Asset Pricing with Misallocation

Winston Wei Dou  Yan Ji  Di Tian  Pengfei Wang*

May 4, 2021

Abstract

Misallocation reduces total factor productivity and economic growth, implying substantial adverse welfare effects. Misallocation measures should therefore provide a more informative empirical stochastic discount factor than aggregate consumption time series in small samples. We find evidence for misallocation-driven low-frequency movements in both aggregate growth and stock returns. We then develop an endogenous growth model with heterogeneous firms, intermediate goods, and financial frictions, in which misallocation emerges analytically as a crucial state variable. In equilibrium, misallocation endogenously generates long-run uncertainty about economic growth by distorting innovation and R&D decisions, leading to significant welfare losses and risk premia in capital markets. 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, Asset pricing.

JEL Classification: L11, O30, O40.

*Dou: University of Pennsylvania (wdou@wharton.upenn.edu). Ji: Hong Kong University of Science and Technology (jii@ust.hk). Tian: University of Pennsylvania (tiandi@sas.upenn.edu). Wang: Peking University (pfwang@phbs.pku.edu.cn). We are grateful for the helpful comments from Hengjie Ai, Jinhui Bai, Paco Buera, Lars Peter Hansen, Vincent Glode, Joao Gomes, Felix Kubler, Vincenzo Quadrini, Adriano Rampini, Michael Roberts, Nick Roussanov, Tom Sargent, Rob Stambaugh, Rob Townsend, Neng Wang, Tom Winberry, and Wei Xiong, as well as the participants in the Macroeconomy and Finance in China Conference, the second and third PKU/PHBS Sargent Institute Workshops, HKUST-JINAN Workshop in Macroeconomics, Singapore Management University, Shanghai University of Finance and Economics, Northern Finance Association Conference, and Wharton. This paper was previously circulated under the title “Granular Uncertainty and Disastrous Misallocation.”
1 Introduction

In the last decade, one of the most important 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 long-run consumption risk (Bansal and Yaron, 2004), a mechanism that quantitatively justifies many asset pricing moments.

This paper studies the connection between misallocation and long-run consumption risk by developing a novel analytically tractable framework, offering insights on the asset pricing implications of misallocation. In our model, 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 which face two financial frictions because of agency conflicts — an equity market constraint for payout and issuance and a collateral constraint for debt. Specifically, the equity market constraint for payout and issuance arises as in Myers (2000) and Lambrecht and Myers (2008, 2012), and the collateral constraint for debt arises as in Buera and Shin (2013) and Moll (2014), among many others. 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 i.i.d.) aggregate shocks can generate persistent shifts in demand for research and development (R&D) through their long-lasting effects on misallocation, which in turn leads to persistent fluctuations in economic growth; as a consequence, our model endogenizes a low-frequency component of economic growth driven by slow-moving misallocation, which has first-order asset pricing implications in capital markets.

Our paper contributes to existing literature in three ways. First, we show that misallocation drives low-frequency movements in R&D intensity and thus economic growth in both the model and data. In our model, 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. 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 long-run consumption risk (Bansal and Yaron, 2004; Hansen, Heaton and Li, 2008). In the data, we find evidence for misallocation-driven low-frequency movements in both aggregate growth 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 (62% and 1.90).
Importantly, future accumulated consumption growth has little explanatory power for
portfolio returns once our misallocation measure is included as a factor.

Third, our model delineates the tight link between firms’ idiosyncratic productivity
shocks and the low-frequency aggregate consumption risk. When firms’ idiosyncratic pro-
ductivity 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 aggregate 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 granular firm-level data, which
helps address the issues of weak identification in the long-run risk literature (Chen, Dou
and Kogan, 2019; Cheng, Dou and Liao, 2020). This highlights the important role of
misallocation measures in estimating the empirical stochastic discount factor (SDF).

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 exist 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. Produc-
tion 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 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 Tikhomirov, 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 an endogenous state variable, capturing the covariance between log capital and log productivity across firms. This state variable determines the economy’s misallocation, based on which both the steady state and transitional dynamics are 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 a worse misallocation predicts declines in R&D intensity and lower growth of aggregate consumption and output. In the cross section, 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.

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). 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, 2019; Cheng, Dou and Liao, 2020).


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.

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, 2008b) suggest that misallocation can play an important role in determining the

1See Eisfeldt and Shi (2018) for a comprehensive survey.
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 $i \in I \equiv [0,1]$ 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 $[t, t+dt)$ using a production technology with constant returns to scale:

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

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

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 $u_{i,t}k_{i,t}\Delta t$ depreciation over $[t, t + dt)$, where $\Delta t$ captures the stochastic depreciation rate,

$$d\Delta t = \delta_k dt + \sigma_k dW_t.$$  \hfill (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.

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

$$da_{i,t} = -\delta_a a_{i,t}dt + \sigma_a a_{i,t}dW_t + dI_{i,t},$$  \hfill (3)

where $\delta_a > 0$ is the constant depreciation rate, and $\sigma_a dW_t$ captures the capital efficiency shock with $\sigma_a > 0$. The modeling of capital efficiency shocks has been widely adopted in the literature.\(^3\) 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). The variable $dI_{i,t}$ is the amount of final goods that is converted to capital over $[t, t + dt)$. Similar to Pástor and Veronesi (2012), we assume that profits are reinvested, so that the firm’s investment rate $dI_{i,t}$ is equal to its profit after paying operation expenses, interests, and dividends.

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

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

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 the variety of intermediate inputs, $N_t$, drives endogenous growth, as in Romer (1987, 1990) and Jones

\(^2\)The capital leasing market is relevant for firms’ production and financial decisions (e.g., Eisfeldt and Rampini, 2008a; Rampini and Viswanathan, 2013; Li and Tsou, 2021; Li and Xu, 2021).

(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

\[
d \ln(z_{i,t}) = -\theta \ln(z_{i,t}) dt + \sigma \sqrt{\theta} dW_{i,t},
\]

(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 \( \theta \) determines the persistence of idiosyncratic productivity \( z_{i,t} \). A higher \( \theta \) makes \( z_{i,t} \) less persistent, implying that firms face higher uncertainty in their future idiosyncratic productivity. Importantly, a change in \( \theta \) does not affect the dispersion in idiosyncratic productivity across firms, because \( \theta \) 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

\[
\max_{p_{j,t}} \pi_{j,t} = p_{j,t} e_{j,t} - e_{j,t},
\]

(6)

subject to the demand curve:

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

(7)

where \( X_t \equiv \int_{i \in \mathcal{J}} x_{i,t} d\tilde{i} \) 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:

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

(8)

where \( \Lambda_t \) is the SDF, and \( \delta_b \) 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:

$$dN_t = \vartheta_t S_t dt - \delta_b N_t dt,$$

where $S_t$ is the aggregate R&D expenditure, and $\vartheta_t$ 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 final-good 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 final-good firm 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 (1992a,b):

$$ U_0 = \mathbb{E}_0 \left[ \int_0^\infty f(C_t, U_t)dt \right], \quad (12) $$

where

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

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 $\delta$ is the subjective discount rate, the parameter $\psi$ is the elasticity of intertemporal substitution (EIS), and the parameter $\gamma$ captures risk aversion.

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

$$ dB_t = (w_t L_t + r_{ft} B_t + D_t - C_t) dt, \quad (14) $$

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

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

where $H_t$ 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 $\tau_{a_{i,t}} dt$ over $[t, t + dt)$ when running the firm $i$.\(^4\)

\(^4\)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, 2008b).
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). In particular, we assume that a fraction $\tau/\rho$ 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 $[t, t + dt]$.

Moreover, the installed capital for production is $k_{i,t} = a_{i,t} + \hat{a}_{i,t}$, and the manager can divert a fraction $1/\lambda$ of leased capital $\hat{a}_{i,t}$ with $\lambda \geq 0$. 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 $\hat{a}_{i,t} \leq \lambda a_{i,t}$. 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 payout/issuance policy is subject to the following equity market constraint:

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

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:

$$- a_{i,t} \leq \hat{a}_{i,t} \leq \lambda a_{i,t}.$$  

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 $[t, t + dt]$. Specifically, the manager of firm $i$ makes leasing ($\tilde{a}_{i,s}$) and production ($u_{i,t}, \ell_{i,t}, x_{i,j,s}$) decisions for all $s \geq t$ to maximize the present value $J_{i,t}$ of his own rents

$$J_{i,t} = \max_{\tilde{a}_{i,s}, u_{i,s}, \ell_{i,s}, x_{i,j,s}} \mathbb{E}_t \left[ \int_t^\infty \frac{\Lambda_s}{\Lambda_t} \tau a_{i,s} ds \right], \quad (18)$$

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

$$dI_{i,t} = y_{i,t} dt - \int_0^N p_{j,i} x_{i,j,t} dj dt - w_t \ell_{i,t} dt - u_{i,t} k_{i,t} d\Delta_t - r_{f,t} \tilde{a}_{i,t} dt - d_{i,t} dt, \quad (19)$$

where the SDF $\Lambda_t$ evolves according to

$$\frac{d\Lambda_t}{\Lambda_t} = -r_{f,t} dt - \eta_t dW_t. \quad (20)$$

The variable $\eta_t$ 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:

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

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:

$$\frac{d\xi_{i,t}}{\xi_{i,t}} = \mu_{\xi,t}(z_{i,t}) dt + \sigma_{\xi,t}(z_{i,t}) dW_t + \sigma_{\xi,t}(z_{i,t}) dW_{i,t}, \quad (22)$$

where $\mu_{\xi,t}(z_{i,t})$, $\sigma_{\xi,t}(z_{i,t})$, and $\sigma_{\xi,t}(z_{i,t})$ 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 $z_t$
for being active:

\[
u_t(z) = \begin{cases} 1, & z \geq z_t \\ 0, & z < z_t \end{cases}, \quad k_t(a, z) = \begin{cases} (1 + \lambda)a, & z \geq z_t \\ 0, & z < z_t \end{cases}
\]  

(23)

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

(24)

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

(25)

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

\[
p_t = \left( \int_0^{N_t} \bar{p}_t^{\frac{\nu}{\nu - 1}} d\bar{j} \right)^{\frac{\nu - 1}{\nu}},
\]

(26)

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

(27)

The productivity cutoff \( z_t \) is determined by:

\[
\bar{z}_t \kappa_t = r_{f, t} + \delta_k + \sigma_k (\sigma_{\xi, t}(\bar{z}_t) - \eta_t).
\]

(28)

where \( \kappa_t \) is

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

(29)

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

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

\[
\frac{dI_{i, t}}{a_{i, t}} = (1 + \lambda) \left( \kappa_t z_{i, t} dt - d\Lambda_t - r_{f, t} dt \right) \mathbb{1}_{z_{i, t} \geq z_t} + (r_{f, t} - \rho) dt.
\]

(30)

As in Moll (2014), the drift in capital 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

\[ D_t = \rho A_t + \int_{j=0}^{N_t} \pi_{j,t} dj - S_t, \]  

(31)

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

\[ A_t = \int_0^\infty \int_0^\infty a \phi_t(a,z) da dz. \]  

(32)

In equation (31), the first term $\rho A_t$ captures the dividend of the final goods sector. The second term $\int_{j=0}^{N_t} \pi_{j,t} dj$ 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

\[ K_t = \int_0^\infty \int_0^\infty k_t(a,z) \phi_t(a,z) da dz. \]  

(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 $\hat{a}_{i,t}$, $u_{i,t}$, $\ell_{i,t}$, and $x_{i,t}$, 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 expenses $S_t$ are determined by equation (10);

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

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

(34)

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

\[ L_t = \int_{z_t}^{\infty} \int_0^{\infty} \ell_t(a,z) \phi_t(a,z) da dz; \]  

(35)
the leased capital market-clearing condition determines households’ bond holdings $B_t$:

$$B_t = \int_0^\infty \int_0^\infty \tilde{a}_t(a,z)\varphi_t(a,z)da\,dz. \quad (36)$$

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

$$K_t = A_t + B_t. \quad (37)$$

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

Because managers’ problem is linear in capital $a_{i,t}$ (see equation (64)), it is not necessary to track the marginal distribution of capital conditional on each productivity type $z$. 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:

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

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

$$\Omega_t(z) \equiv \int_0^z \omega_t(z')dz'. \quad (39)$$

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., $\frac{(1-\nu)e}{\nu(1-\epsilon)} + \alpha = 1$ 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

$$Y_t = (e\nu)^{1-\alpha} Z_t N_t^{1-\alpha} K_t^\alpha L_t^{1-\alpha}, \quad (40)$$

where $Z_t$ is the productivity of the final goods sector, given by

$$Z_t = \left[ \frac{\int_{z_t}^\infty z\omega_t(z)dz}{1 - \Omega_t(z_t)} \right]^\alpha. \quad (41)$$

---

5In 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.
The equilibrium $K_t / A_t$ ratio is determined by the productivity cutoff $z_t$ in equation (28):

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

Factor prices are

$$p_{j,t} = 1 / v \quad \text{and} \quad p_t = N_t^{\nu-1} / v, \tag{43}$$
$$w_t = (1 - \alpha)(1 - \epsilon)Y_t / L_t, \tag{44}$$

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

$$N_t \pi_t = (1 - v)\epsilon Y_t, \tag{45}$$
$$S_t = (\chi v_t)^{1/h} N_t. \tag{46}$$

Equation (40) shows that the economy’s aggregate TFP is $(\epsilon v)^{1/\nu} Z_t N_t^{1-\alpha}$, which depends on the knowledge stock $N_t$ and the productivity $Z_t$ of the final goods sector. The productivity $Z_t$ 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), $Z_t$ is firms’ average productivity $z$ weighted by their capital share $\omega_t(z)$. Similar to Moll (2014), the equilibrium productivity cutoff $z_t$ is determined directly by the CDF of capital share (see equation (42)) due to the bang-bang solution in equation (23). The value of $Z_t$ is higher when more productive firms are associated with more capital, which reflects a more efficient capital allocation in the presence of collateral constraints.

Equation (43) is a direct consequence of homogeneous intermediate goods producers facing the a constant elasticity of substitution, $1 / (1 - v)$. Equation (44) implies that the equilibrium wage is competitive, given by the labor share, $(1 - \alpha)(1 - \epsilon)$, 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 the production function, $\epsilon Y_t$, multiplied by their profitability, as captured by the inverse of the elasticity of substitution $(1 - v)$ among differentiated intermediate goods. In equation (46), innovators’ R&D intensity increases with the value of blueprints $v_t$ with an elasticity of $1/h$. 

15
2.6 Misallocation as a State Variable

The capital share \( \omega_t(z) \) is crucial in determining the final goods sector’s productivity \( Z_t \) 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 B.1, we apply the Berry-Esseen bound (Tikhomirov, 1980; Bentkus, Gotze and Tikhomirov, 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, \( \tilde{a}_{i,t} = \ln(a_{i,t}) \), across firms in the final goods sector approximately follows a normal distribution.

