Financial Intermediation and Capital Reallocation

Hengjie Ai, Kai Li, and Fang Yang*

July 23, 2019

Abstract

To understand the link between financial intermediation activities and the real economy, we build a general equilibrium model in which agency frictions in the financial sector affect the efficiency of capital reallocation across firms and generate aggregate economic fluctuations. We develop a recursive policy iteration approach to fully characterize the nonlinear equilibrium dynamics and the off-steady-state crisis behavior. In our model, adverse shocks to agency frictions exacerbate capital misallocation and manifest themselves as variations in total factor productivity at the aggregate level. Our model endogenously generates countercyclical volatility in the aggregate time series and countercyclical dispersion in the marginal product of capital and asset returns in the cross section.

Keywords: Financial Intermediation, Capital Misallocation, Volatility, Crisis, Limited enforcement

---

*Hengjie Ai (hengjie@umn.edu) is at the Carlson School of Management of University of Minnesota; Kai Li (kaili@ust.hk) is associated with Hong Kong University of Science and Technology; and Fang Yang (fyang@lsu.edu) is at Louisiana State University. We thank Junyan Shen for his participation in the early stage of the paper. We thank Ron Kaniel (co-editor), one anonymous referee, Ravi Bansal, Harjoat Bhamra, Andrea Eiśfeldt, Murray Frank, Peter Gingeleskie, Loukas Karabarbounis, Pablo Kurlat, Alan Moreira, Stavros Panageas, Adriano Rampini, Lukas Schmid, Amir Yaron, and Hongda Zhong, as well as seminar participants at various institutions for their helpful comments. The usual disclaimer applies.
1 Introduction

Motivated by the experience of the Great Recession, interest in understanding the linkage between financial intermediation activities and macroeconomic outcomes has been revived. The purpose of this paper is to study the efficiency of capital reallocation as a mechanism for such linkage. We present a general equilibrium model to link intermediation activities in the financial sector to capital reallocation across non-financial firms in the real sector. We show that shocks originated from the financial sector affect the borrowing constraint of financial intermediaries and can account for a significant fraction of macroeconomic fluctuations.

Two main features distinguish our approach from those of the previous literature. The first is the emphasis on capital reallocation across firms with heterogeneous productivity. The second is the recursive policy function iteration approach, which allows us to obtain global solutions for a general equilibrium model with occasionally binding incentive compatibility constraints.

We focus on a heterogeneous firm setup for two reasons. In the aggregate, the U.S. corporate sector is rarely constrained: it typically has more cash flow than what is needed to finance investment. As is shown in Chari (2012), a typical feature of models with agency frictions is that firms do not pay dividends when financially constrained. However, the net dividend payment of the U.S. corporate sector as a whole is almost always positive, and significantly so most of the time. To understand why some firms are constrained in downturns while others are not, we use a model with heterogeneous firms.

From a quantitative perspective, our model with heterogeneous firms generates capital reallocation, thus allowing financial frictions to play a significant role in generating large economic fluctuations. In models with a representative firm, financial frictions affect the efficiency of intertemporal investment. Previous researchers (for example, Kocherlakota (2000)) have argued that this mechanism alone is unlikely to cause large economic fluctuations because investment is only a small fraction of the total capital stock of the economy.\footnote{In standard real business cycle (RBC) models, annual investment is about 10\% of capital stock, and capital contributes to roughly one-third of the total output. In line with this calculation, the maximum effect of investment on output is about 3.3\%.} In contrast, recent studies on capital misallocation, for example, Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), have found that large efficiency gains, in the order of 30\% − 50\%, can be achieved by improving capital misallocation.

We develop a recursive policy function iteration approach to fully account for the dynamics of the occasionally binding constraints in our model. A prominent feature of major financial crisis is elevated volatility at the aggregate level and sudden increases in the cross-sectional dispersions in prices and quantities. The majority of models with financial frictions have been
solved using local approximation methods, which typically cannot capture the time variation of volatility implied by the model. The recursive policy function iteration method allows us to characterize the variation in the tightness of the incentive compatibility constraints across time and across firms, which is the key feature of our model.

To formalize the link between financial intermediation and capital reallocation, we develop a model of financial intermediation in which firms are subject to idiosyncratic productivity shocks and credit transactions must be intermediated. Because of the heterogeneity in productivity, reallocating capital across firms improves efficiency in production, but requires high-productivity firms to borrow from the rest of the economy. In addition, because of the limited enforcement of lending contracts, the accumulation of intermediaries’ debt or declines in their net worth increase their incentive to default and limit their borrowing capacity. These features of our model have two implications. In the time series, adverse shocks to intermediaries’ net worth weaken their borrowing capacity and slow down the formation of new capital. In the cross section, intermediaries who finance for high-productivity firms are more likely to be affected because they need to borrow more from the rest of the economy and have a higher incentive to default. The latter mechanism amplifies negative primitive shocks by lowering the efficiency of the reallocation of the existing capital stock.

We consider two versions of our model in the calibration: one with total factor productivity (TFP) shocks and the other with financial shocks. We calibrate the volatility of the primitive shocks to match the volatility of output in the U.S. data and evaluate the quantitative importance of financial frictions in both specifications. In our model with TFP shocks, the amplification effect of agency frictions accounts for about 10% of the total volatility of output and is fairly temporary. The magnitude of amplification is modest because of the well-known difficulty for real business cycle (RBC) models to generate large volatilities in asset prices: because productivity shocks could not generate significant variations in asset prices and intermediary net worth, they induce only a limited amount of amplification from financial frictions.

Motivated by the lack of volatility in asset prices in the model with TFP shocks and the finding in the asset pricing literature that a large fraction of asset price variations can be attributed to discount rate shocks, we model financial shocks as exogenous variations in bank managers’ discount rate. This model generates two features distinct from the one with TFP shocks: persistence and asymmetry. A temporary shock to banks’ net worth lowers their borrowing capacity and reduces the efficiency of capital reallocation in the subsequent period. Elevated capital misallocation depresses output and triggers another round of drop in banks’ net worth. This effect propagates over time and has a long-lasting impact on future economic performance. In addition, negative shocks tighten banks’ financing constraints and make the
economy more vulnerable to future shocks, whereas positive shocks relax these constraints and have a smaller impact on capital misallocation. In the extreme case, a sequence of negative shocks depletes the banking sector’s net worth, lowers the borrowing capacity of all banks to suboptimal levels, and sends the economy into a financial crisis marked by heightened macroeconomic volatility, large and persistent drops in output and asset prices, and sharp increases in interest rate spreads.

In our benchmark calibration, the standard deviation of the banker’s discount rate is about 2.3% at the annual level. This is much smaller than the variation in discount rates typically found in the asset pricing literature (see, for example, Campbell and Shiller (1988), and more recently, Lettau and Ludvigson (2014)). Nevertheless, the model matches well the macroeconomic moments in the United States and produces a volatility of aggregate output of 3.0% from the capital reallocation channel. More importantly, it endogenously generates a countercyclical volatility in the time series of aggregate output and consumption, a countercyclical dispersion in the cross section of firm output and stock returns, and a countercyclical efficiency of capital reallocation and capital utilization as in the data.

**Related literature**  Our paper belongs to the literature on that builds macroeconomic models with a financial intermediary sector, for example, Ger, He and Krishnamurthy (2014), and Brunnermeier and Sannikov (2014). More recently, Rampini and Viswanathan (2014) develop a model of financial intermediation where both the net worth of financial intermediaries and that of firms affect economic recoveries. Elenev et al. (2017) study a quantitative model of financial intermediation to evaluate the effects of macro-prudential policy. Gomes et al. (2015) study the asset pricing implications of financial frictions in a model with costly state verification. Our paper builds on the framework of Ger, but we emphasize the capital reallocation channel and solve our model using a global method to account for the occasionally binding constraints.

Our paper builds on the literature that emphasizes the importance of the cyclical properties of capital reallocation and capital misallocation. Eisfeldt and Rampini (2006) provide empirical evidence that the amount of capital reallocation is procyclical and the benefit of capital reallocation is countercyclical. They also present a model in which the cost of capital reallocation is correlated with TFP shocks to rationalize these facts. Eisfeldt and Rampini (2008), Kurlat (2013), Fuchs, Green, and Papanikolaou (2013), and Li and Whited (2014)

---

2There is a vast literature on macro models with credit market frictions, but we do not attempt to summarize it here. A partial list includes Bernanke and Gertler (1989), Carlstrom and Fuerst (1997), Kiyotaki and Moore (1997), Kiyotaki and Moore (2005), Bernanke et al. (1999), Krishnamurthy (2003), Kiyotaki and Moore (2008), Mendoza (2010), Gertler and Karadi (2011), Jermann and Quadrini (2012), He and Krishnamurthy (2012), He and Krishnamurthy (2013), Li (2013), and Bianchi and Bigio (2014). Quadrini (2011) and Brunnermeier et al. (2012) provide comprehensive reviews of this literature.
analyze models of capital reallocation with adverse selection. Shourideh and Zetlin-Jones (2017) develop a model with financial frictions and heterogeneous firms to study the impact of financial shocks. Kehrig (2015) documents the cyclical nature of productivity distribution over the business cycle.

Our paper is also related to the literature that links total factor productivity at the aggregate level to capital misallocation at the firm level (for example, Midrigan and Xu (2014), Moll (2014), and Buera and Moll (2015)). Buera et al. (2011) develop a quantitative model to explain the relationship between aggregate TFP and financial constraints. Gopinath et al. (2015) develop a model with financial frictions to account for the decline in total factor productivity in southern Europe. None of the above papers focus on the effect of financial frictions and capital reallocation on aggregate volatility dynamics as we do.

The idea that shocks may directly originate from the financial sector and affect economic activities is related to the setup of Jermann and Quadrini (2012). Different from that of Jermann and Quadrini (2012), our paper focuses on financial intermediation and capital reallocation and their connections with the macroeconomy.

Our paper also relates to the literature in economics and finance that emphasizes the importance of countercyclical volatility in understanding the macroeconomy and asset markets. Many authors have documented a strong countercyclical relationship between real activity and uncertainty as is proxied by stock market volatility and/or dispersion in firm-level earnings and productivity (see, for example, Bloom (2009), Bloom et al. (2012), Bachmann et al. (2013), and Jurado et al. (2015), among others). A large literature in asset pricing emphasizes the importance of countercyclical volatility in understanding stock market returns (see, for example, Bansal and Yaron (2004), Bansal et al. (2012), and Campbell et al. (2013)). Dou (2017) studies the asset pricing implications of stochastic idiosyncratic volatility in a general equilibrium setup. Our model generates countercyclical volatility as an endogenous equilibrium outcome even though the primitive shocks are homoscedastic.

The rest of the paper is organized as follows. We provide a summary of some stylized facts that motivate the development of our model in Section 2. We describe the model setup in Section 3. In Section 4, we illustrate the link between financial intermediation and capital reallocation of our model in a simple two-period setting. We discuss the construction of the Markov equilibrium of our full model and the recursive policy function iteration approach in Section 5. We calibrate our model and evaluate its quantitative implications on macroeconomic quantities and asset prices in Section 6. Section 7 concludes.
2 Stylized Facts

To motivate our theory that focuses on the efficiency of capital reallocation as a mechanism for the linkage between financial intermediation activities and macroeconomic outcomes, we present several related stylized facts.

Efficiency of capital reallocation requires the marginal product of capital to equalize across firms. Dispersion of the marginal product of capital can therefore be interpreted as capital misallocation. Based on this insight, Hsieh and Klenow (2009) develop a measure of capital misallocation. We follow a similar procedure and compute a measure of the efficiency of capital reallocation using the variance of the cross-sectional distribution of the log marginal product of capital within narrowly defined industries (classified by the four-digit standard industry classification code). Below we document several stylized facts about capital reallocation using this measure. We provide the details of the construction of this measure in Appendix A.

1. Measured total factor productivity (TFP) is highly correlated with a measure of the efficiency of capital reallocation and the rate of capital utilization.\(^3\)

Figure 1: capital reallocation and capital utilization

\(^3\)Capital underutilization can be interpreted as a special form of misallocation.
In the top panel of Figure 1, we plot the time series of log TFP (dashed line) and the measured efficiency of capital reallocation (solid line) in the United States, where all series are HP filtered. The shaded areas indicate NBER-classified recessions. The measured efficiency of capital reallocation closely tracks the time series of log TFP, with a correlation of 0.33, indicating that the efficiency of capital reallocation may account for a significant fraction of variations in measured TFP. In the bottom panel, we plot the time series of log TFP (dashed line) and capital utilization rates (solid line) published by Federal Reserve Bank of St. Louis. Clearly, capital utilization also exhibits pronounced procyclicality, with a correlation of 0.62 with log TFP. Economic downturns are typically associated with sharp declines in capital utilization.

2. The total volume of bank loans is procyclical and positively correlated with the efficiency of capital reallocation and negatively correlated with measures of volatility.

Figure 2: total volume of bank loans

This fact motivates our theory of financial intermediation and its connection with capital reallocation. We calculate the total volume of bank loans of the non-financial corporate sector in the United States from the Flow of Funds Table and plot the time series of changes in the total volume of bank loans (dashed line) and the measured efficiency of capital reallocation (solid line) in the top panel of Figure 2. We also plot
the changes in the total volume of bank loans and the stock market volatility (solid line) in the bottom panel of Figure 2, where stock market volatility is calculated by aggregating the realized variance of monthly returns. The shaded areas in both panels indicate NBER-classified recessions. The total volume of bank loans is strongly procyclical. In addition, the total volume of bank loans is positively correlated with the efficiency of capital reallocation, with a correlation of 0.43, and negatively correlated with stock market volatility, with a correlation of $-0.33$.4

3. The amount of capital reallocation is procyclical, and the cross-sectional dispersion of the marginal product of capital is countercyclical. Those are well-known stylized facts; therefore, we do not provide a detailed discussion here but refer readers to the relevant literature (see, for example, Eisfeldt and Rampini (2006)).

The fourth, fifth, and sixth facts are about the cyclical properties of the volatility of macroeconomic quantities and asset returns and are well known in the macroeconomics literature and the asset pricing literature (see, for example, Bloom (2009), Bansal et al. (2012), and Campbell et al. (2001)).

4. The volatility of macroeconomic quantities, including consumption, investment, and aggregate output, is countercyclical.

5. The volatility of the aggregate stock market return is also countercyclical. The equity premium and interest rate spreads are countercyclical.

6. The volatility of idiosyncratic returns on the stock market is countercyclical.

In the following sections, we set up and analyze a general equilibrium model with financial intermediation and capital reallocation to provide a theoretical and quantitative framework to interpret the above facts.

3 Model Setup

In this section, we describe a general equilibrium model with heterogeneous firms and with agency frictions in the financial intermediation sector.

3.1 Non-financial Firms

There are three types of non-financial firms in our model: final goods producers, intermediate goods producers, and a capital storage firm.

4This negative correlation is robust to other measures of aggregate volatility as well.
Final goods producers Final goods are produced by a representative firm using a continuum of intermediate inputs indexed by $j \in [0, 1]$. Because a final goods producer does not make intertemporal decisions in our model, we suppress time $t$ in this subsection to save notation. We normalize the price of final goods to one and write the profit maximization problem of a final goods producer as

$$\max_{\{Y_j\}} \left\{ Y - \int_{[0,1]} p_j Y_j \, dj \right\}_{Y=\left[\int_{[0,1]} Y_j \, dj\right]^{\frac{\eta-1}{\eta}}} \tag{1},$$

where $p_j$ and $Y_j$ are the price and quantity of input $j$, respectively, and $Y$ stands for the total output of final goods. The parameter $\eta$ is the elasticity of substitution across input varieties.

The Dixit-Stiglitz production technology allows us to quantify the benefit of capital reallocation by calibrating the elasticity of substitution across varieties to the capital misallocation literature (see, for example, Hsieh and Klenow (2009)). In addition, the constant returns to scale of the Dixit-Stiglitz production technology is important for the tractability of the model at the aggregate level.

