}} ||e_{\tau}^{(n)} - e_{t+1}^{(n)}||_{\infty},$$

where e_{τ} denotes the sufficient statistic of the endogenous aggregate state in period τ. Similarly, find $\widetilde{t}^{dn}+1$ where the sufficient statistic of the endogenous aggregate state is closest to the one in period $t+1$ from below, but the shock realization is different from period $t+1$:

$$\widetilde{t}^{dn}+1 = \arg \inf_{\tau \in \mathcal{T}_B \text{ s.t. } e_{\tau}^{(n)} < e_{t+1}^{(n)}} ||e_{\tau}^{(n)} - e_{t+1}^{(n)}||_{\infty}.$$

Then, I have $e^{(n)}_{\widetilde{t}^{up}+1}$ and $e^{(n)}_{\widetilde{t}^{dn}+1}$ that are closest to $e^{(n)}_{t+1}$ from above and below, respectively. Using these two, I compute the weight ω to be used in the convex combination of value functions in the next step:

$$\omega = \frac{e_{t+1}^{(n)} - e^{(n)}_{\widetilde{t}^{dn}+1}}{e^{(n)}_{\widetilde{t}^{up}+1} - e^{(n)}_{\widetilde{t}^{dn}+1}}.$$

2-b′. Compute the expected future value function as follows:

$$\mathbb{E}_t \widetilde{V}_{t+1} = \pi_{G,G} V_{t+1}^{(n)} + \pi_{G,B} \left(\omega V^{(n)}_{\widetilde{t}^{up}+1} + (1-\omega) V^{(n)}_{\widetilde{t}^{dn}+1} \right).$$

Step 2-a' and Step 2-b' construct a synthetic counterfactual conditional value function by the

convex combination of the two value functions that are for the most similar periods to period $t+1$. These adjusted steps help accurately solve the problem in relatively short periods of simulation. For example, the model in Krusell and Smith (1998) can be accurately solved using only $T = 500$ periods of simulation (except for 100 burn-in periods at the beginning and the end of the simulated path).

The step of interpolation after finding the closest periods in terms of sufficient statistics can be understood as a piecewise interpolation, in contrast to the unconditional linear interpolation used in the state space-based approach based on the regression coefficients.

B The method's performance comparison: Krusell and Smith (1998) and Khan and Thomas (2008)

This section compares the performance of the RTM with the existing methods based on the heterogeneous-agent models. First, I compare the equilibrium allocations obtained from the RTM and the ones from the method in Maliar et al. (2010) for the model of Krusell and Smith (1998).[9] The RTM computes the exact level of the Lagrange multiplier for the occasionally binding constraint at the individual level, which enables the accurate computation. The path of the lagrange multipliers is computed by the residuals using the Euler equation as in Rendahl (2014).[10] With both algorithms calibrated to converge in approximately two minutes, their accuracy in terms of square root dynamic consistency differs by less than 10^{-5}. This negligible difference, despite Maliar et al. (2010) not computing Lagrange multipliers for occasionally binding constraints, reflects a model characteristic: the total amount of wealth held by constrained households is negligibly small, not meaningfully contributing to the aggregate capital dynamics. The similar performance between the two algorithms is due to the linearity of the capital dynamics, which makes the state space-based approach and the sequence space-based approach only negligibly different. It takes around 20 minutes for the RTM to converge under the convergence speed parameter $\psi_1 = \psi_2 = \psi_3 = 0.8$, which is a similar computation speed as Maliar et al. (2010) algorithm.

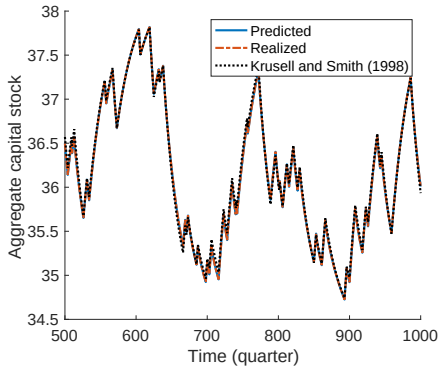
Figure B.1 plots a part of the predicted path $\{K_t^{(n)}\}_{t=0}^T$ and the realized (implied) path $\{K_t^*\}_{t=0}^T$ of aggregate capital K_t obtained from the RTM and the simulated path from the fitted log-linear

[9] The parameters are set as in the benchmark model in Krusell and Smith (1998) without idiosyncratic shocks in the patience parameter β.

