# Repeated Transition Method and the Nonlinear Business Cycle with the Corporate Saving Glut

Lee, Hanbaek

University of Tokyo

December 2022

# Repeated Transition Method and the Nonlinear Business Cycle with the Corporate Saving Glut*

Hanbaek Lee†

December 22, 2022

**Abstract**

This paper develops a novel methodology to globally solve nonlinear dynamic stochastic general equilibrium models with high accuracy. The algorithm is based on the ergodic theorem: if a simulated path of the aggregate shock is long enough, all the possible equilibrium allocations are realized, enabling a complete characterization of the rationally expected future outcomes at each point on the path. The algorithm is applied to a heterogeneous-firm business cycle model where firms hoard cash as a buffer stock. Using the model, I analyze the state-dependent shock sensitivity of consumption over corporate cash stocks and provide empirical evidence.

**Keywords**: Nonlinear business cycle, heterogeneous agents, stochastic dynamic programming, monotone function, state dependence.

**JEL codes**: E32, C63, D25.

*This paper is based on my Ph.D. dissertation at the University of Pennsylvania. This paper previously circulated under the title "Aggregate Uncertainty, Repeated Transition Method and the Aggregate Cash Cycle." I am extremely grateful to Jesús Fernández-Villaverde, Dirk Krueger, Andrew Abel, and Frank Schorfheide for their invaluable guidance and support. I also thank Kosuke Aoki, Florin Bilbiie, Christopher Carroll, Xiang Fang, Zhen Huo, Miles Kimball, Sagiri Kitao, Eunseong Ma, Makoto Nirei, Alessandro Peri, José-Víctor Ríos-Rull, Ludwig Straub, Guanyi Yang, and the participants of the seminars and the conference sessions for their insightful comments and discussions. All errors are my own.

†University of Tokyo. Email: hanbaeklee1@gmail.com

# 1 Introduction

This paper develops a novel algorithm that solves dynamic stochastic general equilibrium models without specifying the law of motion. I name the algorithm repeated transition method. This method solves a broad class of business cycle models that feature rich nonlinear aggregate dynamics globally and accurately.[1] The methodology is particularly useful for heterogeneous-agent models with highly nonlinear aggregate dynamics.

Under the rational expectation, heterogeneous agents are aware of the true law of motion in the aggregate states and correctly predict the future aggregate state. In contrast, there is no specific form of the law of motion known to a researcher. Moreover, it is computationally costly to track the evolution of a distribution of the individual states that is an infinite-dimensional object. To overcome this problem, Krusell and Smith (1998) suggested a log-linear prediction rule of the finite number of moments of the individual state distribution as an approximation to the true law of motion. Afterward, numerous research papers in the literature have found that this prediction rule gives a surprisingly accurate approximation to the true law of motion in the broad class of heterogeneous agent models with aggregate uncertainty.

Still, there are macroeconomic environments where the log-linear rule does not apply. In Fernandez-Villaverde et al. (2022), the nonlinear dynamics due to the endogenous aggregate risk coming out of the interaction between the heterogeneous households and the financial sector make it difficult to approximate the true law of motion using the log-linear specification. Also, the state dependence of the consumption responsiveness to a fiscal stimulus shock in Kaplan and Violante (2014) and the state-dependent aggregate investment dynamics in Lee (2022) are examples of such cases. According to Krusell and Smith (1997, 1998), these problems can be handled by tracking more moments of the state distribution. However, the functional form of

[1] Representative-agent business cycle models featuring highly nonlinear dynamics can also be accurately solved using the methodology.

the prediction rule and selection of the moments still remain an open-ended problem.

The repeated transition method overcomes these problems by relying on the theoretical fact of the ergodic theorem: if a simulated path of stationary shock process is long enough, all the possible allocations should be realized on the simulation path. This fact implies that state-contingent future allocations are obtainable somewhere on the simulation path as a realized outcome. Then, by properly identifying which period has the corresponding outcome to each of the expected future states, an agent's rational expectation at any time on the simulated path can be fully recovered. In the identifying step for the corresponding periods for expected future outcomes, the law of motion does not need to be specified: the only information needed for this step is a measure of similarity among the aggregate states across the periods.

For example, suppose an agent is at time $t$, and a macroeconomist needs to come up with a rationally expected value function of period $t+1$. For each possible exogenous aggregate shock realization $s \in S$ in $t+1$, I find a period $\widetilde{t}_s+1$ in the simulation history where the endogenous aggregate states are the closest to the ones in period $t+1$, and the aggregate shock realization is $s$. Then, I combine the value functions from these periods $\{\widetilde{t}_s+1\}_{s\in S}$ to construct the expected future value function. Due to the ergodic theorem, if the simulation path is long enough, there almost surely exists a period $\widetilde{t}_s+1$ where the endogenous aggregate allocations such as the distribution of individual states are perfectly identical to the ones in period $t+1$ among the periods where the aggregate shock realization is $s$. Therefore, the expected value function can be correctly constructed by combining these state-contingent value functions on the path.

The repeated transition algorithm runs until the time series of the expected allocations and the simulated allocations converge. Therefore, the solution is highly accurate, featuring $R^2$ at unity and mean squared error close to zero, even for highly nonlinear models. This method also provides an accurate global solution for the representative-agent models with aggregate uncertainty. The application to the representative-agent model smoothly follows once the endogenous distribution of the individual states is replaced by the endogenous aggregate allocations. In terms of speed, the repeated transition algorithm outperforms the algorithm of Krusell and Smith (1997) in models with non-trivial market-clearing conditions, as it does not require an extra loop for the market-clearing price.

Using the repeated transition method, I study the role of corporate cash stocks on the business cycle in a heterogeneous-firm business cycle model. In the model, the optimal future cash stock displays a kink after the target cash holding level. On top of this kink, due to the missing general equilibrium force to flatten the dynamics of aggregate cash holdings, the TFP-driven aggregate fluctuations of the cash stocks are highly nonlinear.

In the calibrated model, lagged aggregate cash stock significantly mitigates the aggregate consumption responsiveness to a negative productivity shock and intensifies the responsiveness to a positive productivity shock. Especially, the corporate cash stock gives an asymmetrically stronger insurance effect toward the negative TFP shock than the consumption boosting effect when the positive TFP shock hits. The data counterpart empirically supports this model's prediction of state dependence, and the empirical pattern is observed only after the early 1980s.[2] The fact that corporate cash holding has dramatically increased since the early 1980s partly explains why such a significant nonlinear effect is observed only after the early 1980s.

**Related literature** The repeated transition method builds upon the method utilizing perfect-foresight impulse response suggested by Boppart et al. (2018). In the paper, aggregate allocations' impulse responses are obtained from the transition dynamics induced by MIT shocks to the steady-state distribution. Then, the law of motion of aggregate allocations is locally approximated around the steady state. Therefore, the method assumes certainty equivalence between the expected deterministic

[2]The result is robust over other choices of the cutoff year around 1980.

path and the expected path when the aggregate uncertainty is present. In contrast, the repeated transition method does not assume certainty equivalence and globally solves the model. And it directly computes aggregate allocations and market-clearing prices in each period on the simulation path without specifying the law of motion.

Therefore, the repeated transition algorithm is distinguished from the solution methods based on perturbation and linearization (Reiter, 2009; Boppart et al., 2018; Ahn et al., 2018; Winberry, 2018; Childers, 2018; Auclert et al., 2019). As this method utilizes a single path of simulated aggregate shock that is long enough to fully represents the stochastic process, its approach is closely related to Kahou et al. (2021). Kahou et al. (2021) utilizes the fact that a whole economy's dynamics can be characterized by solving a finite number of agents' problems on a single Monte Carlo draw of individual shocks under the permutation-invariance condition. And the law of motion is nonlinearly computed using the deep-learning algorithm. Instead of the law of motion being characterized as an equilibrium object, the repeated transition algorithm computes the path of equilibrium allocations at each point on the simulated path. Then the law of motion can be backed out from the time series of the realized allocations. My method relies only on relatively simple computational techniques but computes highly accurate solutions. Also, the algorithm is widely applicable as the algorithm does not rely on the particular characteristics of the problem presented in this paper.

**Roadmap** Section 2 explains the repeated transition method based on the model in Krusell and Smith (1998). Section 3 validates the accuracy of the repeated transition method by comparing the computed outcome with the existing well-known results in the literature. Section 4 introduces a heterogeneous-firm business cycle model where firms save cash. Section 5 discusses the business cycle implication of corporate cash holdings predicted by the model compared to the observations from the data. Section 6 concludes.

# 2 Repeated transition method

## 2.1 A model for algorithm introduction: Krusell and Smith (1998)

I explain the repeated transition method based on the heterogeneous agent model with aggregate uncertainty in Krusell and Smith (1998). In this section, I briefly introduce the basic environment of the model.

A measure one of ex-ante homogeneous households consumes and saves. At the beginning of a period, a household is given wealth $a_t$ and an idiosyncratic labor supply shock $z_t$. Households are aware of the distribution of households $\Phi_t$, the aggregate productivity shock $A_t$, and how the aggregate states evolve in the future $G(\Phi_t, A_t, A_{t+1})$. The idiosyncratic shock and the aggregate shock follow the stochastic Markov processes elaborated in Krusell and Smith (1998). The Markov process is specified by the transition matrix $\pi$

$$\pi := \begin{bmatrix} \pi_{uB,uB} & \pi_{uB,eB} & \pi_{uB,uG} & \pi_{uB,eG} \\ \pi_{eB,uB} & \pi_{eB,eB} & \pi_{eB,uG} & \pi_{eB,eG} \\ \pi_{uG,uB} & \pi_{uG,eB} & \pi_{uG,uG} & \pi_{uG,eG} \\ \pi_{eG,uB} & \pi_{eG,eB} & \pi_{eG,uG} & \pi_{eG,eG} \end{bmatrix} = \begin{bmatrix} 0.525 & 0.350 & 0.03125 & 0.09375 \\ 0.035 & 0.84 & 0.0025 & 0.1225 \\ 0.09375 & 0.03125 & 0.292 & 0.583 \\ 0.0099 & 0.1151 & 0.0245 & 0.8505 \end{bmatrix}$$

In each element of the matrix, the first index indicates the current individual state $s \in \{u, e\}$, where $u$ indicates an unemployed status and $e$ indicates an employed status; the second index indicate the current aggregate state $S \in \{B, G\}$ where $B$ indicates a bad aggregate productivity state and $G$ indicates a good aggregate productivity state. The third and fourth indices are the future individual and aggregate states, respectively. For example, $\pi_{uB,uB}$ implies a transition probability that an unemployed worker stays unemployed in the next period when the economy is bad and stays bad in the future period.