Intermediate goods producers There is a continuum of competitive intermediate goods producers, $j \in [0, 1]$, each producing a different variety on a separate island.\footnote{We use the terminology “island” to emphasize that capital cannot move freely among producers of different input varieties. The details of capital market frictions are introduced in Section 3.3.} We use $j$ as the index for both the intermediate input and the island on which it is produced. The production of variety $j$ at time $t$ requires two factors, capital $K_{j,t}$ and labor $L_{j,t}$. We allow the rental price of capital to be island specific but assume a frictionless labor market across the whole economy. Variable $MPK_{j,t}$ denotes the rental price of capital on island $j$, and $MPL_t$ is the economy-wide wage rate. The profit maximization problem for the producer on island $j$ at time $t$ is given by

$$\max \left\{ p_{j,t} Y_{j,t} - MPK_{j,t} \cdot K_{j,t} - MPL_t \cdot L_{j,t} \right\},$$

subject to:

$$Y_{j,t} = \bar{A_t} z_{j,t}^{\frac{1}{1-\alpha}} K_{j,t}^{\alpha} L_{j,t}^{1-\alpha}, \tag{2}$$

where $\bar{A}_t$ is the aggregate productivity common across all islands, $\alpha$ governs the share of capital income in production, and $z_{j,t}$ is an island-$j$-specific idiosyncratic productivity shock that follows:

$$\ln z_{j,t+1} = \ln z_{j,t} + \varepsilon_{j,t+1}. \tag{3}$$
In the above equation, \( \varepsilon_{j,t+1} \) is i.i.d. across firms and over time with \( E[\varepsilon_{j,t+1}] = 1 \). The assumption that \( E[\varepsilon_{j,t+1}] = 1 \) is a normalization, which implies that the average of idiosyncratic productivity of all firms is constant over time, so that aggregate productivity growth comes entirely from \( \bar{A}_t \) and not from the growth of idiosyncratic shocks \( \varepsilon_{j,t} \). We further assume that in equilibrium \( \bar{A}_t = A_t K_t^{1-\alpha} \), where \( A_t \) is a Markov process and \( K_t \) is the aggregate capital stock of the economy. This specification follows Frankel (1962) and Romer (1986). It implies that production is linear in \( K_t \) at the aggregate level and simplifies our analysis.\(^6\)

**Capital Storage Technology** We specify a capital storage technology to allow for variable capital utilization, as in Greenwood et al. (1988). Current period capital, \( K_t \), can be used for two purposes: production of intermediate goods and storage. In the competitive market, the capital storage firm acquires \( K_{S,t} \) at the market price \( Q_t \) and saves the capital for next period through a storage technology \( g(\cdot) \). The capital storage firm chooses \( K_{S,t} \) to maximize profit:

\[
D_{K,t} = \max_{K_{S,t}} \left\{ g\left( \frac{K_{S,t}}{K_t} \right) K_t - Q_t K_{S,t} \right\}.
\]

We use \( u_t \) to denote the capital utilization rate, that is, \( u_t = 1 - \frac{K_{S,t}}{K_t} \). We assume that utilized capital depreciates linearly at rate \( \delta \). Therefore, the law of motion of capital is

\[
K_{t+1} = [g(1 - u_t) + (1 - \delta) u_t] K_t + I_t.
\]

We further assume that \( g(0) = 0 \), \( g' > 1 - \delta \), and \( g'' < 0 \) (which implies a concave storage technology). These assumptions together imply that depreciation is increasing in utilization, and unutilized capital depreciates at a lower rate than utilized capital. Our model is therefore a special case of the variable capital utilization model of Greenwood et al. (1988).\(^7\)

\(^6\)Quantitatively, it is important to allow the parameter \( \alpha \) to be smaller than one for the model to match the capital share in the data. However, this introduces decreasing returns to scale and increases the number of state variables in the construction of the recursive equilibrium. The endogenous growth specification effectively makes the aggregate production to be linear in capital, which reduces the dimension of state variables while allowing us to match the basic business cycle facts on labor and capital share in the data.

\(^7\)The Greenwood et al. (1988) model specifies the law of motion of capital as \( K_{t+1} = (1 - \Delta (u_t)) K_t + I_t \), where \( \Delta (u_t) \) is capital depreciation as a function of the utilization rate, \( u_t \), where \( \Delta' > 0 \) and \( \Delta'' > 0 \). Using the Greenwood et al. (1988) notation, the effective capital depreciation rate in our model is \( \Delta (u) = 1 - u - g(1 - u) + u\delta \). The interpretation is that \( 1 - u - g(1 - u) \) is the depreciation of unutilized capital and \( u\delta \) is the depreciation of utilized capital. Clearly, \( \Delta' (u) > 0 \) and \( \Delta'' (u) > 0 \) under our assumptions. In addition, the assumption \( g(0) = 0 \), \( g' > 1 - \delta \), and \( g'' < 0 \) together imply that \( 1 - u - g(1 - u) < (1 - u) \delta \), that is, utilized capital depreciates faster. Note that the value of one unit of utilized capital depends on the utilization rate in the general model of Greenwood et al. (1988), while our formulations imply it always equals one.
We note that the first-order condition for the capital-goods-producing firm implies

\[ Q_t = g'(1 - u_t). \] (6)

Because the storage technology \( g(\cdot) \) is concave, the market price of capital, \( Q_t \), increases with the capital utilization rate, \( u_t \).

Variable capital utilization is important for the quantitative implications of our model. First, as we show in Figure 1, capital utilization in the data exhibits pronounced procyclicality. Second, allowing capital utilization is important for our model to capture the co-movement of investment across firms. A basic feature of the data is that the investment of all firms tends to drop in recessions. Because capital in our model is one-period predetermined, as in standard neoclassical models, without variable capital utilization, our model implies that in bad times when high-productivity firms are constrained, capital must flow into low-productivity firms, which is counterfactual and inconsistent with the data. In addition, our choice of the functional form of the storage technology is for tractability, as it implies that the price for utilized capital is always one, which as we will show in Section 5, greatly simplifies the construction of the Markov equilibrium of our dynamic model.

### 3.2 Household

There is a continuum of identical households who rank consumption plans according to log preferences:

\[ E \left[ \sum_{t=0}^{\infty} \beta^t \ln C_t \right], \] (7)

where \( \beta \) is the time discount factor.

The household receives labor income and dividend payments from the capital-goods-producing firm and trades the risk-free bonds and equities of banks. In every period \( t \), a typical household purchases \( B_{j,t} \) amount of bonds with a risk-free rate of return \( R_{f,t+1} \) and \( \tau_{j,t} \) shares of equity from bank \( j \). In addition, it receives dividend payments from capital goods producers, \( D_{K,t} \), and labor income \( MPL_t \).\(^8\) The household’s budget constraint is thus

\[ C_t + \int \tau_{j,t} V_{j,t} dj + \int B_{j,t} dj = R_{f,t} \int B_{j,t-1} dj + \int \tau_{j,t-1} (V_{j,t} + D_{B,j,t}) dj + D_{K,t} + MPL_t. \]

In the above setup, under the assumption of log utility, the risk-free return and the return on bank equity must satisfy the standard intertemporal Euler equations, \( E [M_{t+1} R_{f,t+1}] = 1 \)

\(^8\)Since labor supply is inelastic and normalized to one, the total amount of labor income is \( MPL_t \).
and $E[M_{t+1}R_{j,t+1}] = 1$, where $R_{j,t+1} = \frac{V_{j,t+1} + DB_{j,t+1}}{V_{j,t}}$ is the return on bank $j$'s equity and $M_{t+1} = \beta \left( \frac{C_{t+1}}{C_t} \right)^{-1}$ is the stochastic discount factor implied by household consumption.

We will use the terms ”financial intermediaries” and ”banks” interchangeably and assume that the equity and bonds of a bank are not owned by the same household. This assumption allows us to model the agency frictions in banks’ lending contracts but will not affects other aspect of the model because each bank and each household are infinitesimal.

### 3.3 Financial Intermediaries

There is a continuum of islands with i.i.d. productivity shocks, and there is one competitive financial intermediary on each island.\(^9\) Consider a bank on island $j$ who enters into period $t$ with initial net worth $N_{j,t}$. It chooses the total amount of borrowing from the household, $B_{j,t}$, and the total amount of capital stock for the next period, $K_{j,t+1}$. Because there is no capital adjustment cost, the price of capital is one and banks’ budget constraint is

$$K_{j,t+1} = N_{j,t} + B_{j,t}. \quad (8)$$

In our model, the total amount of capital for the next period, $K_{j,t+1}$, is determined at the end of period $t$, before the realization of shocks in the next period. That is, we assume a one-period time to plan, as in standard real business cycle models. However, different from the standard representative firm setup, capital can be reallocated across firms after shocks are realized. The market for capital reallocation opens after the realization of the aggregate productivity shock $\bar{A}_{t+1}$ and the island-$j$-specific idiosyncratic productivity shock, $\varepsilon_{j,t+1}$. Capital is then reallocated among islands with different histories of idiosyncratic productivity shocks, and the capital storage firm.

Figure 3 illustrates the timing of events period $t$ and in period $t+1$. At the end of period $t$, bank $j$ uses total asset holdings, including total net worth $N_{j,t}$ and consumer loans $B_{j,t}$ to purchase $K_{j,t+1}$ amount of capital for period $t + 1$ production before the realization of the productivity shocks in $t + 1$.

Period $t + 1$ is divided into three subperiods. In the first subperiod, the aggregate productivity shock, $\bar{A}_{t+1}$, and the idiosyncratic productivity shock, $\varepsilon_{j,t+1}$, are realized and the capital reallocation market opens. The capital is reallocated across islands. Let $Q_{t+1}$ denote the price of capital on the capital reallocation market at period $t + 1$ and $RA_{j,t+1}$ denote the total amount of capital purchased or sold on the reallocation market by intermediary $j$. The total amount of capital on the island is thus $K_{j,t+1} + RA_{j,t+1}$. In this subperiod, capital

\(^9\)Because financial intermediaries on each island face a competitive capital market, one could also interpret our model as having a continuum of identical financial intermediaries on each island.
on the reallocation market can be purchased only by issuing a within-period interbank loan. This is because the purchase happens before production and the receipt of payment from local firms.

Production happens in the second subperiod, and firms pay back the cost of renting capital to local banks at the end of the second subperiod. Let $Q_{j,t+1}$ denote the price of capital on island $j$. Bank $j$ receives payment, $Q_{j,t+1} [K_{j,t+1} + RA_{j,t+1}]$, from local firms. The price of capital $Q_{j,t+1}$ is island specific because financial constraints may prevent the marginal product of capital from being equalized to the price of capital on the reallocation market when they are binding.

In the third subperiod, after banks receive payment from local firms, but before they pay back loans to creditors, banks have an opportunity to default and set up a new bank to operate on some other island. We assume that following default, bankers can take away all the capital on the island, but they can only sell a fraction $\theta$ of it on the local market.\(^{10}\) Therefore, in the event of default, the total receipt of bankers on island $j$ is $\theta Q_{j,t+1} [K_{j,t+1} + RA_{j,t+1}]$.

If the bank pays back its interbank loans and household deposits, the total net worth is

$$N_{j,t+1} = Q_{j,t+1} [K_{j,t+1} + RA_{j,t+1}] - Q_{t+1} RA_{j,t+1} - R_{f,t+1} B_{j,t}. \quad (9)$$

The possibility of default implies that the contracting between borrowing and lending banks

\(^{10}\)The assumption that banks only sell capital “locally” is made for tractability so that we can collect the common term $Q_{j,t+1} [K_{j,t+1} + RA_{j,t+1}]$ when we combine (9) and (10) into (11). It is relevant only for the quantitative implications of the model but not for the main economic mechanism.

\(^{11}\)Because $RA_{j,t+1}$ depends on aggregate and idiosyncratic shocks, an equivalent way to formulate our model is to allow banks to borrow a loan at time $t$ that provides state-contingent payoffs of $Q_{t+1} RA_{j,t+1} - R_{f,t+1} B_{j,t}$ in period $t + 1$. This alternative formulation does not need to distinguish household loans and interbank loans but yields identical equilibrium allocations. Our formulation makes it transparent to interpret $B_{j,t}$ as the net borrowing between the banking sector and the households and to interpret $RA_{j,t+1}$ as the amount of borrowing between banks.
must respect the following limited enforcement constraint:

\[ N_{j,t+1} \geq \theta Q_{j,t+1} [K_{j,t+1} + RA_{j,t+1}], \forall t \text{ and } \forall j. \] (10)

The incentive compatibility constraint implies that in anticipating the possibility of default, lending banks ensure that the borrowing banks do not have an incentive to default on loans in all possible states of the world. Combining Equations (9) and (10), we can write the limited enforcement constraint as

\[ (1 - \theta) Q_{j,t+1} [K_{j,t+1} + RA_{j,t+1}] - Q_{t+1} RA_{j,t+1} \geq R_{f,t+1} B_{j,t}. \] (11)

Clearly, limited enforcement of contracts imposes a constraint on the amount of borrowing, as in the classical model of Kiyotaki and Moore (1997). Different from Kiyotaki and Moore (1997), the left-hand side of (11) depends on the realization of idiosyncratic shocks, and the constraint has to hold for all possible realizations of idiosyncratic and aggregate shocks. Because more productive firms typically need to borrow from less productive ones, in our model, the financial constraint limits not only the total amount of borrowing of the corporate sector from the household sector, but also, and more importantly, the amount of borrowing within the corporate sector across firms. At the end of the third subperiod, bankers clear their interbank transactions and consumers receive dividend payments from banks and firms, a risk-free return from bank deposits, and make their consumption and saving decisions.

We model banks’ exit and entry as follows. At the end of the third subperiod, after production, each island experiences a liquidation shock with probability \( 1 - \Lambda_{t+1} \), where \( \Lambda_{t+1} \in (0, 1) \). Upon receiving such a shock, the bankers on the island are forced to liquidate, and assets are paid back to the household as a dividend. All firms on the island will lose their production technology and start over with the initial condition \( z_{j,t+2} = 1 \). This specification of entry and exit ensures a stationary cross-sectional distribution of idiosyncratic productivity shocks. In addition, as in the dynamic agency literature (for example, DeMarzo and Sannikov (2006) and DeMarzo and Fishman (2007)), the liquidation probability can be interpreted as the ratio of bankers’ discount rates relative to that of the households. This wedge in discount rates is a parsimonious way to capture the idea that the managers of banks have a shorter investment horizon than the representative household and allows agency frictions to persist in the long run. We also assume that bankers’ net worth moves freely across islands at the end of every period, which implies that the marginal value of bankers’ net worth will be equalized.

Let \( V_i^j (N_{j,t}) \) denote the value function of banker \( i \) at time \( t \). Given that the banker’s net worth is paid back to the household with probability \( 1 - \Lambda_{t+1} \) next period, the banker’s
objective is to maximize
\[ V^j_t (N_{j,t}) = \max_{\{K_{j,t+1}, B_{j,t}, RA_{j,t+1}\}} \mathbb{E}_t \left[ M_{t+1} \left\{ (1 - \Lambda_{t+1}) N_{j,t+1} + \Lambda_{t+1} V^j_{t+1} (N_{j,t+1}) \right\} \right] \] (12)

by choosing total capital stock for the next period \( K_{j,t+1} \), total borrowing from households \( B_{j,t} \), and a state-contingent plan for capital reallocation, \( RA_{j,t+1} \), subject to the budget constraint (8), the law of motion of \( N_{j,t+1} \) given by (9), and the limited enforcement constraint (10). Note that, as in standard neoclassical models, given equilibrium prices \( M_{t+1}, Q_{t+1}, \) and \( Q_{j,t+1} \), the banker’s objective function and constraints are linear in net worth. Therefore, the value function must be linear as well, and we write \( V^j_t (N_{j,t}) = \mu_t N_{j,t} \). Thanks to the assumption that net worth moves freely across islands at the end of every period, the marginal value of net worth must be equalized across islands, and therefore \( \mu_t \) does not depend on the firm index.

### 3.4 Market Clearing

We list the resource constraints and market clearing conditions as follows. First, the amount of capital used for production on all islands must sum to \( u_t K_t \), which is the total amount of utilized capital in the economy:
\[ \int (K_{j,t} + RA_{j,t}) \, dj = u_t K_t. \] (13)

Second, we need a market clearing condition for capital on each island. This market clearing condition implies that in equilibrium, the price of capital on each island must satisfy
\[ Q_{j,t} = MPK_{j,t} + 1 - \delta, \]
where \( MPK_{j,t} \) is the marginal product of capital on island \( j \) in period \( t + 1 \).

Third, the total net worth of the banking sector equals the sum of banks’ net worth across all islands:
\[ N_t = \int N_{j,t} \, dj. \] (14)

Fourth, labor market clearing requires \( \int L_{j,t} \, dj = 1 \) because we assume inelastic labor supply and normalize total labor endowment to one.

Finally, market clearing for final goods requires that total consumption and investment sum to total output: \( C_t + I_t = Y_t \), where \( Y_t \) is total output of final goods defined in (1).
4 The Two-Period Case

In this section, we analyze a two-period version of our fully dynamic model without aggregate uncertainty to demonstrate the key economic mechanism, that is, how equilibrium prices and quantities depend on the initial level of net worth in the banking sector. In the fully dynamic model discussed in Section 5, net worth will be determined by the history of aggregate shocks, but the basic mechanism remains the same.

We make several additional simplifying assumptions. First, there is no capital misallocation or capital under utilization in period 0, so the resource constraint can be written as 
\[ C_0 + I_0 = A_0 K_0. \]

Second, all firms enter in period 0 with 
\[ z_0 = 1, \]
and there is no exit. Third, there are only two possible realizations of idiosyncratic productivity shocks, \( \varepsilon_H \) and \( \varepsilon_L \) with 
\[ \text{Prob}(\varepsilon = \varepsilon_H) = 1 - \text{Prob}(\varepsilon = \varepsilon_L) = \pi. \]

Aggregation of the Product Market

Because the determination of output is a static problem, we suppress time subscript. As we show in Appendix B, under the assumptions of a CES technology and free flow of labor and the normalization condition \( \int L_j dj = 1 \), total output can be computed as

\[
Y = \bar{A} \left\{ \int z_j^{(1-\xi)} (K_j + RA_j)^{\xi} dj \right\} \frac{\xi}{\bar{A}(uK)^{\alpha}}, \tag{15}
\]

where \( \xi = \frac{\alpha \eta}{\alpha + \eta - \alpha} \).

Under our assumption that there are only two possible realizations of idiosyncratic productivity shocks, we can further simplify the aggregate output function. Let \( RA_H \) and \( RA_L \) denote the capital reallocation on high- and low-productivity islands, respectively. Let 
\[ \phi = \frac{K + RA_H}{K + RA_L} \]
be the ratio of firm size across islands. The resource constraint (13) can be written as 
\[ uK = \pi (K + RA_H) + (1 - \pi)(K + RA_L), \]
which, together with the definition of \( \phi \), implies
\[
\frac{K + RA_H}{uK} = \frac{\phi}{\pi \phi + (1 - \pi)}; \quad \frac{K + RA_L}{uK} = \frac{1}{\pi \phi + (1 - \pi)}. \tag{16}
\]
Together with Equation (15), we can write aggregate output as a function of \((\phi, u)\), which is specified in the following proposition.