[10] For some models where an occasionally binding constraint includes the current individual/aggregate state (e.g., irreversible investment), it is necessary to compute the expected future Lagrange multipliers. The RTM can accurately compute this expectation. Section C.1 and C.2 provide further explanations.

law of motion (Krusell and Smith, 1998).[11] As can be seen from all three lines hardly distinguished from each other, the RTM computes almost identical equilibrium allocations as the log-linear law of motion by Krusell and Smith (1998). This is because the log-linear specification almost perfectly captures the actual law of motion in the model.

Figure B.1: Equilibrium aggregate capital dynamics (Krusell and Smith, 1998)



Notes: The figure plots the time series of the aggregate wealth (capital) K_t in the model of Krusell and Smith (1998). The line with a round tick mark is the predicted wealth time series (n^{th} guess) $\{K_t^{(n)}\}_{t=500}^{1000}$. The line with the square tick mark is the realized wealth time series $\{K_t^*\}_{t=500}^{1000}$. The dashed line is the predicted wealth time series implied by the law of motion in Krusell and Smith (1998).

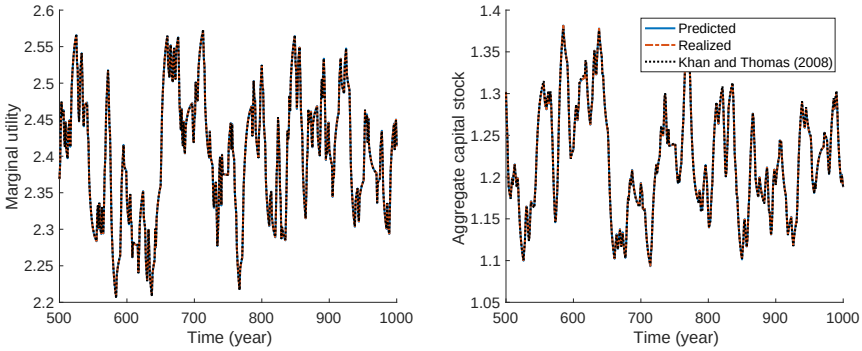
However, when a model features a non-trivial market-clearing condition, as in the model of Khan and Thomas (2008), the RTM substantially outperforms the state space-based approach based on the law of motion Krusell and Smith (1997). This is because their method requires to include an extra loop to find an exact market clearing prices in each period, while the RTM avoids this computational burden by using implied prices that converge to market clearing values in the limit of iterations.

To compare these approaches, I solve the Khan and Thomas (2008), using both the RTM and the Krusell and Smith (1997) algorithm its additional market-clearing loop. The substantial difference in computational efficiency complicates direct timing comparisons. Therefore, I instead evaluate both methods under identical dynamic consistency termination criteria. Figure B.2 plots the dynamics of price p_t (panel (a)) and aggregate capital stock K_t (panel (b)) computed from the RTM and Krusell and Smith (1997) algorithm. For the allocations computed from the RTM, both the predicted time series and the realized time series are plotted. As shown in the figure, all three lines display almost identical dynamics of the price and the aggregate allocations. The mean squared difference in the solutions between the RTM and Khan and Thomas (2008) is less than

[11] This figure is motivated from the fundamental accuracy plot suggested in Den Haan (2010).

10^{-5}.

Figure B.2: Equilibrium allocation paths (Khan and Thomas, 2008)



(a) Price p_t $(= 1/C_t)$ (b) Aggregate capital stocks K_t

Notes: The figure plots the time series of the price p_t the aggregate wealth (capital) K_t in the model of Khan and Thomas (2008). In both panels, the line with a round tick mark is the predicted time series (n^{th} guess) $\{p_t^{(n)}, K_t^{(n)}\}_{t=500}^{1000}$; the line with the square tick mark is the realized time series $\{p_t^*, K_t^*\}_{t=500}^{1000}$; the dashed line is the predicted time series implied by the law of motion.

