Equilibrium Value and Profitability Premiums

Hengjie Ai, Jun E. Li and Jincheng Tong *

September 2026

Abstract

Standard production-based asset pricing models cannot simultaneously explain the value premium and the gross profitability premium. Empirically, we document that growth firms differ from value firms mainly in the persistence of their cash flows and in their investment rates, whereas high-profitability firms differ from low-profitability firms mainly in the level of their current cash flows, suggesting that the two sorts capture productivity components of different persistence. A general equilibrium model in which firm-level productivity has such a two-factor structure accounts for both the gross profitability premium and the value premium. Variable capital utilization makes high-profitability firms more exposed to depreciation risk, generating the profitability premium.

JEL Code: E2, E3, G12

*Hengjie Ai (hengjie.ai@wisc.edu) is at the University of Wisconsin-Madison, Jun E. Li (jun.li@wbs.ac.uk) is affiliated with Warwick Business School, and Jincheng Tong (jincheng.tong@rotman.utoronto.ca) is at the University of Toronto. We thank Winston Dou, Mete Kilic, Ding Luo, Dimitris Papanikolaou, Lukas Schmid, and participants at WFA, MFA, China International Conference in Macroeconomics, FIRS, and University of North Carolina at Chapel Hill for helpful comments.

1 Introduction

Historically, firms with higher profitability ratios, defined as the ratio of gross profit to total assets, earn higher average returns than those with lower profitability ratios. This phenomenon is often referred to as the gross profitability premium (Novy-Marx (2013)). At the same time, stocks with higher book-to-market ratios, or value firms, earn higher average returns than those with lower book-to-market ratios (growth firms). The difference in average returns between value and growth firms is known as the value premium.

The coexistence of the value and profitability premiums presents a challenge to the literature on Q-theory and production-based asset pricing models, for example, Zhang (2005). In these models, productivity is typically specified as an AR(1) process and is the only source of exogenous variation that determines the cross-sectional heterogeneity in firms' market-to-book and profitability ratios. High-productivity firms have high profitability ratios, but they are also growth firms with higher market-to-book ratios. As a result, traditional production-based asset pricing models cannot separate the value and profitability factors. By construction, a profitability premium must imply a value discount. That is, these models can generate either a value premium or a profitability premium, but not both.

In this paper, we develop a general equilibrium model where firm-level productivity follows a two-factor structure with components differing in persistence. We show that our model can not only distinguish the profitability factor from the value factor but also account for the coexistence of the value and profitability premiums.

To distinguish profitability from value empirically, we begin by presenting suggestive evidence for a two-factor structure in firm-level productivity. Consistent with the standard production-based asset pricing literature, we find that both high-profitability and growth firms exhibit higher investment rates and higher cash flows at the time of portfolio formation compared to low-profitability and value firms. More importantly, however, the differences in investment rates are significantly larger for portfolios sorted by the market-to-book ratio than for those sorted by profitability. Additionally, differences in cash flows are substantially more persistent across market-to-book sorted portfolios than profitability-sorted portfolios. These patterns support our hypothesis that shocks with different persistence levels separately drive value and profitability effects: market-to-book ratios capture the permanent productivity component, which strongly affects firms' future growth, whereas profitability measures primarily reflect the transitory component, which mainly affects firms' current profits.

Motivated by the empirical evidence, we develop a general equilibrium model in which firm-level productivity consists of both permanent and transitory components. The transitory component exhibits high conditional volatility but low persistence. In contrast, the

permanent component follows a random walk with low conditional volatility. This structure allows our model to distinguish between the profitability and value factors. A transitory productivity shock raises a firm's contemporaneous profitability but has a limited effect on its investment and market-to-book ratio, because of its lack of persistence. Instead, a permanent productivity shock, though typically smaller in magnitude, has a persistent effect on firms' investment decisions and valuations, but a small impact on current profitability. As a result, from a quantitative perspective, growth firms differ from value firms mainly because they have a higher permanent component of productivity, while gross profitability sorting ranks firms based primarily on the transitory component.

We incorporate two sources of aggregate shocks into our model: total factor productivity (TFP) shocks and shocks to depreciation. The key mechanism in our model that accounts for the profitability premium is variable capital utilization: firms can choose to increase the rate of capital utilization and raise current output at the expense of losing productive capital through accelerated depreciation. In the presence of accelerated capital depreciation resulting from intensive utilization, high-profitability and growth firms respond differently. High-profitability firms, recognizing the transitory nature of their shocks, adjust primarily through capital utilization rather than investment. In contrast, growth firms respond to permanent productivity shocks by raising investment and reducing utilization, because they anticipate that the impact of permanent shocks will persist into the future. In our model, the accelerated capital depreciation through utilization acts as a form of operating leverage, making high-profitability firms more exposed to depreciation risk.

Our model also generates a value premium. Firms with lower permanent components of productivity shocks are identified as value firms. TFP risk, combined with fixed operating costs and convex capital adjustment costs, provides the operating leverage channel that leads to the value premium. This channel is consistent with Zhang (2005). The operating leverage effect amplifies the riskiness of value firms compared to growth firms, where adjustment costs represent a smaller portion of total cash flows.

In summary, our model establishes an endogenous link between value and profitability sorting and the permanent and transitory components of idiosyncratic productivity shocks. This connection is a result of firms' optimal investment and capital utilization decisions in response to shocks with different persistence. At the aggregate level, this mechanism creates distinct risk exposures: value firms become more sensitive to aggregate productivity shocks, while high-profitability firms are more negatively exposed to depreciation shocks, offering an explanation for the coexistence of the value and profitability premiums.

The mechanisms described above emerge endogenously within our heterogeneous-firm general equilibrium framework. Solving such a model is computationally challenging due to

the high-dimensional state space that includes the entire cross-sectional distribution of firms. To address this challenge, we develop an efficient global method that exploits the homogeneity properties of firms' decision problems. We then estimate the model using simulated method of moments. The estimated model successfully captures the differences in investment rates and cash flow persistence across portfolios sorted by market-to-book and profitability. In addition, it generates both the value and profitability premiums, and a negative correlation between the two factors, while fitting aggregate quantities and asset prices reasonably well.

To test the model's implications, we construct an empirical measure of depreciation shocks from abnormal capital losses associated with catastrophic events. The measure delivers evidence consistent with the model's mechanism for the profitability premium: high-profitability firms have more negative exposure to depreciation shocks, and depreciation risk carries a negative price in the cross-section of returns. In addition, the investment of high-profitability firms responds more negatively to depreciation shocks, consistent with their more intensive capital utilization.

Literature review

This paper builds on the large literature on Q-theory-based asset pricing models and the cross-section of equity returns. Zhang (2005) develops a Q-theory-based asset pricing model to explain the value premium. Belo, Lin, and Bazdresch (2014) and İmrohoroğlu and Tüzel (2014) use a similar approach to study the impact of other firm characteristics, such as labor hiring and firm-level productivity, on the cross-section of equity returns. Donangelo, Gourio, Kehrig, and Palacios (2019) examine the relation between firm-level labor share and the cross-section of equity returns. Herskovic, Kind, and Kung (2018) study the low-frequency comovement between value and size premiums using a standard Q-theory-based asset pricing model. Kuehn, Simutin, and Wang (2017) explore the connection between firms' exposures to labor market conditions and stock returns. Kuehn and Schmid (2014) extend the Q-theory asset pricing model to study the pricing of corporate bonds. Eisfeldt and Papanikolaou (2013) incorporate organizational capital into the Q-theory asset pricing paradigm and investigate its impact on firms' stock returns. As pointed out by Novy-Marx (2013), in the above Q-theory-based models, the capital adjustment costs serve as operating leverage and generate a value premium. However, because high-productivity firms are simultaneously high market-to-book and high-profitability ratio firms, these models would also imply a negative profitability premium. In addition, all of the above models are partial equilibrium models that take the pricing kernel as given, while our model is a general equilibrium one, where the pricing kernel

and firms' risk exposure are both endogenously determined in equilibrium.[1]

Several recent papers also study the coexistence of value and profitability premiums. Dou et al. (2021a) develop a model with dynamic strategic competition to explain the profitability premium. Both Dou et al. (2021b) and Kogan et al. (2023) develop equilibrium models where the value premium and the gross profitability premium coexist. Dou et al. (2021b) focus on how industry competition affects the heterogeneity of returns across industries but do not explain the within-industry value and profitability premiums. Like Kogan et al. (2023), our model generates distinct value and profitability factors. Their mechanism emphasizes variable inputs as an operating hedge, whereas ours links permanent and transitory firm-level productivity to endogenous investment and capital utilization.

Our paper also builds on the previous literature that develops tractable models to study the cross-section of expected returns in general equilibrium. Gomes et al. (2003) develop an analytically tractable framework where the cross-section of firm characteristics has a stationary distribution to study the cross-section of expected returns. Ai and Kiku (2013) develop a model with balanced growth and long-run risk and use the technique from the firm dynamics literature (Luttmer (2007)) to solve the aggregate problem. Several works in the production-based asset pricing literature use the Krusell and Smith (1998) method to numerically solve their models.[2] Our aggregation method is different from all of the above. We take advantage of the homogeneity in firms' decision problems and summarize the cross-sectional distribution as a low-dimensional distribution.[3] We exploit the analytical expression for the law of motion of this distribution to solve the model globally. Our methodology could be applied to a larger class of models.

Our paper is also related to the literature that emphasizes the importance of variable capital utilization in understanding business cycle dynamics and asset market valuations. Greenwood et al. (1988) and Jaimovich and Rebelo (2009) show that variable utilization is an important ingredient accounting for the positive correlation between consumption and investment. Garlappi and Song (2017) show that variable utilization plays an important role in understanding the pricing of investment-specific shocks. Grigoris and Segal (2024) find that low capital utilization is associated with high equity returns, and such a pattern is consistent with a standard Q-theory model with variable capital utilization.

[1] A different strand of production-based asset pricing models explores the implications of investment-specific shocks (IST), for example, Papanikolaou (2011), Kogan and Papanikolaou (2013), Kogan and Papanikolaou (2014), and Li (2018). Kogan and Papanikolaou (2012) provide an excellent survey of this literature.

[2] Recent works that approximate firm distribution using distributional moments include Favilukis and Lin (2015), Chen (2018), Bai et al. (2019), and Favilukis et al. (2020).

[3] Other papers that use the same technique to reduce the dimensionality of the firm distribution in general equilibrium setups include Ai and Bhandari (2021).

The rest of the paper is organized as follows. We first present the empirical evidence of a two-factor structure of firm-level productivity in Section 2. Section 3 develops a general equilibrium model that embeds a two-factor structure into firm productivity shocks. We study the asset pricing implications of the model in Section 4 and present our quantitative results in Section 5. Section 6 constructs an empirical measure of depreciation shocks and tests the model's implications for risk exposures, prices of risk, and firm investment. Section 7 concludes.

2 Empirical evidence

In this section, we replicate the profitability premium and its relationship with the value premium as documented by Novy-Marx (2013) and highlight the challenge these findings pose for standard production-based asset pricing models. We also present empirical evidence that motivates the development of a theoretical model featuring a two-factor structure of firm-level productivity.

2.1 Value and profitability premiums

Following Novy-Marx (2013), we define gross profitability as revenue minus the cost of goods sold, scaled by total assets. We construct the market-to-book ratio according to the methodology in Fama and French (2015) and employ standard procedures to sort firms into quintile portfolios based on their profitability and market-to-book ratios. Consistent with standard terminology in the literature, we refer to the portfolio with the lowest market-to-book ratio as the value portfolio and the one with the highest ratio as the growth portfolio.

Table 1 presents the median firm characteristics across these portfolios, including market-to-book ratios, gross profitability, and investment rates. Panel A reports statistics for market-to-book sorted portfolios, while Panel B shows the corresponding metrics for profitability-sorted portfolios. We also document the average excess returns. As documented in the literature (Novy-Marx (2013)), value and profitability premiums cannot be explained by CAPM. Detailed descriptions of variable construction are provided in Appendix A.1.

We observe two key patterns in the data. First, the market-to-book ratio and gross profitability show a strong positive correlation, yet they predict returns in opposite directions. In Panel A, portfolios sorted on market-to-book ratios display monotonically increasing gross profitability from value to growth firms. Similarly, Panel B of Table 1 shows that profitable firms maintain higher market-to-book ratios. Despite these similarities in characteristics, the return patterns diverge sharply: growth firms earn lower average returns than value firms,

Table 1: Firm characteristics and returns

This table compares the median firm characteristics of portfolios sorted by market-to-book (MB) and gross profitability (GP/A). Panels A and B report the firm characteristics of portfolios sorted by MB and GP/A, respectively. We also report the annualized excess returns of value-weighted portfolios $E[R^e]$. The BM sorted portfolios and factors are from Ken French's website. We annualize returns by multiplying by 12. Firm characteristics are at an annual frequency. Numbers in parentheses are Newey–West t-statistics adjusted using 3 lags. The sample period ranges from 1963 to 2024.

Panel A: Market-to-book-sorted portfolios

	Value	2	3	4	Growth	Growth-Value
MB	0.66	1.18	1.80	2.85	6.16	5.51
GP/A	0.22	0.27	0.31	0.35	0.37	0.15
I/K	0.07	0.09	0.11	0.13	0.15	0.08
$E[R^e](\%)$	10.44	8.07	7.48	7.04	7.14	-3.30
(t)	(4.22)	(3.78)	(3.88)	(3.50)	(3.25)	(-1.77)

Panel B: Gross-profitability-sorted portfolios

	Low	2	3	4	High	High-Low
MB	1.54	1.33	1.60	1.91	2.35	0.81
GP/A	0.04	0.18	0.31	0.44	0.69	0.65
I/K	0.08	0.09	0.10	0.11	0.12	0.04
$E[R^e](\%)$	4.57	5.62	7.47	7.04	9.43	4.86
(t)	(2.12)	(2.70)	(3.63)	(3.23)	(4.43)	(2.96)

while high-profitability firms earn higher average returns than low-profitability firms. This pattern poses a significant challenge for production-based asset pricing models, as noted by Novy-Marx (2013). In these frameworks, firm-level productivity shocks typically serve as the only source of exogenous variation driving firm characteristics. Growth firms, those experiencing positive productivity shocks, are also profitable firms. Consequently, models that generate a positive value premium would necessarily imply a negative profitability premium (as seen in Zhang (2005), Belo et al. (2014), İmrohoroğlu and Tüzel (2014), Favilukis and Lin (2015), Li (2018), and Gomes and Schmid (2021)).

Second, we find significant differences in investment behavior. Growth firms invest at rates approximately 8% higher than value firms, while profitable firms' investment rates exceed those of unprofitable firms by only 4%. This distinction aligns with our hypothesis of a two-factor productivity structure: firms with high market-to-book ratios experience positive, persistent productivity shocks. Given the persistence of their productivity advantage, these

firms optimally respond by investing aggressively. In contrast, sorting on gross profitability primarily identifies firms with transitory productivity shocks, leading these firms to increase investment only modestly in response to temporary productivity gains.