**Proposition 1 (Aggregation of the Product Market)**

*The total output of the economy at period 1 is* \( Y = Au^{\alpha} f(\phi) K \), *where the function* \( f(\phi) \) *is* \( \frac{\phi}{\pi \phi + (1 - \pi)} \).
\( f : [1, \hat{\phi}] \rightarrow [0, 1] \) is defined as
\[
f(\phi) = \left\{ \pi e^{(1-\xi)e_H} \left( \frac{\phi}{\pi\phi + (1-\pi)} \right)^\xi + (1-\pi) e^{(1-\xi)e_L} \left( \frac{1}{\pi\phi + (1-\pi)} \right)^\xi \right\}^{\frac{\phi}{\xi}} \tag{17}\]
and \( \hat{\phi} = e^{e_H-e_L} \).

The marginal product of capital on low- and high-productivity islands, \(MPK_L\) and \(MPK_H\), can be written as functions of \((A, \phi, u)\):
\[
MPK_L(A, \phi, u) = \alpha Af(\phi) u^{\alpha-1} \frac{\pi\phi + 1 - \pi}{\pi e^{(1-\xi)(e_H-e_L)}\phi^{\xi} + 1 - \pi}, \tag{18}
\]
\[
MPK_H(A, \phi, u) = MPK_L(A, \phi, u) e^{(1-\xi)(e_H-e_L)} \phi^{1-\xi}. \tag{19}
\]

**Proof.** See Appendix B. ■

It is not hard to show that, given our normalization condition \( E[e^\xi_{j,t+1}] = 1 \), the efficient level of \( \phi \) that implies equalization of the marginal production of capital across all islands is \( \hat{\phi} = e^{e_H-e_L} \) and \( f(\hat{\phi}) = 1 \). The function \( f(\phi) \) is a measure of the efficiency of capital reallocation, since \( \phi \) can be interpreted as a measure of capital reallocation. It is straightforward to show that \( f \) is increasing in \( \phi \).

**Determination of \( u \) and \( \phi \)** Because all firms are identical in period 0, we only need to specify the optimal decisions of a representative bank. The bank’s optimization problem can be written as follows:
\[
V(N_0) = \max_{B_0, K, RA_H, RA_L} \frac{1}{R_f} \left\{ \pi [Q_H(K + RA_H) - QRA_H - R_fB_0] + (1-\pi) [Q_L(K + RA_L) - QRA_L - RB_0] \right\}, \tag{20}
\]
\[
subject \ to : K = N_0 + B_0, \tag{21}
\]
\[
Q_H(K + RA_H) - QRA_H - R_fB_0 \geq \theta Q_H(K + RA_H), \tag{22}
\]
\[
Q_L(K + RA_L) - QRA_L - R_fB_0 \geq \theta Q_L(K + RA_L). \tag{23}
\]

As in the infinite-horizon model, given equilibrium prices, the bank’s problem is linear in \( N_0 \). Of course, equilibrium prices will in general be nonlinear functions of state variables and need to be determined jointly with equilibrium quantities.\(^{13}\)

Note that our model is equivalent to one in which banks are allowed to borrow though

\(^{13}\)Different from the neoclassical models, however, because welfare theorems fail in the presence of frictions, equilibrium prices cannot be constructed from the planner’s problem and have to be solving jointly with equilibrium quantities using equilibrium conditions.
state-contingent loans. A state contingent loan needs to specify the payoff in period 1 as a function of the realizations of productivity shocks $\epsilon_h$ and $\epsilon_l$. In our setup, this payoff can be replicated by $\{QR_{AH} - R_f B_0, QR_{AL} - R_f B_0\}$. Our formulation is convenient as it allows us to map $B_0$ into the net borrowing between households and banks and $\{RA_H, RA_L\}$ into reallocations between banks in the data.

In equilibrium, the price of capital on each island is determined by the marginal product of capital. Because marginal products of capital can be written as functions of $(A, \phi, u)$, we can represent prices of capital as functions of $(A, \phi, u)$:

\begin{align*}
Q_H (A, \phi, u) &= MPK_H (A, \phi, u) + 1 - \delta, \quad (24) \\
Q_L (A, \phi, u) &= MPK_L (A, \phi, u) + 1 - \delta. \quad (25)
\end{align*}

In addition, the first-order condition (6) implies the price of capital on the reallocation market can be written as a function of $u$: $Q (u) = g' (1 - u)$. The above conditions determine the prices of capital as functions of aggregate productivity, capital reallocations, and capital utilization.

Dividing both sides of constraints (22) and (23) by $K$ and using equation (16), we obtain

\begin{align*}
(1 - \theta) Q_H (A, \phi, u) + [(1 - \theta) Q_H (A, \phi, u) - Q(u)] \left[ \frac{\phi u}{\pi \phi + (1 - \pi)} - 1 \right] &\geq s, \quad (26) \\
(1 - \theta) Q_L (A, \phi, u) + [(1 - \theta) Q_L (A, \phi, u) - Q(u)] \left[ \frac{u}{\pi \phi + (1 - \pi)} - 1 \right] &\geq s, \quad (27)
\end{align*}

where we denote $s = \frac{R_f B_0}{K}$ as the ratio of bank liability to capital.

Let $\zeta_H$ and $\zeta_L$ denote the Lagrangian multipliers on the limited enforcement constraints, (22) and (23). The first-order conditions with respect to $RA_j$ can be used to derive a relationship between Lagrangian multipliers and the prices of capital on the high- and low-productivity islands. We use this relationship to define

\begin{align*}
\zeta_H (A, \phi, u) &= \frac{\pi [Q_H (A, \phi, u) - Q(u)]}{Q(u) - (1 - \theta) Q_H (A, \phi, u)} \geq 0, \quad > 0 \implies (26) \text{ holds with } \llap{“} = \\
\zeta_L (A, \phi, u) &= \frac{(1 - \pi) [Q_L (A, \phi, u) - Q(u)]}{Q(u) - (1 - \theta) Q_L (A, \phi, u)} \geq 0 > 0 \implies (27) \text{ holds with } \llap{“} = 
\end{align*}

which imply that

\begin{align*}
Q_H - Q &\geq 0, \quad > 0 \implies (26) \text{ holds with } \llap{“} = \\
Q_L - Q &\geq 0, \quad > 0 \implies (27) \text{ holds with } \llap{“} =.
\end{align*}
If both of the limited enforcement constraints (26) and (27) hold with equality, then they jointly determine $\phi$ and $u$ as functions of $(A,s)$. If neither (26) nor (27) is binding, then $\zeta_H(A,\phi,u) = \zeta_L(A,\phi,u) = 0$, implying $Q_H(A,\phi,u) = Q_L(A,\phi,u) = Q(u)$. Again, $\phi$ and $u$ can be determined as functions of $(A,s)$. In general, Equations (26), (27), (28), (29), and the complementary slackness condition determine $\phi$ and $u$ as functions of $(A,s)$, which we will denote as $\phi(A,s)$ and $u(A,s)$. The following proposition builds on this observation and characterizes the nature of the binding constraints.

**Proposition 2** (Characterization of Binding Constraints) There exist $\hat{s}(A)$ and $\bar{s}(A)$ such that

1. If $s \leq \hat{s}(A)$, then none of the limited commitment constraints bind, and $\phi(A,s)$ and $u(A,s)$ are determined by the equality versions of (30) and (31).

2. If $\hat{s}(A) < s \leq \bar{s}(A)$, then the limited commitment constraint for banks on high productivity islands binds, and $\phi(A,s)$ and $u(s)$ are determined by the equality versions of (26) and (31).

3. If $s > \bar{s}(A)$, then the limited commitment constraint for all banks binds, and $\phi(A,s)$ and $u(A,s)$ are determined by the equality versions of (26) and (27).

4. Both $\hat{s}(A)$ and $\bar{s}(A)$ decrease with $A$.

**Proof.** See Appendix C. ■

The above proposition implies that given both productivity $A$ and $s$, capital reallocation $\phi$ and capital utilization $u$, and therefore the prices, $Q_H$, $Q_L$, and $Q$, are completely determined. When $s$ is below $\hat{s}(A)$, the debt level is low enough and the limited enforcement constraints never bind. As the debt level increases, when $\hat{s}(A) < s \leq \bar{s}(A)$, the limited enforcement constraint binds only if the island receives a high-productivity shock. Efficient capital reallocation requires that banks on high-productivity islands borrow from those on low-productivity islands. Therefore, the limited enforcement constraint is more likely to bind for banks on high-productivity islands. In the region in which $s > \bar{s}(A)$, the banking sector accumulated too much debt and the limited enforcement constraints bind for both realizations of productivity shocks.

Proposition 2 has several important implications. First, in the cross section, the limited enforcement constraint is more likely to bind for intermediaries on high-productivity islands. This is the mechanism for misallocation in our model: when banks are constrained, more productive projects cannot be financed and measured TFP drops. Second, in the time series, the limited enforcement constraint is more likely to bind when banks’ liability (net worth)
is high (low), when aggregate productivity drops. This is the amplification mechanism in our model. Adverse shocks to TFP and banks’ net worth are amplified because they tighten the limited enforcement constraints and exacerbate capital misallocation. Third, the pair of state variables \((A, s)\) fully determine capital reallocation in our model, a fact we will exploit in the construction of the Markov equilibrium in our fully dynamic model.

The previous literature has considered models in which borrowing constraints are not determined by capital stock but by future output. This alternative specification will affect only the quantitative implications of our model but not the qualitative ones. In fact, Proposition 2 applies to a special case of the two-period model, which can be interpreted as one in which borrowing capacity is determined by next-period output. Because of the constant returns to scale technology, \(MPK_{j,t} \cdot K_{j,t} = \alpha Y_{j,t}\), where \(\alpha\) is the capital share parameter. If we set \(\delta = 1\), then \(Q_H = MPK_H\) and \(Q_L = MPK_L\). Under the assumption that \(\alpha = 1\), the limited enforcement constraints (22) and (23) can indeed be written as \(R_f B_0 + QRA_H \leq (1 - \theta) y_H\) and \(R_f B_0 + QRA_L \leq (1 - \theta) y_L\), respectively.

**Model Solution**

To solve the model, we first note that no arbitrage implies that banks must be indifferent between investing in one unit of capital to sell on the reallocation market and saving at the risk-free interest rate. Therefore, \(R_f = Q(u)\). Second, intertemporal optimality on the consumer side implies that \(R_f = \frac{C_1}{\beta C_0}\). Using the aggregate resource constraints to replace \(C_0\) and \(C_1\) and write \(\phi\) and \(u\) as functions of \((A, s)\), we obtain

\[
\frac{Au (A, s) \alpha f (\phi (A, s)) [(1 - \delta) K_0 + I_0]}{A_0 K_0} = Q (u (A, s)).
\] (32)

Third, using the definition of \(s\), we can replace \(B_0\) in banks’ budget constraint (21) by \(B_0 = \frac{s}{R_f} [(1 - \delta) K_0 + I_0]\) and obtain

\[
N_0 = [(1 - \delta) K_0 + I_0] \left(1 - \frac{s}{R_f}\right).
\] (33)

Note that given the initial condition \((A_0, A, K_0, N_0)\), Equations (32) and (33) jointly determine investment \(I_0\) and the bank liability-to-capital ratio, \(s\).

**Implications of the two-period model**

We set parameters in our two-period model to be the same as parameters chosen for the full model, which we discuss in detail in Section 6.1. In particular, we set \(A_0\) and \(A\) to be the mean value, 0.11, and \(K_0 = 1\). We use the policy functions of the two-period model to illustrate the mechanism through which banks’ net worth affects macro quantities, prices, and capital reallocation. We plot equilibrium
quantities and prices as functions of banks’ net worth $N_0$.

Figure 4: macro moments and bank net worth

In Figure 4, we plot aggregate output (top panel) and banking sector leverage (second panel), that is, the banks’ asset-to-net-worth ratio, as functions of banks’ net worth, $N_0$. In the figure, $\hat{N}_0 = N(\hat{s})$ is the level of banks’ net worth above which the limited enforcement constraints do not bind for any banks and capital misallocation does not occur. Further increases in banks’ net worth above $\hat{N}_0$ do not affect output because productivity is constant and capital reallocation stays at its first-best level. As $N_0$ decreases toward $\bar{N}_0 = N(\bar{s})$, only the limited commitment constraint for high-productivity islands, Equation (22), binds. Here, as $N_0$ declines, capital reallocation deteriorates and therefore output decreases. As $N_0$ drops below $\bar{N}_0$, the limited enforcement constraint for both islands binds, and output drops sharply.

Similar to that of Ger, our model features countercyclical leverage. As banks’ net worth declines, banks need to borrow more from the household to finance investment. Consequently, leverage increases. This is the key mechanism in our model that generates countercyclical volatility in aggregate quantities. As we show in Section 5, adverse shocks lower banks’ net worth and raise banking sector leverage. In that event, exogenous shocks are more likely to be amplified, and the volatility in all aggregate quantities increases.

As is shown in Eisfeldt and Rampini (2006), the amount of capital reallocation is pro-
cyclical, and the benefit to capital reallocation is countercyclical. Our model is consistent with this fact. In Figure 4, we plot the prices of capital (third panel), the total amount of capital reallocation (fourth panel), and the capacity utilization rate (fifth panel) as functions of banks’ net worth. As is shown in the third panel, in the region $N_0 \geq \hat{N}_0$, capital reallocation is fully efficient, and marginal products of capital equalize across all islands. As banks’ net worth drops below $\hat{N}_0$, the marginal product of capital on high- and low-productivity islands diverges, but the allocation of capital between low-productivity islands and the storage technology is fully efficient. As banks’ net worth decreases further such that $N_0 < \hat{N}_0$, low-productivity islands become constrained as well, and capital under-utilization occurs: it is invested in the inefficient storage technology despite the fact that reallocation to any productive island could raise its marginal product. As bank net worth declines, the marginal product of capital diverge and the benefit of capital reallocation increases.

In summary, the divergence of the marginal product of capital is echoed by reductions in the total amount of capital reallocation (fourth panel) and decreases in the capital utilization rate (fifth panel). In addition, drops in capital reallocation and capital utilization are much more pronounced in the crisis region in which the limited enforcement constraints bind for all banks.

5 The Dynamic Model

In the dynamic model, banking sector net worth, the key state variable that affects capital reallocation, will be an endogenous equilibrium outcome that generates quantitative predictions on the dynamics of macroeconomic quantities and asset prices. Unlike frictionless economies, where allocations can be solved from the planner’s problem and prices can be constructed from allocations, because of the presence of agency frictions, prices and quantities in our model have to be determined jointly as the fixed point implied by equilibrium conditions. In this section, we develop a policy function iteration approach to solve our model recursively.

We first extend Proposition 1 to represent aggregate production as a function of two variables $(\phi, u)$ that summarize capital misallocation and capital utilization. We then use Proposition 2 to represent $(\phi, u)$ as functions of $s$ and construct a recursive equilibrium using $s$ as a state variable.

**Aggregation of the product market** In our model, thanks to the assumption that banks’ net worth moves freely across islands at the end of every period, the marginal value of banks’ net worth will be equalized. As we show in Appendix B, this feature of the model allows us to construct an equilibrium in which $\frac{N_{j,t}}{z_{j,t}}$ and $\frac{K_{j,t+1}}{z_{j,t}}$ are equalized across islands. It
is convenient to think of all firms that realized the same \( \varepsilon_{j,t+1} \) shock as being in the same sector. Under this interpretation, the marginal product of capital, which is proportional to \( \frac{K_{j,t+1} + RA_{j,t+1}}{z_{j,t+1}} \), must be equalized within each sector.

We further make a simplifying assumption, as in the two-period model, that there are only two possible realizations of idiosyncratic productivity shocks, \( \varepsilon_H \) and \( \varepsilon_L \) with \( \text{Prob}(\varepsilon = \varepsilon_H) = 1 - \text{Prob}(\varepsilon = \varepsilon_L) = \pi \) with a normalization condition \( \pi e^{\varepsilon_H} + (1 - \pi) e^{\varepsilon_L} = 1 \). The total output in Equation (15) can then be written as

\[
Y_{t+1} = \bar{A}_{t+1} \left\{ \int_{\varepsilon_{j,t+1} = \varepsilon_H} z_{j,t+1}^{(1-\xi)} (K_{j,t+1} + RA_{j,t+1})^\xi \, dj + \int_{\varepsilon_{j,t+1} = \varepsilon_L} z_{j,t+1}^{(1-\xi)} (K_{j,t+1} + RA_{j,t+1})^\xi \, dj \right\} ^{\frac{\alpha}{\xi}}.
\]  

(34)

Because \( \frac{K_{j,t+1} + RA_{j,t+1}}{z_{j,t+1}} \) are equalized for all firms that realized the same \( \varepsilon_{j,t+1} \) shock, we can first aggregate within the sectors and then across the sectors to obtain

\[
Y_{t+1} = f(\phi_{t+1}) \bar{A}_{t+1} (u_{t+1} K_{t+1})^{\alpha},
\]

(35)

where

\[
\phi_{t+1} = \frac{K_{t+1} + E[R_{A_j,t+1} | \varepsilon_{j,t+1} = \varepsilon_H]}{K_{t+1} + E[R_{A_j,t+1} | \varepsilon_{j,t+1} = \varepsilon_L]}
\]

(36)

is defined as the ratio of the average size of firms across the two sectors and \( f(\phi) \) is given in Proposition 1.\(^{14}\)

In our model, we obtain an aggregation result that aggregate output and price are independent of the cross-sectional distribution of firms. In Section 6.3 and Appendix E, we show that the distribution of firm size implied by our model follows a power law, and the model implied power law exponent is consistent with the data.