In terms of the computational efficiency, the two methods display a substantial discrepancy. the RTM takes around 9 minutes to converge, while Krusell and Smith (1997) algorithm converge in around 5 to 6 hours on average in MATLAB.[12]

For both models of Krusell and Smith (1998) and Khan and Thomas (2008), I use the RTM using the sufficient statistic approach where the aggregate capital stock (the first moment) is used as the sufficient statistic. For this approach, it is necessary to check whether the firm's individual value (policy) function is strictly monotone in the aggregate capital stock. Figure B.3 and B.4 show the strict monotonicity holds, which validates the aggregate capital stock's qualification for a sufficient statistic.

Figure B.3 plots the levels of the marginal benefit of the Euler equation in the vertical axis and the corresponding aggregate capital stock in the horizontal axis for an individual household with the median-level capital stock and the unemployed status (low idiosyncratic productivity) in the solution of Krusell and Smith (1998). Each panel is for different contemporaneous aggregate productivity levels. The monotonicity stays unaffected regardless of the choice of the individual household.

Figure B.4 plots the level of the value functions in the vertical axis and the the corresponding aggregate capital stock in the horizontal axis for an individual firm with the median-level capital

[12]The convergence speed might change depending on the updating weight.

Figure B.3: Monotonicity of the policy functions in aggregate capital stock

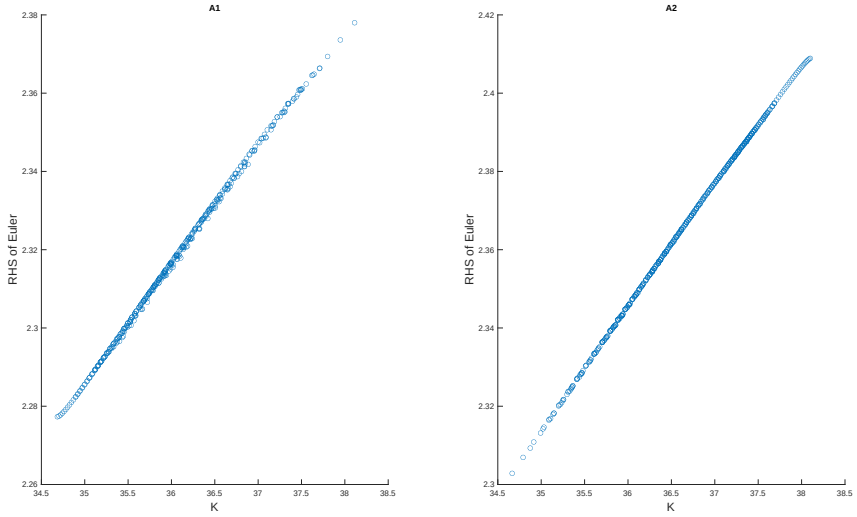
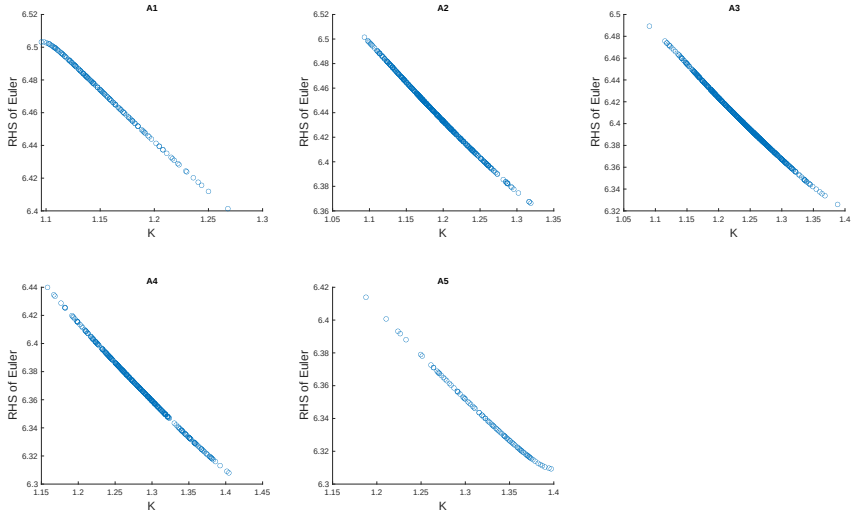


Figure B.4: Monotonicity of the value functions in aggregate capital stock



stock and idiosyncratic productivity in the solution of Khan and Thomas (2008). Each panel is for the different contemporaneous aggregate productivity levels. The monotonicity stays unaffected regardless of the choice of the individual firm.