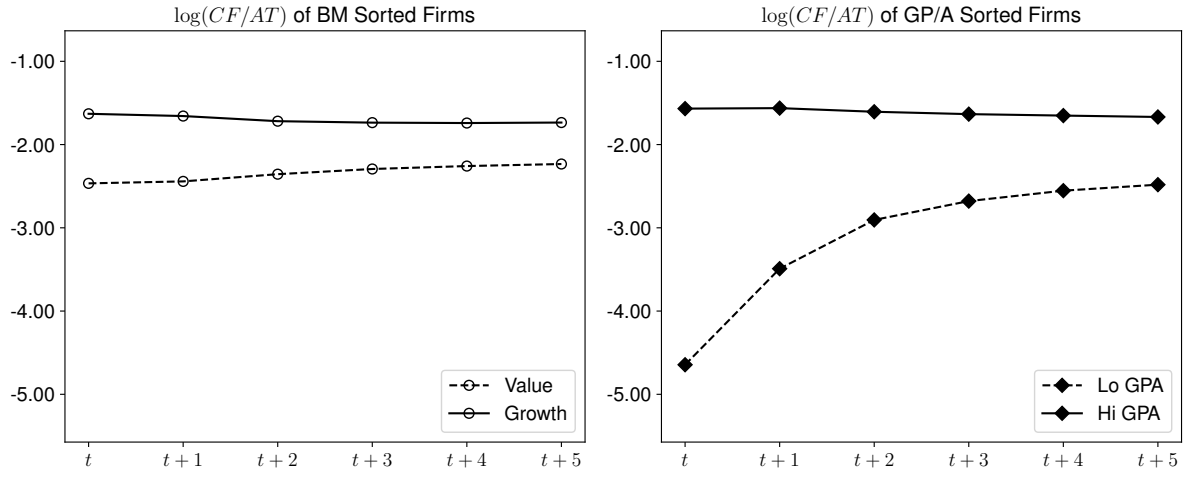
In the next section, we provide additional empirical support for the two-factor structure of firm productivity by examining post-formation patterns in cash flow and investment across portfolios sorted on market-to-book and profitability.

2.2 Cash flow after portfolio formation

Cash flow To examine the persistence of cash flows, we construct portfolios sorted on market-to-book ratios (MB) and profitability (GP/A), then follow these firms for five years after portfolio formation. Figure 1 presents our findings. The left panel tracks the log cash flow-to-asset ratio for value and growth portfolios, while the right panel shows the same for low and high-profitability portfolios. We use operating profit to measure firm cash flows, following Ball et al. (2016). We focus on the median of the cash flow-to-asset ratio in each portfolio.[4]

Figure 1: Cash flows after the portfolio formation year

This figure reports the median cash flows following the portfolio formation year. The horizontal axis indicates years since formation. We plot the log of the median cash-flow-to-asset ratios for portfolios sorted by MB and GP/A. The cash flow is defined as the operating profit, following Ball et al. (2016). The left and right panels display results for portfolios sorted on MB and GP/A, respectively. The sample period runs from 1963 to 2024, at an annual frequency.



[4]The empirically computed standard errors are very small because each portfolio contains around 500–800 firms per year and the estimates use the full sample over a sixty-year period. We therefore omit confidence intervals from the plots.

Figure 1 shows that growth firms consistently have higher cash flows than value firms. While the initial cash flow gap between the two groups is modest, it persists over time. Specifically, the log cash flow difference between growth and value firms is 0.84 at portfolio formation (year 0), decreases slightly to 0.78 after one year, and stands at 0.50 by year 5. This implies a "5-0 ratio", defined as the year-5 gap divided by the formation-year gap, of approximately 0.59.

To interpret this finding within a production-based asset pricing framework, consider a model in which firm output is given by $X^{1-\nu}K^{\nu}$, where X denotes firm-level productivity and K denotes the capital stock. In this setting, the log cash flow-to-asset ratio, $\ln\left(\frac{X^{1-\nu}K^{\nu}}{K}\right)$, corresponds to the log marginal product of capital, $(1-\nu)(\ln X - \ln K)$, which naturally converges toward equilibrium as capital productivity equalizes across firms. If $\ln X$ follows a unit root process, then convergence in the marginal product of capital stems primarily from capital dynamics. In a model with substantial capital adjustment costs and assuming a 10% depreciation rate, the autocorrelation of $\ln K$ is roughly 90% annually, compounding to about 60% over five years. This slow decay pattern aligns with predictions from neoclassical models incorporating capital adjustment costs and permanent productivity shocks.

The profitability-sorted portfolios (right panel) exhibit significantly different convergence dynamics. The initial log cash flow gap between high- and low-profitability portfolios is substantially larger, at around 3. This difference declines rapidly, falling to around 2 after just one year and further to around 0.8 by year 5. The resulting 5-0 CF difference ratio of approximately 0.3 indicates much faster mean reversion than that observed for the value-growth portfolios.

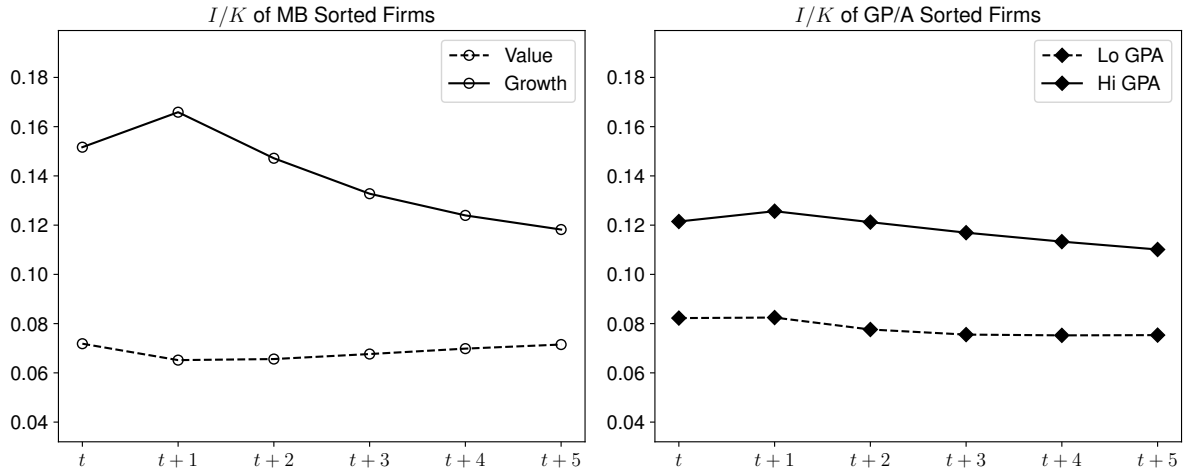
Investment In Figure 2, we plot the median investment rates following portfolio formation for the same market-to-book (left panel) and gross profitability (right panel) sorted portfolios. The median investment rate for a portfolio is defined as the median investment-to-capital ratio of firms within the portfolio.

While cash flow differences between portfolios are initially modest for market-to-book-sorted firms, we observe significant contrasts in investment behavior. At portfolio formation, growth firms invest at rates 8% higher than value firms, with this substantial gap persisting for several years. By comparison, high-profitability firms exceed low-profitability firms' investment rates by just 4% at the portfolio formation year.

This pattern reveals that growth firms consistently maintain higher investment, allowing them to build capital stocks and expand more rapidly than value firms. Our findings complement the work of Gourio (2008) and Byun et al. (2019), who document that firms adjust investment strongly in response to permanent productivity shocks but only modestly

Figure 2: Investment rates after the portfolio formation year

This figure presents the median investment rates after the portfolio formation year for MB and GP/A sorted portfolios. The horizontal axis indicates years since formation. The left and right panels report investment rates for portfolios sorted by MB and GP/A, respectively. The sample period runs from 1963 to 2024, at an annual frequency.



when facing transitory ones.

The distinct patterns of cash flows and investment rates documented above provide supporting evidence for our proposed two-factor productivity structure. In the following section, we develop a general equilibrium model that formalizes this intuition. Using simulated method of moments (SMM), we estimate the key productivity parameters and demonstrate that our model successfully replicates the empirical patterns in returns, cash flows, and investment observed across portfolios.

3 Model setup

Motivated by the empirical evidence in the last section, we build a general equilibrium model where firm-level productivity has a two-factor structure, which includes a transitory and a permanent component. Our model provides a unified framework to study the joint dynamics of macroeconomic quantities and asset prices. Our purpose is to demonstrate that the model provides a quantitative explanation for the coexistence of the value and profitability premiums.

Households Time is infinite and discrete. We assume that the representative household has recursive preferences with risk aversion γ and intertemporal elasticity of substitution

(IES) ψ, as in Epstein and Zin (1989):

$$U_t = \left\{ (1-\beta) C_t^{1-\frac{1}{\psi}} + \beta \left(E_t \left[U_{t+1}^{1-\gamma} \right] \right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right\}^{\frac{1}{1-\frac{1}{\psi}}}, \tag{1}$$

where U_t is time-t utility and C_t is time-t consumption. The household receives both labor and capital income. The household trades equities and a risk-free bond. The aggregate labor supply is inelastic and normalized to one. In every period t, the household purchases B_t units of the risk-free bond, which pays a gross risk-free return $R_{f,t}$ in the next period. The household can purchase $\omega_{j,t}$ shares of the equity of firm j. We denote $V_{j,t}$ as the cum-dividend equity price of firm j at time t and $D_{j,t}$ as the dividend payment of firm j at time t. We use W_t for the real wage at time t and $L_{j,t}$ for the household's labor supply to firm j. The household budget constraint at time t can be written as

$$C_t + \int \omega_{j,t} \left(V_{j,t} - D_{j,t} \right) dj + B_t = W_t \int L_{j,t} dj + R_{f,t-1} B_{t-1} + \int \omega_{j,t-1} V_{j,t} dj, \tag{2}$$

where the integration is with respect to all firms in the economy.

In the recursive equilibrium that we construct, C_t and U_t are functions of a vector of state variables, which we denote as S_t. We postpone the specification of S_t and its law of motion until we introduce the definition of equilibrium later in the section. Under the recursive preference, the stochastic discount factor that prices one unit of consumption goods paid at period $t+1$ (at state S_{t+1}) into period t (at state S_t), with consumption as the numeraire, can be constructed from the household's intertemporal marginal rate of substitution:

$$M_{t+1}(S_{t+1}|S_t) = \beta \left(\frac{C(S_{t+1})}{C(S_t)} \right)^{-\frac{1}{\psi}} \left(\frac{U(S_{t+1})^{1-\gamma}}{E_t \left[U(S_{t+1})^{1-\gamma} \right]} \right)^{\frac{1/\psi-\gamma}{1-\gamma}}. \tag{3}$$

To save notation, we will denote $M_{t+1}(S_{t+1}|S_t)$ as M_{t+1} below.

Production function At time t, firm j produces output $Y_{j,t}$ from capital $K_{j,t}$ and labor $L_{j,t}$ using a Cobb-Douglas production function of the following form:

$$Y_{j,t} = \bar{A}_t \left[X_{j,t}^{1-\nu} Z_{j,t}^{1-\nu} (u_{j,t} K_{j,t})^{\nu} \right]^{\alpha} L_{j,t}^{1-\alpha}, \tag{4}$$

where $1-\alpha$ is the labor share, ν is the span-of-control parameter, as in Atkeson and Kehoe (2005), and $u_{j,t}$ is the utilization rate of physical capital. The term $\bar{A}_t$ is the aggregate productivity of the economy, while $X_{j,t}$ and $Z_{j,t}$ are the permanent component and the

transitory component of firm-specific productivity, respectively:

$$\ln X_{j,t+1} = \mu_X + \ln X_{j,t} + \sigma_X \varepsilon_{j,t+1}, \tag{5}$$

$$\ln Z_{j,t+1} = \rho_Z \ln Z_{j,t} + \sigma_Z \varepsilon^Z_{j,t+1}, \tag{6}$$

where $\varepsilon_{j,t}$ and $\varepsilon^Z_{j,t}$ are i.i.d. standard normal shocks across firms and over time.

Variable capital utilization To address the empirical patterns of the profitability premium, we allow for variable capital utilization and shocks to capital depreciation. Let $I_{j,t}$ denote firm j's investment at time t. The law of motion for the firm's capital accumulation is given by:

$$K_{j,t+1} = \big(1 - \delta(u_{j,t}, \theta_{t+1})\big) K_{j,t} + I_{j,t}, \tag{7}$$

where $\delta(u_{j,t}, \theta_{t+1})$ represents the capital depreciation rate between t and $t+1$, which depends on both utilization $u_{j,t}$ chosen at time t and the aggregate capital depreciation shock θ_{t+1}.[5] Following the variable capital utilization literature, Greenwood et al. (1988), Jaimovich and Rebelo (2009), Garlappi and Song (2017), and Grigoris and Segal (2024), we adopt a constant elasticity capital depreciation function:

$$\delta\left(u_{j,t}, \theta_{t+1}\right) = \delta_0 + \theta_{t+1}\left(\delta_1 + \delta_u f(u_{j,t})\right). \tag{8}$$

Here, δ_0 represents a baseline depreciation rate unaffected by depreciation shocks. θ_{t+1} captures stochastic variations in depreciation. The parameter δ_1 determines the level effect, while $f(u)$ is an increasing and convex function with constant elasticity, which we parameterize and estimate in Section 5. In our specification, $\delta(u, \theta)$ increases with utilization rate u in a convex manner: more intensive capital usage accelerates depreciation and reduces capital accumulation in the subsequent period. Firms optimally choose capital utilization rates to balance current-period production benefits from utilization against loss in production associated with depreciation.

Firms' problem We focus on the recursive equilibrium where profit maximization can be written as a dynamic programming problem. An individual firm is identified by a vector of firm-specific state variables, (K, Z, X). Let $\mathbb{K}\times\mathbb{Z}\times\mathbb{X}$ be the state space of firm characteristics, and let Γ be a density on $\mathbb{K}\times\mathbb{Z}\times\mathbb{X}$ that describes the cross-sectional distribution of firms. Let $S = (A, \theta, \Gamma)$ be the vector of aggregate state variables, including current-period aggregate

[5] The importance of capital depreciation shocks has been emphasized in several studies, including Gertler and Kiyotaki (2010), Gertler and Karadi (2011), Kozlowski et al. (2018), and in continuous-time frameworks, He and Krishnamurthy (2013) and Brunnermeier and Sannikov (2014).

productivity, depreciation shocks, and the distribution of firm-level state variables. In the recursive equilibrium, S_t summarizes aggregate states and determines equilibrium prices through market clearing conditions. At time t, taking the law of motion of aggregate state variables as given, firm j chooses the dividend payout, $D_{j,t}$, investment, $I_{j,t}$, and capital utilization rate, $u_{j,t}$, to maximize its cum-dividend value:

$$V(K_{j,t}, Z_{j,t}, X_{j,t}; S_t) = \max_{I_{j,t}, u_{j,t}, D_{j,t}} \Big\{ D_{j,t} + (1-\kappa_D) E\big[M_{t+1} V(K_{j,t+1}, Z_{j,t+1}, X_{j,t+1}; S_{t+1}) \\ | K_{j,t}, Z_{j,t}, X_{j,t}; S_t \big] \Big\}, \tag{9}$$

subject to the following budget constraint

$$D_{j,t} + I_{j,t} + H(I_{j,t}, K_{j,t}) + FK_{j,t} \leq \Pi(K_{j,t}, Z_{j,t}, X_{j,t}; S_t \| u_{j,t}), \tag{10}$$

and capital accumulation equation (7). Investment entails a standard convex adjustment cost that takes a quadratic form, $H(I_{j,t}, K_{j,t}) = \frac{h}{2}(\frac{I_{j,t}}{K_{j,t}} - a_1)^2 K_{j,t}$, and an operating cost that is proportional to capital stock, $FK_{j,t}$, where $F > 0$ is a scaling parameter. The operating cost is proportional to firm size so that it does not disappear as the firm and the economy grow. A similar setup can be found in Herskovic et al. (2023).

In equation (10), firm j's profit at time t, $\Pi(K_{j,t}, Z_{j,t}, X_{j,t}; S_t \| u_{j,t})$, is given by[6]

$$\Pi\left(K_{j,t}, Z_{j,t}, X_{j,t}; S_t \| u_{j,t}\right) = \max_{L_{j,t}} \bar{A}_t \left[X_{j,t}^{1-\nu} Z_{j,t}^{1-\nu} (u_{j,t} K_{j,t})^{\nu} \right]^{\alpha} L_{j,t}^{1-\alpha} - W_t L_{j,t}. \tag{11}$$

It is convenient to simplify equation (11) by maximizing out labor and using the labor market clearing condition $\int L_{j,t} dj = 1$ to obtain the optimal labor input at firm j:

$$L_{j,t} = \frac{X_{j,t}^{1-\nu} Z_{j,t}^{1-\nu} (u_{j,t} K_{j,t})^{\nu}}{\int X_{i,t}^{1-\nu} Z_{i,t}^{1-\nu} (u_{i,t} K_{i,t})^{\nu} di}. \tag{12}$$

Substituting the optimal labor allocation (12) back into the production function (4), firm output is proportional to the firm's effective capital input $X_{j,t}^{1-\nu} Z_{j,t}^{1-\nu} (u_{j,t} K_{j,t})^{\nu}$, with a factor of proportionality that is common across firms and that each firm takes as given. Taking advantage of the constant income share property of the Cobb-Douglas production function, the labor bill is a fraction $1-\alpha$ of output, so that operating profit is the constant share α of output:

$$\Pi(K_{j,t}, Z_{j,t}, X_{j,t}; S_t \| u_{j,t}) = \alpha\, Y(K_{j,t}, Z_{j,t}, X_{j,t}; S_t \| u_{j,t}). \tag{13}$$

[6] We use the notation $(\cdot \| u_{j,t})$ to emphasize that profit at time t depends not only on time-t state variables, but also on the choice variable $u_{j,t}$.

