Asset Pricing with Misallocation

Winston Wei Dou Yan Ji Di Tian Pengfei Wang*

March 14, 2023

Abstract

We develop an endogenous growth model with heterogeneous firms facing financial frictions, in which misallocation emerges explicitly as a crucial state variable. In equilibrium, misallocation endogenously generates long-run uncertainty about economic growth by distorting innovation decisions, leading to significant welfare losses and risk premia in capital markets. Macroeconomic shocks that affect misallocation are likely to have overly persistent effects on aggregate growth. Using an empirical misallocation measure motivated by the model, we find evidence showing that misallocation captures low-frequency variations in both aggregate growth and stock returns. Empirically, a two-factor model with market and misallocation factors prices size, book-to-market, momentum, and bond portfolios with an R-squared and a mean absolute pricing error close to the Fama-French three-factor model.

Keywords: Agency conflicts, Distribution of firms, Financial frictions, Misallocation, Endogenous growth, Macro finance.

JEL Classification: L11, O30, O40.

*Dou: University of Pennsylvania (wdou@wharton.upenn.edu). Ji: Hong Kong University of Science and Technology (jiy@ust.hk). Tian: University of Pennsylvania (tiandi@sas.upenn.edu). Wang: Peking University (pfwang@phbs.pku.edu.cn). This paper was previously circulated under the title "Granular Uncertainty and Disastrous Misallocation." We are grateful for the helpful comments from Hengjie Ai, Jinhui Bai, Paco Buera, Zhanhui Chen, Nuno Clara, Joel M. David, Andrea Eisfeldt, Lorenzo Garlappi, Vincent Glode, Joao Gomes, Andrei S. Gonçalves, Lars Peter Hansen, Tony He, Ken Judd, Patrick Konermann, Felix Kubler, Hong Liu, Laura Xiaolei Liu, Yang Liu, Ding Luo, Emilio Osambela, Vincenzo Quadrini, Adriano Rampini, Michael Roberts, Nick Roussanov, Tom Sargent, Lukas Schmid, Gill Segal, Dejanir Silva, Rob Stambaugh, Rob Townsend, Neng Wang, Tom Winberry, Haotian Xiang, Wei Xiong, Bo Zhao, Lichen Zhang, and Xiaodong Zhu as well as the participants in the Macroeconomy and Finance in China Conference, PKU/PHBS Sargent Institute Workshops, HKUST-JINAN Workshop in Macroeconomics, Stanford SITE (New Frontier in Asset Pricing), Tsinghua SEM Alumni Conference, Singapore Management University, Shanghai University of Finance and Economics, Purdue University, Peking University (Guanghua), AFA, MFA, FIRS, SFS Cavalcade North America, NFA, CICM, CICF, AsianFA, SAFE Asset Pricing Workshop, SWUFE International Macro-Finance Conference, CUHK Greater Bay Area Finance Conference, HKUST, and University of Pennsylvania (Wharton). Wang acknowledges the financial support from China National Science Foundation (#72125007).

1 Introduction

In the last decade, one of the most import ant developments in the growth literature is the enhanced appreciation of the role of misallocation in helping us understand economic growth (Jones, 2013). The link between misallocation and growth prospects can potentially shed light on the fundamental forces that drive the fluctuation of economic growth in the long run, a mechanism that quantitatively rationalizes many asset pricing moments (e.g., Bansal and Yaron, 2004; Hansen, Heaton and Li, 2008), and that quantitatively justifies large welfare costs of business cycles owing to endogenous fluctuations in consumption growth via capital accumulation (e.g., Barlevy, 2004; Eisfeldt and Rampini, 2006, 2008*b*).

This paper develops a general-equilibrium model with analytical tractability to quantitatively investigate the connection between misallocation and systematic risk that shapes asset prices in capital markets. In our model, a low-frequency component of economic growth emerges endogenously due to slow-moving misallocation as a primitive source of systematic risk faced by investors, shedding novel insights on the asset pricing implications of misallocation.

Specifically, economic growth is driven by endogenous technological advances through the invention of intermediate goods as in standard endogenous growth models (e.g., Romer, 1987, 1990; Jones, 1995). Final goods are produced by heterogeneous firms facing two financial frictions arising from agency conflicts — an equity market constraint for payout and issuance and a collateral constraint for debt.[1] These two financial frictions prevent capital from being reallocated across firms. The misallocation of capital among firms of different productivity emerges analytically as a crucial endogenous state variable, which characterizes the evolution of the economy. In equilibrium, short-run (even white-noise) macroeconomic shocks can generate persistent shifts in demand for research and development (R&D) through their long-lasting effects on misallocation, which can in turn lead to overly persistent fluctuations in economic growth. In other words, our model suggests a novel channel through which business cycle fluctuations are endogenously associated overly persistent fluctuations in consumption growth. Consequently, when investors prefer early resolution of uncertainty and the intertemporal substitution effect dominates according to their preferences, the

[1] More specifically, the equity market constraint for payout and issuance is same as in Myers (2000) and Lambrecht and Myers (2008, 2012), and the collateral constraint for debt is the same as in Buera and Shin (2013) and Moll (2014).

endogenous low-frequency component of economic growth driven by slow-moving misallocation can have first-order asset pricing implications in capital markets and lead to substantial welfare costs of business cycles.

Our paper contributes to the existing literature in four ways. First, we show that misallocation drives low-frequency movements in R&D intensity and thus economic growth in both the model and data. A covariance-type misallocation measure endogenously arises as a sufficient statistic that captures the general-equilibrium effect of the multivariate cross-sectional distribution of heterogeneous firms in the model. Shocks that impact an economy's misallocation can have persistent effects on the economy's growth rate through R&D decisions, providing a misallocation-based explanation for the observed low-frequency covariation in the time series of consumption growth and output growth (e.g., Bansal, Dittmar and Lundblad, 2005; Hansen, Heaton and Li, 2008; Müller and Watson, 2008, 2018). In the data, we find empirical evidence for misallocation-driven low-frequency movements in time series of aggregate growth rates and stock returns.

Second, we show that, as a macroeconomic factor, the misallocation measure motivated by our model has significant cross-sectional asset pricing implications. A two-factor model with market and misallocation factors prices size, book-to-market, momentum, and bond portfolios with an R-squared of 53% and a mean absolute pricing error (MAPE) of 1.82, which are close to those implied by the Fama-French three-factor model with an R-squared of 62% and a MAPE of 1.90. Importantly, future accumulated consumption growth, as a proxy for the low-frequency component of consumption growth that is shown to contain important asset pricing information, becomes of little importance in explaining asset returns once our misallocation measure is taken into account as a factor. We emphasize that the strong pricing power of misallocation factor, as a (macro) nontradable asset pricing factor, is an important, nontrivial empirical finding. As emphasized by Cochrane (2017), it is the sole job of macro-finance to understand what are the primitive sources of systematic risk, by suggesting (macro) nontradable factors, and explain why they earn a premium.[2] However, not many studies find that (macro) nontradable factors motivated by macro-finance models empirically outperform or drive out (ad hoc) tradable factors such as Fama-French factors in explaining the cross section of expected asset returns,[3] partly (not totally)

[2]Other recent reviews on macro-finance models also highlight this point (e.g., Brunnermeier, Eisenbach and Sannikov, 2012; Dou et al., 2020*a*).

[3]A few exceptions include durable consumption growth (Yogo, 2006; Gomes, Kogan and Yogo,

because the measurement error in the nontradable factors causes attenuation bias in the estimates of factor exposures.

Third, our model delineates the tight link between firms' idiosyncratic productivity shocks and the low-frequency aggregate consumption risk. When firms' idiosyncratic productivity is more persistent, the economy's misallocation, which determines the aggregate total factor productivity (TFP) and output, also becomes more persistent. Consequently, this generates more persistent variations in aggregate consumption growth in response to macroeconomic shocks. By connecting the persistence in idiosyncratic productivity with the persistence in aggregate consumption growth, our model implies that long-run risk in aggregate consumption can be estimated based on a vast panel of granular firm-level data, which helps address the issues of weak identification in the long-run risk literature (Chen, Dou and Kogan, 2022; Cheng, Dou and Liao, 2022). This highlights the important role of misallocation measures in estimating the empirical stochastic discount factor (SDF).

Fourth, this paper advances the insight of Eisfeldt and Rampini (2006, 2008*b*) by quantitatively showing that misallocation plays an essential role in generating the welfare costs of business cycles through the endogenous misallocation-driven fluctuations of aggregate consumption growth in the long run, which is extremely costly in the sense that investors are willing to pay a sizeable premium to eliminate such long-run uncertainty about economic growth.

We now elaborate more on the ingredients of our model. There are three sectors in our model economy. The innovation sector uses final goods and existing stock of knowledge to produce new knowledge, which are blueprints for new intermediate goods. An intermediate goods sector uses the designs from the innovation sector together with final goods to produce differentiated goods, which are intermediate goods for final goods production. The final goods sector uses capital, labor, and intermediate goods to produce final goods. There exists a representative household that owns firms in all sectors, a continuum of heterogeneous firms in the final goods sector, and homogeneous firms in intermediate goods and innovation sectors.

Firms in the final goods sector are heterogeneous in productivity and capital. Production takes place using capital, labor, and intermediate goods. Because of agency conflicts, firms face an equity market constraint for payout and issuance and a collateral

2009), expenditure shares of housing (Piazzesi, Schneider and Tuzel, 2007), market liquidity (Pástor and Stambaugh, 2003), intermediary leverages (Adrian, Etula and Muir, 2014; He, Kelly and Manela, 2017), and common fund flows (Dou, Kogan and Wu, 2021), among others.

constraint for debt. The collateral constraint generates capital misallocation among firms as in Buera and Shin (2013) and Moll (2014). A higher misallocation results in a lower productivity in the final goods sector, which reduces the aggregate demand for intermediate goods. This, in turn, motivates innovators to invent new intermediate goods less intensively, leading to a lower growth rate.

Firms endogenously choose their capacity utilization intensity. A higher capacity utilization intensity allows firms to produce more outputs at the cost of bearing a higher depreciation rate of capital. There are aggregate capital depreciation shocks, as in Gourio (2012), Brunnermeier and Sannikov (2017), etc. In equilibrium, because more productive firms use their capital more intensively, aggregate capital depreciation shocks generate endogenous fluctuations in the economy's misallocation.

We show that the misallocation in the final goods sector emerges as an endogenous state variable. Specifically, by applying the Berry-Esseen bound (Tikhomirov, 1980; Bentkus, Gotze and Tikhomoirov, 1997), the capital share of firms of different productivity can be approximated by a log-normal distribution. This parametric functional form implies that the distribution of firms in the cross section is fully summarized by a single endogenous state variable, capturing the covariance between log capital and log productivity across firms. This covariance-type state variable determines the economy's misallocation, based on which both the steady state and transitional dynamics can be characterized in closed form. We show that a calibrated model can quantitatively reproduce the low-frequency components in aggregate consumption growth and the high Sharpe ratio of equity returns as in the data. Short-run i.i.d. shocks can generate persistent effects on the economy's growth because the endogenous misallocation is slow moving. Importantly, the persistence in misallocation largely depends on the persistence in firms' idiosyncratic productivity.

While our main contribution is theoretical, we also empirically test the main predictions of our model. Motivated by the model, we construct a misallocation measure based on the covariance between log productivity and log capital using the U.S. Compustat data. We find evidence that an increase in misallocation predicts declines in R&D intensity and lower growth of aggregate consumption and output over long horizons. In the cross section, we find that the cash flows of value firms load more negatively on misallocation than the cash flows of growth firms. This is consistent with the robust evidence found in the asset pricing literature that the cash flows of value firms load more positively on accumulated consumption growth than those of growth firms (Bansal, Dittmar and Lundblad, 2005; Parker and Julliard, 2005; Hansen, Heaton and Li, 2008; Santos and Veronesi, 2010), providing empirical support for our model prediction that fluctuations in misallocation can drive the low-frequency variations in consumption growth. Further, we show that a two-factor model with market and misallocation factors prices size, book-to-market, momentum, and bond portfolios with an R-squared and a MAPE close to the Fama-French three-factor model. Future accumulated consumption growth has little explanatory power for portfolio returns once our misallocation measure is included as a factor, suggesting that long-run consumption growth affects asset returns through the persistent variation in misallocation. Finally, we provide direct evidence for the core mechanism of our model, which implies that misallocation drives long-run growth through its impact on R&D. We consider the policy shock from the American Jobs Creation Act (AJCA) passed in 2004, which presumably relaxes the financial constraints of firms with pre-tax income from abroad. By exploiting industries' differential exposure to this policy shock in a difference-in-differences (DID) setting, we find that AJCA results in significantly lower industry-level misallocation and higher R&D expenditure in treated industries. Moreover, the impact of AJCA on industry-level R&D expenditure becomes statistically insignificant after controlling for industry-level misallocation.

Related Literature. Our paper is related to three strands of literature. First, we contribute to the long-run risk literature in finance (e.g., Bansal and Yaron, 2004). Various studies try to justify long-run risk with micro foundations (e.g., Ai, 2010; Kaltenbrunner and Lochstoer, 2010; Garleanu, Panageas and Yu, 2012; Kung and Schmid, 2015; Collin-Dufresne, Johannes and Lochstoer, 2016; Ai, Li and Yang, 2020; Gârleanu and Panageas, 2020; Croce, Nguyen and Raymond, 2021). Our paper is mostly related to Kung and Schmid (2015) who show that R&D endogenously drives a small, persistent component in productivity, which generates long-run uncertainty about economic growth. Building on the theoretical framework of Kung and Schmid (2015), we introduce heterogeneous firms to the final goods sector to generate endogenous misallocation as in Moll (2014). The aggregate TFP, which is exogenous in the model of Kung and Schmid (2015), is endogenously in our model, determined by the cross-sectional misallocation. Our theory rationalizes long-run consumption risk through the equilibrium interactions between endogenous misallocation and R&D incentives, which is also supported by the data. Importantly, by connecting the persistence in idiosyncratic productivity with the persistence in aggregate consumption growth, our model implies that long-run risk in aggregate consumption can be estimated based on granular firm-level data, which potentially helps address the issues of weak identification in the long-run risk literature (Chen, Dou and Kogan, 2022; Cheng, Dou and Liao, 2022).

Second, our paper contributes to the large and growing macroeconomics literature that emphasizes the role of misallocation in economic development (e.g., Banerjee and Duflo, 2005; Foster, Haltiwanger and Syverson, 2008; Restuccia and Rogerson, 2008; Buera and Shin, 2011; Jones, 2011; Buera and Shin, 2013; Jones, 2013; Midrigan and Xu, 2014; Moll, 2014; Acemoglu et al., 2018; Peters, 2020), and Hsieh and Klenow (2009); Bartelsman, Haltiwanger and Scarpetta (2013). On the technical side, our model extends the tractable framework of Moll (2014) with intermediate inputs, R&D, and aggregate shocks to generate endogenous stochastic growth. Our paper provides the following insights to this literature. First, our model implies that misallocation in production inputs of final goods producers can affect equilibrium growth because it determines the profits of producing intermediate goods and thus innovators' R&D incentives. Second, misallocation emerges naturally as an endogenous state variable in our model, which motivates an intuitive empirical misallocation measure based on the covariance between firms' log productivity and log capital. Third, our model implies that when idiosyncratic productivity is persistent, investors demand high risk premia because the slow-moving misallocation incubates long-run consumption risk. Our results thus complement the key insight of Moll (2014) who shows that misallocation is less severe in the long-run steady state (without aggregate shocks) when idiosyncratic productivity is more persistent.

Relative to the macroeconomics literature, much less work focuses on the role of misallocation in the corporate finance and asset pricing literature. Important contributions include, for example, Eisfeldt and Rampini (2006, 2008*b*), Opp, Parlour and Walden (2014), Fuchs, Green and Papanikolaou (2016), Ehouarne, Kuehn and Schreindorfer (2017), van Binsbergen and Opp (2019), David, Schmid and Zeke (2019), Ai, Li and Yang (2020), Ai et al. (2020), Lanteri and Rampini (2021), Whited and Zhao (2021).[4] Notably, David, Schmid and Zeke (2019) propose a theory that links misallocation with macroeconomic risk. They show that risk considerations can explain a large proportion of the dispersion in marginal products of capital among U.S. firms,

[4]See Eisfeldt and Shi (2018) for a comprehensive survey.

which suggests that much of the observed dispersion is attributed to efficient sources. Similar to David, Schmid and Zeke (2019), firm decisions in our model are also influenced by their risk exposure, and thus the cross-sectional misallocation is partly driven by firms' risk considerations. The additional insight of our paper is to show that the economy's misallocation itself emerges as a macroeconomic risk factor because the evolution of misallocation determines the evolution of aggregate consumption growth, a primitive determinant of risk premia. In other words, a feedback loop emerges in general equilibrium: an economy's misallocation determines households' SDF and thus affects firm decisions, which in turn determine the misallocation across firms.

Third, our paper is related to the literature on business cycles (e.g., Lucas, 1987). Barlevy (2004) shows that the welfare cost of business cycles is large when fluctuations affect the growth rate of consumption in a model with diminishing returns in investment. The strong procyclical patterns of capital reallocation documented by Eisfeldt and Rampini (2006, 2008*b*) suggest that misallocation can play an important role in determining the welfare costs of business cycles. Through persistent misallocation, our model rationalizes long-run consumption risk and generates a high Sharpe ratio for equity returns, which reflects investors' aversion to aggregate risks. As a result, our model quantifies a large welfare cost of business cycles following the approach of Alvarez and Jermann (2004, 2005).

The outline of the paper is as follows. Section 2 develops a model to depict the equilibrium relation between misallocation and growth. Section 3 calibrates the model to evaluate its quantitative implications. Section 4 provides empirical evidence to support the model's main mechanisms and predictions. Section 5 concludes.

2 Model

There are three sectors: a final goods sector, an intermediate goods sector, and an R&D sector. The R&D sector invents new knowledge (i.e., blueprints for new varieties of intermediate goods) using final goods and existing stock of knowledge, then sells blueprints to the intermediate goods sector. The intermediate goods sector produces differentiated intermediate goods using blueprints created by the R&D sector and final goods, then sells intermediate goods to the final goods sector. The final goods sector uses capital, labor, and intermediate goods to produce final goods. There is a representative household that owns firms in all sectors, a continuum of heterogeneous firms in the final goods sector, and homogeneous firms in the intermediate goods and R&D sectors.

2.1 Final Goods Sector

In the final goods sector, there is a continuum of firms of measure one, indexed by $i \in \mathcal{I} \equiv [0,1]$ and operated by managers. Firms are different from each other in their idiosyncratic productivity $z_{i,t}$ and capital $a_{i,t}$. At each point in time t, the distribution of final goods firms is characterized by the joint probability density function (PDF), $\varphi_t(a,z)$.

The firm produces output at intensity $y_{i,t}$ over $[t, t+\mathrm{d}t)$ using a production technology with constant returns to scale:

$$y_{i,t} = \left[(z_{i,t}u_{i,t}k_{i,t})^{\alpha}\ell_{i,t}^{1-\alpha}\right]^{1-\varepsilon} x_{i,t}^{\varepsilon}, \quad \text{with } \alpha,\ \varepsilon \in (0,1), \tag{1}$$

where labor $\ell_{i,t}$ is hired in a competitive labor market at the equilibrium wage w_t. The variable $k_{i,t} = a_{i,t} + \widehat{a}_{i,t}$ is the capital installed in production, which includes the firm's own capital $a_{i,t}$ and the leased capital $\widehat{a}_{i,t}$ borrowed from a competitive rental market at the equilibrium risk-free rate $r_{f,t}$.[5]

The firm's output increases with its idiosyncratic productivity $z_{i,t}$ and endogenous choice of capacity utilization intensity $u_{i,t} \in [0,1]$. Utilizing capital at intensity $u_{i,t}$ leads to an amount of $u_{i,t}k_{i,t}\mathrm{d}\Delta_t$ depreciation over $[t, t+\mathrm{d}t)$, where $\mathrm{d}\Delta_t$ captures the stochastic depreciation rate,

$$\mathrm{d}\Delta_t = \delta_k \mathrm{d}t + \sigma_k \mathrm{d}W_t. \tag{2}$$

The standard Brownian motion W_t captures the aggregate capital depreciation shock similar in spirit to that of Albuquerue and Wang (2008) and Gourio (2012). The parameters $\delta_k, \sigma_k > 0$ capture the constant and stochastic components of capital depreciation. As we show in Lemma 1 below, in equilibrium, more productive firms utilize capital more intensively by optimally choosing larger $u_{i,t}$. As a result, more productive firms are more exposed to the aggregate depreciation shock $\mathrm{d}W_t$ than less productive firms. The only role of utilization intensity $u_{i,t}$ is to endogenously generate differential exposure to the aggregate shocks in the cross section of firms, so that $\mathrm{d}W_t$

[5]The capital leasing market is relevant for firms' production and financial decisions (e.g., Eisfeldt and Rampini, 2008*a*; Rampini and Viswanathan, 2013; Li and Tsou, 2021; Li and Xu, 2021).

will have an impact on the economy's misallocation. The same results can be obtained if we exogenously specify the aggregate risk exposure across firms.

We assume that the firm's own capital stock evolves according to

$$\mathrm{d}a_{i,t} = -\delta_a a_{i,t}\mathrm{d}t + \sigma_a a_{i,t}\mathrm{d}W_t + \mathrm{d}I_{i,t}, \tag{3}$$

where $\delta_a > 0$ is the constant depreciation rate, and $\sigma_a \mathrm{d}W_t$ captures the capital efficiency shock with $\sigma_a > 0$. The modeling of capital efficiency shocks has been widely adopted in the literature.[6] We assume that a single aggregate shock enters both equations (2) and (3), which implies that improvement in the efficiency of new capital is associated with depreciation of existing capital, capturing the displacement effect of new capital (e.g., Gârleanu, Kogan and Panageas, 2012; Kogan et al., 2017; Kogan, Papanikolaou and Stoffman, 2020). Introducing capital efficiency shock in equation (3) ensures that the aggregate shock $\mathrm{d}W_t$ mainly affects the economy's misallocation without having much effect on the level of aggregate capital stock, because the depreciation of productive capital, captured by the term $\sigma_k \mathrm{d}W_t$ in equation (2), is partially offset by the improved efficiency of new capital, captured by the term $\sigma_a \mathrm{d}W_t$ in equation (3). If we instead eliminate the aggregate shock in equation (3) and only focus on the depreciation shock in equation (2), the time-series variation in both misallocation and aggregate capital will play a significant role in determining the Sharpe ratio of risky assets, which confounds the misallocation channel we emphasize. Under our calibration, the time variation in misallocation plays a determining role to generate a high Sharpe ratio while the time variation in the level of aggregate capital stock does not (see columns (1) and (7) of Table 4).

The variable $\mathrm{d}I_{i,t}$ in equation (3) is the amount of final goods that is converted to capital over $[t, t + \mathrm{d}t)$. Similar to Pástor and Veronesi (2012), we assume that profits are reinvested, so that the firm's investment rate $\mathrm{d}I_{i,t}$ is equal to its profit after paying operation expenses, interests, and dividends (see equation (19) below).

The composite $x_{i,t}$ in equation (1) consists of differentiated intermediate goods, given by the constant elasticity of substitution (CES) aggregation:

$$x_{i,t} = \left(\int_0^{N_t} x_{i,j,t}^{\nu} \mathrm{d}j \right)^{\frac{1}{\nu}}, \tag{4}$$

[6] e.g., Sundaresan (1984), Cox, Ingersoll and Ross (1985), Kogan (2001, 2004), Gourio (2012), Di Tella (2017), and Dou (2017).

where $x_{i,j,t}$ is the quantity of intermediate goods $j \in [0, N_t]$. The elasticity of substitution among differentiated intermediate goods is $1/(1-\nu) > 0$. The economy's stock of knowledge (i.e., the variety of differentiated intermediate goods created based on existing blueprints) at t is N_t. Technological advances through the expansion of N_t drives endogenous growth, as in Romer (1987, 1990) and Jones (1995). Denote by $p_{j,t}$ and p_t the prices of the intermediate good j and the composite of intermediate goods, respectively.

The firm's idiosyncratic productivity $z_{i,t}$ evolves according to

$$\mathrm{d}\ln(z_{i,t}) = -\theta \ln(z_{i,t})\mathrm{d}t + \sigma\sqrt{\theta}\mathrm{d}W_{i,t}, \tag{5}$$

where the standard Brownian motion $W_{i,t}$ captures idiosyncratic shocks to firm i's productivity. The specification of the idiosyncratic process $z_{i,t}$ is similar to that of Moll (2014). The parameter θ determines the persistence of idiosyncratic productivity $z_{i,t}$. A higher θ makes $z_{i,t}$ less persistent, implying that firms face higher uncertainty in their future idiosyncratic productivity. Importantly, a change in θ does not affect the dispersion in idiosyncratic productivity across firms, because θ scales both the drift term and the diffusion term in equation (5).

2.2 Intermediate Goods Sector

There is a continuum of intermediate goods producers, indexed by $j \in [0, N_t]$. They produce intermediate goods using final goods and blueprints created by firms in the R&D sector. Specifically, intermediate goods producer j has monopoly power in setting prices, facing a downward sloping demand for its output. Intermediate good producers buy final goods and transform them to intermediate inputs, based on the blueprints they hold. We assume that one unit of final goods can be transformed into one unit of intermediate goods, meaning that the marginal cost of producing intermediate goods is unity. The producer of intermediate good j solves

$$\max_{p_{j,t}} \pi_{j,t} = p_{j,t}e_{j,t} - e_{j,t}, \tag{6}$$

subject to the demand curve:

$$e_{j,t} = \left(\frac{p_{j,t}}{p_t}\right)^{\frac{1}{\nu-1}} X_t, \tag{7}$$

where $X_t \equiv \int_{i \in \mathcal{I}} x_{i,t} \mathrm{d}i$ is the aggregate demand for the composite of intermediate goods.

The value of a blueprint, denoted by $v_{j,t}$, is the value of owning the exclusive rights to produce intermediate goods j, which is given by the Hamilton-Jacobi-Bellman (HJB) equation:

$$0 = \Lambda_t \left(\pi_{j,t} - \delta_b v_{j,t} \right) \mathrm{d}t + \mathbb{E}_t \left[\mathrm{d}(\Lambda_t v_{j,t}) \right], \tag{8}$$

where Λ_t is the SDF, and δ_b is the hazard rate at which an existing blueprint becomes obsolete. Because of symmetry and homogeneity, all blueprints have identical values, $v_{j,t} \equiv v_t$, and all intermediate good producers make identical flow profits, $\pi_{j,t} \equiv \pi_t$.

2.3 R&D Sector

Intermediate goods producers are competitive and do not make profits in equilibrium. They buy blueprints from innovators at the price $v_{j,t}$. That is, innovators have full bargaining power and seize all the surplus $v_{j,t}$. Thus, $v_{j,t}$ is the value of creating the blueprint for producing the intermediate good j, which shapes the incentive of innovators to create new blueprints.

Blueprints are created by conducting R&D using final goods as in Comin and Gertler (2006). The stock of knowledge N_t evolves as follows:

$$\mathrm{d}N_t = \vartheta_t S_t \mathrm{d}t - \delta_b N_t \mathrm{d}t, \tag{9}$$

where S_t is the aggregate R&D expenditure, and ϑ_t captures the productivity of innovations, which is taken as exogenously given by individual innovators. In equilibrium, the free-entry condition implies that the marginal return of R&D is equal to its marginal cost:

$$v_t \vartheta_t = 1. \tag{10}$$

Following Comin and Gertler (2006) and Kung and Schmid (2015), we specify

$$\vartheta_t = \chi \left(\frac{N_t}{S_t} \right)^h, \tag{11}$$

where $h \in (0, 1)$. Equation (11) implies that there are positive spillovers of the aggregate stock of knowledge (the term N_t^h) as in Romer (1990) and Jones (1995), and that aggregate R&D investment has decreasing marginal returns (the term S_t^{-h}), capturing the congestion effect in developing new blueprints.

2.4 Agents

There is a continuum of households, with workers and managers who consume together. Like in Dou (2017), only managers can manage firms' investments and operations. The managers can be executives, directors, and entrepreneurs; more broadly, they can also be the controlling shareholders who are fully entrenched and have complete control over the firm's investment and payout policies (e.g., Albuquerue and Wang, 2008). Each manager manages a firm in the final goodssector subject to agency problems. Workers lend funds to firms and hold equity claims on all firms. We assume that a full set of Arrow-Debreu securities is available to households, so that idiosyncratic consumption risks can be fully insured and there exists a representative household. The aggregate labor supply is inelastic and normalized to be 1.

Preferences. The representative household has stochastic differential utility as in Duffie and Epstein (1992*a*,*b*):

$$U_0 = \mathbb{E}_0 \left[\int_0^\infty f(C_t, U_t) \mathrm{d}t \right], \tag{12}$$

where

$$f(C_t, U_t) = \left(\frac{1-\gamma}{1-\psi^{-1}} \right) U_t \left[\left(\frac{C_t}{[(1-\gamma)U_t]^{1/(1-\gamma)}} \right)^{1-\psi^{-1}} - \delta \right]. \tag{13}$$

This preference is a continuous-time version of the recursive preferences proposed by Kreps and Porteus (1978), Epstein and Zin (1989), and Weil (1990). The felicity function f is an aggregator over the current consumption rate C_t of final goods and future utility level U_t. The coefficient δ is the subjective discount rate, the parameter ψ is the elasticity of intertemporal substitution (EIS), and the parameter γ captures risk aversion.

The representative household maximizes utility (12) subject to the following budget constraint:

$$\mathrm{d}B_t = \left(w_t L_t + r_{f,t} B_t + D_t - C_t \right) \mathrm{d}t, \tag{14}$$

where $w_t L_t$ is the wage income intensity, with $L_t \equiv 1$, D_t is the dividend intensity of all firms, and B_t is the amount of bonds held by the household at t.

The representative household's SDF is

$$\Lambda_t = \exp\left(\int_0^t f_U(C_s, U_s) ds\right) \rho^{\frac{1-\gamma}{1-\psi^{-1}}} R_t^{\frac{\gamma-\psi^{-1}}{1-\psi^{-1}}} C_t^{-\gamma}, \tag{15}$$

where R_t is the consumption-wealth ratio of the representative household.

Limited Enforcement. An equity market constraint for payout/issuance and a credit market collateral constraint for borrowing endogenously arise from limited enforcement problems of equity and debt contracts.

The manager extracts pecuniary rents $\tau a_{i,t} \mathrm{d}t$ over $[t, t + \mathrm{d}t)$ when running the firm i.[7] These rents represent the cash compensation above the manager's wage (e.g., Myers, 2000; Lambrecht and Myers, 2008, 2012). Shareholders have the option to intervene and take control of the firm by replacing the manager. Intervention is costly because it requires collective action (e.g., Myers, 2000) and can damage the firm's talent-dependent customer capital (e.g., Dou et al., 2020*b*). In particular, we assume that a fraction τ/ρ of capital $a_{i,t}$ is lost upon intervention with $\tau < \rho$, after which shareholders will become the new manager of the firm. In equilibrium, the manager will pay dividend up to the point where shareholders would have no incentive to intervene, implying a payout intensity policy $d_{i,t} = \rho a_{i,t}$ over $[t, t + \mathrm{d}t)$.

Moreover, the installed capital for production is $k_{i,t} = a_{i,t} + \widehat{a}_{i,t}$, and the manager can divert a fraction $1/\lambda$ of leased capital $\widehat{a}_{i,t}$ with $\lambda \geq 1$. As a punishment, the firm would lose his own capital $a_{i,t}$. In equilibrium, the manager is able to borrow up to the point where the manager has no incentive to divert leased capital, implying a collateral constraint $\widehat{a}_{i,t} \leq \lambda a_{i,t}$. The same form of collateral constraints is motivated similarly and adopted widely in the literature (e.g., Banerjee and Newman, 2003; Jermann and Quadrini, 2012; Buera and Shin, 2013; Moll, 2014).

The financial frictions can be summarized in the following proposition.

Proposition 1. *Because of the agency problem with limited enforcement, the firm's pay-*

[7] Managers can extract rents because corporate governance is imperfect. In practice, it is difficult to verify the cash flows generated by firms' assets, even though cash flows are observable and shareholders' property rights to firm assets are protected. For example, it is difficult to distinguish and verify rents and business expenses. The rents here do not include nonpecuniary private benefits, such as prestige from empire building (Eisfeldt and Rampini, 2008*b*).

out/issuance policy is subject to the following equity market constraint:

$$d_{i,t} = \rho a_{i,t}, \tag{16}$$

where $d_{i,t}$ is the dividend flow intensity over $[t, t + dt)$; moreover, the firm's leased capital is subject to the following collateral constraint:

$$-a_{i,t} \leq \widehat{a}_{i,t} \leq \lambda a_{i,t}. \tag{17}$$

Several points are worth further discussions. First, other agency problems can give rise to above equity market and collateral constraints, e.g., Gertler and Kiyotaki (2010); Gertler and Karadi (2011). Second, the equity market constraint is widely studied in the corporate finance literature (e.g., Myers, 2000; Lambrecht and Myers, 2008, 2012). It essentially implies that shareholders cannot freely move funds in and out of firms. Third, our analytically tractable formulation of capital market imperfections captures the fact that external funds available to a firm are limited and costly. Fourth, one specific interpretation of inter-firm borrowing and lending is the existence of a competitive rental market in which firms can rent capital from each other (e.g., Jorgenson, 1963; Hall and Jorgenson, 1969; Buera and Shin, 2013; Rampini and Viswanathan, 2013; Moll, 2014).

Managers' Problem. Similar to Moll (2014), our timeline assumption ensures that the idiosyncratic productivity $z_{i,t}$ is locally deterministic when managers make decisions at t for the production cycle $[t, t + \mathrm{d}t)$. Specifically, the manager of firm i makes leasing ($\widehat{a}_{i,s}$) and production ($u_{i,s}, \ell_{i,s}, x_{i,j,s}$) decisions for all $s \geq t$ to maximize the present value $J_{i,t}$ of his own rents

$$J_{i,t} = \max_{\widehat{a}_{i,s}, u_{i,s}, \ell_{i,s}, x_{i,j,s}} \mathbb{E}_t \left[\int_t^\infty \frac{\Lambda_s}{\Lambda_t} \tau a_{i,s} \mathrm{d}s \right], \tag{18}$$

subject to the equity market constraint (16), the collateral constraint (17), and the intertemporal budget constraint (3) with $\mathrm{d}I_{i,t}$ given by

$$\mathrm{d}I_{i,t} = y_{i,t}\mathrm{d}t - \int_0^{N_t} p_{j,t} x_{i,j,t} \mathrm{d}j\mathrm{d}t - w_t \ell_{i,t}\mathrm{d}t - u_{i,t} k_{i,t} \mathrm{d}\Delta_t - r_{f,t}\widehat{a}_{i,t}\mathrm{d}t - d_{i,t}\mathrm{d}t, \tag{19}$$

where the SDF Λ_t evolves according to

$$\frac{\mathrm{d}\Lambda_t}{\Lambda_t} = -r_{f,t}\mathrm{d}t - \eta_t \mathrm{d}W_t. \tag{20}$$

The variable η_t is the endogenous market price of risk. Because the technology, budget constraint, and collateral constraint are all linear in $a_{i,t}$, the value $J_{i,t}$ is also linear in $a_{i,t}$ with the following form:

$$J_{i,t} \equiv J_t(a_{i,t}, z_{i,t}) = \xi_t(z_{i,t})a_{i,t}, \tag{21}$$

where $\xi_{i,t} \equiv \xi_t(z_{i,t})$ captures the marginal value of capital to the manager, which depends on the firm's idiosyncratic productivity $z_{i,t}$ and the aggregate state of the economy. The variable $\xi_{i,t}$ evolves as follows:

$$\frac{\mathrm{d}\xi_{i,t}}{\xi_{i,t}} = \mu_{\xi,t}(z_{i,t})\mathrm{d}t + \sigma_{\xi,t}(z_{i,t})\mathrm{d}W_t + \sigma_{w,t}(z_{i,t})\mathrm{d}W_{i,t}, \tag{22}$$

where $\mu_{\xi,t}(z_{i,t})$, $\sigma_{\xi,t}(z_{i,t})$, and $\sigma_{w,t}(z_{i,t})$ are endogenously determined in equilibrium.