C Parameter levels in the leading applications

C.1 The leading application I: Krusell and Smith (1998) with endogenous labor supply, investment irreversibility, and fiscal spending shock

The irreversibility constraint parameter is from Guerrieri and Iacoviello (2015). The aggregate TFP process follows the same Markov chain as in Krusell and Smith (1998). The labor disutility parameter is calibrated to match the steady-state labor supply at 0.33. The parameter choices are based on annual frequency.

Table C.1: Parameters for the leading application I

Parameter	Description	Value
α	capital share	0.330
β	discount factor	0.960
δ	depreciation	0.100
ϕ	borrowing constraint parameter	0.975
η	Frisch elasticity	1.000
η	labor disutility	10.850
ρ	idiosyncratic productivity persistence	0.800
σ	idiosyncratic productivity volatility	0.014
$\overline{G}$	steady-state government demand	$0.200 * Y^{ss}$
ρ_A	aggregate TFP persistence	0.900
σ_A	aggregate TFP shock volatility	0.013
ρ_G	government demand persistence	0.800
σ_G	government demand shock volatility	$0.01 * Y^{ss}$

C.2 The leading application II: A heterogeneous-household RBC model of portfolio choice (Krusell and Smith, 1997) with endogenous labor supply

The borrowing limit parameter and the aggregate TFP process are from Krusell and Smith (1997). The labor disutility parameter is calibrated to match the steady-state labor supply at 0.33. The parameter choices are based on quarterly frequency.

Table C.2: Parameters for the leading application II

Parameter	Description	Value
α	capital share	0.360
β	discount factor	0.990
δ	depreciation	0.025
η	Frisch elasticity	1.000
η	labor disutility	8.000
ρ	idiosyncratic productivity persistence	0.900
σ	idiosyncratic productivity volatility	0.050
$\underline{b}$	borrowing limit for the risk-free asset	-2.400
$\overline{B}$	dummy bond term (only for computation)	20.000

D A heterogeneous-firm business cycle model with irreversible investment

In this section, I show that the repeated transition method is powerful in solving nonlinear DSGE models with heterogeneous agents. Specifically, I analyze a heterogeneous-firm real business cycle model where individual firms are subject to the occasionally binding capital irreversibility constraint.

The representative household maximizes the following lifetime utility in the recursive form:

$$\begin{aligned} V(a;X) = \max_{c,a',N} \; & log(c) - \eta N + \beta \mathbb{E} V(a';X') \\ \text{s.t.} \quad & c + \int M(X,X') a'(X') d\Gamma_{X'} = a + w(X)N \\ & X' = G_H(X) \end{aligned}$$

where c is consumption; a is the asset (equity) value before dividend payment; N is labor supply; M is the stochastic discount factor; w is wage; η is labor disutility parameter.[13] The aggregate state X is a bundle of the distribution of individual firms Φ and aggregate TFP A: $X = \{\Phi, A\}$. The representative agent rationally expects the law of motion G_H of the aggregate states. I use the apostrophe to indicate future allocations.

Heterogeneous firms solve the following maximization problem of the present value of the sum

[13] Following Khan and Thomas (2008), I assume the Frisch elasticity of the labor supply is infinity. However, the repeated transition method also applies to the case of the finite Frisch elasticity seamlessly. In the online appendix such examples are provided.

of the dividend stream in the recursive form:

$$\begin{aligned} J(k,z;X) = \max_{k'} \; & d + \mathbb{E}_{z,X} M(X,X') J(k',z';X') \\ \text{s.t. } & d = \pi(k,z;X) + (1-\delta)k - k' \\ & k' \geq \phi I_{ss} + (1-\delta)k \\ & \pi(k,z;X) = \max_{k,n} \; Ak^{\alpha} n^{\gamma} - w(X)n \\ & X' = G_F(X) \end{aligned}$$

where d is dividend; k is individual capital stock; z is the idiosyncratic productivity; n is the labor demand; π is a temporal profit; δ is the depreciation rate; I_{ss} is the steady-state aggregate investment level; ϕ is the irreversibility parameter. G_F is the law of motion of the aggregate state from a firm's perspective, which coincides with G_H under the rational expectation and satisfies the dynamic consistency in the recursive competitive equilibrium. The stochastic discount factor M is determined in the competitive market as follows:

$$M(X,X') = \beta \frac{c(X)}{c(X')}$$

The individual and aggregate log productivities follow AR(1) processes, which are discretized by the standard Tauchen method.