According to equation (5), log individual productivity \( \tilde{z}_{i,t} = \ln z_{i,t} \) also follows a normal distribution, \( \tilde{z}_{i,t} \sim N(0, \sigma^2/2) \), in the stationary equilibrium. Thus, if at some initial point in time \( t_0 \), \( \tilde{z}_{i,t_0} \) and \( \tilde{a}_{i,t_0} \) follow a joint normal distribution, then the distribution of \( \tilde{z}_{i,t} \) and \( \tilde{a}_{i,t} \) 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 \( M_t \equiv -\text{Cov}(\tilde{z}_{i,t}, \tilde{a}_{i,t}) / \text{var}(\tilde{z}_{i,t}) = -2\text{Cov}(\tilde{z}_{i,t}, \tilde{a}_{i,t}) / \sigma^2 \).

Lemma 3 implies that under our approximation, the endogenous variable \( M_t \equiv -\text{Cov}(\tilde{z}_{i,t}, \tilde{a}_{i,t}) / \text{var}(\tilde{z}_{i,t}) \) is a sufficient statistic that characterizes \( \omega_t(z) \). Intuitively, \( M_t \) captures the covariance between \( \tilde{z}_{i,t} \) and \( \tilde{a}_{i,t} \) 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 consistent with
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.2 that our approximation can yield solutions similar to the numerical solutions based on directly tracking the evolution of \( \omega_t(z) \) in both steady states and transitions when the persistence of idiosyncratic productivity, i.e., \( \theta \), is within the range of the empirical estimates of Asker, Collard-Wexler and Loecker (2014) based on U.S. census data. More generally, our log-normal approximation is justified by the detailed numerical analysis of Winberry (2018). In a more general model with aggregate shocks, Winberry (2018) approximates the distribution with a flexible parametric family, which nests our log-normal approximation as a special case. Winberry (2018) solves the model using an efficient perturbation method and shows that for standard calibrations, approximations based on the log-normal approximation can already yield stationary distributions and aggregates during transitions that are virtually indistinguishable from higher degree approximations (see Figures 2 and 3 of Winberry, 2018). Importantly, in the stochastic equilibrium, 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 degree approximations (see Tables 3 and 4 of Winberry, 2018). Given that these variables crucially determine the quantitative asset pricing implications of our model, Winberry (2018)’s findings lend strong support for our analytically tractable approximation.

**Proposition 3.** Under our approximation specified in Lemma 2, the productivity \( Z_t \) of the final goods sector is

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

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

Equation (48) clearly shows that the final goods sector’s productivity \( Z_t \) strictly decreases with the misallocation variable \( M_t \). Thus, a lower \( M_t \) leads to higher \( Z_t \) and aggregate output \( Y_t \).\(^6\) Moreover, a lower misallocation \( M_t \) implies that more produc-

\(^6\)One of the main insights of Hsieh and Klenow (2009) is that the misallocation of resources lowers
tive firms have more capital, leading to a higher productivity cutoff \( z_t \). Under our approximation, equation (41) has a closed-form representation:

\[
    z_t = \exp \left[ -\frac{\sigma^2}{2} M_t - \Phi^{-1} \left( \frac{1}{1 + \lambda A_t} \right) \frac{\sigma}{\sqrt{2}} \right].
\]  

(49)

Thus, a lower misallocation \( M_t \) leads to a higher productivity \( Z_t \) but fewer firms in the final goods sector will be active.\(^7\)

### 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 (42) and (37), 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

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

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

\[
    dM_t = -\theta M_t dt - \frac{\text{Cov}(\tilde{z}_{i,t}, \tilde{a}_{i,t})}{\text{var}(\tilde{z}_{i,t})}, \tag{52}
\]

where is \( \text{Cov}(\tilde{z}_{i,t}, \tilde{a}_{i,t}) \) given by equation (134) in Online Appendix.

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, \( Y_t / A_t \), and the aggregate TFP. Proposition 3 shows that, in our model, the endogenous variable \( M_t \) determines the degree of misallocation as it determines productivity \( Z_t \), 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 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 C, 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 formula (41).

\(^7\)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).
minus capital depreciation, $(\delta_k dt + \sigma_k dW_i) K_t / A_t + \delta_a dt - \sigma_a dW_t$, minus interests on households’ loans, $(K_t / A_t - 1) r_{f,t} dt$, and dividend payout, $\rho dt$.

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 intensity $S_t$ (equation (46)). Importantly, the misallocation $M_t$ determines the economy’s endogenous growth rate over $[t, t + dt)$. This is because $v_t$ equals the present value of profit flow $\pi_t$ (equation (8)), and thus $v_t$ is higher when $\pi_t$ is higher. A lower misallocation $M_t$ increases the final goods sector’s productivity $Z_t$ (equation (48)), leading to a higher aggregate output $Y_t$ (equation (40)) and thus a higher profit flow $\pi_t$ (equation (45)), and ultimately, a higher growth rate of the economy. By linking the final goods sector and the innovation sector through the endogenous productivity $Z_t$, 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 $-\theta M_t dt$ reflects time-varying productivity $z_{i,t}$ evolving according to equation (5). Intuitively, a higher $\theta$ implies a less persistent idiosyncratic productivity $z_{i,t}$, which pushes the misallocation $M_t = -\text{Cov}(\tilde{z}_{i,t}, \tilde{a}_{i,t}) / \text{var}(\tilde{z}_{i,t})$ towards zero. In Section 3, we show that the parameter $\theta$ crucially determines the economy’s long-run consumption risk by affecting the persistence of $M_t$. The second term $\text{Cov}(\tilde{z}_{i,t}, d\tilde{a}_{i,t}) / \text{var}(\tilde{z}_{i,t})$ captures the impact of capital accumulation, $d\tilde{a}_{i,t}$, evolving according to equation (3). Intuitively, a higher $\text{Cov}(\tilde{z}_{i,t}, d\tilde{a}_{i,t})$ implies that more productive firms also accumulate their capital at a higher rate, which reduces misallocation $M_t$.

Importantly, the variable $\text{Cov}(\tilde{z}_{i,t}, d\tilde{a}_{i,t})$ negatively depends on the aggregate shock $dW_t$ (see equation (134) in Online Appendix). Intuitively, a positive shock ($dW_t > 0$) increases the depreciation rate of capital $k_{i,t}$, which reduces the capital accumulation of firms with productivity $z_{i,t}$ above the cutoff $z_t$, without affecting those with productivity below the cutoff due to the optimal solution (23). As a result, a positive shock leads to a lower $\text{Cov}(\tilde{z}_{i,t}, d\tilde{a}_{i,t})$, 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, positive value, indicating that a positive shock reduces $A_t$.

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., \(dW_t \equiv 0\)).

**Proposition 5.** There is a balanced growth path in which \(E_t, M_t, \text{ and } Z_t\) are constant, and aggregate capital \(A_t\), knowledge stock \(N_t\), output \(Y_t\), and consumption \(C_t\) grow at a 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 \((dW_t > 0)\) increases \(M_t\), leading to a lower productivity \(Z_t\) 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 spending, 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.

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 quantitative results. Our model can generate moments on consumption growth and asset prices consistent with the data. In the model, misallocation significantly predicts future consumption growth and R&D intensity. 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 studies the welfare implications of misallocation.

3.1 Calibration

Panel A of Table 1 presents the externally calibrated parameters. Following the standard practice, we set the capital share in the production technology at \(\alpha = 0.33\) and the yearly capital depreciation rate 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
Table 1: Parameter calibration and targeted moments.

### Panel A: Externally determined parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capital share</td>
<td>$\alpha$</td>
<td>0.33</td>
<td>Capital depreciation rate</td>
<td>$\delta_k, \delta_a$</td>
<td>0.04</td>
</tr>
<tr>
<td>Share of intermediate inputs</td>
<td>$\varepsilon$</td>
<td>0.5</td>
<td>Risk aversion</td>
<td>$\gamma$</td>
<td>10</td>
</tr>
<tr>
<td>EIS</td>
<td>$\psi$</td>
<td>1.85</td>
<td>Inverse markup</td>
<td>$\nu$</td>
<td>0.6</td>
</tr>
<tr>
<td>Dividend payout rate</td>
<td>$\rho$</td>
<td>0.025</td>
<td>Knowledge depreciation rate</td>
<td>$\delta_b$</td>
<td>0.25</td>
</tr>
<tr>
<td>1 – R&amp;D elasticity</td>
<td>$h$</td>
<td>0.17</td>
<td>Vol. of idio. productivity</td>
<td>$\sigma$</td>
<td>1.39</td>
</tr>
<tr>
<td>Collateral constraint</td>
<td>$\lambda$</td>
<td>1</td>
<td>Persistence of idio. productivity</td>
<td>$\theta$</td>
<td>0.1625</td>
</tr>
</tbody>
</table>

### Panel B: Internally calibrated parameters and targeted moments

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
<th>Moments</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjective discount rate</td>
<td>$\delta$</td>
<td>0.015</td>
<td>Real risk-free rate (%)</td>
<td>1.21</td>
<td>1.22</td>
</tr>
<tr>
<td>R&amp;D productivity</td>
<td>$\chi$</td>
<td>3.35</td>
<td>Consumption growth rate (%)</td>
<td>1.76</td>
<td>1.77</td>
</tr>
<tr>
<td>Capital depreciation shock</td>
<td>$\sigma_k$</td>
<td>0.13</td>
<td>Consumption growth vol. (%)</td>
<td>1.48</td>
<td>1.49</td>
</tr>
<tr>
<td>Liquidity shock</td>
<td>$\sigma_a$</td>
<td>0.12</td>
<td>$\sigma(\Delta \ln C_t) / \sigma(\Delta \ln Y_t)$</td>
<td>0.61</td>
<td>0.56</td>
</tr>
</tbody>
</table>

Commonly used in recent works in asset pricing, we set the risk aversion at $\gamma = 10$ and the EIS at $\psi = 1.85$ as in Kung and Schmid (2015). The dividend payout rate is set at $\rho = 0.025$. 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 $\delta_b = 0.25$, 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 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 $\delta = 0.015$ to generate a real risk-free rate of about 1.22%. The R&D productivity is a
Table 2: Untargeted moments in data and model.

<table>
<thead>
<tr>
<th>Moments</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>median</td>
<td>2.5%</td>
</tr>
<tr>
<td>Panel A: Consumption moments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$AC_1(\Delta \ln C_t)$ (%)</td>
<td>0.41</td>
<td>0.19</td>
</tr>
<tr>
<td>$AC_2(\Delta \ln C_t)$ (%)</td>
<td>0.02</td>
<td>-0.12</td>
</tr>
<tr>
<td>$AC_4(\Delta \ln C_t)$ (%)</td>
<td>-0.07</td>
<td>-0.22</td>
</tr>
<tr>
<td>$AC_6(\Delta \ln C_t)$ (%)</td>
<td>-0.03</td>
<td>-0.27</td>
</tr>
<tr>
<td>$VR_2(\Delta \ln C_t)$ (%)</td>
<td>1.77</td>
<td>1.18</td>
</tr>
<tr>
<td>$VR_4(\Delta \ln C_t)$ (%)</td>
<td>1.93</td>
<td>1.18</td>
</tr>
<tr>
<td>$VR_6(\Delta \ln C_t)$ (%)</td>
<td>2.21</td>
<td>1.08</td>
</tr>
<tr>
<td>Panel B: Other macroeconomic moments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma(\Delta \ln TFP_t)/\sigma(\Delta \ln Y_t)$ (%)</td>
<td>1.22</td>
<td>0.93</td>
</tr>
<tr>
<td>$\sigma(\Delta \ln S_t)/\sigma(\Delta \ln Y_t)$ (%)</td>
<td>2.10</td>
<td>1.40</td>
</tr>
<tr>
<td>$AC_1(\Delta \ln TFP_t)$ (%)</td>
<td>0.09</td>
<td>0.06</td>
</tr>
<tr>
<td>$AC_1(\Delta \ln S_t)$ (%)</td>
<td>0.21</td>
<td>-0.04</td>
</tr>
<tr>
<td>$AC_1(\Delta \ln Y_t)$ (%)</td>
<td>0.37</td>
<td>0.09</td>
</tr>
<tr>
<td>$AC_1(M_t)$ (%)</td>
<td>0.65</td>
<td>0.68</td>
</tr>
<tr>
<td>Panel C: Asset pricing moments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sharpe ratio of consumption claim</td>
<td>0.40</td>
<td>0.18</td>
</tr>
<tr>
<td>$\sigma_{r_f}$ (%)</td>
<td>0.60</td>
<td>0.36</td>
</tr>
</tbody>
</table>

Note: The variable $\Delta \ln X_t = \ln X_{t+1} - \ln X_t$ represents difference in $\ln X_t$ between year $t$ and year $t - 1$. We use yearly consumption data for the postwar period from 1948 to 2017. Moments in panel A are computed following Beeler and Campbell (2012), who focus on the sample period from 1948 to 2008 (our moments replicate theirs when we focus on the same sample period). $AC_k(\Delta \ln C_t)$ refers to the autocorrelation of consumption growth with a $k$-year lag. $VR_k(\Delta \ln C_t)$ of consumption growth refers to the variance ratio of consumption growth with a $k$-year horizon. Moments in panel B are obtained from Kung and Schmid (2015). 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, 2.5%, and 97.5% of each moment’s distribution across 10,000 independent simulations.

scaling parameter and is set at $\chi = 3.35$ to generate an average consumption growth rate of about 1.77%. We calibrate $\sigma_k = 0.13$ so that the model-implied volatility of consumption growth is about 1.49%. We calibrate $\sigma_a = 0.12$ so the ratio of the volatility of consumption growth and output growth is 0.56.

### 3.2 Quantitative Results

Table 2 presents untargeted moments as a validation test of the model. Panel A shows that the persistence of consumption growth implied by the model is very consistent with
Table 3: Model-implied relationship between misallocation, R&D, and growth.

Panel A: R&D intensity

<table>
<thead>
<tr>
<th></th>
<th>$t$ (current year)</th>
<th>$t + 1$ (next year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>$-0.099$</td>
<td>$-0.078$</td>
</tr>
<tr>
<td></td>
<td>$[-0.007]$</td>
<td>$[-0.010]$</td>
</tr>
<tr>
<td>$R$-squared</td>
<td>$0.745$</td>
<td>$0.464$</td>
</tr>
</tbody>
</table>

Panel B: Consumption growth

<table>
<thead>
<tr>
<th></th>
<th>$t \rightarrow t + 1$</th>
<th>$t \rightarrow t + 2$</th>
<th>$t \rightarrow t + 3$</th>
<th>$t \rightarrow t + 4$</th>
<th>$t \rightarrow t + 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>$-0.061$</td>
<td>$-0.101$</td>
<td>$-0.126$</td>
<td>$-0.140$</td>
<td>$-0.147$</td>
</tr>
<tr>
<td></td>
<td>$[-0.012]$</td>
<td>$[-0.021]$</td>
<td>$[-0.028]$</td>
<td>$[-0.034]$</td>
<td>$[-0.040]$</td>
</tr>
<tr>
<td>$R$-squared</td>
<td>$0.247$</td>
<td>$0.243$</td>
<td>$0.217$</td>
<td>$0.189$</td>
<td>$0.160$</td>
</tr>
</tbody>
</table>

Panel C: Output growth

<table>
<thead>
<tr>
<th></th>
<th>$t \rightarrow t + 1$</th>
<th>$t \rightarrow t + 2$</th>
<th>$t \rightarrow t + 3$</th>
<th>$t \rightarrow t + 4$</th>
<th>$t \rightarrow t + 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>$-0.038$</td>
<td>$-0.035$</td>
<td>$-0.027$</td>
<td>$-0.016$</td>
<td>$-0.003$</td>
</tr>
<tr>
<td></td>
<td>$[-0.026]$</td>
<td>$[-0.042]$</td>
<td>$[-0.054]$</td>
<td>$[-0.063]$</td>
<td>$[-0.071]$</td>
</tr>
<tr>
<td>$R$-squared</td>
<td>$0.031$</td>
<td>$0.016$</td>
<td>$0.014$</td>
<td>$0.017$</td>
<td>$0.018$</td>
</tr>
</tbody>
</table>

Note: We simulate a sample for 160 years with an 80-year burn-in period. In panel A, we regress the R&D intensity 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.

that in the data even though these moments are not directly targeted in our calibration. In Section 3.3, we show that the parameter $\theta$ governing the persistence of idiosyncratic productivity $z_{i,t}$ plays a major role in determining the persistence of consumption growth.