In particular, operating profit is homogeneous of degree ν in capital $K_{j,t}$ and of degree ν in utilization $u_{j,t}$, which is the property that the firm's first-order conditions for investment and utilization rely on.

Aggregation of output and aggregate shocks In our economy, aggregate output can be computed by integrating firm-level output. Here, because labor flows frictionlessly across firms, ν is the effective returns to scale of capital at the firm level. Because adjustment costs generate heterogeneity in the marginal product of capital across firms, we in general need to know the joint distribution of $\{X_{j,t}, Z_{j,t}, K_{j,t}\}$ in order to compute aggregate output. It is convenient to define the effective level of aggregate capital stock index as:

$$\mathbf{K}_t^{\nu} = \int X_{j,t}^{1-\nu} Z_{j,t}^{1-\nu} (u_{j,t} K_{j,t})^{\nu} dj. \tag{14}$$

The denominator of the optimal labor allocation (12) is exactly $\mathbf{K}_t^{\nu}$, so that $L_{j,t} = X_{j,t}^{1-\nu} Z_{j,t}^{1-\nu} (u_{j,t} K_{j,t})^{\nu} \mathbf{K}_t^{-\nu}$. Substituting this back into the production function (4) gives the explicit expressions for firm-level output and, using (13), for operating profit:

$$Y(K_{j,t}, Z_{j,t}, X_{j,t}; S_t \| u_{j,t}) = \bar{A}_t X_{j,t}^{1-\nu} Z_{j,t}^{1-\nu} (u_{j,t} K_{j,t})^{\nu} \mathbf{K}_t^{-\nu(1-\alpha)}, \tag{15}$$

$$\Pi(K_{j,t}, Z_{j,t}, X_{j,t}; S_t \| u_{j,t}) = \alpha \bar{A}_t X_{j,t}^{1-\nu} Z_{j,t}^{1-\nu} (u_{j,t} K_{j,t})^{\nu} \mathbf{K}_t^{-\nu(1-\alpha)}. \tag{16}$$

Integrating (15) across firms, aggregate output is

$$Y_t = \bar{A}_t \left[\int X_{j,t}^{1-\nu} Z_{j,t}^{1-\nu} (u_{j,t} K_{j,t})^{\nu} dj \right]^{\alpha} = \bar{A}_t \mathbf{K}_t^{\alpha\nu}.$$

Our definition of $\mathbf{K}_t$ preserves the returns to scale of $K_{j,t}$, and $\mathbf{K}_t$ can be interpreted as the effective level of capital stock in the economy. Its level also depends on how capital is distributed across firms with different productivity and utilization rates. The index therefore captures the output effects of capital misallocation, in the notion of Hsieh and Klenow (2009), arising from adjustment costs and time-to-build in capital accumulation.

To generate endogenous long-run growth, we follow Romer (1986) and assume knowledge spillovers as a by-product of capital accumulation:

$$\bar{A}_t = A_t \left[\int X_{j,t}^{1-\nu} Z_{j,t}^{1-\nu} (u_{j,t} K_{j,t})^{\nu} dj \right]^{\frac{1-\alpha\nu}{\nu}}, \tag{17}$$

where A_t is an exogenous productivity shock and the term $\left[\int X_{j,t}^{1-\nu} Z_{j,t}^{1-\nu} (u_{j,t} K_{j,t})^{\nu} dj \right]^{\frac{1-\alpha\nu}{\nu}} = \mathbf{K}_t^{1-\alpha\nu}$ captures the externality generated by knowledge spillovers. We choose the power

$1-\alpha\nu$ on $\mathbf{K}_t$ so that aggregate production, as a function of the distribution of firm-level capital, exhibits constant returns to scale. Let $\mathbf{K}(S_t)$ denote the effective aggregate capital index as a function of the aggregate state in equilibrium.

The above construction allows us to write aggregate output conveniently as:

$$Y_t = A_t \mathbf{K}(S_t). \tag{18}$$

Using (17), firm-level operating profit (16) can equivalently be written as

$$\Pi(K_{j,t}, Z_{j,t}, X_{j,t}; S_t \| u_{j,t}) = \alpha A_t \mathbf{K}_t^{1-\nu} X_{j,t}^{1-\nu} Z_{j,t}^{1-\nu} (u_{j,t} K_{j,t})^{\nu}, \tag{19}$$

which is the form used in the proof of Lemma 1.

Finally, aggregate productivity follows a stationary AR(1) process:

$$\ln A_{t+1} = \mu_A + \rho_A (\ln A_t - \mu_A) + \sigma_A \varepsilon_{t+1}^A, \tag{20}$$

where ε_{t+1}^A is a standard normal shock and $|\rho_A| < 1$. Depreciation shocks follow

$$\ln \theta_{t+1} = \rho_\theta \ln \theta_t + \sigma_\theta \varepsilon_{t+1}^\theta, \tag{21}$$

where ε_{t+1}^θ is a standard normal shock. The aggregate state vector S_t collects A_t, θ_t, and the cross-sectional distribution of firms.

Firm entry and exit To complete the description of firm dynamics, we specify entry and exit. At the end of each period t, after dividend payout and investment, incumbent firms receive a random death shock with probability κ_D. Upon receiving the death shock, the firm exits the economy, and its capital stock is removed from production.

At the same time, new firms enter with initial productivity $\left(\bar{Z}, \bar{X}\right)$. We take entrants to be born before the survival draw that determines the set of firms operating in period $t+1$, so the pre-survival mass of entrants is $\bar{\phi} = \kappa_D/(1-\kappa_D)$ and the surviving entrant mass is κ_D. This normalization ensures that the total measure of operating firms is 1 at all times. Newborn firms start to produce one period after they enter the economy. As a result, for a new entrant firm j at time t, the representative household optimally chooses the amount of initial capital $\bar{K}_{j,t}$ to solve the following optimization problem:

$$\begin{aligned} \bar{V}(S_t) = \max_{\bar{K}_{j,t}} \Big\{ & - A_t[\bar{K}_{j,t} + \bar{H}(\bar{K}_{j,t}, \mathbf{K}(S_t))] \\ & + (1-\kappa_D) E\left[M_{t+1} V(\bar{K}_{j,t}, Z_{j,t+1}, X_{j,t+1}; S_{t+1}) | \bar{X}, \bar{Z}, S_t \right] \Big\}. \end{aligned} \tag{22}$$

Here, setting up a new firm requires a convex entry cost $\bar{H}(\bar{K}_{j,t}, \mathbf{K}_t)$, where $\mathbf{K}_t$ is the aggregate effective capital stock defined above. We assume that $\bar{H}$ exhibits constant returns to scale in $(\bar{K}_{j,t}, \mathbf{K}_t)$; together with the scaling of the entry cost by A_t in (22), this ensures the existence of a balanced growth path. With a probability of $(1-\kappa_D)$, new entrants survive to the next period. Let $\bar{K}(S_t)$ be the policy function of the above optimization problem. Conditional on survival, new entrants carry $\bar{K}_{j,t} = \bar{K}(S_t)$ amount of capital to the next period and become incumbent firms to produce output.

Firm distribution and its dynamics The entry and exit dynamics described above allow us to derive a law of motion for the density of the distribution of firm-level state variables, Γ_t. To derive a law of motion for this density, we first represent the law of motion of the firm-level endogenous state variable K as a function of state variables. In our model, firm type is described by a vector of firm-specific state variables (K, Z, X). Let $I(K, Z, X; S)$ and $u(K, Z, X; S)$ be the policy functions of incumbent firms, and use $'$ to represent next period variables in a recursive formulation. Equation (7) implies that in equilibrium, the law of motion of capital for a firm of type (K, Z, X) is given by:

$$K' = \left[1 - \delta\left(u(K, Z, X; S), \theta'\right)\right] K + I\left(K, Z, X; S\right). \tag{23}$$

Given the current aggregate state S and the realization of next period shock θ', for any K', let $\mathbb{I}_{K'}(K, Z, X|S, \theta')$ be the indicator function on the state space $\mathbb{K} \times \mathbb{Z} \times \mathbb{X}$ that takes a value of 1 if (K, Z, X) satisfies Equation (23). Let $\bar{\mathbb{I}}_{K'}(S)$ be the indicator function that takes a value of 1 if $\bar{K}(S) = K'$. The law of motion of Γ can be written as follows. For any $(K', Z', X') \in \mathbb{K} \times \mathbb{Z} \times \mathbb{X}$

$$\begin{aligned}\Gamma_{t+1}(K', Z', X') =& (1-\kappa_D) \int \mathbb{I}_{K'}(K, Z, X|S, \theta') \phi\left(\frac{\ln X' - \mu_X - \ln X}{\sigma_X}\right) \\ & \times \phi\left(\frac{\ln Z' - \rho_Z \ln Z}{\sigma_Z}\right) \Gamma_t(K, Z, X) dK dZ dX \\ & + \kappa_D \bar{\mathbb{I}}_{K'}(S) \phi\left(\frac{\ln X' - \mu_X - \ln \bar{X}}{\sigma_X}\right) \phi\left(\frac{\ln Z' - \rho_Z \ln \bar{Z}}{\sigma_Z}\right),\end{aligned} \tag{24}$$

where ϕ is the density of the standard normal distribution. Here we use the property that the dynamics of (5) and (6) imply that $\frac{\ln X' - \ln X - \mu_X}{\sigma_X}$ and $\frac{\ln Z' - \rho_Z \ln Z}{\sigma_Z}$ follow the standard normal distribution.

The above law of motion has an intuitive interpretation. $\Gamma_{t+1}(K', Z', X')$ is the density of firms at location (K', Z', X') in period $t+1$. They can come from either incumbent firms or new entrants. The first line in (24) accounts for the incumbent firms: if the realization of the

next period shock is such that Equation (23) holds, this firm is counted in the integral, where $\Gamma_t(K, Z, X)$ is the density of such firms in period t, and $\phi\left(\frac{\ln X' - \mu_X - \ln X}{\sigma_X}\right) \cdot \phi\left(\frac{\ln Z' - \rho_Z \ln Z}{\sigma_Z}\right)$ is the probability for such firms to reach (X', Z'). The second line, which accounts for new entrants, can be interpreted similarly.

Recursive competitive equilibrium Given the above description, a recursive equilibrium can be constructed as follows. The aggregate state variables are summarized by $S = (A, \theta, \Gamma)$, where Γ is a distribution on the firm-level state space $\mathbb{K} \times \mathbb{Z} \times \mathbb{X}$. The aggregate capital index $\mathbf{K}(S)$ is not an additional independent state variable; it is implied by S and equilibrium policy functions through Equation (14). Formally, a recursive competitive equilibrium consists of:

- A value function for the representative household's utility, $U(S)$, and the associated consumption policy function, $C(S)$,
- A value function for incumbent firms, $V(K, Z, X; S)$, and the associated policy functions, $I(K, Z, X; S)$ and $u(K, Z, X; S)$,
- A value function for new firms, $\bar{V}(S)$, its normalized counterpart $\bar{v}(S)$, and the associated policy function, $\bar{K}(S)$,
- A stochastic discount factor, $M(S'|S)$, and
- A law of motion of the distribution of firms, Γ,

such that the following conditions are met:

1. given the equilibrium stochastic discount factor and the law of motion of Γ, the value function $V(K, Z, X; S)$ and the policy functions, $I(K, Z, X; S)$ and $u(K, Z, X; S)$, solve the profit maximization problem for incumbent firms, (9),
2. the value function $\bar{V}(S)$ and the policy function $\bar{K}(S)$ solve the profit maximization problem for new entrants, (22),
3. the equilibrium SDF is consistent with household consumption, (3),
4. the law of motion of Γ is consistent with firms' optimal policies, (24), and
5. the goods market clears, that is, $\forall S$,

$$C(S) + \int \Big(I(K, Z, X; S) + H(I(K, Z, X; S), K) + FK\Big)\Gamma(K, Z, X)dKdZdX + \bar{\phi}A\left[\bar{K}(S) + \bar{H}(\bar{K}(S), \mathbf{K}(S))\right] = A\mathbf{K}(S), \tag{25}$$

where $\bar{\phi} = \kappa_D/(1-\kappa_D)$ is the pre-survival mass of entrants, and,

6. asset markets clear so that $\int \omega_{j,t} dj = 1$ for all t, and the labor market clears so that aggregate labor demand is one.[7]

4 Asset pricing implications

We provide a summary of the main asset pricing implications of our model, which serves as a road map for the rest of the paper.

A1 The market-to-book ratio and gross profitability are both increasing functions of X and Z. However, quantitatively, firms' market-to-book ratio is primarily determined by the permanent component, X, and the transitory component Z mainly determines the gross profitability.

A2 While firm-level cash flow responds positively to both X and Z, the impulse responses with respect to X are much more persistent than those with respect to Z.

A3 The firm investment rate is increasing in X, and the capital utilization rate is increasing in Z.

A4 The equity value of high-X firms is less sensitive to productivity shocks A_t than the equity value of low-X firms. The equity value of high-Z firms is more exposed to capital depreciation shocks θ_t than that of low-Z firms.

Taken together, the above asset pricing implications provide an explanation for the coexistence of the value and profitability premiums and the empirical evidence we document in Section 2. In the rest of this section, we use equilibrium conditions to explain the intuition for the above results. In Section 5, we calibrate our model and assess the quantitative importance of the proposed economic mechanisms. In the rest of the paper, we drop the firm-level subscript j without causing confusion.

Market-to-book ratio and investment In this section, we use the firms' optimality condition for investment to explain implications A1 and A3. As we show in Lemma 1 in Appendix B, the firm's value function and policy functions satisfy the following homogeneity property:

[7] To simplify equilibrium conditions, we have imposed labor market clearing to derive the optimal labor allocation equation (12).

$$
\begin{aligned}
V\left(K_t, Z_t, X_t; S_t\right) &= \mathbf{K}_t X_t v\left(k_t, Z_t; m_t, A_t, \theta_t\right), \\
I\left(K_t, Z_t, X_t; S_t\right) &= \mathbf{K}_t X_t i\left(k_t, Z_t; m_t, A_t, \theta_t\right), \\
D\left(K_t, Z_t, X_t; S_t\right) &= \mathbf{K}_t X_t d\left(k_t, Z_t; m_t, A_t, \theta_t\right),
\end{aligned} \tag{26}
$$

where $k_t = \frac{K_t}{X_t \mathbf{K}_t}$ and m_t is constructed from Γ_t.[8] In the cross-section, firm type is summarized by three state variables, (k, Z, X), where $k = \frac{K}{X\mathbf{K}}$. Because value functions and policy functions are homogeneous with respect to X, we can write the normalized value and policy functions as functions of two state variables, (k, Z). More importantly, applying this homogeneity property in aggregation allows us to replace the aggregate state variable Γ, a distribution defined on the three-dimensional space of (k, Z, X), with a lower-dimensional measure, m defined on the state space of (k, Z). In Appendix C.3, we show that the recursive equilibrium defined in the last section can be formulated as a normalized equilibrium system defined on a lower-dimensional state space, where firm-specific state variables are (k, Z) and aggregate state variables are represented by (m, A, θ).