Exploiting the homogeneity of $J_{i,t}$ in capital $a_{i,t}$, we obtain the manager's optimal decisions, summarized in Lemma 1.

Lemma 1. *Factor demands and profits are linear in capital, and there is a productivity cutoff $\underline{z}_t$ for being active:*

$$u_t(z) = \begin{cases} 1, & z \geq \underline{z}_t \\ 0 & z < \underline{z}_t \end{cases}, \qquad k_t(a,z) = \begin{cases} (1+\lambda)a, & z \geq \underline{z}_t \\ 0 & z < \underline{z}_t \end{cases} \tag{23}$$

$$\ell_t(a,z) = \left[\frac{(1-\alpha)(1-\varepsilon)}{w_t}\right]^{\frac{1}{\alpha}} \left(\frac{\varepsilon}{p_t}\right)^{\frac{\varepsilon}{\alpha(1-\varepsilon)}} z u_t(z) k_t(a,z), \tag{24}$$

$$x_{j,t}(a,z) = \left(\frac{p_{j,t}}{p_t}\right)^{\frac{1}{\nu-1}} x_t(a,z), \tag{25}$$

where p_t is the price index and $x_t(a,z)$ is the demand for the composite of intermediate goods,

$$p_t = \left(\int_0^{N_t} p_{j,t}^{\frac{\nu}{\nu-1}} dj \right)^{\frac{\nu-1}{\nu}}, \tag{26}$$

$$x_t(a,z) = \left(\frac{\varepsilon}{p_t} \right)^{\frac{1-(1-\alpha)(1-\varepsilon)}{\alpha(1-\varepsilon)}} \left[\frac{(1-\alpha)(1-\varepsilon)}{w_t} \right]^{\frac{1-\alpha}{\alpha}} z u_t(z) k_t(a,z). \tag{27}$$

The productivity cutoff $\underline{z}_t$ is determined by:

$$\underline{z}_t \kappa_t = r_{f,t} + \delta_k + \sigma_k(\sigma_{\xi,t}(\underline{z}_t) - \eta_t). \tag{28}$$

where κ_t is

$$\kappa_t = \alpha(1-\varepsilon) \left(\frac{\varepsilon}{p_t} \right)^{\frac{\varepsilon}{\alpha(1-\varepsilon)}} \left[\frac{(1-\alpha)(1-\varepsilon)}{w_t} \right]^{\frac{1-\alpha}{\alpha}}. \tag{29}$$

At any point in time t, only firms whose productivity is greater than $\underline{z}_t$ produce, and these firms will rent the maximal amount $\widetilde{a}_{i,t} = \lambda a_{i,t}$ allowed by the collateral constraint. Equations (25) to (27) are standard results of CES aggregation. The productivity cutoff $\underline{z}_t$ is determined by equation (28), where the marginal production return, $\underline{z}_t \kappa_t$, is equal to the marginal cost of leased capital, $r_{f,t} + \delta_k + \sigma_k \left(\sigma_{\xi,t}(\underline{z}_t) - \eta_t \right)$, which includes the locally deterministic user cost of capital and the term $\sigma_k \left(\sigma_{\xi,t}(\underline{z}_t) - \eta_t \right)$ reflecting the firm's exposure to aggregate risks.

Using Lemma 1, equation (19) can be simplified as

$$\frac{\mathrm{d}I_{i,t}}{a_{i,t}} = (1+\lambda) \left(\kappa_t z_{i,t} \mathrm{d}t - \mathrm{d}\Delta_t - r_{f,t} \mathrm{d}t \right) \mathbb{1}_{z_{i,t} \geq \underline{z}_t} + (r_{f,t} - \rho) \mathrm{d}t. \tag{30}$$

As in Moll (2014), the drift term in capital accumulation is proportional to the firm's capital $a_{i,t}$. This is a direct consequence of the constant payout ratio (16) and the constant-returns-to-scale production technology (1) for a fixed N_t. The linear savings policy ensures that $a_{i,t} \geq 0$ for all t.

2.5 Equilibrium and Aggregation

The dividend intensity D_t is given by

$$D_t = \rho A_t + \int_{j=0}^{N_t} \pi_{j,t} \mathrm{d}j - S_t, \tag{31}$$

where A_t is the aggregate capital held by firms in the final goods sector, given by

$$A_t = \int_0^\infty \int_0^\infty a\varphi_t(a,z)\mathrm{d}a\mathrm{d}z. \tag{32}$$

In equation (31), the first term ρA_t captures the dividend of the final goods sector. The second term $\int_{j=0}^{N_t} \pi_{j,t}\mathrm{d}j$ captures the profits from the intermediate goods sector and the third term S_t captures the expenditure on R&D. The aggregate capital K_t in the economy is

$$K_t = \int_0^\infty \int_0^\infty k_t(a,z)\varphi_t(a,z)\mathrm{d}a\mathrm{d}z. \tag{33}$$

Definition 2.1 (Competitive Equilibrium)**.** *At any point in time t, the competitive equilibrium of the economy consists of prices w_t, $r_{f,t}$, and $\{p_{j,t}\}_{j=0}^{N_t}$, and corresponding quantities, such that*

(i) firms in the final goods sector maximize (18) by choosing $\widehat{a}_{i,t}$, $u_{i,t}$, $\ell_{i,t}$, and $x_{i,j,t}$, subject to (16), (17), and (19), given equilibrium prices;

(ii) intermediate goods producers maximize (6) by choosing $p_{j,t}$ for $j \in [0, N_t]$;

(iii) the equilibrium R&D expensiture S_t is determined by equation (10);

(iv) The SDF Λ_t is given by equation (15) and the risk-free rate $r_{f,t}$ is determined by

$$r_{f,t} = -\frac{1}{dt}\mathbb{E}_t\left[\frac{d\Lambda_t}{\Lambda_t}\right]; \tag{34}$$

(v) the labor market-clearing condition determines w_t:

$$L_t = \int_{\underline{z}_t}^\infty \int_0^\infty \ell_t(a,z)\varphi_t(a,z)dadz; \tag{35}$$

(vi) the leased capital market-clearing condition determines households' bond holdings B_t:

$$B_t = \int_0^\infty \int_0^\infty \widetilde{a}_t(a,z)\varphi_t(a,z)dadz. \tag{36}$$

The aggregate capital is the sum of capital in the final goods sector and households' bonds

$$K_t = A_t + B_t. \tag{37}$$

Finally, the resource constraint is automatically satisfied because of Walras's law (see Online Appendix A.4).

Because managers' problem is linear in capital $a_{i,t}$ (see equation (75)), it is not necessary to track the marginal distribution of capital conditional on each productivity type z.[8] We thus follow Moll (2014) and introduce the capital share $\omega_t(z)$ to fully characterize the distribution of firms in the final goods sector:

$$\omega_t(z) \equiv \frac{1}{A_t} \int_0^\infty a\varphi_t(a,z)\mathrm{d}a. \tag{38}$$

Intuitively, the capital share $\omega_t(z)$ plays the role of a density, and it captures the share of firms' capital held by each productivity type z. We define the analogue of the corresponding cumulative distribution function (CDF) as

$$\Omega_t(z) \equiv \int_0^z \omega_t(z')\mathrm{d}z'. \tag{39}$$

To ensure that the equilibrium growth is well behaved, as in standard growth models, we need output Y_t given by equation (40) to be homogenous of degree one in the accumulating factors N_t and K_t, i.e., $\frac{(1-\nu)\varepsilon}{\nu(1-\varepsilon)} + \alpha = 1$ as in Kung and Schmid (2015). For the rest of the paper, we assume this parameter restriction.

Proposition 2. *At any point in time $t \geq 0$, given the capital share $\omega_t(z)$, the equilibrium aggregate output is*

$$Y_t = Z_t K_t^\alpha L_t^{1-\alpha}, \tag{40}$$

where Z_t is the economy's TFP given by

$$Z_t = (\varepsilon\nu)^{\frac{\varepsilon}{1-\varepsilon}} H_t N_t^{1-\alpha} \;\; with \;\; H_t = \left[\frac{\int_{\underline{z}_t}^\infty z\omega_t(z)dz}{1-\Omega_t(\underline{z}_t)}\right]^\alpha. \tag{41}$$

The variable H_t captures the endogenous productivity of the final goods sector. The equilibrium K_t/A_t ratio is determined by the productivity cutoff $\underline{z}_t$ in equation (28):

$$K_t/A_t = (1+\lambda)\left[1-\Omega_t(\underline{z}_t)\right]. \tag{42}$$

[8]In fact, similar to the model of Moll (2014), the marginal distribution of capital is not stationary due to the constant-returns-to-scale production technology.

Factor prices are

$$p_{j,t} = 1/\nu \quad and \quad p_t = N_t^{\frac{\nu-1}{\nu}}/\nu, \tag{43}$$

$$w_t = (1-\alpha)(1-\varepsilon)Y_t/L_t, \tag{44}$$

where κ_t in equation (29) is simplified to $\kappa_t = \alpha(1-\varepsilon)H_t^{-\frac{1}{\alpha}}Y_t/K_t$. The aggregate profits of the intermediate goods sector and R&D expenditure are

$$N_t\pi_t = (1-\nu)\varepsilon Y_t, \tag{45}$$

$$S_t = (\chi v_t)^{\frac{1}{h}} N_t. \tag{46}$$

Equation (40) shows that the economy's aggregate TFP is $(\varepsilon\nu)^{\frac{\varepsilon}{1-\varepsilon}}H_tN_t^{1-\alpha}$, which depends on the knowledge stock N_t and the productivity H_t of the final goods sector. The productivity H_t reflects the degree of misallocation in the economy and determines the growth rate of N_t, and hence the growth rate of aggregate TFP. In equation (41), H_t is firms' average productivity z weighted by their capital share $\omega_t(z)$. Similar to Moll (2014), the equilibrium productivity cutoff $\underline{z}_t$ is determined directly by the CDF of capital share (see equation (42)) due to the bang-bang solution in equation (23). The value of H_t is higher when more productive firms are associated with more capital, which reflects a more efficient capital allocation across firms.

Equation (43) is a direct consequence of homogeneous intermediate goods producers facing the a constant elasticity of substitution, $1/(1-\nu)$. Equation (44) implies that the equilibrium wage is competitive, given by the labor share, $(1-\alpha)(1-\varepsilon)$, in the production function times the aggregate per-capita output, Y_t/L_t.

In the intermediate goods sector, equation (45) implies that the aggregate profit flow, $N_t\pi_t$, equals the share of intermediate goods in aggregate output, εY_t, multiplied by the profitability of intermediate-goods producers, as captured by the inverse of the elasticity of substitution $(1-\nu)$ among differentiated intermediate goods. In equation (46), innovators' R&D expenditure increases with the value of blueprints v_t with an elasticity of $1/h$.

2.6 Misallocation as a State Variable

The capital share $\omega_t(z)$ is crucial in determining the final goods sector's productivity H_t in equation (41), whose value reflects the misallocation of capital. As in the model of Moll (2014) and many other general-equilibrium models with heterogeneous agents, the capital share is an infinite-dimensional object that evolves endogenously.

In this section, we propose an analytical approximation of $\omega_t(z)$. In Online Appendix A.5, we apply the Berry-Esseen bound (Tikhomirov, 1980; Bentkus, Gotze and Tikhomoirov, 1997) to show that in the stationary equilibrium without aggregate shocks, the distribution of capital $a_{i,t}$ across firms in the final goods sector approximately follows a log-normal distribution at any point in time t. This motivates the following lemma.

Lemma 2. *The log capital, $\widetilde{a}_{i,t} = \ln(a_{i,t})$, across firms in the final goods sector approximately follows a normal distribution.*

According to equation (5), log individual productivity $\widetilde{z}_{i,t} = \ln z_{i,t}$ also follows a normal distribution, $\widetilde{z}_{i,t} \sim N(0, \sigma^2/2)$, in the stationary equilibrium. Thus, if at some initial point in time t_0, $\widetilde{z}_{i,t_0}$ and $\widetilde{a}_{i,t_0}$ follow a joint normal distribution, then the distribution of $\widetilde{z}_{i,t}$ and $\widetilde{a}_{i,t}$ will be joint normal for all $t \geq t_0$. This joint-normality allows us to derive a closed-form representation for the capital share $\omega_t(z)$ as follows.

Lemma 3. *For any $t \geq 0$, the capital share $\omega_t(z)$ can be approximated by the PDF of a log-normal distribution,*

$$\omega_t(z) = \frac{1}{z\sigma\sqrt{\pi}} \exp\left[-\frac{(\ln z + M_t\sigma^2/2)^2}{\sigma^2}\right], \tag{47}$$

where $M_t \equiv -Cov(\widetilde{z}_{i,t}, \widetilde{a}_{i,t})/\mathrm{var}(\widetilde{z}_{i,t}) = -2Cov(\widetilde{z}_{i,t}, \widetilde{a}_{i,t})/\sigma^2$.

Lemma 3 implies that under our approximation, the endogenous variable $M_t \equiv -\mathrm{Cov}(\widetilde{z}_{i,t}, \widetilde{a}_{i,t})/\mathrm{var}(\widetilde{z}_{i,t})$ is a sufficient statistic that characterizes $\omega_t(z)$. Intuitively, M_t captures the covariance between $\widetilde{z}_{i,t}$ and $\widetilde{a}_{i,t}$ at t across all firms in the final goods sector. A higher M_t indicates that more productive firms are associated with less capital, reflecting a higher degree of capital misallocation.

The main purpose of our analytical approximation proposed in Lemmas 2 and 3 is to highlight the economy's misallocation as a crucial state variable M_t. This allows us to achieve two results. First, it yields a simple closed-form characterization for the evolution of the economy (see Section 2.7), allowing us to clearly illustrate the key model mechanism that links the persistence of idiosyncratic productivity to that of misallocation. Second, it directly implies an intuitive and theoretically-justified empirical measure of misallocation (see Section 4.1), based on which we provide a set of empirical evidence to support our model predictions. Our idea of using tractable analytical approximations to deliver key model mechanisms is in spirit similar to several important works in the finance literature. For example, Campbell and Shiller (1988) propose log-linear present value approximations to clearly decompose the impact of discount-rate news and cash-flow news on stock valuations. Gabaix (2007, 2012) develops the class of "linearity-generating" processes to achieve analytical convenience when revisiting a set of macro-finance puzzles.

In terms of the accuracy of our approximation, we show in Online Appendix B that our approximation yields solutions sufficiently close to the numerical solutions based on directly tracking the evolution of $\omega_t(z)$ using higher-order approximations in both the balanced growth path and the stochastic steady state under our benchmark calibration. These findings are consistent with the numerical results of Winberry (2018) who approximates the distribution of heterogeneous firms in a DSGE model using a flexible parametric family and shows that the implied dynamics of aggregate variables, such as consumption, output, investment, SDF, and the autocorrelations in the covariance between log capital and log productivity, based on the log-normal approximation are very close to the results based on higher-order approximations, in the stochastic steady state.

Proposition 3. *Under our approximation specified in Lemma 2, the TFP Z_t is*

$$Z_t = (\varepsilon\nu)^{\frac{\varepsilon}{1-\varepsilon}} N_t^{1-\alpha} \left[(1+\lambda)\frac{A_t}{K_t}\exp\left(-\frac{\sigma^2}{2}M_t + \frac{\sigma^2}{4}\right)\Phi\left(\Phi^{-1}\left(\frac{1}{1+\lambda}\frac{K_t}{A_t}\right) + \frac{\sigma}{\sqrt{2}}\right)\right]^{\alpha}, \tag{48}$$

where $\Phi(\cdot)$ represents the CDF of a standard normal variable.

Equation (48) clearly shows that the economy's Z_t strictly decreases with the misallocation variable M_t. Thus, a lower M_t leads to higher Z_t and aggregate output Y_t.[9] Moreover, a lower misallocation M_t implies that more productive firms have more

[9] One of the main insights of Hsieh and Klenow (2009) is that the misallocation of resources lowers aggregate TFP. Proposition 3 shows that, in our model, the endogenous variable M_t determines the degree of misallocation as it determines productivity H_t, and thus TFP. Our analytical formula (48) for Z_t is an approximation for the exact formula (41), which can be linked to the industry-level TFP

capital, leading to a higher productivity cutoff $\underline{z}_t$. Under our approximation, equation (41) has a closed-form representation:

$$\underline{z}_t = \exp\left[-\frac{\sigma^2}{2}M_t - \Phi^{-1}\left(\frac{1}{1+\lambda}\frac{K_t}{A_t}\right)\frac{\sigma}{\sqrt{2}}\right]. \tag{49}$$

Thus, a lower misallocation M_t leads to a higher productivity H_t but fewer firms in the final goods sector will be active.[10]

2.7 Evolution of the Economy

The economy's transitional dynamics are characterized by the evolution of aggregate capital A_t in the final goods sector, knowledge stock N_t, and misallocation M_t. The aggregate capital K_t and bond holdings of households B_t are not state variables because they are determined by equations (37) and (42), given A_t. We summarize the evolution of the economy in the proposition below.

Proposition 4. *For any $t \geq 0$, the economy is fully characterized by the evolution of aggregate capital A_t in the final goods sector, knowledge stock N_t, and misallocation M_t, as follows*

$$\frac{dA_t}{A_t} = \alpha(1-\varepsilon)\frac{Y_t}{A_t}dt - \left[(\delta_k dt + \sigma_k dW_t)\frac{K_t}{A_t} + \delta_a dt - \sigma_a dW_t\right] - \left(\frac{K_t}{A_t} - 1\right) r_{f,t}dt - \rho dt, \tag{50}$$

$$\frac{dN_t}{N_t} = \chi\left(\chi v_t\right)^{\frac{1-h}{h}} dt - \delta_b dt, \tag{51}$$

$$dM_t = -\theta M_t dt - \frac{Cov(\widetilde{z}_{i,t}, d\widetilde{a}_{i,t})}{\text{var}(\widetilde{z}_{i,t})}, \tag{52}$$

where $Cov(\widetilde{z}_{i,t}, d\widetilde{a}_{i,t})$ is given by equation (159) in Online Appendix A.8.

formula derived by Hsieh and Klenow (2009). The key difference is that in our model, firms in the final goods sector produce homogeneous goods. But firms in the model of Hsieh and Klenow (2009) produce differentiated goods. In Online Appendix D, we show that by driving the elasticity of substitution among goods to infinity and wedges to zero, the industry-level TFP formula of Hsieh and Klenow (2009) coincides with our productivity H_t in equation (41).

[10] Banerjee and Moll (2010) show that there could be misallocation on the extensive margin because some productive firms may not run businesses. Our model, like Moll (2014), does not have misallocation on the extensive margin because production does not require upfront fixed costs (i.e., technology is convex).

In equation (50), the evolution of aggregate capital A_t is given by the capital share, $\alpha(1-\varepsilon)$, in the production function times the aggregate output to capital ratio, $Y_t/A_t\mathrm{d}t$, minus capital depreciation, $(\delta_k\mathrm{d}t+\sigma_k\mathrm{d}W_t)K_t/A_t+\delta_a\mathrm{d}t-\sigma_a\mathrm{d}W_t$, minus interests on households' loans, $(K_t/A_t-1)r_{f,t}\mathrm{d}t$, and dividend payout, $\rho\mathrm{d}t$.

In equation (51), the accumulation of knowledge stock N_t increases with the value of blueprints v_t because a higher v_t motivates innovators to increase R&D expenditure S_t (equation (46)). Importantly, the misallocation M_t determines the economy's endogenous growth rate over $[t, t+\mathrm{d}t)$. This is because v_t equals the present value of profit flow π_t (equation (8)), and thus v_t is higher when π_t is higher. A lower misallocation M_t increases the economy's TFP Z_t (equation (48)), leading to a higher aggregate output Y_t (equation (40)) and thus a higher profit flow π_t (equation (45)), and ultimately, a higher growth rate of the economy. By linking the final goods sector and the innovation sector through the endogenous TFP Z_t, the allocation of capital $a_{i,t}$ among firms of different productivity $z_{i,t}$ plays a crucial role in determining economic growth.

Equation (52) shows that the evolution of M_t depends on two terms. The first term $-\theta M_t\mathrm{d}t$ reflects time-varying productivity $z_{i,t}$ evolving according to equation (5). Intuitively, a higher θ implies a less persistent idiosyncratic productivity $z_{i,t}$, which pushes the misallocation $M_t = -\mathrm{Cov}(\widetilde{z}_{i,t}, \widetilde{a}_{i,t})/\mathrm{var}(\widetilde{z}_{i,t})$ towards zero. In Section 3, we show that the parameter θ crucially determines the economy's long-run consumption risk by affecting the persistence of M_t. The second term $\mathrm{Cov}(\widetilde{z}_{i,t}, \mathrm{d}\widetilde{a}_{i,t})/\mathrm{var}(\widetilde{z}_{i,t})$ captures the impact of capital accumulation, $\mathrm{d}\widetilde{a}_{i,t}$, evolving according to equation (3). Intuitively, a higher $\mathrm{Cov}(\widetilde{z}_{i,t}, \mathrm{d}\widetilde{a}_{i,t})$ implies that more productive firms also accumulate their capital at a higher rate, which reduces misallocation M_t.

Importantly, the variable $\mathrm{Cov}(\widetilde{z}_{i,t}, \mathrm{d}\widetilde{a}_{i,t})$ negatively depends on the aggregate shock $\mathrm{d}W_t$ (see equation (159) in Online Appendix A.8). Intuitively, a positive shock ($\mathrm{d}W_t > 0$) increases the depreciation rate of capital $k_{i,t}$, which reduces the capital accumulation of firms with productivity $z_{i,t}$ above the cutoff $\underline{z}_t$, without affecting those with productivity below the cutoff because these firms do not produce (see equation (23)). As a result, a positive shock leads to a lower $\mathrm{Cov}(\widetilde{z}_{i,t}, \mathrm{d}\widetilde{a}_{i,t})$, increasing the misallocation M_t. Aggregate shocks can have positive or negative effects on the aggregate capital A_t in the final goods sector, depending on the sign of $\sigma_k K_t/A_t - \sigma_a$. Under our calibration, $\sigma_k K_t/A_t - \sigma_a$ has a small magnitude. Thus, aggregate shocks $\mathrm{d}W_t$ do not directly drive large variations in A_t. In Section 3.4, we conduct counterfactual experiments to show that our model generates endogenous long-run consumption risk and a high Sharpe ratio mainly through the time variation in misallocation M_t, rather than time variation in aggregate capital stock.

Define $E_t = N_t/A_t$ as the aggregate knowledge stock to capital ratio. Because the economy is homogeneous of degree one in A_t, the three state variables (A_t, N_t, M_t) can be reduced to two state variables (E_t, M_t).

2.8 Balanced Growth Path

We characterize the economy's balanced growth path in the absence of aggregate shocks (i.e., $\mathrm{d}W_t \equiv 0$).

Proposition 5. *There is a balanced growth path in which E_t, M_t, and H_t are constant, and aggregate capital A_t, knowledge stock N_t, output Y_t, TFP Z_t, and consumption C_t grow at the same constant rate.*

In the presence of aggregate shocks, the economy's growth rate is time varying, depending on its misallocation M_t. A positive shock ($\mathrm{d}W_t > 0$) increases M_t, leading to a lower TFP Z_t through equation (48). Per our discussion in Section 2.7, a higher Z_t increases aggregate output Y_t, and thus the value of blueprints v_t. This motivates innovators to increase their R&D expenditure, leading to a higher growth rate of knowledge stock N_t. Because the economy's misallocation M_t is persistent (see equation (52)), i.i.d. shocks that affect misallocation M_t can generate persistent effects on both aggregate output and consumption growth. Thus, the economy's misallocation M_t not only determines contemporaneous growth but also predicts future growth of output and consumption.

Proposition 6 presents the relationship between R&D to capital ratio S_t/A_t and misallocation M_t in the balanced growth path.

Proposition 6. *The R&D to capital ratio, S_t/A_t, is negatively related to misallocation M_t in the balanced growth path,*

$$\ln\left(\frac{S_t}{A_t}\right) = -\frac{\alpha\sigma^2}{2h}M_t + \frac{\alpha\sigma^2}{4h} + \frac{1}{h}\ln(\chi) + \frac{\alpha}{h}\ln(1+\lambda) + \left(\frac{\alpha}{h}-1\right)\ln\left(\frac{A_t}{N_t}\right) + \frac{1}{h}\ln\left(\frac{(1-\nu)\varepsilon(\varepsilon\nu)^{\frac{\varepsilon}{1-\varepsilon}}}{r_{f,t}+\delta_b}\right) + \frac{\alpha}{h}\ln\left(\Phi\left(\Phi^{-1}\left(\frac{1}{1+\lambda}\frac{K_t}{A_t}\right) + \frac{\sigma}{\sqrt{2}}\right)\right). \quad (53)$$

3 Quantitative Analysis

In this section, we evaluate whether misallocation can quantitatively generate low-frequency movements in aggregate consumption growth (Bansal and Yaron, 2004). Section 3.1 calibrates the model. Section 3.2 presents the untargeted moments. Our model can generate moments of consumption growth and asset prices consistent with the data. In the model, misallocation significantly predicts R&D expenditure and future growth. Section 3.3 illustrates the important role of the persistence of idiosyncratic productivity in determining the persistence of aggregate consumption growth through the slow-moving misallocation variable. Section 3.4 conduct counterfactual and sensitivity analyses to illustrate the key mechanisms of the model. Importantly, we show that in our model, it is the time variation in misallocation M_t that generates endogenous long-run consumption risk and asset-pricing moments consistent with the data. Section 3.5 studies the welfare implications of misallocation.

3.1 Calibration

Panel A of Table 1 presents externally calibrated parameters. Following the standard practice, we set the capital share in the production technology at $\alpha = 0.33$. We set the yearly capital depreciation rates at $\delta_k = \delta_a = 0.04$. We set the share of intermediate inputs at $\varepsilon = 0.5$ according to the estimates of Jones (2011, 2013). The inverse markup is set at $\nu = 0.6$ to guarantee the existence of a balanced growth path. Recursive preferences are commonly used in recent works of asset pricing, we set the risk aversion at $\gamma = 10$ and the EIS at $\psi = 1.85$ as in Kung and Schmid (2015). We set the payout ratio at $\rho = 0.06$ and the rent extraction rate at $\tau = 0.01$. These two parameters imply that the average dividend payout to shareholders is about 2.5% of the market value in equilibrium. We set $h = 0.17$ so that the elasticity of new blueprints with respect to R&D is 0.83, following the calibration of Kung and Schmid (2015). We set the depreciation rate of knowledge stock at $\delta_b = 0.2$, which is within the range of the standard values used by the Bureau of Labor Statistics (BLS) in the R&D stock calculations. We set the volatility of idiosyncratic productivity at $\sigma = 1.39$ according to the calibration of Moll (2014). We set the persistence of idiosyncratic productivity at $\theta = 0.1625$, which implies that firms' idiosyncratic productivity has a yearly autocorrelation of $\exp(-\theta) = 0.85$, consistent with the estimate of Asker, Collard-Wexler and Loecker (2014) based on U.S. census data as well as the calibration

Table 1: Parameter calibration and targeted moments.

Panel A: Externally determined parameters					
Parameter	Symbol	Value	Parameter	Symbol	Value
Capital share	α	0.33	Capital depreciation rate	δ_k	0.04
Share of intermediate inputs	ε	0.5	Capital depreciation rate	δ_a	0.04
EIS	ψ	1.85	Risk aversion	γ	10
Inverse markup	ν	0.6	Rent extraction rate	τ	0.01
Dividend payout ratio	ρ	0.06	Knowledge depreciation rate	δ_b	0.2
$1-$ R&D elasticity	h	0.17	Vol. of idio. productivity	σ	1.39
Collateral constraint	λ	1	Persistence of idio. productivity	θ	0.1625

Panel B: Internally calibrated parameters and targeted moments					
Parameter	Symbol	Value	Moments	Data	Model
Subjective discount rate	δ	0.01	Real risk-free rate (%)	0.86	1.47
R&D productivity	χ	1.93	Consumption growth rate (%)	1.80	1.88
Capital depreciation shock	σ_k	0.19	Consumption growth vol. (%)	2.93	2.59
Capital efficiency shock	σ_a	0.18	$\sigma(\Delta \ln C_t)/\sigma(\Delta \ln Y_t)$	0.52	0.51

Note: In panel B, when constructing the model moments, we simulate a sample for 160 years with an 80-year burn-in period. We then compute the model-implied moments as we do for the data in each simulation. For each moment, the table reports the median of the distribution across 10,000 independent simulations. The data moments for the real risk-free rate, consumption growth rate, and consumption growth volatility are from Bansal and Yaron (2004), and the data moment for the ratio of the volatility of consumption growth and output growth ($\sigma(\Delta \ln C_t)/\sigma(\Delta \ln Y_t)$) is from Kaltenbrunner and Lochstoer (2010).

in the macroeconomics literature (e.g., Khan and Thomas, 2008; Moll, 2014; Winberry, 2018, 2021). We set the collateral constraint parameter at $\lambda = 1$, which is within the range of the calibration in the macroeconomics literature (e.g., Buera and Shin, 2013; Jermann and Quadrini, 2012; Midrigan and Xu, 2014; Moll, 2014; Dabla-Norris et al., 2021).

The remaining parameters are calibrated by matching the relevant moments summarized in Panel B of Table 1. When constructing the model moments, we simulate a sample for 160 years with an 80-year burn-in period. We then compute the model-implied moments as we do for the data in each simulation. For each moment, the table reports the median of the distribution across 10,000 independent simulations. The discount rate is set at $\delta = 0.01$ to generate a real risk-free rate of about 1.47%. The R&D productivity is a scaling parameter and is set at $\chi = 1.93$ to generate an average consumption growth rate of about 1.88%. We calibrate $\sigma_k = 0.19$ so that the

Table 2: Untargeted moments in data and model.

Moments	Data	Model	Moments	Data	Model
Panel A: Consumption moments					
$AC1(\Delta \ln C_t)$ (%)	0.49	0.48	$AC2(\Delta \ln C_t)$ (%)	0.15	0.28
$AC5(\Delta \ln C_t)$ (%)	-0.08	0.08	$AC10(\Delta \ln C_t)$ (%)	0.05	0.01
$VR2(\Delta \ln C_t)$ (%)	1.61	1.48	$VR5(\Delta \ln C_t)$ (%)	2.01	2.31
$AC1(\Delta \ln S_t)$ (%)	0.21	0.26	$AC1(M_t)$ (%)	0.65	0.72
Panel B: Asset pricing moments					
Sharpe ratio	0.33	0.36	σ_{r_f} (%)	0.97	1.40

Note: The notation $\Delta \ln X_t = \ln X_{t+1} - \ln X_t$ represents difference in $\ln X_t$ between year t and year $t-1$. The consumption data moments in panel A and the data moments in panel B are from Bansal and Yaron (2004). $ACk(\Delta \ln C_t)$ refers to the autocorrelation of consumption growth with a k-year lag. $VRk(\Delta \ln C_t)$ refers to the variance ratio of consumption growth with a k-year horizon. The data moment $AC1(\Delta \ln S_t)$ for yearly autocorrelation in R&D expenditure is from Kung and Schmid (2015). The data moment $AC1(M_t)$ for the misallocation measure is constructed in our sample. When constructing the model moments, we simulate a sample for 160 years with an 80-year burn-in period. We then compute the model-implied moments as we do for the data in each simulation. For each moment, the table reports the median of the distribution across 10,000 independent simulations.

model-implied volatility of consumption growth is about 2.59%. These moments are similar to the data moments estimated by Bansal and Yaron (2004). Finally, we calibrate $\sigma_a = 0.18$ so that the ratio of the volatility of consumption growth and output growth is 0.51, which is consistent with the data moment estimated by Kaltenbrunner and Lochstoer (2010).

3.2 Untargeted Moments and Predicative Regressions

Table 2 presents untargeted moments as a validation test of the model. Panel A shows that the moments reflecting the persistence of consumption growth implied by the model is very consistent with that in the data even though these moments are not directly targeted in our calibration. Specifically, the yearly autocorrelation of consumption growth ($AC1(\Delta \ln C_t)$) is 0.49 in the data and 0.48 in the model. Consumption growth is endogenously generated by innovators' R&D expenditure S_t, which depends on the degree of misallocation M_t. The last row of Table 2 shows that the yearly autocorrelation of R&D expenditure S_t and misallocation M_t also have similar values in the model and data. In Section 3.3, we show that the parameter θ governing the persistence of idiosyncratic productivity $z_{i,t}$ plays a major role in

Table 3: Model-implied relationship between misallocation, R&D, and growth.

Panel A: R&D intensity (S_t/N_t)		
	t	$t+1$
β	-0.035	-0.037
	$[-0.002]$	$[-0.002]$
R-squared	0.754	0.850

Panel B: Consumption growth ($\Delta \ln C_t$)					
	$t \to t+1$	$t \to t+2$	$t \to t+3$	$t \to t+4$	$t \to t+5$
β	-0.093	-0.220	-0.285	-0.333	-0.372
	$[-0.011]$	$[-0.009]$	$[-0.016]$	$[-0.025]$	$[-0.033]$
R-squared	0.472	0.892	0.791	0.694	0.619

Panel C: Output growth ($\Delta \ln Y_t$)					
	$t \to t+1$	$t \to t+2$	$t \to t+3$	$t \to t+4$	$t \to t+5$
β	-0.192	-0.232	-0.245	-0.281	-0.285
	$[-0.029]$	$[-0.035]$	$[-0.042]$	$[-0.049]$	$[-0.057]$
R-squared	0.341	0.344	0.297	0.284	0.242

Note: We simulate a sample for 160 years with an 80-year burn-in period. In panel A, we regress the R&D intensity (S_t/N_t) in year t and $t+1$ on misallocation M_t. In panels B and C, we regress the cumulative growth of consumption and output from year t to $t+\tau$ ($\tau = 1, 2, ..., 5$) on M_t, respectively. We report the median of each statistic across 10,000 independent simulations.

determining the persistence of consumption growth.

Panel B of Table 2 shows that our model implies a smooth risk-free rate and a high Sharpe ratio for the consumption claim, consistent with both the data and the moments implied by the long-run risk model of Bansal and Yaron (2004). Thus, the endogenous low-frequency movements in aggregate consumption growth implied by our model have reasonable implications for asset prices.

Table 3 studies the relationship between misallocation, R&D and growth in our model. Panel A shows that misallocation M_t in year t is negatively correlated with contemporaneous R&D intensity S_t/N_t. The misallocation M_t also negatively predicts the R&D intensity S_{t+1}/N_{t+1} in the next year.

Panel B of Table 3 shows that misallocation M_t significantly negatively predicts future consumption growth over time horizons of one year to five years. The coefficients are more negative for longer horizons as consumption growth is persistent. Misallocation can predict future consumption growth in our model because it is the persistence in misallocation that generates persistent consumption growth through endogenous R&D. Because our model only has one aggregate shock, the conditional consumption growth is strongly correlated with innovations in misallocation. As a result, both the t-statistic and the R-sqaured are large. Panel C of Table 3 shows that misallocation M_t also negatively predicts future output growth. The regression coefficients are statistically significant, though the standard errors are larger than those of the coefficients predicting consumption growth.