**Determination of capital reallocation**  As in the two-period model, Proposition 2 can be applied to represent \( (\phi, u) \) as functions of the state variable \( s \). Using the fact that firms in the same sector have the same \( \frac{RA_{j,t+1}}{K_{j,t+1}} \) ratio, we can divide both sides of the resource constraint (13) by \( K_{t+1} \) and aggregate within sectors to obtain

\[
\pi \left( 1 + \frac{E[R_{A_j,t+1} | \varepsilon_{j,t+1} = \varepsilon_H]}{K_{t+1}} \right) + (1 - \pi) \left( 1 + \frac{E[R_{A_j,t+1} | \varepsilon_{j,t+1} = \varepsilon_L]}{K_{t+1}} \right) = u_{t+1}.
\]

(37)

\(^{14}\)The conditional expressions \( E[R_{A_j,t+1} | \varepsilon_{j,t+1} = \varepsilon] \) in Equation (36) should be interpreted as the average \( RA_{j,t+1} \) across all firms within the sector that realized the same value of the \( \varepsilon \) shock.
Together with the definition of $\phi_{t+1}$, the above equation can be used to represent $E[RA_{j,t+1}|\varepsilon_{j,t+1}=\varepsilon]/K_{t+1}$ as a function of $\phi_{t+1}$ and $u_{t+1}$:

$$E[RA_{j,t+1}|\varepsilon_{j,t+1}=\varepsilon_H]/K_{t+1} = \frac{u_{t+1}\phi_{t+1}}{\pi \phi_{t+1} + 1 - \pi} - 1;$$

$$E[RA_{j,t+1}|\varepsilon_{j,t+1}=\varepsilon_L]/K_{t+1} = \frac{u_{t+1}}{\pi \phi_{t+1} + 1 - \pi} - 1.$$

The limited commitment constraint (11) can be simplified to

$$(1 - \theta) Q_{j,t+1} K_{j,t+1} + [(1 - \theta) Q_{j,t+1} - Q_{t+1}] RA_{j,t+1} \geq R_{f,t+1} B_{j,t}.$$  \hspace{1cm} (38)

Dividing both sides of constraint (38) by $K_{j,t+1}$, it is clear that the above constraint is identical for all firms within the same sector. We can therefore consolidate the above inequality and represent it as a constraint for $(\phi_{t+1}, u_{t+1})$:

$$
\begin{align*}
(1 - \theta) Q_{H,t+1} + [(1 - \theta) Q_{H,t+1} - Q_{t+1}] & \left( \frac{u_{t+1} \phi_{t+1}}{\pi \phi_{t+1} + 1 - \pi} - 1 \right) \geq s_{t+1}, \\
(1 - \theta) Q_{L,t+1} + [(1 - \theta) Q_{L,t+1} - Q_{t+1}] & \left( \frac{u_{t+1}}{\pi \phi_{t+1} + 1 - \pi} - 1 \right) \geq s_{t+1},
\end{align*}
$$

where $s_{t+1} = \frac{R_{f,t+1} B_{j,t}}{K_{j,t+1}}$, which is common across all islands, and $Q_{H,t+1}$ and $Q_{L,t+1}$ are prices of capital that are common across islands within the same sector.

As in the two-period model, competition on the capital markets within the same island implies (24) and (25). That is, the sector-specific capital prices, $Q_{H,t+1}$ and $Q_{L,t+1}$, depend on the marginal products of capital, which are only functions of $A_{t+1}$, $\phi_{t+1}$, and $u_{t+1}$. Inequalities (39) and (40), together with the complementary slackness conditions, determine $\phi_{t+1} = \phi(A_{t+1}, s_{t+1})$ and $u_{t+1} = u(A_{t+1}, s_{t+1})$, where the functions $\phi(A_{t+1}, s_{t+1})$ and $u(A_{t+1}, s_{t+1})$ are the same as described in Proposition 2.

**Recursive formulation** The recursive formulation of the model requires an appropriate choice of state variables that allows us to construct the global solution to the Markov equilibrium in two steps. The first step solves for policies as functions of the state variables, and the second step obtains the law of motion of state variables from equilibrium conditions. Although it is common to use net worth as the state variable in the literature, the most convenient choice of state variable in our setup is total banking sector liability normalized by the capital stock, $s_t$. In principle, any one-to-one transformation can be used as state variables in the construction of the Markov equilibrium. Our choice is conceptually and computationally simpler.

To describe the recursive equilibrium in its full generality, we allow both productivity $A$
and the firm survival rate \( \Lambda \) to be subject to exogenous shocks. We denote \( x_t = (A_t, \Lambda_t) \) as the exogenous state variables and \( \omega_t = (x_t, s_t) \) as the vector of state variables for the Markov equilibrium. We choose \( s_t \) as the state variable for two reasons. First, as shown in Proposition 2, once \( s_t \) is known, capital reallocation \( \phi_t \) and capital utilization \( u_t \) can be determined for each possible realization of \( A_t \) without referencing any other state variables or equilibrium dynamics. Second, \( s_{t+1} = \frac{R_{f,t+1}B_{t+1}}{K_{t+1}} \) depends only on period-\( t \) information but not on the realizations of \( x_{t+1} \). This implies that the construction of the law of motion of this endogenous state variable involves solving a single equation for \( s_{t+1} \) instead of a functional equation for \( s_{t+1} \) as a function of \( x_{t+1} \). In what follows, we provide a construction of the Markov equilibrium of our dynamic model. Because our construction uses a recursive procedure, it naturally leads to an iterative procedure to numerically solve the model, which we describe in detail in Appendix D.

Thanks to the assumption \( \bar{A}_t = A_t K_t^{1-\alpha} \), equilibrium quantities are homogeneous of degree one in \( K \) and equilibrium prices do not depend on \( K \). It is therefore convenient to work with normalized quantities. We define

\[
c = \frac{C}{K}, \quad i = \frac{I}{K}, \quad n = \frac{N}{K}, \quad b_f = \frac{B_f}{K}.
\]

Because of the linear property of the banker’s optimization problem (12) with respect to net worth, banks’ value function is linear in net worth: \( V_j^t(N_{j,t}) = \mu(\omega_t) N_{j,t} \) for some function \( \mu(\omega_t) \). In what follows, we describe a procedure that determines quantities \( c(\omega) \) and \( i(\omega) \) and prices \( \mu(\omega) \) and \( R_f(\omega) \) as functions of \( \omega \). We will call \( \{c(\omega), i(\omega), \mu(\omega), R_f(\omega)\}_\omega \) the equilibrium functionals. Given a set of equilibrium functionals, other equilibrium prices and quantities can be constructed accordingly.

To construct a Markov equilibrium, we first use the law of motion of individual banks’ net worth, Equation (9), to derive an expression for the law of motion for aggregate net worth. Let \( N \) be the aggregate banking sector net worth in the current period. In the next period, a \( \pi \) fraction of all banks realizes \( \varepsilon_{J,t+1} = \varepsilon_H \), and their total net worth becomes

\[
\pi \left[ Q_{H,t+1} [K_{t+1} + RA_{H,t+1}] - Q_{t+1}R_{A,H,t+1} - R_{f,t+1}B_t \right] .
\]

A \( 1 - \pi \) fraction realizes a \( \varepsilon_L \) shock and their total net worth becomes

\[
(1 - \pi) \left[ Q_{L,t+1} [K_{t+1} + RA_{L,t+1}] - Q_{t+1}R_{A,L,t+1} - R_{f,t+1}B_t \right] .
\]

Taking into account the fact that a fraction \( 1 - \Lambda_{t+1} \) of all banks are forced to liquidate, we
then have
\[ N_{t+1} = \Lambda_{t+1} \left\{ [\pi Q_{H,t+1} (K_{t+1} + RA_{H,t+1}) + (1 - \pi) Q_{L,t+1} (K_{t+1} + RA_{L,t+1})] \right. \]
\[ \left. - Q_{t+1} [\pi RA_{H,t+1} + (1 - \pi) RA_{L,t+1}] - R_{f,t+1} B_t \right\}. \] (42)

To construct a recursive equilibrium, we think of \( t \) as the current period and \( t + 1 \) as the next period, and use the above relationship to express equilibrium quantities not as functions of time but rather as functions of relevant state variables. We use lowercase letters to denote current period normalized quantities and lowercase letters with a prime symbol (′) to denote next period normalized quantities. Using capital prices derived in Equations (24) and (25) and the fact that the capital income share is \( \alpha \), we obtain
\[ \pi Q_{H,t+1} (K_{t+1} + RA_{H,t+1}) + (1 - \pi) Q_{L,t+1} (K_{t+1} + RA_{L,t+1}) = \alpha Y_{t+1} + (1 - \delta) K_{t+1}. \]

Using the resource constraint (13) and dividing both sides of Equation (42) by \( K_{t+1} \), we obtain an expression for next period net worth normalized by the capital stock, \( n' \):
\[ n' = \Lambda' \left\{ \alpha A' (u')^\eta f (\phi') + (1 - u') MPK (u') + (1 - \delta) - s \right\}, \] (43)
where \( MPK (u) \) is defined by
\[ MPK (u) = Q (u) - (1 - \delta), \]
with \( Q (u) \) given in Equation (6).

Next, we derive the law of motion for the endogenous state variable, \( s_{t+1} \). Because the ratio \( \frac{R_{f,t+1} B_t}{K_{t+1}} \) is equalized across all firms, \( s' = \frac{[R_f(\omega)b_f(\omega)]}{K_{t+1}} \). Using Equation (5), we derive
\[ s' = \frac{R_f b_f}{g (1 - u) + (1 - \delta) u + i}. \] (44)

To obtain an expression for \( b_f \), we can divide both sides of the bank budget constraint (8) by \( K \):
\[ g (1 - u) + (1 - \delta) u + i = n + b_f \] (45)
and use Equation (45) to replace \( b_f \) in (44) to obtain
\[ s' = \frac{g (1 - u) + (1 - \delta) u + i - n}{g (1 - u) + (1 - \delta) u + i} R_f. \] (46)

Using Equation (43) to replace \( n \) in (46), we obtain
\[ s' = R_f (\omega) \left\{ 1 - \frac{\Lambda \left\{ \alpha \left(1 - \frac{1}{\eta} \right) A (\omega) f (\phi (\omega)) + (1 - u (\omega)) MPK (u (\omega)) + (1 - \delta) - s \right\}}{g (1 - u (\omega)) + (1 - \delta) u (\omega) + i (\omega)} \right\}. \] (47)

Equation (47) is the law of motion for the state variable \( s \). This choice for the state
variable allows us to efficiently obtain the global solution to a GE model with agency frictions and occasionally binding constraints through two steps. The first step determines $\phi$ and $u$ as functions of $(A, s)$ using Proposition 2. This step is a static problem and does not involve any dynamic programming. The second step is to use Equation (47) to solve the law of motion for $s$. Two of our modeling choices imply that the two steps can be computed separately and greatly improve the numerical efficiency of the solution. First, using $s$ as the state variable implies that the law of motion (47) is a single equation to solve, whereas the law of motion of $n$, (43), is a functional equation for $n'$ as a function of $A'$. Second, thanks to our choice of the functional form of the capital utilization technology, Tobin’s Q does not depend on capital utilization and is always equal to one. Equation (47) can therefore be solved once $\phi(A, s)$ and $u(A, s)$ are determined. In general, the term $\delta$ in (47) should depend on capital utilization $u$. In this more general case, the two steps cannot be computed separately, and all equilibrium prices and quantities must be jointly determined.

Together, the two steps determine the law of motion of state variables and the equilibrium functionals as the fixed point of a set of equilibrium conditions. In Appendix C, we show that the equilibrium functionals must satisfy the following conditions. First, the first-order condition for households’ optimal investment decision leads to the usual intertemporal Euler equation,

$$E[M(\omega, \omega')] R_f(\omega) = 1,$$

where $M(\omega, \omega')$ denotes the stochastic discount factor of households:

$$M(\omega, \omega') = \frac{\beta [A_w(\omega) f(\phi(\omega)) - i(\omega)]}{c(\omega') [g(1 - u(\omega)) + (1 - \delta) u(\omega) + i(\omega)].}$$

Second, banks’ optimal choice for intertemporal investment implies

$$\mu(\omega) = E \left[ \tilde{M}(\omega, \omega') \{1 + (\zeta_H(A', \phi(\omega'), u(\omega')) + \zeta_L(A', \phi(\omega'), u(\omega'))\} Q(u') \right],$$

where $\tilde{M}(\omega, \omega')$ is defined as

$$\tilde{M}(\omega, \omega') = M(\omega, \omega') \{1 - \Lambda' + \Lambda' \mu(\omega')\},$$

and the terms $\zeta_H(A', \phi(\omega'), u(\omega'))$ and $\zeta_L(A', \phi(\omega'), u(\omega'))$ denote the Lagrangian multipliers on the limited enforcement constraints, as derived in (28) and (29) in the two-period model.
Third, the envelope condition on banks’ optimization problem is

\[ \mu(\omega) = E \left[ \tilde{M}(\omega, \omega') \{ 1 + \zeta_H(A', \phi(\omega'), u(\omega')) + \zeta_L(A', \phi(\omega'), u(\omega')) \} \right] R_f(\omega). \quad (52) \]

Fourth, and finally, we note that the resource constraint requires

\[ c(\omega) + i(\omega) = Au^a(\omega) f(\phi(\omega)). \quad (53) \]

Note that the four unknown equilibrium functionals, \( c(\omega), i(\omega), \mu(\omega), \) and \( R_f(\omega), \) can be determined by the four functional equations (48), (50), (52), and (53). Our construction of the Markov equilibrium is formally summarized by the following proposition:

**Proposition 3** Suppose there exists a set of equilibrium functionals, \( \{c(\omega), i(\omega), \mu(\omega), R_f(\omega)\}_z \), such that with the law of motion of \( s \) given by (47), \( \{c(\omega), i(\omega), \mu(\omega), R_f(\omega)\}_z \) satisfies the functional equations (48), (50), (52), and (53), then \( \{c(\omega), i(\omega), \mu(\omega), R_f(\omega)\}_z \) constitutes a Markov equilibrium.

**Proof.** See Appendix C. \( \blacksquare \)

The recursive method used in our construction of the Markov equilibrium naturally leads to a numerically efficient way to compute the solution of the model. We detail this in Appendix D.

### 6 Quantitative Results

In this section, we evaluate quantitatively the impact of financial frictions under two specifications of our model: a specification with TFP shocks only and a specification with shocks to agency frictions only. In the model with TFP shocks only, productivity shocks are the only source of primitive shocks and are amplified by financial frictions. Consistent with the previous literature (for example, Kocherlakota (2000) and Chen and Song (2013)), we find that financial frictions do amplify TFP shocks but the effect is quantitatively small. Amplification accounts for about 9% of the macroeconomic fluctuations in the model with TFP shocks. In addition, the economy rarely runs to the crisis region in which the limited enforcement constraint binds for all banks because TFP shocks do not generate large enough variations in banks’ net worth.

In the second specification of the model, we introduce financial shocks modeled as stochastic shocks to bankers’ discount rates. From an economics point of view, because the discount rate \( \Lambda \) measures the relative patience between bankers and households and affects the severity of the agency friction of limited commitment, it can be interpreted as shocks to agency
frictions. From a quantitative point of view, the asset pricing literature (for a survey, see Cochrane (2011)) found large shocks to discount rates, and our calibration interprets these shocks as primitive shocks. The literature on macro models with financial frictions typically needs shocks that directly affect the net worth of financial intermediaries to generate a quantitatively significant impact, for example, the capital quality shocks used in Ger and the shocks to cross-sectional dispersion in firm-level productivity growth, as in Elenev et al. (2018). This discount rate shocks in our model play a similar role, although the magnitude of the shocks needed in our model is much smaller than that documented in the above asset pricing literature.

We find that relatively small financial shocks generate large fluctuations in capital misallocation and can account for most of the macroeconomic fluctuations in the U.S. economy. We show that this model endogenously generates countercyclical volatility at the aggregate level and countercyclical dispersion in the cross section. In addition, this version of the model captures several salient features of the recent financial crisis, such as spikes in macroeconomic volatility, sharp drops in capital reallocation and capital utilization, and sudden increases in interest rate spreads. Our purpose here is not to argue that financial shocks are the only source of economic fluctuations, but merely to analyze their impact on the efficiency of capital reallocation and the macroeconomy.

6.1 Calibration

We calibrate our model at the quarterly frequency. To facilitate the comparison, we choose the same parameters, except for the volatility of exogenous shocks, for both specifications of our model. This approach guarantees that both the model with productivity shocks and that with financial shocks have the same deterministic steady state. We then calibrate the volatility of the exogenous shocks in both models to match the volatility of aggregate output in the data and evaluate the model’s implications on the dynamics of macroeconomic quantities and asset prices.

The common parameters of both calibrations can be divided into three groups. The first group is the standard preference and technology parameters that can be chosen from the previous literature, which we list in the top panel of Table 1. We set the quarterly discount rate to $\beta = 0.99$, and the quarterly depreciation rate to $\delta = 2\%$. We set the capital share to $\alpha = 0.33$, as in the standard RBC models, and the elasticity of substitution across varieties to $\eta = 4$, which is consistent with the value used in Hsieh and Klenow (2009).

The second group comprises technology parameters specific to our model, but can be chosen directly to match a relevant moment in the data. We set the capital storage technology
in the constant elasticity form:
\[ g(x) = a_0 + \frac{b_0}{\nu}x^\nu. \]

We set \(a_0\) and \(b_0\) to match the time-series average of the capital utilization rate of 80% in the data and the average capital depreciation rate of 10%, a standard value in RBC models. We set the elasticity parameter to \(\nu = 0.99\) so that the volatility of capital utilization rate in our model with financial shocks matches that in the data: 5% per year.\(^{15}\)

This procedure leaves five remaining parameters in the third group: \(E[A]\), \(E[\Lambda]\), \(\theta\), \(\varepsilon_H - \varepsilon_L\), and \(\pi\).\(^{16}\) We choose these five parameters so that the deterministic steady state of our model best matches seven moments in the data: the mean aggregate economic growth rate, the capital-reallocation-to-investment ratio, the banking sector leverage ratio, the investment-to-output ratio, the external-financing-to-total-investment ratio, the power law slope of the size distribution, and the ratio of output in high- and low-productivity firms.\(^{17}\) The last is a normalization condition that high- and low-productivity islands each account for 50% of total output.