Panel B of Table 2 shows that several key macroeconomic moments implied by the model are also roughly aligned with the data. These moments include the volatility of TFP growth ($\sigma(\Delta \ln TFP_t)$) and R&D expenditures growth ($\sigma(\Delta \ln S_t)$) relative to the volatility of output growth ($\sigma(\Delta \ln Y_t)$). The yearly autocorrelations in TFP growth ($AC1(\Delta \ln TFP_t)$), R&D expenditures growth ($AC1(\Delta \ln S_t)$), output growth ($AC1(\Delta \ln Y_t)$), and our misallocation measure ($AC1(M_t)$).

Panel C 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 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 the misallocation $M_t$ in year $t$ is negatively correlated with
contemporaneous R&D intensity (i.e., $S_t/A_t$). The misallocation $M_t$ also negatively predicts the R&D intensity in the next year, $t + 1$.

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 conditional misallocation growth. As a result, both the $t$-statistic and the $R^2$ are larger for the regression coefficients over shorter horizons.

Panel C of Table 3 shows that misallocation $M_t$ also negatively predicts future output growth. However, the regression coefficients are not statistically significant as output growth implied by the model is less persistent compared to consumption growth (see Table 2).

### 3.3 Inspection of Model Mechanisms

We now illustrate how the persistence of idiosyncratic productivity and other key parameters determine the persistence of aggregate consumption growth and the quantitative implications of the model.

**Impulse Response Functions.** Consider a stationary economy in the balanced growth path. At $t = 0$, there is a one-time unexpected shock (i.e., $dW_t < 0$ over $[0, dt]$) that reduces the misallocation $M_t$ by $\sigma_k$. The blue solid line in each panel of Figure 1 plots the transitional dynamics in our baseline calibration.

The blue solid line in panel A illustrates the evolution of the misallocation $M_t$ after the shock, which follows equation (52). In the absence of aggregate shocks, aggregate consumption would be $C_0 \exp(gt)$, growing at a constant rate $g = 1.77\%$ 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 1.025 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, defined by $dC_t/(C_t dt)$. The blue solid line shows that the conditional consumption growth increases dramatically to about 3% when the shock
Note: Consider an unexpected shock that reduces misallocation $M_t$ by $\sigma_k$ at $t = 0$. Panels A, B, and C plot the transitional dynamics of misallocation $M_t$, excess consumption $C_t/(C_0 e^{\theta t})$, and consumption growth $dC_t/(C_0 dt)$ for three economies with different $\theta$. For each economy, we calibrate the parameter $\chi$ so that the consumption growth rate in the balanced growth path is the same as our baseline 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 $Z_t$, R&D-capital ratio $S_t/A_t$, and knowledge-capital ratio $N_t/A_t$ for the baseline case with $\exp(-\theta) = 0.85$.

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

hits at $t = 0$. This is because the reduction in misallocation $M_t$ immediately increases the productivity $Z_t$ of the final goods sector (panel D). A higher $Z_t$ increases the profits of innovators, motivating them to spend more on R&D (panel E), which consequently leads to a 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-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 $t = 0$.

In panels A, B and C, we further compare our baseline calibration with two economies of higher 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 below the black dashed line, which is below the blue solid 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 equilibrium.

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 2.72, 3.74, and 6.20 years for the blue solid ($\exp(-\theta) = 0.85$), black dashed ($\exp(-\theta) = 0.9$), and red dash-dotted lines ($\exp(-\theta) = 0.95$), respectively, indicating that $M_t$ is more persistent when $\theta$ 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 $M_t$.

The 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, $dC_t/(C_t\,dt)$. The connection between the persistence of these two variables is delivered through the endogenously persistent misallocation $M_t$. 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 in idiosyncratic productivity with the persistence in aggregate consumption growth, our model provides a way to estimate long-run risk in 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, 2019; Cheng, Dou and Liao, 2020).

**Inspection of Key Parameters.** Table 4 shows how the main variables of our model respond to changes in key parameters.

Column (1) presents the baseline calibration. In column (2), we consider a less persistent idiosyncratic productivity by increasing $\theta$ from 0.1625 to 0.22, which corresponds to a yearly autocorrelation of 0.80 in $\ln z_{i,t}$. Compared with the baseline calibration, the average misallocation $M_t$ increases from $-0.51$ to $-0.42$ because productive firms
Table 4: Inspection of key parameters.

<table>
<thead>
<tr>
<th></th>
<th>(1) Baseline</th>
<th>(2) $\theta = 0.22$</th>
<th>(3) $\lambda = 1.2$</th>
<th>(4) $\chi = 3.45$</th>
<th>(5) $\sigma_k = 0.12$</th>
<th>(6) $\sigma_a = 0.13$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{E}[M_t]$ (%)</td>
<td>-0.51</td>
<td>-0.42</td>
<td>-0.52</td>
<td>-0.52</td>
<td>-0.50</td>
<td>-0.50</td>
</tr>
<tr>
<td>$\mathbb{E}[Z_t]$ (%)</td>
<td>1.63</td>
<td>1.59</td>
<td>1.70</td>
<td>1.64</td>
<td>1.63</td>
<td>1.63</td>
</tr>
<tr>
<td>$\mathbb{E}[\Delta \ln C_t]$ (%)</td>
<td>1.77</td>
<td>0.96</td>
<td>2.46</td>
<td>2.48</td>
<td>1.64</td>
<td>1.66</td>
</tr>
<tr>
<td>$\sigma(\Delta \ln C_t)$ (%)</td>
<td>1.49</td>
<td>1.43</td>
<td>1.57</td>
<td>1.50</td>
<td>1.22</td>
<td>1.26</td>
</tr>
<tr>
<td>$\sigma(\Delta \ln C_t)/\sigma(\Delta \ln Y_t)$</td>
<td>0.56</td>
<td>0.54</td>
<td>0.57</td>
<td>0.56</td>
<td>0.41</td>
<td>0.39</td>
</tr>
<tr>
<td>$AC1(\Delta \ln C_t)$ (%)</td>
<td>0.41</td>
<td>0.35</td>
<td>0.40</td>
<td>0.40</td>
<td>0.43</td>
<td>0.43</td>
</tr>
<tr>
<td>$AC1(\Delta \ln TFP_t)$ (%)</td>
<td>0.27</td>
<td>0.22</td>
<td>0.27</td>
<td>0.27</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>$AC1(\Delta \ln Y_t)$ (%)</td>
<td>0.31</td>
<td>0.26</td>
<td>0.31</td>
<td>0.32</td>
<td>0.29</td>
<td>0.28</td>
</tr>
<tr>
<td>$AC1(M_t)$ (%)</td>
<td>0.82</td>
<td>0.78</td>
<td>0.82</td>
<td>0.82</td>
<td>0.83</td>
<td>0.83</td>
</tr>
<tr>
<td>Sharpe ratio (%)</td>
<td>0.41</td>
<td>0.30</td>
<td>0.44</td>
<td>0.41</td>
<td>0.32</td>
<td>0.33</td>
</tr>
</tbody>
</table>

Note: 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.

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 $Z_t$ decreases from 1.63 to 1.59. The average consumption growth rate decreases to 0.96% and the volatility of consumption growth decreases to 1.43%. A lower persistence of idiosyncratic productivity reduces the yearly autocorrelation in consumption to 0.35; moreover, aggregate TFP, output, and misallocation all become less persistent. Because of the decrease in consumption persistence, the Sharpe ratio declines from 0.41 in the baseline calibration to 0.30 in column (2).

In column (3), we consider a more relaxed collateral constraint by increasing $\lambda$ from 1 to 1.2. The average misallocation $M_t$ stays roughly unchanged at $-0.52$ compared to the baseline calibration. 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 $\text{Cov}(\bar{z}_{i,t}, \bar{a}_{i,t})$ in equation (52)). A change in $\lambda$ does not affect this difference much because a higher $\lambda$ scales up the revenue of both high-productivity and low-productivity firms. However, a higher $\lambda$ does lead to a higher productivity $Z_t$ in the final-goods sector because $\lambda$ directly increases $Z_t$ in equation (48), reflecting the instantaneous reallocation of capital through the capital leasing market. The higher $Z_t$ increases the average consumption growth rate to 2.46% and the volatility of consumption growth to 1.57%, without affecting the persistence of consumption growth and other macro variables. The Sharpe ratio is increased to 0.44.

In column (4), we consider a higher productivity of R&D by increasing $\chi$ from 3.35 to
3.45. Compared with our baseline calibration in column (1), column (4) shows that all the variables remain roughly unchanged, except for a higher consumption growth rate (2.48% vs. 1.77% in the baseline). The higher growth rate is not due to a better allocation of capital among firms because $Z_t$ is roughly unchanged. As discussed in Section 3.1, the parameter $\chi$ can be thought of as a pure scaling factor that determines the equilibrium growth rate.

In columns (5), we consider a smaller magnitude of capital depreciation shocks by reducing $\sigma_k$ from 0.13 to 0.12. Compared with our baseline calibration in column (2), the volatility of consumption growth decreases from 1.49% to 1.22% and the ratio of consumption growth volatility to output growth volatility decreases to 0.41. The Sharpe ratio decreases from 0.41 to 0.32 as the SDF becomes less volatile. The other variables reported in Table 2 remain roughly unchanged compared with the baseline calibration.

Finally, in column (6), we consider a larger magnitude of liquidity shocks by increasing $\sigma_a$ from 0.12 to 0.13. Compared with our baseline calibration in column (1), the volatility of consumption growth decreases from 1.49% to 1.26%, because liquidity shocks do not affect the evolution of misallocation (equation (52)), and are negatively correlated with capital depreciation shocks (50). As a result, the ratio of consumption growth volatility to output growth volatility decreases significantly to 0.39. The Sharpe ratio decreases to 0.33 as the SDF becomes less volatile.

### 3.4 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.
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)).

Our model implies that eliminating business cycles leads to a welfare gain of 0.58%, with a 95% confidence interval of [0.23%, 1.10%]. 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 $\chi$ in an economy without aggregate shocks to achieve the same growth rate of aggregate consumption as our baseline 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 245.25%, with a 95% confidence interval of [222.71%, 268.72%], 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 construct a measure of misallocation implied by our model. We show that misallocation in the U.S. data 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. We thus identify shocks to the misallocation as a proxy for shocks to the low-frequency components of consumption growth and the pricing kernel. Lastly, in Section 4.3, we study the cross-sectional asset pricing implications of misallocation.

---

8The 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.
Note: The red solid line plots the time series of our misallocation measure $\text{MisAlloc}_t$ (corresponding to the left y-axis). The pink bars represent the change in $\text{MisAlloc}_t$ (corresponding to the right y-axis). The shaded areas represent recessions or severe financial crises.

Figure 2: Time series plot of our misallocation measure $\text{MisAlloc}_t$.

4.1 Misallocation Measure

Motivated by our theory, we construct a measure of misallocation based on the U.S. Compustat data:

$$\text{MisAlloc}_t = \text{the HP filtered time series of } -\beta_t^\text{Alloc}. \quad (53)$$

The variable $\beta_t^\text{Alloc}$ captures the cross-sectional alignment between log productivity and log assets at time $t$, obtained from the following panel regression using firm-year observations from 1965 to 2016:

$$a_{i,t} = \alpha_t + \beta_t^\text{Alloc} z_{i,t} + \epsilon_{i,t}, \quad (54)$$

where the independent variable $a_{i,t} \equiv T^{-1} \sum_{\tau=1}^{T} \ln(ppent_{i,t+1-\tau})$ is the average log net property, plant and equipment of firm $i$ in the past $T$ years, i.e., from year $t+1-T$ to year $t$. The dependent variable $z_{i,t} \equiv T^{-1} \sum_{\tau=1}^{T} \ln(sales_{i,t+1-\tau}/production\_capital_{i,t+1-\tau})$ is a

\footnote{All our empirical results are robust if we use tangible net worth to construct $a_{i,t}$ in the data. A firm’s tangible net worth is constructed using its current assets plus net physical plant, property, and equipment plus other assets minus total liabilities. As emphasized in a seminal work by Chava and Roberts (2008),}
Table 5: Misallocation and R&D intensity.

<table>
<thead>
<tr>
<th>Sample period</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>$-0.076$</td>
<td>$-0.080$</td>
<td>$-0.076$</td>
<td>$-0.079$</td>
<td>$-0.073$</td>
<td>$-0.071$</td>
</tr>
<tr>
<td></td>
<td>$[-0.033]$</td>
<td>$[-0.032]$</td>
<td>$[-0.032]$</td>
<td>$[-0.031]$</td>
<td>$[-0.029]$</td>
<td>$[-0.029]$</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.089</td>
<td>0.102</td>
<td>0.102</td>
<td>0.116</td>
<td>0.121</td>
<td>0.117</td>
</tr>
</tbody>
</table>

measure of the average log productivity of firm $i$ in the past $T$ years, where productivity in year $t$ is constructed as the ratio of firm $i$’s sales to its assets for production in year $t$. We construct the capital for production by adding the rented capital to the net property, plant and equipment of firm $i$. Following the standard accounting practice and the literature (e.g., Rauh and Sufi, 2011; Rampini and Viswanathan, 2013), we capitalize the rental expenses to obtain an approximation of the amount of rented capital. Specifically, we capitalize the rental expense by a factor 10 and capped by a fraction, 0.25, of the net property, plant and equipment. To construct $a_{i,t}$ and $z_{i,t}$, we set $T = 3$. All empirical results are robust for alternative choices of $T$.

Figure 2 plots the time series of our misallocation measure $\text{MisAlloc}_t$. Sharp spikes in $\text{MisAlloc}_t$ are observed during periods of economic downturns, including economic recessions and three finance 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 by Figure 2 is consistent with our model’s prediction that a large increase in misallocation generally represents a period of time with macroeconomic recession and financial turmoil.

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 lenders commonly use a firm’s tangible net worth to assess the borrower’s ability to support and pay back loans (i.e., borrowing capacity). Naturally, tangible net worth, as a borrowing capacity measure, is widely reflected in loan covenants (e.g., DeAngelo, DeAngelo and Wruck, 2002; Roberts and Sufi, 2009; Sufi, 2009; Prilmeier, 2017).