3.3 Impulse Response Functions

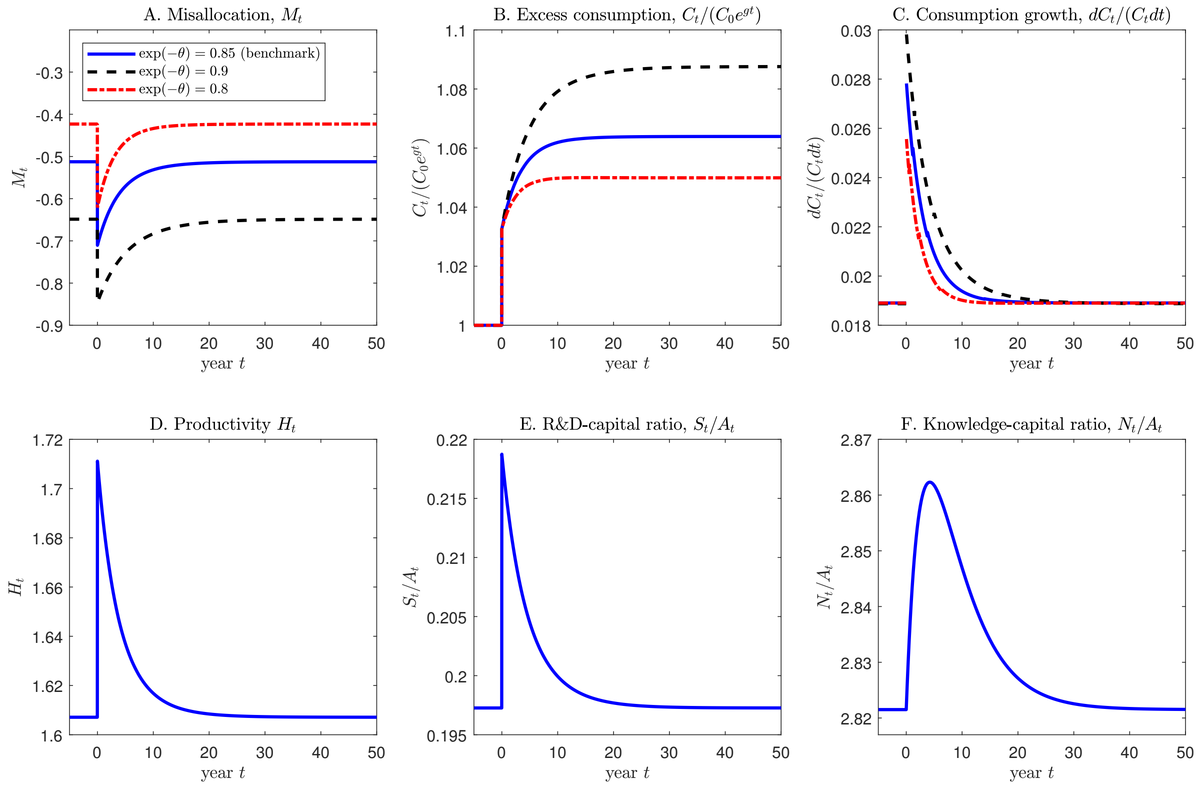
We illustrate how the persistence of idiosyncratic productivity determines the persistence of aggregate consumption growth and the quantitative implications of the model.

Response to a One-Time Shock. Consider a stationary economy in the balanced growth path. At $t = 0$, there is a one-time unexpected shock (i.e., $\mathrm{d}W_t < 0$ over $[0, \mathrm{d}t)$) that reduces the misallocation M_t by σ_k. The blue solid line in each panel of Figure 1 plots the transitional dynamics under our benchmark calibration.

Panel A plots the evolution of the misallocation M_t after the shock, which follows equation (52). The blue solid line shows that misallocation M_t declines immediately at $t = 0$ and slowly recovers and reaches the steady-state level after about 20 years.

In the absence of aggregate shocks, aggregate consumption would be $C_0 \exp(gt)$, growing at a constant rate $g = 1.88\%$ for all $t \geq 0$. We take out the trend effect in C_t by focusing on excess consumption, defined by $C_t/(C_0 \exp(gt))$. The blue solid line in panel B shows that excess consumption $C_t/(C_0 \exp(gt))$ stays at one before the shock, and it immediately jumps to about 1.03 when the shock hits at $t = 0$, and continues to increase until reaching the balanced growth path. Even though the shock is transitory, the economy converges to a steady state with permanently higher consumption due to the endogenous accumulation of capital A_t and knowledge stock N_t.

Panel C illustrates a similar idea by plotting the conditional consumption growth rate, defined by $\mathrm{d}C_t/(C_t \mathrm{d}t)$. The blue solid line shows that the conditional consumption growth rate increases dramatically to about 2.8% when the shock hits at $t = 0$. This is because the reduction in misallocation M_t immediately increases the productivity H_t of the final goods sector (panel D). A higher H_t increases the profits of innovators, motivating them to spend more on R&D (panel E), which consequently leads to a



Note: Consider an unexpected shock that reduces misallocation M_t by σ_k at $t = 0$. Panels A, B, and C plot the transitional dynamics of misallocation M_t, excess consumption $C_t/(C_0 e^{gt})$, and conditional consumption growth rate $\mathrm{d}C_t/(C_t\mathrm{d}t)$ in three economies with different θ. For each economy, we calibrate the parameter χ so that the consumption growth rate in the balanced growth path is the same as our benchmark calibration. All other parameters are set according to our calibration in Table 1. Panels D, E, and F plot the transitional dynamics of final goods sector's productivity H_t, R&D to capital ratio S_t/A_t, and knowledge stock to capital ratio $E_t = N_t/A_t$ for the benchmark economy with $\exp(-\theta) = 0.85$.

Figure 1: Transitional dynamics after a one-time shock in misallocation M_t.

higher growth rate of the economy. Crucially, it is the persistence in misallocation M_t (panel A) that results in persistent excess consumption growth relative to the balanced growth path (panels B and C). Panel F plots the evolution of the knowledge stock to capital ratio $E_t = N_t/A_t$, which has hump-shaped dynamics because we only introduce a one-time shock in M_t at $t = 0$.

Role of the Persistence of Idiosyncratic Productivity. In panels A, B and C, we further compare our benchmark calibration with two economies of different persistence of idiosyncratic productivity $z_{i,t}$. The yearly autocorrelation in $\ln z_{i,t}$ is $\text{corr}(\ln z_{i,t}, \ln z_{i,t+1}) = \exp(-\theta)$ according to equation (5). Panel A shows that the economy with a higher persistence of $z_{i,t}$ is associated with lower misallocation in the balanced growth path (i.e., the red dash-dotted line is above the blue solid line, which is above the black dashed line). This follows the insight of Buera and Shin (2011) and Moll (2014): More productive firms accumulate more capital relative to less productive firms over time. Thus, the covariance between capital and productivity across firms increases with the persistence of productivity, resulting in less misallocation in the balanced growth path.

In addition, we find that the convergence speed of M_t decreases with the persistence of idiosyncratic productivity (see equation (52)). Specifically, we compute the half-life of transitions, defined as the time required for M_t to recover to half of its value in the balanced growth path after the shock. The half life of M_t is 2.2, 2.8, and 3.8 years for the red dash-dotted ($\exp(-\theta) = 0.8$), blue solid ($\exp(-\theta) = 0.85$), and black dashed lines ($\exp(-\theta) = 0.9$), respectively, indicating that M_t is more persistent when θ is smaller. Comparing the three curves in panels B and C, it is clear that the economy with a higher persistence of $z_{i,t}$ has more persistent consumption growth after the shock in M_t (i.e., it takes more time for $C_t/(C_0 \exp(gt))$ and $\mathrm{d}C_t/(C_t\mathrm{d}t)$ to converge to the levels in the balanced growth path).

One key insight of our model is that the persistence of the level of idiosyncratic productivity, $z_{i,t}$, plays an important role in determining the persistence of the growth rate of aggregate consumption, $\mathrm{d}C_t/(C_t\mathrm{d}t)$. The persistence of these two variables is connected with each other via the persistent endogenous misallocation M_t. Specifically, when idiosyncratic productivity becomes more persistent, the persistence of misallocation M_t increases (panel A of Figure 1). As a result, the consumption growth rate also becomes more persistent (panel C of Figure 1) because misallocation M_t directly determines the economy's aggregate output Y_t and consumption C_t through its effect on TFP Z_t.

Our insight is related to that of Moll (2014), who shows that transitions to steady states are slow when idiosyncratic productivity shocks are persistent. Different from Moll (2014), we show that the persistence of idiosyncratic productivity not only determines the transition speed of the level of output and TFP, but also the growth rate of aggregate consumption in a model with endogenous growth. This allows our theory to generate endogenous low-frequency movements in aggregate consumption growth through persistent misallocation, thereby rationalizing asset prices in the capital market. Crucially, by linking the persistence of idiosyncratic productivity with the persistence of aggregate consumption growth, our model provides a way to estimate the long-run risk of aggregate consumption growth based on granular firm-level data, which helps address the issues of weak identification in the long-run risk literature (Chen, Dou and Kogan, 2022; Cheng, Dou and Liao, 2022).

3.4 Inspection of Key Parameters and Mechanisms

We conduct counterfactual and sensitivity analyses to illustrate the key mechanisms of the model. Importantly, we show that in our model, it is the time variation in misallocation M_t, rather than time variation in aggregate capital stock, that generates endogenous long-run consumption risk and asset-pricing moments consistent with the data. This differentiates our theoretical mechanism from those in the literature (e.g., Kaltenbrunner and Lochstoer, 2010; Kung and Schmid, 2015). Table 4 shows how the main variables of our model respond to changes in key parameters and variables.

Column (1) presents the baseline case of our full model. In column (2), we consider a less persistent idiosyncratic productivity by increasing θ from 0.1625 to 0.22, which corresponds to a reduction in the yearly autocorrelation of $\ln z_{i,t}$ from 0.85 to 0.8. Compared with the baseline, the average misallocation M_t increases from -0.53 to -0.43 because productive firms are more likely to be unproductive when productivity is more transitory, weakening the self-financing channel through capital accumulation. As a result, the final-goods sector's productivity H_t decreases from 1.63 to 1.56. The average consumption growth rate decreases to 1.05%. A lower persistence of idiosyncratic productivity reduces the yearly autocorrelation of consumption growth to 0.42; moreover, aggregate TFP, output, and misallocation all become less persistent. Because of the decrease in consumption persistence, the Sharpe ratio declines from 0.36 in the baseline to 0.31 in column (2).

In column (3), we consider a more restrictive collateral constraint by reducing λ from 1 to 0.9. The average misallocation M_t stays unchanged compared to the baseline. This is because the equilibrium misallocation is mainly determined by firms' differential speed of capital accumulation across different productivity $z_{i,t}$ (i.e., the term $\mathrm{Cov}(\widetilde{z}_{i,t}, \mathrm{d}\widetilde{a}_{i,t})$ in equation (52)). A change in λ does not affect this difference much because a lower λ scales down the revenue of both high-productivity and low-productivity firms. However, reducing λ directly leads to a lower TFP Z_t in equation

Table 4: Inspection of key parameters.

	(1) Baseline	(2) $\theta = 0.22$	(3) $\lambda = 0.9$	(4) $\chi = 1.85$	(5) $\sigma_k = 0$	(6) $\sigma_a = 0$	(7) Fix M_t
$\mathbb{E}[M_t]$	−0.53	−0.43	−0.53	−0.53	−0.57	−0.68	−0.53
$\mathbb{E}[H_t]$	1.63	1.56	1.61	1.61	1.57	1.62	1.62
$\mathbb{E}[\Delta \ln C_t]$ (%)	1.88	1.05	1.53	1.09	1.73	3.87	1.47
$\sigma(\Delta \ln C_t)$ (%)	2.59	2.62	2.46	2.72	3.79	6.64	0.51
$\sigma(\Delta \ln C_t)/\sigma(\Delta \ln Y_t)$	0.56	0.46	0.51	0.46	0.16	0.24	0.13
$AC1(\Delta \ln C_t)$	0.48	0.42	0.48	0.49	0.52	0.47	0.69
$AC1(M_t)$	0.72	0.66	0.73	0.72	0.97	0.59	1.00
$\sigma(M_t)$ (%)	20.44	19.37	19.90	20.78	9.06	18.57	0
Sharpe ratio	0.36	0.31	0.35	0.37	0.40	0.72	0.09

Note: The notation $\Delta \ln X_t = \ln X_{t+1} - \ln X_t$ represents difference in $\ln X_t$ between year t and year $t-1$. $AC1(\Delta \ln C_t)$ and $AC1(M_t)$ refers to the yearly autocorrelation of consumption growth and misallocation. When constructing the model moments, we simulate a sample for 160 years with an 80-year burn-in period. We then compute the model-implied moments as we do for the data in each simulation. We report the median of each moment's distribution across 10,000 independent simulations.

(48), reflecting the instantaneous reallocation of capital through the capital leasing market. The lower H_t reduces the average consumption growth rate to 1.53% and the volatility of consumption growth to 2.46%, without affecting the persistence of consumption growth or other macro variables much. The Sharpe ratio is reduced slightly to 0.35.

In column (4), we consider a lower productivity of R&D by reducing χ from 1.93 to 1.85. Compared with our baseline in column (1), column (4) shows that all variables remain roughly unchanged, except for a lower consumption growth rate (1.09% vs. 1.88% in the baseline). The lower growth rate is determined by the productivity of R&D, rather than a better allocation of capital among firms because H_t is roughly unchanged. As discussed in Section 3.1, the parameter χ can be thought of as a pure scaling factor that determines the equilibrium growth rate.

In column (5), we turn off capital depreciation shocks by setting $\sigma_k = 0$. Thus, the economy's only aggregate shock is the capital efficiency shock $\sigma_a \mathrm{d}W_t$ entering equation (50). The main difference between columns (1) and (5) are the following: i). consumption growth volatility increases substantially from 2.59% to 3.79% after eliminating capital depreciation shocks. The reason is that $-\sigma_k K_t/A_t \mathrm{d}W_t$ and $\sigma_a \mathrm{d}W_t$ in equation (50) roughly offset each other under our benchmark calibration. Once we eliminate capital depreciation shocks by setting $\sigma_k = 0$, the aggregate capital growth dA_t/A_t in the final goods sector becomes much more volatile, resulting in a higher volatility of consumption growth; ii). M_t becomes very persistent, with a yearly autocorrelation of 0.97, and much less volatile, with an annualized volatility of 9.06%, as opposed to 0.72 and 20.44% in our baseline (column (1)). Intuitively, capital efficiency shocks do not directly affect the evolution of M_t through equation (52), because the term $\text{Cov}(\widetilde{z}_{i,t}, d\widetilde{a}_{i,t})$ only depends on $\sigma_k dW_t$ but not on $\sigma_a dW_t$. The mild time variation in M_t is driven by several slow-moving variables (i.e., $\underline{z}_t$, $r_{f,t}$, and κ_t, which determine $\text{Cov}(\widetilde{z}_{i,t}, d\widetilde{a}_{i,t})$ through equation (159)); and iii). The Sharpe ratio remains high in column (5) because of the volatile and persistent consumption growth. However, the volatile and persistent consumption growth dynamics are not driven by time-varying misallocation M_t, instead, they are mechanically determined by the persistent knowledge stock to capital ratio E_t.[11]

Complementary to the experiment in column (5), we eliminate capital efficiency shocks in column (6) by setting $\sigma_a = 0$. Compared with our baseline (column (1)), this leads to substantially higher consumption growth rate, consumption growth volatility, and Sharpe ratio. The main reason is that a positive shock dW_t reduces aggregate capital A_t in the final goods sector through equation (50) and increases misallcoation M_t through equation (52), both of which will substantially reduce aggregate output and consumption. As a result, consumption growth becomes much more volatile, resulting in a much higher Sharpe ratio and average consumption growth rate (due to Jensen's inequality). By contrast, in our baseline (column (1)), a positive shock dW_t does not affect aggregate capital A_t much because capital efficiency shocks $\sigma_a dW_t$ largely offset capital depreciation shocks $-\sigma_k K_t/A_t dW_t$ in equation (50). Thus, the quantitative implications of our baseline mostly reflect the impact of time-varying misallocation M_t.

Finally, we echo the arguments above by exogenously fixing M_t at its long-run mean in column (7). In this case, the time variation in aggregate variables is purely driven by the term $(-\sigma_k K_t/A_t + \sigma_a)dW_t$ in equation (50). Because our benchmark calibration roughly has the property of $\sigma_k \overline{K_t/A_t} \approx \sigma_a$, the volatility of consumption growth is only 0.51% in column (6), which is much smaller than 2.59% in our baseline (column (1)). As a result, the Sharpe ratio is merely 0.09 despite the high persistence in consumption growth.

[11] This mechanism is related to Kaltenbrunner and Lochstoer (2010), who emphasize households' consumption smoothing motives in generating persistent consumption growth. In our model, households make loan decisions (i.e., B_t) instead of corporate investment decisions.

Taken together, columns (5) to (7) show that our model generates endogenous long-run consumption risk and a high Sharpe ratio mainly through the time variation in misallocation M_t, rather than time variation in aggregate capital stock. Our calibration of σ_k and σ_a roughly ensures that the aggregate shock $\mathrm{d}W_t$ mainly directly drives the evolution of misallocation M_t, rather than the evolution of aggregate capital A_t in the final goods sector. As a key theoretical and conceptual point, we emphasize the role of misallocation in generating low-frequency movements in aggregate consumption growth, which differentiates our paper from the seminal work of Kung and Schmid (2015). In Section 4, we show that the misallocation-based story is also supported by the data.

3.5 Welfare Implications of Misallocation

In our model, fluctuations in aggregate quantities reflect changes in misallocation, implying that the cost of misallocation can be evaluated by measuring the cost of business cycles. We provide some suggestive quantitative evaluation through the lens of the model.

Because aggregate consumption includes both transitory fluctuations and permanent fluctuations, it is important to define business cycles to comprise only transitory fluctuations (Alvarez and Jermann, 2004, 2005). Kung and Schmid (2015) emphasize the importance of distinguishing business cycles from growth cycles (i.e., the low-frequency movements in consumption) in endogenous stochastic growth models with long-run consumption risk.

Following the approach of Alvarez and Jermann (2004), we use the model-generated consumption time series to calculate the potential benefits of eliminating business cycles (i.e., transitory fluctuations). Specifically, business cycles are defined as fluctuations that last up to eight years as in Burns and Mitchell (1946) and Alvarez and Jermann (2004). In our simulated consumption time series, we use a one-sided moving average to represent a low-pass filter that lets pass frequencies that correspond to cycles of eight years or more (Baxter and King, 1999).[12] The estimated moving average coefficients allow us to calculate the costs of business cycles based on the model-implied consumption risk premium (Alvarez and Jermann, 2004, equations (4) and (6)).

[12]The frequency response function is one for frequencies lower than eight years and zero otherwise. Section III of Alvarez and Jermann (2004) provides details on estimating the moving average coefficients.

Our model implies that eliminating business cycles leads to a welfare gain of 0.51%, with a 95% confidence interval of $[0.15\%, 1.01\%]$. The magnitude of our estimate is similar to that of Alvarez and Jermann (2004), which is directly obtained from asset price data. We also estimate the potential benefits of eliminating all consumption uncertainty. Specifically, we calibrate the parameter χ in an economy without aggregate shocks to achieve the same growth rate of aggregate consumption as our benchmark calibration. Because the representative household's utility (12) is homogeneous of degree one in aggregate consumption, the consumption-equivalent welfare gain of eliminating all consumption uncertainty is equal to the percentage increase in the value of utility U_0, when the agent moves from the baseline economy to the economy without aggregate shocks at $t = 0$. Our model implies that eliminating all consumption uncertainty leads to a welfare gain of as large as 35.14%, with a 95% confidence interval of $[20.23\%, 51.90\%]$, which is also in the ballpark range estimated by Alvarez and Jermann (2004). Per the insight of Alvarez and Jermann (2004, 2005), our model implies a large gain from the elimination of all consumption uncertainty because consumption and the pricing kernel have large permanent components. This further confirms our main results that misallocation endogenously drives low-frequency movements in aggregate consumption growth in our model.

4 Empirical Results

In this section, we conduct empirical analyses to test our model's main predictions. In Section 4.1, we use the U.S. data to construct an empirical measure of misallocation implied by our model. We show that misallocation becomes more severe during economic recessions and financial crises. In Section 4.2, we provide evidence for the effects of misallocation on R&D intensity and economic growth over long horizons. We thus identify shocks to the misallocation as a proxy for shocks to the low-frequency components of consumption growth and the pricing kernel. In Section 4.3, we provide cross-sectional evidence to support the model's key prediction that fluctuations in misallocation drive the low-frequency component of consumption growth. In Section 4.4, we study the cross-sectional asset pricing implications of misallocation, as a macroeconomic risk factor. Finally, in Section 4.5, we provide empirical support for the model's core mechanism that firms' financial constraints result in misallocation, which in turn determines firms' R&D expenditure.

4.1 Empirical Measure of Misallocation

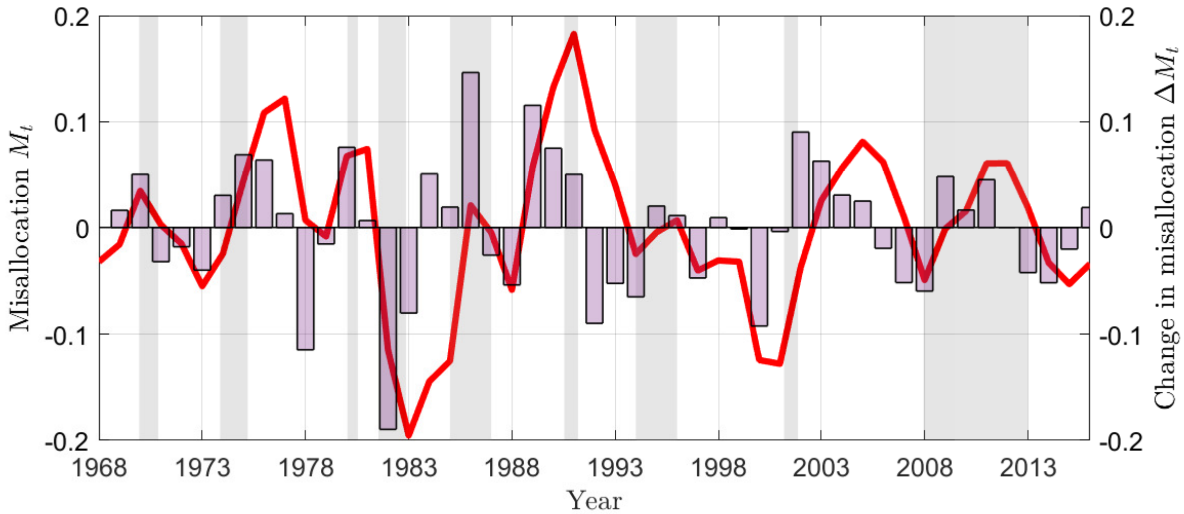
In our model, the economy's misallocation can be fully captured by the state variable $M_t \equiv -\mathrm{Cov}(\widetilde{z}_{i,t}, \widetilde{a}_{i,t})/\mathrm{var}(\widetilde{z}_{i,t})$, defined in Lemma 3. The term $\mathrm{Cov}(\widetilde{z}_{i,t}, \widetilde{a}_{i,t})/\mathrm{var}(\widetilde{z}_{i,t})$ can be estimated by a single variable linear regression that regresses log capital $\widetilde{a}_{i,t}$ on log productivity $\widetilde{z}_{i,t}$ using the cross section of firms in each year t:

$$\widetilde{a}_{i,t} = \alpha_t + \beta_t \widetilde{z}_{i,t} + \varepsilon_{i,t}, \tag{54}$$

where the estimated coefficient $\widehat{\beta}_t$ directly captures $\mathrm{Cov}(\widetilde{z}_{i,t}, \widetilde{a}_{i,t})/\mathrm{var}(\widetilde{z}_{i,t})$. Thus, motivated by our theory, we construct an empirical measure of M_t using the HP-filtered time-series of the minus of the estimated coefficient $\widehat{\beta}_t$ from 1965 to 2016. The HP filter allows us to extract the cyclical component of the time series of $-\widehat{\beta}_t$, following the literature (e.g., Eisfeldt and Rampini, 2006). Figure 2 plots the time series of the empirical measure M_t. Sharp spikes in M_t are observed during periods of economic downturns, including economic recessions and three financial crises: the savings and loan (S&L) crisis from January 1986 to December 1987, the Mexican peso crisis from January 1994 to December 1995, and the European sovereign debt crisis from September 2008 to December 2012. The stylized pattern shown in Figure 2 is consistent with our model's prediction that a large increase in misallocation generally represents a period of time with macroeconomic recessions and financial turmoil.

We now elaborate on the construction of independent and dependent variables $\widetilde{a}_{i,t}$ and $\widetilde{z}_{i,t}$ in specification (54). These two variables are empirical measures of log capital and productivity for firm i in year t, respectively, constructed using the U.S. Compustat data. Specifically, following the standard practice, we exclude firms from financial industry, utility industry, and public administration (SIC codes between $6,000 - 6,999$, $4,900 - 4,999$, and $9,000 - 9,999$). We construct the independent variable $\widetilde{a}_{i,t}$ using the average log capital of firm i over the past T years, i.e., $\widetilde{a}_{i,t} \equiv T^{-1} \sum_{\tau=1}^{T} \ln(capital_{i,t+1-\tau})$, with $T = 3$. The empirical results are robust to alternative choices of T. Firm i's capital is measured by its net property, plant and equipment, i.e., $capital_{i,t} = ppent_{i,t}$.[13]

[13] All the empirical results are robust if we use a firm's tangible net worth to construct its capital, i.e., $capital_{i,t} = tangible\_net\_worth_{i,t}$. A firm's tangible net worth is constructed as $tangible\_net\_worth_{i,t} = ppent_{i,t} + current\_assets_{i,t} + other\_assets_{i,t} - total\_liabilities_{i,t}$, which is the firm's net property, plant and equipment plus current assets plus other assets minus total liabilities. As emphasized by the seminal work of Chava and Roberts (2008), lenders commonly use a firm's tangible net worth to assess its ability to support and pay back loans. Naturally, tangible net worth, as a measure for firms' borrowing capacity, is widely reflected in loan covenants (e.g., DeAngelo, DeAngelo and Wruck, 2002; Roberts and Sufi,



Note: The red solid line plots the time series of our empirical misallocation measure M_t (corresponding to the left y-axis). The pink bars represent its year-on-year changes ΔM_t (corresponding to the right y-axis). The shaded areas represent recessions or severe financial crises.

Figure 2: Time series plot of our empirical misallocation measure M_t.

We construct the dependent variable $\widetilde{z}_{i,t}$ using the average log productivity of firm i over the past T years, i.e., $\widetilde{z}_{i,t} \equiv T^{-1} \sum_{\tau=1}^{T} \ln(productivity_{i,t+1-\tau})$. In our model, firms' technology has constant returns to scale, and thus firms' output and sales are linear in capital after substituting out labor and intermediate inputs using Lemma 1, i.e., $y_{i,t} = q_t z_{i,t} k_{i,t}$ for $z_{i,t} \geq \underline{z}_t$, where $q_t = (\varepsilon/p_t)^{\frac{\varepsilon}{\alpha(1-\varepsilon)}} \left[(1-\alpha)(1-\varepsilon)/w_t\right]^{\frac{1-\alpha}{\alpha}}$. This equation suggests that the ratio of sales ($y_{i,t}$) to production capital $k_{i,t}$ is $q_t z_{i,t}$, which captures firms' productivity $z_{i,t}$ in our model. Motivated by this structural relationship, we thus measure firm i's productivity using $productivity_{i,t} = sales_{i,t}/production\_capital_{i,t}$.[14] Following the model, $production\_capital_{i,t}$ is measured by the sum of the firm's own capital ($ppent_{i,t}$) and rented capital. We measure the amount of rented capital by capitalizing rental expenses, following standard accounting practice and the literature (e.g., Rauh and Sufi, 2011; Rampini and Viswanathan, 2013). Specifically, firm i's rented

2009; Sufi, 2009; Prilmeier, 2017). Thus, measuring a firm's capital based on its tangible net worth could be more consistent with the specification of borrowing constraint (17) in our model. In Figure OA.2 of Online Appendix C.1, we show that the misallocation measure constructed using tangible net worth has similar cyclical patterns to that constructed using net property, plant and equipment (see Figure 2).

[14]Note that although this productivity measure depends on q_t according to our model, the estimated coefficient β_t in specification (54) will not be affected because the independent variable $\widetilde{z}_{i,t}$ is the log of productivity, and thus the aggregate variable q_t is absorbed by the constant term α_t when estimating (54).

Table 5: Misallocation, R&D, and growth in the data.

Panel A: R&D intensity		
	t	$t+1$
β	-0.076^{**}	-0.079^{**}
	[0.032]	[0.031]
R-squared	0.102	0.116

Panel B: Consumption growth					
	$t \to t+1$	$t \to t+2$	$t \to t+3$	$t \to t+4$	$t \to t+5$
β	-0.035	-0.060	-0.100^{*}	-0.154^{**}	-0.208^{***}
	[0.024]	[0.040]	[0.052]	[0.060]	[0.066]
R-squared	0.040	0.042	0.069	0.120	0.173

Panel C: Output growth					
	$t \to t+1$	$t \to t+2$	$t \to t+3$	$t \to t+4$	$t \to t+5$
β	-0.034	-0.062	-0.092	-0.153^{*}	-0.216^{***}
	[0.040]	[0.062]	[0.075]	[0.082]	[0.083]
R-squared	0.015	0.020	0.029	0.067	0.123

Note: The sample period is 1965-2016. Standard errors are reported in brackets. *, **, and *** indicate statistical significance at 10%, 5%, and 1%, respectively.

capital in year t is its total rental expenses in the year multiplied by a factor 10 and capped by a fraction, 0.25, of $ppent_{i,t}$.[15]

4.2 Growth Forecasts

Time-varying growth prospects in consumption are at the core of the long-run risk literature following Bansal and Yaron (2004). However, the empirical evidence regarding this channel is still controversial. Few instruments have been shown to successfully predict consumption growth over long horizons. Our model implies that the degree of misallocation predicts R&D intensity, which determines future economic growth. Thus, the empirical misallocation measure M_t motivated by our model naturally provides an economically meaningful predictor for R&D intensity and aggregate growth rates. In this subsection, we present empirical evidence for this model implication.

We obtain data on private business R&D investment from the National Science

[15] All empirical results are robust if we use a factor 5, 6, or 8 to capitalize rental expenses or cap the amount by a fraction, 0, 0.5, 1, or 2, of $ppent_{i,t}$.

Foundation (NSF) and the R&D stock from the Bureau of Labor Statistics (BLS). Following Kung and Schmid (2015), these two time series are considered as empirical counterparts for S_t and N_t, respectively. The ratio of the two (i.e., S_t/N_t) is our empirical measure for R&D intensity. We regress R&D intensity in the current year (t) and the next year ($t+1$) on the misallocation measure M_t, as follows:

$$\frac{S_{t+h}}{N_{t+h}} = \alpha + \beta M_t + \nu_t, \tag{55}$$

where $h = 0, 1$. Panel A of Table 5 presents the empirical results, which indicate that a higher misallocation is associated with a decline in contemporaneous R&D intensity and predicts a lower R&D intensity in the next year.

Next, we empirically test whether misallocation negatively predicts future consumption growth. We obtain annual data for consumption from the Bureau of Economic Analysis (BEA). Consumption is measured as per-capita expenditures on nondurable goods and services. The nominal variables are converted to real terms using the consumer price index (CPI), obtained from the Center for Research in Security Prices (CRSP). Inspired by Kung and Schmid (2015), we run the following regression:

$$\Delta \ln C_{t,t+1} + \cdots + \Delta \ln C_{t+h-1,t+h} = \alpha + \beta M_t + \nu_{t,t+h}, \tag{56}$$

where $h = 1, \cdots, 5$ and $\Delta \ln C_{t+h-1,t+h}$ is the one-year log consumption growth from year $t+h-1$ to $t+h$. Panel B of Table 5 presents the results of projecting future consumption growth over horizons of one to five years on the misallocation measure M_t. The slope coefficients are negative and decreasing with horizons. Specifically, the slope coefficients are statistically significant in the last two columns of panel B, which correspond to consumption growth over horizons of three to five years. The R-squared monotonically increases from 0.069 to 0.173 when time horizon increases from $t \to t+3$ to $t \to t+5$.

We further run regressions similar to (56) using future log output growth as the dependent variable, where output is measured by CPI-deflated GDP per capita, obtained from BEA. Panel C of Table 5 presents the results of projecting future output growth over horizons of one to five years on the misallocation measure M_t. The slope coefficients are negative and decreasing with horizons, and become statistically significant in the last two columns with an R-squared of 0.067 and 0.123, respectively. These empirical findings are robust for an alternative sample period from 1970 to 2016 (see Table OA.3 in Online Appendix C.1).

Taken together, we find evidence that the aggregate growth rates of consumption and output can be predicted by our misallocation measure M_t over long horizons. Our findings lend empirical support to the notion of misallocation-driven low-frequency variation in growth, consistent with the implications of the model. Our model thus helps rationalize and identify misallocation as an economic source of long-run risks in the data.

4.3 Cross-Sectional Evidence

A core implication of our model is that fluctuations in misallocation drive the low-frequency component of consumption growth. We provide cross-sectional evidence to support this prediction.

Our starting point is the robust evidence found in the asset pricing literature (Bansal, Dittmar and Lundblad, 2005; Parker and Julliard, 2005; Hansen, Heaton and Li, 2008; Santos and Veronesi, 2010): the cash flows of value firms load more positively on accumulated consumption growth than those of growth firms. Given that a higher misallocation results in a lower consumption growth in our model, we expect the cash flows of value firms to load more negatively on misallocation if fluctuations in misallocation indeed drive the low-frequency variations in consumption growth, as implied by our model.

To test this prediction, we follow the empirical strategy of Santos and Veronesi (2010). In June of each year t, we sort firms into quintiles based on their book-to-market ratios $BE_{i,t-1}/ME_{i,t-1}$ in year $t-1$, where $BE_{i,t-1}$ is the book equity from Compustat and $ME_{i,t-1}$ is the market equity from CRSP. For each quintile portfolio, we compute the value-weighted return on equity (ROE) across all firms within the portfolio, where a firms' ROE is its income before extraordinary items divided by its common equity. Let $ROE^{p}_{t+j,j+1}$ denote the value-weighted ROE at year $t+j$ of the portfolio p, which was formed $j+1$ years earlier, i.e., in year $t-1$. We run a regression similar to the specification adopted by Santos and Veronesi (2010), except for including accumulated misallocation shocks as an additional independent variable:

$$\sum_{j=0}^{4} \rho^j ROE^{p}_{t+j,j+1} = \beta^p_0 + \beta^p_1 \sum_{j=0}^{4} \rho^j \Delta M_{t+j} + \beta^p_2 \sum_{j=0}^{4} \rho^j ROE^{Mkt}_{t+j} + \nu_t, \tag{57}$$

Table 6: Misallocation, R&D, and growth in the data.

$BE_{i,t-1}/ME_{i,t-1}$	Q1 (low)	Q2	Q3	Q4	Q5 (high)
β_1^p	-0.179	-0.242^{*}	-0.352^{***}	-0.373^{**}	-0.692^{**}
	[0.270]	[0.145]	[0.126]	[0.163]	[0.340]

Note: In June of each year t, we sort firms into quintiles on their book-to-market ratios $BE_{i,t-1}/ME_{i,t-1}$ in year $t-1$. For each quintile portfolio, we estimate β_1^p according to specification (57). The sample period is 1965-2016. Standard errors are reported in brackets. *, **, and *** indicate statistical significance at 10%, 5%, and 1%, respectively.

where $\rho = 0.95$ is a constant as in Santos and Veronesi (2010). The variable ΔM_t is the year-on-year changes in M_t and the variable ROE_t^{Mkt} is the ROE of the market portfolio. The coefficient of interest is β_1^p, which captures the loadings of accumulated ROE on accumulated misallocation shocks.

Table 6 presents the results. The accumulated ROE of firms with high book-to-market ratios (i.e., value firms in the quintile group 5 labeled as Q5) is significantly more negatively exposed to accumulated misallocation shocks than that of firms with low book-to-market ratios (i.e., growth firms in the quintile group 1 labeled as Q1). The loadings monotonically decreases from -0.179 to -0.692 as the book-to-market ratio increases from Q1 to Q5.

4.4 Asset Pricing Implications

Our model implies that the misallocation M_t plays a significant role in determining the SDF of representative agent through its effects on aggregate consumption growth. Thus, we should expect innovations in misallocation M_t, as a macroeconomic risk factor, to explain cross-sectional asset returns in the data. In this subsection, we study the cross-sectional asset pricing implications of misallocation to lend further support to our model.