Since the adjustment cost function $H\left(I, K\right)$ exhibits constant returns to scale, we can write $H\left(I, K\right) = h\left(\frac{I}{K}\right) K$. For notational convenience, we denote $h_K\left(\frac{I}{K}\right) = \frac{\partial}{\partial K}\left[h\left(\frac{I}{K}\right) K\right] = h\left(\frac{I}{K}\right) - h'\left(\frac{I}{K}\right) \frac{I}{K}$ as the partial derivative of the adjustment cost function with respect to capital. Additionally, we define $M_{t,t+s} = \prod_{\tau=1}^{s} M_{t+\tau}$ as the s-period-ahead pricing kernel at date t, and $\chi_{t,t+s} = (1-\kappa_D)^s \prod_{\tau=1}^{s-1}\Big(1 - \delta(u_{t+\tau}, \theta_{t+\tau+1})\Big)$ as the survival-adjusted capital carryover factor for investment made at date t and used in production at date $t+s$. It combines survival over s periods with the fraction of capital that remains after depreciation over the intervening $s-1$ periods. For $s = 1$, the product is empty, so $\chi_{t,t+1} = (1-\kappa_D)$.

The following proposition establishes the standard Q-theory relationship between investment rate and the discounted value of the marginal product of capital.

Proposition 1. *(Q-theory of investment)*

The optimal investment-to-capital ratio satisfies

$$
\Big[1 + h'\Big(\frac{I_t}{K_t}\Big)\Big] = \mathbb{E}_t\Big[\sum_{s=1}^{\infty} M_{t,t+s}\chi_{t,t+s}\Big(\alpha\nu\frac{Y_{t+s}}{K_{t+s}} - h_K\Big(\frac{I_{t+s}}{K_{t+s}}\Big)\Big)\Big], \tag{27}
$$

Proof. See Appendix B. □

Equation (27) represents the first-order optimality condition for investment. The left-hand side shows the marginal cost of investment, while the right-hand side captures the

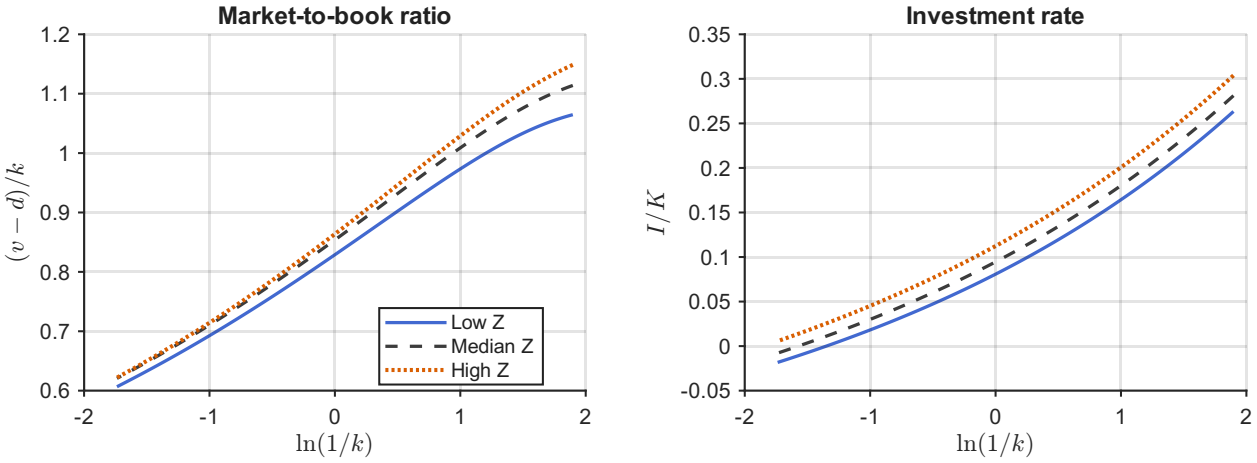
[8] See Equation (C19) in Appendix C.3 for this construction.

marginal benefit: the present value of future marginal product of capital plus the savings in adjustment costs from an additional unit of investment. In our Cobb-Douglas production function with decreasing returns to scale, factor income maintains constant shares of output: labor receives $1-\alpha$, capital receives $\alpha\nu$, and profits account for $\alpha(1-\nu)$. Consequently, the marginal product of capital is proportional to its average product by a factor of $\alpha\nu$, representing capital's share of total output.

As noted in asset pricing implication A1, the firm-level market-to-book ratio is primarily driven by the permanent productivity component X rather than the transitory component Z. Positive shocks to both X and Z raise future productivity. Due to adjustment costs, investment increases cannot fully offset productivity shocks' impact on capital's marginal product, causing both investment rates and market-to-book ratios to rise. However, X's impact persists long-term, while Z's effect is temporary.

Figure 3: Investment and market-to-book ratio

This figure plots firms' optimal policy for the investment rate and market-to-book ratio, as functions of normalized permanent productivity X and transitory productivity Z, at the deterministic steady state. We define the normalized permanent productivity as $\ln(\frac{\mathbf{K}_t X}{K}) = \ln(\frac{1}{k})$, which is the log of the inverse of state variable k. The solid, dashed, and dotted lines denote firms with low, medium, and high levels of transitory productivity Z, respectively. "low Z" represents one conditional standard deviation below the median value of transitory productivity, while "high Z" means one standard deviation above the median.



To illustrate the different impacts of X and Z, Figure 3 plots the market-to-book ratio (left panel) and the investment rate $\frac{I}{K}$ (right panel) for three levels of Z: low, medium, and high. Instead of using X itself, we use $\ln(\frac{1}{k}) = \ln(\frac{\mathbf{K}X}{K})$ as the horizontal axis, because the impact of X on the market-to-book ratio and the investment rate is fully summarized by the state variable k. For a fixed K, positive shocks to X move firms to the right in the $\ln(\frac{1}{k})$ space.

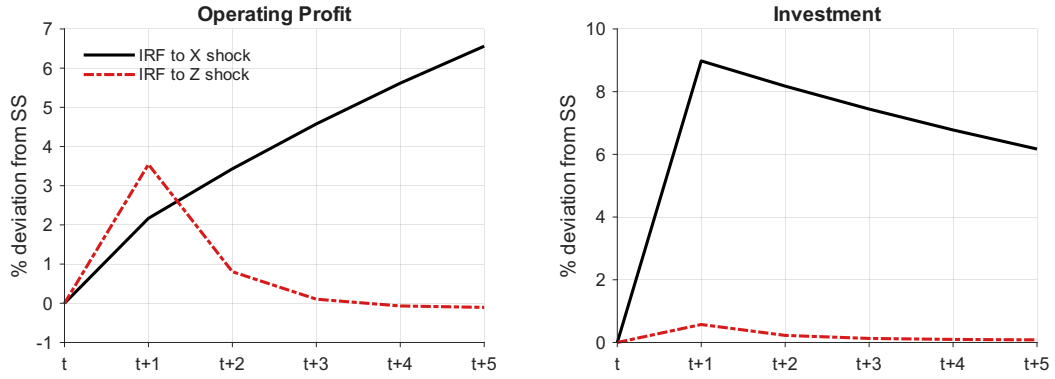
Clearly, both investment rates and market-to-book ratios increase with X and Z. In Figure 3, "low Z" represents one conditional standard deviation below the median value of

transitory productivity, while "high Z" means one standard deviation above the median. As shown in Figure 3, the variations in investment rates and market-to-book ratios are much more pronounced across different values of permanent productivity X than across different values of transitory productivity Z. The contrast is quantitatively stark: across the plotted range of $\ln(\frac{1}{k})$, variation in market-to-book ratios and investment rates associated with X is roughly an order of magnitude larger than the corresponding variation associated with Z.

In the model, individual firm characteristics are captured by two state variables: k and Z. Since Z has a relatively minor impact on the market-to-book ratio, sorting firms by this ratio primarily separates them along the k dimension. Firms on the right of the $\frac{1}{k}$ space have a higher level of X relative to their capital stock. They have a higher market-to-book ratio and exhibit a significantly higher investment rate than low-market-to-book firms on the left side of the $\frac{1}{k}$ space. This feature aligns with the evidence in Figure 2, where market-to-book sorted portfolios show significant differences in investment rates. In contrast, sorting by profitability ratio primarily separates firms along the Z dimension, as transitory shocks Z have a greater magnitude than permanent shocks X in our model.

Figure 4: Impulse response functions for transitory and permanent shocks

This figure compares the impulse responses of a firm's operating profit (left panel) and investment rate (right panel) to a permanent productivity shock X and a transitory productivity shock Z. Both impulses equal one standard deviation of the permanent shock σ_X and are realized between t and $t+1$. The horizontal axis is years after the shock and the vertical axis is the percentage deviation from the firm's balanced-growth path without the shock. The solid black line is the response to X, the dash-dotted red line the response to Z.



To illustrate asset pricing implications A2 and A3, Figure 4 plots a firm's responses to permanent and transitory productivity shocks. The left panel shows operating profit (cash flow), and the right panel shows the investment rate. Both shocks raise log productivity by the same magnitude on impact, so the comparison holds the initial shock size fixed. Responses are measured relative to the corresponding no-shock path, using the parameter values reported in Table 2.[9]

[9]The impulse response function depends on the firm-level state variable (k, Z). In Figure 4, we set Z to

The left panel of Figure 4 illustrates asset pricing implication A2. Both X and Z shocks raise current productivity, but their effects on future productivity differ. A positive Z shock mainly raises current output, and firms respond by using existing capital more intensively. This utilization response amplifies the initial increase in cash flow. However, as Z mean reverts, the productivity advantage disappears and the cash flow response fades quickly. A positive X shock raises both current and future productivity. The cash flow response therefore persists and grows as the firm accumulates capital in response to permanently higher productivity. This difference in persistence is consistent with the evidence in Figure 1, where cash flow-to-asset differences across market-to-book-sorted portfolios decay more slowly than those across profitability-sorted portfolios.

The right panel of Figure 4 illustrates the investment response described in asset pricing implication A3. The investment condition (27) equates the marginal cost of investment to its expected discounted marginal benefit. A positive X shock raises the marginal product of capital over many future periods, increasing the benefit of investment. A positive Z shock also raises productivity, but much of this gain disappears before new capital becomes productive. Its effect on the benefit of investment is therefore limited. As a result, firms increase investment much more strongly in response to X than to Z, even when the initial shocks have the same size. In the right panel, the investment rate rises by about 9.0% after the X shock and 0.6% after the Z shock, and the response to X remains large over the plotted horizon. This pattern is consistent with Figure 3, where investment rates and market-to-book ratios depend primarily on the permanent productivity component.

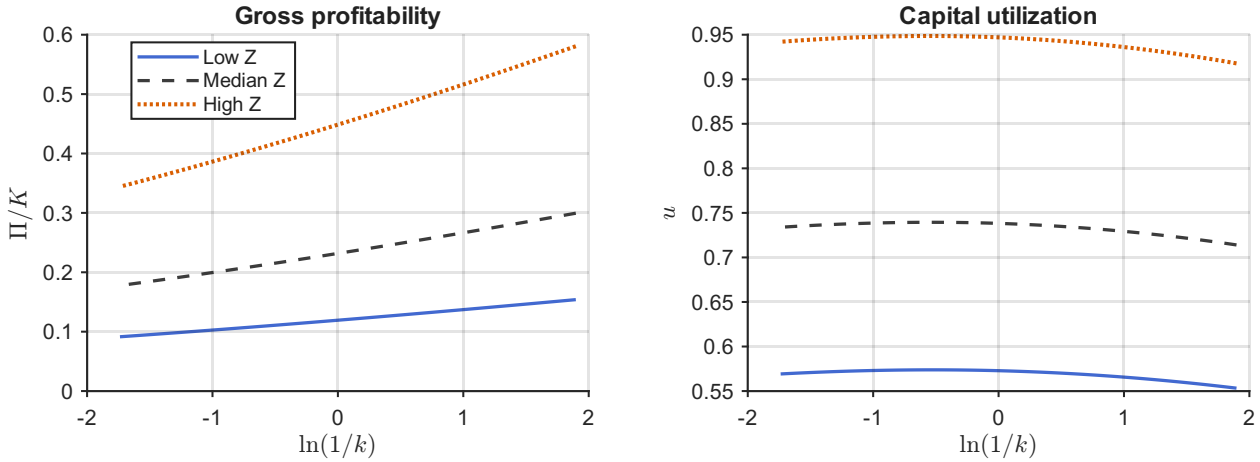
Capital utilization and profitability In this section, we examine the optimality condition for capital utilization and discuss the intuition behind related asset pricing implications. A firm's profitability ratio is defined as gross profit divided by capital, $\frac{\Pi_{i,t}}{K_{i,t}}$. By Equation (16), this ratio can be written as $\frac{\Pi_{i,t}}{K_{i,t}} = \alpha A_t Z_{i,t}^{1-\nu} u_{i,t}^{\nu} k_{i,t}^{\nu-1}$. This formulation reveals that the profitability ratio increases with the transitory productivity component Z. It also increases with the permanent productivity component X, since normalized capital k is inversely related to X. However, from a quantitative perspective, cross-sectional heterogeneity in gross profitability is predominantly determined by differences in the transitory productivity component Z, as its conditional volatility is substantially larger than that of X in our model.

To illustrate how profitability depends on state variables, Figure 5 plots the gross profitability ratio (left panel) and the optimal capital utilization rate (right panel) as functions of normalized permanent productivity, $\ln\left(\frac{1}{k}\right)$, for three levels of transitory Z shock. The

its median and k to its firm-level deterministic steady state, the level at which the firm's capital grows at the economy's balanced-growth rate.

Figure 5: Utilization and profitability

This figure plots firms' optimal policy for capital utilization and gross profitability as functions of normalized permanent productivity and transitory productivity Z, at the deterministic steady state. We define the normalized permanent productivity as $\ln(\frac{\mathbf{K}_t X}{K}) = \ln(\frac{1}{k})$, which is the log of the inverse of state variable k. The solid, dashed, and dotted lines denote firms with low, medium, and high levels of transitory productivity Z, respectively. "low Z" represents one conditional standard deviation below the median value of transitory productivity, while "high Z" means one standard deviation above the median.



left panel shows substantial differences in gross profitability across different levels of Z, while profitability remains almost flat across different levels of the permanent component of productivity X. These patterns indicate that, qualitatively, profitability increases with both X and Z in our model. However, quantitatively, profitability is primarily driven by the state variable Z rather than X. This contrasts sharply with the pattern of market-to-book ratios shown in Figure 3, where heterogeneity mainly stems from X.

Together, the left panels of Figures 3 and 5 confirm asset pricing implication A1: sorting on market-to-book ratio identifies heterogeneity in the permanent component of productivity X, while sorting on gross profitability identifies the transitory component of productivity Z.

Asset pricing implication A3 states that investment responds strongly to the permanent component of productivity shock X and weakly to the transitory component Z, while capital utilization increases sharply with Z but responds weakly or even negatively to X. To understand this relationship, we first present a proposition about the optimality condition for capital utilization.

Proposition 2. *Under the utilization functional form introduced in Section 5.2, the optimal utilization policy is given by*

$$u_t^{1+\lambda} = \frac{\nu \Pi_t}{(1-\kappa_D) E_t[M_{t+1}\theta_{t+1}\tilde{q}_{t+1}]\delta_u K_t}, \tag{28}$$

where $\tilde{q}_{t+1}$ is defined recursively in equation (B10) in Appendix B.

In our model, higher utilization rates increase current-period output at the cost of accelerated capital depreciation. Firms optimally choose capital utilization to balance these tradeoffs. The numerator of equation (28) represents current-period capital income (a constant fraction of output), capturing the marginal benefit of utilization. The denominator represents the marginal benefit of preserving capital: the present value of the marginal benefit of investment in the next period. The variable $\tilde{q}_{t+1}$ is the marginal value of capital at the beginning of period $t+1$ before depreciation (see equation (B10)). The term $E_t[M_{t+1}\theta_{t+1}\tilde{q}_{t+1}]$ includes θ_{t+1} because utilization costs depend on the next period's depreciation rate, which is a function of θ_{t+1}.