All the empirical results are robust if we use a factor 5, 6, or 8; moreover, the results are also robust if we use a fraction 0, 0.5, 1, or 2.
of misallocation predicts R&D intensity, which determines consumption growth. Thus, misallocation measures could be economically meaningful predictors of R&D intensity and aggregate growth rates. In this subsection, we present empirical evidence supporting this prediction.

In Table 5, we regress the R&D intensity in the current year \( t \) and the next year \( t + 1 \) on the misallocation measure \( MisAlloc_t \). Columns (1), (3), and (5) show that a higher misallocation is associated with a contemporaneous decline in R&D intensity, which is robust across different sample periods. Columns (2), (4), and (6) of Table 5 further show that a higher misallocation predicts a further decline in R&D intensity in the next year.

Table 6 presents the results of projecting future consumption growth over horizons of one to five years on the misallocation measure \( MisAlloc_t \). The slope coefficients are negative and decreasing with horizons, and these results are robust across different sample periods. Specifically, the slope coefficients are statistically significant in columns (3) to (5) of panel B, which correspond to consumption growth over horizons of three to five years. The \( R^2 \) monotonically increases from 0.040 to 0.173 when time horizon increases from \( t \rightarrow t + 3 \) to \( t \rightarrow t + 5 \).

Table 7 presents the results of projecting future output growth over horizons of one to five years on the misallocation measure \( MisAlloc_t \). The slope coefficients are negative and decreasing with horizons, and become statistically significant in columns (4) and (5) with an \( R^2 \) of 0.078 and 0.132, respectively. These results are robust across different sample periods.

Taken together, we find evidence that the aggregate growth rates of consumption and output can be predicted by our misallocation measure over long horizons. Our regression results in Tables 6 and 7 lend empirical support to the notion of misallocation-driven low-
frequency variation in consumption and output 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 Asset Pricing Implications

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 misallocation measure $\text{MisAlloc}_t$ is a factor that is significantly priced in the cross section of asset returns for standard test portfolios. The test portfolios include 25 size-sorted and book-to-market-sorted portfolios, 10 momentum-sorted portfolios, and 6 maturity-sorted Treasury bond portfolios. The results are presented in Table 8 and visualized in Figures 3 and 4, which plot the realized mean excess returns of test portfolios against their predicted mean excess returns based on various factor models.

To elaborate, panel A of Figure 3 presents the predicted mean excess returns based on the CAPM as a benchmark, which fails to price our test portfolios ($R^2 = 0.30$). Panel B presents the results based on a two-factor model with market returns and $\text{MisAlloc}_t$. Panel C presents the results of the Fama-French three-factor model. Comparing panels B and C, our two-factor model with market returns and $\text{MisAlloc}_t$ produces a cross-sectional fit ($R^2 = 0.53$) comparable with the Fama-French three-factor model ($R^2 = 0.62$). Notably, the two-factor model better explains the expected returns of the 10 momentum-sorted portfolios than the Fama-French three-factor model, which is known to have a poor explanatory power for momentum-sorted portfolio returns. Once we include $\text{MisAlloc}_t$ in the Fama-French three-factor model in panel D, not surprisingly, the improvement
Note: We plot 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 mean excess returns predicted by various linear factor asset pricing models. We use yearly data from 1961 to 2016.

Figure 3: Realized versus predicted mean excess returns with MisAlloc.

mainly lies in the explanatory power for momentum portfolio returns ($R^2 = 0.68$).

Our theory suggests that the main channel through which long-run expected economic growth, especially consumption growth, affects asset returns 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 measure MisAlloc is already included as a factor. Following Parker and Julliard (2005), we use accumulated future consumption growth to approximate long-run expected consumption growth. 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 ($R^2 = 0.55$). In panel B, we augment this two-factor model with MisAlloc 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 asset returns. However, the coefficient on accumulated future consumption growth becomes statistically insignificant after including MisAlloc as a factor, which
Note: We plot 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 mean excess returns predicted by various linear factor asset pricing models. We use yearly data from 1961 to 2016.

Figure 4: Realized versus predicted mean excess returns with $MisAlloc_t$ and accumulated future consumption growth.

has a statistically significant coefficient (see column (6) of Table 8). Similar patterns are shown in panels C and D when we include $MisAlloc_t$ in the factor model that contains Fama-French three factors and accumulated future consumption growth.

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 heterogeneous 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
Table 8: Portfolio returns and model fit.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MKT</td>
<td>3.34</td>
<td>2.96</td>
<td>2.67</td>
<td>2.25</td>
<td>2.42</td>
<td>2.83</td>
<td>2.31</td>
<td>2.18</td>
</tr>
<tr>
<td>MKT, M</td>
<td>3.60</td>
<td>3.20</td>
<td>3.99</td>
<td>3.55</td>
<td>2.59</td>
<td>3.05</td>
<td>3.51</td>
<td>3.37</td>
</tr>
<tr>
<td>FF</td>
<td>3.47</td>
<td>1.46</td>
<td>3.65</td>
<td>1.60</td>
<td>1.72</td>
<td>1.50</td>
<td>2.57</td>
<td>1.55</td>
</tr>
<tr>
<td>FF, M</td>
<td>1.73</td>
<td>1.36</td>
<td>1.42</td>
<td>1.91</td>
<td>1.64</td>
<td>1.38</td>
<td>1.80</td>
<td>1.99</td>
</tr>
<tr>
<td>FF, C</td>
<td>1.28</td>
<td>0.58</td>
<td>0.98</td>
<td>0.79</td>
<td>0.94</td>
<td>0.62</td>
<td>1.08</td>
<td>0.83</td>
</tr>
<tr>
<td>FF, M, C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.12</td>
<td>-0.12</td>
<td>-0.10</td>
<td>-0.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mkt</td>
<td>4.92</td>
<td>3.82</td>
<td>3.66</td>
<td>4.86</td>
<td>4.63</td>
<td>3.88</td>
<td>4.59</td>
<td>5.05</td>
</tr>
<tr>
<td>Mkt, t-FM</td>
<td>[1.73]</td>
<td>[1.36]</td>
<td>[1.42]</td>
<td>[1.91]</td>
<td>[1.64]</td>
<td>[1.38]</td>
<td>[1.80]</td>
<td>[1.99]</td>
</tr>
<tr>
<td>Mkt, t-Shanken</td>
<td>[1.28]</td>
<td>[0.58]</td>
<td>[0.98]</td>
<td>[0.79]</td>
<td>[0.94]</td>
<td>[0.62]</td>
<td>[1.08]</td>
<td>[0.83]</td>
</tr>
<tr>
<td>MisAlloc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MisAlloc, t-FM</td>
<td>[-4.66]</td>
<td>[-5.74]</td>
<td>[-4.52]</td>
<td>[-5.54]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MisAlloc, t-Shanken</td>
<td>[-2.10]</td>
<td>[-2.54]</td>
<td>[-2.20]</td>
<td>[-2.41]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMB</td>
<td>3.01</td>
<td>2.40</td>
<td>2.61</td>
<td>2.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMB, t-FM</td>
<td>[1.54]</td>
<td>[1.23]</td>
<td>[1.35]</td>
<td>[1.31]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMB, t-Shanken</td>
<td>[1.05]</td>
<td>[0.52]</td>
<td>[0.80]</td>
<td>[0.55]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HML</td>
<td>4.36</td>
<td>4.05</td>
<td>4.40</td>
<td>4.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HML, t-FM</td>
<td>[2.16]</td>
<td>[2.00]</td>
<td>[2.18]</td>
<td>[2.07]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HML, t-Shanken</td>
<td>[1.46]</td>
<td>[0.84]</td>
<td>[1.30]</td>
<td>[0.87]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG</td>
<td>0.02</td>
<td>0.01</td>
<td>0.02</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG, t-FM</td>
<td>[3.32]</td>
<td>[0.67]</td>
<td>[4.06]</td>
<td>[1.07]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG, t-Shanken</td>
<td>[2.13]</td>
<td>[0.32]</td>
<td>[2.71]</td>
<td>[0.48]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Test diagnostics

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total MAPE</td>
<td>2.76</td>
<td>1.82</td>
<td>1.90</td>
<td>1.70</td>
<td>2.03</td>
<td>1.78</td>
<td>1.95</td>
<td>1.71</td>
</tr>
<tr>
<td>Size and B/M 25</td>
<td>2.77</td>
<td>1.77</td>
<td>1.30</td>
<td>1.68</td>
<td>1.50</td>
<td>1.68</td>
<td>1.39</td>
<td>1.68</td>
</tr>
<tr>
<td>Momentum 10</td>
<td>3.30</td>
<td>2.62</td>
<td>3.72</td>
<td>2.30</td>
<td>3.59</td>
<td>2.34</td>
<td>3.69</td>
<td>2.33</td>
</tr>
<tr>
<td>Bond 6</td>
<td>1.85</td>
<td>1.26</td>
<td>1.36</td>
<td>0.74</td>
<td>1.61</td>
<td>1.28</td>
<td>1.43</td>
<td>0.78</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.30</td>
<td>0.53</td>
<td>0.62</td>
<td>0.68</td>
<td>0.55</td>
<td>0.60</td>
<td>0.63</td>
<td>0.69</td>
</tr>
</tbody>
</table>

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 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 can potentially help address the issues of weak identification in the long-run risk literature (Chen, Dou and Kogan, 2019; Cheng, Dou and Liao, 2020).

References


Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta. 2013. “Cross-Country Diff-


Croce, Mariano Max, Thien T. Nguyen, and Steve Raymond. 2021. “Persistent Govern-

38


Khan, Aubhik, and Julia K. Thomas. 2008. “Idiosyncratic Shocks and the Role of Non-


A Proofs

A.1 Proof of Proposition 1

To ensure that shareholders do not intervene, the manager pays a dividend flow of $\zeta a_{i,t}dt$ to shareholders over $[t, t + dt]$ where $\zeta$ is to be determined. Thus, firm $i$’s total dividend payment is $(\tau + \zeta)a_{i,t}dt$ over $[t, t + dt)$, 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 $\zeta$, $\rho$, and $\tau$. Related to our discussion in Online Appendix A.2, the manager’s value is proportional to the firm’s capital, given by $\tilde{\xi}_{i,t}a_{i,t}$, where $\tilde{\xi}_{i,t}$ 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 $\zeta/\tau$ of the manager’s private benefits. Thus, shareholders’ value is $(\zeta/\tau)\tilde{\xi}_{i,t}a_{i,t}$. If shareholders intervene, the firm’s value will drop to $(1 - \tau/\rho)\tilde{\xi}_{i,t}a_{i,t}$ due to the loss of capital. But, because shareholders now are also managers, they will have the claim to all dividends, which generate a value of $(1 + \zeta/\tau)(1 - \tau/\rho)\tilde{\xi}_{i,t}a_{i,t}$. Thus, the manager chooses the intensity $\zeta$ of dividends to shareholders such that shareholders are indifferent about having an intervention or not:

$$\frac{\zeta}{\tau} \tilde{\xi}_{i,t}a_{i,t} = (1 + \zeta/\tau)(1 - \tau/\rho)\tilde{\xi}_{i,t}a_{i,t}, \quad (55)$$

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

$$\tau + \zeta = \rho. \quad (56)$$

A.2 Proof of Lemma 1

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

\[ \ell_{i,t} = \left[ \frac{w_t}{(1-\alpha)(1-\varepsilon)} \right]^{\frac{1}{(1-\alpha)(1-\varepsilon)-1}}. \quad (57) \]

Substituting equations (2), (16), (19) and (57) into (3):

\[
d\alpha_{i,t} = - \int_0^{N_i} p_{j,t}x_{i,j,t} \, dj \, dt - u_{i,t}k_{i,t} (\delta_k \, dt + \sigma_k \, dW_t) + a_{i,t} (-\delta_a \, dt + \sigma_a \, dW_t) - r_{f,t} \hat{a}_{i,t} \, dt - \rho \alpha_{i,t} \, dt
+ \left[ 1 - (1-\alpha)(1-\varepsilon) \right] \left[ \frac{w_t}{(1-\alpha)(1-\varepsilon)} \right]^{\frac{1}{(1-\alpha)(1-\varepsilon)-1}} (z_{i,t}u_{i,t}k_{i,t})^{\frac{1}{(1-\alpha)(1-\varepsilon)-1}} x_{i,t}^{\frac{\alpha(1-\varepsilon)}{1-\alpha(1-\varepsilon)}} \, dt. \quad (58)\]

Taking the first-order condition with respect to \( x_{i,j,t} \) in the right-hand side of equation (58), we derive firm \( i \)'s optimal demand for intermediate goods \( j \in [0, N_i] \):

\[
x_{i,j,t} = \left[ \frac{\varepsilon}{p_{j,t}} \right]^{\frac{1-\alpha(1-\varepsilon)}{\alpha(1-\varepsilon)}} \left[ \frac{(1-\alpha)(1-\varepsilon)}{w_t} \right]^{\frac{1-\alpha}{\alpha}} z_{i,t}u_{i,t}k_{i,t}. \quad (59)\]

Substituting into equation (4), we derive \( x_{i,t} \):

\[
x_{i,t} = \left( \frac{\varepsilon}{p_t} \right)^{\frac{1-\alpha(1-\varepsilon)}{\alpha(1-\varepsilon)}} \left[ \frac{(1-\alpha)(1-\varepsilon)}{w_t} \right]^{\frac{1-\alpha}{\alpha}} z_{i,t}u_{i,t}k_{i,t}, \quad (60)\]

where the price index \( p_t \) is given by

\[
p_t = \left( \int_0^{N_i} p_{j,t} \, dj \right)^{\frac{\nu-1}{\nu}}. \quad (61)\]

Substituting equation (60) into (57), we obtain (24). Substituting equation (60) into (59), we obtain (25). Substituting equation (60) and \( \int_0^{N_i} p_{j,t}x_{i,j,t} \, dj = p_t x_{i,t} \) into (58), we obtain

\[
d\alpha_{i,t} = - u_{i,t}k_{i,t} (\delta_k \, dt + \sigma_k \, dW_t) + a_{i,t} (-\delta_a \, dt + \sigma_a \, dW_t) - r_{f,t} \hat{a}_{i,t} \, dt - \rho \alpha_{i,t} \, dt
+ \alpha (1-\varepsilon) \left( \frac{\varepsilon}{p_t} \right)^{\frac{1-\alpha(1-\varepsilon)}{\alpha(1-\varepsilon)}} \left[ \frac{(1-\alpha)(1-\varepsilon)}{w_t} \right]^{\frac{1-\alpha}{\alpha}} z_{i,t}u_{i,t}k_{i,t} \, dt. \quad (62)\]

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

\[ 0 = \max_{\tilde{a}_{i,t}, \tilde{u}_{i,t}} \tau a_{i,t} \, dt + \mathbb{E}_t \left[ \frac{d\Lambda_t}{\Lambda_t} J_{i,t} + df_{i,t} + \frac{d\Lambda_t}{\Lambda_t} dJ_{i,t} \right]. \]  

(63)

subject to the budget constraint (62). 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:

\[ J_{i,t} \equiv J_t(a_{i,t}, z_{i,t}) = \tilde{\xi}_t(z_{i,t}) a_{i,t}, \]  