Specifically, we study whether our empirical misallocation measure M_t is a risk factor significantly priced in the cross section of standard test assets, including 25 size-sorted and book-to-market-sorted portfolios, 10 momentum-sorted portfolios, and 6 maturity-sorted Treasury bond portfolios. For each asset i, we estimate the factor loadings using the following time-series regression:

$$R_{i,t}^e = c_i + \sum_k \beta_{i,k} f_{k,t} + \nu_{i,t}, \tag{58}$$

where $R^e_{i,t} = R_{i,t} - R_{f,t}$ is the excess return of asset i over the risk-free rate and $f_{k,t}$ represents risk factor k. We then estimate the cross-sectional price of risk associated with the factors $f_{k,t}$ by running a cross-sectional regression of time-series average excess returns, $\mathbb{E}[R^e_{i,t}]$, on risk factor exposures estimated in equation (58) as follows,

$$\mathbb{E}[R^e_{i,t}] = \alpha + \sum_k \widehat{\beta}_{i,k}\lambda_k + \epsilon_i, \tag{59}$$

where the estimated $\widehat{\lambda}_k$ is the price of risk for factor k and $\widehat{\alpha}$ is the average cross-sectional pricing error or zero-beta rate.

The above estimation procedure is implemented using different linear factor models. The results are presented in Table 7 and visualized in Figures 3 and 4. As a benchmark, column (1) of Table 7 reports the results of CAPM, which includes market excess returns as the single risk factor. It clearly shows that the exposure to market risk cannot explain the spread in average returns across portfolios. The cross-sectional intercept is statistically significant and the factor price of risk is statistically insignificant. The pricing errors are large, with a high total MAPE of 2.76% and a low adjusted R-squared of 0.30. column (2) of Table 7 presents the results based on a two-factor model that includes the year-on-year changes in the empirical misallocation measure, ΔM_t, as an additional risk factor. The price of risk for ΔM_t is -0.12, which is negative and statistically significant as predicted by our model.[16] Relative to CAPM, the R-squared increases significantly to 0.53 and the total MAPE declines significantly to 1.82%. The test assets are lined up very close to the 45-degree line in the two factor model (panel B of Figure 3), which is in sharp contrast to the prediction of CAPM (panel A of Figure 3).

As another benchmark, column (3) of Table 7 presents the results of Fama-French three-factor (FF3) models. Comparing columns (2) and (3) of Table 7, the FF3 model achieves a higher R-squared of 0.62. However, the two-factor model with market returns and the misallocation factor ΔM_t has a lower total MAPE. In addition, we break up the MAPE by asset class. In the cross section of 25 size-sorted and book-to-market-sorted portfolios, the FF3 model generates a lower MAPE of 1.30% while the two-factor model generates a MAPE of 1.77%. This is not surprising given that the excess returns of small caps over big caps and of value stocks over growth stocks

[16] The magnitude of the price of risk for ΔM_t does not represent the risk premium of ΔM_t because the misallocation factor ΔM_t does not lie in the space of excess returns.

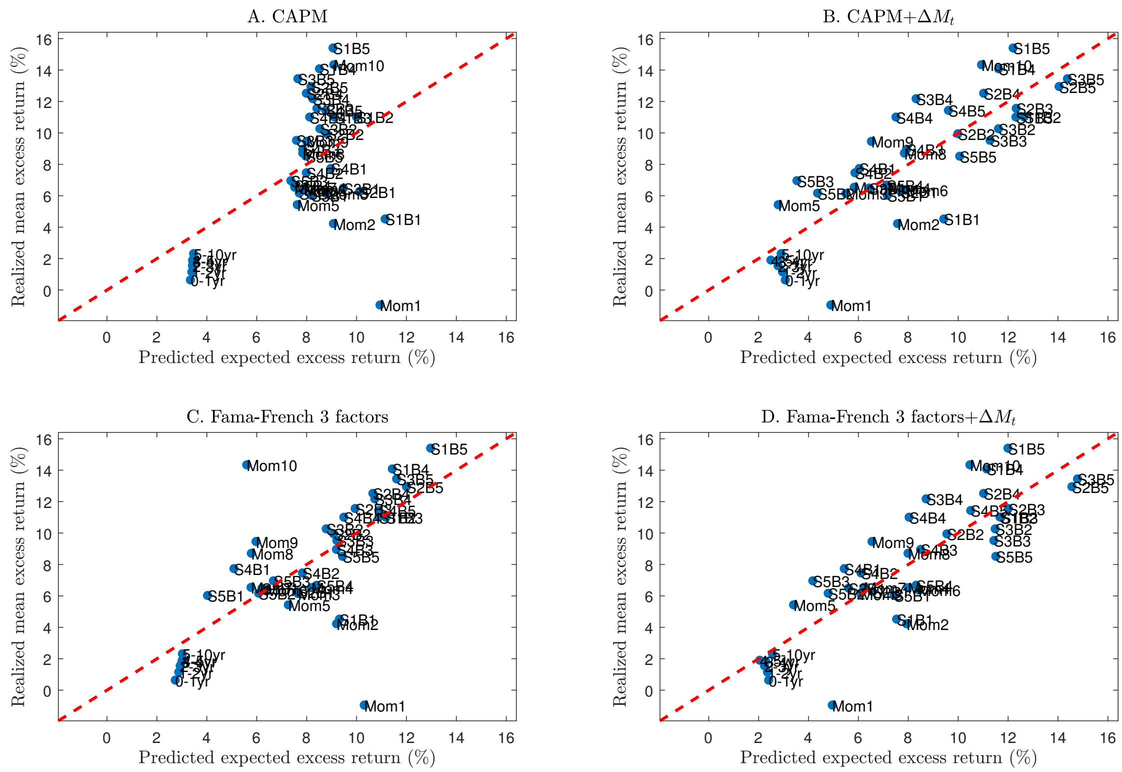
Table 7: Portfolio returns and model fit.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Mkt	$Mkt, \Delta M$	FF	$FF, \Delta M$	Mkt, CG	$Mkt, \Delta M, CG$	FF, CG	$FF, \Delta M, CG$
Panel A: Prices of risk								
Intercept	3.34***	2.96***	2.67***	2.25***	2.42***	2.83***	2.31***	2.18***
t-FM	[3.60]	[3.20]	[3.99]	[3.55]	[2.59]	[3.05]	[3.51]	[3.37]
t-Shanken	[3.47]	[1.46]	[3.65]	[1.60]	[1.72]	[1.50]	[2.57]	[1.55]
Mkt	4.92*	3.82	3.66	4.86*	4.63	3.88	4.59*	5.05**
t-FM	[1.73]	[1.36]	[1.42]	[1.91]	[1.64]	[1.38]	[1.80]	[1.99]
t-Shanken	[1.28]	[0.58]	[0.98]	[0.79]	[0.94]	[0.62]	[1.08]	[0.83]
ΔM		−0.12***		−0.12***		−0.10***		−0.11***
t-FM		[−4.66]		[−5.74]		[−4.52]		[−5.34]
t-Shanken		[−2.10]		[−2.54]		[−2.20]		[−2.41]
SMB			3.01	2.40			2.61	2.55
t-FM			[1.54]	[1.23]			[1.35]	[1.31]
t-Shanken			[1.05]	[0.52]			[0.80]	[0.55]
HML			4.36**	4.05**			4.40**	4.18**
t-FM			[2.16]	[2.00]			[2.18]	[2.07]
t-Shanken			[1.46]	[0.84]			[1.30]	[0.87]
CG					0.02***	0.01	0.02***	0.01
t-FM					[3.32]	[0.67]	[4.06]	[1.07]
t-Shanken					[2.13]	[0.32]	[2.71]	[0.48]
Panel B: Test diagnostics								
Total MAPE	2.76	1.82	1.90	1.70	2.03	1.78	1.95	1.71
Size and B/M 25	2.77	1.77	1.30	1.68	1.50	1.68	1.39	1.68
Momentum 10	3.30	2.62	3.72	2.30	3.59	2.34	3.69	2.33
Bond 6	1.85	1.26	1.36	0.74	1.61	1.28	1.43	0.78
Adjusted R-squared	0.30	0.53	0.62	0.68	0.55	0.60	0.63	0.69

Note: This table presents pricing results for 41 test assets, including 25 size-sorted and book-to-market-sorted portfolios, 10 momentum-sorted portfolios, and 6 maturity-sorted Treasury bond portfolios. Each model is estimated using equation (59). Mkt is the market's excess return over the risk-free rate. ΔM is the misallocation factor, which is the year-on-year changes in the empirical misallocation measure M_t. SMB and HML are the two factors in the FF3 model, capturing the excess returns of small caps over big caps and of value stocks over growth stocks, repectively. Panel A reports the prices of risk with Fama-McBeth and Shanken t-statistics. Panel B reports test diagnostics, including MAPE and R-squared. All numbers are in annualized percentage unit. The sample is yearly and spans the period from 1965 to 2016. *, **, and *** indicate statistical significance according to t-FM at 10%, 5%, and 1%, respectively.

are the two factors in the FF3 model. We find that the two-factor model outperforms the FF3 model in terms of total MAPE mainly because it generates lower MAPE for the 10 momentum-sorted portfolios. Specifically, in this cross section, the two-factor model implies a MAPE of 2.62% while the FF3 model implies a MAPE of 3.72%. It is well known that the FF3 model has a poor explanatory power for momentum-sorted portfolio returns. The cross-sectional fit is clearly displayed in panels B and C of Figure 3, which shows that the two-factor model outperforms the FF3 model mainly due to the improved fit for momentum-sorted portfolios. In column (4) of Table 7, we further include the misallocation factor ΔM_t to the FF3 model to construct a four-factor model. Compared with the FF3 model, the cross-sectional fit further improves as shown by the lower total MAPE and higher adjusted R-squared in the four-factor model. The improvement is mainly due to improved explanatory power for momentum-sorted portfolio.

Our theory suggests that the main channel through which the low-frequency component of consumption growth prices assets is the persistent variation in misallocation. Thus, we expect long-run expected consumption growth to have little explanatory power for portfolio returns if the misallocation factor ΔM_t is already included in the cross-sectional regression. Following Parker and Julliard (2005), we use accumulated future consumption growth to approximate long-run expected consumption growth. Column (5) of Table 7 and panel A of Figure 4 shows that the two-factor model with market returns and accumulated future consumption growth can fit the returns of our test portfolios well (R-squared$= 0.55$). In column (6) of Table 7 and panel B of Figure 4, we augment this two-factor model with the misallocation factor ΔM_t to construct a three-factor model. We find that the relation between realized mean excess returns and predicted mean excess returns across our test portfolios stays almost unchanged, implying that expected consumption growth and misallocation are indeed similarly priced in the cross section of test assets. However, the coefficient on accumulated future consumption growth becomes statistically insignificant after including ΔM_t as a factor, which has a statistically significant coefficient. Similar patterns are shown in columns (7) and (8) of Table 7 and panels C and D of Figure 4, when we include the misallocation factor ΔM_t in a four-factor model that contains the Fama-French three factors and accumulated future consumption growth.



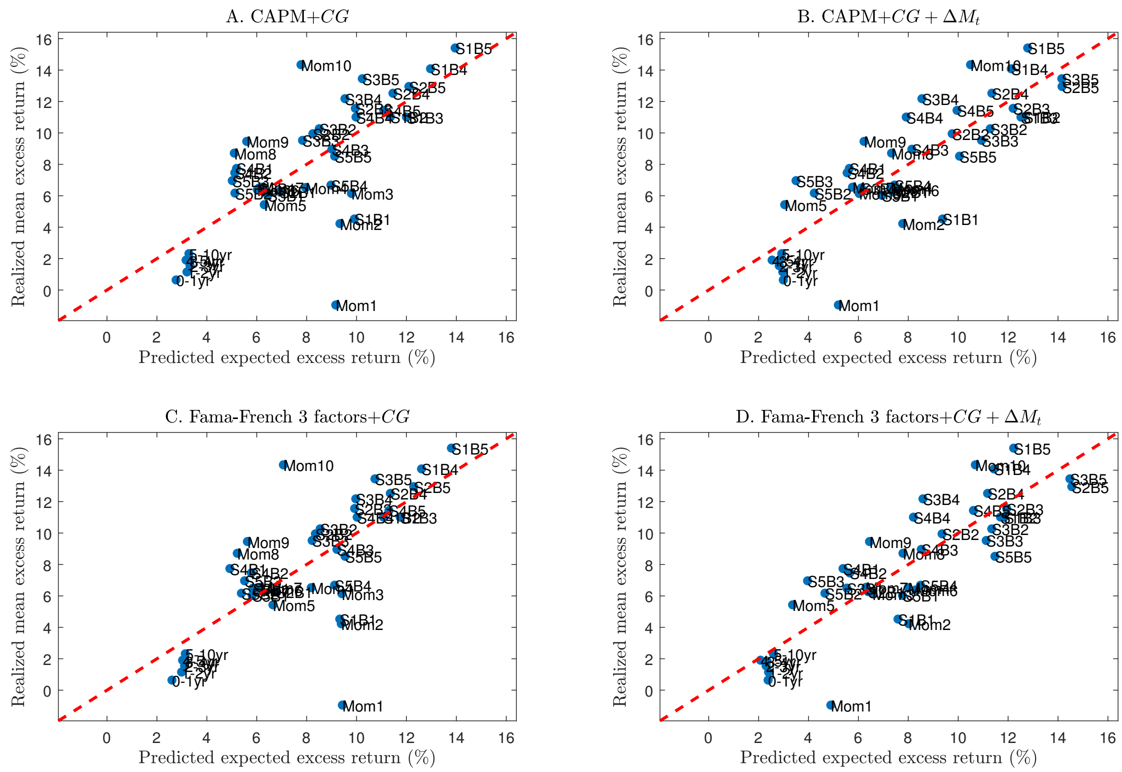
Note: This figure plots the realized mean excess returns of 35 equity portfolios (25 size-sorted and book-to-market-sorted portfolios and 10 momentum-sorted portfolios) and 6 maturity-sorted Treasury bond portfolios against the expected excess returns predicted by various linear factor asset pricing models. The sample is yearly and spans the period from 1965 to 2016.

Figure 3: Realized versus predicted mean excess returns in factor models with M_t.

4.5 Empirical Test on the Model's Core Mechanism

The core mechanism of our model is that misallocation drives long-run growth through its impact on R&D. In this subsection, we provide evidence for this mechanism by examining industry-level responses to a policy shock that alleviates firms' financial constraints.

The AJCA passed in 2004 allows domestic firms in the U.S. to repatriate their foreign profits at a tax rate of 5.25%, whereas the tax rate is 35% under the prior law. The passage of this law effectively relaxes the financial constraints of treated firms, significantly boosting the investments of firms that are financially constrained (Faulkender and Petersen, 2012). According to our model, relaxed financial constraints would lead to lower misallocation, providing firms more incentive to conduct R&D. To



Note: This figure plots the realized mean excess returns of 35 equity portfolios (25 size-sorted and book-to-market-sorted portfolios and 10 momentum-sorted portfolios) and 6 maturity-sorted Treasury bond portfolios against the expected excess returns predicted by various linear factor asset pricing models. The sample is yearly and spans the period from 1965 to 2016.

Figure 4: Realized versus predicted mean excess returns in factor models with M_t and accumulated future consumption growth.

test this prediction, we estimate the impact of AJCA on industry-level misallocation and R&D expenditure by exploiting industries' differential exposure to AJCA using a DID method.

Specifically, we construct industry-level measures for misallocation, R&D-capital ratio, and AJCA exposure using U.S. Compustat data. We use three-digit SIC codes (SIC3) to define industries. For each industry, we construct the industry-level misallocation following the procedures described in Section 4.1, except for running regression (54) based on firms within each industry. The industry-level R&D-capital ratio is constructed as the ratio of the total R&D expenditure to total capital of firms within the industry, where a firm's capital is measured by its net property, plant and equipment. To capture an industry's exposure to AJCA, we construct an industry-level measure for foreign business intensity, which is the average proportion of pre-tax income from abroad during the 3-year period prior to AJCA (i.e., from 2001 to 2003) across all firms within each industry. We consider industries with foreign business intensity above 30% as treated industries and the other industries as untreated industries. Treated industries are matched with untreated industries using the nearest neighbor matching method (see Online Appendix C.2 for details) based on eight industry-level characteristics.[17] All industry-level characteristics are averaged over the 3-year period prior to AJCA.

We run the following regression using industry-year observations for the period of 2001-2007:

$$Y_{s,t} = \alpha Treat_s \times Post_{t \geq 2004} + \beta_1 Treat_s + \beta_2 Post_t + \epsilon_{s,t}, \tag{60}$$

where $Treat_s = 1$ if industry s is a treated industry and $Treat_s = 0$ otherwise. The variable $Post_{t \geq 2004}$ is a time indicator that equals one for years after 2004. The coefficient of interest is α, which estimates the average effect of AJCA on the outcome variable $Y_{s,t}$ of treated industries. Our interested outcome variables are industry-level misallocation ($M_{s,t}$) and R&D-capital ratio ($RD_{s,t}$). The estimated coefficients are presented in column (1) of panels A and B in Table 8. Our results indicate that AJCA results in significantly lower misallocation and higher R&D-capital ratio in treated industries.

Next, we estimate the effect of AJCA in each year for the period of 2000-2007 by running the following regression:

$$Y_{s,t} = \sum_{\tau=-4}^{3} \alpha_\tau Treat_s \times Year_t^\tau + \beta_1 Treat_s + \sum_{\tau=-4}^{3} \beta_{2,\tau} Year_t^\tau + \epsilon_{s,t}, \tag{61}$$

where $Year_t^\tau$ is an indicator variable that captures the time difference relative to year 2004, and it equals 1 if $t = 2004 + \tau$ and 0 otherwise. The coefficient α_τ estimates the impact of AJCA on the outcome variable $Y_{s,t}$ of treated industries τ year after (or before if $\tau < 0$) the year (i.e., 2004) in which this policy was implemented. The estimated impacts on industry-level misallocation and R&D-capital ratio are presented in columns (2) to (8) of panels A and B in Table 8, respectively, and visualized in Figure 5. The leading terms of the estimated treatment effects are close to 0 and

[17] The eight industry-level characteristics are the Herfindahl index computed using firms' market shares in terms of sales, average total sales of firms, mean and standard deviation of firms' profit margin, mean and standard deviation of firms' ROE, mean and standard deviation of firms' Tobin's Q. We construct a firm's (net) profit margin using its income before extraordinary items divided by its sales as in Dou, Ji and Wu (2021), and a firm's Tobin's Q as $Tobin\_Q_{i,t} = (total\_assets_{i,t} + market\_equity_{i,t} - book\_equity_{i,t}) / total\_asset_{i,t}$, following Gompers, Ishii and Metrick (2003).

Table 8: Impacts of AJCA on misallocation and R&D.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Industry-level misallocation							
α	α_{-4}	α_{-3}	α_{-2}	α_0	α_1	α_2	α_3
-2.936^{**}	-0.809	-0.590	0.685	-1.847^{**}	-2.673^{***}	-3.377^{***}	-3.713^{***}
[1.231]	[2.335]	[2.392]	[0.701]	[0.776]	[1.005]	[1.211]	[1.325]
Panel B: Industry-level R&D-capital ratio							
α	α_{-4}	α_{-3}	α_{-2}	α_0	α_1	α_2	α_3
0.009^{**}	0.000	-0.001	-0.003	0.004^{***}	0.009^{**}	0.012^{**}	0.007
[0.004]	0.005	[0.003]	[0.002]	[0.001]	[0.004]	[0.005]	[0.006]
Panel C: Industry-level R&D-capital ratio controlling for misallocation							
α	α_{-4}	α_{-3}	α_{-2}	α_0	α_1	α_2	α_3
0.002	-0.001	-0.004	-0.006^{*}	0.000	0.002	0.004	-0.004
[0.008]	0.006	[0.005]	[0.003]	[0.005]	[0.008]	[0.010]	[0.012]

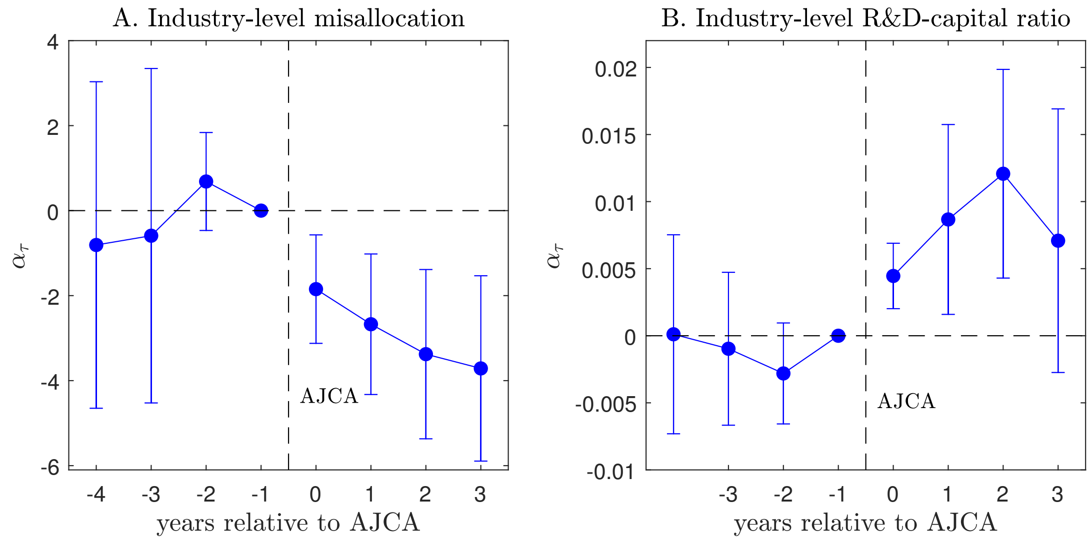
Note: Panel A estimates the impacts of AJCA, which alleviates the financial constraints of treated firms, on industry-level misallocation. Column (1) reports the estimated $\widehat{\alpha}$ in specification (60). Columns (2) to (8) report the estimated α_τ in specification (61) for $\tau = -4, -3, -2, 0, 1, 2, 3$. All coefficients are normalized relative to $\tau = -1$. Panel B estimates the impacts of AJCA on industry-level R&D-capital ratio. Panel C estimates the impacts of AJCA on industry-level R&D-capital ratio, controlling for industry-level misallocation. Standard errors are reported in brackets. *, **, and *** indicate statistical significance at 10%, 5%, and 1%, respectively.

statistically insignificant, suggesting that the parallel trend assumption is satisfied in the years before 2004. Regarding the three years after 2004, we estimate that AJCA has significant negative effects on industry-level misallocation and significant positive effects on industry-level R&D-capital ratio.

Furthermore, we provide evidence that the positive impact of AJCA on the industry-level R&D-capital ratio is achieved through the change in industry-level misallocation. Specifically, we modify specification (60) as follows:

$$RD_{s,t} = \alpha Treat_s \times Post_{t \geq 2004} + \beta_1 Treat_s + \beta_2 Post_t + \beta_3 M_{s,t} + \beta_4 Treat_s \times M_{s,t} + \epsilon_{s,t}, \tag{62}$$

which controls for industry-level misallocation $M_{s,t}$ and its interaction term with $Treat_s$. The estimated coefficient is presented in column (1) of panel C in Table 8, which indicates that AJCA no longer has a significant effect on R&D-capital ratio after controlling for industry-level misallocation.



Note: The solid lines visualize the empirical estimates in columns (2) to (8) of panels A and B in Table 8, respectively, in which all coefficients are normalized relative to $\tau = -1$. The vertical bars represent the corresponding 95% confidence intervals.

Figure 5: Impacts of AJCA on misallocation and R&D.

We further estimate the impact of AJCA on industry-level R&D-capital ratio in each year, controlling for industry-level misallocation, by running the following regression:

$$RD_{s,t} = \sum_{\tau=-4}^{3} \alpha_\tau Treat_s \times Year_t^\tau + \beta_1 Treat_s + \sum_{\tau=-4}^{3} \beta_{2,\tau} Year_t^\tau + \beta_3 M_{s,t} + \beta_4 Treat_s \times M_{s,t} + \epsilon_{s,t}, \tag{63}$$

Columns (2) to (8) of panel C in Table 8 report the estimates in each year, which indicate that the impacts of AJCA on industry-level R&D-capital ratio are statistically insignificant after controlling for industry-level misallocation.

Complementary to the DID specification, we also consider an alternative empirical specification and show that our findings above are robust. Specifically, we run the following cross-sectional regression:

$$\Delta \overline{RD}_s = \alpha Treat_s + \beta \overline{X}_s + \epsilon_s, \tag{64}$$

where the independent variable $\overline{X}_s$ is a vector of average industry-level characteristics over the 3-year period prior to AJCA, including industry s's Herfindahl index, average total sales of firms, mean and standard deviation of firms' profit margin, and mean and standard deviation of firms' ROE. The dependent variable $\Delta \overline{RD}_s$ is the change in industry-level average R&D-capital ratio between the 3-year period prior to AJCA and the 3-year period after AJCA, i.e., $\Delta\overline{RD}_s = \frac{1}{3}\sum_{t=2005}^{2007} RD_{s,t} - \frac{1}{3}\sum_{t=2001}^{2003} RD_{s,t}$. The estimated coefficient $\widehat{\alpha}$ in specification (64) is 0.010, with a p-value of 0.037, indicating that AJCA significantly increases the R&D-capital ratio of treated industries relative to untreated industries.

Moreover, we estimate the impact of AJCA on R&D-capital ratio, controlling for changes in industry-level misallocation by running the following cross-sectional regression:

$$\Delta\overline{RD}_s = \alpha Treat_s + \beta\overline{X}_s + \beta_M\Delta\overline{M}_s + \epsilon_s, \tag{65}$$

where $\overline{M}_s = \frac{1}{3}\sum_{t=2005}^{2007} M_{s,t} - \frac{1}{3}\sum_{t=2001}^{2003} M_{s,t}$. The estimated coefficient $\widehat{\alpha}$ in specification (65) is 0.009, with a p-value of 0.072, suggesting that AJCA no longer significantly increases the R&D-capital ratio of treated industries relative to untreated industries, after controlling for changes in industry-level misallocation. In other words, our results suggest that AJCA has positive impacts on treated industries' R&D-capital ratio mainly through the channel of reducing industry-level misallocation.

5 Conclusion

This paper provides a misallocation-based explanation for long-run consumption risk, a mechanism that quantitatively justifies many asset pricing moments. We develop a novel analytically tractable growth model with heterogenous firms, in which misallocation emerges as an endogenous state variable.

The model delineates the tight link between an economy's misallocation and its growth prospects. We show that short-run i.i.d. shocks that impact the economy's misallocation can have persistent effect on the economy's aggregate consumption growth, thereby generating endogenous long-run consumption risk. In the data, we construct a misallocation measure implied by the model and find evidence that the aggregate growth rates of consumption and output can be predicted by misallocation over long horizons. Moreover, as an asset pricing factor, misallocation explains the cross-sectional asset returns of standard test portfolios. By connecting the persistence of idiosyncratic productivity with the persistence of aggregate consumption growth, our model implies that long-run risk in aggregate consumption can be estimated based on granular firm-level data, which can potentially help address the issues of weak identification in the long-run risk literature (Chen, Dou and Kogan, 2022; Cheng, Dou and Liao, 2022).

References

Acemoglu, Daron, Ufuk Akcigit, Harun Alp, Nicholas Bloom, and William Kerr. 2018. "Innovation, Reallocation, and Growth." *American Economic Review*, 108(11): 3450–3491.

Adrian, Tobias, Erkko Etula, and Tyler Muir. 2014. "Financial intermediaries and the cross-section of asset returns." *The Journal of Finance*, 69(6): 2557–2596.

Ai, Hengjie. 2010. "Information Quality and Long-Run Risk: Asset Pricing Implications." *Journal of Finance*, 65(4): 1333–1367.

Ai, Hengjie, Jun E Li, Kai Li, and Christian Schlag. 2020. "The Collateralizability Premium." *Review of Financial Studies*, 33(12): 5821–5855.

Ai, Hengjie, Kai Li, and Fang Yang. 2020. "Financial intermediation and capital reallocation." *Journal of Financial Economics*, 138(3): 663–686.

Albuquerue, Rui, and Neng Wang. 2008. "Agency Conflicts, Investment, and Asset Pricing." *Journal of Finance*, 63(1): 1–40.

Algan, Yann, Olivier Allais, and Wouter J. Den Haan. 2008. "Solving heterogeneous-agent models with parameterized cross-sectional distributions." *Journal of Economic Dynamics and Control*, 32(3): 875–908.

Alvarez, Fernando, and Urban J. Jermann. 2004. "Using Asset Prices to Measure the Cost of Business Cycles." *Journal of Political Economy*, 112(6): 1223–1256.

Alvarez, Fernando, and Urban J. Jermann. 2005. "Using Asset Prices to Measure the Persistence of the Marginal Utility of Wealth." *Econometrica*, 73(6): 1977–2016.

Andrews, Donald W.K. 1983. "First Order Autoregressive Processes and Strong Mixing." Cowles Foundation for Research in Economics, Yale University Cowles Foundation Discussion Papers 664.

Angeletos, George-Marios. 2007. "Uninsured Idiosyncratic Investment Risk and Aggregate Saving." *Review of Economic Dynamics*, 10(1): 1–30.

Asker, John, Allan Collard-Wexler, and Jan De Loecker. 2014. "Dynamic Inputs and Resource (Mis)Allocation." *Journal of Political Economy*, 122(5): 1013–1063.

Banerjee, Abhijit, and A. Newman. 2003. "Inequality, Growth, and Trade Policy."

Banerjee, Abhijit V., and Benjamin Moll. 2010. "Why Does Misallocation Persist?" *American Economic Journal: Macroeconomics*, 2(1): 189–206.

Banerjee, Abhijit V., and Esther Duflo. 2005. "Growth Theory through the Lens of Development Economics." In *Handbook of Economic Growth*. Vol. 1 of *Handbook of Economic Growth*, , ed. Philippe Aghion and Steven Durlauf, Chapter 7, 473–552. Elsevier.

Bansal, Ravi, and Amir Yaron. 2004. "Risks for the long run: A potential resolution of asset pricing puzzles." *Journal of Finance*, 59(4): 1481–1509.

Bansal, Ravi, Robert F. Dittmar, and Christian T. Lundblad. 2005. "Consumption, Dividends, and the Cross Section of Equity Returns." *The Journal of Finance*, 60(4): 1639–1672.

Barlevy, Gadi. 2004. "The Cost of Business Cycles Under Endogenous Growth." *American Economic Review*, 94(4): 964–990.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta. 2013. "Cross-Country Differences in Productivity: The Role of Allocation and Selection." *American Economic Review*, 103(1): 305–334.

Baxter, Marianne, and Robert G. King. 1999. "Measuring Business Cycles: Approximate Band-Pass Filters For Economic Time Series." *The Review of Economics and Statistics*, 81(4): 575–593.

Bentkus, V., F. Gotze, and A. Tikhomoirov. 1997. "Berry-Esseen bounds for statistics of weakly dependent samples." *Bernoulli*, 3(3): 329–349.

Brunnermeier, Markus K., and Yuliy Sannikov. 2017. "Macro, Money and Finance: A Continuous-Time Approach." In *Handbook of Macroeconomics*. 1497–1546. North-Holland.

Brunnermeier, Markus, Thomas Eisenbach, and Yuliy Sannikov. 2012. "Macroeconomics with Financial Frictions: A Survey." National Bureau of Economic Research, Inc NBER Working Papers 18102.

Buera, Francisco J., and Yongseok Shin. 2011. "Self-insurance vs. self-financing: A welfare analysis of the persistence of shocks." *Journal of Economic Theory*, 146(3): 845–862.

Buera, Francisco J., and Yongseok Shin. 2013. "Financial Frictions and the Persistence of History: A Quantitative Exploration." *Journal of Political Economy*, 121(2): 221–272.

Burns, Authur F., and Wesley C. Mitchell. 1946. "Measuring Busines Cycles."

Campbell, John Y., and Robert J. Shiller. 1988. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." *Review of Financial Studies*, 1(3): 195–228.

Chava, Sudheer, and Michael Roberts. 2008. "How Does Financing Impact Investment?

The Role of Debt Covenants." *Journal of Finance*, 63(5): 2085–2121.

Cheng, Xu, Winston Wei Dou, and Zhipeng Liao. 2022. "Macro-Finance Decoupling: Robust Evaluations of Macro Asset Pricing Models." *Econometrica*, 90(2): 685–713.

Chen, Hui, Winston Wei Dou, and Leonid Kogan. 2022. "Measuring "Dark Matter" in Asset Pricing Models." *Journal of Finance*, forthcoming.

Cochrane, John H. 2017. "Macro-Finance." *Review of Finance*, 21(3): 945–985.

Collin-Dufresne, Pierre, Michael Johannes, and Lars A. Lochstoer. 2016. "Parameter Learning in General Equilibrium: The Asset Pricing Implications." *American Economic Review*, 106(3): 664–98.

Comin, Diego, and Mark Gertler. 2006. "Medium-Term Business Cycles." *American Economic Review*, 96(3): 523–551.

Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross. 1985. "An Intertemporal General Equilibrium Model of Asset Prices." *Econometrica*, 53(2): 363–384.

Croce, Mariano Max, Thien T. Nguyen, and Steve Raymond. 2021. "Persistent Government Debt and Aggregate Risk Distribution." *Journal of Financial Economics*, forthcoming.

Dabla-Norris, Era, Yan Ji, Robert M. Townsend, and D. Filiz Unsal. 2021. "Distinguishing constraints on financial inclusion and their impact on GDP, TFP, and the distribution of income." *Journal of Monetary Economics*, 117(C): 1–18.

David, Joel, Lukas Schmid, and David Zeke. 2019. "Risk-Adjusted Capital Allocation and Misallocation."

DeAngelo, Harry, Linda DeAngelo, and Karen H. Wruck. 2002. "Asset liquidity, debt covenants, and managerial discretion in financial distress:: the collapse of L.A. Gear." *Journal of Financial Economics*, 64(1): 3–34.

Di Tella, Sebastian. 2017. "Uncertainty Shocks and Balance Sheet Recessions." *Journal of Political Economy*, 125(6): 2038–2081.

Dou, Winston, and Yan Ji. 2021. "External financing and customer capital: A financial theory of markups." *Management Science*, 67(9): 5569–5585.

Dou, Winston Wei. 2017. "Embrace or Fear Uncertainty:Growth Options, Limited Risk Sharing, and Asset Prices."

Dou, Winston Wei, Leonid Kogan, and Wei Wu. 2021. "Common Fund Flows: Flow Hedging and Factor Pricing." Working Papers.

Dou, Winston Wei, Xiang Fang, Andrew W. Lo, and Harald Uhlig. 2020*a*. "Macro-Finance Models with Nonlinear Dynamics." Becker Friedman Institute for Economics at University of Chicago Working Papers.

Dou, Winston Wei, Yan Ji, and Wei Wu. 2021. "Competition, profitability, and discount rates." *Journal of Financial Economics*, 140(2): 582–620.

Dou, Winston, Yan Ji, David Reibstein, and Wei Wu. 2020*b*. "Inalienable customer capital, corporate liquidity, and stock returns." *Journal of Finance*, forthcoming.

Duffie, Darrell, and Larry G. Epstein. 1992*a*. "Asset pricing with stochastic differential utility." *Review of Financial Studies*, 5(3): 411–36.

Duffie, Darrell, and Larry G. Epstein. 1992*b*. "Stochastic differential utility." *Econometrica*, 60(2): 353–94.

Ehouarne, Cedric, Lars Kuehn, and David Schreindorfer. 2017. "Misallocation Cycles."

Eisfeldt, Andrea L., and Adriano A. Rampini. 2006. "Capital reallocation and liquidity." *Journal of Monetary Economics*, 53(3): 369–399.

Eisfeldt, Andrea L., and Adriano A. Rampini. 2008*a*. "Leasing, Ability to Repossess, and Debt Capacity." *The Review of Financial Studies*, 22(4): 1621–1657.

Eisfeldt, Andrea L., and Adriano A. Rampini. 2008*b*. "Managerial incentives, capital reallocation, and the business cycle." *Journal of Financial Economics*, 87(1): 177 – 199.

Eisfeldt, Andrea L., and Yu Shi. 2018. "Capital Reallocation." *Annual Review of Financial Economics*, 10: 361–386.

Epstein, Larry G., and Stanley E. Zin. 1989. "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework." *Econometrica*, 57(4): 937–969.

Faulkender, Michael, and Mitchell Petersen. 2012. "Investment and Capital Constraints: Repatriations Under the American Jobs Creation Act." *Review of Financial Studies*, 25(11): 3351–3388.

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2008. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review*, 98(1): 394–425.