The income sources of a household are labor work and capital stock. The budget constraint of the household is as follows:

$$c_t + a_{t+1} = w_t z_t + (1 + r_t)a_t$$

The wage $w_t$ and capital rent $r_t$ are determined at the competitive input factor markets. Households are subject to a borrowing constraint $a_{t+1} \geq 0$. I close the model by introducing a representative firm producing output from a constant returns-to-scale production function. The recursive formulation of the model is as follows:

$$\begin{aligned}
\text{(Household)} \quad & v(a, s; S, \Phi) = \max_{c,a'} \quad log(c) + \beta \mathbb{E}(v(a', s'; S', \Phi')) \\
& \text{s.t.} \quad c + a' = w(S, \Phi)z(s) + (1 + r(S, \Phi))a \\
& a' \geq 0, \quad \Phi' = G(\Phi, S, S') \\
\text{(Production sector)} \quad & \max_{K,L} \quad A(S)K^{\alpha}L^{1-\alpha} - w(S, \Phi)L - (r(S, \Phi) + \delta)K \\
\text{(Market clearing)} \quad & \hat{K}(S, \Phi) = \int a d\Phi \\
& \hat{L}(S, \Phi) = \int z d\Phi \\
\text{(Shock processes)} \quad & \mathbb{P}(s', S'|s, S) = \pi_{sS,s'S'}, \quad s, s' \in \{u, e\}, \quad S, S' \in \{B, G\}
\end{aligned}$$

All the variables with an apostrophe indicate variables in the future period. Following the original model assumption, $z = 0.25$ when $s = u$ and $z = 1$ when $s = e$. If $S = B$, I assume $A = 0.99$, and when $S = G$, $A = 1.01$.[3]

## 2.2 Intuition behind the methodology

In this section, I explain the basic intuition behind the methodology. For this, I first briefly describe the methodology. Suppose I simulate $T$ periods of aggregate shocks $\{A_t\}_{t=0}^{T}$, and hypothetically the simulated path is long enough to make almost

[3]For brevity, I omit the explanation of the other parameter levels.

all the possible equilibrium allocations happen on the simulated path.[4] Then, I start from guessing the following three time series: 1) value functions, $\{V_t^{(0)}\}_{t=0}^T$, 2) distributions of individual states $\{\Phi_t^{(0)}\}_{t=0}^T$, and 3) prices $\{p_t^{(0)}\}_{t=0}^T$. Using these guesses, I solve the allocations backward from the terminal period $T$ to obtain $\{V_t^*\}_{t=0}^T$, and simulate the economy forward using the solution. The forward simulation generates the time series of the distribution of individual states $\{\Phi_t^*\}_{t=0}^T$ and prices $\{p_t^*\}_{t=0}^T$ from the market-clearing conditions. Using these, I update the guess to move on to the next iteration, $\{V_t^{(1)}, \Phi_t^{(1)}, p_t^{(1)}\}_{t=0}^T$.

Now, suppose that I've run the $n^{th}$ iteration and that I am now at the $(n+1)^{th}$ iteration at period $t$ after solving the problem backward from the terminal period $T$ until period $(t+1)$. On the simulated aggregate state path, suppose that the shock realization at period $t+1$ is $G$, $S_{t+1} = G$ ($A_{t+1} = 1.01$). For the problem of an agent at $t$, a macroeconomist needs to construct a rationally expected future value function $\mathbb{E}_t\widetilde{V}_{t+1}$. However, this is a difficult task because only $V_{t+1}(\cdot, S = G)$ is available from the backward solution, while $V_{t+1}(\cdot, S = B)$ is not. This is natural as only one shock can be realized in a period. I define this unobserved values $V_{t+1}(\cdot, S = B)$ as a counterfactual conditional value function.

In the standard state-space-based approach, this problem is handled by replacing the time index with the distribution or sufficient statistics and specifying a law of motion in these aggregate states. Then, the counterfactual conditional value function is obtained by interpolating the unconditional value function at the predicted future state. Therefore, the accuracy of this predicted future state from the law of motion determines the accuracy of the solution. However, before obtaining the solution and simulating the economy based on the solution, it is hardly known whether the law of motion is correctly specified or not. Then, if the law of motion turns out to be incorrect, a researcher needs to restart solving the problem from scratch, coming up

[4]In theory, an infinitely-long simulation needs to be considered, but for the illustrative purpose, I consider a $T$-period long simulation. Later in the application, a long-enough finite simulation is used as an approximation for the infinitely-long ergodic simulation.

with a new guess about the law of motion. However, a proper guess is difficult to obtain, as there is an infinite degree of freedom in the new guess. Particularly, there are two types of difficulties in this step. One is about which statistics to include in the law of motion; the other is about what parametric forms to choose for the law of motion. Unless the aggregate dynamics are well-known to be log-linear, as in Krusell and Smith (1998), this problem cannot be easily resolved.

Then, I consider a new approach where the counterfactual conditional value function is obtained from the value function of another period $\widetilde{t}+1$ in which the endogenous aggregate state is exactly the same as the period $t+1$, but the counterfactual shock is realized:

$$\Phi^{(n)}_{\widetilde{t}+1} = \Phi^{(n)}_{t+1}$$
$$S_{\widetilde{t}+1} = B \neq G = S_{t+1}$$

Then, all the aggregate states of the realized state of period $\widetilde{t}+1$ are identical to the ones in the counterfactual state of period $t+1$. Thus, I am given that

$$V^{(n)}_{\widetilde{t}+1}(\cdot, S = B) = V^{(n)}_{t+1}(\cdot, S = B).$$

Importantly, $V^{(n)}_{\widetilde{t}+1}(\cdot, S = B)$ is the observed *factual* conditional value function available in the $n$th iteration. As both $V^{(n)}_{t+1}(\cdot, S = G)$ and $V^{(n)}_{t+1}(\cdot, S = B)(= V^{(n)}_{\widetilde{t}+1}(\cdot, S = B))$ are available, the rationally expected future value function $\mathbb{E}_t\widetilde{V}_{t+1}$ can be correctly constructed. Even when the aggregate shock process is discretized finer than two grid points, the rationally expected future value function can be obtained using the same procedures. Due to the ergodic theorem, if a simulated path is long enough, the existence of such period $\widetilde{t}+1$ is almost surely guaranteed.

In this new approach, a law of motion does not need to be specified to construct the rational expected future value function. As long as the period $\widetilde{t}+1$ that mimics the counterfactual realization of $t+1$ is identified, the problem can be solved. For this step, tracking $\{\Phi_t^{(n)}\}_{t=0}^T$ is important, as it allows us to identify the period $\widetilde{t}+1$. In the following section, I elaborate on the detailed steps to implement the repeated transition method.

## 2.3 Algorithm

I simulate a single path of exogenous aggregate TFP shocks for a long-enough period $T$, $\mathbb{A} = \{A_t\}_{t=0}^T$, using the aggregate transition matrix $\pi^A$. So, I also have the time series of the corresponding aggregate states, $\mathbb{S} = \{S_t\}_{t=0}^T$, where $S_t \in \{B, G\}$. The aggregate transition matrix is as follows:[5]

$$\pi^A = \begin{bmatrix} \pi_{B,B} & \pi_{B,B} \\ \pi_{G,B} & \pi_{G,B} \end{bmatrix} = \begin{bmatrix} 0.875 & 0.125 \\ 0.125 & 0.875 \end{bmatrix}$$

For the brevity of notation, I define a price vector $p_t := (w_t, r_t)$. I define a time partition $\mathcal{T}(S)$ that groups periods with the same shock 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 repeated transition method is as follows:

Step 1. Guess on the paths of the value functions, the state distributions, and the prices. $\{V_t^{(n)}, \Phi_t^{(n)}, p_t^{(n)}\}_{t=0}^T$.[6]

Step 2. Solve the model backward from the terminal period $T$ in the following sub-steps. The explanation is based on an arbitrary period $t$. Without a loss of generality, 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

[5] The transition matrix is from Krusell and Smith (1998).

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

$t+1$ ($S_{\widetilde{t}+1} = B$):[7]

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

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.

Step 3. Using $\{a_{t+1}^*\}_{t=0}^T$, simulate forward the time series of the distribution of the individual states $\{\Phi_t^*\}_{t=0}^T$ starting from $\Phi_0^* = \Phi_0^{(n)}$.[8]

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 implied prices $\{p_t^*\}_{t=0}^T$.[9]

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

$$\sup_{BurnIn \le t \le T - BurnIn} ||p_t^* - p_t^{(n)}||_\infty < tol$$

[7]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.

[8]In this step, I use the non-stochastic simulation method (Young, 2010).

[9]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. In Section 3, I use this algorithm to solve the model in Khan and Thomas (2008). In the algorithm of Khan and Thomas (2008), 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 algorithm, 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 repeated transition method significantly saves computation time. I discuss further the computational gain in Section 3.