(64)

where \( \tilde{\xi}_t(z_{i,t}) = \tilde{\xi}_t(z_{i,t}) \) 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. Substituting equations (20) and (64) into (63), we obtain

\[ 0 = \max_{\tilde{a}_{i,t}, \tilde{u}_{i,t}} \tau a_{i,t} \, dt + \mathbb{E}_t \left[ (1 - r_{f,t} \, dt - \eta_t dW_t) (\xi_t a_{i,t} + \tilde{\xi}_t(t) a_{i,t} + d\tilde{\xi}_t a_{i,t}) \right]. \]  

(65)

The variable \( \tilde{\xi}_t \) evolves as follows:

\[ \frac{d\tilde{\xi}_t}{\tilde{\xi}_t} = \mu_{\tilde{\xi}_t} dt + \sigma_{\tilde{\xi}_t} dW_t + \sigma_{w,t} dW_{i,t}, \]  

(66)

where \( \mu_{\tilde{\xi}_t} \equiv \mu_{\tilde{\xi}_t}(z_{i,t}), \sigma_{\tilde{\xi}_t} \equiv \sigma_{\tilde{\xi}_t}(z_{i,t}), \) and \( \sigma_{w,t} \equiv \sigma_{w,t}(z_{i,t}) \) are endogenously determined in equilibrium. Using equations (62) and (66), and the properties that \((dW_t)^2 = dt, \mathbb{E}_t[dW_{i,t}] = \mathbb{E}_t[dW_t] = \mathbb{E}_t[dW_i dW_{i,t}] = 0\), we obtain the following equations after omitting higher-order terms:

\[ \mathbb{E}_t \left[ (1 - r_{f,t} \, dt - \eta_t dW_t) \tilde{\xi}_t a_{i,t} \right] = -r_{f,t} a_{i,t} \tilde{\xi}_t dt, \]  

(67)

\[ \mathbb{E}_t \left[ (1 - r_{f,t} \, dt - \eta_t dW_t) d\tilde{\xi}_t a_{i,t} \right] = \mu_{\tilde{\xi}_t} a_{i,t} \tilde{\xi}_t dt - \eta_t \sigma_{\tilde{\xi}_t} a_{i,t} \tilde{\xi}_t dt. \]  

(68)

\[ \mathbb{E}_t \left[ (1 - r_{f,t} \, dt - \eta_t dW_t) \tilde{\xi}_t d\tilde{\xi}_t a_{i,t} \right] = \mathbb{E}_t \left[ (1 + \sigma_{\tilde{\xi}_t} dW_t - \eta_t dW_t) \tilde{\xi}_t d\tilde{\xi}_t a_{i,t} \right]
\]

\[ = \sigma_k (\eta_t - \sigma_{\tilde{\xi}_t} - \delta_k) u_{i,t} k_{i,t} \tilde{\xi}_t dt - \sigma_{\alpha} (\eta_t - \sigma_{\tilde{\xi}_t}) \tilde{\xi}_t a_{i,t} dt - r_{f,t} \tilde{\alpha} a_{i,t} \tilde{\xi}_t dt - \left( \delta_t + \sigma_{\alpha} a_{i,t} \right) \tilde{\xi}_t dt + a(1 - \varepsilon) \left( \frac{\varepsilon}{P_t} \right)^{\frac{1}{\alpha}} \frac{(1 - \alpha)(1 - \varepsilon)}{w_t} \tilde{\xi}_t dt. \]  

(69)
Substituting equations (67), (68), (69) into (65), we obtain

\[ 0 = \max_{\tilde{a}_{i,t}, u_{i,t}} \left( r_{f,t} a_{i,t} \tilde{\epsilon}_{i,t} dt - \mu_{\tilde{\epsilon}_{i,t}} a_{i,t} \tilde{\epsilon}_{i,t} dt + \eta_i \sigma_{\tilde{\epsilon}_{i,t}} a_{i,t} \tilde{\epsilon}_{i,t} dt - \eta_i \sigma_{\tilde{\epsilon}_{i,t}} a_{i,t} \tilde{\epsilon}_{i,t} dt \right) \\
\] 

\[ + \alpha(1 - \varepsilon) \left( \frac{\varepsilon}{p_t} \right)^{\frac{\varepsilon}{\alpha(1 - \varepsilon)}} \left( \frac{1 - \alpha}{w_t} \right) \frac{1}{z_{i,t} u_{i,t} \tilde{a}_{i,t} \tilde{\epsilon}_{i,t} dt}. \]  

(70)

Using \( k_{i,t} = a_{i,t} + \tilde{a}_{i,t} \), we can see that maximizing equation (70) is essentially the same as maximizing

\[ 0 = \max_{\tilde{a}_{i,t}, u_{i,t}} \left[ \sigma_k (\eta_i - \sigma_{\tilde{\epsilon}_{i,t}}) - \delta_k u_{i,t} \tilde{a}_{i,t} \tilde{\epsilon}_{i,t} dt - r_{f,t} \tilde{a}_{i,t} \tilde{\epsilon}_{i,t} dt \right] \\
\] 

\[ + \alpha(1 - \varepsilon) \left( \frac{\varepsilon}{p_t} \right)^{\frac{\varepsilon}{\alpha(1 - \varepsilon)}} \left( \frac{1 - \alpha}{w_t} \right) \frac{1}{z_{i,t} u_{i,t} \tilde{a}_{i,t} \tilde{\epsilon}_{i,t} dt}. \]  

(71)

Because a positive shock (\( dW_t > 0 \)) increases misallocation through higher capital depreciation of productive firms, we have \( \eta_i < 0 \) in equilibrium. Moreover, because \( \tilde{\epsilon}_{i,t} \) is not affected by the manager’s choice of \( \tilde{a}_{i,t} \). The objective function (71) is linear in both \( \tilde{a}_{i,t} \) and \( u_{i,t} \). Thus, conditional on \( u_{i,t} = 1 \), we can characterize the productivity cutoff \( z_t \) that makes the manager indifferent about leasing capital as follows:

\[ z_t \kappa_t = r_{f,t} + \delta_k + \sigma_k (\sigma_{\tilde{\epsilon}_{i,t}} - \eta_i), \]  

(72)

where

\[ \kappa_t = \alpha(1 - \varepsilon) \left( \frac{\varepsilon}{p_t} \right)^{\frac{\varepsilon}{\alpha(1 - \varepsilon)}} \left( \frac{1 - \alpha}{w_t} \right) \frac{1}{z_{i,t} u_{i,t} \tilde{a}_{i,t} \tilde{\epsilon}_{i,t} dt}. \]  

(73)

Because \( r_{f,t} > 0 \) in equilibrium, it is clear that firms will optimally choose \( u_{i,t} = 1 \) for \( z_{i,t} \geq z_t \). The optimal leasing amount follows a bang-bang solution:

\[ \tilde{a}_t(a, z) = \begin{cases} 
\lambda a, & z \geq z_t \\
-a, & z < z_t 
\end{cases}, \]  

(44)

which leads to the bang-bang solution in capital:

\[ k_t(a, z) = \begin{cases} 
(1 + \lambda) a, & z \geq z_t \\
0, & z < z_t \end{cases} \].  

(75)

\(^{11}\)In other words, the bang-bang cutoff productivity for \( u_{i,t} \) is lower than the cutoff productivity for \( \tilde{a}_{i,t} \) when \( r_{f,t} > 0 \).
The optimal capacity utilization intensity is given by

\[ u_t(z) = \begin{cases} 
1, & z \geq z_t \\
0, & z < z_t 
\end{cases} \tag{76} \]

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

### A.3 Proof of Proposition 2

Define the productivity \( Z_t \) of the final goods sector as

\[ Z_t = \left[ \frac{1}{K_t} \int_{z_t}^{\infty} \int_{0}^{\infty} z u_t(z) k_t(a,z) \varphi_t(a,z) dk dz \right]^\alpha, \tag{77} \]

Using equations (23), (32), (38), and \( k_t(a,z) = a + \bar{a}_t(a,z) \), \( Z_t \) can be written as

\[ Z_t = \left[ (1 + \lambda) \frac{A_t}{K_t} \int_{z_t}^{\infty} z \omega_t(z) dz \right]^\alpha. \tag{78} \]

Substituting equation (75) into the capital market-clearing condition (33), we obtain

\[ (1 + \lambda) \int_{z_t}^{\infty} \int_{0}^{\infty} a \varphi_t(a,z) dk dz = K_t. \tag{79} \]

Given the definition of capital share (38), the left-hand side of equation (79) can be simplified as

\[ (1 + \lambda) \int_{z_t}^{\infty} \int_{0}^{\infty} a \varphi_t(a,z) dk dz = (1 + \lambda) A_t \int_{z_t}^{\infty} \omega_t(z) dz = (1 + \lambda) A_t (1 - \Omega_t(z_t)). \tag{80} \]

Thus, we have the following equation

\[ (1 + \lambda) (1 - \Omega_t(z_t)) = \frac{K_t}{A_t}, \tag{81} \]

which determines the equilibrium \( K_t / A_t \). Substituting equation (81) into (78), we obtain (41).

Substituting equation (7) into (6), we obtain

\[ \pi_{j,t} = \max_{p_{j,t}} (p_{j,t} - 1) \left( \frac{p_{j,t}}{p_t} \right)^{\frac{1}{\gamma_t}} X_t, \tag{82} \]
Taking the first-order condition, we obtain

\[ p_{j,t} = \frac{1}{\nu} \quad \text{for all } j. \]  

(83)

Substituting equation (83) into the price index (26), we obtain

\[ p_t = N_t^{\nu} \frac{p_{j,t}}{p_t} = N_t^{\nu} \frac{1}{\nu}. \]  

(84)

Substituting equation (24) into (35) and using (77), we obtain

\[ L_t = \left( \frac{\varepsilon}{p_t} \right)^{\frac{\alpha}{1-\alpha}} \left[ \frac{(1-\alpha)(1-\varepsilon)}{w_t} \right]^{\frac{1}{\alpha}} \int_{z_t}^{\infty} \int_{0}^{\infty} zu_t(z)k_t(a,z)q_t(a,z)dadz \]

\[ = \left( \frac{\varepsilon}{p_t} \right)^{\frac{\alpha}{1-\alpha}} \left[ \frac{(1-\alpha)(1-\varepsilon)}{w_t} \right]^{\frac{1}{\alpha}} Z_t^{\frac{1}{\alpha}} K_t. \]  

(85)

Substituting equation (84) into (85), we derive the equilibrium wage \( w_t \):

\[ w_t = (1-\alpha)(1-\varepsilon)(\varepsilon)^{\frac{\alpha}{1-\epsilon}} N_t^{\frac{1-\alpha}{(1-\epsilon)}} Z_t(K_t/L_t)^{\alpha}. \]  

(86)

By definition, the aggregate output \( Y_t \) is

\[ Y_t = \int_{z_t}^{\infty} \int_{0}^{\infty} \left[ (zu_t(z)k_t(a,z))^\alpha \ell_t(a,z)^{1-\alpha} \right]^{1-\varepsilon} x_t(a,z)^{\varepsilon} q_t(a,z)dadz. \]  

(87)

Substituting equations (24) and (27) into (87), we obtain

\[ Y_t = \left( \frac{\varepsilon}{p_t} \right)^{\frac{\alpha}{1-\alpha}} \left[ \frac{(1-\alpha)(1-\varepsilon)}{w_t} \right]^{\frac{1}{\alpha}} \int_{z_t}^{\infty} \int_{0}^{\infty} zu_t(z)k_t(a,z)q_t(a,z)dadz. \]  

(88)

Further substituting equations (77), (84) and (86) into the above equation, we obtain

\[ Y_t = (\varepsilon)^{\frac{\alpha}{1-\alpha}} Z_t N_t^{\frac{1-\alpha}{(1-\epsilon)}} K_t^\alpha L_t^{1-\alpha}. \]  

(89)

Using equation (89), the equilibrium wage \( w_t \) in (86) can be simplified as

\[ w_t = (1-\alpha)(1-\varepsilon) \frac{Y_t}{L_t}. \]  

(90)
Equation (29) can be simplified by substituting equations (84) and (90) into (29):

\[ \kappa_t = \alpha (1 - \varepsilon) (\varepsilon \nu) \left( \varepsilon + \alpha - 1 \right) - \alpha (1 - \varepsilon) Z_t \frac{1}{K_t}. \]

(91)

Substituting equations (27), (83) and (84) into (82) and using (77), we obtain

\[ \pi_t = \frac{1 - \nu}{\nu} (\varepsilon \nu) \left( \varepsilon + \alpha - 1 \right) - \frac{1}{\nu} \left( \varepsilon + \alpha - 1 \right) \left( \frac{1 - \alpha}{\bar{w}_t} \right) \left[ \frac{1 + \alpha}{1 - \nu} \right] - \frac{1}{\nu} \left( \varepsilon + \alpha - 1 \right) N_t \pi_t = \left( 1 - \nu \right) \frac{Y_t}{N_t}. \]

(92)

Further, substituting equation (86) into the above equation and using (89), we obtain

\[ \pi_t = \frac{1 - \nu}{\nu} (\varepsilon \nu) \left( \varepsilon + \alpha - 1 \right) - \frac{1}{\nu} \left( \varepsilon + \alpha - 1 \right) \left( \frac{1 - \alpha}{\bar{w}_t} \right) \left[ \frac{1 + \alpha}{1 - \nu} \right] - \frac{1}{\nu} \left( \varepsilon + \alpha - 1 \right) N_t \pi_t = \left( 1 - \nu \right) \frac{Y_t}{N_t}. \]

(93)

Thus, we have

\[ \int_{j=0}^{N_t} \pi_t dj = N_t \pi_t = (1 - \nu) \varepsilon Y_t. \]

(94)

Substituting equation (10) into (11), we obtain

\[ S_t = \left( \chi \nu \right) \frac{1}{\pi} N_t. \]

(95)

### A.4 Resource Constraint

By definition, the aggregate output \( Y_t dt \) is

\[ Y_t dt = \int_{t}^{\infty} \int_{0}^{\infty} y_t(a, z) dt \varphi_t(a, z) daca = \int_{0}^{\infty} \int_{0}^{\infty} y_t(a, z) dt \varphi_t(a, z) daca. \]

(96)

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

\[ Y_t dt = dA_t + (\delta_a dt - \sigma_a dW_t)A_t + w_t L_t dt + (\delta_k dt + \sigma_k dW_t)K_t + r_{f, t} B_t dt + \rho A_t dt \]

\[ + \int_{0}^{\infty} \int_{0}^{\infty} \left( \int_{0}^{N_t} p_{j, t} x_j(t, a, z) dja \right) \varphi_t(a, z) daca, \]

(97)

where the last term is the revenue of the intermediate goods sector. Using equations (7), (25) and the definition \( X_t \equiv \int_{i \in I} x_{i, t} di = \int_{0}^{\infty} \int_{0}^{\infty} x_t(a, z) \varphi_t(a, z) daca, \) it can be simplified
as follows

\[
\int_0^\infty \int_0^\infty \left( \int_0^{N_i} p_{j,t} x_{j,t}(a,z) \, dj \, dt \right) \varphi_t(a,z) \, da \, dz = \int_0^N_i \left( \int_0^\infty \int_0^\infty p_{j,t} x_{j,t}(a,z) \varphi_t(a,z) \, da \, dz \right) \, dj \, dt = \int_0^N_i p_{j,t} e_{j,t} \, dj \, dt = \int_0^N_i \pi_{j,t} \, dj \, dt + \int_0^N_i e_{j,t} \, dj \, dt. \tag{98}
\]