Fuchs, William, Brett Green, and Dimitris Papanikolaou. 2016. "Adverse selection, slow-moving capital, and misallocation." *Journal of Financial Economics*, 120(2): 286–308.

Gabaix, Xavier. 2007. "Linearity-Generating Processes: A Modelling Tool Yielding Closed Forms for Asset Prices." National Bureau of Economic Research, Inc NBER Working Papers 13430.

Gabaix, Xavier. 2012. "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance." *The Quarterly Journal of Economics*, 127(2): 645–700.

Garleanu, Gârleanu B., Stavros Panageas, and Jianfeng Yu. 2012. "Technological Growth and Asset Pricing." *Journal of Finance*, 67(4): 1265–1292.

Gârleanu, Nicolae B., and Stavros Panageas. 2020. "Heterogeneity and Asset Prices: A Different Approach." National Bureau of Economic Research, Inc.

Gârleanu, Nicolae, Leonid Kogan, and Stavros Panageas. 2012. "Displacement risk and asset returns." *Journal of Financial Economics*, 105(3): 491–510.

Gertler, Mark, and Nobuhiro Kiyotaki. 2010. "Chapter 11 - Financial Intermediation and Credit Policy in Business Cycle Analysis." In . Vol. 3 of *Handbook of Monetary Economics*, , ed. Benjamin M. Friedman and Michael Woodford, 547 – 599. Elsevier.

Gertler, Mark, and Peter Karadi. 2011. "A model of unconventional monetary policy." *Journal of Monetary Economics*, 58(1): 17 – 34.

Gomes, João., Leonid Kogan, and Motohiro Yogo. 2009. "Durability of Output and Expected Stock Returns." *Journal of Political Economy*, 117(5): 941–986.

Gompers, Paul, Joy Ishii, and Andrew Metrick. 2003. "Corporate Governance and Equity Prices." *The Quarterly Journal of Economics*, 118(1): 107–156.

Gourio, Francois. 2012. "Disaster Risk and Business Cycles." *American Economic Review*, 102(6): 2734–2766.

Hall, Robert E., and Dale W. Jorgenson. 1969. "Tax Policy and Investment Behavior: Reply and Further Results." *The American Economic Review*, 59(3): 388–401.

Hansen, Lars Peter, John C. Heaton, and Nan Li. 2008. "Consumption Strikes Back? Measuring Long-Run Risk." *Journal of Political Economy*, 116(2): 260–302.

He, Zhiguo, Bryan Kelly, and Asaf Manela. 2017. "Intermediary asset pricing: New evidence from many asset classes." *Journal of Financial Economics*, 126(1): 1 – 35.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *The Quarterly Journal of Economics*, 124(4): 1403–1448.

Jermann, Urban, and Vincenzo Quadrini. 2012. "Macroeconomic Effects of Financial Shocks." *American Economic Review*, 102(1): 238–271.

Ji, Yan, Songyuan Teng, and Robert M. Townsend. 2021. "Branch Expansion versus Digital Banking: The Dynamics of Growth and Inequality in a Spatial Equilibrium Model." NBER Working Papers.

Jones, Charles I. 1995. "R&D-Based Models of Economic Growth." *Journal of Political Economy*, 103(4): 759–784.

Jones, Charles I. 2011. "Intermediate Goods and Weak Links in the Theory of Economic Development." *American Economic Journal: Macroeconomics*, 3(2): 1–28.

Jones, Charles I. 2013. "Misallocation, Economic Growth, and Input-Output Economics." , ed. Daron Acemoglu, Manuel Arellano and Eddie Dekel Vol. II of *Advances in Economics and Econometrics*, 419–454. Cambridge University Press.

Jorgenson, Dale. 1963. "Capital Theory and Investment Behavior." *American Economic Review*, 53(2): 247–259.

Kaltenbrunner, Georg, and Lars A. Lochstoer. 2010. "Long-Run Risk through Consumption Smoothing." *Review of Financial Studies*, 23(8): 3190–3224.

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

Kogan, Leonid. 2001. "An equilibrium model of irreversible investment." *Journal of Financial Economics*, 62(2): 201–245.

Kogan, Leonid. 2004. "Asset prices and real investment." *Journal of Financial Economics*, 73(3): 411–431.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman. 2017. "Technological Innovation, Resource Allocation, and Growth*." *The Quarterly Journal of Economics*, 132(2): 665–712.

Kogan, Leonid, Dimitris Papanikolaou, and Noah Stoffman. 2020. "Left Behind: Creative Destruction, Inequality, and the Stock Market." *Journal of Political Economy*, 128(3): 855–906.

Kreps, David M., and Evan L. Porteus. 1978. "Temporal resolution of uncertainty and dynamic choice theory." *Econometrica*, 46(1): 185–200.

Krusell, Per, and Anthony A. Smith. 1998*a*. "Income and Wealth Heterogeneity in the Macroeconomy." *Journal of Political Economy*, 106(5): 867–896.

Krusell, Per, and Anthony A. Smith. 1998*b*. "Income and Wealth Heterogeneity in the Macroeconomy." *Journal of Political Economy*, 106(5): 867–896.

Kung, Howard, and Lukas Schmid. 2015. "Innovation, Growth, and Asset Prices." *Journal of Finance*, 70(3): 1001–1037.

Lambrecht, Bart M., and Stewart C. Myers. 2008. "Debt and managerial rents in a real-options model of the firm." *Journal of Financial Economics*, 89(2): 209 – 231.

Lambrecht, Bart M., and Stewart C. Myers. 2012. "A Lintner Model of Payout and Managerial Rents." *The Journal of Finance*, 67(5): 1761–1810.

Lanteri, Andrea, and Adriano A. Rampini. 2021. "Constrained-Efficient Capital Reallocation." National Bureau of Economic Research, Inc NBER Working Papers 28384.

Li, Kai, and Chi-Yang Tsou. 2021. "Leasing as a risk-sharing mechanism." Working papers.

Li, Kai, and Yiming Xu. 2021. "Leasing as a Mitigation Channel of Capital Misallocation." Working papers.

Lucas, Robert E. 1987. *Models of business cycles.* Oxford: Basil Blackwell.

Midrigan, Virgiliu, and Daniel Yi Xu. 2014. "Finance and Misallocation: Evidence from Plant-Level Data." *American Economic Review*, 104(2): 422–458.

Moll, Benjamin. 2014. "Productivity losses from financial frictions: Can self-financing undo capital misallocation?" *American Economic Review*, 104(10): 3186–3221.

Müller, Ulrich K., and Mark W. Watson. 2008. "Testing Models of Low-Frequency Variability." *Econometrica*, 76(5): 979–1016.

Müller, Ulrich K., and Mark W. Watson. 2018. "Long-Run Covariability." *Econometrica*, 86(3): 775–804.

Myers, Stewart C. 2000. "Outside Equity." *The Journal of Finance*, 55(3): 1005–1037.

Opp, Marcus M., Christine A. Parlour, and Johan Walden. 2014. "Markup cycles, dynamic misallocation, and amplification." *Journal of Economic Theory*, 154(C): 126–161.

Parker, Jonathan A., and Christian Julliard. 2005. "Consumption risk and the cross section of expected returns." *Journal of Political Economy*, 113(1): 185–222.

Pástor, Ľuboš, and Pietro Veronesi. 2012. "Uncertainty about Government Policy and Stock Prices." *The Journal of Finance*, 67(4): 1219–1264.

Pástor, L'uboš, and Robert Stambaugh. 2003. "Liquidity Risk and Expected Stock Returns." *Journal of Political Economy*, 111(3): 642–685.

Peters, Michael. 2020. "Heterogeneous mark-ups, growth and endogenous misallocation." *Econometrica*, 88(5): 2037–2073.

Piazzesi, Monika, Martin Schneider, and Selale Tuzel. 2007. "Housing, consumption and asset pricing." *Journal of Financial Economics*, 83(3): 531–569.

Prilmeier, Robert. 2017. "Why do loans contain covenants? Evidence from lending relationships." *Journal of Financial Economics*, 123(3): 558–579.

Rampini, Adriano A., and S. Viswanathan. 2013. "Collateral and Capital Structure." *Journal of Financial Economics*, 109(2): 466 – 492.

Rauh, Joshua D., and Amir Sufi. 2011. "Explaining Corporate Capital Structure: Product Markets, Leases, and Asset Similarity*." *Review of Finance*, 16(1): 115–155.

Reiter, Michael. 2009. "Solving heterogeneous-agent models by projection and perturbation." *Journal of Economic Dynamics and Control*, 33(3): 649–665.

Restuccia, Diego, and Richard Rogerson. 2008. "Policy Distortions and Aggregate Productivity with Heterogeneous Plants." *Review of Economic Dynamics*, 11(4): 707–720.

Roberts, Michael R., and Amir Sufi. 2009. "Renegotiation of financial contracts: Evidence from private credit agreements." *Journal of Financial Economics*, 93(2): 159–184.

Romer, Paul M. 1987. "Growth Based on Increasing Returns Due to Specialization." *American Economic Review*, 77(2): 56–62.

Romer, Paul M. 1990. "Endogenous Technological Change." *Journal of Political Economy*, 98(5): 71–102.

Santos, Tano, and Pietro Veronesi. 2010. "Habit formation, the cross section of stock returns and the cash-flow risk puzzle." *Journal of Financial Economics*, 98(2): 385–413.

Steen, N. M., G. D. Byrne, and E. M. Gelbard. 1969. "Gaussian Quadratures for the Integrals." *Mathematics of Computation*, 23(107): 661–671.

Sufi, Amir. 2009. "Bank Lines of Credit in Corporate Finance: An Empirical Analysis." *Review of Financial Studies*, 22(3): 1057–1088.

Sundaresan, M. 1984. "Consumption and Equilibrium Interest Rates in Stochastic Production Economies." *The Journal of Finance*, 39(1): 77–92.

Tikhomirov, A. N. 1980. "On the Convergence Rate in The Central Limit Theorem for Weakly Dependent Random Variables." *Theory of Probability and Its Applications*, XXV(4): 790–809.

van Binsbergen, Jules H., and Christian C. Opp. 2019. "Real Anomalies." *Journal of Finance*, 74(4): 1659–1706.

Weil, Philippe. 1990. "Nonexpected utility in macroeconomics." *Quarterly Journal of Economics*, 105(1): 29–42.

Whited, Toni M., and Jake Zhao. 2021. "The Misallocation of Finance." *Journal of Finance*, forthcoming.

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

Winberry, Thomas. 2021. "Lumpy Investment, Business Cycles, and Stimulus Policy." *American Economic Review*, 111(1): 364–96.

Yogo, Motohiro. 2006. "A Consumption-Based Explanation of Expected Stock Returns." *The Journal of Finance*, 61(2): 539–580.

Online Appendix

A Proofs

A.1 Proof of Proposition 1

To ensure that shareholders do not intervene, the manager pays a dividend flow of $\zeta a_{i,t}\mathrm{d}t$ to shareholders over $[t, t+\mathrm{d}t)$ where ζ is to be determined. Thus, firm i's total dividend payment is $(\tau+\zeta)a_{i,t}\mathrm{d}t$ over $[t, t+\mathrm{d}t)$, which includes the rents paid to the manager and the dividends paid to shareholders. If the manager follows this payout policy consistently, shareholders will be willing to defer intervention continuously. Payouts and rents evolve in lockstep.

We now derive the relation between ζ, ρ, and τ. Per our discussion in Online Appendix A.2, the manager's value is proportional to the firm's capital, given by $\xi_{i,t}a_{i,t}$, where $\xi_{i,t}$ depends on the firm's idiosyncratic productivity $z_{i,t}$ and the aggregate state of the economy. If shareholders do not intervene, they receive a dividend payment that is a fraction ζ/τ of the manager's private benefits. Thus, shareholders' value is $(\zeta/\tau)\xi_{i,t}a_{i,t}$. If shareholders intervene, the firm's value will drop to $(1-\tau/\rho)\xi_{i,t}a_{i,t}$ due to the loss of capital. However, because shareholders now are also managers, they will have the claim to all dividends, which generate a value of $(1+\zeta/\tau)(1-\tau/\rho)\xi_{i,t}a_{i,t}$. Thus, the manager chooses the intensity ζ of dividends to shareholders such that shareholders are indifferent about having an intervention or not:

$$\frac{\zeta}{\tau}\xi_{i,t}a_{i,t} = (1+\zeta/\tau)(1-\tau/\rho)\xi_{i,t}a_{i,t}, \tag{66}$$

which implies $\zeta = (1-\tau/\rho)\tau/[1-(1-\tau/\rho)]$. The firm's total dividend payout ratio is

$$\tau + \zeta = \rho. \tag{67}$$

A.2 Proof of Lemma 1

Given capital $k_{i,t} = a_{i,t} + \widehat{a}_{i,t}$, utilization intensity $u_{i,t}$, and intermediate composite $x_{i,t}$, firm i solves a static maximization problem when choosing $\ell_{i,t}$ and $x_{i,j,t}$. Taking the first-order condition with respect to $\ell_{i,t}$ on the right-hand side of equation (19) of the main text, we obtain the optimal labor demand:

$$\ell_{i,t} = \left[\frac{w_t}{(1-\alpha)(1-\varepsilon)(z_{i,t}u_{i,t}k_{i,t})^{\alpha(1-\varepsilon)} x_{i,t}^{\varepsilon}} \right]^{\frac{1}{(1-\alpha)(1-\varepsilon)-1}}. \tag{68}$$

Substituting equations (2), (16), and (19) of the main text and equation (68) into equation (3) of the main text:

$$\begin{aligned} \mathrm{d}a_{i,t} = & -\int_0^{N_t} p_{j,t}x_{i,j,t}\mathrm{d}j\mathrm{d}t - u_{i,t}k_{i,t}\left(\delta_k \mathrm{d}t + \sigma_k \mathrm{d}W_t\right) + a_{i,t}(-\delta_a \mathrm{d}t + \sigma_a \mathrm{d}W_t) - r_{f,t}\widehat{a}_{i,t}\mathrm{d}t - \rho a_{i,t}\mathrm{d}t \\ & + [1-(1-\alpha)(1-\varepsilon)] \left[\frac{w_t}{(1-\alpha)(1-\varepsilon)} \right]^{\frac{(1-\alpha)(1-\varepsilon)}{(1-\alpha)(1-\varepsilon)-1}} (z_{i,t}u_{i,t}k_{i,t})^{\frac{\alpha(1-\varepsilon)}{1-(1-\alpha)(1-\varepsilon)}} x_{i,t}^{\frac{\varepsilon}{1-(1-\alpha)(1-\varepsilon)}}\mathrm{d}t. \end{aligned} \tag{69}$$

Taking the first-order condition with respect to $x_{i,j,t}$ in the right-hand side of equation (69), we derive firm i's optimal demand for intermediate goods $j \in [0, N_t]$:

$$x_{i,j,t} = \left[\frac{\varepsilon}{p_{j,t}} \left[\frac{w_t}{(1-\alpha)(1-\varepsilon)} \right]^{\frac{(1-\alpha)(1-\varepsilon)}{(1-\alpha)(1-\varepsilon)-1}} (z_{i,t}u_{i,t}k_{i,t})^{\frac{\alpha(1-\varepsilon)}{1-(1-\alpha)(1-\varepsilon)}} x_{i,t}^{1-\nu-\frac{\alpha(1-\varepsilon)}{1-(1-\alpha)(1-\varepsilon)}} \right]^{\frac{1}{1-\nu}}. \tag{70}$$

Substituting into equation (4) of the main text, we derive $x_{i,t}$:

$$x_{i,t} = \left(\frac{\varepsilon}{p_t} \right)^{\frac{1-(1-\alpha)(1-\varepsilon)}{\alpha(1-\varepsilon)}} \left[\frac{(1-\alpha)(1-\varepsilon)}{w_t} \right]^{\frac{1-\alpha}{\alpha}} z_{i,t}u_{i,t}k_{i,t}, \tag{71}$$

where the price index p_t is given by

$$p_t = \left(\int_0^{N_t} p_{j,t}^{\frac{\nu}{\nu-1}} \mathrm{d}j \right)^{\frac{\nu-1}{\nu}}. \tag{72}$$

Substituting equation (71) into (68), we obtain equation (24) of the main text. Substituting equation (71) into (70), we obtain (25). Substituting equation (71) and

$\int_0^{N_t} p_{j,t} x_{i,j,t} \mathrm{d}j = p_t x_{i,t}$ into (69), we obtain

$$\begin{aligned} \mathrm{d}a_{i,t} = & - u_{i,t} k_{i,t} \left(\delta_k \mathrm{d}t + \sigma_k \mathrm{d}W_t \right) + a_{i,t} (- \delta_a \mathrm{d}t + \sigma_a \mathrm{d}W_t) - r_{f,t} \widehat{a}_{i,t} \mathrm{d}t - \rho a_{i,t} \mathrm{d}t \\ & + \alpha (1-\varepsilon) \left(\frac{\varepsilon}{p_t} \right)^{\frac{\varepsilon}{\alpha(1-\varepsilon)}} \left[\frac{(1-\alpha)(1-\varepsilon)}{w_t} \right]^{\frac{1-\alpha}{\alpha}} z_{i,t} u_{i,t} k_{i,t} \mathrm{d}t. \end{aligned} \tag{73}$$

Thus, the manager's problem (18) can be simplified and characterized recursively as follows:

$$0 = \max_{\widehat{a}_{i,t}, u_{i,t}} \quad \tau a_{i,t} \mathrm{d}t + \mathbb{E}_t \left[\frac{\mathrm{d}\Lambda_t}{\Lambda_t} J_{i,t} + \mathrm{d}J_{i,t} + \frac{\mathrm{d}\Lambda_t}{\Lambda_t} \mathrm{d}J_{i,t} \right]. \tag{74}$$

subject to the budget constraint (73). Because the technology, budget constraint, and collateral constraint are all linear in $a_{i,t}$, the value $J_{i,t}$ is also linear in $a_{i,t}$ with the following form:

$$J_{i,t} \equiv J_t(a_{i,t}, z_{i,t}) = \xi_t(z_{i,t}) a_{i,t}, \tag{75}$$

where $\xi_{i,t} \equiv \xi_t(z_{i,t})$ captures the marginal value of capital to firm i's manager, which depends on the firm's idiosyncratic productivity $z_{i,t}$ and the aggregate state of the economy. Substituting equation (20) of the main text and equation (75) into (74), we obtain

$$\begin{aligned} 0 = \max_{\widehat{a}_{i,t}, u_{i,t}} \quad & \tau a_{i,t} \mathrm{d}t + \mathbb{E}_t \left[(-r_{f,t} \mathrm{d}t - \eta_t \mathrm{d}W_t) \xi_{i,t} a_{i,t} \right] \\ & + \mathbb{E}_t \left[(1 - r_{f,t} \mathrm{d}t - \eta_t \mathrm{d}W_t)(\mathrm{d}\xi_{i,t} a_{i,t} + \xi_{i,t} \mathrm{d}a_{i,t} + \mathrm{d}\xi_{i,t} \mathrm{d}a_{i,t}) \right]. \end{aligned} \tag{76}$$

The variable $\xi_{i,t}$ evolves as follows:

$$\frac{\mathrm{d}\xi_{i,t}}{\xi_{i,t}} = \mu_{\xi,i,t} \mathrm{d}t + \sigma_{\xi,i,t} \mathrm{d}W_t + \sigma_{w,i,t} \mathrm{d}W_{i,t}, \tag{77}$$

where $\mu_{\xi,i,t} \equiv \mu_{\xi,t}(z_{i,t})$, $\sigma_{\xi,i,t} \equiv \sigma_{\xi,t}(z_{i,t})$, and $\sigma_{w,i,t} \equiv \sigma_{w,t}(z_{i,t})$ are endogenously determined in equilibrium. Using equations (73) and (77), and the properties that $(\mathrm{d}W_t)^2 = \mathrm{d}t$, $\mathbb{E}_t[\mathrm{d}W_{i,t}] = \mathbb{E}_t[\mathrm{d}W_t] = \mathbb{E}_t[\mathrm{d}W_t \mathrm{d}W_{i,t}] = 0$, we obtain the following equations after omitting higher-order terms:

$$\mathbb{E}_t \left[(-r_{f,t} \mathrm{d}t - \eta_t \mathrm{d}W_t) \xi_{i,t} a_{i,t} \right] = - r_{f,t} a_{i,t} \xi_{i,t} \mathrm{d}t, \tag{78}$$

$$\mathbb{E}_t \left[(1 - r_{f,t} \mathrm{d}t - \eta_t \mathrm{d}W_t) \mathrm{d}\xi_{i,t} a_{i,t} \right] = \mu_{\xi,i,t} a_{i,t} \xi_{i,t} \mathrm{d}t - \eta_t \sigma_{\xi,i,t} a_{i,t} \xi_{i,t} \mathrm{d}t. \tag{79}$$

$$\mathbb{E}_t\left[(1-r_{f,t}\mathrm{d}t-\eta_t\mathrm{d}W_t)(\xi_{i,t}\mathrm{d}a_{i,t}+\mathrm{d}\xi_{i,t}\mathrm{d}a_{i,t})\right]=\mathbb{E}_t\left[(1+\sigma_{\xi,i,t}\mathrm{d}W_t-\eta_t\mathrm{d}W_t)\xi_{i,t}\mathrm{d}a_{i,t}\right]$$
$$=[\sigma_k(\eta_t-\sigma_{\xi,i,t})-\delta_k]u_{i,t}k_{i,t}\xi_{i,t}\mathrm{d}t-\sigma_a(\eta_t-\sigma_{\xi,i,t})\xi_{i,t}a_{i,t}\mathrm{d}t-r_{f,t}\widehat{a}_{i,t}\xi_{i,t}\mathrm{d}t-(\rho+\delta_a)a_{i,t}\xi_{i,t}\mathrm{d}t$$
$$+\alpha(1-\varepsilon)\left(\frac{\varepsilon}{p_t}\right)^{\frac{\varepsilon}{\alpha(1-\varepsilon)}}\left[\frac{(1-\alpha)(1-\varepsilon)}{w_t}\right]^{\frac{1-\alpha}{\alpha}}z_{i,t}u_{i,t}k_{i,t}\xi_{i,t}\mathrm{d}t. \tag{80}$$

Substituting equations (78), (79), (80) into (76), we obtain

$$0=\max_{\widehat{a}_{i,t},u_{i,t}}\quad \tau a_{i,t}\mathrm{d}t-r_{f,t}a_{i,t}\xi_{i,t}\mathrm{d}t+\mu_{\xi,i,t}a_{i,t}\xi_{i,t}\mathrm{d}t-\eta_t\sigma_{\xi,i,t}a_{i,t}\xi_{i,t}\mathrm{d}t$$
$$[\sigma_k(\eta_t-\sigma_{\xi,i,t})-\delta_k]u_{i,t}k_{i,t}\xi_{i,t}\mathrm{d}t-\sigma_a(\eta_t-\sigma_{\xi,i,t})\xi_{i,t}a_{i,t}\mathrm{d}t-r_{f,t}\widehat{a}_{i,t}\xi_{i,t}\mathrm{d}t-(\rho+\delta_a)a_{i,t}\xi_{i,t}\mathrm{d}t$$
$$+\alpha(1-\varepsilon)\left(\frac{\varepsilon}{p_t}\right)^{\frac{\varepsilon}{\alpha(1-\varepsilon)}}\left[\frac{(1-\alpha)(1-\varepsilon)}{w_t}\right]^{\frac{1-\alpha}{\alpha}}z_{i,t}u_{i,t}k_{i,t}\xi_{i,t}\mathrm{d}t. \tag{81}$$

Using $k_{i,t}=a_{i,t}+\widehat{a}_{i,t}$, we can see that maximizing equation (81) is essentially the same as maximizing

$$0=\max_{\widehat{a}_{i,t},u_{i,t}}\quad [\sigma_k(\eta_t-\sigma_{\xi,i,t})-\delta_k]u_{i,t}\widehat{a}_{i,t}\xi_{i,t}\mathrm{d}t-r_{f,t}\widehat{a}_{i,t}\xi_{i,t}\mathrm{d}t$$
$$+\alpha(1-\varepsilon)\left(\frac{\varepsilon}{p_t}\right)^{\frac{\varepsilon}{\alpha(1-\varepsilon)}}\left[\frac{(1-\alpha)(1-\varepsilon)}{w_t}\right]^{\frac{1-\alpha}{\alpha}}z_{i,t}u_{i,t}\widehat{a}_{i,t}\xi_{i,t}\mathrm{d}t. \tag{82}$$

Because a positive shock ($\mathrm{d}W_t>0$) increases misallocation through higher capital depreciation of productive firms, we have $\eta_t<0$ in equilibrium. Moreover, because $\xi_{i,t}$ is not affected by the manager's choice of $\widehat{a}_{i,t}$. The objective function (82) is linear in both $\widehat{a}_{i,t}$ and $u_{i,t}$. Thus, conditional on $u_{i,t}=1$, we can characterize the productivity cutoff $\underline{z}_t$ that makes the manager indifferent about leasing capital as follows:

$$\underline{z}_t\kappa_t=r_{f,t}+\delta_k+\sigma_k(\sigma_{\xi,t}(\underline{z}_t)-\eta_t), \tag{83}$$

where

$$\kappa_t=\alpha(1-\varepsilon)\left(\frac{\varepsilon}{p_t}\right)^{\frac{\varepsilon}{\alpha(1-\varepsilon)}}\left[\frac{(1-\alpha)(1-\varepsilon)}{w_t}\right]^{\frac{1-\alpha}{\alpha}}. \tag{84}$$

Because $r_{f,t}>0$ in equilibrium, it is clear that firms will optimally choose $u_{i,t}=1$

for $z_{i,t} \geq \underline{z}_t$.[18] The optimal leasing amount follows a bang-bang solution:

$$\widehat{a}_t(a,z) = \begin{cases} \lambda a, & z \geq \underline{z}_t \\ -a & z < \underline{z}_t \end{cases}, \tag{85}$$

which leads to the bang-bang solution in capital:

$$k_t(a,z) = \begin{cases} (1+\lambda)a, & z \geq \underline{z}_t \\ 0 & z < \underline{z}_t \end{cases}. \tag{86}$$

The optimal capacity utilization intensity is given by

$$u_t(z) = \begin{cases} 1, & z \geq \underline{z}_t \\ 0 & z < \underline{z}_t \end{cases}. \tag{87}$$

In fact, any utilization intensity $u_{i,t} \in [0,1]$ is optimal when $z_{i,t} < \underline{z}_t$ because $k_{i,t} = 0$. We set its value to zero without loss of generality.

A.3 Proof of Proposition 2

Define the productivity H_t of the final goods sector as

$$H_t = \left[\frac{1}{K_t} \int_{\underline{z}_t}^{\infty} \int_0^{\infty} z u_t(z) k_t(a,z) \varphi_t(a,z) \mathrm{d}k \mathrm{d}z \right]^{\alpha}, \tag{88}$$

Using equations (23), (32), and (38) of the main text and $k_t(a,z) = a + \widehat{a}_t(a,z)$, H_t can be written as

$$H_t = \left[(1+\lambda) \frac{A_t}{K_t} \int_{\underline{z}_t}^{\infty} z \omega_t(z) \mathrm{d}z \right]^{\alpha}. \tag{89}$$

Substituting equation (86) into the capital market-clearing condition (33) of the main text, we obtain

$$(1+\lambda) \int_{\underline{z}_t}^{\infty} \int_0^{\infty} a \varphi_t(a,z) \mathrm{d}k \mathrm{d}z = K_t. \tag{90}$$

Given the definition of capital share in (38)of the main text, the left-hand side of

[18] In other words, the (latent) cutoff productivity for $u_{i,t}$ is lower than the cutoff productivity $\underline{z}_t$ for $\widehat{a}_{i,t}$ when $r_{f,t} > 0$.

equation (90) can be simplified as

$$(1+\lambda)\int_{\underline{z}_t}^{\infty}\int_0^{\infty} a\varphi_t(a,z)\mathrm{d}k\mathrm{d}z = (1+\lambda)A_t\int_{\underline{z}_t}^{\infty}\omega_t(z)\mathrm{d}z = (1+\lambda)A_t(1-\Omega_t(\underline{z}_t)). \quad (91)$$

Thus, we have the following equation

$$(1+\lambda)[1-\Omega_t(\underline{z}_t)] = \frac{K_t}{A_t}, \quad (92)$$

which determines the equilibrium K_t/A_t. Substituting equation (92) into (89), we obtain

$$H_t = \left[\frac{\int_{\underline{z}_t}^{\infty} z\omega_t(z)\mathrm{d}z}{1-\Omega_t(\underline{z}_t)}\right]^{\alpha}. \quad (93)$$

Substituting equation (7) of the main text into equation (6) of the main text, we obtain

$$\pi_{j,t} = \max_{p_{j,t}} \quad (p_{j,t}-1)\left(\frac{p_{j,t}}{p_t}\right)^{\frac{1}{\nu-1}} X_t, \quad (94)$$

Taking the first-order condition, we obtain

$$p_{j,t} = \frac{1}{\nu} \text{ for all } j. \quad (95)$$

Substituting equation (95) into the price index (26) of the main text, we obtain

$$p_t = N_t^{\frac{\nu-1}{\nu}} p_{j,t} = N_t^{\frac{\nu-1}{\nu}}/\nu. \quad (96)$$

Substituting equation (24) of the main text into (35) of the main text and using (88), we obtain

$$\begin{aligned} L_t &= \left(\frac{\varepsilon}{p_t}\right)^{\frac{\varepsilon}{\alpha(1-\varepsilon)}}\left[\frac{(1-\alpha)(1-\varepsilon)}{w_t}\right]^{\frac{1}{\alpha}}\int_{\underline{z}_t}^{\infty}\int_0^{\infty} zu_t(z)k_t(a,z)\varphi_t(a,z)\mathrm{d}a\mathrm{d}z \\ &= \left(\frac{\varepsilon}{p_t}\right)^{\frac{\varepsilon}{\alpha(1-\varepsilon)}}\left[\frac{(1-\alpha)(1-\varepsilon)}{w_t}\right]^{\frac{1}{\alpha}} H_t^{\frac{1}{\alpha}}K_t. \end{aligned} \quad (97)$$

Substituting equation (96) into (97), we derive the equilibrium wage w_t:

$$w_t = (1-\alpha)(1-\varepsilon)(\varepsilon\nu)^{\frac{\varepsilon}{1-\varepsilon}} N_t^{\frac{(1-\nu)\varepsilon}{\nu(1-\varepsilon)}} H_t(K_t/L_t)^{\alpha}. \quad (98)$$

By definition, the aggregate output Y_t is

$$Y_t = \int_{\underline{z}_t}^{\infty} \int_0^{\infty} \left[(zu_t(z)k_t(a,z))^{\alpha} (\ell_t(a,z))^{1-\alpha} \right]^{1-\varepsilon} x_t(a,z)^{\varepsilon} \varphi_t(a,z) \mathrm{d}a \mathrm{d}z. \tag{99}$$

Substituting equations (24) and (27) of the main text into (99), we obtain

$$Y_t = \left(\frac{\varepsilon}{p_t} \right)^{\frac{\varepsilon}{\alpha(1-\varepsilon)}} \left[\frac{(1-\alpha)(1-\varepsilon)}{w_t} \right]^{\frac{1-\alpha}{\alpha}} \int_{\underline{z}_t}^{\infty} \int_0^{\infty} zu_t(z)k_t(a,z)\varphi_t(a,z) \mathrm{d}a \mathrm{d}z. \tag{100}$$

Further, substituting equations (88), (96) and (98) into the above equation, we obtain

$$\begin{aligned} Y_t =& (\varepsilon\nu)^{\frac{\varepsilon}{1-\varepsilon}} H_t N_t^{\frac{(1-\nu)\varepsilon}{\nu(1-\varepsilon)}} K_t^{\alpha} L_t^{1-\alpha} \\ =& (\varepsilon\nu)^{\frac{\varepsilon}{1-\varepsilon}} H_t N_t^{1-\alpha} K_t^{\alpha} L_t^{1-\alpha}. \end{aligned} \tag{101}$$

Using equation (101), the equilibrium wage w_t in (98) can be simplified as

$$w_t = (1-\alpha)(1-\varepsilon)\frac{Y_t}{L_t}. \tag{102}$$

Equation (29) of the main text can be simplified by substituting out w_t and p_t using equations (96) and (102):

$$\begin{aligned} \kappa_t =& \alpha(1-\varepsilon)(\varepsilon\nu)^{\frac{\varepsilon}{1-\varepsilon}} H_t^{\frac{\alpha-1}{\alpha}} N_t^{\frac{(1-\nu)\varepsilon}{\nu(1-\varepsilon)}} K_t^{\alpha-1} L_t^{1-\alpha} \\ =& \alpha(1-\varepsilon) H_t^{-\frac{1}{\alpha}} \frac{Y_t}{K_t}. \end{aligned} \tag{103}$$

Substituting equations (27) of the main text and equations (95) and (96) into (94) and using (88), we obtain

$$\pi_t = \frac{1-\nu}{\nu} (\varepsilon\nu)^{\frac{1-(1-\alpha)(1-\varepsilon)}{\alpha(1-\varepsilon)}} H_t^{\frac{1}{\alpha}} N_t^{\frac{1-\nu}{\nu}\left[\frac{1-(1-\alpha)(1-\varepsilon)}{\alpha(1-\varepsilon)} - \frac{1}{1-\nu}\right]} \left[\frac{(1-\alpha)(1-\varepsilon)}{w_t} \right]^{\frac{1-\alpha}{\alpha}} K_t. \tag{104}$$

Further, substituting equation (98) into the above equation and using (101), we obtain

$$\begin{aligned} \pi_t =& \frac{1-\nu}{\nu} (\varepsilon v)^{\frac{1}{1-\varepsilon}} H_t N_t^{\frac{\varepsilon-\nu}{\nu(1-\varepsilon)}} K_t^{\alpha} L_t^{1-\alpha} \\ =& (1-\nu)\varepsilon \frac{Y_t}{N_t}. \end{aligned} \tag{105}$$

Thus, we have

$$\int_{j=0}^{N_t} \pi_t \mathrm{d}j = N_t \pi_t = (1-\nu)\varepsilon Y_t. \tag{106}$$

Substituting equation (10) of the main text into (11) of the main text, we obtain

$$S_t = (\chi v_t)^{\frac{1}{h}} N_t. \tag{107}$$

A.4 Resource Constraint

By definition, the aggregate output $Y_t \mathrm{d}t$ is

$$Y_t \mathrm{d}t = \int_{\underline{z}_t}^{\infty} \int_0^{\infty} y_t(a,z)\mathrm{d}t \varphi_t(a,z)\mathrm{d}a\mathrm{d}z = \int_0^{\infty}\int_0^{\infty} y_t(a,z)\mathrm{d}t\varphi_t(a,z)\mathrm{d}a\mathrm{d}z. \tag{108}$$

Substituting equations (3) and (19) of the main text into the above equation and using (16), (23), (32), (33), (35), and (36) of the main text, we obtain

$$\begin{aligned} Y_t \mathrm{d}t =& \mathrm{d}A_t + (\delta_a \mathrm{d}t - \sigma_a \mathrm{d}W_t)A_t + w_t L_t \mathrm{d}t + (\delta_k \mathrm{d}t + \sigma_k \mathrm{d}W_t)K_t + r_{f,t}B_t\mathrm{d}t + \rho A_t \mathrm{d}t \\ &+ \int_0^{\infty}\int_0^{\infty}\left(\int_0^{N_t} p_{j,t}x_{j,t}(a,z)\mathrm{d}j\mathrm{d}t\right)\varphi_t(a,z)\mathrm{d}a\mathrm{d}z, \end{aligned} \tag{109}$$

where the last term is the revenue of the intermediate goods sector. Using equations (7) and (25) of the main text and the definition $X_t \equiv \int_{i\in\mathcal{I}} x_{i,t}\mathrm{d}i = \int_0^\infty\int_0^\infty x_t(a,z)\varphi_t(a,z)\mathrm{d}a\mathrm{d}z$, it can be simplified as follows

$$\begin{aligned} \int_0^{\infty}\int_0^{\infty}\left(\int_0^{N_t} p_{j,t}x_{j,t}(a,z)\mathrm{d}j\mathrm{d}t\right)\varphi_t(a,z)\mathrm{d}a\mathrm{d}z &= \int_0^{N_t}\left(\int_0^{\infty}\int_0^{\infty} p_{j,t}x_{j,t}(a,z)\varphi_t(a,z)\mathrm{d}a\mathrm{d}z\right)\mathrm{d}j\mathrm{d}t \\ &= \int_0^{N_t} p_{j,t}e_{j,t}\mathrm{d}j\mathrm{d}t \\ &= \int_0^{N_t} \pi_{j,t}\mathrm{d}j\mathrm{d}t + \int_0^{N_t} e_{j,t}\mathrm{d}j\mathrm{d}t. \end{aligned} \tag{110}$$

Substituting equation (110) into (109), we obtain

$$\begin{aligned} Y_t \mathrm{d}t =& \mathrm{d}A_t + (\delta_a \mathrm{d}t - \sigma_a \mathrm{d}W_t)A_t + w_t L_t \mathrm{d}t + (\delta_k \mathrm{d}t + \sigma_k \mathrm{d}W_t)K_t + r_{f,t}B_t\mathrm{d}t + \rho A_t \mathrm{d}t \\ &+ \int_0^{N_t} \pi_{j,t}\mathrm{d}j\mathrm{d}t + \int_0^{N_t} e_{j,t}\mathrm{d}j\mathrm{d}t. \end{aligned} \tag{111}$$

Substituting equations (14) and (31) of the main text into (111), we obtain the resource constraint

$$Y_t\mathrm{d}t = \underbrace{\mathrm{d}A_t + (\delta_a\mathrm{d}t - \sigma_a\mathrm{d}W_t)A_t\mathrm{d}t + (\delta_k\mathrm{d}t + \sigma_k\mathrm{d}W_t)K_t}_{\text{investment in the final goods sector}}$$
$$+ \underbrace{S_t\mathrm{d}t + \int_0^{N_t} e_{j,t}\mathrm{d}j\mathrm{d}t}_{\text{R\&D and intangible goods production}} + C_t\mathrm{d}t - \mathrm{d}B_t. \tag{112}$$

Note that the resource constraint (112) holds by Walras's law in equilibrium. This can be proved by substituting equations (44) and (50) of the main text into (111), and using the condition below

$$\int_0^\infty \int_0^\infty \left(\int_0^{N_t} p_{j,t}x_{j,t}(a,z)\mathrm{d}j\mathrm{d}t \right) \varphi_t(a,z)\mathrm{d}k\mathrm{d}z = \varepsilon Y_t\mathrm{d}t, \tag{113}$$

which simply says that the cost of purchasing intangible goods is equal to a share ε of Y_t (the derivation is similar to equation (45) of the main text).