Proposition 2 has an intuitive interpretation: capital utilization increases with current-period profit but decreases with the present value of future investment benefits. An increase in Z raises both current-period output (numerator) and the marginal benefit of investment (denominator). As previously explained, Z's impact on market-to-book ratio is minimal due to its low persistence, but its impact on current-period output is substantial due to its large magnitude. The right panel of Figure 5 shows that optimal utilization rates increase strongly with Z but remain almost flat across different levels of X. Together, the right panels of Figures 3 and 5 confirm asset pricing implication A3: investment rates increase strongly with the permanent productivity X, while capital utilization responds much more strongly to the transitory productivity Z.

Value and profitability premiums The two-factor productivity structure enables us to distinguish between value/growth firms and profitability-sorted firms. In our model, sorting on market-to-book ratio identifies firms based on their permanent productivity component and reveals the state variable k. Conversely, sorting on profitability distinguishes firms according to their transitory productivity component and identifies the state variable Z. Consequently, value and profitability factors capture distinct aspects of firm-level fundamentals, which explains asset pricing implication A4 of our model.

First, exposure to capital depreciation shocks becomes more negative as Z increases. Because depreciation risk carries a negative price, the more negative exposure of high-profitability firms raises their expected returns. From equation (7), capital depreciation represents a negative cash flow to firms and provides a form of operating leverage. As previously discussed, high-Z firms maintain higher capital utilization rates, resulting in faster capital depreciation. When a positive depreciation shock to θ occurs, it depletes the aggregate capital stock. High-Z firms experience larger losses due to their higher capital depreciation. In contrast, low-Z firms are relatively insulated from these shocks because of their lower utilization and depreciation rates. The profitability premium emerges naturally in our model

because the profitability ratio increases with Z.[10]

Notably, permanent productivity X has minimal impact on optimal capital utilization, as illustrated in the right panel of Figure 5. This occurs because an increase in X produces two counterbalancing effects: it enhances current-period productivity (encouraging utilization) while simultaneously raising expected future productivity (incentivizing lower utilization to reduce depreciation). Consequently, firms across different book-to-market ratio portfolios exhibit similar capital utilization rates and comparable exposure to capital depreciation shocks.

Meanwhile, exposure to aggregate productivity shocks increases with k, generating a value premium. The mechanism driving higher risk exposure to aggregate productivity shocks resembles that described by Zhang (2005): capital adjustment costs $H(I, K)$ and, more critically, fixed operating costs F create operating leverage.[11] As shown in Figure 3, firms on the left of the $\frac{1}{k}$ space have a lower market-to-book ratio and at the same time, a lower level of cash flow per unit of capital. As a result, convex and fixed costs constitute a higher fraction of the present value of their cash flow, resulting in higher operating leverage and a greater exposure to aggregate TFP shocks in equilibrium.

This operating leverage channel also affects firms with different levels of Z. Low-Z firms (low-profitability firms) similarly experience higher operating leverage with respect to aggregate TFP shocks. The combination of investment adjustment costs and TFP shocks therefore dampens the profitability premium, a key challenge identified by Novy-Marx (2013) for Q-theory-based asset pricing models. However, quantitatively, this mechanism is insufficient to completely offset the profitability premium we discussed previously. As shown in Figure 3, investment rate dispersion is substantially higher along the X dimension than the Z dimension. Since adjustment costs are a function of investment rates, the operating leverage channel arising from the investment channel is much more pronounced across market-to-book-sorted portfolios, and fairly weak across profitability-sorted portfolios. Additionally, as we demonstrate in the next section, this offsetting mechanism enables our model to capture the empirical observation that value and profitability factors are negatively correlated.

[10] Grigoris and Segal (2024) find that high utilization firms earn a lower average return than low utilization firms when sorting directly on utilization rates. The utilization premium they document is likely due to a different risk factor because the authors find it to be unrelated to firm productivity differences or the profitability risk premium.

[11] Without fixed operating costs, convex capital adjustment costs alone may or may not generate a value premium, depending on parameter values. See Kogan and Papanikolaou (2012) for a discussion on the importance of fixed costs in generating operating leverage.

5 Quantitative results

In this section, we estimate our model using simulated method of moments (SMM) and evaluate its ability to jointly account for the cross-section of investment dynamics and expected returns.

5.1 Numerical solution

To facilitate estimation, we develop an efficient global solution method based on Krusell and Smith (1998). We make two modifications to their approach that allow us to avoid the computationally costly simulation step. First, we show that, in equilibrium, the summary measure m follows a mixture of lognormal distributions. This distributional characterization allows us to compute moments of the cross-sectional distribution analytically, without simulation. Second, we solve firms' policy functions on a discrete grid of the state space and approximate them using cubic splines to ensure accuracy. When evaluating aggregate moments, we translate the cubic-spline approximations of the utilization policy into log-polynomials. Combined with the mixture-of-lognormals representation, this log-polynomial representation allows us to evaluate the cross-sectional integrals in closed form using moments of lognormal distributions. As a result, the aggregate law of motion can be computed without the simulation step required by the standard Krusell and Smith (1998) algorithm. A detailed explanation of the numerical method is provided in Appendix D.

5.2 Estimation procedure

We divide the model parameters into two groups. The first group consists of standard parameters commonly used in production-based asset pricing literature, which can be calibrated by matching standard macroeconomic moments or following previous literature. The second group contains parameters unique to our model, which we estimate using simulated method of moments (SMM).

For the calibrated parameters, we set risk aversion to $\gamma = 20$, which falls in the range estimated by Malloy et al. (2009). The intertemporal elasticity is set to $\psi = 1.7$, consistent with the long-run risk literature. We choose a time discount factor $\beta = 0.954$, which generates a realistically low interest rate: the model delivers a risk-free rate of 0.2% per year against 1.3% in the data. Growth is endogenous in our economy, so we do not calibrate a trend growth rate directly; instead we target the average growth rate of output, 1.9% per year, which matches the average growth rate of the U.S. economy in the post-war period.

For production technology parameters, we set the capital share parameter α to 0.3, reflecting the average capital share in the U.S. economy. We calibrate the span of control parameter ν to 0.85, consistent with estimates in the literature such as Atkeson and Kehoe (2005). The exogenous firm death rate κ_D is set at 5% per year, in line with the firm-dynamics literature, such as Luttmer (2007). We normalize new entrant firms' initial productivity $\bar{X}$ so that the integral over the cross-sectional distribution, $E[X] = 1$. We calibrate the exogenous aggregate productivity process in equation (20) to $\rho_A = 0.985$ and $\sigma_A = 0.051$; these values are close to the calibration in Kung and Schmid (2015). The unconditional mean of log productivity, $\mu_A = -0.021$, is a normalization parameter. We approximate the log productivity process using the Rouwenhorst method.

We specify the capital depreciation shock in Equation (8) as a log AR(1) process:

$$\ln \theta_t = \mu_\theta + \rho_\theta(\ln \theta_{t-1} - \mu_\theta) + \sigma_\theta \varepsilon_t^\theta.$$

We normalize $\mu_\theta = 0$ so that the mean of θ_t is one at the deterministic steady state, while leaving σ_θ for estimation. We estimate these parameters using macroeconomic quantity moments and asset pricing moments.

We set the function $f(u) = \frac{u^{1+\lambda} - \bar{u}^{1+\lambda}}{1+\lambda}$ with a constant slope parameter $\delta_u > 0$ and a constant elasticity parameter $\lambda > 0$. In our model simulation, we verify that the average firm-level utilization rate equals $\bar{u}$. To simplify estimation, we parameterize δ_1 in Equation (8) as $\bar{\delta} - \delta_0$. Since the steady-state value of $\theta_t = 1$, $\bar{\delta}$ represents the long-run average depreciation rate, which we set to $\bar{\delta} = 8\%$ per year, consistent with Jorgenson (1996). We calibrate the baseline depreciation rate to $\delta_0 = 6\%$, so that $\delta_1 = \bar{\delta} - \delta_0 = 2\%$ is the part of average depreciation that is exposed to the depreciation shock, and we leave the slope parameter δ_u and the elasticity parameter λ for estimation. Three remaining auxiliary parameters are fixed rather than estimated because the targeted moments are not sufficiently informative about them: the fixed operating cost $F = 0.065$, the drift of the permanent idiosyncratic productivity component $\mu_X = 0.015$, and the linear term of the investment adjustment cost $a_1 = 0.05$. All calibrated parameters are listed in the top panel of Table 2.

The eight parameters are estimated using the eighteen moments listed in Table 3, providing ten overidentifying moments. These moments fall into three categories. First, macroeconomic moments include the mean and volatility of output growth, the volatilities of consumption and investment growth relative to output growth, the correlation between consumption and investment growth, the mean and volatility of the risk-free rate, and the mean and volatility of the aggregate depreciation rate. These moments are particularly informative about production technology parameters, such as the adjustment cost function curvature h, capital depreciation technology parameters, and the persistence and volatility

Table 2: Model Parameters

This table reports model parameters. Panel A lists the calibrated parameter values. Panel B lists the eight parameters estimated by simulated method of moments, with standard errors. Standard errors follow Lee and Ingram (1991).

Panel A: Calibrated Parameters

Parameter	Symbol	Value
Time discount factor	β	0.954
Relative risk aversion	γ	20.000
IES	ψ	1.700
Capital share in production	α	0.300
Span of control parameter	ν	0.850
Persistence of aggregate productivity	ρ_A	0.985
Volatility of aggregate productivity	σ_A	0.051
Mean of aggregate productivity (log)	μ_A	-0.021
Mean depreciation shock	μ_θ	0.000
Baseline depreciation rate	δ_0	0.060
Long-run average depreciation rate	$\bar{\delta}$	0.080
Fixed operating cost	F	0.065
Drift of idio. permanent productivity	μ_X	0.015
Capital adj. cost (linear term)	a_1	0.050
Exit probability	κ_D	0.050

Panel B: Estimated Parameters

Parameter	Symbol	Estimate	S.E.
Utilization elasticity parameter	λ	1.499	0.411
Utilization slope parameter	δ_u	0.381	0.087
Capital adj. cost parameter	h	2.641	0.301
Persistence of depreciation shock	ρ_θ	0.944	0.019
Volatility of depreciation shock	σ_θ	0.200	0.059
Volatility of idio. permanent productivity shock	$(1-\nu)\sigma_X$	0.024	0.005
Persistence of idio. transitory productivity shock	ρ_Z	0.286	0.026
Volatility of idio. transitory productivity shock	$(1-\nu)\sigma_Z$	0.444	0.048

of aggregate depreciation shocks (ρ_θ and σ_θ). Second, the investment rates of market-to-book and profitability-sorted portfolios are informative about firm-level productivity process parameters (σ_X, ρ_Z, and σ_Z), as the investment rate spreads across these portfolios reveal the two-factor productivity structure shown in Figure 2. We also target the profitability spreads produced by both sorts, which pin down how sharply each sorting variable separates firms on the transitory component of productivity. These profitability-related moments are informative about ρ_Z and σ_Z. Third, the expected returns of market-to-book- and profitability-sorted portfolios, along with the market excess return, the correlation between value and profitability factors and the cash-flow dynamics of characteristic-sorted portfolios, capture key asset pricing implications of our model. Since our model does not explicitly model financial leverage, we lever up model-generated returns by assuming a leverage ratio of 2, following Croce (2014).

We estimate all parameters jointly while imposing the model's general equilibrium restrictions using SMM (Lee and Ingram (1991)). Let Θ denote the vector of parameters to be estimated. The SMM procedure minimizes the weighted distance between empirical moments Φ^E and model-simulated moments $\Phi^M(\Theta)$ generated from the model:

$$\hat{\Theta} = \text{argmin}_\Theta \left[\Phi^E - \Phi^M(\Theta)\right]' W \left[\Phi^E - \Phi^M(\Theta)\right] \tag{29}$$

where W is the weighting matrix. For J empirical moments, we set $W = \text{diag}\left(\Phi^E \Phi^{E\prime}\right)^{-1}$, so that each moment enters as a relative deviation and the objective is scale-free. We solve the minimization problem using a simulated annealing method.

All data moments in Table 3 are computed at an annual frequency. Macroeconomic variables come from the Federal Reserve Bank of St. Louis for the period 1963-2024. Real consumption is measured as personal consumption expenditures divided by the CPI. Real output is calculated by dividing nominal GDP by the GDP deflator, and real investment is obtained by deflating fixed private investment with the investment deflator. Market returns, risk-free rate, and book-to-market (BM) sorted portfolio returns are from Ken French's website. Gross-profitability-sorted portfolios are constructed following Novy-Marx (2013).

5.3 Estimation results

As shown in Table 2, the estimated elasticity of the capital depreciation function is $\lambda = 1.50$ and the estimated slope is $\delta_u = 0.38$, so depreciation rises with capital utilization; both are precisely estimated. The estimated adjustment cost parameter, $h = 2.64$, is close to the value of 2 commonly used in the literature studying convex and non-convex investment costs, including Bloom (2009) and Cooper and Haltiwanger (2006). The depreciation shock is both

persistent and volatile, $\rho_\theta = 0.94$ and $\sigma_\theta = 0.20$. These two parameters allow us to match the average and volatility of the depreciation process in the data. The three parameters δ_u, ρ_θ, and σ_θ that govern the dynamics of depreciation are primarily informed by the cross-sectional dynamics of investment and the volatility of the macroeconomic moments. All eight parameters are estimated with t-statistics above three.

Table 3: Moments Used in SMM Estimation

This table reports the eighteen moments used in the SMM estimation of the model. Panel A reports the aggregate moments and Panel B the cross-sectional moments. We solve and simulate the model annually and report all statistics at an annual frequency; values are reported as decimals rather than percentages. The Model column is the population moment at the estimated parameters, computed on a panel of 5,000 firms over 800 years after a 400-year burn-in. In Panel B each block reports the same three statistics for one sort. All quintile spreads are high-minus-low, so $E[R]$ in the market-to-book block is the excess return of the highest market-to-book quintile minus that of the lowest quintile and is negative when there is a value premium; model returns are levered at a ratio of two. $corr(R^{MB}, R^{GP/A})$ is the correlation between the two high-minus-low return spreads, which is minus the correlation between the value factor and the profitability factor. CF^{rel} compares the rates at which the two sorts' cash flow spreads decay: for each sort let r be the fraction of the year-0 high-minus-low log cash flow-to-asset spread that survives to year 5, computed from the median log cash flow-to-asset ratio in the highest and lowest quintiles, and let $CF^{rel} = \ln r^{MB} / \ln r^{GP/A}$. A value below one means the market-to-book spread in terms of cash flow is the more persistent of the two.

Panel A: Aggregate Moments

Moment	Data	Model	Moment	Data	Model
$E[\Delta y]$	0.019	0.025	$\sigma(\Delta y)$	0.029	0.034
$E[R^f]$	0.013	0.002	$\sigma(R^f)$	0.022	0.009
$E[\delta_t]$	0.076	0.073	$\sigma(\delta_t)$	0.010	0.004
$E[R^M] - E[R^f]$	0.073	0.071	$\sigma(\Delta c)/\sigma(\Delta y)$	0.710	0.852
$corr(\Delta c, \Delta i)$	0.745	0.869	$\sigma(\Delta i)/\sigma(\Delta y)$	2.627	1.511

Panel B: Cross-sectional Moments

Market-to-book sort			Gross profitability sort		
Moment	Data	Model	Moment	Data	Model
I/K (Hi−Lo)	0.080	0.101	I/K (Hi−Lo)	0.040	0.052
GP/A (Hi−Lo)	0.150	0.169	GP/A (Hi−Lo)	0.650	0.445
$E[R]$ (Hi−Lo)	-0.033	-0.015	$E[R]$ (Hi−Lo)	0.048	0.016
$corr(R^{MB}, R^{GP/A})$	0.523	0.390	CF^{rel} (MB rel. GP/A)	0.390	0.440

As shown in Panel A of Table 3, our model successfully matches the moments of macroeconomic quantities. The preference and technology parameters align with previous production-based asset pricing models such as Croce (2014) and Ai et al. (2013). Output

grows at 2.5% per year against 1.9% in the data, with a volatility of 3.4% against 2.9%. Consumption growth is smoother than output growth and investment growth is more volatile, with relative volatilities of 0.85 and 1.51 against 0.71 and 2.63 in the data, and the two are strongly positively correlated, 0.87 against 0.75. The average depreciation rate is 7.3% per year, against 7.6% in the data and close to the average economic depreciation reported in Jorgenson (1996). The model-simulated depreciation rate is smoother than its empirical counterpart, 0.4% against 1.0%. This is partly because the slope parameter δ_u is small.