Substituting equation (98) into (97), we obtain

\[
Y_t \, dt = dA_t + (\delta_a \, dt - \sigma_a \, dW_t) A_t + w_t L_t \, dt + (\delta_k \, dt + \sigma_k \, dW_t) K_t + r_f \, B_t \, dt + \rho A_t \, dt \\
+ \int_0^N_i \pi_{j,t} \, dj \, dt + \int_0^N_i e_{j,t} \, dj \, dt, \tag{99}
\]

Substituting equations (14) and (31) into (99), we obtain the resource constraint

\[
Y_t \, dt = dA_t + (\delta_a \, dt - \sigma_a \, dW_t) A_t + (\delta_k \, dt + \sigma_k \, dW_t) K_t \\
+ \text{investment in the final goods sector} + S_t \, dt + \int_0^N_i e_{j,t} \, dj \, dt + C_t \, dt - dB_t. \tag{100}
\]

Note that the resource constraint (100) holds by Walras’s law in equilibrium. This can be proved by substituting equations (44) and (50) into (99), and using the condition below

\[
\int_0^\infty \int_0^\infty \left( \int_0^{N_i} p_{j,t} x_{j,t}(a,z) \, dj \, dt \right) \varphi_t(a,z) \, da \, dz = \epsilon Y_t \, dt, \tag{101}
\]

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

### A.5 Proof of Lemma 3

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

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

\[
\varphi_t(a,z) = |J| \psi_t(\tilde{a}, \tilde{z}), \tag{102}
\]
where \( J \) is defined by

\[
J = \begin{pmatrix}
\partial \tilde{a} / \partial a & \partial \tilde{a} / \partial z \\
\partial \tilde{z} / \partial a & \partial \tilde{z} / \partial z
\end{pmatrix}.
\] (103)

Thus, we have

\[
\varphi_t(a, z) = \frac{1}{az} \psi_t(\tilde{a}, \tilde{z}).
\] (104)

Using equation (104), the term \( \int_0^\infty a \varphi_t(a, z) da \) in equation (38) can be written as

\[
\int_0^\infty a \varphi_t(a, z) da = \int_{-\infty}^\infty \frac{a}{z} \psi_t(\tilde{a}, \tilde{z}) d\tilde{a}.
\] (105)

Let \( f(\tilde{z}) \) be the PDF of \( \tilde{z} \), which follows a normal distribution, \( N(0, \sigma^2/2) \), in the stationary equilibrium. Thus, equation (105) can be written as

\[
\int_0^\infty a \varphi_t(a, z) da = \int_{-\infty}^\infty \frac{a}{z} \psi_t(\tilde{a}, \tilde{z}) f(\tilde{z}) d\tilde{a} = \mathbb{E}[\exp(\tilde{a}_{i,t})/z].
\] (106)

Using equation (104), the variable \( A_t \) defined in (32) can be written as

\[
A_t = \int_{-\infty}^\infty \int_{-\infty}^\infty a \frac{1}{az} \psi_t(\tilde{a}, \tilde{z}) az d\tilde{a} d\tilde{z} = \mathbb{E}[\exp(\tilde{a}_{i,t})].
\] (107)

Substituting equations (106) and (107) into (38), we obtain

\[
\omega_t(z) = \frac{\mathbb{E}[\exp(\tilde{a}_{i,t})/z] f(\tilde{z})}{\mathbb{E}[\exp(\tilde{a}_{i,t})]/z}.
\] (108)

Because \( \tilde{a}_{i,t} \) and \( \tilde{z}_{i,t} \) follow a joint normal distribution with covariance \( \Gamma_t \), we have

\[
\mathbb{E}[\exp(\tilde{a}_{i,t})/z] = \exp\left( \mathbb{E}[\tilde{a}_{i,t}]/z + \frac{1}{2} \text{var}(\tilde{a}_{i,t}/z) \right),
\] (109)

\[
\mathbb{E}[\exp(\tilde{a}_{i,t})] = \exp\left( \mathbb{E}[\tilde{a}_{i,t}] + \frac{1}{2} \text{var}(\tilde{a}_{i,t}) \right),
\] (110)

where

\[
\mathbb{E}(\tilde{a}_{i,t}|z) = \mathbb{E}(\tilde{a}_{i,t}) + 2\tilde{z}_t/\sigma^2,
\] (111)

\[
\text{var}(\tilde{a}_{i,t}|z) = \text{var}(\tilde{a}_{i,t}) - 2\Gamma_t^2/\sigma^2.
\] (112)
Substituting equations (109) to (112) into (108), we obtain

\[
\omega_t(z) = \frac{f(\tilde{z})}{z} \exp \left(2\tilde{z}\frac{\Gamma_t}{\sigma^2}\right) \exp \left(-\frac{\Gamma_t^2}{\sigma^2}\right)
\]

\[
= \frac{1}{z\sigma\sqrt{\pi}} \exp \left(-\frac{(\ln z - \Gamma_t)^2}{\sigma^2}\right).
\] (113)

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

### A.6 Proof of Proposition 3

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

\[
\frac{1}{1 + \lambda} \frac{K_t}{A_t} = 1 - \Omega_t(\tilde{z}_t) = \int_{\tilde{z}_t}^{\infty} \frac{1}{\sigma\sqrt{\pi}} \exp \left(-\frac{(\tilde{z} - \Gamma_t)^2}{2\sigma^2}\right) d\tilde{z} = \Phi \left(\frac{\Gamma_t - \tilde{z}_t}{\sigma / \sqrt{2}}\right).
\] (114)

Rearranging the above equation, we obtain \(\tilde{z}_t\):

\[
\tilde{z}_t = \exp \left(\Gamma_t - \Phi^{-1}\left(\frac{1}{1 + \lambda} \frac{K_t}{A_t}\right) \frac{\sigma}{\sqrt{2}}\right).
\] (115)

The term \(\int_{\tilde{z}_t}^{\infty} z\omega_t(z)dz\) in equation (41) can be simplified using (47), as follows

\[
\int_{\tilde{z}_t}^{\infty} z\omega_t(z)dz = \int_{\tilde{z}_t}^{\infty} \frac{1}{z\sigma\sqrt{\pi}} \exp \left(-\frac{(\tilde{z} - \Gamma_t)^2}{2\sigma^2}\right) dz
\]

\[
= \int_{\tilde{z}_t}^{\infty} \frac{1}{\sigma\sqrt{\pi}} \exp \left(-\frac{(\tilde{z} - \Gamma_t - \sigma^2/2)^2}{2\sigma^2}\right) \exp \left(\Gamma_t + \frac{\sigma^2}{4}\right) d\tilde{z}
\]

\[
= \exp \left(\Gamma_t + \frac{\sigma^2}{4}\right) \Phi \left(\frac{\Gamma_t + \sigma^2/2 - \tilde{z}_t}{\sigma / \sqrt{2}}\right).
\] (116)

Substituting equation (116) into (41), we obtain

\[
Z_t = \left[\left(1 + \lambda\right) \frac{A_t}{K_t} \exp \left(\Gamma_t + \frac{\sigma^2}{4}\right) \Phi \left(\frac{\Gamma_t + \sigma^2/2 - \tilde{z}_t}{\sigma / \sqrt{2}}\right)\right]^a.
\] (117)

Further, substituting equation (115) into the above equation, we obtain

\[
Z_t = \left[\left(1 + \lambda\right) \frac{A_t}{K_t} \exp \left(\Gamma_t + \frac{\sigma^2}{4}\right) \Phi \left(\Phi^{-1}\left(\frac{1}{1 + \lambda} \frac{K_t}{A_t}\right) + \frac{\sigma}{\sqrt{2}}\right)\right]^a.
\] (118)
Thus, omitting the higher-order term $d\Gamma_t = -M_t \text{var}(\tilde{z}_{i,t}) = -M_t \sigma^2 / 2$, we obtain equation (48) in the main text.

### A.7 Proof of Proposition 4

Equation (32) implies

$$A_{t+dt} - A_t = \int_0^\infty \int_0^\infty d\alpha_t(a,z) \varphi_t(a,z) dz \, da_t.$$

(119)

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

$$A_{t+dt} - A_t = (1 + \lambda) \kappa_t \int_0^\infty \int_\mathbb{R}_+ z \, d\alpha_t(a,z) dz \, da_t - (1 + \lambda) r_f \int_0^\infty \int_\mathbb{R}_+ a \, d\varphi_t(a,z) dz \, da_t$$

$$- (1 + \lambda) (\delta_k dt + \sigma_k dW_t) \int_0^\infty \int_\mathbb{R}_+ a \varphi_t(a,z) dz \, da_t + \sigma_a A_t dW_t + (r_f - \rho - \delta_a) A_t dt.$$  (120)

Using equations (77), (79), and (91), the above equation can be simplified as

$$dA_t = \alpha (1 - \epsilon) Y_t dt - (r_f + \delta_k) K_t dt - (\rho + \delta_a - r_f) A_t dt + (\sigma_a A_t - \sigma K_t) dW_t.$$  (121)

Substituting equations (11) and (46) into (9), we obtain

$$\frac{dN_t}{N_t} = \chi (\chi^\prime_t)^{1-h} dt - \delta_b dt.$$  (122)

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

$$\tilde{z}_{i,t+dt} = \tilde{z}_{i,t} - \theta \tilde{z}_{i,t} dt + \sigma \sqrt{\theta} dW_{i,t}.$$  (123)

Thus,

$$d\Gamma_t = \text{Cov} \left( \tilde{a}_{i,t} + d\tilde{a}_{i,t}, \tilde{z}_{i,t} - \theta \tilde{z}_{i,t} dt + \sigma \sqrt{\theta} dW_{i,t} \right) - \Gamma_t$$

$$= (1 - \theta dt) \text{Cov}(\tilde{a}_{i,t}, \tilde{z}_{i,t}) - \Gamma_t$$

$$= - \theta \Gamma_t dt + (1 - \theta dt) \text{Cov}(\tilde{z}_{i,t}, d\tilde{a}_{i,t}).$$  (124)

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

$$d\Gamma_t = - \theta \Gamma_t dt + \text{Cov}(\tilde{z}_{i,t}, d\tilde{a}_{i,t}).$$  (125)

Substituting out $\Gamma_t = -M_t \text{var}(\tilde{z}_{i,t}) = -M_t \sigma^2 / 2$, we obtain equation (52) in the main text.
We now derive the expression for \( \text{Cov}(\tilde{z}_{i,t}, d\tilde{a}_{i,t}) \) under Lemma 2. Using Ito’s lemma
\[
d\tilde{a}_{i,t} = \frac{1}{a_{i,t}} da_{i,t} - \frac{1}{2a_{i,t}^2} (da_{i,t})^2. \tag{126}
\]
Substituting equation (2), (3), and (30) into the above equation, we obtain the evolution of \( \tilde{a}_{i,t} \). In particular, for \( z_{i,t} < \tilde{z}_t \), we have
\[
d\tilde{a}_{i,t} = (r_{f,t} - \rho - \delta_a) dt + \sigma_a dW_t. \tag{127}
\]
For \( z_{i,t} \geq \tilde{z}_t \), we have
\[
d\tilde{a}_{i,t} = (1 + \lambda) [\kappa z_{i,t} dt - (\delta_k dt + \sigma_k dW_t) - r_{f,t} dt] + (r_{f,t} - \rho - \delta_a) dt + \sigma_a dW_t. \tag{128}
\]
Because \( \mathbb{E}[\tilde{z}_{i,t}] = 0 \), we have
\[
\text{Cov}(\tilde{z}_{i,t}, d\tilde{a}_{i,t}) = \mathbb{E}[\tilde{z}_{i,t} d\tilde{a}_{i,t}]. \tag{129}
\]
Substituting equations (127) and (128) into (129), we obtain
\[
\text{Cov}(\tilde{z}_{i,t}, d\tilde{a}_{i,t}) = (1 + \lambda) \kappa \int_{\tilde{z}_t}^{\infty} z f(\tilde{z}) d\tilde{z} - (1 + \lambda) [(r_{f,t} + \delta_k) dt + \sigma_k dW_t] \int_{\tilde{z}_t}^{\infty} \tilde{z} f(\tilde{z}) d\tilde{z},
\]
where \( f(\tilde{z}) \) is the PDF of \( \tilde{z} \), which follows a normal distribution, \( N(0, \sigma^2/2) \), in the stationary equilibrium.

Substituting out \( f(\tilde{z}) \), the term \( \int_{\tilde{z}_t}^{\infty} \tilde{z} f(\tilde{z}) d\tilde{z} \) in equation (130) can be simplified as follows:
\[
\int_{\tilde{z}_t}^{\infty} \tilde{z} f(\tilde{z}) d\tilde{z} = \int_{\tilde{z}_t}^{\infty} \tilde{z} \frac{1}{\sigma \sqrt{\pi}} \exp \left( -\frac{\tilde{z}^2}{\sigma^2} \right) d\tilde{z} = -\frac{\sigma}{2 \sqrt{\pi}} \int_{\tilde{z}_t}^{\infty} d \exp \left( -\frac{\tilde{z}^2}{\sigma^2} \right) = \frac{\sigma}{2 \sqrt{\pi}} \exp \left( -\frac{\tilde{z}_t^2}{\sigma^2} \right). \tag{131}
\]
The term $\int_{\tilde{z}}^{\infty} \tilde{z} f(\tilde{z}) d\tilde{z}$ in equation (130) becomes:

$$
\int_{\tilde{z}}^{\infty} \tilde{z} f(\tilde{z}) d\tilde{z} = \frac{1}{\sigma \sqrt{\pi}} \int_{\tilde{z}}^{\infty} \tilde{z} \exp \left( -\sigma^2 \tilde{z}^2 \right) d\tilde{z}
$$

$$
= \frac{1}{\sigma \sqrt{\pi}} \exp \left( \frac{\sigma^2}{4} \right) \int_{\tilde{z}}^{\infty} \tilde{z} \exp \left( -\frac{1}{\sigma^2} \left( \tilde{z} - \frac{\sigma^2}{2} \right)^2 \right) d\tilde{z}
$$

$$
= \frac{1}{\sigma \sqrt{\pi}} \exp \left( \frac{\sigma^2}{4} \right) \left[ \frac{1}{\sqrt{\pi}} \exp \left( -\frac{1}{\sigma^2} \left( \tilde{z} - \frac{\sigma^2}{2} \right)^2 \right) + \sigma \Phi \left( \frac{\sigma^2 / 2 - \tilde{z}}{\sigma / \sqrt{2}} \right) \right].
$$

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

$$
\int_{\tilde{z}}^{\infty} \tilde{z} f(\tilde{z}) d\tilde{z} = -\frac{\sigma}{2 \sqrt{\pi}} \exp \left( \frac{\sigma^2}{4} \right) \int_{\tilde{z}}^{\infty} \tilde{z} \exp \left( -\frac{1}{\sigma^2} \left( \tilde{z} - \frac{\sigma^2}{2} \right)^2 \right) d\tilde{z} + \frac{\sigma^2}{2} \exp \left( \frac{\sigma^2}{4} \right) \Phi \left( \frac{\sigma^2 / 2 - \tilde{z}}{\sigma / \sqrt{2}} \right).
$$