Note that the distance is measured after excluding the burn-in periods at the beginning and the end of the simulation path. This is an adjustment to handle a potential bias from the imperfect guesses on the terminal period's value function $V_T^{(n)}$ and the initial period's distribution $\Phi_0^{(n)}$.

If the distance is smaller than the tolerance level, the algorithm is converged. Otherwise, I make the following updates on the guess:[10]

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

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

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

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.

$(\psi_1, \psi_2, \psi_3)$ are the parameters of convergence speed in the algorithm. If $\psi_i$ 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), I found uniformly setting $\psi_i$ around 0.8 guarantees convergence at a fairly high convergence speed. However, if a model is highly nonlinear, as in the baseline model in Section 4, the convergence speed needs to be controlled to be much 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

[10] 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:

$$log(p_t^{(n+1)}) = log(p_t^{(n)})\psi_1 + log(p_t^*)(1-\psi_1)$$

particularly more nonlinear than the others.

As can be seen from the convergence criterion in Step 5, the algorithm stops only when the expected allocation paths are close enough to the simulated allocation paths. Therefore, once the convergence is achieved, the accuracy of the solution is guaranteed. If the accuracy is measured in $R^2$ or in the mean-squared errors, as in Krusell and Smith (1998), the repeated transition method features $R^2$ of unity, and its mean-squared error becomes negligibly different than zero.

After the equilibrium allocations are computed over the in-sample path $\mathbb{A}$, I estimate the implied law of motion from the in-sample allocations. The law of motion can potentially take any nonlinear form. Then, using the fitted law of motion, equilibrium allocations are computed over out-of-sample paths of simulated aggregate shocks.

## 2.4 A sufficient statistic approach

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

However, if there are sufficient statistics that can perfectly represent a period's endogenous aggregate state, such as aggregate capital stock in Krusell and Smith (1998), the computational efficiency can be substantially improved.[11] This is because I can find period $\widetilde{t}+1$ by only comparing the distance between these sufficient statistics instead of the distributions. For example, in Krusell and Smith (1998), the aggregate capital is the sufficient statistics, which makes Step 2-a easier:

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

As the algorithm relies on the ergodic theorem, a sufficiently long period of simu-

[11] I discuss under which condition the sufficient statistics approach can be used in Section 2.5

lation is needed for accurate computation. However, in practice, the simulation still ends in finite periods. Therefore, the period $\widetilde{t}+1$ that shares exactly identical sufficient statistics 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 endogenous aggregate allocation is closest to the one in period $t+1$ from above, but the shock realization is different from period $t+1$:

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

Similarly, find $\widetilde{t}^{dn}+1$ where the endogenous aggregate allocation is closest to the one in period $t+1$ from below, but the shock realization is different from period $t+1$:

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

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

$$\omega = \frac{K_{t+1}^{(n)} - K_{\widetilde{t}^{dn}+1}^{(n)}}{K_{\widetilde{t}^{up}+1}^{(n)} - K_{\widetilde{t}^{dn}+1}^{(n)}}$$

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

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

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 simulation path).

## 2.5 A sufficient condition for the sufficient statistics

In this section, I analyze under which condition the sufficient statistic can replace the entire distribution in the repeated transition method to allow the sufficient statistics approach (Section 2.4). In Krusell and Smith (1998), the law of motion in the entire distribution is sharply approximated by the law of motion in the aggregate capital stock. This is one example where a sufficient statistic can completely represent the infinite-dimensional object. Likewise, various research in the literature has considered sufficient statistics to overcome the curse of dimensionality, but there has been little theoretical explanation of when such an approximation can be used. Proposition 1 provides a sufficient condition for using sufficient statistics in the repeated transition method.

**Proposition 1** (A sufficient condition for the sufficient statistics)**.**
*For a sufficiently large $T$, if there exists a time series of an aggregate allocation $\{x_t\}_{t=0}^{T}$ such that for each time partition $\mathcal{T}_S = \{t|S_t = S\}$, $\forall S \in \{B, G\}$ and for $\forall(a, z)$,*

$$(i) \quad x_{\tau_0} < x_{\tau_1} \iff V^{(n)}_{\tau_0}(a, z) < V^{(n)}_{\tau_1}(a, z) \textit{ for any } \tau_0, \tau_1 \in \mathcal{T}_S$$

*or*

$$(ii) \quad x_{\tau_0} < x_{\tau_1} \iff V^{(n)}_{\tau_0}(a, z) > V^{(n)}_{\tau_1}(a, z) \textit{ for any } \tau_0, \tau_1 \in \mathcal{T}_S$$

*then $x_t$ is the sufficient statistics of the endogenous aggregate state $\Phi_t$ for $\forall t$. In other words, for $\forall t \in \mathcal{T}_S$,*

$$\arg \inf_{\tau \in \mathcal{T}_S} ||\Phi_\tau^{(n)} - \Phi_t^{(n)}||_\infty = \arg \inf_{\tau \in \mathcal{T}_S} ||x_\tau - x_t||_\infty.$$

*Proof.*

See Online Appendix. ■

Proposition 1 states that if a time series $\{x_t\}_{t=0}^T$ monotonically ranks the level of the corresponding period's value function for each individual state, $x_t$ is the sufficient statistic of time period $t$ in the repeated transition method. The intuition behind the proposition is as follows. Suppose a situation where a researcher is searching for a value function to build a rationally expected future value function. If a time index of the correct counterfactual period to use is explicitly given as $\tau$ to a researcher, then the researcher can easily identify which value function to use, as all value functions are indexed by time. So, in this case, $V_\tau$ is trivially the one to use.

Now instead of $\tau$, suppose the level of $x_\tau$ is known to the researcher. Then, similar to the prior situation where $\tau$ is known, the researcher can identify which value function to use because the ranking information uniquely pins down the corresponding value function due to the strict monotonicity. For example, if two periods $\tau_0$ and $\tau_1$ share the same level of $x_t$, thus $x_{\tau_0} = x_{\tau_1}$, then the strict monotonicity says $V_{\tau_0} = V_{\tau_1}$. If this is not the case ($V_{\tau_0} \neq V_{\tau_1}$), then either the ergodicity or strict monotonicity assumption is violated, and this is the key idea of the proof.

To summarize the theoretical results in this section, once the ranking information across the different periods' value functions is known, one can exactly pin down which period's value function to use. This is also a practically desired feature for the implementation, as the strict monotonicity of value functions in the sufficient statistic makes it feasible to smoothly interpolate the value functions along the sufficient statistic (2-a' and 2-b' in Section 2.4).

The sufficient condition provides a theoretical ground to understand how a sufficient statistic approach works in the repeated transition method. In the quantitative analysis of the baseline model in Section 5.4, the monotonicity is quantitatively validated for the converged solution. However, the sufficient condition is not constructive for the algorithm as it cannot be checked prior to the implementation: the condition can be verified only after the solution converges. Also, the sufficient statistics in the repeated transition method do not imply that these statistics are only allocations to be considered in the law of motion in the state-space-based approach. This is because the former may not include sufficient information about the inter-temporal dynamics in the endogenous aggregate state variables.[12]

# 3 Accuracy of the repeated transition method

This section compares the equilibrium allocations obtained from the repeated transition method and the ones from the methods of Krusell and Smith (1998) and Krusell and Smith (1997). In the computation, parameters are set as in the benchmark model in Krusell and Smith (1998) without idiosyncratic shocks in the patience parameter $\beta$. Both of the algorithms are designed to stop when the largest absolute difference between the simulated average capital stock and the expected average capital stock is less than $10^{-6}$.

In the converged solution, the mean squared difference in the solutions between the repeated transition method and Krusell and Smith (1998) algorithm is around $2 * 10^{-4}$. It takes around 20 minutes for the repeated transition method to converge under the convergence speed parameter $\psi_1 = \psi_2 = \psi_3 = 0.8$; it takes around 20 mins for Krusell and Smith (1998) algorithm.[13] The convergence speed might change

[12] When I fit the nonlinear aggregate dynamics of sufficient statistics obtained from the repeated transition method to the parametric/non-parametric law of motion in Section 5.3, the fittest specification includes multiple lagged terms of the sufficient statistics. However, the sufficient statistics for each time period in the repeated transition method is just a single-dimensional aggregate allocation.

[13] This computation is done in 2015 MacBook Pro laptop with a 2.2 GHz quad-core processor

depending on the updating weight.

Figure 1 plots the expected path (Predicted) and the simulated path (Realized) of aggregate capital $K_t$ obtained from the repeated transition method and the simulated path from Krusell and Smith (1998).[14] The expected path refers to $\{V_t^{(n)}, \Phi_t^{(n)}, p_t^{(n)}\}_{t=0}^T$ in Section 2.3, and the simulated path indicates $\{V_t^*, \Phi_t^*, p_t^*\}_{t=0}^T$. As can be seen from all three lines hardly distinguished from each other, the repeated transition method computes almost identical equilibrium allocations as Krusell and Smith (1998) algorithm at a similar speed. This is because the log-linear specification almost perfectly captures the actual law of motion in Krusell and Smith (1998). Thus, their algorithm with the log-linear specification can accurately compute the solution at high speed.

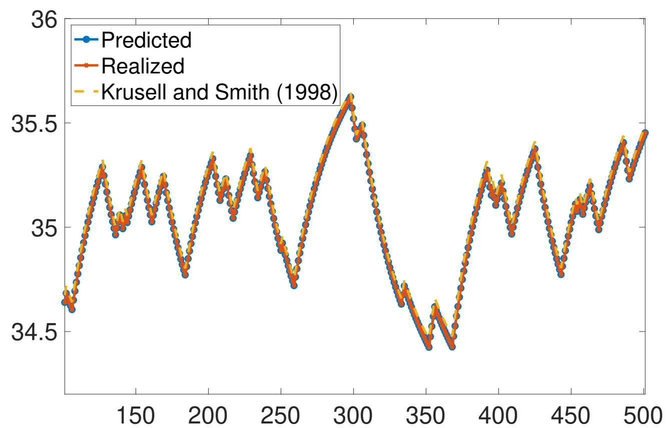



Figure 1: Computed dynamics in aggregate wealth (Krusell and Smith, 1998)

However, the repeated transition method outperforms Krusell and Smith (1997) algorithm when the market-clearing condition is not trivial, as in the model of Khan and Thomas (2008).[15] This is because the non-trivial market-clearing condition requires an extra loop to find an exact market-clearing condition in each iteration, while the repeated transition method does not.