A.5 Proof of Lemma 2

We provide a heuristic proof for Lemma 2, that is, the actual distribution of $\widetilde{a}_{i,t} \equiv \ln a_{i,t}$ is approximately normal. In the absence of aggregate shocks, consider the balanced growth path. Thus, all equilibrium prices are constant as shown in the proof of Proposition 5. The productivity cutoff $\underline{z}$ determined by equation (83) becomes:

$$\underline{z}\kappa = r_f + \delta_k. \tag{114}$$

Rewriting equations (3) and (30) of the main text using (114) as follows:

$$\frac{\mathrm{d}a_{i,t}}{\mathrm{d}t} = s(z_{i,t})a_{i,t}, \tag{115}$$

where

$$s(z) = (1+\lambda)\kappa \max\{z - \underline{z}, 0\} + r_f - \rho - \delta_a, \tag{116}$$

and κ is given by equation (103), with $H_t \equiv H$ and Y_t/K_t being a constant in the balanced growth path, as follows:

$$\kappa = \alpha(1-\varepsilon)H^{-\frac{1}{\alpha}}\frac{Y_t}{K_t}. \tag{117}$$

To better illustrate intuitions, we rewrite equation (115) in discrete time with a time interval $\Delta t \approx 0$:

$$a_{i,t+\Delta t} = [1 + s(z_{i,t})\Delta t]\, a_{i,t}. \tag{118}$$

We denote $a_{i,n} \equiv a_{i,n\Delta t}$ and $z_{i,n} \equiv z_{i,n\Delta t}$ for $n = 1, 2, ...$ Then, it follows that

$$a_{i,n+1} = [1 + s(z_{i,n})\Delta t]\, a_{i,n}. \tag{119}$$

Define $\xi_{i,n} \equiv \ln(1 + s(z_{i,n})\Delta t) - \bar{\xi}$ with $\bar{\xi} \equiv \mathbb{E}\left[\ln(1 + s(z_{i,n})\Delta t)\right]$, thus equation (119) can be written as

$$\ln a_{i,n+1} = \ln a_{i,n} + \bar{\xi} + \xi_{i,n}. \tag{120}$$

For a large $T > 0$, suppose we set $N_T = T/\Delta t$ (without loss of generality, we assume that N_T is an integer), then equation (120) implies

$$\ln a_{i,t} = \ln a_{i,1} + (N_T - 1)\bar{\xi} + \sum_{n=1}^{N_T-1} \xi_{i,n}. \tag{121}$$

In the balanced growth path, $z_{i,n}$ follows a stationary process evolving according to equation (5) of the main text. Thus, the process $\xi_{i,n}$ is also stationary.

The evolution of $\ln z_{i,n}$ can be directly obtained from equation (5) of the main text, as follows:

$$\ln z_{i,n+1} = e^{-\theta\Delta t}\ln z_{i,n} + \sigma_\Delta \varepsilon_{i,n+1}, \tag{122}$$

where $\varepsilon_{i,n+1}$ is a standard normal variable and

$$\sigma_\Delta = \sigma\sqrt{\frac{1-e^{-2\theta\Delta t}}{2}}. \tag{123}$$

According to Andrews (1983), the process $z_{i,n}$ is strong mixing with mixing coefficients dominated by an exponentially declining sequence. Let

$$\sigma_{N_T}^2 = \mathbb{E}\left[\xi_{i,1}^2\right] + 2\sum_{n=1}^{N_T-1}\left(1-\frac{n}{N_T}\right)\mathbb{E}\left[\xi_{i,1}\xi_{i,n}\right]. \tag{124}$$

Using the Berry-Esseen bound developed by Tikhomirov (1980) and Bentkus, Gotze and Tikhomoirov (1997), we obtain

$$\sup_x \left|\mathbb{P}\left\{\sum_{n=1}^{N_T-1}\xi_{i,n} \leq \sigma_{N_T}x\right\} - \Phi(x)\right| \leq AN_T^{-1/2}\ln^2 N_T, \tag{125}$$

where $\Phi(x)$ is the CDF of a standard normal random variable, and A is a constant that depends on model parameters.

A.6 Proof of Lemma 3

Let $\psi_t(\widetilde{a},\widetilde{z})$ be the joint distribution of $\widetilde{a}$ and $\widetilde{z}$. Define $\Gamma_t \equiv \mathrm{Cov}(\widetilde{a}_{i,t},\widetilde{z}_{i,t})$. Under Lemma 2, $\psi_t(\widetilde{a},\widetilde{z})$ is the PDF of a joint normal distribution, with the covariance between $\widetilde{a}$ and $\widetilde{z}$ being Γ_t.

The PDF $\varphi_t(a,z)$ is related to $\psi_t(\widetilde{a},\widetilde{z})$ through the Jacobian matrix J, as follows:

$$\varphi_t(a,z) = |J|\psi_t(\widetilde{a},\widetilde{z}), \tag{126}$$

where J is defined by

$$J = \begin{pmatrix} \partial\widetilde{a}/\partial a & \partial\widetilde{a}/\partial z \\ \partial\widetilde{z}/\partial a & \partial\widetilde{z}/\partial z \end{pmatrix}. \tag{127}$$

Thus, we have

$$\varphi_t(a,z) = \frac{1}{az}\psi_t(\widetilde{a},\widetilde{z}). \tag{128}$$

Using equation (128), the term $\int_0^\infty a\varphi_t(a,z)\mathrm{d}a$ in equation (38) of the main text can be written as

$$\int_0^\infty a\varphi_t(a,z)\mathrm{d}a = \int_{-\infty}^\infty \frac{a}{z}\psi_t(\widetilde{a},\widetilde{z})\mathrm{d}\widetilde{a}. \tag{129}$$

Let $f(\widetilde{z})$ be the PDF of $\widetilde{z}$, which follows a normal distribution, $N(0,\sigma^2/2)$, in the stationary equilibrium. Thus, equation (129) can be written as

$$\int_0^\infty a\varphi_t(a,z)\mathrm{d}a = \int_{-\infty}^\infty \frac{a}{z}\psi_t(\widetilde{a}|\widetilde{z})f(\widetilde{z})\mathrm{d}\widetilde{a} = \mathbb{E}\left[\exp(\widetilde{a}_{i,t})|\widetilde{z}\right]\frac{f(\widetilde{z})}{z}. \tag{130}$$

Using equation (128), the variable A_t defined in (32) can be written as

$$A_t = \int_{-\infty}^\infty \int_{-\infty}^\infty a\frac{1}{az}\psi_t(\widetilde{a},\widetilde{z})az\mathrm{d}\widetilde{a}\mathrm{d}\widetilde{z} = \mathbb{E}\left[\exp(\widetilde{a}_{i,t})\right]. \tag{131}$$

Substituting equations (130) and (131) into equation (38) of the main text, we obtain

$$\omega_t(z) = \frac{\mathbb{E}[\exp(\widetilde{a}_{i,t})|\widetilde{z}]}{\mathbb{E}[\exp(\widetilde{a}_{i,t})]}\frac{f(\widetilde{z})}{z}. \tag{132}$$

Because $\widetilde{a}_{i,t}$ and $\widetilde{z}_{i,t}$ follow a joint normal distribution with covariance Γ_t, we have

$$\mathbb{E}[\exp(\widetilde{a}_{i,t})|\widetilde{z}] = \exp\left(\mathbb{E}[\widetilde{a}_{i,t}|\widetilde{z}] + \frac{1}{2}\mathrm{var}(\widetilde{a}_{i,t}|\widetilde{z})\right), \tag{133}$$

$$\mathbb{E}[\exp(\widetilde{a}_{i,t})] = \exp\left(\mathbb{E}[\widetilde{a}_{i,t}] + \frac{1}{2}\mathrm{var}(\widetilde{a}_{i,t})\right), \tag{134}$$

where

$$\mathbb{E}[\widetilde{a}_{i,t}|\widetilde{z}] = \mathbb{E}[\widetilde{a}_{i,t}] + 2\widetilde{z}\Gamma_t/\sigma^2, \tag{135}$$

$$\mathrm{var}(\widetilde{a}_{i,t}|\widetilde{z}) = \mathrm{var}(\widetilde{a}_{i,t}) - 2\Gamma_t^2/\sigma^2. \tag{136}$$

Substituting equations (133) to (136) into (132), we obtain

$$\begin{aligned}\omega_t(z) =& \frac{f(\widetilde{z})}{z}\exp\left(\frac{2\widetilde{z}\Gamma_t}{\sigma^2}\right)\exp\left(-\frac{\Gamma_t^2}{\sigma^2}\right) \\ =& \frac{1}{z\sigma\sqrt{\pi}}\exp\left(-\frac{(\ln z - \Gamma_t)^2}{\sigma^2}\right).\end{aligned} \tag{137}$$

This formula turns out to be the same as equation (29) of Moll (2014). Substituting out $\Gamma_t = -M_t\mathrm{var}(\widetilde{z}_{i,t})$, we obtain equation (47) in the main text.

A.7 Proof of Proposition 3

Define $\Gamma_t \equiv \mathrm{Cov}(\widetilde{a}_{i,t}, \widetilde{z}_{i,t}) = -M_t \mathrm{var}(\widetilde{z}_{i,t}) = -M_t \sigma^2/2$. Substituting equation (47) of the main text into (42) of the main text, we obtain the equation that determines the productivity cutoff $\underline{z}_t$ under our approximation of $\omega_t(z)$:

$$\frac{1}{1+\lambda}\frac{K_t}{A_t} = 1 - \Omega_t(\underline{z}_t) = \int_{\underline{\widetilde{z}}_t}^{\infty} \frac{1}{\sigma\sqrt{\pi}} \exp\left(-\frac{(\widetilde{z}-\Gamma_t)^2}{\sigma^2}\right) \mathrm{d}\widetilde{z} = \Phi\left(\frac{\Gamma_t - \underline{\widetilde{z}}_t}{\sigma/\sqrt{2}}\right). \tag{138}$$

Rearranging the above equation, we obtain $\underline{z}_t$:

$$\underline{z}_t = \exp\left(\Gamma_t - \Phi^{-1}\left(\frac{1}{1+\lambda}\frac{K_t}{A_t}\right)\frac{\sigma}{\sqrt{2}}\right). \tag{139}$$

The term $\int_{\underline{z}_t}^{\infty} z\omega_t(z)\mathrm{d}z$ can be simplified using (47) of the main text, as follows

$$\begin{aligned}\int_{\underline{z}_t}^{\infty} z\omega_t(z)\mathrm{d}z &= \int_{\underline{z}_t}^{\infty} z\frac{1}{z\sigma\sqrt{\pi}} \exp\left(-\frac{(\widetilde{z}-\Gamma_t)^2}{\sigma^2}\right) \mathrm{d}z \\ &= \int_{\underline{\widetilde{z}}_t}^{\infty} \frac{1}{\sigma\sqrt{\pi}} \exp\left(-\frac{(\widetilde{z}-\Gamma_t-\sigma^2/2)^2}{\sigma^2}\right) \exp\left(\Gamma_t + \frac{\sigma^2}{4}\right) \mathrm{d}\widetilde{z} \\ &= \exp\left(\Gamma_t + \frac{\sigma^2}{4}\right)\Phi\left(\frac{\Gamma_t + \sigma^2/2 - \underline{\widetilde{z}}_t}{\sigma/\sqrt{2}}\right).\end{aligned} \tag{140}$$

Substituting equation (140) into (41) of the main text, we obtain

$$H_t = \left[(1+\lambda)\frac{A_t}{K_t}\exp\left(\Gamma_t + \frac{\sigma^2}{4}\right)\Phi\left(\frac{\Gamma_t + \sigma^2/2 - \underline{\widetilde{z}}_t}{\sigma/\sqrt{2}}\right)\right]^{\alpha}. \tag{141}$$

Further, substituting equation (139) into the above equation, we obtain

$$H_t = \left[(1+\lambda)\frac{A_t}{K_t}\exp\left(\Gamma_t + \frac{\sigma^2}{4}\right)\Phi\left(\Phi^{-1}\left(\frac{1}{1+\lambda}\frac{K_t}{A_t}\right) + \frac{\sigma}{\sqrt{2}}\right)\right]^{\alpha}. \tag{142}$$

Substituting out $\Gamma_t = -M_t \mathrm{var}(\widetilde{z}_{i,t}) = -M_t\sigma^2/2$ and using equation (41) of the main text, we obtain equation (48) of the main text,

$$Z_t = (\varepsilon\nu)^{\frac{\varepsilon}{1-\varepsilon}} N_t^{1-\alpha}\left[(1+\lambda)\frac{A_t}{K_t}\exp\left(-\frac{\sigma^2}{2}M_t + \frac{\sigma^2}{4}\right)\Phi\left(\Phi^{-1}\left(\frac{1}{1+\lambda}\frac{K_t}{A_t}\right) + \frac{\sigma}{\sqrt{2}}\right)\right]^{\alpha}. \tag{143}$$

A.8 Proof of Proposition 4

Equation (32) of the main text implies

$$A_{t+\mathrm{d}t} - A_t = \int_0^\infty \int_0^\infty \mathrm{d}a_t(a,z)\varphi_t(a,z)\mathrm{d}a\mathrm{d}z. \tag{144}$$

Substituting equations (2), (3), and (30) of the main text into the above equation, we obtain

$$\begin{aligned} A_{t+\mathrm{d}t} - A_t = &(1+\lambda)\kappa_t \int_{\underline{z}_t}^\infty \int_0^\infty za\mathrm{d}t\varphi_t(a,z)\mathrm{d}a\mathrm{d}z - (1+\lambda)r_{f,t}\int_{\underline{z}_t}^\infty \int_0^\infty a\mathrm{d}t\varphi_t(a,z)\mathrm{d}a\mathrm{d}z \\ &- (1+\lambda)\left(\delta_k\mathrm{d}t + \sigma_k\mathrm{d}W_t\right)\int_{\underline{z}_t}^\infty \int_0^\infty a\varphi_t(a,z)\mathrm{d}a\mathrm{d}z + \sigma_a A_t\mathrm{d}W_t + (r_{f,t} - \rho - \delta_a)A_t\mathrm{d}t. \end{aligned} \tag{145}$$

Using equations (88), (90), and (103), the above equation can be simplified as

$$\mathrm{d}A_t = \alpha(1-\varepsilon)Y_t\mathrm{d}t - (r_{f,t} + \delta_k)K_t\mathrm{d}t - (\rho + \delta_a - r_{f,t})A_t\mathrm{d}t + (\sigma_a A_t - \sigma_k K_t)\mathrm{d}W_t. \tag{146}$$

Substituting equations (11) and (46) of the main text into equation (9) of the main text, we obtain

$$\frac{\mathrm{d}N_t}{N_t} = \chi\left(\chi v_t\right)^{\frac{1-h}{h}}\mathrm{d}t - \delta_b\mathrm{d}t. \tag{147}$$

Define $\Gamma_t \equiv \mathrm{Cov}(\widetilde{a}_{i,t}, \widetilde{z}_{i,t}) = -M_t\mathrm{var}(\widetilde{z}_{i,t}) = -M_t\sigma^2/2$. Next, we derive the evolution of Γ_t. By definition, $\Gamma_{t+\mathrm{d}t} \equiv \mathrm{Cov}(\widetilde{a}_{i,t+\mathrm{d}t}, \widetilde{z}_{i,t+\mathrm{d}t})$. According to equation (5) of the main text, we have

$$\widetilde{z}_{i,t+\mathrm{d}t} = \widetilde{z}_{i,t} - \theta\widetilde{z}_{i,t}\mathrm{d}t + \sigma\sqrt{\theta}\mathrm{d}W_{i,t}. \tag{148}$$

Thus,

$$\begin{aligned} \mathrm{d}\Gamma_t =& \mathrm{Cov}\left(\widetilde{a}_{i,t} + \mathrm{d}\widetilde{a}_{i,t}, \widetilde{z}_{i,t} - \theta\widetilde{z}_{i,t}\mathrm{d}t + \sigma\sqrt{\theta}\mathrm{d}W_{i,t}\right) - \Gamma_t \\ =&(1-\theta\mathrm{d}t)\mathrm{Cov}(\widetilde{a}_{i,t} + \mathrm{d}\widetilde{a}_{i,t}, \widetilde{z}_{i,t}) - \Gamma_t \\ =&-\theta\Gamma_t\mathrm{d}t + (1-\theta\mathrm{d}t)\mathrm{Cov}(\widetilde{z}_{i,t}, \mathrm{d}\widetilde{a}_{i,t}). \end{aligned} \tag{149}$$

Omitting the higher-order term $\mathrm{Cov}(\widetilde{z}_{i,t}, \mathrm{d}\widetilde{a}_{i,t})\mathrm{d}t$, we obtain

$$\mathrm{d}\Gamma_t = -\theta\Gamma_t\mathrm{d}t + \mathrm{Cov}(\widetilde{z}_{i,t}, \mathrm{d}\widetilde{a}_{i,t}). \tag{150}$$

Substituting out $\Gamma_t = -M_t \text{var}(\widetilde{z}_{i,t}) = -M_t \sigma^2/2$, we obtain equation (52) of the main text.

We now derive the expression for $\text{Cov}(\widetilde{z}_{i,t}, \text{d}\widetilde{a}_{i,t})$ under Lemma 2. Using Ito's lemma

$$\text{d}\widetilde{a}_{i,t} = \frac{1}{a_{i,t}}\text{d}a_{i,t} - \frac{1}{2a_{i,t}^2}(\text{d}a_{i,t})^2. \tag{151}$$

Substituting equation (2), (3), and (30) of the main text into the above equation, we obtain the evolution of $\widetilde{a}_{i,t}$. In particular, for $z_{i,t} < \underline{z}_t$, we have

$$\text{d}\widetilde{a}_{i,t} = (r_{f,t} - \rho - \delta_a)\text{d}t + \sigma_a \text{d}W_t. \tag{152}$$

For $z_{i,t} \geq \underline{z}_t$, we have

$$\text{d}\widetilde{a}_{i,t} = (1+\lambda)\left[\kappa_t z_{i,t}\text{d}t - (\delta_k \text{d}t + \sigma_k \text{d}W_t) - r_{f,t}\text{d}t\right] + (r_{f,t} - \rho - \delta_a)\text{d}t + \sigma_a \text{d}W_t. \tag{153}$$

Because $\mathbb{E}[\widetilde{z}_{i,t}] = 0$, we have

$$\text{Cov}(\widetilde{z}_{i,t}, \text{d}\widetilde{a}_{i,t}) = \mathbb{E}[\widetilde{z}_{i,t}\text{d}\widetilde{a}_{i,t}]. \tag{154}$$

Substituting equations (152) and (153) into (154), we obtain

$$\text{Cov}(\widetilde{z}_{i,t}, \text{d}\widetilde{a}_{i,t}) = (1+\lambda)\kappa_t \text{d}t \int_{\underline{\widetilde{z}}_t}^{\infty} \widetilde{z} z f(\widetilde{z})\text{d}\widetilde{z} - (1+\lambda)\left[(r_{f,t} + \delta_k)\text{d}t + \sigma_k \text{d}W_t\right] \int_{\underline{\widetilde{z}}_t}^{\infty} \widetilde{z} f(\widetilde{z})\text{d}\widetilde{z}, \tag{155}$$

where $f(\widetilde{z})$ is the PDF of $\widetilde{z}$, which follows a normal distribution, $N(0, \sigma^2/2)$, in the stationary equilibrium.

Substituting out $f(\widetilde{z})$, the term $\int_{\underline{\widetilde{z}}_t}^{\infty} \widetilde{z} f(\widetilde{z})\text{d}\widetilde{z}$ in equation (155) can be simplified as follows:

$$\int_{\underline{\widetilde{z}}_t}^{\infty} \widetilde{z} f(\widetilde{z})\text{d}\widetilde{z} = \int_{\underline{\widetilde{z}}_t}^{\infty} \widetilde{z} \frac{1}{\sigma\sqrt{\pi}} \exp\left(-\frac{\widetilde{z}^2}{\sigma^2}\right)\text{d}\widetilde{z} = -\frac{\sigma}{2\sqrt{\pi}} \int_{\underline{\widetilde{z}}_t}^{\infty} \text{d}\exp\left(-\frac{\widetilde{z}^2}{\sigma^2}\right) = \frac{\sigma}{2\sqrt{\pi}} \exp\left(-\frac{\underline{\widetilde{z}}_t^2}{\sigma^2}\right). \tag{156}$$

The term $\int_{\underline{\widetilde{z}}_t}^{\infty} \widetilde{z} z f(\widetilde{z}) \mathrm{d}\widetilde{z}$ in equation (155) can be simplified as follows:

$$
\begin{aligned}
&\int_{\underline{\widetilde{z}}_t}^{\infty} \widetilde{z} z f(\widetilde{z}) \mathrm{d}\widetilde{z} = \frac{1}{\sigma\sqrt{\pi}} \int_{\underline{\widetilde{z}}_t}^{\infty} \widetilde{z} \exp\left(\widetilde{z} - \frac{\widetilde{z}^2}{\sigma^2}\right) \mathrm{d}\widetilde{z} \\
=&\frac{1}{\sigma\sqrt{\pi}} \exp\left(\frac{\sigma^2}{4}\right) \int_{\underline{\widetilde{z}}_t}^{\infty} \widetilde{z} \exp\left(-\frac{1}{\sigma^2}\left(\widetilde{z} - \frac{\sigma^2}{2}\right)^2\right) \mathrm{d}\widetilde{z} \qquad (157)\\
=&\frac{1}{\sigma\sqrt{\pi}} \exp\left(\frac{\sigma^2}{4}\right) \left[\int_{\underline{\widetilde{z}}_t}^{\infty} \left(\widetilde{z} - \frac{\sigma^2}{2}\right) \exp\left(-\frac{1}{\sigma^2}\left(\widetilde{z} - \frac{\sigma^2}{2}\right)^2\right) \mathrm{d}\widetilde{z} + \int_{\underline{\widetilde{z}}_t}^{\infty} \frac{\sigma^2}{2} \exp\left(-\frac{1}{\sigma^2}\left(\widetilde{z} - \frac{\sigma^2}{2}\right)^2\right) \mathrm{d}\widetilde{z}\right].
\end{aligned}
$$

Integrating both terms on the right-hand side of the above equation, we obtain

$$
\begin{aligned}
&\int_{\underline{\widetilde{z}}_t}^{\infty} \widetilde{z} z f(\widetilde{z}) \mathrm{d}\widetilde{z} \\
=&-\frac{\sigma}{2\sqrt{\pi}} \exp\left(\frac{\sigma^2}{4}\right) \int_{\underline{\widetilde{z}}_t}^{\infty} \mathrm{d} \exp\left(-\frac{1}{\sigma^2}\left(\widetilde{z} - \frac{\sigma^2}{2}\right)^2\right) + \frac{\sigma^2}{2} \exp\left(\frac{\sigma^2}{4}\right) \Phi\left(\frac{\sigma^2/2 - \underline{\widetilde{z}}_t}{\sigma/\sqrt{2}}\right) \\
=&\frac{\sigma}{2} \exp\left(\frac{\sigma^2}{4}\right) \left[\frac{1}{\sqrt{\pi}} \exp\left(-\frac{1}{\sigma^2}\left(\underline{\widetilde{z}}_t - \frac{\sigma^2}{2}\right)^2\right) + \sigma\Phi\left(\frac{\sigma^2/2 - \underline{\widetilde{z}}_t}{\sigma/\sqrt{2}}\right)\right]. \qquad (158)
\end{aligned}
$$

Substituting equations (156) and (158) into (155), we obtain

$$
\begin{aligned}
\mathrm{Cov}(\widetilde{z}_{i,t}, \mathrm{d}\widetilde{a}_{i,t}) =&\frac{(1+\lambda)\sigma^2\kappa_t}{2} \exp\left(\frac{\sigma^2}{4}\right) \Phi\left(\frac{\sigma^2/2 - \underline{\widetilde{z}}_t}{\sigma/\sqrt{2}}\right) \mathrm{d}t \\
&+ \frac{(1+\lambda)\sigma}{2\sqrt{\pi}} [(\underline{z}_t\kappa_t - r_{f,t} - \delta_k)\mathrm{d}t - \sigma_k \mathrm{d}W_t] \exp\left(-\frac{\underline{\widetilde{z}}_t^2}{\sigma^2}\right). \qquad (159)
\end{aligned}
$$

A.9 Proof of Proposition 5

In the absence of aggregate shocks, the evolution of aggregate capital A_t (equation (50) of the main text) becomes

$$
\frac{\mathrm{d}A_t}{A_t} = \alpha(1-\varepsilon)\frac{Y_t}{A_t}\mathrm{d}t - (r_{f,t} + \delta_k)\frac{K_t}{A_t}\mathrm{d}t - (\rho + \delta_a - r_{f,t})\mathrm{d}t. \qquad (160)
$$

In the balanced growth path, aggregate output Y_t, consumption C_t, capital A_t, and knowledge stock N_t all grow at a constant rate g:

$$
\frac{\mathrm{d}Y_t}{Y_t} = \frac{\mathrm{d}C_t}{C_t} = \frac{\mathrm{d}A_t}{A_t} = \frac{\mathrm{d}N_t}{N_t} = g\mathrm{d}t. \qquad (161)
$$

The variables E_t, K_t/A_t, Y_t/A_t, Γ_t, H_t, $\underline{z}_t$, κ_t, $r_{f,t}$, and v_t are all constant. From now on, we omit the subscript t for these variables. The risk-free rate r_f is determined by the representative agent's first-order condition

$$\frac{\mathrm{d}C_t}{C_t} = \psi(r_f - \delta)\mathrm{d}t. \tag{162}$$

Substituting equation (161) into (162), (51) of the main text, and (160), respectively, we obtain

$$r_f = g/\psi + \delta, \tag{163}$$

$$g = \chi(\chi v)^{\frac{1-h}{h}} - \delta_b, \tag{164}$$

$$\frac{Y_t}{A_t} = \frac{g + \rho + \delta_a - r_f + (r_f + \delta_k)K_t/A_t}{\alpha(1-\varepsilon)}. \tag{165}$$

Dividing both sides of equation (40) of the main text by A_t and using $L_t \equiv 1$, we obtain

$$E = \left[\frac{1}{(\varepsilon\nu)^{\frac{\varepsilon}{1-\varepsilon}} H} \frac{Y_t}{A_t} \left(\frac{A_t}{K_t}\right)^{\alpha}\right]^{\frac{1}{1-\alpha}}, \tag{166}$$

The flow profit π to each intermediate-goods producer is a constant and given by equation (45) of the main text,

$$\pi = (1-\nu)\varepsilon\frac{Y_t}{N_t} = (1-\nu)\varepsilon\frac{1}{E}\frac{Y_t}{A_t}. \tag{167}$$

Substituting equation (167) into equation (8) of the main text, we obtain the value of blue prints v

$$v = \frac{\pi}{r_f + \delta_b}. \tag{168}$$

Substituting equations (163), (166), and (167) into (168), we obtain v

$$v = \frac{(1-\nu)\varepsilon(\varepsilon\nu)^{\overline{(1-\varepsilon)(1-\alpha)}^{\varepsilon}}}{g/\psi + \delta + \delta_b} \left(\frac{A_t}{Y_t}\frac{K_t}{A_t}\right)^{\frac{\alpha}{1-\alpha}} H^{\frac{1}{1-\alpha}}. \tag{169}$$

Define $\Gamma_t \equiv \mathrm{Cov}(\widetilde{a}_{i,t}, \widetilde{z}_{i,t}) = -M_t \mathrm{var}(\widetilde{z}_{i,t}) = -M_t\sigma^2/2$. The steady-state value of H is given by $Z_t/((\varepsilon\nu)^{\varepsilon/(1-\varepsilon)} N_t^{1-\alpha})$ according to equation (41) of the main text. Using equation (48) of the main text, we obtain H as follows:

$$H = \left[(1+\lambda) \frac{A_t}{K_t} \exp\left(\Gamma + \frac{\sigma^2}{4} \right) \Phi \left(\Phi^{-1} \left(\frac{1}{1+\lambda} \frac{K_t}{A_t} \right) + \frac{\sigma}{\sqrt{2}} \right) \right]^{\alpha}, \tag{170}$$

where the covariance Γ in the balanced growth path is obtained by setting $\mathrm{d}M_t = 0$ in equation (52) of the main text:

$$\Gamma = \frac{\mathrm{Cov}(\widetilde{z}_{i,t}, \mathrm{d}\widetilde{a}_{i,t})}{\theta \mathrm{d}t}. \tag{171}$$

In the balanced growth path, $\mathrm{Cov}(\widetilde{z}_{i,t}, \mathrm{d}\widetilde{a}_{i,t})$ given by equation (159) becomes

$$\begin{aligned} \mathrm{Cov}(\widetilde{z}_{i,t}, \mathrm{d}\widetilde{a}_{i,t}) =& \frac{(1+\lambda)\sigma^2\kappa}{2} \exp\left(\frac{\sigma^2}{4} \right) \Phi \left(\frac{\sigma^2/2 - \underline{\widetilde{z}}}{\sigma/\sqrt{2}} \right) \mathrm{d}t \\ &+ \frac{(1+\lambda)\sigma}{2\sqrt{\pi}} (\underline{z}\kappa - r_f - \delta_k) \exp\left(-\frac{\underline{\widetilde{z}}^2}{\sigma^2} \right) \mathrm{d}t. \end{aligned} \tag{172}$$

Substituting equation (172) into (171):

$$\Gamma = \frac{(1+\lambda)\sigma^2\kappa}{2\theta} \exp\left(\frac{\sigma^2}{4} \right) \Phi \left(\frac{\sigma^2/2 - \ln \underline{z}}{\sigma/\sqrt{2}} \right) + \frac{(1+\lambda)\sigma}{2\theta\sqrt{\pi}} (\underline{z}\kappa - r_f - \delta_k) \exp\left(-\frac{\underline{\widetilde{z}}^2}{\sigma^2} \right). \tag{173}$$

When solving above equations, we need to know K_t/A_t, $\underline{z}$, and κ. They are given by equation (139), (28) of the main text (setting $\sigma_k = 0$), and (103), respectively, as follows

$$\underline{z} = \exp\left(\Gamma - \Phi^{-1} \left(\frac{1}{1+\lambda} \frac{K_t}{A_t} \right) \frac{\sigma}{\sqrt{2}} \right), \tag{174}$$

$$\underline{z}\kappa = r_f + \delta_k, \tag{175}$$

$$\kappa = \alpha(1-\varepsilon) H^{-\frac{1}{\alpha}} \frac{Y_t}{A_t} \frac{A_t}{K_t}. \tag{176}$$

A.10 Proof of Proposition 6

Substituting equation (167) into (168), we obtain

$$v_t = \frac{(1-\nu)\varepsilon}{r_{f,t} + \delta_b} \frac{Y_t}{N_t}. \tag{177}$$

Substituting equations (40) and (48) of the main text into (177), and using $L_t \equiv 1$, we obtain

$$\ln(v_t) = -\frac{\alpha\sigma^2}{2}M_t + \frac{\alpha\sigma^2}{4} + \ln\left(\frac{(1-\nu)\varepsilon(\varepsilon\nu)^{\frac{\varepsilon}{1-\varepsilon}}}{r_{f,t}+\delta_b}\right) + \alpha\ln(1+\lambda) + \alpha\ln\left(\frac{A_t}{N_t}\right) \\ + \alpha\ln\left(\Phi\left(\Phi^{-1}\left(\frac{1}{1+\lambda}\frac{K_t}{A_t}\right) + \frac{\sigma}{\sqrt{2}}\right)\right). \tag{178}$$

Substituting equation (178) into (46) of the main text, we obtain

$$\ln\left(\frac{S_t}{A_t}\right) = -\frac{\alpha\sigma^2}{2h}M_t + \frac{\alpha\sigma^2}{4h} + \frac{1}{h}\ln(\chi) + \frac{\alpha}{h}\ln(1+\lambda) + \left(\frac{\alpha}{h}-1\right)\ln\left(\frac{A_t}{N_t}\right) \\ + \frac{1}{h}\ln\left(\frac{(1-\nu)\varepsilon(\varepsilon\nu)^{\frac{\varepsilon}{1-\varepsilon}}}{r_{f,t}+\delta_b}\right) + \frac{\alpha}{h}\ln\left(\Phi\left(\Phi^{-1}\left(\frac{1}{1+\lambda}\frac{K_t}{A_t}\right) + \frac{\sigma}{\sqrt{2}}\right)\right). \tag{179}$$

B Assessment of Our Analytical Approximation Method

In this appendix section, we evaluate the accuracy of our analytical approximation method. In Subsection B.1, we review the existing numerical approximation methods and point out their differences from our analytical approximation method in the end. In Subsection B.2, we exactly solve the model in the absence of aggregate shocks and show that the implied variables and distributions in the balanced growth path are similar to the solutions obtained using our analytical approximation method. These results are also theoretically supported by our heuristic proof in Online Appendix A.5.

In the presence of aggregate shocks, it is extremely difficult to accurately solve the model using existing numerical methods. This is because agents need to base their optimal decisions today on the evolution of future prices that are consistent with the economy's law of motion. Because of the heterogeneity across agents, future prices will, through market-clearing conditions, depend on the future distribution of agents, which is infinite dimensional. Thus, the cross-sectional distribution of agents becomes part of the aggregate state space, substantially increasing the computational complexity. In principle, it is impossible to "exactly solve" dynamic stochastic general equilibrium models with heterogeneous agents, which features three crucial parts, heterogeneous agents, aggregate shocks, and endogenous equilibrium prices.[19] Despite the difficulties,

[19] It is the combination of all three features that makes it impossible to solve the model exactly. For example, without heterogeneous agents, the aggregate state space has a finite dimension. The model

in Subsection B.3, we take an off-the-shelf higher-degree numerical approximation method to solve the model. We show that this method yields results similar to those obtained based on our analytical approximation method.