The levered market excess return of 7.1% per year is close to 7.3% in the data. The model achieves this, however, with a risk-free rate that is too low and too smooth: 0.2% per year against 1.3% in the data, with a volatility of 0.9% against 2.2%. With no shocks to the intertemporal substitution motive, there is little in the model to move the short rate around, and the level of the equity premium is obtained in part by depressing the risk-free rate rather than by raising the average return on capital. Like most production-based asset pricing models, aggregate market return volatility is also low in our model. First, we do not include investment-specific shocks as in Papanikolaou (2011), which would generate a wedge between the relative prices of investment and consumption. Second, in the data, a significant fraction of stock market volatility can arise from reallocation between capital income and labor income, not just from aggregate productivity or depreciation shocks. Our aggregate model might understate stock market volatility due to the absence of any sectoral reallocation mechanism.

The key premise of our theory is a two-factor structure of firm-level productivity. Quantitatively, this structure is governed by the productivity parameters for permanent and transitory shocks: σ_X, σ_Z, and ρ_Z.[12] While permanent shocks have smaller contemporaneous volatility, they persist and have long-term impacts on firm cash flows. The transitory shocks have larger contemporaneous volatility, but they decay quickly and have limited long-term effects. The estimated σ_Z helps us to match the dispersion of profitability across firms.

Because the market-to-book ratio is forward-looking, portfolios sorted by market-to-book load heavily on the persistent component of productivity, whereas profitability sorting primarily distinguishes firms based on the transitory component. As demonstrated in Section 2.2, market-to-book-sorted portfolios exhibit modest but very persistent cash flow differences, while profitability-sorted portfolios display substantial initial differences that rapidly converge. Our model reproduces the ordering of these empirical patterns. Specifically, in the data, the difference in log cash flow-to-asset ratios between the growth and value portfolios decays much more slowly than the difference between the low- and high-profitability portfolios. We use the relative five-year decay of the two spreads to summarize the persistence

[12]In the model, the X and Z shocks are scaled by $(1-\nu)$ when entering the production function. We report their rescaled values.

of the two factors.

For each sort, let r denote the fraction of the year-0 high-minus-low log cash flow-to-asset spread that remains in year 5. We compare the average log decay rates of the two spreads using $CF^{rel} = \ln r^{MB} / \ln r^{GP/A}$. A value below one indicates that market-to-book cash flow differences decay more slowly than profitability differences. In the data, $CF^{rel} = 0.39$, so the market-to-book spread decays at about 0.39 times the rate of the profitability spread. At these average rates, the same proportional reduction takes roughly two and a half times as long for the market-to-book spread, as measured by $1/CF^{rel}$. The model produces $CF^{rel} = 0.44$, close to the empirical value, and therefore captures the relative persistence of the two spreads.

The two-factor productivity structure also manifests in the investment behavior of portfolios sorted on market-to-book and profitability. In Panel B of Table 3, we report investment rates for both types of portfolios. Consistent with Figure 2, investment increases with both market-to-book ratio and gross profitability. However, the spread between investment rates is significantly larger for market-to-book-sorted portfolios than for profitability-sorted portfolios. This feature is also prominent in our model. In the model the investment spread is 0.10 across market-to-book portfolios and 0.05 across profitability portfolios, against 0.08 and 0.04 in the data, and symmetrically the profitability spread is far larger across profitability portfolios than across market-to-book portfolios, 0.44 against 0.17. Our model generates similar patterns for profitability and investment characteristics across value- and profitability-sorted portfolios.

5.4 Value and profitability premiums

5.4.1 Expected returns

We report the characteristics (top panel), average returns (middle panel), and exposure to fundamental shocks (bottom panel) of portfolios sorted by market-to-book and profitability in Table 4. The moments are computed from a long-sample simulation, t-statistics are omitted. We make three observations. First, our model generates a significant value premium as well as a profitability premium. The expected return decreases monotonically from value to growth, with a levered spread between the value and growth portfolios of 1.5% per year. Similarly, the expected return increases monotonically across gross-profitability-sorted portfolios in Panel B, with a levered spread between high- and low-profitability portfolios of 1.6% per year, as reported in the Model column of Panel B of Table 3. Both premia have the right sign and are of the right order of magnitude. The model generates about one-half of the empirical value premium of 3.30% per year and one-third of the empirical profitability premium of

4.86% per year. It is notoriously challenging for general equilibrium production-based asset pricing models to generate sizable cross-sectional premiums. With macroeconomic dynamics disciplined by the data, both the volatility of the pricing kernel and the prices of aggregate risks are limited. Our unlevered returns are both slightly above 0.75%; they are quantitatively larger than the magnitudes of unlevered returns in Favilukis and Lin (2015), which is also a production-based model in general equilibrium.

Table 4: Cross-Sectional Moments: Model-Simulated Data

This table presents the moments of model-simulated data. Panels A and B report the firm characteristics for portfolios sorted by the market equity-to-physical capital ratio (V/K) and gross profitability (Π/K), respectively. We report the median firm characteristics, including V/K, Π/K, and I/K. Annual excess returns of value-weighted portfolios ($E[R^e]$) are reported in percentages. We also report portfolio returns' exposure to the TFP shock ε^A and depreciation shock ε^θ. All statistics are population moments computed from a long sample of model-simulated data; no t-statistics are reported.

Panel A: Market-to-book-sorted portfolios						
	Value	2	3	4	Growth	Growth-Value
V/K	0.69	0.75	0.80	0.88	1.06	0.37
Π/K	0.16	0.22	0.26	0.31	0.33	0.17
I/K	0.05	0.07	0.09	0.11	0.15	0.10
Returns						
$E[R^e](\%)$	7.57	7.57	7.48	7.18	6.05	-1.52
Risk Exposures to Fundamental Shocks						
ε^A	0.76	0.65	0.53	0.34	0.02	-0.75
ε^θ	-3.46	-3.69	-3.87	-4.05	-3.83	-0.37

Panel B: Gross-profitability-sorted portfolios						
	Low	2	3	4	High	High-Low
V/K	0.74	0.82	0.81	0.86	0.95	0.21
Π/K	0.09	0.14	0.21	0.30	0.53	0.45
I/K	0.07	0.09	0.09	0.10	0.12	0.05
Returns						
$E[R^e](\%)$	6.52	6.53	7.14	7.54	8.12	1.60
Risk Exposures to Fundamental Shocks						
ε^A	0.54	0.41	0.51	0.48	0.36	-0.18
ε^θ	-3.17	-3.44	-3.69	-3.96	-4.65	-1.47

Second, as in the data, the valuation ratio (V/K) and the profitability ratio (Π/K) are positively correlated. We interpret the market-to-book ratio in the data as a proxy for V/K and the gross profitability ratio as a measure of Π/K. In our model, market-to-book sorting primarily loads on the permanent component of firm-level productivity X, while profitability is primarily determined by the transitory component of productivity Z. However, market-to-book also increases with the transitory component Z, and profitability also increases with the permanent component X. As a result, as shown in Panel A, gross profitability generally increases from value to growth portfolios. Similarly, in Panel B, the market-to-book ratio tends to increase across gross-profitability-sorted portfolios.[13]

Third, the bottom rows of Panel A of Table 4 show that exposure to aggregate productivity shocks ε^A declines sharply from value to growth portfolios, while exposure to depreciation shocks ε^θ declines only moderately. Although both aggregate shocks affect future growth through endogenous capital accumulation, the value premium primarily reflects value stocks' higher exposure to aggregate productivity risk, which carries a positive price of risk. In contrast, exposure to aggregate productivity shocks decreases only slightly from low- to high-profitability portfolios, whereas exposure to depreciation shocks becomes substantially and monotonically more negative. Because depreciation risk carries a negative price, the more negative exposure of high-profitability portfolios generates the profitability premium.

The portfolio exposures in Table 4 also help explain the negative correlation between the value and profitability factors. The table reports the market-to-book spread as Growth-minus-Value and the profitability spread as High-minus-Low. Reversing the first spread, the Value-minus-Growth factor has positive exposure to both aggregate productivity and depreciation shocks, whereas the High-minus-Low profitability factor has negative exposure to both shocks, with an especially large negative exposure to depreciation risk. Thus, the two factor portfolios load with opposite signs on the model's primitive aggregate shocks. In the model, the correlation between the two factors is -0.39, compared with -0.52 in the data. The model therefore reproduces the negative correlation and accounts for roughly three quarters of its empirical magnitude.

The two-factor structure of productivity is key to explaining these facts. In single-factor models, both gross profitability and market-to-book would load positively on firm-level productivity, resulting in an almost perfect positive correlation between the market-to-book ratio and profitability. Such models cannot explain the coexistence of value and profitability premiums and the negative correlation between value and profitability factors.

[13]The cross-sectional heterogeneity in V/K in our model is much smaller than the corresponding heterogeneity in the data because market-to-book ratios are noisy proxies for the model object. Several factors may contribute to this noise. In our model, capital K represents physical capital used in production and does not directly correspond to book equity, which varies much more than physical capital in the data.

6 Measuring and testing depreciation shocks

As discussed in the previous section, the profitability premium in the model is mainly driven by the depreciation shock. In this section, we construct an empirical measure of depreciation shocks and use it to test three implications of our model. First, consistent with the model's central prediction, high-profitability portfolios have more negative exposure to depreciation shocks in the data. Second, depreciation risk carries a negative price. Third, we show that depreciation shocks have a more negative impact on investments of high-profitability firms, consistent with their more intensive capital utilization implied by our model.

6.1 Measuring depreciation shocks

Depreciation shocks are difficult to measure directly. In our model, a positive depreciation shock reduces the capital stock available for production. Catastrophic events can cause similar losses by destroying productive capital beyond its normal depreciation and scheduled retirement. This motivates an empirical proxy for depreciation shocks based on losses from catastrophic events.

We obtain data on catastrophic events from NOAA/NCEI's Billion-Dollar Weather and Climate Disasters database, which records U.S. events with total losses exceeding $1 billion.[14] The database covers 1980–2024 and contains roughly 400 events across several hazard types, including floods, hurricanes, droughts, etc., with total direct losses reported in dollars. We aggregate losses across U.S. events and scale them by the national capital stock to construct a proxy for aggregate depreciation shocks. Economic linkages among firms allow disaster-related disruptions to propagate beyond directly affected firms, affecting the aggregate economy.

For each calendar quarter, we sum the CPI-adjusted losses from all events beginning in that quarter and divide this sum by the real stock of nonresidential fixed assets at the end of the previous calendar year. This denominator measures capital stock in place before the losses occur. We convert nominal capital stocks to real terms using the deflator for nonresidential fixed assets.[15] The resulting quarterly series provides a proxy for abnormal capital losses relative to the existing capital stock, denoted ε_t^{cap}.

[14] The data can be accessed here.

[15] Alternative capital-stock denominators produce depreciation-shock series with correlations above 0.99 with the baseline series.

6.2 Asset pricing tests

In our model, high-profitability firms command higher expected returns because they have more negative exposure to negatively priced depreciation risk. We test this mechanism using the empirical measure constructed above.

Risk exposures We first examine the risk exposures of gross-profitability-sorted portfolios to the empirical depreciation-shock measure. The results are reported in Table 5. In Panel A, we report the exposures of profitability-sorted portfolios to the primitive depreciation shock, ε^{θ}, using model-simulated data and controlling for the TFP shock.[16] In Panel B, we report the exposures to abnormal capital loss, ε^{cap}, controlling for both the market factor and the TFP shock. The results are quantitatively robust to including either control individually or no controls.

As discussed in Section 5.4, high-profitability portfolios have more negative exposure to depreciation shocks in the model. Panel A shows the results in the model-simulated data. Panel B shows a similar pattern in the data: exposure to the empirical depreciation shock decreases almost monotonically from low- to high-profitability portfolios. This pattern supports the model's prediction that high-profitability portfolios are more negatively exposed to depreciation shocks.

Price of risk test In our model, a positive depreciation shock destroys productive capital , reduces aggregate consumption and slows down the economic growth. Therefore, such a shock raises the marginal utility of the representative agent. Assets with positive exposure to it provide a hedge against these adverse states and consequently earn lower expected returns, implying a negative price of depreciation risk. We test this prediction using a two-stage Fama–MacBeth procedure.

We implement a two-stage Fama–MacBeth procedure to examine how the empirical depreciation shock is priced in the cross-section of stock returns. Our test assets are the 25 Fama–French portfolios sorted on size and book-to-market and the 10 portfolios sorted on gross profitability. We also use the 25 portfolios sorted on operating profitability and investment from Ken French's website for robustness tests. In the first stage, we estimate each portfolio's exposures to the market factor and the empirical depreciation-shock measure using time-series regressions. In the second stage, for each quarter, we run a cross-sectional regression of portfolio excess returns on the estimated exposures. The estimated price of each risk is the time-series average of the corresponding quarter-specific

[16] The model features two primitive aggregate shocks: TFP and depreciation shocks.

Table 5: Risk exposures

This table presents the risk exposures of gross-profitability-sorted portfolios. The variable MKT_t is the market factor. Panel A reports the regression results from the model-simulated data. The variable ε_t^θ refers to the primitive depreciation shock. We control for the aggregate TFP shock ε^A. The statistics for the model-simulated data are computed from a long simulation and should be interpreted as population moments; the corresponding t-statistics are not reported. In Panel B, we perform the same test at a quarterly frequency using the empirically measured depreciation shock, ε^{cap}, controlling for the aggregate TFP shock and the market factor. Both ε^θ and ε^{cap} are standardized to have a mean of zero and a standard deviation of one. Numbers in parentheses are Newey-West t-statistics adjusted with 3 lags.

Panel A: Risk exposures in the model

	GPA1	GPA2	GPA3	GPA4	GPA5	High-Low
ε^θ	-3.17	-3.44	-3.69	-3.96	-4.65	-1.47
ε^A	0.54	0.41	0.51	0.48	0.36	-0.18

Panel B: Risk exposures in the data

	GPA1	GPA2	GPA3	GPA4	GPA5	High-Low
ε^{cap}	0.56	0.18	0.19	-0.24	-0.56	-1.12
(t)	(2.08)	(1.21)	(0.84)	(-1.29)	(-4.58)	(-3.45)
ε^A	0.04	0.11	-0.03	-0.09	-0.08	-0.13
(t)	(0.51)	(1.53)	(-0.40)	(-1.86)	(-0.90)	(-0.87)
MKT	0.94	0.98	0.98	1.01	0.99	0.04
(t)	(19.09)	(19.29)	(39.17)	(31.81)	(22.67)	(0.51)

slope estimates. Table 6 reports a significantly negative estimated price of depreciation risk in both specifications. Combined with high-profitability portfolios' more negative exposures, this implies higher compensation for depreciation risk, consistent with the model's mechanism for the profitability premium.

6.3 Depreciation shock and firm-level investments

Beyond the two-factor structure of firm-level productivity shocks, one important channel of our model for the profitability premium is the interaction between depreciation shocks and capital utilization decisions. Our model predicts that high-profitability firms utilize their capital more intensively, leading to faster capital depreciation and greater exposure to depreciation shocks. This implies that high-profitability firms' investments should respond more negatively to depreciation shocks than those of low-profitability firms.