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

[15]Krusell and Smith (1997) algorithm is a variant of the algorithm in Krusell and Smith (1998), which is applicable to models with non-trivial market-clearing conditions. Khan and Thomas (2008) uses this algorithm.

I solve the model in Khan and Thomas (2008) using both the repeated transition method and the Krusell and Smith (1997) algorithm with an external loop for the non-trivial market-clearing condition. Both of the algorithms are designed to stop when the following criterion is satisfied:[16]

$$\max\{\sup_t\{||p_t^* - p_t^{(n)}||\}, \sup_t\{||K_t^* - K_t^{(n)}||\}\} < 10^{-6}$$

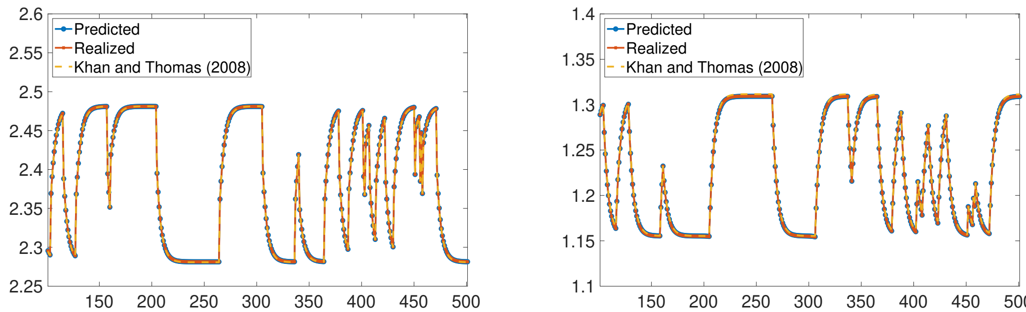
Figure 2 plots the dynamics of price $p_t$ and aggregate capital stock $K_t$ computed from the repeated transition method and Krusell and Smith (1997) algorithm. For the allocations computed from the repeated transition method, both the predicted time series and the realized time series are plotted. 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 repeated transition method and Khan and Thomas (2008) is less than $10^{-5}$.

In the application of the repeated transition method, I use $\psi_1 = \psi_2 = \psi_3 = 0.9$ for the updating rule, which is higher than the previous application. The reason for using this conservative updating rule is because the model in Khan and Thomas (2008) features a strong general equilibrium effect; dramatic updates in the price might lead to divergence. The repeated transition method took around 20 minutes to converge on average, while Krusell and Smith (1997) algorithm converged in around 5 to 6 hours on average. The convergence speed might change depending on the updating weight.

# 4 Application: Real business cycle model with the corporate saving glut

In this section, I analyze the business cycle implications of the rising corporate

[16]The terminal condition is slightly different from the one in Step 5 of Section 2.3. Likewise, the terminal condition can be flexibly adjusted based on different combinations of $V_t^{(n)}$, $\Phi_t^{(n)}$, and $p_t^{(n)}$.




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

Figure 2: Computed dynamics in aggregate capital stocks (Khan and Thomas, 2008)

cash holding using the heterogeneous-firm real business cycle model. There are two reasons why the algorithm is applied to this particular model. The first is the rising importance of the corporate cash holding in the U.S. economy. Figure 3 plots the time series of the aggregate cash holding to GDP ratio, where nominal cash holding is from the Flow of Fund data, and the nominal GDP is from NIPA.[17] As seen from the sharply rising trend in the ratio, the corporate cash holding has risen substantially faster than the output.[18] Have these rising corporate cash holdings affected the business cycles in the U.S.? This paper investigates the answer to this question through the lens of the business cycle model with the corporate saving glut.

The second reason is the sharp contrast in the saving patterns between the models with the heterogeneous firms and the heterogeneous households. Due to the external financing cost, which generates a precautionary motivation for holding cash stock, the firms' saving pattern partly mimics the household's savings. However, as the internal financing does not trigger any adjustment cost, the contemporaneous component of the objective function is concave only in the limited part of the domain, unlike the counterpart in the household's utility maximization problem. Therefore, a satiation point exists for hoarding cash, leaving a kink in the inter-temporal saving decision.

[17]The calculation of the aggregate corporate cash stock is explained in Appendix C.

[18]The corporate finance literature has investigated reasons for rising corporate cash holding. However, analyzing these different reasons is out of the scope of this paper.

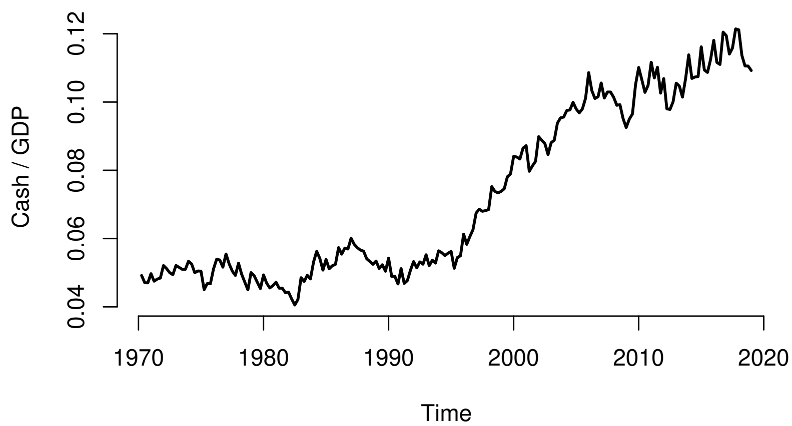



Figure 3: The time-series of the corporate cash-to-GDP ratio

This kink is the starkest difference between the baseline model and the heterogeneous-household models, which is to be investigated further in detail.

### 4.1 Technology

There is a continuum of measure one of ex-ante homogeneous firms that hoard cash and produces business outputs. For simplicity, I assume a firm operates using only labor input. This can be understood as an equivalent setup to a model where a firm uses both capital and labor inputs, but the optimal capital demand decision is already frictionlessly internalized in the labor demand decision. Consistent with this explanation, I set the labor share (equivalent to the span of control parameter) at $\gamma = 0.85$ in the quantitative analysis, which is greater than the standard labor share and captures internalized capital demand decision.

The business output is produced by the following Cobb-Douglas production function:

$$f(n_{it}, z_{it}; A_t) = z_{it} A_t n_{it}^{\gamma}$$

where $n_{it}$ is labor demand; $\gamma < 1$ is the span of control parameter; $z_{it}$ and $A_t$ are idiosyncratic and aggregate productivities, respectively. Each firm needs to pay a fixed operation cost $\xi > 0$ in each period.

The log of idiosyncratic productivity shock process $\{z_{it}\}$ follows an AR(1) process:

$$log(z_{it+1}) = \rho_z log(z_{it}) + \epsilon_{it+1}, \quad \epsilon_{it+1} \sim_{i.i.d} N(0, \sigma_z)$$

For computation, the idiosyncratic productivity process is discretized by the Tauchen method.[19] The stochastic aggregate productivity process is from (Krusell and Smith, 1998):[20]

$$\Gamma_A = \begin{bmatrix} 0.8750 & 0.1250 \\ 0.1250 & 0.8750 \end{bmatrix}$$
$$A_t \in \{A_B, A_G\} = \{0.99, 1.01\}.$$

## 4.2 External financing cost

A firm earns operating profit and decides how much to distribute as a dividend $d_{it}$ to equity holders (a representative household). The remaining part of the operating profit after the dividend payout is used to adjust cash holding, $ca_{t+1}/(1 + r^{ca}) - ca_t$. The future cash holding is discounted at an internal discount rate $r^{ca} > 0$ as cash is not traded in the market across the firms. $r^{ca}$ is an exogenous parameter and is assumed to be lower than the market interest rate $r_t$. The cash holding level is assumed to be non-negative $ca_t \geq 0$. Thus, the model features a standard incomplete market assumption with the borrowing limit as in Aiyagari (1994).

If a dividend is determined to be negative, then a firm is financing through an external source, which incurs extra pecuniary cost $C(d_{it})$ (Jermann and Quadrini,

[19] I discretize it using equally-spaced nine grid points within the two-standard deviation range around the mean.

[20] The repeated transition method works for a finer discretization than two grid points. For example, Lee (2022) uses a finer discretization (five grid points) for the repeated transition method. However, to preserve the symmetry between the corporate cash-holding model and the household saving model (Krusell and Smith, 1998), I assume the same aggregate productivity process.

2012; Riddick and Whited, 2009). This external financing cost is specified as follows:

$$C(d_{it}) := \frac{\mu}{2}\mathbb{I}\{d_{it} < 0\}d_{it}^2$$

Thus, the net dividend is $d_{it} - \frac{\mu}{2}\mathbb{I}\{d_{it} < 0\}d_{it}^2$. It is worth noting that this net dividend function belongs to $\mathbb{C}^2$ class as it smoothly changes the slope at $d_{it} = 0$ without a kink. Therefore, the standard theory of the concave utility of the household seamlessly applies to the model.