Although all the parameters are jointly determined by the over identifying moment conditions, certain moments are relatively more responsible for pinning down certain parameters. For example, the mean productivity level, \(E[A] = 0.11\), is largely determined by a mean growth rate of the U.S. economy of 2% per year. We set \(e^{\varepsilon_H-\varepsilon_L} = 12.84\) and \(\pi = 0.19\) so that the capital-reallocation-to-investment ratio is 22%, as is reported in Eisfeldt and Rampini (2006), and high- and low-productivity islands each account for 50% of total output. The banker’s discount rate, \(\Lambda\), affects the banking sector’s total net worth. The fraction of assets that bankers can abscond with, \(\theta\), has an impact on the borrowing capacity of the banking sector for a given level of net worth and therefore the total investment in the economy. We calibrate \(E[\Lambda]\) and \(\theta\) to jointly match a banking sector leverage ratio of 4.5, which is broadly consistent with Ger; and set the investment-to-output ratio to 25%, as in the U.S. postwar data; and set the mean external-financing-to-total-investment ratio to 22%, as in Shourideh and Zetlin-Jones (2017). Finally, the power law coefficient of the firm size distribution is jointly determined by the distribution of \(\varepsilon_j\) and the exit rate \(\Lambda\). Our model generates a power law coefficient of 1.07, as reported in Luttmer (2007). The calibrated parameters and targeted moments are listed in Table 1.

We set both \(A_t\) and \(\Lambda_t\) to be i.i.d. over time and examine the propagation of shocks

\(^{15}\)Elasticity \(\nu\) is the only technology parameter pinned down by a second moment in the data. We choose \(\nu\) so that the volatility of capital utilization matches our preferred model with financial shocks.

\(^{16}\)The normalization condition, \(E[e^{\varepsilon_j + 1}] = 1\), implies that \(\varepsilon_H - \varepsilon_L\) can be treated as one parameter to determine both \(\varepsilon_H\) and \(\varepsilon_L\).

\(^{17}\)Our model generates a power law in the right tail of the distribution of firm size. We provide a detailed derivation of the power law implied by our model in Appendix E.
endogenously generated by the model. In the model with TFP shocks only, we assume $\ln A_t$ follows a normal distribution and calibrate the volatility of $\ln A_t$ to match the volatility of aggregate output of 3% in U.S. postwar data. To make sure that $\Lambda \in (0,1)$, for parsimony, we set $\Lambda_t = \frac{\exp(\lambda_t)}{\exp(\lambda_t) + \exp(-\lambda_t)}$ and assume $\lambda_t$ follows a normal distribution. This specification allows us to choose one parameter in the model with financial shocks, the standard deviation of $\lambda_t$, to match the volatility of aggregate output.

6.2 Impulse responses

To understand the different implications of TFP shocks and discount rate shocks in our model with financial frictions, we use the policy function iteration method introduced in Section 5 to numerically solve the model and plot in Figure 5 the impulse functions for one-standard-deviation shocks to $\ln A$ (left panel) and those for one-standard-deviation shocks to $\lambda$ (right panel). The horizontal axis represents the number of quarters after the initial shock, and the vertical axis is log deviations from the steady state. In both panels, we use dashed lines for positive shocks and solid lines for negative shocks.

Figure 5: impulse responses to productivity and financial shocks

Notice that the effects of TFP shocks on the economy differ sharply from the effects of
discount rate shocks. More specifically, they differ along three dimensions: the magnitude of amplification, the persistency, and the symmetry.

First, in our model, financial frictions do amplify primitive shocks, but the amplification effect is quantitatively small in the model with productivity shocks and much larger in the model with financial shocks. The amplification effect can be measured by the variation in the efficiency of capital reallocation, \( u^\alpha f(\phi) \), (bottom panels) relative to that in output, \( y = Au^\alpha f(\phi) \), (fourth panels). In the model with productivity shocks, following impact, output drops by about 5% (fourth panel on the left), and most of the change comes from the direct effect of productivity, \( \ln A \) (top panel on the left), not from the efficiency of capital reallocation (bottom panel on the left). In contrast, the changes in output in the model with financial shocks (right panels) are completely due to the efficiency of capital reallocation.

Second, shocks to \( \lambda \) have much more persistent effects than shocks to productivity, even though both shocks occur for one period and immediately return to steady state afterward. The key intuition behind this sharp contrast is that these two types of shocks have very different quantitative impacts on banks’ net worth and, in turn, the limited commitment constraints. The effects from productivity shocks are transitory because productivity shocks have two offsetting effects on bank net worth. On the one hand, a negative productivity shock lowers capital income and therefore reduces bank’s net worth. This is the standard amplification channel in models with financial frictions because a negative shock to banks’ net worth lowers their borrowing capacity. On the other hand, negative productivity shocks also reduce aggregate investment and therefore the total financing needs of the corporate sector. The second effect implies that the financial constraints are less likely to bind. In our model, the tightness of the financial constraint is determined by the state variable \( n \), total bank net worth normalized by total capital stock (i.e., total financing demand). As shown in Figure 5, a one-standard-deviation (5%) increase in the productivity shock increases banks’ net worth by about 0.2% on impact (second panel on the left). As a result, the tightness in the limited commitment constraints is almost unaffected. Because productivity returns back to steady state immediately, and because the tightness in the limited commitment constraint hardly reacts to these shocks, their effect on macroeconomic quantities virtually disappears after one period.

Conversely, the model with \( \lambda \) shocks generates a very persistent impact on the economy. Following a one-standard-deviation financial shock, although the initial reaction of output is fairly modest (less than 1%), it propagates over time and lasts for many periods.\(^{18}\) The reason

\(^{18}\)Note that both models are calibrated to match a standard deviation of output growth rate of 3% per year. The model with financial shocks requires much smaller shocks and a much smaller initial impact of these shocks because financial shocks provide a powerful propagation mechanism thanks to the feedback effects between the real economy and the banking sector’s net worth.
is, as we show in Figure 5 in the second panel on the right, a negative shock in the bankers’
discount rate, which raises the current period’s dividend payment, immediately lowers banks’
net worth. As banks’ borrowing capacity reduces, so does the efficiency of capital reallocation
(bottom panel), for two reasons. The first is that net worth is a persistent state variable and
takes time to recover. In addition, because of the lower efficiency of capital reallocation,
the capital rental rate, and therefore bank income, drops, making it more difficult for banks
to rebuild their net worth. These two effects reinforce each other to generate long-lasting
effects. In fact, as we show in the forth panel on the right of Figure 5, with a 1% reduction
in $\lambda$, total output drops by about 1%, and the effect is so persistent that the system is still
far from its steady state even after twenty quarters.

Third, the effects of productivity shocks are largely symmetric, whereas the effects of
financial shock are asymmetric. As shown in the left panels of Figure 5, positive and nega-
tive shocks in productivity result in changes in quantities and prices of similar magnitude.
Qualitatively, as we have seen in the policy functions in the two-period model (Figure 4),
negative shocks to net worth have a larger impact on capital misallocation than do positive
ones, especially in the “crisis” region. Quantitatively, however, productivity shocks induce
very modest changes in banks’ net worth, and the asymmetry and countercyclical volatility
generated by the model is insignificant.

In contrast, the asymmetry in the impulse responses of quantity and prices with respect
to shocks to agency frictions is apparent in the right panels of Figure 5. A positive shock to
$\lambda$ relaxes the limited enforcement constraint, lowers leverage (third panel on the right), and
reduces the effect of future shocks. A negative shock to $\lambda$ tightens the limited enforcement
constraint and makes the system more sensitive to additional disturbances. As a result,
negative shocks are amplified and positive shocks are dampened, leading to endogenous
countercyclical volatility in our model.

6.3 Simulation

To understand the quantitative implications of the two specifications of our model, we sim-
ulate the model for 800 quarters and discard the first 400 quarters. We then aggregate
the quarterly quantities in the remaining part of the simulation into annual quantities and
compute moments for annualized quantities.

Unconditional moments We report the moments of macroeconomic quantities (top
panel), those of interest rates (second panel), and the cyclicality of variables of interest
(bottom panel) in Table 2. Both versions of our model are consistent with the basic features
of the data in terms of the relatively low volatility in consumption growth, the high volatility
in investment growth, and the comovement between consumption growth and investment growth.\footnote{Relative to the data, both versions of our model produce too much volatility in investment and too little volatility in consumption. This counterfactual implication of the model can be corrected by adding an investment adjustment cost.} Both feature a low risk-free interest rate. The model with TFP shocks generates very little variation in interest rates. The model with financial shocks produces a significant volatility in the interest rate, as in the data.

Focusing on the bottom panel of Table 2, we see that the data feature pronounced countercyclical volatility in aggregate output and the stock market return, where output is calculated using quarterly industrial production and the stock market return is computed as the realized standard deviation of the monthly market return. In Table 2, $Corr[\Delta lnY, Vol(\Delta lnY')]$ measures the correlation between output growth and the volatility in growth rates in the following year, where $Vol(\Delta lnY')$ is computed as the realized standard deviation in the growth rates of output in the next year.\footnote{We interpret all production as industrial production in the model.} $Corr[\Delta lnY_t, Vol(R'_M)]$ measures the correlation between output growth and the realized volatility in the stock market return in the next year. As we explained in Section 6.2, positive financial shocks lower banking sector leverage and reduce the effect of future shocks, whereas negative financial shocks raise leverage and enhance the amplification mechanism. This asymmetry translates into significant countercyclical volatility in our model with financial shocks. Conversely, volatility is almost acyclical in the model with TFP shocks because these shocks have little effect on the banking sector’s net worth.

To quantify the impact of capital reallocation, in the last row of Table 2, we compute the fraction of measured TFP shocks attributed to variations in the efficiency of capital reallocation. In our model, the measured log TFP equals $ln \bar{A}_t + ln u^\alpha_t f(\phi_t)$, where the component $ln u^\alpha_t f(\phi_t)$ is the efficiency of capital reallocation. We compute the following variance ratio as a measure of the amplification effect of financial frictions:

$$\bar{\Delta} = \frac{Var \left[ ln \left( u^\alpha_{t+1} f(\phi_{t+1}) \right) - ln \left( u^\alpha_t f(\phi_t) \right) \right]}{Var \left[ ln A_{t+1} - ln \bar{A}_t + ln \left( u^\alpha_{t+1} f(\phi_{t+1}) \right) - ln \left( u^\alpha_t f(\phi_t) \right) \right]}.$$

In the model with TFP shocks only, $\bar{\Delta} = 0.09$; that is, financial frictions account for 9% of the total variation in TFP. In the model with financial shocks only, $\bar{\Delta} = 0.98$; thus, the efficiency of capital reallocation accounts for virtually all of the macroeconomic fluctuations.

**Capital reallocation and financial intermediation** To further investigate the role of capital reallocation and capital utilization, we report the related statistics in Table 3. The model with financial shocks exhibits significantly more time variation in the dispersion of the marginal product of capital (23%) than the model with TFP shocks (1%) and a significantly
higher volatility in capital utilization (5%) than the model with TFP shocks (0.5%). The intuition for the sharp contrast between the two models is that financial shocks modeled as the discount rate shocks lead to significantly larger fluctuations in banks’ net worth and the tightness of limited commitment constraints, and therefore generate larger variations in the efficiency of capital reallocation than productivity shocks.

Our model of financial shocks is also consistent with empirical evidence of the procyclicality of capital reallocation and the countercyclicality of the benefit of capital reallocation documented by Eisfeldt and Rampini (2006), as shown in the bottom panel of Table 3. In the model with financial shocks only, the correlation between the capital-reallocation-to-investment ratio and measured TFP is 0.29, which is close to its empirical counterpart. As in the data, the benefit of capital reallocation, measured by the dispersion in the marginal product of capital, has a negative correlation with measured TFP of −0.33, which is of similar magnitude as in the data. As we have explained in our discussion of the impulse response functions, an adverse shock to banks’ net worth lowers banks’ borrowing capacity and limits capital reallocation across firms. As a result, the dispersion in the marginal product of capital rises. This mechanism of our model is confirmed by the positive correlation between changes in the total amount of bank loans and capital reallocation, 0.47 and the negative correlation between changes in bank loans and the variance in the marginal product of capital of −0.57, both of which are consistent with the patterns in the data. Both versions of the model generate a positive correlation between capital utilization and the measure of TFP, similar in magnitude to that observed in the data.

The model with TFP shocks also produces a procyclical capital reallocation. A positive productivity shock implies that the production technology becomes more efficient. As a result, capital moves from the storage technology to the productive sectors. However, a positive productivity shock is associated with a relatively small increase in banks’ net worth but a surge in total investment. As a result, most of the new investment flows into the unconstrained low-productivity firms, and the cross-sectional dispersion in the marginal product of capital increases. The model with TFP shocks therefore generates a counterfactually procyclical benefit of capital reallocation: the dispersion in the marginal product of capital is positively correlated with aggregate TFP, as shown in Table 3.

**Capital reallocation and financial flows of corporate sector** Our model has predictions not only on how investment activities (including new investment and capital reallocation) vary over the business cycle but also on how the real and financial sides of investment are correlated.21 To see this, adding $RA_{j,t+1}$ to both sides of the banks’ budget constraint.

---

21We thank an anonymous referee for encouraging us to study this issue.
(8), we have
\[ N_{j,t} + B_{j,t} + RA_{j,t+1} = (1 - \delta) K_{j,t} + I_{j,t} + RA_{j,t+1}. \]

Note that the right side of the above equation is the total amount of capital used by firm \( j \).

We can further expand the left side of the above equation to obtain
\[
(N_{j,t} + R_{f,t}B_{j,t-1}) + (B_{j,t} - R_{f,t}B_{j,t-1}) + RA_{j,t+1} = (1 - \delta) K_{j,t} + I_{j,t} + RA_{j,t+1} \tag{54}
\]

The left side is the financial side of investment: the term \( N_{j,t} + R_{f,t}B_{j,t-1} \) is the total assets at time \( t \) (equity plus debt), \( B_{j,t} - R_{f,t}B_{j,t-1} \) is external financing (newly issued debt) obtained from households, and \( RA_{j,t+1} \) is interbank loans obtained for the purpose of capital reallocation. On the right side, \( I_{j,t} \) is new investment and \( RA_{j,t+1} \) is capital reallocation. Equation (54) therefore provides an accounting identity that links the real and financial sides of capital reallocation and allows us to test the implications of our model against the relevant empirical evidence provided in the literature.

Chari (2012) and Shourideh and Zetlin-Jones (2017) define the external-financing-to-total-investment ratio as the total amount of external financing of all firms divided by the total investment of all firms. In the context of Equation (54), \( B_{j,t} - R_{f,t}B_{j,t-1} + RA_{j,t+1} \) is the total amount of external financing through newly issued debt and interbank loans, and total investment includes both new investment and capital reallocation, \( I_{j,t}^{tot} = I_{j,t} + RA_{j,t+1} \). Consistent with the construction in Chari (2012) and Shourideh and Zetlin-Jones (2017), the external-financing-to-total-investment ratio in our model can be computed as
\[
\frac{\sum_j \max[(B_{j,t} - R_{f,t}B_{j,t-1}) + RA_{j,t+1}, 0]}{\sum_j I_{j,t}^{tot}}.
\]

Empirically, we replicate Shourideh and Zetlin-Jones (2017) to construct the external-financing-to-total-investment ratio in the data and in our calibrated model (construction details are provided in Appendix B). As shown in Table 3, quantitatively, our model reproduces an average external-financing-to-total-investment ratio of 0.17, broadly in line with a ratio of 0.22 in the data.

In addition, Eisfeldt and Shi (2018) document a positive correlation between external financing and capital reallocation. Our model also generates a positive correlation of 0.62, closely matching its empirical counterpart, 0.78. This evidence suggests that external financing is an important funding source for capital reallocation and supports our model mechanism that financial frictions can disrupt external financing to firms, thereby impeding capital reallocation efficiency.
Crisis dynamics  To further understand the implications of our model on volatility dynamics and economic recessions, in Table 4, we report the moments of macroeconomic quantities in the data and those in our model with financial shocks for recession periods and non-recession periods, respectively. For simplicity, we use a rule-of-thumb classification and define a recession as years with two consecutive quarters of declines in real GDP in both the data and the model. Our definition yields about 15% of the sample being classified as a recession in both the data and the model simulation, which is similar to the NBER’s definition of recession.

As shown in Table 4, in our model, as in the data, the volatility in output, consumption, and investment is significantly more volatile in recession periods because of a stronger amplification mechanism from the banking sector during bad time, so our model matches the countercyclical pattern in the volatility of aggregate output quite well. Because of the absence of adjustment costs, most of the volatility in output is absorbed by investment, not by consumption. In addition, our model generates the empirical pattern that in recessions the level of capital utilization rate drops, its volatility increases, and the spread between the interbank lending rate and the household deposit rate widens. All the above features are the endogenous outcomes of the financial frictions in the model.

7 Conclusions

We have presented a general equilibrium model with financial intermediaries and capital reallocation. Our model emphasizes the role of financial intermediaries in reallocating capital across firms with heterogeneous productivity. Shocks to financial frictions alone may account for a large fraction of the fluctuations in measured TFP and aggregate output. Our calibrated model is consistent with the salient features of business cycle variations in macroeconomic quantities and asset prices. In particular, our model successfully generates countercyclical volatility in aggregate consumption and output, and countercyclical dispersion in the cross section.