Table 6: Price of Risk

This table reports estimated prices of risk from two-stage Fama–MacBeth regressions. The test assets are the 25 Fama–French portfolios sorted on size and book-to-market, the 10 portfolios sorted on gross profitability, and the 25 portfolios sorted on operating profitability and investment. The first-stage time-series regressions estimate each portfolio's exposures to the market factor and the empirical depreciation-shock measure. In the second stage, for each quarter, portfolio excess returns are regressed cross-sectionally on the estimated exposures. The reported prices of risk are the time-series averages of the slope estimates. Returns are in percentage. The depreciation shock measure is standardized to have a mean of zero and a standard deviation of one. The symbol λ denotes the estimated price of risk. Numbers in parentheses are t-statistics. Standard errors are estimated using a Newey–West estimator with a lag of 3. The sample is 1980Q1-2024Q4, at a quarterly frequency.

	MEBM25 + GP10	MEBM25 + OPINV25
$\lambda_{\varepsilon^{cap}}$	-0.87	-1.62
(t)	(-3.29)	(-5.34)
λ_{Mkt}	2.15	2.25
(t)	(3.62)	(3.61)

To test this implication, we regress capital growth rates on the interaction term between gross profitability and the depreciation shock, using both model-simulated and empirical data. The regressions are performed at an annual frequency. We construct the annual version of ε^{cap} by summing the CPI-adjusted losses from events beginning in each calendar year and dividing this sum by the real stock of nonresidential fixed assets at the end of the preceding year. The results are presented in Table 7, with standard errors clustered by firm and year.

The results in Table 7 show that the interaction term between gross profitability and the depreciation shock is negative and statistically significant in both model-simulated and empirical data. This suggests that high-profitability firms' investments are more negatively affected by depreciation shocks, consistent with our model's predictions. The magnitude of the effect is also comparable between the model and the data.

Additionally, we present empirical evidence on the positive correlation between profitability and utilization in Appendix A.2.

7 Conclusion

In this paper, we develop a production-based general equilibrium model with heterogeneous firms to jointly explain the value and profitability premiums. We emphasize the importance of a two-factor structure of productivity shocks to distinguish value and profitability as

Table 7: Depreciation shock and firm-level investments

This table presents the impact of a depreciation shock on firms with different levels of gross profitability. The regression specification is

$$\Delta \ln K_{i,t+1} = a_i + c_{s,t} + \beta GPA_{it} \times \text{Depr Shock}_t + bGPA_{it} + \gamma Ctrl_{it} + e_{i,t},$$

where a_i denotes firm fixed effects, $c_{s,t}$ denotes industry-by-year fixed effects, and $Ctrl_{i,t}$ is a vector of control variables comprising the book-to-market ratio, book leverage, the log of market capitalization, and sales growth. Column (1) replaces industry-by-year fixed effects with year fixed effects. Industries are defined using the Fama–French 12-industry classification. In the data, $\Delta \ln K_{i,t+1}$ is the growth rate of gross property, plant, and equipment (PPEGT). The variable Depr Shock$_t$ denotes the empirical abnormal capital loss measure, ε_t^{cap}, in the data and the primitive depreciation shock, ε_t^{θ}, in the model. Columns (1) and (2) report estimates using the empirical data, while column (3) reports the corresponding regression using model-simulated data. The model statistics are computed from a long simulation and are interpreted as population moments; therefore, the corresponding t-statistics are not reported. The empirical sample covers 1980–2024 at an annual frequency. Numbers in parentheses are t-statistics. Standard are clustered by firm and year.

	Data		Model
	(1)	(2)	(3)
$GPA_{i,t} \times$ Depr Shock$_t$	-0.014	-0.014	-0.011
	(-2.38)	(-2.48)	
Observations	160,061	160,061	3,780,308
R^2	0.205	0.213	0.926
Firm FE	Yes	Yes	Yes
Year FE	Yes	No	Yes
Industry × Year FE	No	Yes	No
Controls	Yes	Yes	Yes
Cluster SE	Firm, Year	Firm, Year	Firm, Year

distinct firm-level characteristics. In addition, our model features two independent sources of aggregate shocks: productivity shocks and capital depreciation shocks. In equilibrium, the value factor endogenously loads more on productivity shocks, and the profitability factor loads more on capital depreciation shocks. Our model not only matches the key features of macroeconomic quantities and the dynamics of the aggregate asset market, but, more importantly, it captures well the coexistence of the value and profitability premiums, as well as the different investment patterns across market-to-book- and profitability-sorted portfolios. Our empirical evidence supports the model's mechanism for the profitability premium.

References

Ai, H. and A. Bhandari (2021): "Asset Pricing With Endogenously Uninsurable Tail Risk," Econometrica, 89, 1471–1505.

Ai, H., M. M. Croce, and K. Li (2013): "Toward a quantitative general equilibrium asset pricing model with intangible capital," Review of Financial Studies, 26, 491–530.

Ai, H. and D. Kiku (2013): "Growth to value: Option exercise and the cross section of equity returns," Journal of Financial Economics, 107, 325–349.

Ai, H., D. Kiku, R. Li, and J. Tong (2021): "A unified model of firm dynamics with limited commitment and assortative matching," The Journal of Finance, 76, 317–356.

Atkeson, A. and P. J. Kehoe (2005): "Modeling and Measuring Organization Capital," Journal of Political Economy, 113, 1026–53.

Bai, H., K. Hou, H. Kung, E. X. Li, and L. Zhang (2019): "The CAPM strikes back? An equilibrium model with disasters," Journal of Financial Economics, 131, 269–298.

Ball, R., J. Gerakos, J. T. Linnainmaa, and V. Nikolaev (2016): "Accruals, cash flows, and operating profitability in the cross section of stock returns," Journal of Financial Economics, 121, 28–45.

Belo, F., X. Lin, and S. Bazdresch (2014): "Labor Hiring, Investment, and Stock Return Predictability in the Cross Section," Journal of Political Economy, 122, 129–177.

Bloom, N. (2009): "The impact of uncertainty shocks," econometrica, 77, 623–685.

Brunnermeier, M. K. and Y. Sannikov (2014): "A macroeconomic model with a financial sector," American Economic Review, 104, 379–421.

Byun, S. K., V. Polkovnichenko, and M. J. Rebello (2019): "Composition of Cash Flow Shocks, Firm Investment, and Cash Holdings," Available at SSRN 2852545.

Chen, A. Y. (2018): "A general equilibrium model of the value premium with time-varying risk premia," The Review of Asset Pricing Studies, 8, 337–374.

Cooper, R. W. and J. C. Haltiwanger (2006): "On the Nature of Capital Adjustment Costs," The Review of Economic Studies, 73, 611–633.

Croce, M. M. (2014): "Long-run productivity risk: A new hope for production-based asset pricing?" Journal of Monetary Economics, 66, 13–31.

Donangelo, A., F. Gourio, M. Kehrig, and M. Palacios (2019): "The cross-section of labor leverage and equity returns," Journal of Financial Economics, 132, 497–518.

Dou, W. W., Y. Ji, and W. Wu (2021a): "Competition, profitability, and discount rates," Journal of Financial Economics, 140, 582–620.

——— (2021b): "The Oligopoly Lucas Tree," The Review of Financial Studies, 35, 3867–3921.

Eisfeldt, A. l. and D. Papanikolaou (2013): "Organization Capital and the Cross-Section of Expected Returns," The Journal of Finance, 68, 1365–1406.

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

Fama, E. F. and K. R. French (2015): "A five-factor asset pricing model," Journal of Financial Economics, 116, 1–22.

Favilukis, J. and X. Lin (2015): "Wage rigidity: A quantitative solution to several asset pricing puzzles," Review of Financial Studies, 29, 148–192.

Favilukis, J., X. Lin, and X. Zhao (2020): "The Elephant in the Room: The Impact of Labor Obligations on Credit Markets," American Economic Review, 110, 1673–1712.

Garlappi, L. and Z. Song (2017): "Capital utilization, market power, and the pricing of investment shocks," Journal of Financial Economics, 126, 447–470.

Gertler, M. and P. Karadi (2011): "A Model of Unconventional Monetary Policy," Journal of Monetary Economics, 58, 17–34.

Gertler, M. and N. Kiyotaki (2010): "Financial intermediation and credit policy in business cycle analysis," Handbook of Monetary Economics, 3, 547–599.

Gomes, J., L. Kogan, and L. Zhang (2003): "Equilibrium Cross Section of Returns," Journal of Political Economy, 111, 693–732.

Gomes, j. f. and l. Schmid (2021): "Equilibrium Asset Pricing with Leverage and Default," The Journal of Finance, 76, 977–1018.

Gourio, F. (2008): "Estimating Firm-Level Risk," working paper.

Greenwood, J., Z. Hercowitz, and G. W. Huffman (1988): "Investment, Capacity Utilization, and the Real Business Cycle," The American Economic Review, 78, 402–417.

Grigoris, F. and G. Segal (2024): "The utilization premium," Management Science, 70, 207–224.

He, Z. and A. Krishnamurthy (2013): "Intermediary Asset Pricing," American Economic Review, 103, 732–70.

Herskovic, B., T. Kind, and H. Kung (2018): "Micro uncertainty and asset prices," Available at SSRN 3220825.

——— (2023): "Micro uncertainty and asset prices," Journal of Financial Economics, 149, 27–51.

Hsieh, C.-T. and P. J. Klenow (2009): "Misallocation and manufacturing TFP in China and India," The Quarterly journal of economics, 124, 1403–1448.

İmrohoroğlu, A. and Ş. Tüzel (2014): "Firm-level productivity, risk, and return," Management Science, 60, 2073–2090.

Jaimovich, N. and S. Rebelo (2009): "Can News about the Future Drive the Business Cycle?" American Economic Review, 99, 1097–1118.

Jorgenson, D. W. (1996): "Empirical studies of depreciation," Economic Inquiry, 34, 24–42.

Kogan, L., J. Li, and H. H. Zhang (2023): "Operating hedge and gross profitability premium," The Journal of Finance, 78, 3387–3422.

Kogan, L. and D. Papanikolaou (2012): "Economic Activity of Firms and Asset Prices," Annual Review of Financial Economics, 4, 1–24.

——— (2013): "Firm Characteristics and Stock Returns: The Role of Investment-Specific Shocks," The Review of Financial Studies, 26, 2718–2759.

——— (2014): "Growth opportunities, technology shocks, and asset prices," Journal of Finance, 69, 675–718.

Kopecky, K. A. and R. M. H. Suen (2010): "Finite State Markov-chain Approximations to Highly Persistent Processes," Review of Economic Dynamics, 13, 701–714.

Kozlowski, J., L. Veldkamp, and V. Venkateswaran (2018): "The Tail that Keeps the Riskless Rate Low," .

Krusell, P. and A. A. Smith, Jr (1998): "Income and wealth heterogeneity in the macroeconomy," Journal of Political Economy, 106, 867–896.

Kuehn, L.-A. and L. Schmid (2014): "Investment-Based Corporate Bond Pricing," The Journal of Finance.

Kuehn, L.-a., M. Simutin, and J. J. Wang (2017): "A Labor Capital Asset Pricing Model," The Journal of Finance, 72, 2131–2178.

Kung, H. and L. Schmid (2015): "Innovation, growth, and asset prices," Journal of Finance, 70, 1001–1037.

Lee, B.-S. and B. F. Ingram (1991): "Simulation estimation of time-series models," Journal of Econometrics, 47, 197–205.

Li, J. (2018): "Explaining Momentum and Value Simultaneously," Management Science, 64, 4239–4260.

Luttmer, E. G. (2007): "Selection, growth, and the size distribution of firms," Quarterly Journal of Economics, 122, 1103–1144.

Malloy, C. J., T. J. Moskowitz, and A. Vissing-Jørgensen (2009): "Long-run stockholder consumption risk and asset returns," The Journal of Finance, 64, 2427–2479.

Novy-Marx, R. (2013): "The other side of value: The gross profitability premium," Journal of Financial Economics, 108, 1–28.

Papanikolaou, D. (2011): "Investment shocks and asset prices," Journal of Political Economy, 119, 639–685.

Romer, P. M. (1986): "Increasing Returns and Long-run Growth," Journal of Political Economy, 94, 1002 – 1037.

Rouwenhorst, K. G. (1995): "Asset Pricing Implications of Equilibrium Business Cycle Models," in Frontiers of Business Cycle Research, ed. by T. F. Cooley, Princeton, NJ: Princeton University Press, 294–330.

Zhang, L. (2005): "The value premium," Journal of Finance, 60, 67–103.

Internet Appendix

A Appendix on empirical evidence

A.1 Data construction

We obtain firm-level balance sheet data from the Compustat Annual Database and monthly stock price data from CRSP for 1963 to 2024. Following standard practice in the literature, we exclude utility firms (SIC codes 4900-4999) and financial firms (SIC codes 6000-6999). We deflate firm-level variables using the GDP deflator from the Bureau of Economic Analysis (BEA). Table A.8 provides detailed variable definitions.

Table A.8: Variable definitions

Variables	Definition	Sources
BE	Book value of equity, computed as the book value of stockholders' equity, plus balance sheet deferred taxes and investment tax credit (if available), minus the book value of preferred stock. Depending on availability, we use the redemption, liquidation, or par value (in that order) as the book value of preferred stock.	Compustat
ME	Market value of equity is computed as price per share times the number of shares outstanding.	Compustat
BM	Book equity divided by the market capitalization.	Compustat
GP/A	Compustat item REVT minus COGS divided by AT.	Compustat
I/K	Compustat item CAPX minus SPPE divided by lagged PPEGT.	Compustat
Operating Profit	Revenue (REVT) less cost of goods sold (COGS) and selling, general, and administrative expenses net of R&D (XSGA-XRD). See Ball et al. (2016).	Compustat

A.2 Profitability and utilization

We provide additional evidence on the relation between firm-level gross profitability and utilization. A direct test at the firm level is not feasible because utilization is observed only by industry and covers a limited set of sectors.[17] We assign each firm the utilization rate of the most detailed industry for which data are available, assuming that firms within an

[17]The available utilization data cover manufacturing (NAICS 31–33), mining (NAICS 21), and electric and gas utilities (NAICS 2211 and 2212), leaving many industries in the Compustat universe uncovered.

industry share the same rate. This evidence is consistent with the model's mechanism: high-profitability firms use their capital more intensively and consequently have greater exposure to depreciation shocks.

Table A.9 reports regressions of firm-level gross profitability on industry-level utilization, controlling for firm fixed effects and other firm characteristics. The positive correlation between utilization and profitability holds across specifications. Alternative industry definitions, such as the Fama–French 12-industry classification, produce quantitatively similar results.

Table A.9: Profitability and utilization

This table reports the regression results of firm-level gross profitability on utilization. The regression specification is

$$GPA_{it} = \alpha_i + c_t + \beta \text{Utilization}_{i,t} + \gamma Z_{it} + e_{i,t},$$

α_i and c_t are firm and year fixed effects. Columns (2) and (3) include lagged gross profitability, the log book-to-market ratio, book leverage, and log total assets; column (1) excludes these controls. Utilization is expressed as a fraction. Standard errors are clustered by firm and year in columns (1) and (2), and by two-digit NAICS industry and year in column (3). Numbers in parentheses are t-statistics. The sample period is 1963–2024.

	(1)	(2)	(3)
Utilization	0.089	0.057	0.057
	(4.014)	(3.781)	(4.437)
Observations	122,616	108,603	108,603
R-squared	0.741	0.842	0.842
Controls	No	Yes	Yes
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Cluster SE	Firm, Year	Firm, Year	Industry, Year

B Proof of Propositions

We first provide a lemma that characterizes the homogeneity property of our economy.