If there is no external financing cost, hoarding cash is not the desired option for a firm because it is more expensive than receiving the dividend $\left(\frac{1}{1+r^{ca}} > \frac{1}{1+r_t}\right)$. However, due to the presence of an external financing cost, a firm has a precautionary motivation to hoard cash, saving for rainy days (low $z_t$ or low $A_t$). Therefore, the firms smooth their dividend payout in equilibrium. Consistent with this, rich empirical evidence has been documented for corporate dividend smoothing behavior in the corporate finance literature (Leary and Michaely, 2011; Bliss et al., 2015). Especially, Leary and Michaely (2011) empirically showed that cash-rich firms smoothen their dividend significantly more than others. The equilibrium patterns in this model can match these empirical patterns.

## 4.3 Recursive formulation

At the beginning of each period, a firm $i$ is given with a cash holding $ca_{it}$ and an idiosyncratic productivity level $z_{it}$. Thus, the individual state variable $x_{it}$ is as follows:

$$x_{it} = \{ca_{it}, z_{it}\}$$

All firms rationally expect the future and are aware of the full distribution of the firm-level state variables. The aggregate state variable $X_t$ is as follows:

$$X_t = \{A_t, \Phi_t\}$$

where $A_t$ is the aggregate productivity, and $\Phi_t$ is the distribution of the individual state variable $x_{it}$.

The recursive formulation of a firm's problem is as follows:

$$\begin{aligned}
\text{[Firm]} \quad & J(ca, z; X) = \max_{ca', d} \quad d - C(d) + \mathbb{E}(q(X, X')J(ca', z'; X')) \\
& \text{s.t.} \quad d + \frac{ca'}{1 + r^{ca}} = \pi(z; A, \Phi) + ca \\
& \qquad ca' \geq 0, \quad \Phi' = G(\Phi, A) \\
\text{[Operating profit]} \quad & \pi(z; A, \Phi) := \max_n zAn^\gamma - w(A, \Phi)n - \xi \\
\text{[Idiosyncratic productivity]} \quad & z' = G_z(z) \text{ (AR(1) process)} \\
\text{[External financing cost]} \quad & C(d) := \frac{\mu}{2}\mathbb{I}(d < 0)d^2 \\
\text{[Aggregate state]} \quad & X := \{A, \Phi\}
\end{aligned}$$

where $J$ is the value function of a firm; $ca$ and $z$ are cash holding and idiosyncratic productivity, respectively; $A$ is the aggregate productivity; $\Phi$ is the distribution of the individual state variables; $w$ and $q$ are wage and stochastic discount factor which are functions of aggregate state variables $X = \{A, \Phi\}$.

## 4.4 Equilibrium

I close the model by introducing a stand-in household that holds equity as wealth and saves on equity. The household consumes and supplies labor and rationally expects the future aggregate states. The income sources of the household are labor work and dividends from equity.

The recursive formulation of the representative household's problem is as follows:

$$
\begin{aligned}
V(a;X) = \max_{c,a',l_H} \quad & log(c) - \eta l_H + \beta \mathbb{E}^{A'} V(a';X') \\
\text{s.t. } & c + \int \Gamma_{A,A'} a' q(X,X') dX' = w(X) l_H + a \\
& G(X) = \Phi' \\
& G_A(A) = A' \quad \text{(AR(1) process)}
\end{aligned}
$$

where $V$ is the value function of the household; $a$ is wealth; $c$ is consumption; $a'$ is a future saving level; $l_H$ is labor supply; $w$ is wage, and $q$ is the stochastic discount factor. The household is holding the equity of firms as their wealth.

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

$$
\begin{aligned}
\text{(Labor market)} \quad l_H(X) &= \int n(ca, z; X) d\Phi \\
\text{(Equity market)} \quad a(X) &= \int (J(ca, z; X) + C(d(ca, z; X))) d\Phi
\end{aligned}
$$

The external financing costs and the aggregate firm values jointly form the supply of equity. In the market clearing condition, the supply meets the demand for equity in the form of household wealth.

The model does not assume a centralized market for cash holding. Therefore, $r^{ca}$ is not endogenously determined in the market. This is a realistic assumption, as a firm's cash holding is not tradable across firms. I interpret this setup as the cash holding return is determined by each firm's idiosyncratic financing status independently from the centralized capital market condition. $r^{ca}$ is the average level of the idiosyncratic financing cost.[21]

[21]For simplicity, the model is abstract from the heterogeneity in the financing cost. All the exogenous heterogeneity is loaded on the heterogeneous productivities.

## 4.5 The role of market incompleteness and the financial constraint: Comparison with Aiyagari (1994)

In this section, I study the individual firm's cash hoarding patterns in the stationary equilibrium. This analysis is essential to understand why the model would feature highly nonlinear dynamics under aggregate uncertainty. Due to the external financing cost, a firm saves cash out of precautionary motivation. However, there exists a target cash level, which the optimal cash holding does not go beyond. This is because after the target level of cash, holding cash becomes costlier: if a firm holds a large stock of cash, then the future risk of external financing is almost perfectly hedged while carrying additional cash bears only a smaller return than the market interest rate. Proposition 2 states the existence of the target cash level.[22]

**Proposition 2** (The existence of the target cash-holding level)**.**
*Suppose policy functions are non-trivial:* $ca'(ca, z) > 0$ *and* $d(ca, z) > 0$ *for some* $ca > 0$*, given* $z$*. Then, there exists* $\overline{ca}(z) > 0$ *such that* $ca'(ca, z) \leq \overline{ca}(z)$ *for* $\forall ca \geq 0$*.*

*Proof.* See Online Appendix. ■

Therefore, the future cash holding policy function features a kink due to the flat region after the target cash holding level.[23] Figure 4 shows the future cash holding policy function for the lowest and highest productivity firms. For the highest productivity firms, due to the persistence of the productivity shock, the target cash holding level is lowest: they are the least concerned firms about the future external financing cost. As can be seen from the figure, the policy function crosses the 45-degree ray. That is, the highest productivity firms with little cash increase the cash

[22] It is worth noting that the implication of Proposition 2 is different from Proposition 4 of Aiyagari (1993). Proposition 4 of Aiyagari (1993) implies that a household with a excessively large wealth would gradually decrease the wealth. In contrast, Proposition 2 implies that a firm with a excessively large cash stock would immediately reduce the cash stock to the target level.

[23] The existence of the target cash holding level is similar to the prediction of the consumption buffer stock model (Carroll, 1997).

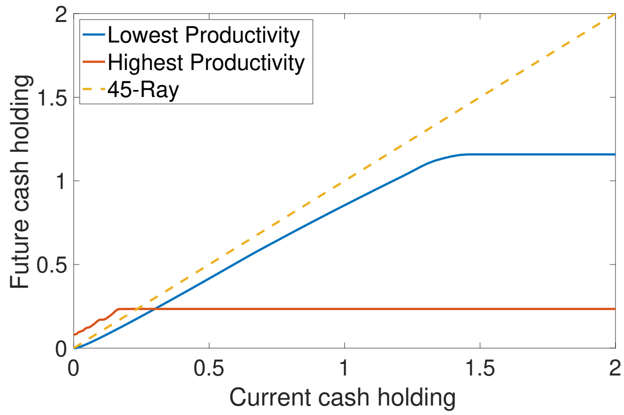



Figure 4: Cash holding policy function of individual firms

holding until they reach the target level. On the other hand, the lowest productivity firms' target cash holding level is the highest, but their low profit makes them decrease their future cash holding. Therefore, the lowest productivity firms' policy function does not cross the 45-degree ray.

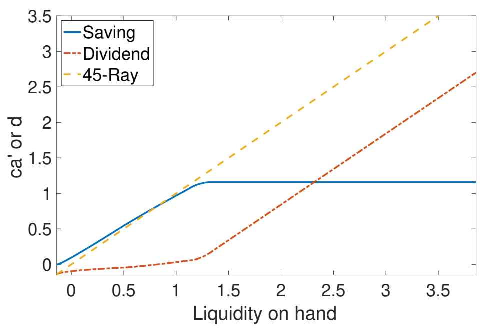
The kinked cash-holding policy function becomes starkly contrasted with the wealth accumulation pattern of households in Aiyagari (1994). To make a direct comparison feasible, I define the liquidity on hand, which is a firm-side counterpart of the total resource in Aiyagari (1994):

$$\text{Liquidity on hand}_t := \underbrace{\pi_t}_{\text{Liquidity from operating profit}} + \underbrace{ca_t}_{\text{Cash}}.$$

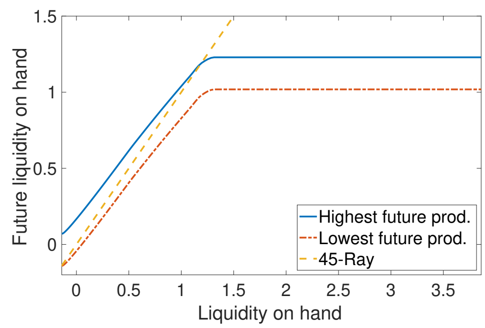
Figure 5 plots the saving and dividend policies in panel (a) and the future liquidity on hand in panel (b) as functions of the liquidity on hand. This figure is the firm-side counterpart of Figure I in Aiyagari (1994).[24] For a sharp illustration, I only plot the policy functions for a firm at the lowest productivity ($z = min\mathbb{Z}$). For both firms in my model and households in Aiyagari (1994), there are cases where the borrowing limit is bound, leaving the saving policy flat near the borrowing limit. However, that

[24]Dividend $d$ is the counterpart of consumption $c$, and future cash holding $ca'$ is the counterpart of future wealth $a_{t+1}$.

region spans only a tiny range for the firm side, so it is not visible in panel (a) of Figure 5.[25] If a firm does not hold enough liquidity on hand, an extra dollar increase in liquidity goes to both saving and dividend. Especially a greater amount goes to the saving than dividends. However, if a firm holds enough liquidity on hand, additional liquidity is solely paid out as a dividend, as can be seen from the parallel dividend curve to the 45-degree line for the high liquidity on hand region. In panel (b), the future liquidity on hand displays a kinked pattern over the current liquidity on hand, similar to the dynamics of the cash policy function in Figure 4.