B.1 Discussions of Existing Numerical Approximation Methods

The literature has proposed various approximation algorithms since the influential work of Krusell and Smith (1998*a*). Generally speaking, there are three methods classified by how we approximate the cross-sectional distribution and solve for aggregate dynamics. The first method is proposed by Algan, Allais and Den Haan (2008), who approximate the cross-sectional distribution using a flexible parametric family, which reduces the infinite-dimensional distribution to a finite number of parameters. They then solve for the dynamics of these parameters using a globally accurate projection technique. The second method is proposed by Reiter (2009), who approximates the distribution with a fine histogram, which again reduces the infinite dimensional distribution to a finite number of parameters representing the probability mass of each discretized agent type. Because it requires a large number of discretized bins to approximate the distribution, Reiter (2009) needs to solve the model using locally accurate approximations with respect to the aggregate state vector. The third method is recently proposed by Winberry (2018), who approximates the cross-sectional distribution using the parametric family proposed by Algan, Allais and Den Haan (2008). Instead of solving the model using a global projection method, Winberry (2018) solves the model using a locally accurate perturbation method, similar to Reiter (2009).

Our proposed analytical approximation method in Section 2.6 of the main text is related to the numerical approximation methods discussed above with one crucial difference. In our method, the cross-sectional distribution of agents is assumed to be joint log normal (in $z_{i,t}$ and $a_{i,t}$) at any point in time, which allows us to derive "analytical formulas" for the evolution of the parameters (i.e., the misallocation measure

can be solved using standard iteration or projection methods based on discretized grids of the aggregate state space. Moreover, without aggregate shocks, there exists a deterministic steady state with stationary cross-sectional distributions, which can be solved using a standard shooting algorithm (e.g., Buera and Shin, 2013; Dabla-Norris et al., 2021; Dou and Ji, 2021; Ji, Teng and Townsend, 2021). Finally, without endogenous equilibrium prices (i.e., a model in partial equilibrium), agents make decisions based on the exogenously specified evolution of aggregate prices, and the future cross-sectional distribution is irrelevant from agents' perspective. Thus, the distribution of agents is not part of the aggregate state space when we solve for the optimal policy functions, making standard iteration and projection methods tractable.

M_t) that characterize the distribution. By contrast, in the numerical approximation methods of Algan, Allais and Den Haan (2008) and Winberry (2018), the assumed parametric function is used to "numerically fit" the cross-sectional distribution of agents, and as a result, analytical formulas are not derived.

Broadly speaking, our idea of approximating the joint distribution of firms using the covariance between log productivity and log capital is related to several seminal papers in the macroeconomics literature. In the model of Krusell and Smith (1998*b*), agents face idiosyncratic labor productivity shocks (employed or unemployed). Because the covariance between individual productivity and individual wealth does not affect aggregate output, Krusell and Smith (1998*b*) find that the behavior of macroeconomic aggregates can be almost perfectly described by the mean of the wealth distribution. By contrast, in our model, the covariance between firm productivity and firm capital determines the degree of misallocation, the aggregate TFP, and the growth rate. Thus, both the mean of the capital distribution and the covariance have to be included to describe the behavior of macroeconomic aggregates. Angeletos (2007) considers idiosyncratic investment risk and endognenous wealth accumulation. In his model, the covariance between individual productivity and individual wealth affects the aggregate output. However, by assuming that productivity shocks are i.i.d. over time, the covariance is no longer a state variable; and the mean of wealth serves as a single state variable that sufficiently captures the economy's dynamics. Our model is closest to that of Moll (2014), who considers persistent idiosyncratic productivity shocks and endogenous wealth accumulation in a general equilibrium model with heterogeneous agents. Moll (2014) solves the model without aggregate shocks using the finite difference method for partial differential equations (PDE). Building on Moll (2014)'s characterization of wealth shares, we propose an analytically tractable approximation for the infinite-dimensional joint distribution using the covariance between log productivity and log capital across firms. Thus, misallocation naturally emerges as a crucial endogenous state variable. Our approximation method remains tractable even in the presence of aggregate shocks.

B.2 Deterministic Steady States without Aggregate Shocks

We begin by evaluating the accuracy of our analytical approximation method in the balanced growth path. Because there are no aggregate shocks, the capital share distribution, $\omega(z)$, is stationary in the balanced growth path. We are able to solve

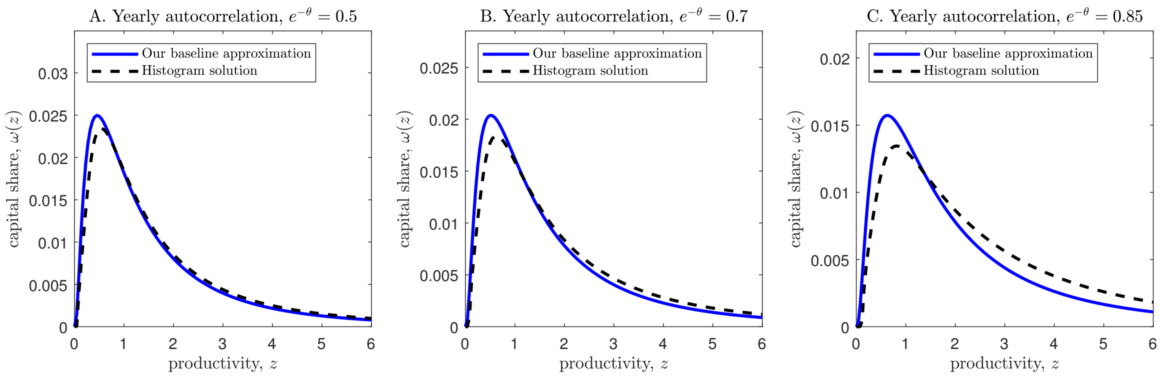
Table OA.1: Accuracy of our analytical approximation in the balanced growth path.

Variables	$\exp(-\theta) = 0.5$		$\exp(-\theta) = 0.7$		$\exp(-\theta) = 0.85$	
	baseline	histogram	baseline	histogram	baseline	histogram
Firm profitability, κ	0.049	0.049	0.044	0.044	0.036	0.036
Productivity cutoff, $\underline{z}$	1.085	1.084	1.219	1.213	1.483	1.470
Value of blueprints, v	0.274	0.274	0.290	0.290	0.319	0.319
Productivity, H	1.449	1.424	1.507	1.475	1.607	1.549
Flow profit of innovators, π	0.060	0.060	0.064	0.064	0.070	0.070
Wage-capital ratio, w/A	0.331	0.341	0.331	0.342	0.331	0.351
Dividend-capital ratio, D/A	0.061	0.061	0.061	0.061	0.061	0.061
R&D-capital ratio, S/A	0.197	0.203	0.197	0.203	0.197	0.209
Knowledge stock-capital ratio, N/A	3.292	3.393	3.105	3.210	2.822	2.999
Balanced growth rate, g (%)	1.893	1.888	1.886	1.855	1.894	1.844
Risk-free rate, r_f (%)	2.023	2.021	2.020	2.003	2.024	1.997

Note: This table evaluates the accuracy of our analytical approximation method in the balanced growth path (i.e., no aggregate shocks) for different persistence of idiosyncratic productivity (θ). The columns labeled "baseline" present the steady-state values of corresponding variables solved by our analytical approximation method. The columns labeled "histogram" present the steady-state values solved by the histogram method with $1,001$ equal-spaced grids for productivity $z_{i,t}$. All the parameter values are taken from our benchmark calibration in Table 1 except for θ (we consider three values, $\exp(-\theta) = 0.5, 0.7, 0.85$) and ψ (we set its value to match a growth rate of about 1.88% for corresponding θ).

it accurately by approximating $\omega(z)$ nonparametrically using a fine histogram, as in Buera and Shin (2013) and Moll (2014). To ensure accuracy, we choose $1,001$ equal-spaced grids for idiosyncratic productivity z over the interval $[z_{min}, z_{max}]$, with $z_{min} = 0$ and $z_{max} = 10$. We verify that the solution does not change when the number of grids is further increased.

Table OA.1 compares the solution of our analytical approximation method and the solution based on the histogram method for various key endogenous aggregate variables. When the yearly autocorrelation in idiosyncratic productivity is 0.5 (i.e., $\exp(-\theta) = 0.5$), our analytical approximation method yields solutions almost identical to those of the histogram method. When idiosyncratic productivity becomes more persistent, the accuracy of method becomes slightly worse. Importantly, in the benchmark calibration with $\exp(-\theta) = 0.85$, our analytical approximation method still yields results very similar to the histogram solution. Not only for these aggregate variables, our baseline approximation method also well captures the endogenous capital share distribution. Figure OA.1 compares the capital share distribution $\omega(z)$



Note: This figure compares the capital share $\omega(z)$ in the balanced growth path (i.e., no aggregate shocks) solved by our analytical approximation method (see equation (47) of the main text) and that solved by the histogram method with 1001 equal-spaced grids for productivity $z_{i,t}$. The blue solid line in each panel represents the our analytical approximation method and the black dashed line represents the histogram method. All the parameter values are taken from our benchmark calibration in Table 1 except for θ (consider three values, $\exp(-\theta) = 0.5, 0.7, 0.85$) and ψ (set its value to match a 1.89% growth rate for corresponding θ).

Figure OA.1: Accuracy of capital share distributions in the balanced growth path.

solved by our analytical approximation method (blue solid line) and that solved by the histogram method (black dashed line). Panels A and B show that when the yearly autocorrelation of idiosyncratic productivity $z_{i,t}$ is 0.5 or 0.7, the two curves are almost overlapping with each other, indicating that our analytical approximation method provides extremely high accuracy. Panel C shows that when the yearly autocorrelation is 0.85, as in our benchmark calibration, the endogenous capital share distributions solved by the two solution methods remain very similar.

If idiosyncratic productivity becomes more persistent, the capital share distribution solved by our analytical approximation method will further diverge from that solved by the histogram method. Intuitively, an extremely high persistence of idiosyncratic productivity will endogenously generate a fat right tail for the capital share distribution because productive firms will accumulate significant amounts of capital in the balanced growth path. This fat right tail cannot be well approximated by a log-normal distribution specified in equation (47) of the main text, resulting in relatively large numerical errors from our analytical approximation method. Despite the potential bad performance at extremely low values of θ (i.e., high values of persistence, $\exp(-\theta)$), Figure OA.1 shows that our analytical approximation method is sufficiently accurate for a wide range of empirically relevant θ estimated by (Asker, Collard-Wexler and Loecker, 2014) based on the U.S. data.

Table OA.2: Accuracy of our analytical approximation in stochastic steady states.

Variables	Baseline	2nd-order	3rd-order	4th-order
Firm profitability, $\mathbb{E}[\kappa_t]$	0.035	0.035	0.034	0.035
Productivity cutoff, $\mathbb{E}[\log(\underline{z}_t)]$	0.437	0.410	0.465	0.390
Value of blueprints, $\mathbb{E}[v_t]$	0.332	0.331	0.331	0.330
Productivity, $\mathbb{E}[\ln(H_t)]$	0.486	0.494	0.514	0.488
Flow profit of innovators, $\mathbb{E}[\pi_t]$	0.071	0.071	0.072	0.071
Wage-capital ratio, $\mathbb{E}[\ln(w_t/A_t)]$	-1.120	-1.155	-1.152	-1.168
Dividend-capital ratio, $\mathbb{E}[\ln(D_t/A_t)]$	-2.892	-2.837	-2.830	-2.798
Consumption-capital ratio, $\mathbb{E}[\ln(C_t/A_t)]$	-0.963	-0.984	-0.980	-0.988
R&D-capital ratio, $\mathbb{E}[\ln(S_t/A_t)]$	-1.611	-1.662	-1.662	-1.688
Knowledge stock-capital ratio, $\mathbb{E}[\ln(N_t/A_t)]$	1.011	0.977	0.967	0.964
Capital growth, $\mathbb{E}[\Delta \ln A_t]$ (%)	1.883	1.617	1.789	1.564
Consumption growth, $\mathbb{E}[\Delta \ln C_t]$ (%)	1.882	1.616	1.788	1.564
Volatility of consumption growth, $\text{var}[\Delta \ln C_t]$ (%)	2.585	2.206	1.725	1.828
Autocorrelation in consumption, $AC1(\Delta \ln C_t)$	0.484	0.529	0.503	0.459
Risk-free rate, $\mathbb{E}[r_{f,t}]$ (%)	1.473	1.503	1.733	1.536
Return on wealth, $\mathbb{E}[\ln(R_{w,t})]$ (%)	2.502	2.208	2.183	1.919
Volatility of return on wealth, $\text{var}[\ln(R_{w,t})]$ (%)	2.765	2.454	1.936	1.869

Note: This table evaluates the accuracy of our analytical approximation method in the full stochastic steady states with aggregate shocks. The column labeled "baseline" presents the steady-state values of corresponding variables solved by our analytical approximation method. The columns labeled "2nd-order", "3rd-order", and "4th-order" present the results corresponding to the second-, third-, and fourth-order numerical approximation methods for the capital share distribution $\omega_t(z)$ based on equation (241). All the parameter values are taken from our benchmark calibration in Table 1.

B.3 Stochastic Steady States with Aggregate Shocks

Per our discussions in Subsection B.1, there are three numerical approximation methods that we can benchmark to when assessing the accuracy of our analytical approximation method. Among these three methods, we choose the most recent method proposed by Winberry (2018) for its numerical tractability and stability.[20] The main advantages of this method are the fast computation speed and flexibility in the degree of approximation for the cross-sectional distribution.

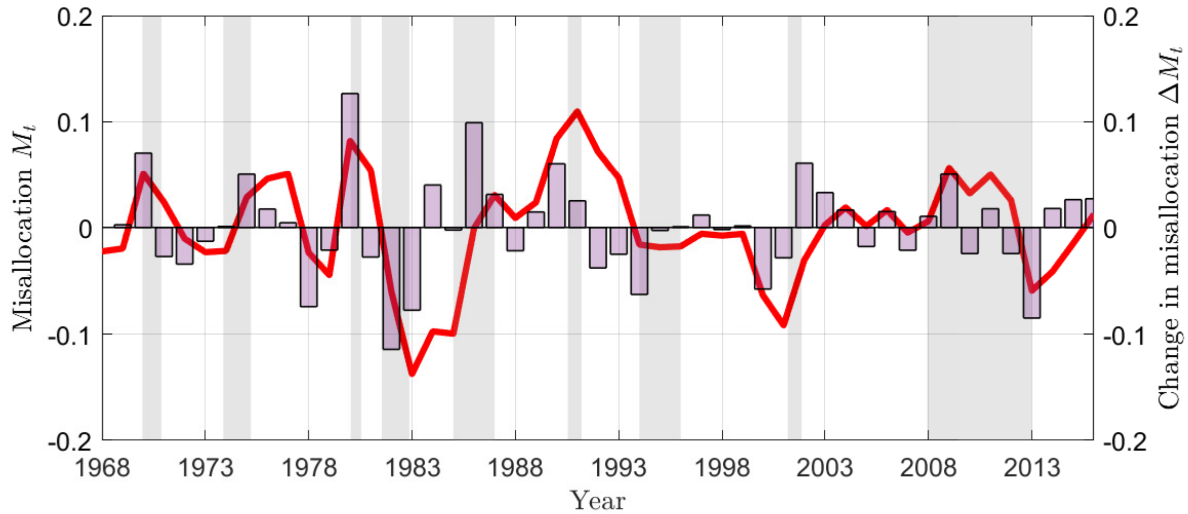
[20]Because of our model has large nonlinearities introduced by the endogenous SDF, the projection method of Algan, Allais and Den Haan (2008) is difficult to implement at higher-order approximations. Moreover, our model features kinked decision rules introduced by the endogenous time-varying productivity cutoff $\underline{z}_t$. Thus, the method of Reiter (2009) is difficult to be implemented accurately as the fixed grids of histograms generate large numerical errors around the cutoff $\underline{z}_t$ when using local pertubation approaches to approximate the evolution of aggregate state space.

In particular, we use a flexible parametric family to approximate the capital share $\omega_t(z)$ at any point in time t. With a 4th-order approximation, this boils down to using 4 parameters to match the mean, variance, skewness, and kurtosis of $\omega_t(z)$. The evolution of these 4 parameters can be computed using numerical integration methods (Gauss-Hermite and Gauss Legendre quadratures) based on Kolmogorov forward equations. The implementation details of this method are presented in Online Appendix G. Consistent with Winberry (2018), we find that a 4th-order approximation for the distribution can sufficiently capture complicated shapes, and further increasing the degree of approximation will not change the model-implied dynamics much.

In Table OA.2, we compare the key variables solved by our analytical approximation method and those solved by the higher-degree numerical approximation methods. To ensure that the results obtained from the two solution methods are comparable, we implement both methods using the perturbation approach around the corresponding deterministic steady state. Table OA.2 presents the results of our baseline analytical approximation method and the 2nd-order, 3rd-order, and 4th-order numerical approximation methods. The results of our analytical approximation method are very similar to the results of the 2nd-order numerical approximation method, because our analytical approximation method essentially keeps track of the first and second moments of $\omega_t(z)$.[21] However, the results of our analytical approximation method do not exactly match the results of the 2nd-order numerical approximation method because there are fundamental differences when implementing the two methods. A discussion is provided in Online Appendix G.

Table OA.2 also shows that the differences between our analytical approximation method and the 4th-order numerical approximation method are generally within 15% for most variables. Notably, the persistence in aggregate consumption growth, a key variable of our interest, has a yearly autocorrelation of 0.484, which is quite close to the value of 0.459 implied by the 4th-order numerical approximation method .

[21]The first moment is $m_{1,t} = -M_t\sigma^2/2$ and the second moment is $m_{2,t} = \sigma^2/2$, which is a constant under our analytical approximation method.



Note: The red solid line plots the time series of our empirical misallocation measure M_t (corresponding to the left y-axis). The measure M_t is constructed according to Section 4.1 of the main text, where $\widetilde{a}_{i,t}$ is constructed using firms' tangible net worth (see footnote 13). The pink bars represent its year-on-year changes ΔM_t (corresponding to the right y-axis). The shaded areas represent recessions or severe financial crises.

Figure OA.2: Time series plot of our empirical misallocation measure M_t.

C Supplemental Material for Empirical Analyses

C.1 Robustness of Empirical Results

Figure OA.2 plots the time series of our misallocation measure $MisAlloc_t$. The measure $MisAlloc_t$ is constructed according to Section 4.1 of the main text, where $a_{i,t}$ is constructed using firms' tangible net worth, i.e., firms' current assets plus net physical plant, property, and equipment plus other assets minus total liabilities.

Table OA.3 shows that the empirical findings in Table 5 of the main text are robust for an alternative sample period from 1970 to 2016.

C.2 Procedure for Nearest Neighbor Matching

For each SIC-3 industry s, we calculate the average industry characteristics during the 3-year period before AJCA (i.e., from 2001 to 2003), $\overline{X}_s = \frac{1}{3}\sum_{t=2001}^{2003} X_{s,t}$, where $X_{s,t}$ is a vector of eight industry characteristics, including the Herfindahl index computed using firms' market shares in terms of sales, average total sales of firms, mean and standard deviation of firms' profit margin, mean and standard deviation of firms' ROE, mean

Table OA.3: Misallocation, R&D, and growth in the data.

Panel A: R&D intensity		
	t	$t+1$
β	-0.073^{**}	-0.071^{**}
	[0.029]	[0.029]
R-squared	0.121	0.117

Panel B: Consumption growth					
	$t \to t+1$	$t \to t+2$	$t \to t+3$	$t \to t+4$	$t \to t+5$
β	-0.033	-0.059	-0.100^{*}	-0.156^{***}	-0.205^{***}
	[0.024]	[0.041]	[0.052]	[0.059]	[0.064]
R-squared	0.039	0.044	0.076	0.139	0.190

Panel C: Output growth					
	$t \to t+1$	$t \to t+2$	$t \to t+3$	$t \to t+4$	$t \to t+5$
β	-0.031	-0.061	-0.098	-0.159^{*}	-0.214^{**}
	[0.040]	[0.064]	[0.077]	[0.083]	[0.084]
R-squared	0.013	0.020	0.035	0.078	0.132

Note: The sample period is 1970-2016. Standard errors are reported in brackets. *, **, and *** indicate statistical significance at 10%, 5%, and 1%, respectively.

and standard deviation of firms' Tobin's Q. We construct a firm's (net) profit margin using its income before extraordinary items divided by its sales as in Dou, Ji and Wu (2021), and a firm's Tobin's Q as $Tobin\_Q_{i,t} = (total\_assets_{i,t} + market\_equity_{i,t} - book\_equity_{i,t})/total\_asset_{i,t}$, following Gompers, Ishii and Metrick (2003).

Next, we match each treated industry with $n = 5$ untreated industries which have the shortest Mahalanobis distances from the treated industry. The Mahalanobis distance between any two industries s and r is given by $\sqrt{(\overline{X}_s - \mu)'\Omega^{-1}(\overline{X}_r - \mu)}$, where $\overline{X}_s$ and $\overline{X}_r$ represent the vectors of the eight characteristics of industries s and r, and μ and Ω represent the mean vector and covariance matrix of the eight characteristics. This matching process is performed with replacement in untreated industries.

The DID specifications (60) to (63) are estimated with the following weights. Each treated industry is assigned with a weight of 1 and each of the 5 untreated industries matched to it is assigned with a weight of 1/5. Because we allow for replacement, some untreated industries could be matched with multiple treated industries. The weight for such industries is the sum of weights across matches. For example, if an untreated industry is matched with k treated industries, its weight is $k/5$. If an untreated industry is not matched with any treated industry, its weight is 0.

D TFP Formulas

Our TFP Z_t in equation (41) of the main text depends on the final goods sector's productivity H_t (can be thought of as the TFP of the final goods sector) and the economy's aggregate knowledge stock N_t. In this appendix section, we show that our formula for H_t is consistent with the TFP formula of Hsieh and Klenow (2009) when goods are homogeneous and the industry is not distorted by wedges.

The final-goods sector's productivity H_t given by equation (41) of the main text is equivalent to (88) with $u_t(z)$ being set at its optimal value, 1:

$$H_t = \left[\frac{1}{K_t} \int_{\underline{z}_t}^{\infty} \int_0^{\infty} z k_t(a,z) \varphi_t(a,z) \mathrm{d}k \mathrm{d}z \right]^{\alpha}, \tag{180}$$

The above equation is equivalent to

$$H_t = \left[\frac{1}{K_t} \int_0^{\infty} \int_0^{\infty} z k_t(a,z) \varphi_t(a,z) \mathrm{d}k \mathrm{d}z \right]^{\alpha}, \tag{181}$$

because $k_t(a,z) = 0$ for $z \leq \underline{z}_t$ according to equation (23) of the main text. Without loss of generality, we rewrite equation (181) to focus on a countable number of firms,

$$H_t = \left(\frac{1}{K_t} \sum_i z_i k_i \right)^{\alpha}, \tag{182}$$

We do a change of variables by replacing z_i with $z_i^{1/\alpha}$ (this is because the firm-level productivity in equation (1) of the main text is z_i^{α} not z_i), equation (182) becomes

$$H_t = \left(\frac{1}{K_t} \sum_i z_i^{1/\alpha} k_i \right)^{\alpha}, \tag{183}$$

Next, we show that equation (183) is consistent with the industry-level TFP formula used by Hsieh and Klenow (2009) when goods are homogeneous and the industry is not distorted by wedges. In the model of Hsieh and Klenow (2009), there are s

industries and each industry has M_s firms. They define a single industry's TFP as

$$TFP_s = \frac{Y_s}{K_s^{\alpha_s} L_s^{1-\alpha_s}}. \tag{184}$$

To be consistent with our model setup, we focus on deriving TFP_s in one single industry in the model of Hsieh and Klenow (2009), which corresponds to our final goods sector. Moreover, without loss of generality, we also normalize the aggregate labor in industry s to one, i.e., $L_s = 1$, as in our model. We derive the formula of TFP_s using the original notations of Hsieh and Klenow (2009).

Substituting $L_s = 1$ and equation (3) of Hsieh and Klenow (2009) into (184),

$$TFP_s = \frac{1}{K_s^{\alpha_s}} \left(\sum_{i=1}^{M_s} Y_{si}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}. \tag{185}$$

Substituting equation (4) of Hsieh and Klenow (2009) into the above equation,

$$TFP_s = \frac{1}{K_s^{\alpha_s}} \left[\sum_{i=1}^{M_s} \left(A_{si} K_{si}^{\alpha_s} L_{si}^{1-\alpha_s} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}. \tag{186}$$

Using the first-order condition, labor L_{si} in the model of Hsieh and Klenow (2009) can be solved as follows

$$L_{si} = \left[\frac{(1-\tau_{Ysi}) P_{si} A_{si} K_{si}^{\alpha_s} (1-\alpha_s)}{w} \right]^{\frac{1}{\alpha_s}}. \tag{187}$$

Substituting equation (187) into (186), we obtain

$$TFP_s = \frac{1}{K_s^{\alpha_s}} \left(\frac{1-\alpha_s}{w} \right)^{\frac{1-\alpha_s}{\alpha_s}} \left[\sum_{i=1}^{M_s} \left[A_{si}^{1/\alpha_s} K_{si} \left[(1-\tau_{Ysi}) P_{si} \right]^{\frac{1-\alpha_s}{\alpha_s}} \right]^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}. \tag{188}$$

The labor market clearing condition in the model of Hsieh and Klenow (2009) implies

$$\sum_{i=1}^{M_s} \left[\frac{(1-\tau_{Ysi}) P_{si} A_{si} K_{si}^{\alpha_s} (1-\alpha_s)}{w} \right]^{\frac{1}{\alpha_s}} = 1. \tag{189}$$

Substituting (189) into (188),

$$TFP_s = \frac{1}{K_s^{\alpha_s}} \frac{\left[\sum_{i=1}^{M_s} \left[A_{si}^{1/\alpha_s} K_{si} \left[(1-\tau_{Ysi})P_{si}\right]^{\frac{1-\alpha_s}{\alpha_s}}\right]^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}}{\left[\sum_{i=1}^{M_s} A_{si}^{1/\alpha_s} K_{si} \left[(1-\tau_{Ysi})P_{si}\right]^{1/\alpha_s}\right]^{1-\alpha_s}}. \tag{190}$$

Let $\sigma \to \infty$ and $\tau_{Ysi} = 0$, then P_{si} is equalized across all i, i.e., $P_{si} \equiv P_s$. This assumption allows us to simplify equation (190) as follows,

$$TFP_s = \frac{1}{K_s^{\alpha_s}} \frac{\sum_{i=1}^{M_s} A_{si}^{1/\alpha_s} K_{si}}{\left(\sum_{i=1}^{M_s} A_{si}^{1/\alpha_s} K_{si}\right)^{1-\alpha_s}} = \left[\frac{1}{K_s} \sum_{i=1}^{M_s} A_{si}^{1/\alpha_s} K_{si}\right]^{\alpha_s}. \tag{191}$$

Except for notational differences, the formula (191) is identical to (183).

E Numerical Algorithm

Our model can be solved either using a local perturbation approach or a global approach based on value function iterations.[22] In this appendix section, we present the numerical algorithm for the global approach based on value function iterations.

We discretize the model with time interval Δt. The Brownian motion shock $\mathrm{d}W_t$ takes two value, $\sqrt{\Delta t}$ and $-\sqrt{\Delta t}$, with equal probabilities. Define $\Gamma_t \equiv \mathrm{Cov}(\widetilde{a}_{i,t}, \widetilde{z}_{i,t}) = -M_t \mathrm{var}(\widetilde{z}_{i,t}) = -M_t \sigma^2/2$. The economy is summarized by the evolution of two endogenous state variables, $E_t \equiv N_t/A_t$ and Γ_t.

We use superscripts $+$ and $-$ to denote variables at $t + \Delta t$, corresponding to $\mathrm{d}W_t = \sqrt{\Delta t}$ and $\mathrm{d}W_t = -\sqrt{\Delta t}$, respectively. The endogenous state variable Γ_t evolves according to equation (150):

$$\Gamma_{t+\Delta t} = \Gamma_t - \theta \Gamma_t \mathrm{d}t + \mathrm{Cov}(\widetilde{z}_{i,t}, \mathrm{d}\widetilde{a}_{i,t}), \tag{192}$$

[22] Because the aggregate dynamics do not feature occasionally binding constraints or region-dependent policy rules, the local perturbation approach can be easily implemented in *dynare*.

where $\mathrm{Cov}(\widetilde{z}_{i,t}, \mathrm{d}\widetilde{a}_{i,t})$ is given by equation (159), as follows:

$$\begin{aligned}\mathrm{Cov}(\widetilde{z}_{i,t}, \mathrm{d}\widetilde{a}_{i,t}) =& \frac{(1+\lambda)\sigma^2\kappa_t}{2} \exp\left(\frac{\sigma^2}{4}\right) \Phi\left(\frac{\sigma^2/2 - \widetilde{\underline{z}}_t}{\sigma/\sqrt{2}}\right) \Delta t \\ &+ \frac{(1+\lambda)\sigma}{2\sqrt{\pi}} [(\underline{z}_t \kappa_t - r_{f,t} - \delta_k)\Delta t - \sigma_k \mathrm{d}W_t] \exp\left(-\frac{\widetilde{\underline{z}}_t^2}{\sigma^2}\right). \end{aligned} \tag{193}$$

Let $\Gamma^+_{t+\Delta t}$ and $\Gamma^-_{t+\Delta t}$ be the value of $\Gamma_{t+\Delta t}$ corresponding to $\mathrm{d}W_t = \sqrt{\Delta t}$ and $\mathrm{d}W_t = -\sqrt{\Delta t}$, respectively. In equation (193), the variables κ_t, $\widetilde{\underline{z}}_t$, and $r_{f,t}$ are given by equations (103), (139), and the SDF, respectively, as follows:

$$\kappa_t = \alpha(1-\varepsilon) H_t^{-\frac{1}{\alpha}} \frac{Y_t}{A_t} \frac{A_t}{K_t}, \tag{194}$$

$$\widetilde{\underline{z}}_t = \Gamma_t - \Phi^{-1}\left(\frac{1}{1+\lambda} \frac{K_t}{A_t}\right) \frac{\sigma}{\sqrt{2}}, \tag{195}$$

$$r_{f,t} = -\frac{1}{\Delta t} \ln\left(\mathbb{E}_t\left[\frac{\Lambda_{t+\Delta t}}{\Lambda_t}\right]\right), \tag{196}$$

where Y_t/A_t, H_t, and K_t/A_t are functions of state variables E_t and Γ_t, given by equations (101), (142), and (28), respectively, as follows:

$$\frac{Y_t}{A_t} = (\varepsilon\nu)^{\frac{\varepsilon}{1-\varepsilon}} H_t E_t^{1-\alpha} \left(\frac{K_t}{A_t}\right)^{\alpha}, \tag{197}$$

$$H_t = \left[(1+\lambda)\frac{A_t}{K_t} \exp\left(\Gamma_t + \frac{\sigma^2}{4}\right) \Phi\left(\Phi^{-1}\left(\frac{1}{1+\lambda}\frac{K_t}{A_t}\right) + \frac{\sigma}{\sqrt{2}}\right)\right]^{\alpha}, \tag{198}$$

$$\underline{z}_t \kappa_t = r_{f,t} + \delta_k + \sigma_k(\sigma_{\xi,t}(\underline{z}_t) - \eta_t). \tag{199}$$

The endogenous state variable E_t evolves according to

$$\frac{\Delta E_t}{E_t} = \frac{\Delta N_t}{N_t} - \frac{\Delta A_t}{A_t}. \tag{200}$$

Substituting equations (146) and (147) into the above equation, we obtain

$$\begin{aligned}\frac{E_{t+\Delta t}}{E_t} =& 1 + \chi\left(\chi v_t\right)^{\frac{1-h}{h}} \Delta t - \alpha(1-\varepsilon)\frac{Y_t}{A_t}\Delta t + (r_{f,t} + \delta_k)\frac{K_t}{A_t}\Delta t \\ &+ (\rho + \delta_a - \delta_b - r_{f,t})\Delta t - \left(\sigma_a - \sigma_k \frac{K_t}{A_t}\right) \mathrm{d}W_t. \end{aligned} \tag{201}$$

Let $E^+_{t+\Delta t}$ and $E^-_{t+\Delta t}$ be the value of $E_{t+\Delta t}$ corresponding to $\mathrm{d}W_t = \sqrt{\Delta t}$ and $\mathrm{d}W_t = -\sqrt{\Delta t}$, respectively.

In equation (201), the variable $v_t = v(E_t, \Gamma_t)$ is given by equation (8); it is a function of state variables (E_t, Γ_t) and can be solved recursively as follows

$$\begin{aligned} v(E_t, \Gamma_t) =& \frac{1}{1+\delta_b \Delta t}\left(\pi_t \Delta t + \mathbb{E}_t\left[\frac{\Lambda_{t+\Delta t}}{\Lambda_t} v(E_{t+\Delta t}, \Gamma_{t+\Delta t})\right]\right) \\ =& \frac{1}{1+\delta_b \Delta t}\left(\pi_t \Delta t + \frac{1}{2}\frac{\Lambda^+_{t+\Delta t}}{\Lambda_t} v(E^+_{t+\Delta t}, \Gamma^+_{t+\Delta t}) + \frac{1}{2}\frac{\Lambda^-_{t+\Delta t}}{\Lambda_t} v(E^-_{t+\Delta t}, \Gamma^-_{t+\Delta t})\right], \end{aligned} \tag{202}$$

where π_t is given by equation (105):

$$\pi_t = \frac{(1-\nu)\varepsilon}{E_t}\frac{Y_t}{A_t}. \tag{203}$$

Epstein and Zin (1989) show that the SDF in equation (15) is equivalent to

$$\frac{\Lambda_{t+\Delta t}}{\Lambda_t} = e^{-\frac{\delta(1-\gamma)}{1-1/\psi}\Delta t}\left(\frac{C_{t+\Delta t}}{C_t}\right)^{-\frac{1-\gamma}{\psi(1-1/\psi)}}(1+R_{w,t+\Delta t}\Delta t)^{\frac{1/\psi-\gamma}{1-1/\psi}}, \tag{204}$$

where $R_{w,t+\Delta t}$ the net return on wealth

$$1 + R_{w,t+\Delta t}\Delta t = \frac{W_{t+\Delta t}}{W_t - C_t \Delta t}. \tag{205}$$

We have

$$\mathbb{E}_t\left[\frac{\Lambda_{t+\Delta t}}{\Lambda_t}(1+R_{w,t+\Delta}\Delta t)\right] = 1. \tag{206}$$

Substituting equations (204) and (205) into (206), we obtain

$$1 = \mathbb{E}_t\left[e^{-\frac{\delta(1-\gamma)}{1-1/\psi}\Delta t}\left(\frac{C_{t+\Delta t}}{C_t}\right)^{-\frac{1-\gamma}{\psi-1}}\left(\frac{W_{t+\Delta t}}{C_{t+\Delta t}}\frac{C_{t+\Delta t}}{C_t}\frac{1}{W_t/C_t - \Delta t}\right)^{\frac{1-\gamma}{1-1/\psi}}\right]. \tag{207}$$

Rearranging the above equation, we obtain

$$\frac{W_t}{C_t} = \Delta t + e^{-\delta \Delta t}\mathbb{E}_t\left[\left(\frac{C_{t+\Delta t}}{C_t}\right)^{1-\gamma}\left(\frac{W_{t+\Delta t}}{C_{t+\Delta t}}\right)^{\frac{1-\gamma}{1-1/\psi}}\right]^{\frac{1-1/\psi}{1-\gamma}}. \tag{208}$$

The wealth-consumption ratio W_t/C_t is a function of state variables, denoted by $WC_t \equiv WC(E_t, \Gamma_t)$. Let $C^+_{t+\Delta t}$ and $C^-_{t+\Delta t}$ be the value of $C_{t+\Delta t}$ corresponding to $\mathrm{d}W_t = \sqrt{\Delta t}$ and $\mathrm{d}W_t = -\sqrt{\Delta t}$, respectively. We can rewrite equation (208) as

$$WC_t = \Delta t + e^{-\delta \Delta t} \left[\frac{1}{2} \left(\frac{C^+_{t+\Delta t}}{C_t} \right)^{1-\gamma} (WC^+_{t+\Delta t})^{\frac{1-\gamma}{1-1/\psi}} + \frac{1}{2} \left(\frac{C^-_{t+\Delta t}}{C_t} \right)^{1-\gamma} (WC^-_{t+\Delta t})^{\frac{1-\gamma}{1-1/\psi}} \right], \tag{209}$$

where

$$WC^+_{t+\Delta t} = WC(E^+_{t+\Delta t}, \Gamma^+_{t+\Delta t}), \tag{210}$$

$$WC^-_{t+\Delta t} = WC(E^-_{t+\Delta t}, \Gamma^-_{t+\Delta t}). \tag{211}$$

The aggregate consumption is given by equation (14):

$$\begin{aligned} \frac{C_t}{A_t} =& \frac{w_t}{A_t} + \frac{D_t}{A_t} + r_{f,t} \frac{B_t}{A_t} - \left(\frac{B_{t+\Delta t}}{A_{t+\Delta t}} \frac{A_{t+\Delta t}}{A_t} - \frac{B_t}{A_t} \right) \frac{1}{\Delta t} \\ =& \frac{w_t}{A_t} + \frac{D_t}{A_t} + r_{f,t} \left(\frac{K_t}{A_t} - 1 \right) - \left[\left(\frac{K_{t+\Delta t}}{A_{t+\Delta t}} - 1 \right) \frac{A_{t+\Delta t}}{A_t} - \left(\frac{K_t}{A_t} - 1 \right) \right] \frac{1}{\Delta t}. \end{aligned} \tag{212}$$

Because C_t is known (i.e., $\mathrm{d}B_t/B_t$ is locally deterministic), theoretically we have

$$\left(\frac{K^+_{t+\Delta t}}{A^+_{t+\Delta t}} - 1 \right) \frac{A^+_{t+\Delta t}}{A_t} = \left(\frac{K^-_{t+\Delta t}}{A^-_{t+\Delta t}} - 1 \right) \frac{A^-_{t+\Delta t}}{A_t}, \tag{213}$$

where $K^+_{t+\Delta t}$, $A^+_{t+\Delta t}$ and $K^-_{t+\Delta t}$, $A^-_{t+\Delta t}$ are the values of $K_{t+\Delta t}$, $A_{t+\Delta t}$ corresponding to $\mathrm{d}W_t = \sqrt{\Delta t}$ and $\mathrm{d}W_t = -\sqrt{\Delta t}$, respectively. Because of property (213), the numerical error caused by discretization is minimized by using $0.5 \left(\frac{K^+_{t+\Delta t}}{A^+_{t+\Delta t}} - 1 \right) \frac{A^+_{t+\Delta t}}{A_t} + 0.5 \left(\frac{K^-_{t+\Delta t}}{A^-_{t+\Delta t}} - 1 \right) \frac{A^-_{t+\Delta t}}{A_t}$ to approximate $\left(\frac{K_{t+\Delta t}}{A_{t+\Delta t}} - 1 \right) \frac{A_{t+\Delta t}}{A_t}$ in equation (212). Thus, the term $C_t/A_t \equiv CA(E_t, \Gamma_t)$ in equation (212) can be solved as a function of state variables E_t and Γ_t.