Substituting equations (131) and (133) into (130), we obtain

$$
\text{Cov}(\tilde{z}_{i,t}, d\tilde{a}_{i,t}) = \left( 1 + \lambda \right) \frac{\sigma^2 \kappa_t}{2} \exp \left( \frac{\sigma^2}{4} \right) \Phi \left( \frac{\sigma^2 / 2 - \tilde{z}}{\sigma / \sqrt{2}} \right) dt
$$

$$
+ \frac{(1 + \lambda)\sigma}{2 \sqrt{\pi}} \left[ (\tilde{z}_t \kappa_t - r_{f,t} \delta_k) dt - \sigma dW_t \right] \exp \left( -\frac{\tilde{z}_t^2}{\sigma^2} \right).
$$

A.8 Proof of Proposition 5

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

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

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 $g$:

$$
\frac{dY_t}{Y_t} = \frac{dC_t}{C_t} = \frac{dA_t}{A_t} = \frac{dN_t}{N_t} = g dt.
$$
The variables $E_t, K_t/A_t, Y_t/A_t, \Gamma_t, Z_t, z_t, \kappa_t, r_f, t$, 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

$$\frac{dC_t}{C_t} = \psi(r_f - \delta)dt. \quad (137)$$

Substituting equation (136) into (137), (51), and (135), we obtain

$$r_f = \frac{g}{\psi} + \delta, \quad (138)$$

$$g = \chi(h^b)\frac{1}{\psi} - \delta_b, \quad (139)$$

$$\frac{Y_t}{A_t} = \frac{g + \rho + \delta_a - r_f + (r_f + \delta_k)K_t/A_t}{\alpha(1 - \epsilon)}. \quad (140)$$

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

$$E = \left[1 - \frac{1}{\epsilon \nu} Z \frac{Y_t}{A_t} \left(\frac{K_t}{A_t}\right)^{1-\alpha}\right]^{1-\alpha}. \quad (141)$$

The flow profit $\pi$ to each intermediate-goods producer is a constant and given by equation (45),

$$\pi = (1 - \nu)\epsilon \frac{Y_t}{N_t} = (1 - \nu)\epsilon \frac{1}{E} \frac{Y_t}{A_t}. \quad (142)$$

Substituting equation (142) into (8), we obtain the value of blue prints $v$

$$v = \frac{\pi}{r_f + \delta_b}. \quad (143)$$

Substituting equations (138), (141), and (142) into (143), we obtain $v$

$$v = (1 - \nu)\epsilon \frac{\epsilon \nu^{1-\alpha}}{g/\psi + \delta + \delta_b} \left(\frac{A_t}{Y_t} \frac{K_t}{A_t}\right)^{\alpha} \frac{1}{Z^{1-\alpha}}. \quad (144)$$

Define $\Gamma_t \equiv \text{Cov}(\tilde{a}_{i,t}, \tilde{z}_{i,t}) = -M_t\text{var}(\tilde{z}_{i,t}) = -M_t\sigma^2/2$. The steady-state value of $Z$ is given by equation (48), as follows:

$$Z = \left[(1 + \lambda) \frac{A_t}{K_t} \exp \left(\Gamma + \frac{\sigma^2}{4}\right) \Phi \left(\Phi^{-1} \left(1 - \frac{1}{\lambda} \frac{K_t}{A_t}\right) + \frac{\sigma}{\sqrt{2}}\right)\right]^\alpha, \quad (145)$$

where the covariance $\Gamma$ in the balanced growth path is obtained by setting $dM_t = 0$ in
equation (52):
\[ \Gamma = \frac{\text{Cov}(\tilde{z}_{i,t}, d\tilde{a}_{i,t})}{\theta dt}. \] (146)

In the balanced growth path, \( \text{Cov}(\tilde{z}_{i,t}, d\tilde{a}_{i,t}) \) given by equation (134) becomes
\[ \text{Cov}(\tilde{z}_{i,t}, d\tilde{a}_{i,t}) = \frac{(1 + \lambda)\sigma^2\kappa}{2} \exp\left(\frac{\sigma^2}{4}\right) \Phi\left(\frac{\sigma^2/2 - \tilde{z}}{\sigma/\sqrt{2}}\right) dt \]
\[ + \frac{(1 + \lambda)\sigma}{2\sqrt{\pi}}(z\kappa - r_f - \delta_k) \exp\left(-\frac{\tilde{z}^2}{\sigma^2}\right) dt. \] (147)

Substituting equation (147) into (146):
\[ \Gamma = \frac{(1 + \lambda)\sigma^2\kappa}{2\theta} \exp\left(\frac{\sigma^2}{4}\right) \Phi\left(\frac{\sigma^2/2 - \ln \tilde{z}}{\sigma/\sqrt{2}}\right) + \frac{(1 + \lambda)\sigma}{2\theta \sqrt{\pi}}(z\kappa - r_f - \delta_k) \exp\left(-\frac{\tilde{z}^2}{\sigma^2}\right). \] (148)

When solving above equations, we need to know \( K_t/A_t, \tilde{z}, \) and \( \kappa. \) They are given by equation (115), (28, setting \( \sigma_k = 0), \) and (91), as follows
\[ \tilde{z} = \exp\left(\Gamma - \Phi^{-1}\left(\frac{1}{1 + \lambda} \frac{K_t}{A_t}\right) \frac{\sigma}{\sqrt{2}}\right), \] (149)
\[ z\kappa = r_f + \delta_k, \] (150)
\[ \kappa = \alpha(1 - \epsilon)Z^{-\frac{1}{2}}Y_t \frac{A_t}{A_t K_t}. \] (151)

B Discussions on the Analytical Approximation

In this appendix, we provide more discussions to justify the approximation of capital share \( \omega_t(z) \) in Section 2.7 of the main text. In Section B.1, we show that the distribution of \( \ln a_{i,t} \) can be approximated by a normal distribution, which justifies our Lemma 2 in the main text. In Section B.2, we evaluate the accuracy of our analytical approximation. We show that our approximation of \( \omega_t(z) \) can yield solutions similar to the numerical solutions of various variables in both steady states and transitions for a large range of parameter values.

B.1 Theoretical Property on the Approximation of \( \omega_t(z) \)

We prove Lemma 2, that is, the actual distribution of \( \tilde{a}_{i,t} \equiv \ln a_{i,t} \) is approximately normal. In the absence of aggregate shocks, consider the balanced growth path with \( T \approx \infty. \)
Thus, all equilibrium prices are constant as shown in the proof of Proposition 5. The productivity cutoff $z$ determined by equation (72) becomes:

$$z\kappa = r_f + \delta_k.$$  \hfill (152)

Rewriting equations (3) and (30) using (152) as follows:

$$\frac{da_{i,t}}{dt} = s(z_{i,t})a_{i,t},$$  \hfill (153)

where

$$s(z) = (1 + \lambda)\kappa \max\{z - \bar{z}, 0\} + r_f - \rho - \delta_a,$$  \hfill (154)

and $\kappa$ is given by equation (151). To better illustrate intuitions, we rewrite equation (153) in discrete time with a time interval $\Delta t \approx 0$:

$$a_{i,t+\Delta t} = \left[1 + s(z_{i,t})\Delta t\right]a_{i,t}. \hfill (155)$$

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

$$a_{i,n+1} = \left[1 + s(z_{i,n})\Delta t\right]a_{i,n}. \hfill (156)$$

Define $\bar{\zeta}_{i,n} \equiv \ln(1 + s(z_{i,n})\Delta t) - \bar{\xi}$ with $\bar{\xi} \equiv \mathbb{E} \left[\ln(1 + s(z_{i,n})\Delta t)\right]$, thus equation (156) can be written as

$$\ln a_{i,n+1} = \ln a_{i,n} + \bar{\xi} + \bar{\zeta}_{i,n}. \hfill (157)$$

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 (157) implies

$$\ln a_{i,t} = \ln a_{i,1} + (N_T - 1)\bar{\xi} + \sum_{n=1}^{N_T-1} \bar{\zeta}_{i,n}. \hfill (158)$$

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

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

$$\ln z_{i,n+1} = e^{-\theta\Delta t} \ln z_{i,n} + \sigma\Delta \varepsilon_{i,n+1}, \hfill (159)$$

OA-16
where $\varepsilon_{i,n+1}$ is standard normal variable and

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

(160)

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

$$
\sigma_N^2 = \mathbb{E} \left[ \sigma_{i,1}^2 \right] + 2 \sum_{n=1}^{N_T-1} \left( 1 - \frac{n}{N_T} \right) \mathbb{E} \left[ \bar{\zeta}_{i,1} \bar{\xi}_{i,n} \right].
$$

(161)

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

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

(162)

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

### B.2 Assessment of Approximation Errors

Without our approximation of $\omega_t(z)$, it is impossible to “exactly” solve the model numerically because $\omega_t(z)$ is an infinite dimensional object and the model is written in general equilibrium. Thus, to assess the accuracy of our analytical approximation of $\omega_t(z)$, we focus on the case without aggregate shocks (i.e., set $dW_t \equiv 0$). Our assessment considers the model’s approximation in both steady states and transitions, which presumably reflect the goodness of fit of our approximation “in response to shocks.”

In particular, we compare our analytical approximation with the numerical solutions of $\omega_t(z)$ (obtained by solving ordinary differential equations, ODE, and partial differential equations, PDE) in both steady states and transitions, respectively. We find that our approximation of $\omega_t(z)$ can closely capture the capital share in steady states as well as the economy’s transitional dynamics when the process (5) of idiosyncratic productivity is not very persistent. Specifically, our analytical approximation is very close to the numerical solutions obtained by solving the whole distribution of $\omega_t(z)$ using a standard ODE/PDE method for empirically relevant values of the persistence of idiosyncratic productivity, i.e., $\exp(-\theta) \leq 0.95$ according to the estimate of Asker, Collard-Wexler and Loecker (2014) based on U.S. census data.

We start by evaluating the accuracy of the capital share, $\omega(z)$, in the steady state for
Note: This figure compares the steady-state capital share $\omega(z)$ between our approximation and the ODE solution under different yearly autocorrelation in idiosyncratic productivity $z_{i,t}$. The ODE solution is obtained by directly solving the steady-state ODE characterizing $\omega(z)$. The analytical approximation is given by equation (47) of the main text. The blue solid line in each panel plots the ODE solution of $\omega(z)$ and the black dashed line plots the solution of $\omega(z)$ under our approximation.

Figure OA.1: Capital shares and autocorrelation.

different values of the persistence parameter $\theta$ in equation (5). The yearly autocorrelation in $\ln z_{i,t}$ is given by $\text{corr}(\ln z_{i,t}, \ln z_{i,t+1}) = \exp(-\theta)$. In Figure OA.1, the blue solid line plots the ODE solution of $\omega(z)$. The black dashed line plots the approximated $\omega(z)$ given by equation (47). Panels A, B, and C show that our approximation of $\omega(z)$ is very close to the ODE solution when the yearly autocorrelation in productivity is below 0.95. Our approximation also implies that the capital share becomes more right skewed when the persistence of idiosyncratic productivity increases, which is consistent with the ODE solution. Panel D shows that the approximation becomes worse when the yearly autocorrelation is 0.99. This is because when firms' productivity becomes more
Note: This figure compares the transitional dynamics of productivity, aggregate capital (normalized by trend growth), and aggregate output (normalized by trend growth) between our approximation and the PDE solutions under different yearly autocorrelation in idiosyncratic productivity $z_i$. The PDE solutions of transitional dynamics (blue solid lines) are obtained by directly solving the PDEs that characterize the evolution of $\omega_t(z)$ using the finite-difference algorithm. The approximated transitional dynamics (black dashed lines) are solved based on the evolution of $M_t$ (equation (52)). Initial capital shares are given by equation (47) with $\text{Cov}(\tilde{z}_i,0,\tilde{a}_i,0) = -0.5$.

Figure OA.2: Transition dynamics from a distorted initial capital distribution.

persistent, the approximation error characterized by equation (162) will be larger due to serial dependence. Thus, our approximation works best for when the persistence of idiosyncratic productivity is not extremely high.

The steady-state capital share summarizes the cross-sectional distribution of firms, which determines the steady-state productivity $Z_t$, aggregate output $Y_t$, and aggregate capital $K_t$. Given that our approximation can reasonably capture the capital share, it is straightforward to verify that the steady-state productivity, aggregate output, and aggregate capital under our approximation are also close to their ODE solutions.

Next, we check whether our approximation can generate reasonably accurate tran-
sitional dynamics. We follow the exercise of Moll (2014) by starting from a distorted initial allocation, in which firms’ capital and productivity are negatively correlated, i.e., \( \text{Cov}(\tilde{z}_{i,0}, \tilde{a}_{i,0}) = -0.5 \). Figure OA.2 compares the transitional dynamics of productivity, aggregate capital (normalized by trend growth), and aggregate output (normalized by trend growth) between our approximation and the PDE solutions for different values of yearly autocorrelation in idiosyncratic productivity \( z_{i,t} \). Panels A to C show that when the yearly autocorrelation is zero, our approximated transitional dynamics are almost identical to the transitional dynamics solved by the PDE method. Specifically, our approximation can capture the immediate jump in productivity and aggregate output at \( t = 0 \). Panels D to F show that our approximated transitional dynamics are also very close to the transitional dynamics solved by the PDE method when yearly autocorrelation is 0.85. In panels G to I, we consider the case with yearly autocorrelation equal to 0.97. It is shown that even with very persistent process of idiosyncratic productivity, our approximation is reasonably close to the transitional dynamics solved by the PDE method.

C TFP Formulas

We show that our formula for the final goods sector’s productivity \( Z_t \) is consistent with the formula of Hsieh and Klenow (2009) when goods are homogeneous and the industry is not distorted by wedges.

The final-goods sector’s productivity \( Z_t \) given by equation (41) of the main text is equivalent to (77) with \( u_t(z) \) being set at its optimal value, 1:

\[
Z_t = \left[ \frac{1}{K_t} \int_{0}^{\infty} \int_{0}^{\infty} z k_t(a, z) \varphi_t(a, z) dk dz \right]^\alpha, \tag{163}
\]

The above equation is equivalent to

\[
Z_t = \left[ \frac{1}{K_t} \int_{0}^{\infty} \int_{0}^{\infty} z k_t(a, z) \varphi_t(a, z) dk dz \right]^\alpha, \tag{164}
\]

because \( k_t(a, z) = 0 \) for \( z \leq z_t \) according to equation (23). Without loss of generality, we rewrite equation (164) to focus on a countable number of firms,

\[
Z_t = \left( \frac{1}{K_t} \sum_i z_i k_i \right)^\alpha, \tag{165}
\]

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) is $z_t^\alpha$ not $z_t$), equation (165) becomes

$$Z_t = \left(\frac{1}{K_t} \sum_i z_t^{1/\alpha_k} k_i\right)^{\alpha},$$

(166)

Next, we show that equation (166) 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

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

(167)

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., $L_s = 1$. 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 (167),

$$TFP_s = \frac{1}{K_s^{\alpha_s}} \left( \sum_{i=1}^{M_s} Y_s^{1/\sigma} \right)^{\frac{\sigma}{\sigma-1}}.$$  

(168)

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

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

(169)

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

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

(170)

Substituting equation (170) into (169), we obtain

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

(171)
The labor market clearing condition in the model of Hsieh and Klenow (2009) implies

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

Substituting (172) into (171),

$$TFP_s = \frac{1}{K_s^{\alpha_s}} \left[ \sum_{i=1}^{M_s} \left[ \frac{A_{si}^{1/\alpha_s} K_{si} \left[ (1 - \tau_{Ysi})P_{si} \right]^{1-\alpha_s}}{\sigma} \right]^{\frac{\sigma-1}{\sigma-1}} \sigma^\frac{\sigma-1}{\sigma-1} \right](173)$$

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 (173) as follows,

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

(174)

Except for notational differences, the formula (174) is identical to (166).