(a) Allocation of liquidity on hand




(b) Evolution of liquidity on hand

Figure 5: Liquidity on hand and the policy functions (when $z = min\mathbb{Z}$)

The kinked policy function plays an important role in the business cycle, as it generates nonlinearity in the business cycle. Depending on the fraction of firms that are located over the flat policy region, the aggregate cash dynamics vary. In other words, the aggregate fluctuations in this model crucially depend on the endogenous state $\Phi$, the distribution of individual states.

# 5 Quantitative analysis

In this section, I quantitatively analyze the recursive competitive equilibrium al-

[25]Households behave in a similar pattern in Krusell and Smith (1998). However, in both Aiyagari (1994) and Krusell and Smith (1998), the fraction of these constrained households are small as well. Especially, this is one of the major reasons why the aggregate dynamics in Krusell and Smith (1998) display only a negligible nonlinearity.

locations computed from the repeated transition method. For easier computation, I first normalize the firm's value function by contemporaneous consumption $c_t$ following Khan and Thomas (2008). I define a price $p_t := 1/c_t$ and the normalized value function $\tilde{J}_t := p_t J_t$. From the intra-temporal and inter-temporal optimality conditions of households, I have $w_t = \eta/p_t$ and $q_t = \beta p_{t+1}/p_t$. Thus, $p_t$ is the only price to characterize the equilibrium. I take the sufficient statistics approach described in Section 2.4, and the aggregate cash holdings $CA_t$ (the first moment of the distribution of cash holding) is the sufficient statistics. I validate this approach by showing Proposition 1 is satisfied in Section 5.4.

## 5.1 Calibration

The model's key parameters are the external financing cost parameter $\mu$ and the operating cost parameter $\xi$. The external financing cost $\mu$ is identified from the corporate cash-to-output ratio. In the moment calculation, the aggregate cash stock is obtained from the Flow of Funds of the Federal Reserve Board, and the aggregate output (GDP) is from the National Income and Product Accounts (NIPA) from Bureau of Economic Analysis (BEA).[26] In the model, as $\mu$ increases, the corporate cash-to-output ratio increases due to the heightened precautionary motivation.

| Parameters | Target Moments | Data | Model | Level |
|---|---|---|---|---|
| $\mu$ | Corporate cash holding/Output (%) | 10.00 | 9.28 | 0.40 |
| $\xi$ | Consumption/Output (%) | 66.00 | 64.02 | 0.15 |
| $\eta$ | Labor supply hours | 0.33 | 0.34 | 3.90 |

Table 1: Calibration target and parameters

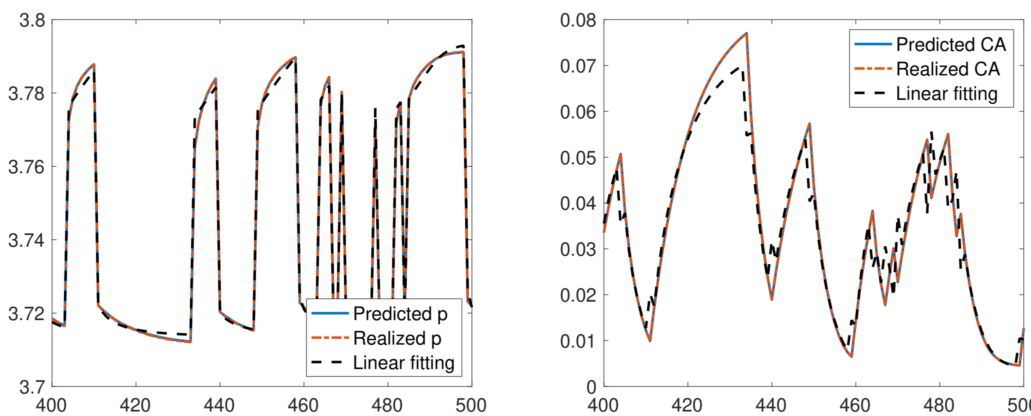
The identifying moment of the operating cost parameter $\xi$ is the consumption-to-output ratio. The consumption data is from NIPA.[27] As operating cost increases,

[26]The detailed definition of aggregate cash holding is available in Appendix C.
[27]Consumption includes both durable and non-durable consumptions.

the ratio decreases due to the reduced dividends. The calibrated parameters and the corresponding moments are summarized in Table 1. The other fixed parameters are summarized in Appendix B.

## 5.2 Nonlinear business cycle

Using the repeated transition method, I compute the recursive competitive equilibrium allocations over the simulated path of aggregate shocks. Using the aggregate cash stock, I take the sufficient statistics approach described in Section 2.4. The dynamics of the aggregate cash stocks are highly nonlinear for two reasons. First, the individual firm's cash holding policy function features a kink, as described in Section 4.5. Second, the general equilibrium effect does not strongly affect each firm's cash holding demand. It is because the price of cash holding is exogenously fixed at $r^{ca}$ as the cash is not allowed to be traded across the firms.




(a) Price $p_t$ $(= 1/C_t)$ (b) Aggregate cash stocks $CA_t$

Figure 6: Aggregate fluctuations in the economy

Figure 6 plots a part of the simulated path of the price $p_t$ (panel (a)) and aggregate corporate cash holding $CA_t$ (panel (b)) obtained from both the repeated transition method and the log-linear specification of the law of motion. The solid line plots the expected allocations (guess from the $n^{th}$ iteration), and the dot-dashed line plots the realized allocations (simulation based on the policy in $(n+1)^{th}$ iteration) in the repeated transition method. The dashed line represents the dynamics of the allocations in the log-linear specification of the law of motion. To obtain the parameters in the log-linear specification, I fit the equilibrium allocations from the repeated transition method into the log-linear specification, and the result is as follows:

$$log(CA_{t+1}) = -0.5742 + 0.9061 * log(CA_t), \quad \text{if } A_t = A_B, \text{ and } R^2 = 0.9971, \; MSE = 0.0017$$
$$log(CA_{t+1}) = -0.8949 + 0.6829 * log(CA_t), \quad \text{if } A_t = A_G, \text{ and } R^2 = 0.9823, \; MSE = 0.0039$$
$$log(p_t) = 1.3232 - 0.0018 * log(CA_t), \quad \text{if } A_t = A_B, \text{ and } R^2 = 0.8828, \; MSE = 0.0000$$
$$log(p_t) = 1.3093 - 0.0011 * log(CA_t), \quad \text{if } A_t = A_G, \text{ and } R^2 = 0.8928, \; MSE = 0.0000$$

In the repeated transition method, the expectation path and the realized path converge as can be seen from the figure. Therefore, $R^2$s are unity and the mean squared errors are as small as $10^{-6}$ for both $p_t$ and $CA_t$. On the other hand, the log-linear fitting results in low $R^2$ and high mean squared errors. This results show that the aggregate fluctuations in this economy are highly nonlinear.

One important reason for the nonlinearity is the missing general equilibrium effect on cash. If the price of cash is determined in the competitive market, the dynamics of aggregate cash stocks are smoothed. For example, when there is a surge of cash holding demand, the price of cash holding goes up to mitigate the surge and vice versa for the case of decreasing cash holding demand. In many of the models in the literature, this flattening force from the general equilibrium has been proven to be powerful enough to guarantee the log-linear specification as the true law of motion. One example is Khan and Thomas (2008), where the micro-level lumpiness is smoothed out by real interest rate dynamics. However, due to the missing general equilibrium effect, the log-linear prediction rule fails to capture the true law of motion in this paper.

On top of the nonlinearity, there is another complication in the model that the prototype algorithm of Krusell and Smith (1998) cannot simply address: there is a non-trivial market-clearing condition with respect to price $p_t$. Krusell and Smith (1997) suggests an algorithm to solve this problem by considering an external loop in the algorithm that solves market-clearing price $p_t$ in each iteration. This algorithm is known to successfully solve the log-linear models with non-trivial market-clearing conditions, such as Khan and Thomas (2008). However, due to the extra loop in each iteration, the algorithm entails high computation costs. In contrast, the repeated transition method tracks the implied price instead of the market clearing price on the simulated path. Therefore, the method does not require an extra loop for computing market-clearing price, so it saves a great amount of computation time. In the baseline model, computation time is reduced by a factor of 10.[28]

## 5.3 Recovering the true nonlinear law of motion

In this section, I recover the true law of motion from the converged equilibrium outcomes over the simulated path. Then, I test the validity of the true law of motion by fitting the law of motion into the out-of-sample simulation path.

Specifically, the following laws of motion are studied:

$$CA_{t+1} = G_{CA}(CA_t, CA_{t-1}, CA_{t-2}, \dots, CA_{t-n}; A_t)$$
$$p_t = G_p(CA_t, CA_{t-1}, CA_{t-2}, \dots, CA_{t-n}; A_t)$$

Table 2 reports the goodness of fitness ($R^2$) of the different specifications. The first five rows report the fitness when only the contemporaneous aggregate cash stock $CA_t$ is considered up to different polynomial orders. When a single argument is considered without the higher-order polynomials, $R^2$ gets as low as 0.8956 for $G_{CA}$ if the contemporaneous productivity state is $G$.[29] As the more higher-order polynomials are included, the better the fitness becomes. However, the fitness of the specification

[28]The KS algorithm takes around five hours to compute a converged solution when the simulation length is T = 500 and the cross-sectional grid of cash holding is 50 points. However, in the repeated transition method, it takes only around 30 minutes to make a convergence. For the fair comparison, the initial guess of the KS algorithm is from the log-linear relationship implied in the initial guess of the repeated transition method.

[29]This pure linear specification is different from the log-linear specification in Section 5.2.

of $G_p$ stops improving after a certain threshold. This shows that the true law of motion can be recovered only by including further historical allocations.