Our model represents a first step in understanding the link between financial market frictions and capital reallocation. Several related issues may be important for future research. First, we have followed the Ger setup and attributed all external financing to bank loans in our model. In practice, the public equity market also plays an important role. Modeling financial frictions through both channels requires a model that distinguishes bank net worth, which affects banks’ borrowing capacity, and firm net worth, which determines firms’

\footnote{Although we do not extend our model to allow for adjustment costs for parsimony, the fit of our model can be improved by allowing for such an extension.}
borrowing capacity on public capital markets, such as the Rampini and Viswanathan (2019) model. Explicitly modeling capital reallocation in setups similar to that of Rampini and Viswanathan (2019) can provide a better understanding of the relative importance of banks and the public capital market. Second, an important next step is to infer or impute shocks to financial frictions from the data and investigate whether our model can account for the realized variations in macroeconomic quantities and asset prices once these shocks are fed into the model. We leave this step for future research.
Table 1
Calibrated Parameters and Targeted Moments

<table>
<thead>
<tr>
<th>Group 1: Parameters chosen from previous literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>discount rate</td>
</tr>
<tr>
<td>depreciation</td>
</tr>
<tr>
<td>capital share</td>
</tr>
<tr>
<td>markup</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group 2: Parameters chosen by matching a relevant moment</th>
</tr>
</thead>
<tbody>
<tr>
<td>storage technology</td>
</tr>
<tr>
<td>storage technology</td>
</tr>
<tr>
<td>storage technology</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group 3: Parameters jointly determined by targeted moments</th>
</tr>
</thead>
<tbody>
<tr>
<td>relative productivity</td>
</tr>
<tr>
<td>prob. of $\varepsilon_H$</td>
</tr>
<tr>
<td>aggregate productivity</td>
</tr>
<tr>
<td>bankers’ discount rate</td>
</tr>
<tr>
<td>bankers’ outside option</td>
</tr>
</tbody>
</table>

Targeted moments that jointly determine parameters in Group 3

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>mean capital reallocation</td>
</tr>
<tr>
<td>output ratio ($\frac{\pi y_H}{(1-\pi)y_L}$)</td>
</tr>
<tr>
<td>mean aggregate growth</td>
</tr>
<tr>
<td>leverage of banking sector</td>
</tr>
<tr>
<td>investment-output ratio</td>
</tr>
<tr>
<td>mean external financing to total investment</td>
</tr>
<tr>
<td>power law slope of size distribution</td>
</tr>
</tbody>
</table>

Table 1 lists the parameter values we use in our model and the macroeconomic moments used to calibrate these parameter values.
Table 2
Unconditional Moments

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>TFP Shocks</th>
<th>Financial Shocks</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Macroeconomic moments</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E[\Delta \ln Y]$</td>
<td>2.00%</td>
<td>2.00%</td>
<td>2.00%</td>
</tr>
<tr>
<td>$Std[\Delta \ln Y]$</td>
<td>3.00%</td>
<td>3.00%</td>
<td>3.00%</td>
</tr>
<tr>
<td>$Std[\Delta \ln C]$</td>
<td>2.04%</td>
<td>0.64%</td>
<td>0.83%</td>
</tr>
<tr>
<td>$Std[\Delta \ln I]$</td>
<td>6.53%</td>
<td>10.18%</td>
<td>12.51%</td>
</tr>
<tr>
<td>$Corr[\Delta \ln C, \Delta \ln I]$</td>
<td>0.40</td>
<td>0.59</td>
<td>0.17</td>
</tr>
<tr>
<td><strong>Interest rates</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E[\ln R_f]$</td>
<td>0.86%</td>
<td>0.24%</td>
<td>0.16%</td>
</tr>
<tr>
<td>$Std[\ln R_f]$</td>
<td>2.29%</td>
<td>0.01%</td>
<td>1.03%</td>
</tr>
<tr>
<td><strong>Cyclical properties</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Corr[\Delta \ln Y, Vol(\Delta \ln Y')]$</td>
<td>-0.19</td>
<td>0.00</td>
<td>-0.19</td>
</tr>
<tr>
<td>$Corr[\Delta \ln Y, Vol(R'_M)]$</td>
<td>-0.37</td>
<td>0.00</td>
<td>-0.22</td>
</tr>
<tr>
<td>$Var[\Delta uf(\phi)]/Var[\Delta TFP]$</td>
<td>0.09</td>
<td>0.98</td>
<td></td>
</tr>
</tbody>
</table>

Table 2 documents moments of macroeconomic quantities and interest rates in the United States (1930-2009), those generated by our model with TFP shocks, and those generated by our model with financial shocks. $Corr[\Delta \ln Y, Vol(\Delta \ln Y')]$ stands for the correlation of output growth and the variance of future output growth. The latter is calculated as the realized variance of quarterly output growth for the next two years. $Var[\Delta uf(\phi)]/Var[\Delta TFP]$ stands for the fraction of variance in output that can be accounted for by changes in the efficiency of capital reallocation.
Table 3
Capital Reallocation and Financial Intermediation

<table>
<thead>
<tr>
<th>Moments</th>
<th>Data</th>
<th>TFP Shocks</th>
<th>Financial Shocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capital reallocation and capital utilization</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E [RA/I]$</td>
<td>0.22</td>
<td>0.22</td>
<td>0.43</td>
</tr>
<tr>
<td>$E [u]$</td>
<td>80%</td>
<td>80%</td>
<td>77%</td>
</tr>
<tr>
<td>$E [EF/I^{tot}]$</td>
<td>0.22</td>
<td>0.15</td>
<td>0.17</td>
</tr>
<tr>
<td>$E [Var (ln MPK)]$</td>
<td>0.95</td>
<td>0.22</td>
<td>0.18</td>
</tr>
<tr>
<td>$Std [Var (ln MPK)]$</td>
<td>0.25</td>
<td>0.01</td>
<td>0.23</td>
</tr>
<tr>
<td>$Vol [u]$</td>
<td>5%</td>
<td>0.50%</td>
<td>5%</td>
</tr>
<tr>
<td>Cyclical properties</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Corr [\Delta \ln TFP, RA/I]$</td>
<td>0.25</td>
<td>0.14</td>
<td>0.29</td>
</tr>
<tr>
<td>$Corr [\Delta \ln TFP, Var (ln MPK)]$</td>
<td>−0.24</td>
<td>0.52</td>
<td>−0.33</td>
</tr>
<tr>
<td>$Corr [\Delta \ln TFP, \ln u]$</td>
<td>0.56</td>
<td>0.62</td>
<td>0.57</td>
</tr>
<tr>
<td>$Corr [\Delta \ln BL, RA/I]$</td>
<td>0.27</td>
<td>0.44</td>
<td>0.47</td>
</tr>
<tr>
<td>$Corr [\Delta \ln BL, Var (ln MPK)]$</td>
<td>−0.37</td>
<td>0.70</td>
<td>−0.57</td>
</tr>
<tr>
<td>$Corr [EF/I^{tot}, RA/I]$</td>
<td>0.78</td>
<td>0.72</td>
<td>0.62</td>
</tr>
</tbody>
</table>

Table 3 documents the moments of capital reallocation and capital utilization in the data, and those generated by our models. Our construction of capital reallocation series follows Eisfeldt and Rampini (2006), and that of the external-financing-to-total-investment ratio follows Shourideh and Zetlin-Jones (2017). Construction details, as well as the calculation of the cross-sectional dispersion in the log marginal product of capital (ln MPK), can be found in Appendix A. The capacity utilization rate ($u$) is published by the Federal Reserve Bank of St. Louis.
Table 4
Crisis Dynamics

<table>
<thead>
<tr>
<th>Moments</th>
<th>Non-Recession Years</th>
<th>Recession Years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Data</td>
<td>Model</td>
</tr>
<tr>
<td>$Vol[\Delta \ln Y]$</td>
<td>2.29%</td>
<td>2.39%</td>
</tr>
<tr>
<td>$Vol[\Delta \ln C]$</td>
<td>1.53%</td>
<td>0.61%</td>
</tr>
<tr>
<td>$Vol[\Delta \ln I]$</td>
<td>5.25%</td>
<td>11.14%</td>
</tr>
<tr>
<td>$E [u]$</td>
<td>81.3%</td>
<td>77.12%</td>
</tr>
<tr>
<td>$Vol [u]$</td>
<td>3.68%</td>
<td>5.00%</td>
</tr>
</tbody>
</table>

Table 4 documents the first and second moments of macroeconomic quantities and interest rates in recession and non-recession periods in the data and in our model with financial shocks. In both the model and the data, a recession is classified as two consecutive quarters of decline in real GDP.
Appendix: Financial Intermediation and Capital Reallocation

A  Data Construction

Measurement of the efficiency of capital reallocation  We first derive an aggregation result that is similar to Hsieh and Klenow (2009) and Hopenhayn and Neumeyer (2008). In fact, the product market of our model is a special case of the above papers without labor market distortions. Consider the maximization problem in Equation (2), first order conditions with respect to $K_j$ and $L_j$ imply

$$ (1 - \alpha) p_j Y_j = MPL \cdot L_j, \quad \alpha p_j Y_j = MPK_j \cdot K_j $$

Together, the above imply

$$ \frac{K_j}{L_j} = \frac{MPL}{MPK_j} \frac{\alpha}{1 - \alpha} $$

(56)

To save notation, we denote $A_j = \bar{A} z_j^{\frac{1}{1-\eta}}$ in this section. Note also that the total output of firm $j$ can be written as

$$ Y_j = A_j K_j^\alpha L_j^{1 - \alpha} = A_j \left[ \frac{K_j}{L_j} \right]^\alpha L_j $$

(57)

$$ = A_j \left[ \frac{L_j}{K_j} \right]^{1 - \alpha} K_j. $$

(58)

We can use Equations (56) and (57) to write $L_j$ as a function of $Y_j$, and use Equations (56) and (58) to write $K_j$ as a function of $Y_j$

$$ L_j = \frac{Y_j}{A_j} \left[ \frac{\alpha MPL}{(1 - \alpha) MPK_j} \right]^{-\alpha}, \quad K_j = \frac{Y_j}{A_j} \left[ \frac{\alpha MPL}{(1 - \alpha) MPK_j} \right]^{1 - \alpha}. $$

(59)

Note that the final goods producer’s optimality condition implies that $Y_j = p_j^{-\eta} Y$. Using this equation to replace $Y_j$ in the above equation and integrate across all $j$, we have

$$ \bar{K} = \int \frac{p_j^{-\eta}}{A_j} \left[ \frac{1}{MPK_j} \right]^{1 - \alpha} dj \left[ \frac{\alpha MPL}{1 - \alpha} \right]^{1 - \alpha} Y $$

$$ \bar{L} = \int \frac{p_j^{-\eta}}{A_j} \left[ \frac{1}{MPK_j} \right]^{-\alpha} dj \left[ \frac{\alpha MPL}{1 - \alpha} \right]^{-\alpha} Y, $$

(60)

(61)

where $\bar{K}$ and $\bar{L}$ stands for the total capital and total labor employed for production, respec-
tively. Together, Equations (60) and (61) imply
\[ Y = \frac{\bar{K}^\alpha \bar{L}^{1-\alpha}}{\left[ \int \frac{p_j^{-\eta}}{A_j} \left[ \frac{1}{MPK_j} \right]^{-\alpha} dj \right]^\alpha \left[ \int \frac{p_j^{-\eta}}{A_j} \left[ \frac{1}{MPK_j} \right]^{-\alpha} dj \right]^{1-\alpha}}. \]  

(62)

We next express \( p_j \) in Equation (62) as a function of productivity and prices. Note that Equation (55) implies
\[ MPK_j \cdot K_j + MPL \cdot L_j = p_j Y_j. \]  

(63)

Also, Equation (59) implies
\[ MPK_j \cdot K_j + MPL \cdot L_j = \frac{Y_j}{A_j} \left[ \frac{MPL}{1-\alpha} \right]^{1-\alpha} \left[ \frac{MPK_j}{\alpha} \right]^\alpha. \]  

(64)

Combining Equations (63) and (64), we have:
\[ p_j = \frac{1}{A_j} \left[ \frac{MPL}{1-\alpha} \right]^{1-\alpha} \left[ \frac{MPK_j}{\alpha} \right]^\alpha. \]  

(65)

Note that because the price of the final goods is normalized to one, we have \( \int p_j^{-\eta} dj \right]^{1-\eta} = 1. \) Integrating Equation (65) over \( j \), we have:
\[ \left[ \frac{MPL}{(1-\alpha)} \right]^{1-\alpha} = \left\{ \int \left[ \frac{1}{A_j} \left[ \frac{MPK_j}{\alpha} \right]^{\alpha} \right]^{1-\eta} dj \right\}^{-\frac{1}{1-\eta}}. \]  

(66)

Together, (65) and (66) imply
\[ p_j = \frac{1}{A_j} \left[ \frac{MPK_j}{\alpha} \right]^{\alpha} \left\{ \int \left[ \frac{1}{A_j} \left[ \frac{MPK_j}{\alpha} \right]^{\alpha} \right]^{1-\eta} dj \right\}^{-\frac{1}{1-\eta}}. \]  

(67)

Replacing \( p_j \) in Equation (62) with Equation (67), and using \( A_j = A z_j^{\frac{1}{\eta-1}} \), we can write
\[ Y = TFP \bar{K}^\alpha \bar{L}^{1-\alpha}, \]  

where \( TFP = A \times EF \), and
\[ EF = \frac{\left\{ \int \frac{z_j}{MPK_j^{\alpha\eta-1}} dj \right\}^{1+\alpha\eta-\alpha}}{\alpha^{1+\alpha\eta-\alpha}} \left[ \int \frac{z_j}{MPK_j^{1+\alpha\eta-\alpha}} dj \right]^\alpha. \]  

(68)
Thanks to the normalization assumptions that \( z_0 = 1 \) and \( E[e^{ij}] = 1 \), we obtain \( \int z_j dj = \int e^{ij} dj = 1 \). It is therefore straightforward to show that \( TFP = A \) if \( MPK_j = MPK \) for all \( j \). The \( EF \) defined above is our measure of the efficiency measure of capital reallocation. Under the assumption \( \ln z_j \) and \( \ln MPK_j \) are jointly normally distributed, we can show that

\[
\ln EF = -\frac{1}{2} [\alpha(\eta - 1) + 1] \alpha \sigma^2, \tag{69}
\]

where \( \sigma^2 \) is the cross-sectional variance of the log marginal product of capital.

**Misallocation and TFP** In Figure 1, we plot the measured efficiency of capital reallocation and total factor productivity. We measure the cross-sectional dispersion of TFPR following Hsieh and Klenow (2009). In the context of our model, Equation (55) implies

\[
MPK_j = \alpha \frac{p_j Y_j}{K_j}.
\]

Following Chen and Song (2013), we measure \( MPK_j \) by the ratio of operating income before depreciation (OIBDP) to one-year-lag net plant, property and equipment (PPENT). As in Hsieh and Klenow (2009), we focus on the manufacturing sector and compute the cross-sectional dispersion measure within narrowly defined industries (as classified by the 4-digit standard industry classification code). Specifically, for firm \( j \) in industry \( h \), we compute

\[
\frac{MPK_{j,h}}{MPK_h} = \frac{\alpha \frac{p_{j,h} Y_{j,h}}{K_{j,h}}}{\alpha \frac{p_{h} Y_{h}}{K_{h}}} = \frac{p_{j,h} Y_{j,h}}{p_{h} Y_{h}},
\]

where \( \frac{p_{h} Y_{h}}{K_{h}} \) is measured at the industry level. We then compute the variance of \( \frac{MPK_{j,h}}{MPK_h} \) within each industry \( h \) and take average across all the industries for each year. This is our empirical measure of \( \sigma^2 \) in Equation (69). We use the first order approximation in Equation (69) to construct the time series of the efficiency measure of capital reallocation, which is the solid line in Figure 1. The measure of total factor productivity is directly taken from the published TFP series on the U.S. Bureau of Labor Statistics website. Both series are HP filtered.

**Capital reallocation** We follow Eisfeldt and Rampini (2006) to construct the capital reallocation measure from the annual firm level data in Compustat. Reallocation is defined as the sum of acquisitions and sales of property, plant and equipment. When we calculate the its ratio to investment, investment is defined as capital expenditures plus acquisitions. We deflate all the series using CPI from Federal Reserve Bank of St. Louis web site to remove any effects from variation in nominal prices.
**External financing** We follow Shourideh and Zetlin-Jones (2017) to construct the external financing to total investment ratio from the annual firm level data in Compustat. Our data construction follows Shourideh and Zetlin-Jones (2017). We first construct available funds (AF) as OANCF (Operating Activities/Net Cash Flow) if the firm reports SCF its statement of cash flows using format code 7 (SGF = 7); otherwise, we define AF as FOPT (Funds from Operations – Total). The total investment or gross investment is constructed as the sum of CAPX (Capital Expenditure) and AQC (Acquisitions) minus SPPE (Sale of Property). The external financing (EF) is then defined as the difference between total investment and the available funds (AF) if it is positive.

**Total volume of bank loans** We measure the total volume of bank loans of non-financial corporate sector through the aggregate balance sheet of nonfinancial corporate business (Table B.103) as reported in the U.S. Flow of Funds Table. In particular, the bank loan is calculated as the difference between total credit market liability, the sum of debt securities (Line 26) and loans (Line 30), and corporate bond (Line 26). Under this construction, bank loans consist of the following credit market liability items: commercial paper (Line 27), municipal securities (Line 28), depository institution loans (Line 31), other loans and advances (Line 32) and mortgages (Line 33). The bank loan measure is deflated using CPI from Federal Reserve Bank of St. Louis website.