The recursive competitive equilibrium is defined based on the following market-clearing conditions:

$$\begin{aligned} \text{(Labor market)} \quad & N(X) = \int n(k,z;X) d\Phi \\ \text{(Equity market)} \quad & a(X) = \int J(k,z;X) d\Phi. \end{aligned}$$

In the market clearing condition, the supply of equity meets the demand in the form of household wealth.

For computation, I use the standard parameter levels in the literature, which are available in Appendix C. For easier computation, I normalize the firm's value function by contemporaneous consumption $c(S)$ following Khan and Thomas (2008). Then, I define a price $p(S) := 1/c(S)$ and the normalized value function $\tilde{J}(k,z;S) := p(S)J(k,z;S)$. From the intra-temporal and inter-temporal optimality conditions of households, $w(S) = \eta/p(S)$ and $M(S,S') = \beta p(S')/p(S)$. Thus, $p(S)$ is

the only price to characterize the equilibrium. Then, the equilibrium price $p(S)$ is determined from the following variant of the non-trivial market clearing condition:[14]

$$p = \arg_{\widetilde{p}} \left\{ 1/\widetilde{p} - \int \left[d(x; X, \widetilde{p}) + w(X, \widetilde{p}) n(x; X, \widetilde{p}) \right] d\Phi = 0 \right\}.$$

where the consumption is $c = 1/\widetilde{p}$. In the problem above, the individual dividend policy d, wage w, and labor demand n functions are augmented by an argument $\widetilde{p}$, indicating that these objects are calculated assuming the price is at the level of $\widetilde{p}$. This is a fixed-point problem and computationally costly to solve, as consumption $c = 1/\widetilde{p}$ affects the firms' inter-temporal decision, which in turn affects the dividend, consumption, and thus, $\widetilde{p}$. Instead of the market clearing price, the repeated transition method uses the implied price p^*, which is obtained as follows:

$$\begin{aligned} p^* &= \arg_{\widetilde{p}} \left\{ 1/\widetilde{p} - \int \left[d(x; X, p^{(n)}) + w(X, p^{(n)}) l(x; X, p^{(n)}) \right] d\Phi = 0 \right\} \\ &= 1/\int \left[d(x; X, p^{(n)}) + w(X, p^{(n)}) l(x; X, p^{(n)}) \right] d\Phi, \end{aligned}$$

where $p^{(n)}$ is the guessed price in the n^{th} iteration. By bypassing the costly fixed-point problem, the method dramatically improves the speed of the computation.

The first-order condition of a firm's problem is as follows:[15]

$$1 = \mathbb{E}_{z,X} M(X, X') J_1(k', z'; X') + \lambda(k, z; X) \tag{1}$$

where λ is the Lagrange multiplier of the occasionally binding constraint. The envelope condition of a firm's problem is as follows:

$$\begin{aligned} J_1(k, z; X) &= \pi_1(k, z; X) + (1 - \delta) - \lambda(k, z; X)(1 - \delta) \\ &= \pi_1(k, z; X) + (1 - \delta)(1 - \lambda(k, z; X)) \end{aligned} \tag{2}$$