The bottom seven rows of Table 2 report the fitness of the law of motion when the additional lagged terms of the aggregate cash stock are considered. Up to the third order polynomials are included for each of lagged terms on top of the polynomial terms of the contemporaneous cash stocks up to the fifth order.[30] As more lagged terms are considered in the law of motion, the fitness improves, especially in $G_p$. However, only after the polynomials of the seven-period lagged aggregate cash stock are included in the law of motion, the accurate law of motion is recovered.

| | | | Goodness of fitness: $R^2$ | | | |
|---|---|---|---|---|---|---|
| | # of lagged | order | $CA_{t+1}: Good$ | $CA_{t+1}: Bad$ | $p_t: Good$ | $p_t: Bad$ |
| **Contemp.** | 0 | 1 | 0.8956 | 0.9452 | 0.9922 | 0.9966 |
| | 0 | 2 | 0.9839 | 0.9952 | 0.9927 | 0.9976 |
| | 0 | 3 | 0.9973 | 0.9995 | 0.9930 | 0.9976 |
| | 0 | 4 | 0.9993 | 0.9999 | 0.9932 | 0.9976 |
| | 0 | 5 | 0.9996 | 1.0000 | 0.9933 | 0.9976 |
| **Add. history** | 1 | 3 | 0.9999 | 1.0000 | 0.9987 | 0.9979 |
| | 2 | 3 | 0.9999 | 1.0000 | 0.9997 | 0.9984 |
| | 3 | 3 | 0.9999 | 1.0000 | 0.9998 | 0.9987 |
| | 4 | 3 | 0.9999 | 1.0000 | 0.9998 | 0.9991 |
| | 5 | 3 | 0.9999 | 1.0000 | 0.9998 | 0.9994 |
| | 6 | 3 | 0.9999 | 1.0000 | 0.9998 | 0.9996 |
| | 7 | 3 | 0.9999 | 1.0000 | 0.9998 | 0.9997 |

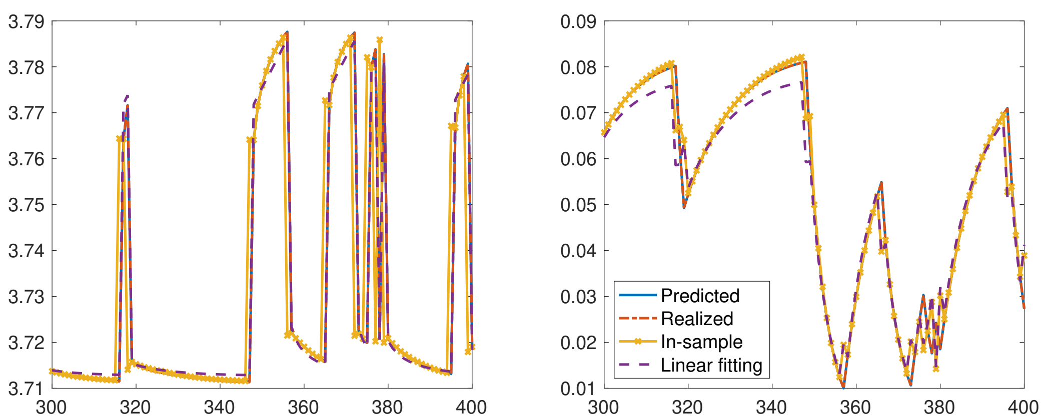
Table 2: The fitness of law of motion across different specifications

This exercise starkly shows the substantial nonlinearity of the aggregate fluctuations in this model. In the repeated transition method, the contemporaneous aggregate cash stock is used as a sufficient statistic of each period's cross-section. However, this does *not* imply that the true aggregate law of motion is a function of only the contemporaneous cash stock. The contemporaneous cash stock is rather a labeling of each period that correctly sorts the rankings of the value functions across the periods in the repeated transition method (Proposition 1). In Section 5.4, the monotonicity

[30]The results only negligibly change over different order choices.

of value functions in the contemporaneous cash stock is verified.

I validate the recovered law of motion by fitting it on the out-of-sample simulation path.[31] Specifically, I solve the model on another simulation path to obtain the converged equilibrium dynamics using the repeated transition method and compare the dynamics with the implied dynamics in the recovered true law of motion on the in-sample path. Figure 7 plots $p_t$ (panel (a)) and $CA_t$ (panel (b)) for 1) predicted time series (solid line), 2) realized time series (dot-dashed line), 3) time series implied by the recovered in-sample law of motion (solid line with ticks), and 4) time series implied by the linear law of motion (dashed line). The predicted time series and the realized time series are indistinguishably close to each other due to the convergence requirement of the repeated transition method. The time series implied by the recovered law of motion also closely track the converged equilibrium dynamics, validating the specification. The goodness of fitness ($R^2$) in the time series implied by the recovered law of motion is greater than 0.999 for both $G_{CA}$ and $G_p$ in all shock realizations.



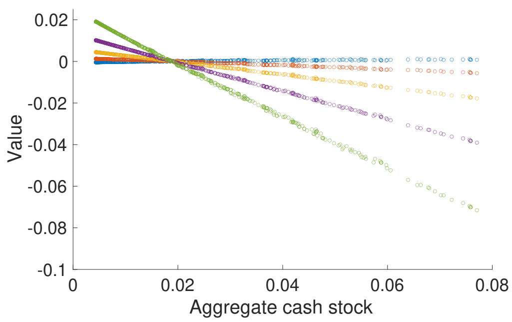

(a) Price $p_t$ $(= 1/C_t)$ (b) Aggregate cash stocks $CA_t$

Figure 7: Fitting into the out-of-sample path

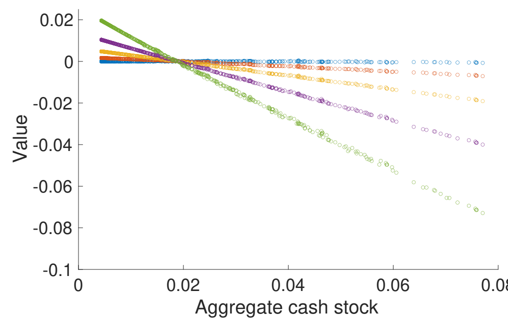
[31]The recovered true of motion refers to the specification that considers up to the seven-period lagged aggregate cash stocks, where $R^2$ is the highest in Table 2.

## 5.4 Monotonicity of the value function

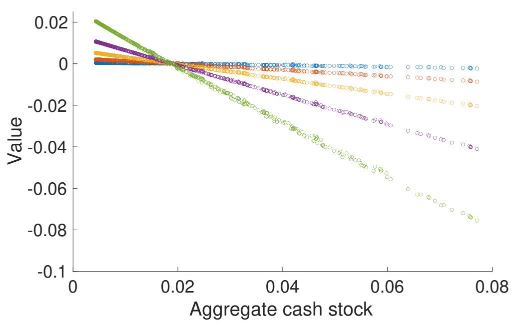
In this section, I check the monotonicity of the value function in the aggregate cash stock. The monotonicity is the sufficient condition for the aggregate cash stock to be used as a sufficient statistic (Proposition 1). Specifically, I check whether the value function is strictly monotone in the aggregate cash stock $CA_t$ for individual state variables $(a_t, z_t)$ and aggregate shock realizations $A_t$.



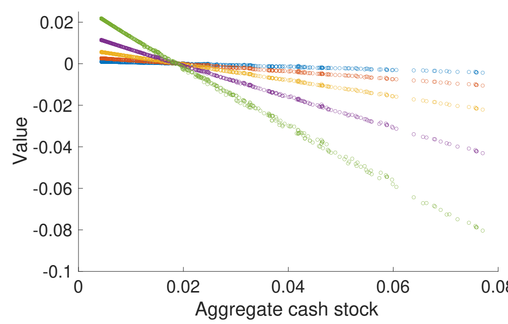

(a) Individual cash = 1e-8




(b) Individual cash = 0.09




(c) Individual cash = 0.23




(d) Individual cash = 0.47

Figure 8: Monotonicity of value function in aggregate cash stock

Figure 8 plots the level of the value function for different individual states for the aggregate shock realization $A_t = A_B$. Four different levels of individual cash stock $ca_t$ are considered, and the corresponding figures are panels (a),(b),(c), and (d).[32] Each panel is the scatter plot of value function levels for the five different levels of the idiosyncratic productivity $z_t$.[33] The horizontal axis is the aggregate cash stock $CA_t$

[32]The four different levels of $ca_t$ are 1e-8, 0.09, 0.23, and 0.47

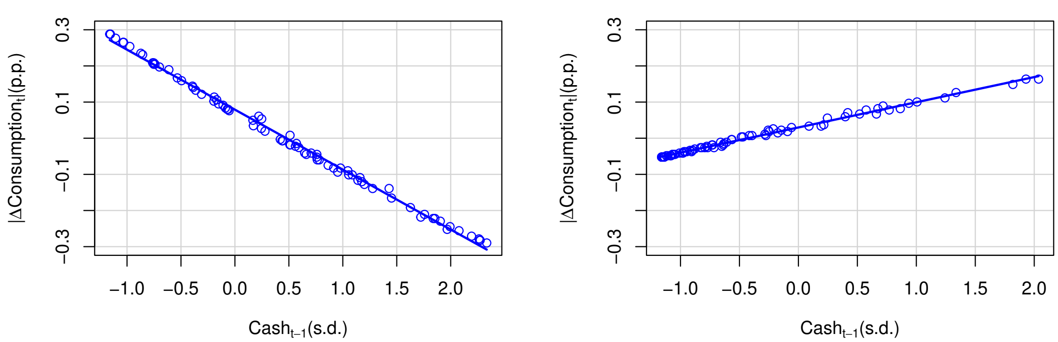
[33]The five different levels of $z_t$ are 0.7950, 0.8916, 1, 1.1215, and 1.2579.

and the vertical axis is the level of value function $V_t$. As can be seen from the figure, the scatter plot of the values forms five different linear lines. This shows that the value functions are strictly monotone in the aggregate cash, validating the qualification of the aggregate cash as the sufficient statistics in the repeated transition method.