The consumption growth terms in equation (209) are given by

$$\frac{C^+_{t+\Delta t}}{C_t} = \frac{CA(E^+_{t+\Delta t}, \Gamma^+_{t+\Delta t})}{CA(E_t, \Gamma_t)} \frac{A_{t+\Delta t}}{A_t}, \tag{214}$$

$$\frac{C^-_{t+\Delta t}}{C_t} = \frac{CA(E^-_{t+\Delta t}, \Gamma^-_{t+\Delta t})}{CA(E_t, \Gamma_t)} \frac{A_{t+\Delta t}}{A_t}. \tag{215}$$

The variables w_t/A_t and D_t/A_t are given by equations (44) and (31) of the main text:

$$\frac{w_t}{A_t} \equiv wA(E_t, \Gamma_t) = (1-\alpha)(1-\varepsilon)\frac{Y_t}{A_t}, \tag{216}$$
$$\frac{D_t}{A_t} \equiv DA(E_t, \Gamma_t) = \rho + (1-\nu)\varepsilon\frac{Y_t}{A_t} - \frac{S_t}{A_t}, \tag{217}$$

where S_t/A_t is given by equation (46) of the main text:

$$\frac{S_t}{A_t} = \frac{S_t}{N_t}E_t = \left(\chi v(E_t, \Gamma_t)\right)^{\frac{1}{h}} E_t. \tag{218}$$

The variables $A_{t+\Delta t}/A_t$ is given by equation (146):

$$\frac{A_{t+\Delta t}}{A_t} = 1 + \alpha(1-\varepsilon)\frac{Y_t}{A_t}\Delta t - (r_{f,t} + \delta_k)\frac{K_t}{A_t}\Delta t - (\rho + \delta_a - r_{f,t})\Delta t + \left(\sigma_a - \sigma_k\frac{K_t}{A_t}\right)\mathrm{d}W_t. \tag{219}$$

After solving the $WC(E_t, \Gamma_t)$ ratio from equation (209), substituting into the equation (204) to obtain the SDF:

$$\frac{\Lambda^+_{t+\Delta t}}{\Lambda_t} = e^{-\frac{\delta(1-\gamma)}{1-1/\psi}\Delta t}\left(\frac{C^+_{t+\Delta t}}{C_t}\right)^{-\gamma}\left(\frac{WC(E^+_{t+\Delta t}, \Gamma^+_{t+\Delta t})}{WC(E_t, \Gamma_t) - \Delta t}\right)^{\frac{1/\psi-\gamma}{1-1/\psi}}, \tag{220}$$

$$\frac{\Lambda^-_{t+\Delta t}}{\Lambda_t} = e^{-\frac{\delta(1-\gamma)}{1-1/\psi}\Delta t}\left(\frac{C^-_{t+\Delta t}}{C_t}\right)^{-\gamma}\left(\frac{WC(E^-_{t+\Delta t}, \Gamma^-_{t+\Delta t})}{WC(E_t, \Gamma_t) - \Delta t}\right)^{\frac{1/\psi-\gamma}{1-1/\psi}}. \tag{221}$$

Welfare. In discrete time, the preference specified in equation (12) of the main text is

$$U_t = \left[(1-e^{-\delta\Delta t})C_t^{1-1/\psi} + e^{-\delta\Delta t}\left(\mathbb{E}_t\left[(U_{t+\Delta t})^{1-\gamma}\right]\right)^{\frac{1-1/\psi}{1-\gamma}}\right]^{\frac{1}{1-1/\psi}}. \tag{222}$$

Dividing both sides by C_t,

$$\left(\frac{U_t}{C_t}\right)^{1-1/\psi} = (1-e^{-\delta\Delta t}) + e^{-\delta\Delta t}\left(\mathbb{E}_t\left[\left(\frac{C_{t+\Delta t}}{C_t}\frac{U_{t+\Delta t}}{C_{t+\Delta t}}\right)^{1-\gamma}\right]\right)^{\frac{1-1/\psi}{1-\gamma}}. \tag{223}$$

Steps of Implementing the Numerical Algorithm. Following the standard practice, we discretize the state variables (E_t, Γ_t) into dense grids. The values that not fall on any grid are obtained by linear interpolation or extrapolation. We then solve the model in the steps listed below. Because we need to solve a large number of nonlinear equations, we use the commercial nonlinear solver *knitro*.[23] All the programs are written in C++ with parallel computing in a state-of-the-art server of 56 cores.

(1) Guess $v(E_t, \Gamma_t) = 0.1$ for all states.

(2) Guess $\sigma_\xi(\underline{z}_t, E_t, \Gamma_t) = 0$ for all states.

(3) Guess $\eta(E_t, \Gamma_t) = 0$ for all states.

(4) Solve the evolution of endogenous state variables E_t and Γ_t.

(5) Solve equation (208) using *knitro* to obtain the wealth-consumption ratio as a function of state variables, i.e., $WC(E_t, \Gamma_t)$.

(6) Solve equations (220) and (221) to obtain the SDF as a function of state variables, i.e.,

$$\frac{\Lambda^+_{t+\Delta t}}{\Lambda_t} \equiv SDF(E^+_{t+\Delta t}, \Gamma^+_{t+\Delta t}), \tag{224}$$

$$\frac{\Lambda^-_{t+\Delta t}}{\Lambda_t} \equiv SDF(E^-_{t+\Delta t}, \Gamma^-_{t+\Delta t}). \tag{225}$$

Next, calculate the market price of risk η_t in equation (20) of the main text as follows

$$\widehat{\eta}(E_t, \Gamma_t) = -\frac{SDF(E^+_{t+\Delta t}, \Gamma^+_{t+\Delta t}) - SDF(E^-_{t+\Delta t}, \Gamma^-_{t+\Delta t})}{2\sqrt{\Delta t}}. \tag{226}$$

If $\max|\widehat{\eta}(E_t, \Gamma_t) - \eta(E_t, \Gamma_t)| < 10^{-9}$, stop. Otherwise, jump to step (4) using $\widehat{\eta}(E_t, \Gamma_t)$ as the initial guess of $\eta(E_t, \Gamma_t)$.

(7) Solve managers' problem in equation (18) of the main text to obtain $\sigma_\xi(\underline{z}_t, E_t, \Gamma_t)$. This is achieved in the following substeps.

(7.1) Problem (18) of the main text can be simplified because it is linear in $a_{i,t}$ as in equation (75). This means that we only need to solve $\xi(z_{i,t}, E_t, \Gamma_t)$ recursively

[23] See https://www.artelys.com/solvers/knitro for more details.

as follows

$$\xi(z_{i,t}, E_t, \Gamma_t) = \tau\Delta t + \mathbb{E}_t \left[\frac{\Lambda_{t+\Delta t}}{\Lambda_t} \frac{a_{i,t+\Delta t}}{a_{i,t}} \xi(z_{i,t+\Delta t}, E_{t+\Delta t}, \Gamma_{t+\Delta t}) \right] \tag{227}$$

The evolution $a_{i,t+\Delta t}/a_{i,t}$ is given by equations (2), (3), (30) of the main text:

$$\begin{aligned} \frac{a_{i,t+\Delta t}}{a_{i,t}} =& 1 + (1+\lambda)\left(\kappa_t z_{i,t}\mathrm{d}t - \delta_k \mathrm{d}t + \sigma_k \mathrm{d}W_t - r_{f,t}\mathrm{d}t\right) \mathbb{1}_{z_{i,t} \geq \underline{z}_t} \\ & + (r_{f,t} - \rho - \delta_a)\mathrm{d}t + \sigma_a \mathrm{d}W_t, \end{aligned} \tag{228}$$

Substituting equation (228) into (227), we obtain

$$\begin{aligned} \xi_{i,t} =& \tau\Delta t + \mathbb{E}_t \left[\frac{\Lambda_{t+\Delta t}}{\Lambda_t} \left[1 + (1+\lambda)\left(\kappa_t z_{i,t}\mathrm{d}t - \delta_k \mathrm{d}t + \sigma_k \mathrm{d}W_t - r_{f,t}\mathrm{d}t\right) \mathbb{1}_{z_{i,t} \geq \underline{z}_t}\right] \xi_{i,t+\Delta t} \right] \\ & + \mathbb{E}_t \left[\frac{\Lambda_{t+\Delta t}}{\Lambda_t} \left[(r_{f,t} - \rho - \delta_a)\mathrm{d}t + \sigma_a \mathrm{d}W_t\right] \xi_{i,t+\Delta t} \right] \end{aligned} \tag{229}$$

(7.2) Calculate $\widehat{\sigma}_\xi(z_{i,t}, E_t, \Gamma_t)$ as follows

$$\widehat{\sigma}_\xi(z_{i,t}, E_t, \Gamma_t) = \frac{\underline{\xi}^+_{t+\Delta t} - \underline{\xi}^-_{t+\Delta t}}{2\xi(z_{i,t}, E_t, \Gamma_t)\sqrt{\Delta t}}, \tag{230}$$

where

$$\underline{\xi}^+_{t+\Delta t} = \mathbb{E}_t \left[\xi(z_{i,t+\Delta t}, E^+_{t+\Delta t}, \Gamma^+_{t+\Delta t})\right], \tag{231}$$

$$\underline{\xi}^-_{t+\Delta t} = \mathbb{E}_t \left[\xi(z_{i,t+\Delta t}, E^-_{t+\Delta t}, \Gamma^-_{t+\Delta t})\right]. \tag{232}$$

The expectation is taken with respect to idiosyncratic shocks in $z_{i,t+\Delta t}$.

(7.3) Solve $\underline{z}(E_t, \Gamma_t)$ using equation (199), and then find the value of $\widehat{\sigma}_\xi(\underline{z}_t, E_t, \Gamma_t)$.

(7.4) If $\max\left|\widehat{\sigma}_\xi(\underline{z}_t, E_t, \Gamma_t) - \sigma_\xi(\underline{z}_t, E_t, \Gamma_t)\right| < 10^{-9}$, stop. Otherwise, jump to step (3) using $\widehat{\sigma}_\xi(\underline{z}_t, E_t, \Gamma_t)$ as the initial guess for $\sigma_\xi(\underline{z}_t, E_t, \Gamma_t)$.

(8) Solve equation (202) to obtain $\widehat{v}(E_t, \Gamma_t)$.

(9) If $\max|\widehat{v}(E_t, \Gamma_t) - v(E_t, \Gamma_t)| < 10^{-9}$, stop. Otherwise, jump to step (2) using $\widehat{v}(E_t, \Gamma_t)$ as the initial guess for $v(E_t, \Gamma_t)$.

F Household Budget Constraint

Consider a household h with wealth W_t^h at t. The budget constraint is

$$W_{t+\mathrm{d}t}^h = W_t^h - C_t^h \mathrm{d}t + w_t L_t^h \mathrm{d}t + (Q_{t+\mathrm{d}t} - Q_t)\mathbb{Z}_t^h + D_t \mathbb{Z}_t^h \mathrm{d}t + r_{f,t} B_t^h \mathrm{d}t, \tag{233}$$

where $C_t^h \mathrm{d}t$ is the household's consumption over $[t, t+\mathrm{d}t)$, which is assumed to be locally deterministic. The variable $w_t L_t^h \mathrm{d}t$ is the labor income over $[t, t+\mathrm{d}t)$. The variable $(Q_{t+\mathrm{d}t} - Q_t)\mathbb{Z}_t^h$ is the change in the household's stock value, where Q_t is the stock market value per share and $\mathbb{Z}_t^h$ is the number of shares held by the household at t. The variable $D_t \mathbb{Z}_t^h \mathrm{d}t$ is the dividend and $r_{f,t} B_t^h \mathrm{d}t$ is the interest earnings over $[t, t+\mathrm{d}t)$.

The wealth W_t^h consists bonds B_t^h and a share $\mathbb{Z}_t^h$ of the stock market:

$$W_t^h = Q_t \mathbb{Z}_t^h + B_t^h, \tag{234}$$

Substituting equations (234) into (233), we obtain

$$Q_{t+\mathrm{d}t}\mathbb{Z}_{t+\mathrm{d}t}^h + B_{t+\mathrm{d}t}^h = -C_t^h \mathrm{d}t + w_t L_t^h \mathrm{d}t + Q_{t+\mathrm{d}t}\mathbb{Z}_t^h + D_t \mathbb{Z}_t^h \mathrm{d}t + (1 + r_{f,t}\mathrm{d}t) B_t^h. \tag{235}$$

Aggregating equation (235) over all households, we obtain

$$C_t \mathrm{d}t + Q_{t+\mathrm{d}t}\mathbb{Z}_{t+\mathrm{d}t} + B_{t+\mathrm{d}t} = w_t L_t \mathrm{d}t + Q_{t+\mathrm{d}t}\mathbb{Z}_t + D_t \mathbb{Z}_t \mathrm{d}t + (1 + r_{f,t}\mathrm{d}t) B_t. \tag{236}$$

In equilibrium, the total share is normalized to be one:

$$\mathbb{Z}_t \equiv 1 \quad \text{for all} \quad t. \tag{237}$$

Thus, equation (236) becomes

$$\mathrm{d}B_t = w_t L_t \mathrm{d}t + D_t \mathrm{d}t + r_{f,t} B_t \mathrm{d}t - C_t \mathrm{d}t, \tag{238}$$

which is the equation (14) of the main text. Equation (238) shows that $\mathrm{d}B_t/\mathrm{d}t$ is locally deterministic because of our assumption that C_t is locally deterministic.

To compute $\mathrm{d}B_t$, we use equation (42) of the main text,

$$B_t = K_t - A_t = A_t\left(\frac{K_t}{A_t} - 1\right) = [\lambda - (1+\lambda)\Omega_t(\underline{z}_t)]A_t. \tag{239}$$

Thus,

$$\mathrm{d}B_t = [\lambda - (1+\lambda)\Omega_t(\underline{z}_{t+\mathrm{d}t})]A_{t+\mathrm{d}t} - [\lambda - (1+\lambda)\Omega_t(\underline{z}_t)]A_t. \tag{240}$$

The diffusion term on the right-hand side of equation (240) will cancel out because $\mathrm{d}B_t$ is locally deterministic.

G Higher-Degree Approximation

Following Algan, Allais and Den Haan (2008), we use the following functional form to approximate the capital share distribution $\omega_t(\widetilde{z})$ defined in equation (38) of the main text with $\widetilde{z} = \ln(z)$:

$$\omega_t(\widetilde{z}) \approx g_{0,t} \exp\left(g_{1,t}(\widetilde{z} - m_{1,t}) + \sum_{i=2}^{n}\left[g_{i,t}(\widetilde{z} - m_{1,t})^i - m_{i,t}\right]\right), \tag{241}$$

where $m_{1,t}$, ..., $m_{n,t}$ correspond to the 1st, ..., nth moments of $\omega_t(\widetilde{z})$, given by

$$m_{1,t} = \int_{-\infty}^{\infty} \widetilde{z}\omega_t(\widetilde{z})\mathrm{d}\widetilde{z}, \tag{242}$$

$$m_{i,t} = \int_{-\infty}^{\infty} (\widetilde{z} - m_{1,t})^i \omega_t(\widetilde{z})\mathrm{d}\widetilde{z} \text{ for } i = 2, ..., n. \tag{243}$$

When $n = 2$, the approximation based on equation (241) is similar to our analytical approximation method in equation (47) of the main text, with $m_{1,t} = -M_t\sigma^2/2$ and $m_{2,t} = \sigma^2/2$.[24]

[24]Even when $n = 2$, the numerical approximation method does not produce identical results as our analytical approximation method (see Table OA.2). This is because there is a subtle difference between the two methods. In our analytical approximation method, by assuming a normal distribution for $\widetilde{a}$, we derive a normal-density function of $\omega_t(\widetilde{z})$ for all $t \geq 0$. Then we track the evolution of $\Gamma_t = -M_t\sigma^2/2$, which captures the first moment $m_{1,t}$ of $\omega_t(\widetilde{z})$. In the numerical approximation method with $n = 2$, we fit $\omega_t(\widetilde{z})$ at t using a normal density function as specified by equation (241), and then we compute the non-parametric distribution of $\omega_{t+\Delta t}(\widetilde{z})$ at $t + \Delta t$ based on the evolution of $\widetilde{z}_{i,t}$ and $\widetilde{a}_{i,t}$. Next, we fit $\omega_{t+\Delta t}(\widetilde{z})$ using a normal density function by matching the first and second moments, $m_{1,t+\Delta t}$ and $m_{2,t+\Delta t}$, implied by $\omega_{t+\Delta t}(\widetilde{z})$. Because of this subtle difference, the results of our analytical approximation method are slightly different from those of the numerical approximation method with $n = 2$ (see Table OA.2).

The evolution of $m_{i,t}$ for $i = 1, 2, ..., n$ can be derived as follows. Consider a small time interval $[t, t + \Delta t)$, equation (5) of the main text implies that $\widetilde{z}_{i,t+\Delta t}$ is given by

$$\widetilde{z}_{i,t+\Delta t} = (1 - \theta\Delta t)\widetilde{z}_{i,t} + \sigma\sqrt{\theta\Delta t}\varepsilon_{i,t}, \quad \text{with } \varepsilon_{i,t} \sim N(0,1). \tag{244}$$

Thus, conditioning on $\tilde{z}_{i,t}$ at t, the probability of having $z_{i,t+\Delta t}$ falling in a small interval $[\widetilde{z}, \widetilde{z} + \Delta\widetilde{z}]$ at $t + \Delta t$ is given by

$$\begin{aligned} P(z_{i,t+\Delta t} \in [\widetilde{z}, \widetilde{z} + \Delta\widetilde{z}] | \widetilde{z}_{i,t}) =& \Phi\left(\frac{\widetilde{z} + \Delta\widetilde{z} - (1 - \theta\Delta t)\widetilde{z}_{i,t}}{\sigma\sqrt{\theta\Delta t}}\right) - \Phi\left(\frac{\widetilde{z} - (1 - \theta\Delta t)\widetilde{z}_{i,t}}{\sigma\sqrt{\theta\Delta t}}\right) \\ =& \phi\left(\frac{\widetilde{z} - (1 - \theta\Delta t)\widetilde{z}_{i,t}}{\sigma\sqrt{\theta\Delta t}}\right)\frac{\Delta\widetilde{z}}{\sigma\sqrt{\theta\Delta t}}. \end{aligned} \tag{245}$$

Equations (2), (3), and (30) of the main text imply that $a_{i,t+\Delta t}$ is given by

$$\frac{a_{i,t+\Delta t}}{a_{i,t}} = [1 + (r_{f,t} - \delta_a - \rho)\Delta t] + \sigma_a\Delta W_t + (1 + \lambda)\left[(\kappa_t z_{i,t} - \delta_k - r_{f,t})\Delta t - \sigma_k\Delta W_t)\right]\mathbb{1}_{z_{i,t} \geq \underline{z}_t}. \tag{246}$$

Thus, conditioning on $\widetilde{z}_{i,t}$ at t and given the aggregate shock ΔW_t, to have $a_{i,t+\Delta t} \in [a, a + \Delta a]$, we need

$$a_{i,t} \in \left[\frac{a}{\Psi_t(\widetilde{z}_{i,t})}, \frac{a + \Delta a}{\Psi_t(\widetilde{z}_{i,t})}\right], \tag{247}$$

where

$$\Psi_t(\widetilde{z}_{i,t}) = [1 + (r_{f,t} - \delta_a - \rho)\Delta t] + \sigma_a\Delta W_t + (1 + \lambda)\left[(\kappa_t z_{i,t} - \delta_k - r_{f,t})\Delta t - \sigma_k\Delta W_t)\right]\mathbb{1}_{z_{i,t} \geq \underline{z}_t}. \tag{248}$$

Thus, the density $\varphi_{t+\Delta t}(a, \tilde{z})$ is given by

$$\varphi_{t+\Delta t}(a, \widetilde{z})\Delta a\Delta\tilde{z} = \int_{-\infty}^{\infty}\varphi_t(a/\Psi_t(\widetilde{z}_{i,t}), \tilde{z}_{i,t})\frac{\Delta a}{\Psi_t(\widetilde{z}_{i,t})}P(z_{i,t+\Delta t} \in [\widetilde{z}, \widetilde{z} + \Delta\widetilde{z}] | \widetilde{z}_{i,t})\mathrm{d}\widetilde{z}_{i,t}. \tag{249}$$

Substituting equation (245) into (249), we obtain

$$\varphi_{t+\Delta t}(a, \widetilde{z}) = \frac{1}{\sigma\sqrt{\theta\Delta t}}\int_{-\infty}^{\infty}\frac{1}{\Psi_t(\widetilde{x})}\varphi_t(a/\Psi_t(\widetilde{x}), \widetilde{x})\phi\left(\frac{\widetilde{z} - (1 - \theta\Delta t)\widetilde{x}}{\sigma\sqrt{\theta\Delta t}}\right)\mathrm{d}\widetilde{x}. \tag{250}$$

By definition of equation (38) of the main text, the capital share at $t + \Delta t$ is

$$\omega_{t+\Delta t}(\widetilde{z}) = \frac{1}{A_{t+\Delta t}}\int_0^{\infty}a\varphi_{t+\Delta t}(a, \widetilde{z})\mathrm{d}a. \tag{251}$$

Substituting equation (250) into (251), we obtain

$$\omega_{t+\Delta t}(\widetilde{z}) = \frac{1}{\sigma\sqrt{\theta\Delta t}}\frac{1}{A_{t+\Delta t}}\int_{-\infty}^{\infty}\left(\int_{0}^{\infty}\frac{a}{\Psi_t(\widetilde{x})}\varphi_t(a/\Psi_t(\widetilde{x}),\widetilde{x})\mathrm{d}a\right)\phi\left(\frac{\widetilde{z}-(1-\theta\Delta t)\widetilde{x}}{\sigma\sqrt{\theta\Delta t}}\right)\mathrm{d}\widetilde{x}. \tag{252}$$

Define $a' = a/\Psi_t(\widetilde{x})$. Using the definition of (38), the term $\int_0^\infty \frac{a}{\Psi_{i,t}}\varphi_t(a/\Psi_t(\widetilde{x}),\widetilde{x})\mathrm{d}a$ in equation (252) can be written as

$$\int_{0}^{\infty}\frac{a}{\Psi_t(\widetilde{x})}\varphi_t(a/\Psi_t(\widetilde{x}),\widetilde{x})\mathrm{d}a = \Psi_t(\widetilde{x})\int_0^\infty a'\varphi_t(a',\widetilde{x})\mathrm{d}a' = \Psi_t(\widetilde{x})\omega_t(\widetilde{x})A_t. \tag{253}$$

Substituting equation (253) into (252), we obtain

$$\omega_{t+\Delta t}(\widetilde{z}) = \frac{1}{\sigma\sqrt{\theta\Delta t}}\frac{A_t}{A_{t+\Delta t}}\int_{-\infty}^{\infty}\Psi_t(\widetilde{x})\omega_t(\widetilde{x})\phi\left(\frac{\widetilde{z}-(1-\theta\Delta t)\widetilde{x}}{\sigma\sqrt{\theta\Delta t}}\right)\mathrm{d}\widetilde{x}, \tag{254}$$

where $\Psi_t(\widetilde{x})$ is defined in equation (248) with $\widetilde{x} = \ln(x)$.

Using $\omega_{t+\Delta t}(\widetilde{z})$ in equation (254), we can compute the moments at $t+\Delta t$ as follows

$$m_{1,t+\Delta t} = \int_{-\infty}^{\infty}\widetilde{z}\omega_{t+\Delta t}(\widetilde{z})\mathrm{d}\widetilde{z}, \tag{255}$$

$$m_{i,t+\Delta t} = \int_{-\infty}^{\infty}(\widetilde{z}-m_{1,t+\Delta t})^i\omega_{t+\Delta t}(\widetilde{z})\mathrm{d}\widetilde{z} \;\text{ for } i = 2,...,n, \tag{256}$$

which can be numerically integrated using Gauss-Legendre quadratures.

Implementation Details. Equation (254) cannot be directly computed if we use a local perturbation approach because the function $\Psi_t(\widetilde{x})$ has a kink at $\widetilde{x} = \underline{\widetilde{z}}_t$. Substituting out $\Psi_t(\widetilde{x})$ using (248), we rewrite equation (254) as follows:

$$\begin{aligned}\omega_{t+\Delta t}(\widetilde{z}) =& \frac{1}{\sigma\sqrt{\theta\Delta t}}\frac{A_t}{A_{t+\Delta t}}\left[\int_{-\infty}^{\underline{\widetilde{z}}_t}\Psi_{l,t}\omega_t(\widetilde{x})\phi\left(\frac{\widetilde{z}-(1-\theta\Delta t)\widetilde{x}}{\sigma\sqrt{\theta\Delta t}}\right)\mathrm{d}\widetilde{x}\right.\\ &\left.+\int_{\underline{\widetilde{z}}_t}^{\infty}\Psi_{h,t}(\widetilde{x})\omega_t(\widetilde{x})\phi\left(\frac{\widetilde{z}-(1-\theta\Delta t)\widetilde{x}}{\sigma\sqrt{\theta\Delta t}}\right)\mathrm{d}\widetilde{x}\right],\end{aligned} \tag{257}$$

where

$$\Psi_{l,t} = [1+(r_{f,t}-\delta_a-\rho)\Delta t]+\sigma_a\Delta W_t, \tag{258}$$

$$\Psi_{h,t}(\widetilde{x}) = [1+(r_{f,t}-\delta_a-\rho)\Delta t]+\sigma_a\Delta W_t+(1+\lambda)\left[(\kappa_t\exp(\widetilde{x})-\delta_k-r_{f,t})\Delta t-\sigma_k\Delta W_t)\right]. \tag{259}$$

By doing a change of variables, equation (257) can be rewritten as

$$\omega_{t+\Delta t}(\widetilde{z}) = \frac{1}{\sigma\sqrt{\theta\Delta t}}\frac{A_t}{A_{t+\Delta t}}\left[\int_0^\infty \Psi_{l,t}\omega_t(\underline{\widetilde{z}}_t-\widetilde{x})\phi\left(\frac{\widetilde{z}-(1-\theta\Delta t)(\underline{\widetilde{z}}_t-\widetilde{x})}{\sigma\sqrt{\theta\Delta t}}\right)\mathrm{d}\widetilde{x}\right.$$
$$\left. + \int_0^\infty \Psi_{h,t}(\underline{\widetilde{z}}_t+\widetilde{x})\omega_t(\underline{\widetilde{z}}_t+\widetilde{x})\phi\left(\frac{\widetilde{z}-(1-\theta\Delta t)(\underline{\widetilde{z}}_t+\widetilde{x})}{\sigma\sqrt{\theta\Delta t}}\right)\mathrm{d}\widetilde{x}\right], \quad (260)$$

so that the integration regions are $[0,\infty)$. Further, we make another change of variables by defining $\widetilde{e} = \frac{(1-\theta\Delta t)\widetilde{x}}{\sigma\sqrt{2\theta\Delta t}}$. Equation (260) becomes

$$\omega_{t+\Delta t}(\widetilde{z}) = \frac{\sqrt{2}}{1-\theta\Delta t}\frac{A_t}{A_{t+\Delta t}}\left[\int_0^\infty \Psi_{l,t}\omega_t\left(\underline{\widetilde{z}}_t-\frac{\sigma\sqrt{2\theta\Delta t}\widetilde{e}}{1-\theta\Delta t}\right)\phi\left(\frac{\widetilde{z}-(1-\theta\Delta t)\underline{\widetilde{z}}_t}{\sigma\sqrt{\theta\Delta t}}+\sqrt{2}\widetilde{e}\right)\mathrm{d}\widetilde{e}\right.$$
$$\left. + \int_0^\infty \Psi_{h,t}\left(\underline{\widetilde{z}}_t+\frac{\sigma\sqrt{2\theta\Delta t}\widetilde{e}}{1-\theta\Delta t}\right)\omega_t\left(\underline{\widetilde{z}}_t+\frac{\sigma\sqrt{2\theta\Delta t}\widetilde{e}}{1-\theta\Delta t}\right)\phi\left(\frac{\widetilde{z}-(1-\theta\Delta t)\underline{\widetilde{z}}_t}{\sigma\sqrt{\theta\Delta t}}-\sqrt{2}\widetilde{e}\right)\mathrm{d}\widetilde{e}\right]. \quad (261)$$

Substituting the PDF of the standard normal distribution for $\phi(\cdot)$, equation (261) becomes

$$\omega_{t+\Delta t}(\widetilde{z}) = \frac{1}{\sqrt{\pi}(1-\theta\Delta t)}\exp\left(-\frac{1}{2}\left[\frac{\widetilde{z}-(1-\theta\Delta t)\underline{\widetilde{z}}_t}{\sigma\sqrt{\theta\Delta t}}\right]^2\right)\frac{A_t}{A_{t+\Delta t}}\left[\int_0^\infty f_{l,t}(\widetilde{e}|\widetilde{z})\exp(-\widetilde{e}^2)\mathrm{d}\widetilde{e}\right.$$
$$\left. + \int_0^\infty f_{h,t}(\widetilde{e}|\widetilde{z})\exp(-\widetilde{e}^2)\mathrm{d}\widetilde{e}\right], \quad (262)$$

where

$$f_{l,t}(\widetilde{e}|\widetilde{z}) = \Psi_{l,t}\omega_t\left(\underline{\widetilde{z}}_t-\frac{\sigma\sqrt{2\theta\Delta t}\widetilde{e}}{1-\theta\Delta t}\right)\exp\left(-\frac{\sqrt{2}(\widetilde{z}-(1-\theta\Delta t)\underline{\widetilde{z}}_t)}{\sigma\sqrt{\theta\Delta t}}\widetilde{e}\right), \quad (263)$$

$$f_{h,t}(\widetilde{e}|\widetilde{z}) = \Psi_{h,t}\left(\underline{\widetilde{z}}_t+\frac{\sigma\sqrt{2\theta\Delta t}\widetilde{e}}{1-\theta\Delta t}\right)\omega_t\left(\underline{\widetilde{z}}_t+\frac{\sigma\sqrt{2\theta\Delta t}\widetilde{e}}{1-\theta\Delta t}\right)\exp\left(\frac{\sqrt{2}(\widetilde{z}-(1-\theta\Delta t)\underline{\widetilde{z}}_t)}{\sigma\sqrt{\theta\Delta t}}\widetilde{e}\right). \quad (264)$$

Equation (262) can be computed using one-sided Gauss-Hermite quadrature (Steen, Byrne and Gelbard, 1969).[25] The results presented in Table OA.2 are computed based on this integration method.

[25]The term $\exp\left(\frac{\sqrt{2}(\widetilde{z}-(1-\theta\Delta t)\underline{\widetilde{z}}_t)}{\sigma\sqrt{\theta\Delta t}}\widetilde{e}\right)$ can introduce large numerical errors if it is too large. Thus, the choice Δt cannot be too small.

Alternatively, we can compute an approximation of equation (254) by changing the timing assumption of our model. Note that equation (254) is obtained based on the timing assumption that the shock to idiosyncratic productivity $z_{i,t+\Delta t}$ at $t + \Delta t$ occurs after capital accumulation over $[t, t + \Delta t)$ based on productivity $z_{i,t}$. When $\Delta t \approx 0$, this timing assumption yields similar results to an alternative timing assumption under which the value of idiosyncratic productivity $z_{i,t+\Delta t}$ at $t + \Delta t$ is realized at the beginning of $[t, t + \Delta t)$. Then capital accumulation over $[t, t + \Delta t)$ is based on $z_{i,t+\Delta t}$. In this case, equation (254) becomes

$$\omega_{t+\Delta t}(\widetilde{z}) = \frac{\Psi_t(\widetilde{z})}{\sigma\sqrt{\theta \Delta t}} \frac{A_t}{A_{t+\Delta t}} \int_{-\infty}^{\infty} \omega_t(\widetilde{x}) \phi \left(\frac{\widetilde{z} - (1 - \theta \Delta t)\widetilde{x}}{\sigma\sqrt{\theta \Delta t}} \right) \mathrm{d}\widetilde{x}, \tag{265}$$

which can be rewritten as (by making a change of variable $\widetilde{e} = \frac{\widetilde{z} - (1-\theta\Delta t)\widetilde{x}}{\sigma\sqrt{2\theta\Delta t}}$):

$$\omega_{t+\Delta t}(\widetilde{z}) = \frac{\Psi_t(\widetilde{z})}{\sqrt{\pi}(1 - \theta \Delta t)} \frac{A_t}{A_{t+\Delta t}} \int_{-\infty}^{\infty} \omega_t \left(\frac{\widetilde{z} - \sigma\sqrt{2\theta\Delta t}\widetilde{e}}{1 - \theta \Delta t} \right) \exp(-\widetilde{e}^2) \mathrm{d}\widetilde{e}. \tag{266}$$

Equation (266) can be easily implemented in dynare because the kinked function $\Psi_t(\widetilde{z})$ is not part of the integrand. Thus, the integral in equation (266) can be easily computed using Gauss-Hermite quadratures. The integral in equations (255) and (256) can then be computed as the sum of two integrals over $[-\infty, \widetilde{\underline{z}})$ and $[\widetilde{\underline{z}}, +\infty)$, respectively. The integral in each interval can be computed using Gauss-Legendre quadratures.