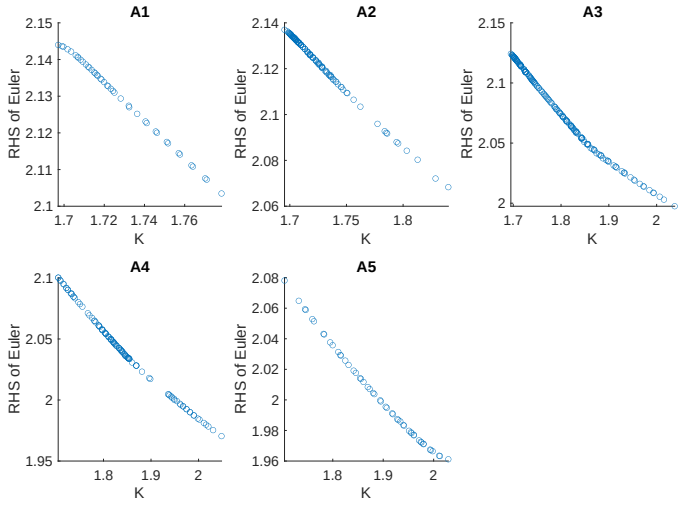
Combining equations (1) and (2), we have the following inter-temporal optimality condition:

$$1 - \lambda(k, z; X) = \underbrace{\mathbb{E}_{z,X} M(X, X')(\pi_1(k', z'; X') + (1 - \delta)(1 - \lambda(k', z'; X')))}_{\text{Expected marginal benefit}},$$

[14] This condition is derived from combining the household's budget constraint and the equity market clearing condition.

[15] The subscript denotes the partial derivative with respect to the argument in the corresponding argument order.

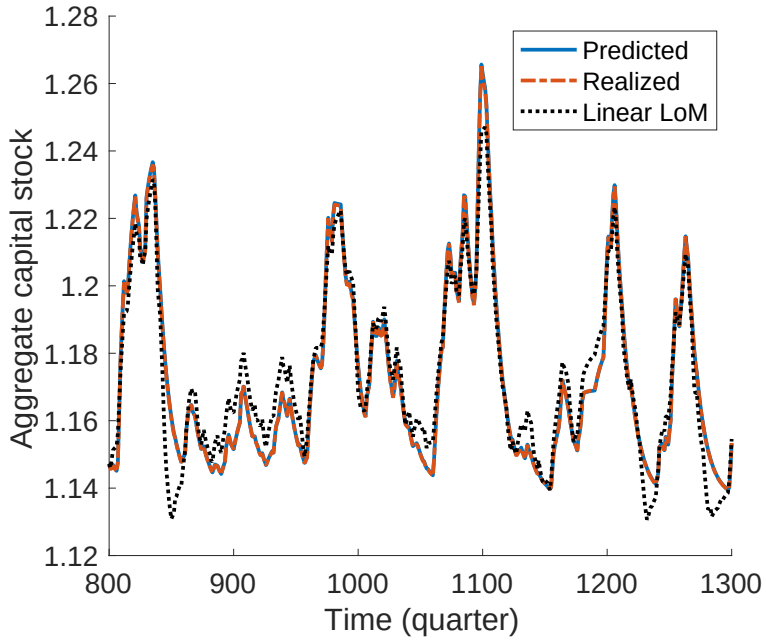
Figure D.5: Strict monotonicity of the marginal benefit in the aggregate capital stock



The expected marginal benefit requires the computation of state-contingent allocations (X' − *dependent*) of marginal profit π_1 and Lagrange multiplier λ. For this problem, I take the sufficient statistic approach, and the aggregate capital stock $K(S)$ (the first moment of the distribution of the firm-level capital stocks) is the sufficient statistic. I validate this approach by showing the monotonicity condition of Proposition 1 is satisfied. In particular, I show that each individual's marginal benefit $\pi_1 + (1-\delta)(1-\lambda)$ is strictly monotone in K for all aggregate exogenous state realization (TFP level). Figure D.5 plots the marginal benefit in the vertical axis and the corresponding aggregate equilibrium capital stock in the horizontal axis for different TFP levels ($A1$, $A2$, $A3$, ..., $A7$) given the individual state fixed at the median levels of capital stock and the firm-level productivity. In the unreported tests, which are available in the sample code, I confirm that monotonicity holds regardless of the choice of an individual firm.