## 5.5 Macroeconomic implications and empirical evidence

In this section, I analyze the role of corporate cash holdings on aggregate consumption fluctuations using the baseline model and support the model prediction from the empirical evidence.

In the model, the aggregate productivity $A_t$ can take one of two values $\{A_B, A_G\}$, and it follows a persistent process. I define the negative aggregate productivity shock as a TFP shift from $A_G$ to $A_B$ and the positive aggregate productivity shock as a TFP shift from $A_B$ to $A_G$. In the baseline model, depending on the cash stock a firm holds, the responsiveness of the firm-level dividends to the exogenous aggregate TFP shock changes. For example, when a firm is short of cash, a negative TFP shock in productivity makes a firm reduce dividends further than it would do when it has abundant cash stocks. It is because the firm with little cash needs to not only pay out dividends but save cash out of concern for the future.




(a) Negative shock (b) Positive shock

Figure 9: State-dependent shock responses of consumption

Then, due to this dividend channel, the responsiveness of household consumption becomes dependent on the aggregate cash stock. Figure 9 plots the relationship between the consumption responsiveness over the aggregate cash stocks separately for negative aggregate shock (panel (a)) and positive aggregate shock (panel (b)). In this model, the magnitude of aggregate shock is uniform at $|A_G - A_B|$(= 2% TFP shock).[34] Therefore, If the consumption shock responses are different across the periods, it is due to the endogenous state dependence of the responsiveness rather than the shock magnitude variation. Then, I separately collect the periods of negative aggregate shock and positive aggregate shock and compute the consumption $c_t$ and one-period-ahead aggregate cash stock $CA_{t-1}$ for each period. As can be seen from Figure 9, the consumption responsiveness decreases in the aggregate cash stock for the negative aggregate shock. From a similar intuition with the opposite direction, the consumption responsiveness increases in the aggregate cash stock for the positive aggregate shock.

| | Dep. Var.: $\|log(c_t)\|$ $(p.p.)$ | |
|---|---|---|
| | Neg. (1) | Pos. (2) |
| $Cash_{t-1}(s.d.)$ | -0.166 (0.001) | 0.07 (0.001) |
| Constant | Yes | Yes |
| Observations | 83 | 84 |
| $R^2$ | 0.996 | 0.994 |

Table 3: State-dependent consumption response to negative and positive shocks

Table 3 reports the regression coefficient when the observations in Figure 9 are fitted into the linear regression. The numbers in the bracket are the standard errors. When the lagged aggregate cash stock increases by one standard deviation, the consumption responsiveness to the negative aggregate TFP shock (-2% TFP shock) significantly decreases by 0.17 percentage points. For the positive aggregate TFP shock (+2% TFP shock), the consumption responsiveness decreases by 0.07 percent-

[34]The aggregate shock is defined as a shift from one productivity to the other.

age points when the lagged aggregate cash stock increases by one standard deviation.

Therefore, the aggregate cash holding gives a consumption buffer against a negative aggregate shock by smoothing the dividend stream in the simulated data. Also, the aggregate cash holding helps a positive productivity shock solely pass down to the consumption. And the consumption buffer effect against the negative TFP shock is significantly stronger than the consumption boosting effect for the positive TFP shock in the model. I support this model prediction from the macro-level data. The data is the quarterly frequency and covers from 1951 to 2018. Consumption and the total dividend of the corporate sector are from NIPA; the aggregate cash holding and the total asset holding are obtained from the Flow of Funds.[35]

| | Dependent variables: | | | |
|---|---|---|---|---|
| | $\|log(c_t)\|$ $(p.p.)$ before 1980 | | $\|log(c_t)\|$ $(p.p.)$ after 1980 | |
| | Neg. (1) | Pos. (2) | Neg. (3) | Pos. (4) |
| $Cash_{t-1}(s.d.)$ | -0.108 | 0.036 | -0.226 | 0.164 |
| | (0.09) | (0.072) | (0.085) | (0.09) |
| Constant | Yes | Yes | Yes | Yes |
| Observations | 63 | 49 | 77 | 79 |
| $R^2$ | 0.023 | 0.005 | 0.086 | 0.041 |

Table 4: Sensitivity of consumption to aggregate TFP shock contingent on corporate cash holdings

Table 4 reports the data counterpart of the results in Table 3, separately for the periods before 1980 (the first two columns) and after 1980 (the last two columns).[36] For this analysis, the consumption variations derived from the TFP variation are controlled by consumption residualization over the polynomials of TFP up to the fourth order.[37] As can be seen from Figure 3, since around 1980, the corporate cash stock has rapidly increased. This rising corporate cash holding has brought a change

[35] All time series are detrended using the HP filter with a frequency parameter at 1600.

[36] The result is not sensitive to the choice of the cutoff year.

[37] I use the utilization-adjusted TFP from Fernald (2014).

in the relationship between corporate cash stock and consumption. In the pre-1980 periods, consumption responsiveness was not dependent on the one-period-ahead corporate cash stock. In contrast, in the post-1980 periods, consumption responsiveness becomes significantly dependent on the corporate cash stock. When the lagged aggregate cash stock increases by one standard deviation, the consumption responsiveness to the negative aggregate shock significantly decreases by 0.23 percentage points. For the positive aggregate shock, the consumption responsiveness increases by 0.16 percentage points when the lagged aggregate cash stock increases by one standard deviation.[38] In Appendix A, I show that this effect in the data is driven by the dividend channel. These results tightly support the model prediction.

# 6 Concluding remarks

This paper develops a novel algorithm to solve general equilibrium models with highly nonlinear aggregate fluctuations. The method utilizes the ergodic theorem's prediction that if the simulated path is long enough, all the possible equilibrium outcomes are realized on the path. Therefore, the rationally expected future value function can be perfectly recovered at each period on the path by identifying counterfactual state-contingent outcomes on the path. The algorithm runs until the expected path converges to the realized path, so the solution is highly accurate. Using this methodology, I study the role of the corporate cash stock on the consumption response to the TFP shock in a heterogeneous-firm business cycle model. When the economy is with large corporate cash stocks, consumption responds less sensitively to a negative TFP shock and more strongly to a positive TFP shock than it would otherwise do. This model prediction is well supported by the macro-level data.

When a model considers the aggregate fluctuations of the equilibrium allocations separated from the (log-)linearly specified exogenous shock processes, the business

[38] These two estimates are statistically not different from the model-side estimates in Table 3.

cycle in the model necessarily features nonlinearity. In such models, depending on the endogenous states, the economy can respond differently to the same shock as in this paper's baseline model. Then, possibly a recession (or a boom) is not merely a product of an enormous exogenous negative shock but rather a product of the interaction between the endogenous economic pre-condition and the moderate exogenous shock. The repeated transition method can provide accurate and efficient solutions for the research that identifies such pre-conditions of an economy through the nonlinear business cycle models.

# References


Ahn, S., G. Kaplan, B. Moll, T. Winberry, and C. Wolf (2018): "When Inequality Matters for Macro and Macro Matters for Inequality," *NBER Macroeconomics Annual*, 32, 1–75, _eprint: https://doi.org/10.1086/696046.

Aiyagari, S. R. (1993): "Uninsured idiosyncratic risk and aggregate saving," Working Papers 502, Federal Reserve Bank of Minneapolis.

——— (1994): "Uninsured Idiosyncratic Risk and Aggregate Saving," *The Quarterly Journal of Economics*, 109, 659–684, publisher: Oxford University Press.

Auclert, A., B. Bardóczy, M. Rognlie, and L. Straub (2019): "Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models," Tech. Rep. w26123, National Bureau of Economic Research, Cambridge, MA.

Bliss, B. A., Y. Cheng, and D. J. Denis (2015): "Corporate payout, cash retention, and the supply of credit: Evidence from the 2008–2009 credit crisis," *Journal of Financial Economics*, 115, 521–540.

Boppart, T., P. Krusell, and K. Mitman (2018): "Exploiting MIT shocks in heterogeneous-agent economies: the impulse response as a numerical derivative," *Journal of Economic Dynamics and Control*, 89, 68–92.

Carroll, C. D. (1997): "Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis," *The Quarterly Journal of Economics*, 112, 56.

Childers, D. (2018): "Solution of Rational Expectations Models with Function Valued States," *Working Paper*, 126.

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

Fernald, J. (2014): "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity," *Federal Reserve Bank of San Francisco, Working Paper Series*, 01–28.

Fernandez-Villaverde, J., S. Hurtado, and G. Nuno (2022): "Financial Frictions and the Wealth Distribution," *Working Paper*, 69.

Jermann, U. and V. Quadrini (2012): "Macroeconomic Effects of Financial Shocks," *American Economic Review*, 102, 238–271.

Kahou, M. E., J. Fernandez-Villaverde, J. Perla, and A. Sood (2021): "Exploiting Symmetry in High-Dimensional Dynamic Programming," *Working Paper*, 51.

Kaplan, G. and G. L. Violante (2014): "A Model of the Consumption Response to Fiscal Stimulus Payments," *Econometrica*, 82, 1199–1239.

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

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

——— (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106, 867–896.

Leary, M. T. and R. Michaely (2011): "Determinants of Dividend Smoothing: Empirical Evidence," *Review of Financial Studies*, 24, 3197–3249.

Lee, H. (2022): "Striking While the Iron Is Cold: Fragility after a Surge of Lumpy Investments," *Working Paper*, 68.

Reiter, M. (2009): "Solving heterogeneous-agent models by projection and perturbation," *Journal of Economic Dynamics and Control*, 33, 649–665.

Riddick, L. A. and T. M. Whited (2009): "The Corporate Propensity to Save," *The Journal of Finance*, 64, 1729–1766.

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

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