D Numerical Algorithm

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

We use superscripts $+$ and $-$ to denote variables at $t + \Delta t$, corresponding to $dW_t = \sqrt{\Delta t}$ and $dW_t = -\sqrt{\Delta t}$, respectively. The endogenous state variable $\Gamma_t$ evolves according to equation (125):

$$\Gamma_{t+\Delta t} = \Gamma_t - \theta \Gamma_t dt + \text{Cov}(\tilde{z}_{i,t}, d\tilde{a}_{i,t}),$$

(175)

where $\text{Cov}(\tilde{z}_{i,t}, d\tilde{a}_{i,t})$ is given by equation (134), as follows:

$$\text{Cov}(\tilde{z}_{i,t}, d\tilde{a}_{i,t}) = \frac{(1 + \lambda) \sigma^2 \kappa_t}{2} \exp \left( \frac{\sigma^2}{4} \right) \Phi \left( \frac{\sigma^2 / 2 - \tilde{z}_t}{\sigma / \sqrt{2}} \right) \Delta t$$

$$+ \frac{(1 + \lambda) \sigma}{2 \sqrt{\pi}} \left[ (z_{it} h_t - r_{f,t} - \delta_k) \Delta t - \sigma_k dW_t \right] \exp \left( -\frac{\tilde{z}_t^2}{\sigma^2} \right).$$

(176)
Let $\Gamma_{t+\Delta t}$ and $\Gamma_{t-\Delta t}$ be the value of $\Gamma_{t+\Delta t}$ corresponding to $dW_t = \sqrt{\Delta t}$ and $dW_t = -\sqrt{\Delta t}$, respectively. In equation (176), the variables $\kappa_t$, $\tilde{z}_t$, and $r_{f,t}$ are given by equations (91), (115), and the SDF, respectively, as follows:

\[
\kappa_t = \alpha(1-\varepsilon)Z_t^{-\frac{1}{2}} \frac{Y_t A_t}{A_t K_t},
\]

\[
\tilde{z}_t = \Gamma_t - \Phi^{-1}\left(\frac{1}{1+\lambda} \frac{K_t}{A_t}\right) \frac{\sigma}{\sqrt{2}},
\]

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

where $Y_t/A_t$, $Z_t$, and $K_t/A_t$ are functions of state variables $E_t$ and $\Gamma_t$, given by equations (89), (118), and (28), respectively, as follows:

\[
\frac{Y_t}{A_t} = (\varepsilon v)^{e \tau} Z_t E_t^{1-a} \left(\frac{K_t}{A_t}\right)^a,
\]

\[
Z_t = \left[(1+\lambda) \frac{A_t}{K_t} \exp\left(\Gamma_t + \frac{\sigma^2}{4}\right) \Phi\left(\Phi^{-1}\left(\frac{1}{1+\lambda} \frac{K_t}{A_t}\right) + \frac{\sigma}{\sqrt{2}}\right)\right]^a.
\]

Let $E_{t+\Delta t}^+$ and $E_{t-\Delta t}^-$ be the value of $E_{t+\Delta t}$ corresponding to $dW_t = \sqrt{\Delta t}$ and $dW_t = -\sqrt{\Delta t}$, respectively.

In equation (184), the variable $v_t = v(E_t, \Gamma_t)$ is given by equation (8); it is a function of
state variables \((E_t, \Gamma_t)\) and can be solved recursively as follows

\[
v(E_t, \Gamma_t) = \frac{1}{1 + \delta_b \Delta t} \left( \pi_t \Delta t + \mathbb{E}_t \left[ \frac{\Lambda_{t+\Delta t}^+}{\Lambda_t} v(E_{t+\Delta t}, \Gamma_{t+\Delta t}) \right] \right) = \frac{1}{1 + \delta_b \Delta t} \left( \pi_t \Delta t + \frac{1}{2} \frac{\Lambda_{t+\Delta t}^+}{\Lambda_t} v(E_{t+\Delta t}, \Gamma_{t+\Delta t}) + \frac{1}{2} \frac{\Lambda_{t+\Delta t}^-}{\Lambda_t} v(E_{t+\Delta t}, \Gamma_{t+\Delta t}) \right), \tag{185}
\]

where \(\pi_t\) is given by equation (93):

\[
\pi_t = \frac{(1 - \nu) e Y_t}{E_t A_t}. \tag{186}
\]

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

\[
\frac{\Lambda_{t+\Delta t}}{\Lambda_t} = e^{-\delta_{t+\Delta t}} \left( \frac{C_{t+\Delta t}}{C_t} \right)^{1-\gamma} \left( 1 + R_{w,t+\Delta t} \right) \frac{1}{1-1/\psi}, \tag{187}
\]

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

\[
1 + R_{w,t+\Delta t} \Delta t = \frac{W_{t+\Delta t}}{W_t - C_t \Delta t}. \tag{188}
\]

We have

\[
\mathbb{E}_t \left[ \frac{\Lambda_{t+\Delta t}}{\Lambda_t} (1 + R_{w,t+\Delta t}) \right] = 1. \tag{189}
\]

Substituting equations (187) and (188) into (189), we obtain

\[
1 = \mathbb{E}_t \left[ e^{-\delta_{t+\Delta t}} \left( \frac{C_{t+\Delta t}}{C_t} \right)^{1-\gamma} \left( \frac{W_{t+\Delta t} C_{t+\Delta t}}{W_t / C_t - \Delta t} \right)^{1-\gamma} \right]. \tag{190}
\]

Rearranging the above equation, we obtain

\[
\frac{W_t}{C_t} = \Delta t + e^{-\delta \Delta t} \mathbb{E}_t \left[ \left( \frac{C_{t+\Delta t}}{C_t} \right)^{1-\gamma} \left( \frac{W_{t+\Delta t} C_{t+\Delta t}}{C_{t+\Delta t}} \right)^{1-\gamma} \right]. \tag{191}
\]

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 \(dW_t = \sqrt{\Delta t}\) and
\[ dW_t = -\sqrt{\Delta t}, \] respectively. We can rewrite equation (191) as
\[
WC_t = \Delta t + e^{-\delta \Delta t} \left[ \frac{1}{2} \left( \frac{C_t^{+\Delta t}}{C_t} \right)^{1-\gamma} (WC_t^{+\Delta t})^{1-\gamma} \varpi + \frac{1}{2} \left( \frac{C_t^{-\Delta t}}{C_t} \right)^{1-\gamma} (WC_t^{-\Delta t})^{1-\gamma} \right],
\]
(192)
where
\[
WC_t^{+\Delta t} = WC(E_t^{+\Delta t}, \Gamma_t^{+\Delta t}),
\]
(193)
\[
WC_t^{-\Delta t} = WC(E_t^{-\Delta t}, \Gamma_t^{-\Delta t}).
\]
(194)
The aggregate consumption is given by equation (14):
\[
\frac{C_t}{A_t} = \frac{w_t}{A_t} + \frac{D_t}{A_t} + r_{f,t} \frac{B_t}{A_t} - \left( \frac{B_{t+\Delta t}}{A_{t+\Delta t}} \frac{A_{t+\Delta t}}{A_t} - \frac{B_t}{A_t} \right) \frac{1}{\Delta t}
\]
\[= \frac{w_t}{A_t} + \frac{D_t}{A_t} + r_{f,t} \left( \frac{K_t}{A_t} - 1 \right) - \left[ \left( \frac{K_{t+\Delta t}}{A_{t+\Delta t}} - 1 \right) \frac{A_{t+\Delta t}}{A_t} - \left( \frac{K_t}{A_t} - 1 \right) \right] \frac{1}{\Delta t}. \]
(195)
Because \( C_t \) is known (i.e., \( dB_t / B_t \) is locally deterministic), theoretically we have
\[
\left( \frac{K_t^{+\Delta t}}{A_t} - 1 \right) \frac{A_t^{+\Delta t}}{A_t} = \left( \frac{K_t^{-\Delta t}}{A_t} - 1 \right) \frac{A_t^{-\Delta t}}{A_t},
\]
(196)
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 \( dW_t = \sqrt{\Delta t} \) and \( dW_t = -\sqrt{\Delta t} \), respectively. Because of property (196), the numerical error caused by discretization is minimized by using
\[
0.5 \left( \frac{K_t^{+\Delta t}}{A_t} - 1 \right) \frac{A_t^{+\Delta t}}{A_t} + 0.5 \left( \frac{K_t^{-\Delta t}}{A_t} - 1 \right) \frac{A_t^{-\Delta t}}{A_t}
\]
to approximate \( \left( \frac{K_t^{+\Delta t}}{A_t} - 1 \right) \frac{A_t^{+\Delta t}}{A_t} \) in equation (195). Thus, the term \( C_t / A_t \equiv CA(E_t, \Gamma_t) \) in equation (195) can be solved as a function of state variables \( E_t \) and \( \Gamma_t \).

The consumption growth terms in equation (192) are given by
\[
\frac{C_t^{+\Delta t}}{C_t} = \frac{CA(E_t^{+\Delta t}, \Gamma_t^{+\Delta t})}{CA(E_t, \Gamma_t)} \frac{A_{t+\Delta t}}{A_t},
\]
(197)
\[
\frac{C_t^{-\Delta t}}{C_t} = \frac{CA(E_t^{-\Delta t}, \Gamma_t^{-\Delta t})}{CA(E_t, \Gamma_t)} \frac{A_{t+\Delta t}}{A_t}. \]
(198)
The variables \( w_t/A_t \) and \( D_t/A_t \) are given by equations (44) and (31):

\[
\frac{w_t}{A_t} \equiv wA(E_t, \Gamma_t) = (1 - \alpha)(1 - \varepsilon) \frac{Y_t}{A_t},
\]

(199)

\[
\frac{D_t}{A_t} \equiv DA(E_t, \Gamma_t) = \rho + (1 - \nu)\varepsilon \frac{Y_t}{A_t} - \frac{S_t}{A_t},
\]

(200)

where \( S_t/A_t \) is given by equation (46)

\[
\frac{S_t}{A_t} = \left( \frac{S_t}{N_t} \right) E_t = \left( \chi v(E_t, \Gamma_t) \right)^{\frac{1}{\psi}} E_t.
\]

(201)

The variables \( A_t + \Delta t/A_t \) is given by equation (121):

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

(202)

After solving the WC\((E_t, \Gamma_t)\) ratio from equation (192), substituting into the equation (187) to obtain the SDF:

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

(203)

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

(204)

Welfare. The preference (12) in discrete time is

\[
U_t = \left( 1 - e^{-\delta \Delta t} \right) C_t^{1-1/\psi} + e^{-\delta \Delta t} \left( E_t \left[ \left( U_{t+\Delta t} \right)^{1-\gamma} \right] \right)^{1-1/\psi} \left( 1-1/\psi \right).
\]

(205)

Dividing both sides by \( A_t \) and define \( \tilde{U}_t = (U_t/A_t)^{1-1/\psi} \), we obtain

\[
\tilde{U}_t = \left( 1 - e^{-\delta \Delta t} \right) \left( \frac{C_t}{A_t} \right)^{1-1/\psi} + e^{-\delta \Delta t} \left( E_t \left[ \left( \frac{A_{t+\Delta t}}{A_t} \right)^{1-\gamma} \tilde{U}_{t+\Delta t}^{1-1/\psi} \right] \right)^{1-1/\psi} \left( 1-1/\psi \right).
\]

(206)

We obtain \( \tilde{U}_t \) by solving the above equation. We then obtain \( U_t = \tilde{U}_t^{1-1/\psi} A_t \), which is homogeneous of degree one in \( C_t \).
Steps of Implementing the Numerical Algorithm. Following the standard practice, we discretize the state variables \((E_t, \Gamma_t)\) into dense grids. The values that not fall on any grid are obtained by linear interpolation/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*.\(^{12}\) 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 \(\sigma_{\xi}(z_t, E_t, \Gamma_t) = 0\) 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 \(\Gamma_t\).
5. Solve equation (191) using *knitro* to obtain the wealth-consumption ratio as a function of state variables, i.e., \(WC(E_t, \Gamma_t)\).
6. Solve equations (203) and (204) to obtain the SDF as a function of state variables, i.e.,

\[
\frac{\Lambda^+_{t+\Delta t}}{\Lambda_t} \equiv SDF(E^+_{t+\Delta t}, \Gamma^+_{t+\Delta t}),
\]

(207)

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

(208)

Next, calculate the market price of risk \(\eta_t\) in equation (20) as follows

\[
\hat{\eta}(E_t, \Gamma_t) = -\frac{SDF(E^+_{t+\Delta t}, \Gamma^+_{t+\Delta t}) - SDF(E^-_{t+\Delta t}, \Gamma^-_{t+\Delta t})}{2\sqrt{\Delta t}}.
\]

(209)

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

7. Solve managers’ problem (18) to obtain \(\sigma_{\xi}(z_t, E_t, \Gamma_t)\). This is achieved in the following substeps.

7.1 Problem (18) can be simplified because it is linear in \(a_{i,t}\) as in equation (64).

This means that we only need to solve \(\xi(z_{i,t}, E_t, \Gamma_t)\) recursively as follows

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

(210)

\(^{12}\)See https://www.artelys.com/solvers/knitro for more details.
The evolution $a_{i,t+\Delta t}/a_{i,t}$ is given by equation (3).

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

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

(211)

where

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

(212)

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

(213)

The expectation is taken with respect to idiosyncratic shocks in $z_{i,t+\Delta t}$.

(7.3) Solve $\xi(E_t, \Gamma_t)$ using equation (182), and then find the value of $\hat{\sigma}_\xi(z_{i,t}, E_t, \Gamma_t)$.

(7.4) If $\max |\hat{\sigma}_\xi(\xi_{i,t}, E_t, \Gamma_t) - \sigma_\xi(\xi_{i,t}, E_t, \Gamma_t)| < 10^{-9}$, stop. Otherwise, jump to step (3) using $\hat{\sigma}_\xi(z_{i,t}, E_t, \Gamma_t)$ as the initial guess for $\sigma_\xi(\xi_{i,t}, E_t, \Gamma_t)$.

(8) Solve equation (185) to obtain $\hat{v}(E_t, \Gamma_t)$.

(9) If $\max |\hat{v}(E_t, \Gamma_t) - v(E_t, \Gamma_t)| < 10^{-9}$, stop. Otherwise, jump to step (2) using $\hat{v}(E_t, \Gamma_t)$ as the initial guess for $v(E_t, \Gamma_t)$. 

OA-28