For the computation, I run a simulation of 1,000 periods of aggregate TFP shock, where the TFP is discretized by the Tauchen method, covering the two standard deviation ranges. The computed aggregate capital path is highly nonlinear due to the occasionally binding constraint. Figure D.6 plots a part of the predicted path $\{K_t^{(n)}\}_{t=0}^T$ and the realized (implied) path $\{K_t^*\}_{t=0}^T$ of aggregate capital K_t obtained from the repeated transition method and the simulated path from the fitted log-linear law of motion. While the predicted path and the realized path coincide at the equilibrium path, the log-linear prediction significantly deviates from the others.

Figure D.6: The equilibrium path of aggregate capital stock



D.1 Nonlinearity and aggregation

In this section, I compare the nonlinearity implied in the heterogeneous firm model with the one in the representative firm model, which is obtained by simply muting the heterogeneous firm-level productivity.[16] For a valid comparison, I feed the same exogenous aggregate TFP path for both models and compute the equilibrium using the repeated transition method. Figure D.7 plots a part of the equilibrium capital path of the heterogeneous firm model (solid line) and the representative firm model (dash-dotted line) in log deviation from each model's steady state.

The volatility of the aggregate capital stock is significantly greater in the heterogeneous firm model, which is the by-product of the greater volatility in the aggregate investment. Table D.3 reports the business cycle statistics of the two models. The output and investment are around 10 percent more volatile in the heterogeneous firm model than the other, while consumption volatilities are at a similar level. The skewness of output is greater, and the skewness of consumption and investment is lower in the heterogeneous firm model.

The representative firm model fails to represent the heterogeneous-firm model over the business cycle. The reason is the nonlinearity at the firm-level capital dynamics.[17] To see this, I compute

[16] All the parameters are assumed at the same level except for the firm-level productivity.

[17] This result is specific to this model. For example, Khan and Thomas (2008) shows that the general equilibrium effect washes out the firm-level nonlinearity in their model.

Figure D.7: Equilibrium dynamics comparison: Heterogeneous vs. representative

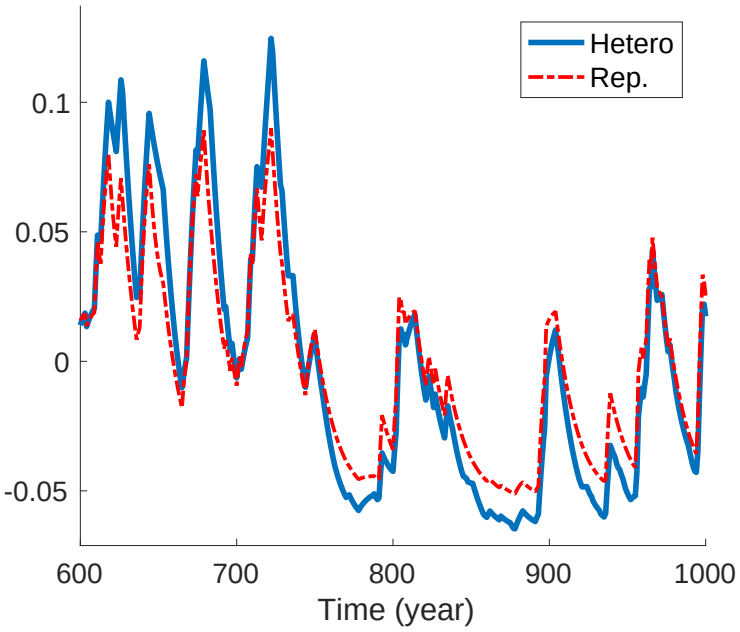
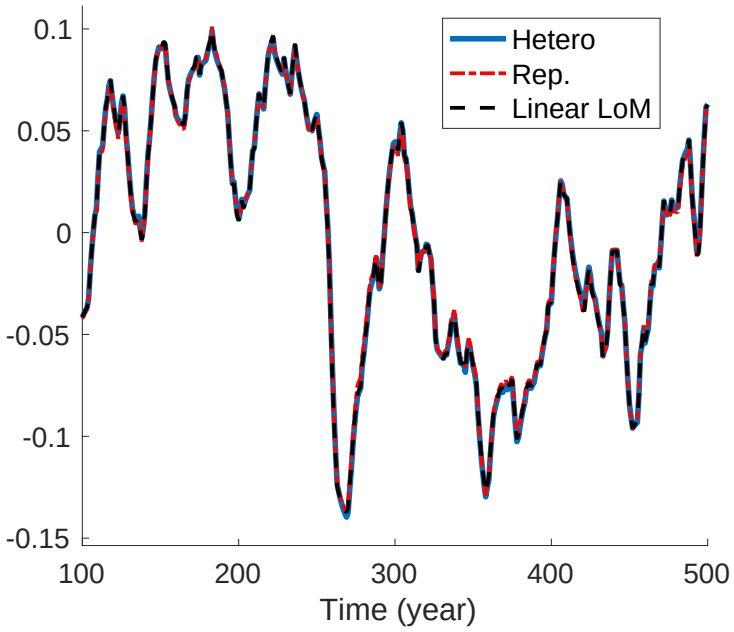