### B Misallocation and Aggregation on the Product Market

**Proof of Proposition 1** Using the fact that \( Y_j = p_j^{-\eta} Y \) and Equation (56), we can write

\[
    p_j Y_j = Y_j^{1-\frac{1}{\eta}} Y_j^{\frac{1}{\eta}} = [A_j K_j^\alpha L_j^{1-\alpha}]^{1-\frac{1}{\eta}} Y_j^{\frac{1}{\eta}}. \tag{70}
\]

Combining Equation (55) and Equation (70), we can write the marginal product of capital as

\[
    MPK_j = \alpha A_j^{1-\frac{1}{\eta}} K_j^{\alpha(1-\frac{1}{\eta})-1} L_j^{(1-\alpha)(1-\frac{1}{\eta})} Y_j^{\frac{1}{\eta}}, \tag{71}
\]

and write the marginal product of labor as

\[
    (1 - \alpha) [A_j K_j^\alpha]^{1-\frac{1}{\eta}} Y_j^{\frac{1}{\eta}} = MPLL_j^{1-(1-\alpha)(1-\frac{1}{\eta})}. \tag{72}
\]

Equation (72) implies \( L_j \propto [A_j K_j^\alpha]^{1-(1-\alpha)(1-\frac{1}{\eta})} \). Using the resource constraint, \( \int L_j d_j = \)
\[ L_j = \frac{[A_j K_j^a]^{\frac{1}{1-\delta}}}{\int [A_j K_j^a]^{\frac{1}{1-\delta}}(1-\delta)\,dj} \bar{L} \]  

(73)

Together, Equations (71) and (73) and the assumption that \( A_j = \bar{A} z_j^{\frac{1}{1-\delta}} \) imply that \( MPK_j \) can be written as a function of \( (z_j, K_j) \):

\[ MPK_j = \alpha \bar{A} z_j^{(1-\xi)K_j^{-\xi}} \left\{ \int z_j^{(1-\xi)K_j^{-\xi}}dj \right\}^{\frac{1}{(\eta-1)(1-\xi)^{-1}}} \]  

(74)

where we normalize total labor supply \( \bar{L} = 1 \) and denote \( \xi = \frac{\alpha \eta - \alpha}{1 + \alpha \eta - \alpha} \) as in the proposition. Using Equation (74) to replace \( MPK_j \) in the expression of \( EF \) in Equation (68), to get

\[ EF = \frac{\left\{ \int z_j^{1-\xi}K_j^{-\xi}dj \right\}^{\frac{\alpha}{\xi}}}{\left\{ \int K_jdj \right\}^{\alpha}}. \]  

(75)

Using the fact that the total amount of capital utilized in production is \( \int K_jdj = uK \) and total labor supply \( \bar{L} = 1 \), we can derive total output \( Y \) as:

\[ Y = \bar{A} \times EF \times (uK)^\alpha, \]

\[ = \bar{A} \left\{ \int z_j^{1-\xi}K_j^{-\xi}dj \right\}^{\frac{\alpha}{\xi}} \left\{ \int z_j^{1-\xi} \left( \frac{K_j}{uK} \right) dj \right\}^{\frac{\alpha}{\xi}} (uK)^\alpha. \]

This gives Equation (15) in the paper.

In the special case that there are only two possible realizations of idiosyncratic productivity shocks, \( \bar{\varepsilon}_H \) and \( \bar{\varepsilon}_L \), as we consider in our paper, under the assumption that \( \bar{A} = \bar{A} K^{1-\alpha} \), plugging in the relation in Equation (16) together with the definition of \( \phi \), we can derive the expression for the total output and the marginal product of capital in Proposition 1.

**Aggregation in the fully dynamic model**  In this subsection, we provide details why Proposition 1 applies to the fully dynamic model. Because \( \frac{z_j+1}{K_{j+1}^{1-\alpha} + RA_{j+1}^{1-\alpha}} \) must be equalized on all islands in the same sector, the individual ratios must equal to the average ratio of the sector \( \frac{E[z_{j+1}^{1-\alpha} | \varepsilon_{j+1} = \varepsilon]}{E[K_{j+1}^{1-\alpha} + RA_{j+1}^{1-\alpha} | \varepsilon_{j+1} = \varepsilon]} \). For \( \varepsilon = \bar{\varepsilon}_H, \bar{\varepsilon}_L \), the integrals in Equation (34) can therefore be
written as
\[
\int_{\varepsilon_{j,t+1} = \varepsilon}^{z_{j+1}} \left( \frac{K_{j,t+1} + RA_{j,t+1}}{K_{j,t+1} + RA_{j,t+1}} \right)^{(1-\xi)} (K_{j,t+1} + RA_{j,t+1}) \, dj
\]
\[
= \left( \frac{E [z_{j,t+1} | \varepsilon_{j,t+1} = \varepsilon]}{E [K_{j,t+1} + RA_{j,t+1} | \varepsilon_{j,t+1} = \varepsilon]} \right)^{1-\xi} \int_{\varepsilon_{j,t+1} = \varepsilon}^{z_{j+1}} (K_{j,t+1} + RA_{j,t+1}) \, dj.
\]  (76)

Note that \( E [z_{j,t+1} | \varepsilon_{j,t+1} = \varepsilon] = E [z_{j,t} e^{\varepsilon_{j,t+1}} | \varepsilon_{j,t+1} = \varepsilon] = e^{\varepsilon} \), as \( \varepsilon_{j,t+1} \) is independent of \( z_{i,t} \) and \( E [z_{j,t}] = 1 \). Also, if we define \( \phi_{t+1} = \frac{E [K_{j,t+1} + RA_{j,t+1} | \varepsilon_{j,t+1} = \varepsilon]}{E [K_{j,t+1} + RA_{j,t+1} | \varepsilon_{j,t+1} = \varepsilon]} \) as the ratio of the average size of firms in the two sectors, then \( \frac{E [K_{j,t+1} + RA_{j,t+1} | \varepsilon_{j,t+1} = \varepsilon]}{\pi u_{t+1} K_{t+1}} = \pi \phi_{t+1} (1-\pi) \) as \( u_{t+1} K_{t+1} \) is the average size of all firms in the economy. In addition, because the total measure of the \( \varepsilon_H \) sector is \( \pi \),
\[
\int_{\varepsilon_{j,t+1} = \varepsilon_H} (K_{j,t+1} + RA_{j,t+1}) \, dj = \pi E [K_{j,t+1} + RA_{j,t+1} | \varepsilon_{j,t+1} = \varepsilon_H]
\]
\[
= \frac{\phi_{t+1}}{\pi \phi_{t+1} + (1-\pi) u_{t+1} K_{t+1}}.
\]  (77)

We can use Equations (76) and (77) and write
\[
\int_{\varepsilon_{j,t+1} = \varepsilon_H} z_{j+1}^{(1-\xi)} (K_{j,t+1} + RA_{j,t+1}) \, dj = \pi \left( \frac{e^{\varepsilon_H}}{\phi_{t+1}} \frac{\phi_{t+1}}{\pi \phi_{t+1} + (1-\pi) u_{t+1} K_{t+1}} \right)^{1-\xi} \left( \frac{\phi_{t+1}}{\pi \phi_{t+1} + (1-\pi)} \right)^{\xi} (u_{t+1} K_{t+1})^{\xi}.
\]

We can simplify \( \int_{\varepsilon_{j,t+1} = \varepsilon_H} z_{j+1}^{(1-\xi)} (K_{j,t+1} + RA_{j,t+1}) \, dj \) similarly and write Equation (34) as:
\[
Y_{t+1} = \tilde{A}_{t+1} \left\{ \pi \epsilon^{(1-\xi)\varepsilon_H} (\frac{\phi_{t+1}}{\pi \phi_{t+1} + (1-\pi)})^{\xi} + (1-\pi) \epsilon^{(1-\xi)\varepsilon_L} (\frac{1}{\pi \phi_{t+1} + (1-\pi)})^{\xi} \right\} (u_{t+1} K_{t+1})^{\xi},
\]

where \( f (\phi) \) is defined in Proposition 1.

C Proof of Proposition 2

Note the policy functions for \( \phi (A,s) \) and \( u (A,s) \) are determined by conditions (26), (27), (28) and (29).

We note that the difference between the LHS of Equation (26) and the LHS of Equation
is:

\[(1 - \theta) Q_L - [Q - (1 - \theta) Q_L] \left( \frac{u}{\pi \phi + 1 - \pi} - 1 \right) \]

\[- \left\{ (1 - \theta) Q_H - [Q - (1 - \theta) Q_H] \left( \frac{u \phi}{\pi \phi + 1 - \pi} - 1 \right) \right\}, \]

\[= \frac{u}{\pi \phi + 1 - \pi} \{(1 - \theta) Q_L - (1 - \theta) \phi Q_H + (\phi - 1) Q\}. \]

To study the nature of the binding constraint, it is convenient to define \(\Delta\) as

\[\Delta (A, \phi, u) = (1 - \theta) Q_L (A, \phi, u) - (1 - \theta) \phi Q_H (A, \phi, u) + (\phi - 1) Q (u), \]

\[= (1 - \theta) MPK_L (A, \phi, u) - (1 - \theta) \phi MPK_H (A, \phi, u) + (\phi - 1) MPK (u) \]

\[+ \theta (\phi - 1) (1 - \delta), \]

in which market clearing condition on the reallocation market implies \(Q (u) = MPK (u) + 1 - \delta.\)

Note that we have three cases:

- Only the constraint on high productivity island (26) binds \(\implies\) \(\Delta > 0.\)
- Both constraints (26) and (27) bind \(\implies\) \(\Delta = 0.\)
- Only the constraint on low productivity island (27) binds \(\implies\) \(\Delta < 0.\)

We also define the LHS of constraint on the high productivity island as a function of \((A, u, \phi):\)

\[\Psi (A, \phi, u) = (1 - \theta) Q_H (A, \phi, u) - [Q (u) - (1 - \theta) Q_H (A, \phi, u)] \frac{RA_H}{K}, \]  

\[= (1 - \theta) [MPK_H (A, \phi, u) + (1 - \delta)] \]

\[- [MPK (u) - (1 - \theta) MPK_H (A, \phi, u) + \theta (1 - \delta)] \left( \frac{u \phi}{\pi \phi + 1 - \pi} - 1 \right). \]

**First best case, no constraint binds**

\[Q_H = Q_L = Q = \alpha u^{\alpha - 1} A + 1 - \delta. \]

The optimal capital utilization Equation (6) implies

\[\alpha u^{\alpha - 1} A + 1 - \delta = b_0 (1 - u)^{\nu - 1}. \]  

49
Equation (79) defines the capital utilization rate as a function of productivity, \( \hat{u}(A) \). We define \( \hat{Q} \) as the price of capital in the first best case given the productivity \( A \):
\[
\hat{Q}(A) = \alpha \hat{u}(A)^{\alpha-1} A + 1 - \delta.
\]

Also define \( \hat{s}(A) \) to be the highest level of \( s \) such that there is no capital misallocation:
\[
\hat{s}(A) = \Psi \left( A, \hat{\phi}, \hat{u}(A) \right),
\]
in which \( \hat{\phi} \) is defined as \( \hat{\phi} = e^{\varepsilon_H - \varepsilon_L} \), the first best capital reallocation ratio which equalize the marginal products of capital across all islands.

One simplifies the expression (78) and shows:
\[
\Psi \left( A, \hat{\phi}, \hat{u}(A) \right) = \frac{\hat{Q}(A)}{\pi \hat{\phi} + 1 - \pi} \left\{ (1 - \theta \hat{u}(A)) \hat{\phi} - (1 - \pi)(\hat{\phi} - 1) \right\}.
\]

**Claim 1** If \( s_1 \leq \hat{s}(A) \) then the optimal policy is given by:
\[
\phi(A, s) = \hat{\phi}, \quad u(A, s) = \hat{u}.
\] (80)

**Proof.** We need to show that Equations (26), (27), (30) and (31) are satisfied with appropriate choices of the Lagrangian multipliers. Under the proposed policies and prices, the LHS of Equation (26) is
\[
(1 - \theta) \left[ MPK_H \left( A, \hat{\phi}, \hat{u}(A) \right) + (1 - \delta) \right]
\]
\[
- \left[ MPK_H \left( A, \hat{\phi}, \hat{u}(A) \right) - (1 - \theta) MPK_H \left( A, \hat{\phi}, \hat{u}(A) \right) + \theta (1 - \delta) \right] \left( \frac{\hat{u}(A) \hat{\phi}}{\pi \hat{\phi} + (1 - \pi) - 1} \right),
\]
\[
= \Psi \left( A, \hat{\phi}, \hat{u}(A) \right) = \hat{s}(A) \geq s.
\]

Also,
\[
\Delta = (1 - \theta) MPK_L \left( A, \hat{\phi}, \hat{u}(A) \right) - (1 - \theta) \phi MPK_H \left( A, \hat{\phi}, \hat{u}(A) \right) + (\phi - 1) MPK \left( \hat{u}(A) \right),
\]
\[
= \theta (\phi - 1) MPK_H \left( A, \hat{\phi}, \hat{u}(A) \right) > 0.
\]

Therefore, both Equations (26) and (27) are satisfied. Finally, note that Equation (80) implies that
\[
MPK_H \left( A, \hat{\phi}, \hat{u}(A) \right) = MPK_L \left( A, \hat{\phi}, \hat{u}(A) \right) = MPK \left( \hat{u}(A) \right) = \alpha \hat{u}(A)^{\alpha-1} A,
\]
and therefore \( \xi_H \left( A, \hat{\phi}, \hat{u}(A) \right) = \xi_L \left( A, \hat{\phi}, \hat{u}(A) \right) = 0 \). As a result, the Kuhn-Tucker conditions
(30) and (31) for optimality are satisfied. ■

**Only the constraint on high productivity islands bind:** In this case:

\[ Q_H (A, \phi, u) > Q_L (A, \phi, u) = Q (A, \phi, u) \]

where

\[
Q_H (A, \phi, u) = \alpha A u^{\alpha - 1} f (\phi) \frac{\pi \phi + 1 - \pi}{\pi \hat{\phi}^{1 - \xi} \phi^\xi + 1 - \pi} + 1 - \delta, \\
Q_L (A, \phi, u) = \alpha A u^{\alpha - 1} f (\phi) \frac{\pi \phi + 1 - \pi}{\pi \hat{\phi}^{1 - \xi} \phi^\xi + 1 - \pi} \left( \frac{\hat{\phi}}{\phi} \right)^{1 - \xi} + 1 - \delta,
\]

in which \( \hat{\phi} \) is defined as \( \hat{\phi} = e^{\varepsilon H - \varepsilon L} \).

The optimality condition for capital utilization implies:

\[
\alpha A u^{\alpha - 1} f (\phi) \frac{\pi \phi + 1 - \pi}{\pi \hat{\phi}^{1 - \xi} \phi^\xi + 1 - \pi} + (1 - \delta) = b_0 (1 - u)^{\nu - 1}
\]

The above equation defines the capital utilization rate as a function of \((A, \phi)\), which we will define as \( u_L (A, \phi) \). Let \( \overline{\phi} (A) \) be the unique solution to

\[
\Delta (A, \overline{\phi} (A), u_L (A, \overline{\phi} (A))) = 0,
\]

and define \( \overline{s} (A) \) to be the highest level of \( s \) such that only the constraints on high productivity islands are binding

\[
\overline{s} = \Psi (A, \overline{\phi} (A), u_L (A, \overline{\phi} (A))\).
\]

Given the definition of \( u_L (A, \phi) \), we can show that

\[
\Psi (A, \phi, u_L (A, \phi)) = \frac{MPK_L (A, \phi, u_L (A, \phi))}{\pi \phi + 1 - \pi} \\
\left\{ (1 - \theta) u_L (A, \phi) \hat{\phi}^{1 - \xi} \phi^\xi \\
- [(\phi - 1) (1 - \pi) - \phi (1 - u_L (A, \phi))] \right\} \\
+ \frac{1 - \delta}{\pi \phi + 1 - \pi} \left\{ (1 - \theta) u_L (A, \phi) \phi - \left[ \begin{array}{c} (\phi - 1) (1 - \pi) \\
- \phi (1 - u_L (A, \phi)) \end{array} \right] \right\},
\]

51
\[
\Delta(A, \phi, u_L(A, \phi)) = \operatorname{MPK}_L(A, \phi, u_L(\phi)) \left[ (\phi - 1) - (1 - \theta) \left( \frac{\phi^{1-\xi} \phi^\xi - 1}{\phi - 1} \right) \right] + (1 - \delta) \theta (\phi_1 - 1).
\]

Using the above expressions, we can prove that \(\Psi(A, \phi, u_L(A, \phi))\) is strictly decreasing in \(\phi\) and \(\Delta(A, \phi, u_L(A, \phi))\) is strictly increasing functions of \(\phi\). As a result, i) \(\phi \geq \bar{\phi}(A)\) if and only if \(\Delta(A, \phi, u_L(A, \phi)) \geq 0\); ii) \(\phi \geq \bar{\phi}(A)\) if and only if \(\Delta(A, \phi, u_L(A, \phi)) \leq \Delta(A, \bar{\phi}, u_L(A, \bar{\phi}))\). We can now prove the second part of Proposition 2 by verifying the following claim.

**Claim 2** If \(\hat{s}(A) \leq s \leq \bar{s}(A)\) then the optimal policy \(\phi(A, s)\) is implicitly defined by the unique solution to
\[
\Psi(A, \phi, u_L(A, \phi)) = s. \tag{81}
\]

Given \(\phi(A, s)\), the optimal policy \(u(A, s)\) is given by
\[
u(A, s) = u_L(A, \phi(A, s)). \tag{82}
\]

**Proof.** First, by construction, \(\Psi(A, \phi, u_L(A, \phi)) = s\) and constraint (26) holds with equality. Also, the assumption that \(s \leq \hat{s}(A)\) implies \(\phi \geq \hat{\phi}(A)\) and \(\Delta(A, \phi, u_L(A, \phi)) \geq 0\); therefore, (26) is satisfied. Finally, condition (82) implies \(\operatorname{MPK}(A, s) = \operatorname{MPK}_L(A, \phi(A, s))\) and \(\xi_L(A, s) = 0\); therefore, the Kuhn-Tucker conditions (30) and (31) are satisfied.