Table D.3: Business cycle statistics: Heterogeneous vs. representative

	Heterogeneous	Representative
Volatility		
$log(Output)$	0.042	0.039
$log(Consumption)$	0.034	0.034
$log(Investment)$	0.083	0.077
Skewness		
$log(Output)$	0.672	0.638
$log(Consumption)$	-0.049	-0.02
$log(Investment)$	1.757	1.926

the same heterogeneous and representative firm models without the occasionally binding constraint (fully reversible investment), which is the source of the nonlinearity. I refer to this version as the frictionless benchmark. Figure D.8 plots the aggregate capital dynamics of the frictionless benchmark of both models in log deviation from the steady state and the predicted path by the log-linear law of motion. These three lines perfectly coincide indicating that firm-level linearity

Figure D.8: Equilibrium dynamics comparison - frictionless: Heterogeneous vs. representative



makes the perfect representation that is also linear.[18]

D.2 Parameters used for the heterogeneous-firm business cycle model with irreversible investment

All the parameters are from Khan and Thomas (2008) except for the irreversibility constraint parameter, which I used the level of Guerrieri and Iacoviello (2015).

[18]The relevance of the firm-level nonlinearity in this model is contrasted with the neutrality result of Veracierto (2002). However, the two results are based on different firm-level setups, so the direct comparison is limited: the leading application is based on the occasionally-binding irreversibility constraint, while Veracierto (2002) features an (S, s) cycle in the firm-level capital stock.

Table D.4: Parameters for the heterogeneous-firm business cycle model with irreversible investment

Parameter	Description	Value
α	capital share	0.256
γ	labor share	0.640
δ	depreciation	0.069
ϕ	borrowing constraint parameter	0.975
β	discount factor	0.977
η	labor disutility	2.400
ρ	idiosyncratic productivity persistence	0.859
σ	idiosyncratic productivity volatility	0.022
ρ_A	aggregate TFP persistence	0.859
σ_A	aggregate TFP volatility	0.014

References

Den Haan, W. J. (2010, January). Assessing the accuracy of the aggregate law of motion in models with heterogeneous agents. *Journal of Economic Dynamics and Control 34*(1), 79–99.

Guerrieri, L. and M. Iacoviello (2015, March). OccBin: A toolkit for solving dynamic models with occasionally binding constraints easily. *Journal of Monetary Economics 70*, 22–38.

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

Krusell, P. and A. A. Smith, Jr. (1997, June). Income and Wealth Heterogeneity, Portfolio Choice, and Equilibrium Asset Returns. *Macroeconomic Dynamics 1*(02).

Krusell, P. and A. A. Smith, Jr. (1998, October). Income and Wealth Heterogeneity in the Macroeconomy. *Journal of Political Economy 106*(5), 867–896.

Maliar, L., S. Maliar, and F. Valli (2010, January). Solving the incomplete markets model with aggregate uncertainty using the Krusell–Smith algorithm. *Journal of Economic Dynamics and Control 34*(1), 42–49.

Rendahl, P. (2014, 06). Inequality Constraints and Euler Equation-Based Solution Methods. *The Economic Journal 125*(585), 1110–1135.

Veracierto, M. L. (2002, March). Plant-level irreversible investment and equilibrium business cycles. *American Economic Review 92*(1), 181–197.

Young, E. R. (2010, January). Solving the incomplete markets model with aggregate uncertainty using the Krusell-Smith algorithm and non-stochastic simulations. *Journal of Economic Dynamics and Control 34*(1), 36–41.