**Both constraints bind:** In the case where both constraints are binding, from \(\Delta = 0\), we can express \(\operatorname{MPK}\) as a function of \((A, \phi, u)\):
\[
\operatorname{MPK}(A, \phi, u) = (1 - \theta) \operatorname{MPK}_L(A, \phi, u) \frac{\phi^{1-\xi} \phi^\xi - 1}{\phi - 1} - \theta (1 - \delta).
\]

The optimality condition for capital utilization implies
\[
(1 - \theta) \operatorname{MPK}_L(A, \phi, u) \frac{\phi^{1-\xi} \phi^\xi - 1}{\phi - 1} - \theta (1 - \delta) = b_0 (1 - u)^{\nu - 1}, \tag{83}
\]
Equation (83) defines \(u\) as a function of \((A, \phi)\), which we will denote as \(u_{HL}(A, \phi)\).

The fact that the constraint for \(H\) type island binds implies
\[
\Psi(A, \phi, u_{HL}(A, \phi)) = s.
\]
Part three of Proposition 2 can therefore be proved as the result of the following claim.

**Claim 3**  For \( s > \bar{s}(A) \), then the optimal policy \( \{ \phi(A, s), u(A, s) \} \) are jointly determined by:

\[
\Psi(A, \phi, u_{HL}(A, \phi)) = s, \quad u(A, s) = u_{HL}(A, \phi(A, s)).
\] (84)

**Proof.** Clearly, by construction, both constrains (26) and (27) hold with equality. Also, we can show that \( u(A, s) < u_L(A, \phi(A, s)) \); therefore, \( MPK(u) < MPK_j(A, \phi(A, s), u) \), with \( u = u_{HL}(A, \phi(A, s)) \) for \( j = H, L \). As a result, the Kuhn-Tucker conditions (30) and (31) are satisfied with and \( \zeta_j(A, \phi, u_{HL}(A, \phi)) > 0 \) for \( j = H, L \).  

D  Recursive policy function iteration

In this section, we describe an operator that maps the space of equilibrium functionals into itself such that if a fixed point for the operator exists, it constitutes a Markov equilibrium described in Proposition 3 of the paper. Although the construction of the operator may not be unique, our procedure is aimed toward numerical efficiency, because it leads naturally to a recursive approach to compute the equilibrium functionals.

**D.1 The iteration procedure**

We denote \( x = (A, \Lambda) \) as the exogenous state variables and \( \omega = (x, s) \) as the vector of state variables for the Markov equilibrium. First, we observe that Proposition 2 allows us to determine the policy functions \( \phi(\omega) \) and \( u(\omega) \) without any iteration. Second, given an initial guess of next period consumption, \( c(\omega) \) and the value of bank net worth, \( \mu(\omega) \), we can use the intertemporal Euler Equation (50) to determine the current period consumption and investment policies and use the envelop condition (52) to determine the current period value of bank net worth. At the same time, we need to verify that the policy functions and the law of motion of the state variable, Equation (47) are consistent with each other. Because both Equations (50) and (52) are discounting relationships, it is reasonable to expect that if we iterate this procedure, the policy functions, \( c(\omega) \) and \( \mu(\omega) \) will converge. Below are the details.

1. Using Proposition 2 to construct the policy functions \( \phi(A, s) \) and \( u(A, s) \).

   Note that the discussion in Appendix C implies that the policy functions for \( \phi \) and \( u \) are only functions of \( (A, s) \) and can be computed independently of the iterations.

2. Starting from an initial guess of the equilibrium functionals \( \{c^0(\omega), \mu^0(\omega)\} \).
3. Given a set of equilibrium functionals, \( \{ c^n(\omega), \mu^n(\omega) \} \), we use the equilibrium conditions (48), (50), (52), and (53) to solve the optimal investment policy \( i(x, s) \) and the endogenous law of motion of the state variable \( s' = \Gamma(s, x') \). Update the equilibrium functional using the above policy function to compute \( \{ c^{n+1}(\omega), \mu^{n+1}(\omega) \} \). We provide the details of the step in the next section.

4. Iterate on step 3 until the error is smaller than a preset convergence criteria, \( \varepsilon \), i.e.,
\[
\sup_x |c^{n+1}(\omega) - c^n(\omega)| + \sup_x |\mu^{n+1}(\omega) - \mu^n(\omega)| < \varepsilon.
\]

The advantage of our approach is that it makes full use of the first order optimality conditions to improve numerical efficiency. In addition, thanks to the simplification of Proposition 2, the dependence of policy functions on the occasionally binding limited enforcement constraints is fully determined before any iteration.

D.2 Updating policy functions

The first order and envelope conditions Given the policy functions \( \phi(A, s) \) and \( u(A, s) \), we can represent the prices for capital and the Lagrangian multipliers as functions of state variables \( (A, s) \), by using Equations (6), (18), (19), (26), (27), (28) and (29) and the market clearing condition on the reallocation market implies \( Q(A, s) = MPK(A, s) + 1 - \delta \). With a slight abuse of notation, we denote these pricing functionals as \( \{ MPK_j(A, s) \}_{j=H,L}, \{ Q_j(A, s) \}_{j=H,L}, Q(A, s) \) and \( \{ \zeta_j(A, s) \}_{j=H,L} \). Using the above pricing functionals, we can combine the first order condition (50) and the envelope condition (52) as
\[
E \left[ \tilde{M}' \left\{ 1 + \zeta_H(A', s') + \zeta_L(A', s') \right\} \right] R_f(\omega) = E \left[ \tilde{M}' \left\{ 1 + [\zeta_H(A', s') + \zeta_L(A', s')] \right\} Q(A', s') \right],
\]
where \( \tilde{M} \) is defined in Equation (51).

We start with an initial guess of policy functions of normalized consumption and marginal value of net worth, \( c^n(x, s) \) and \( \mu^n(x, s) \) and construct the SDF using \( c^n(x, s) \) as the next period consumption policy:
\[
M(\omega, \omega') = \beta \frac{C(\omega)}{C(\omega')} = \beta \frac{c^n(x, s) K}{c^n(x', s') K'} = \frac{\beta [m(A, s) - i]}{c^n(x', s') [1 - \delta(A, s) + i]},
\]
where we denote
\[
m(A, s) = Au^\alpha(A, s) f(\phi(A, s)),
\]
\[
1 - \delta(A, s) = g(1 - u(A, s)) + (1 - \delta) u(A, s).
\]
Because the risk-free interest rate $R_f (\omega)$ satisfies Equation (48), we have:

$$R_f (\omega) = \frac{1}{E \left[ \frac{\beta[m(A,s)-i]}{c^n(x',s')[1-\delta(A,s)+i]} \right]}.$$ (86)

Assuming that $\mu^n (x, s)$ is the marginal value of net worth in the next period, the left-hand side of Equation (85) can now be written as:

$$E \left[ \frac{\beta[m(A,s)-i]}{c^n(x',s')[1-\delta(A,s)+i]} \right] \left\{ (1 - \Lambda') + N' \mu^n (x', s') \right\} \left\{ 1 + \zeta_H (A', s') + \zeta_L (A', s') \right\}$$

Similarly, the RHS of Equation (85) is

$$E \left[ \frac{\beta[m(A,s)-i]}{c^n(x',s')[1-\delta(A,s)+i]} \right] \left\{ 1 + \zeta_H (A', s') + \zeta_L (A', s') \right\}$$

Therefore, Equation (85) can be written as:

$$E \left[ \frac{\beta[m(A,s)-i]}{c^n(x',s')[1-\delta(A,s)+i]} \right] \left\{ (1 - \Lambda') + N' \mu^n (x', s') \right\} \left\{ 1 + \zeta_H (A', s') + \zeta_L (A', s') \right\} Q (A', s')$$

or

$$m (A, s) - i \frac{\beta[m(A,s)-i]}{[1-\delta(A,s)+i]} = \beta E \left[ \frac{\{(1-\Lambda') + N' \mu^n (x', s')\}}{c^n(x',s')} \right] \left\{ 1 + \zeta_H (A', s') + \zeta_L (A', s') \right\} Q (A', s') \times E \left[ \frac{1}{c^n(x',s')} \right].$$ (88)

**The law of motion of state variables** Let $N$ denote the aggregate net worth of the banking sector in the current period. Because only a fraction $\Lambda$ of the banks survive to the next period,
total bank net worth in the period, \( N' = N' \{ \pi N_H' + (1 - \pi) N_L' \} \). Using Equations (9),

\[
\pi N_H' + (1 - \pi) N_L' = \pi \{ Q_H (\omega') [K' + RA_H (\omega')] - Q(\omega')RA_H (\omega') - R_F (\omega)B_f \} \\
+ (1 - \pi) \{ Q_L (\omega') [K' + RA_L (\omega')] - Q(\omega')RA_L (\omega') - R_F (v)B_f \} \\
= \pi Q_H (\omega') [K' + RA_H (\omega')] + (1 - \pi) Q_L (\omega') [K' + RA_L (\omega')] \\
- Q(\omega') [\pi RA_H (\omega') + (1 - \pi) RA_L (\omega')] - R_F (\omega)B_f
\] (89)

Using \( K'_H = K' + RA_H (\omega') \) and \( K'_L = K' + RA_L (\omega') \), we write the first two terms of the above equation as

\[
\pi Q_H (\omega') K'_H + (1 - \pi) Q_L (\omega') K'_L \\
= \pi M PK_H (\omega') K'_H + (1 - \pi) M PK_L (\omega') K'_L + (1 - \delta) [\pi K'_H + (1 - \pi) K'_L].
\] (90)

Using the fact that capital share is \( \alpha \), we derive \( \pi M PK_H (\omega') K'_H + (1 - \pi) M PK_L (\omega') K'_L = \alpha Y' \). Together with \( \pi K'_H + (1 - \pi) K'_L = u'K' \), Equation (90) can be written as:

\[
\pi Q_H (\omega') K'_H + (1 - \pi) Q_L (\omega') K'_L = \alpha Y' + (1 - \delta) u'K'.
\]

The resource constraint (13) implies \( \pi RA_H (\omega') + (1 - \pi) RA_L (\omega') = (u' - 1) K' \). We can combine the first three terms in Equation (89) as:

\[
\pi Q_H (\omega') K'_H + (1 - \pi) Q_L (\omega') K'_L - Q(Z') [\pi RA_H (\omega') + (1 - \pi) RA_L (\omega')] \\
= \alpha Y' + (1 - \delta) u'K' - Q(\omega') (u' - 1) K' \\
= \alpha Y' + (1 - \delta) u'K' + (1 - \delta) (1 - u') K'.
\] (91)

Because

\[
-Q(z') (u' - 1) K' = [M PK (\omega') + 1 - \delta] (1 - u') K' \\
= M PK (\omega') (1 - u') K' + (1 - \delta) (1 - u') K',
\]

Equation (91) can now be simplified to

\[
\pi Q_H (\omega') K'_H + (1 - \pi) Q_L (\omega') K'_L - Q(\omega') [\pi RA_H (\omega') + (1 - \pi) RA_L (\omega')] \\
= \alpha Y' + (1 - \delta) M PK (\omega') K' + (1 - \delta) K'.
\]

Therefore, Equation (89) can be simplified as:

\[
\pi N_H' + (1 - \pi) N_L' = \alpha Y' + (1 - u') M PK (\omega') K' + (1 - \delta) K' - R_f (\omega) B_f.
\]
We have:

\[ N' = \Lambda [\alpha Y' + (1 - u') MPK(\omega') K' + (1 - \delta) K' - R_f(\omega) B_f]. \]  

(92)

Banks’ budget constraint (8) implies

\[ B_f = K' - N. \]

Therefore, by the definition of \( s' \),

\[ s' = R_f(\omega) \frac{B_f}{K'} = R_f(\omega) \left(1 - \frac{N}{K'}\right) = R_f(\omega) \left(\frac{K'}{K} - \frac{N}{K}\right). \]  

(93)

Using Equation (92),

\[ N = \Lambda [\alpha Y + (1 - u) MPK(\omega) K + (1 - \delta) K - R_f(\omega_{-1}) B_{f,-1}], \]

we can express all the terms on the right-hand side of the above equation as functions of the state variables:

\[ Y = m(A, s) K, \quad u = u(A, s), \quad R_f(\omega_{-1}) B_{f,-1} = sK. \]

Therefore, \( n = \frac{N}{K} \) can be written as a function of the state variable \((A, s)\):

\[ \frac{N}{K} = \Lambda [\alpha m(A, s) + (1 - u(A, s)) MPK(A, s) + (1 - \delta) - s] \]  

(94)

Now we can combine Equations (93) and (94), and use equation (86) to derive the law of motion of the state variable \( s \):

\[ s' = \left\{ [1 - \delta (A, s) + \beta] - \Lambda [\alpha m(A, s) + (1 - u(A, s)) MPK(A, s) + (1 - \delta) - s] \right\}, \]

\[ E \left[ \frac{\beta [m(A, s) - \beta]}{e^{\gamma(x')\beta(s) + \beta}} \right] [1 - \delta (A, s) + \beta]. \]

or

\[ s' = \left\{ [1 - \delta (A, s) + \beta] - \Lambda [\alpha m(A, s) + (1 - u(A, s)) MPK(A, s) + (1 - \delta) - s] \right\} \]

\[ \beta \left[ m(A, s) - \beta \right] E \left[ \frac{1}{e^{\gamma(x')\beta(s) + \beta}} \right] \]  

(95)

Given the policy functions \( \{c^a(x, s), \mu^a(x, s)\} \), Equations (88) and (95) are two nonlinear equations with two unknowns, \( i \) and \( s' \). For each point in the state space, \((x, s)\), we solve these two equations for the policy function \( i(x, s) \) and the law of motion of next period \( s' \) as
a function of \((s, x)\). We update the policy functions for the next iteration by setting
\[
c_{n+1}(x, s) = m(A, s) - i(x, s),
\]
\[
\mu^{n+1}(x, s) = E \left[ \frac{\{(1 - \Lambda') + \Lambda'\mu^n(x', s')\}}{c^n(x', s')} \{1 + \zeta_H(A', s') + \zeta_L(A', s')\} \right] / E \left[ \frac{1}{c^n(x', s')} \right].
\]

**E Power law in firm size**

Here we derive the slope of the power law distribution of firm size implied by our model.

Recall that distribution of random variable \(z\) obeys a power law if its cumulative distribution function of \(z\) is given by:
\[
P(z > X) \propto X^{-\zeta},
\]
for some constant \(\zeta > 0\). The parameter \(\zeta\) is called the power-law exponent. That is, the cumulative distribution of a power-law variable is log linear with the slope \(-\zeta\).

The assumption that bankers’ net worth moves freely across islands at the end of every period implies that \(\frac{N_{j,t}}{z_{j,t}}\) and \(\frac{K_{j,t+1}}{z_{j,t}}\) must be equalized across islands. In another words, firm size \(K_{j,t+1}\) is proportional to idiosyncratic productivity shock \(z_{j,t}\). Here, we provide a computation for the power law of the distribution of \(z_{j,t}\). In our model, \(z_{j,t+1} = z_{j,t} e^{\varepsilon_{j,t+1}}\), and \(\varepsilon_{j,t+1}\) is i.i.d. across firms and over time with two possible realizations: \(\text{Prob}(\varepsilon = \varepsilon_H) = 1\) – \(\text{Prob}(\varepsilon = \varepsilon_L) = \pi\) and a normalization condition \(\pi e^{\varepsilon_H} + (1 - \pi) e^{\varepsilon_L} = 1\). To save notation, we suppress the firm index \(j\) and define \(\tilde{z} = \ln(z)\), so that the law of motion of \(z\) can be equivalently characterized as \(\tilde{z}_{t+1} = \tilde{z}_t + \varepsilon_{t+1}\).

According to firms’ exit and entry as described in Section 3.3, existing firms die with probability \(1 - \Lambda\) after production and start over with the initial condition \(z_{j,t+1} = 1\). Let \(\phi(\tilde{z})\) be the density of \(\tilde{z}\). For \(\tilde{z}\) large enough, we have
\[
\phi_{t+1}(\tilde{z}) = \Lambda \left[ \pi \phi_t(\tilde{z} - \varepsilon^H) + (1 - \pi) \phi_t(\tilde{z} - \varepsilon^L) \right].
\]

Assume that \(\phi(\tilde{z}) = k e^{-\zeta \tilde{z}}\) for large \(\tilde{z}\), the above equation implies that the stationary distribution must satisfy:
\[
k e^{-\zeta \tilde{z}} = \Lambda \left[ \pi k e^{-\zeta(\tilde{z} - \varepsilon^H)} + (1 - \pi) k e^{-\zeta(\tilde{z} - \varepsilon^L)} \right].
\]

Therefore, \(\zeta\) must be a solution to the following equation:
\[
1 = \Lambda \left[ \pi e^{\varepsilon^H} + (1 - \pi) e^{\varepsilon^L} \right].
\]
Then $\Pr (z > X) = \Pr (\tilde{z} > \ln X)$ can be computed as:

$$
\int_{\ln X}^{\infty} \phi (\tilde{z}) d\tilde{z} = \int_{\ln X}^{\infty} ke^{-\zeta \tilde{z}} d\tilde{z} \\
= \frac{1}{\zeta} X^{-\zeta} \propto X^{-\zeta}
$$

Therefore, the parameter $\zeta$ is the power-law slope of firm size distribution.

References


Dou, W. W. 2017. Embrace or fear uncertainty: Growth options, limited risk sharing, and asset pricies.