Lemma 1. *(Homogeneity) The value function and policy functions for the firm's maximization problem satisfy the following homogeneity property:*

$$\begin{aligned} V\left(K_t, Z_t, X_t; S_t\right) &= \mathbf{K}_t X_t v\left(k_t, Z_t; m_t, A_t, \theta_t\right), \\ I\left(K_t, Z_t, X_t; S_t\right) &= \mathbf{K}_t X_t i\left(k_t, Z_t; m_t, A_t, \theta_t\right), \\ D\left(K_t, Z_t, X_t; S_t\right) &= \mathbf{K}_t X_t d\left(k_t, Z_t; m_t, A_t, \theta_t\right), \end{aligned} \tag{B1}$$

where $k_t = \frac{K_t}{X_t \mathbf{K}_t}$*, for some normalized value functions and policy functions,* v*,* i*, and* d *defined on a lower-dimensional space. Here* A_t *and* θ_t *are the two exogenous aggregate state variables, aggregate productivity and the depreciation shock. Below we abbreviate the normalized aggregate state as* $s_t \equiv (m_t, A_t, \theta_t)$*. The definition and the law of motion of the summary measure* m *are given in Appendix C.3.*

Proof of Lemma 1 We first establish homogeneity of operating profit. Substituting the spillover relation $\bar{A}_t = A_t \mathbf{K}_t^{1-\alpha\nu}$ from (17) into (16) gives

$$\Pi_t = \alpha A_t \mathbf{K}_t^{1-\alpha\nu} \mathbf{K}_t^{-\nu(1-\alpha)} X_t^{1-\nu} Z_t^{1-\nu} (u_t K_t)^{\nu} = \alpha A_t \mathbf{K}_t^{1-\nu} X_t^{1-\nu} Z_t^{1-\nu} (u_t K_t)^{\nu}. \tag{B2}$$

Using the notation $k_t = \frac{K_t}{X_t \mathbf{K}_t}$,

$$\Pi_t = \alpha A_t X_t^{1-\nu} Z_t^{1-\nu} \left(u_t k_t X_t \mathbf{K}_t\right)^{\nu} \mathbf{K}_t^{1-\nu} = \alpha A_t Z_t^{1-\nu} \left(u_t k_t\right)^{\nu} \mathbf{K}_t X_t,$$

so that Π_t is homogeneous of degree one in $\mathbf{K}_t X_t$, and the corresponding normalized profit flow, per unit of $\mathbf{K}_t X_t$, is

$$\pi(k_t, Z_t; s_t) = \frac{\Pi_t}{\mathbf{K}_t X_t} = \alpha A_t Z_t^{1-\nu} (u_t k_t)^{\nu},$$

which depends on the firm only through (k_t, Z_t).

We next verify that the firm's constraints are homogeneous. Dividing the budget constraint (10) by $\mathbf{K}_t X_t$ and using the degree-one homogeneity of the adjustment cost function H gives

$$d_t + i_t + H(i_t, k_t) + F k_t \leq \pi(k_t, Z_t; s_t),$$

and dividing the capital accumulation equation (7) by $\mathbf{K}_t X_t$ gives

$$\mathbf{G}_{t,t+1}\frac{X_{t+1}}{X_t}\,k_{t+1} = \big(1-\delta(u_t,\theta_{t+1})\big)k_t + i_t, \tag{B3}$$

where

$$\mathbf{G}_{t,t+1} \equiv \frac{\mathbf{K}_{t+1}}{\mathbf{K}_t} \tag{B4}$$

is the gross growth rate of the aggregate capital index (14). Firms' investment and utilization decisions determine it in equilibrium, and each firm takes it as given. The normalized constraints depend on (k_t, i_t, d_t, u_t) and the aggregate state, but not on the scale $\mathbf{K}_t X_t$.

The value function satisfies (B1), where $v\,(k_t, Z_t; s_t)$ is defined as the fixed point of the following Bellman equation:

$$v(k_t, Z_t; s_t) = \max_{i_t, u_t} d_t + (1-\kappa_D)E_t\left[M_{t+1}v(k_{t+1}, Z_{t+1}; s_{t+1})\,\mathbf{G}_{t,t+1}e^{\mu_X+\sigma_X\varepsilon_{t+1}}\right], \tag{B5}$$

where $\mathbf{G}_{t,t+1}$ is the growth rate of the aggregate capital index defined in (B4) and $e^{\mu_X+\sigma_X\varepsilon_{t+1}} = \frac{X_{t+1}}{X_t}$ is the gross growth rate of the firm's permanent productivity component.

Proof of Propositions 1 and 2 We write the firm's optimization problem sequentially in Lagrangian form. We drop the subscript j for simplicity. Throughout the proof, we consider an interior optimum with $u_t > 0$ and $\delta_u > 0$, and we impose the transversality conditions

$$\lim_{s\to\infty} E_t\big[\chi_{t,t+s}M_{t,t+s}\tilde{q}_{t+s}\big] = 0, \qquad \lim_{s\to\infty} E_t\big[(1-\kappa_D)^s M_{t,t+s}\tilde{q}_{t+s}K_{t+s}\big] = 0, \tag{B6}$$

where the survival-adjusted capital-retention factor $\chi_{t,t+s}$ is defined below.

The Lagrangian associated with the dynamic programming problem (9) is

$$\mathcal{L} = E_t\Big[\sum_{s=0}^{\infty}(1-\kappa_D)^s M_{t,t+s}\Big(D_{t+s} + \tilde{q}_{t+s}\left(\left(1-\delta(u_{t+s-1},\theta_{t+s})\right)K_{t+s-1} + I_{t+s-1} - K_{t+s}\right)\Big)\Big],$$

where $D_{t+s} = \Pi_{t+s} - I_{t+s} - H(I_{t+s}, K_{t+s}) - FK_{t+s}$ is the dividend defined by the budget constraint (10) and the Lagrangian multiplier for the capital accumulation equation is denoted as $\tilde{q}_t$. The firm takes the fixed cost as given when choosing investment and utilization, so it does not enter the first-order conditions below. It affects the level of the firm's dividend, and hence the value of the firm, but not its marginal decisions.

- first-order condition for investment I_t

$$1 + H_I(I_t, K_t) = (1 - \kappa_D) E_t\big[M_{t+1} \tilde{q}_{t+1}\big], \tag{B7}$$

- first-order condition for utilization u_t

$$\Pi_{u,t} = (1 - \kappa_D)\delta_u f'(u_t) K_t E_t\big[M_{t+1}\theta_{t+1}\tilde{q}_{t+1}\big], \tag{B8}$$

where $\delta_u f'(u_t)$ is the marginal increase in depreciation exposure implied by (8). Operating profit is homogeneous of degree ν in utilization, so $\Pi_{u,t} = \nu \Pi_t / u_t$. Under the functional form $f(u) = (u^{1+\lambda} - \bar{u}^{1+\lambda})/(1+\lambda)$ introduced in Section 5.2, $f'(u) = u^{\lambda}$. The utilization condition therefore becomes

$$\nu \frac{\Pi_t}{u_t} = (1 - \kappa_D)\delta_u u_t^{\lambda} K_t E_t\big[M_{t+1}\theta_{t+1}\tilde{q}_{t+1}\big]. \tag{B9}$$

- first-order condition for K_{t+1}

$$\tilde{q}_{t+1} = \frac{\partial \Pi_{t+1}}{\partial K_{t+1}} - H_K(I_{t+1}, K_{t+1}) + (1-\kappa_D) E_{t+1}\big[M_{t+1,t+2}(1 - \delta(u_{t+1}, \theta_{t+2}))\tilde{q}_{t+2}\big], \tag{B10}$$

where the marginal product of capital is $MPK_{t+1} = \frac{\partial \Pi_{t+1}}{\partial K_{t+1}}$.

Combining equations (B7) and (B10), we have

$$\begin{aligned} 1 + H_I(I_t, K_t) =& (1 - \kappa_D) E_t\Big\{ M_{t+1}\Big[\frac{\partial \Pi_{t+1}}{\partial K_{t+1}} - H_K(I_{t+1}, K_{t+1}) \\ &+ (1-\kappa_D) E_{t+1}\big[M_{t+1,t+2}\big(1 - \delta(u_{t+1}, \theta_{t+2})\big)\tilde{q}_{t+2}\big]\Big]\Big\}, \end{aligned} \tag{B11}$$

Using the law of iterated expectations to iterate $\tilde{q}$ forward with equation (B10), and applying the first transversality condition in (B6), we obtain

$$1 + H_I(I_t, K_t) = E_t\Big[\sum_{s=1}^{\infty} \chi_{t,t+s} M_{t,t+s} MPK_{t+s}\Big] - E_t\Big[\sum_{s=1}^{\infty} \chi_{t,t+s} M_{t,t+s} H_K(I_{t+s}, K_{t+s})\Big],$$

where

$$\chi_{t,t+s} = (1 - \kappa_D)^s \prod_{\tau=1}^{s-1} \Big(1 - \delta(u_{t+\tau}, \theta_{t+\tau+1})\Big),$$

with the convention that the empty product equals one, so that $\chi_{t,t+1} = 1 - \kappa_D$.

Because the convex adjustment cost function is homogeneous of degree one, $H(I,K) = h(I/K)K$, so that $H_I(I_t, K_t) = h'(I_t/K_t)$ and $H_K(I_{t+s}, K_{t+s}) = h_K(I_{t+s}/K_{t+s})$, with

$h_K(x) = h(x) - xh'(x)$ as defined in Section 4. Moreover, because operating profit is homogeneous of degree ν in capital and $\Pi_{t+s} = \alpha Y_{t+s}$ by (16), the marginal product of capital satisfies $MPK_{t+s} = \nu\Pi_{t+s}/K_{t+s} = \alpha\nu Y_{t+s}/K_{t+s}$. Substituting these expressions into the display above delivers equation (27). This completes the proof of Proposition 1. □

Proof of Proposition 2 The result follows directly from the utilization first-order condition (B9): multiplying both sides by u_t and solving for $u_t^{1+\lambda}$ yields

$$u_t^{1+\lambda} = \frac{\nu\Pi_t}{(1-\kappa_D)E_t[M_{t+1}\theta_{t+1}\tilde{q}_{t+1}]\delta_u K_t}.$$

□

C Normalized recursive equilibrium and the summary measure

We construct the recursive competitive equilibrium in several steps, using normalized value and policy functions.

C.1 Normalization of X

The measure of operating firms is one, so $E[X]$ denotes the cross-sectional mean of permanent productivity. Each period, a mass κ_D of surviving entrants replaces exiting firms. We choose their initial productivity $\bar{X}$ so that the cross-sectional mean remains one when initialized at one:

$$E[X] = \int X_{j,t+1}dj = \int_{\text{survivors}} X_{j,t}e^{\mu_X+\sigma_X\varepsilon_{j,t+1}}dj + \int_{\text{entrants}} \bar{X}e^{\mu_X+\sigma_X\varepsilon_{j,t+1}}dj = 1.$$

This condition determines the initial productivity of entrants, $\bar{X}$. Since the next period's shock, $\varepsilon_{j,t+1}$, is independent of the current level of X_t, the law of iterated expectations gives

$$\begin{aligned}\int X_{j,t+1}dj &= (1-\kappa_D)\int X_{j,t}djE[e^{\mu_X+\sigma_X\varepsilon_{j,t+1}}] + \kappa_D\bar{X}E[e^{\mu_X+\sigma_X\varepsilon_{j,t+1}}] \\ &= (1-\kappa_D)e^{\mu_X+\frac{1}{2}\sigma_X^2}\int X_{j,t}dj + \kappa_D e^{\mu_X+\frac{1}{2}\sigma_X^2}\bar{X}.\end{aligned}$$

In the steady-state equilibrium, this simplifies to:

$$E[X] = (1-\kappa_D)e^{\mu_X+\frac{1}{2}\sigma_X^2}E[X] + \kappa_D e^{\mu_X+\frac{1}{2}\sigma_X^2}\bar{X}. \tag{C12}$$

To ensure that $E[X] = 1$, we impose:

$$\bar{X} = \frac{1 - (1 - \kappa_D)e^{\mu_X + \frac{1}{2}\sigma_X^2}}{\kappa_D e^{\mu_X + \frac{1}{2}\sigma_X^2}}.$$

Similarly, in our calibration, we set $E[\ln Z] = 0$. These restrictions are particularly useful when introducing the summary measure in Appendix C.

C.2 Labor market clearing

Labor is chosen within the period, and aggregate labor supply is fixed. We can therefore solve the labor allocation and impose market clearing before solving the firm's dynamic problem. We use the normalized capital stock

$$k_{j,t} = \frac{K_{j,t}}{X_{j,t}\mathbf{K}_t}, \tag{C13}$$

the same normalization used in Lemma 1 and in Section 4, where $\mathbf{K}_t$ is the aggregate capital index defined in (14).

Firms' optimality condition with respect to labor implies that the choice of labor must be proportional to $X_{j,t}^{1-\nu} Z_{j,t}^{1-\nu} (u_{j,t} K_{j,t})^{\nu}$. Imposing labor market clearing and using the definition of the aggregate capital index (14), firm-level labor input is

$$L_{j,t} = \frac{X_{j,t}^{1-\nu} Z_{j,t}^{1-\nu} (u_{j,t} K_{j,t})^{\nu}}{\int X_{i,t}^{1-\nu} Z_{i,t}^{1-\nu} (u_{i,t} K_{i,t})^{\nu} di} = X_{j,t}^{1-\nu} Z_{j,t}^{1-\nu} (u_{j,t} K_{j,t})^{\nu} \mathbf{K}_t^{-\nu},$$

where $S_t = (A, \theta, \Gamma)$ is the vector of aggregate state variables. Substituting $L_{j,t}$ into the firm-level production function (4) and writing $K_{j,t} = k_{j,t} X_{j,t} \mathbf{K}_t$, firm-level output and operating profit become

$$Y_{j,t} = A_t Z_{j,t}^{1-\nu} (u_{j,t} k_{j,t})^{\nu} X_{j,t} \mathbf{K}_t, \qquad \Pi_{j,t} = \alpha A_t Z_{j,t}^{1-\nu} (u_{j,t} k_{j,t})^{\nu} X_{j,t} \mathbf{K}_t,$$

where we have used the externality (17). Both are homogeneous of degree one in $X_{j,t}\mathbf{K}_t$, which is the property Lemma 1 exploits, and the normalized profit flow per unit of $X_{j,t}\mathbf{K}_t$ is

$$\pi(k_{j,t}, Z_{j,t}; s_t \| u_{j,t}) = \frac{\Pi_{j,t}}{X_{j,t}\mathbf{K}_t} = \alpha A_t Z_{j,t}^{1-\nu} (u_{j,t} k_{j,t})^{\nu}. \tag{C14}$$

Integrating firm-level output and using (14) gives $Y_t = A_t \mathbf{K}_t$, as in (18). Thus output per unit of the aggregate capital index, $Y_t / \mathbf{K}_t = A_t$, is stationary.

C.3 Construction of normalized recursive equilibrium

We write a state variable as K, its state space as $\mathbb{K}$, and a subset of that space as $\boldsymbol{K}$. In the original formulation, the aggregate state is $S = (A, \theta, \Gamma)$, where Γ is the distribution of firms over (K, Z, X). Homogeneity allows us to replace Γ with a measure m on (k, Z) and write the normalized aggregate state as $s = (m, A, \theta)$. We construct this equilibrium following Ai and Bhandari (2021) and Ai et al. (2021).

The homogeneity properties in Lemma 1 allow us to express the Bellman equations (9) and (22) in terms of normalized value functions. For incumbent firms, the profit maximization problem is

$$v(k_t, Z_t; s_t) = \max_{i_t, u_t} d_t + (1 - \kappa_D) E_t \left[M_{t+1} v(k_{t+1}, Z_{t+1}; s_{t+1}) \, \mathbf{G}_{t,t+1} e^{\mu_X + \sigma_X \varepsilon_{t+1}} \right], \tag{C15}$$

subject to

$$d_t + i_t + H(i_t, k_t) + F k_t \leq \pi(k_t, Z_t; s_t), \tag{C16}$$

$$\mathbf{G}_{t,t+1} e^{\mu_X + \sigma_X \varepsilon_{t+1}} k_{t+1} = (1 - \delta(u_t, \theta_{t+1})) k_t + i_t, \tag{C17}$$

where $\mathbf{G}_{t,t+1} = \mathbf{K}_{t+1}/\mathbf{K}_t$ is the endogenous gross growth rate of the aggregate capital index defined in (B4), π is the normalized profit flow (C14), and the depreciation function δ is given by equation (8).

For entrants, we normalize the value function and capital policy by $\bar{X}\mathbf{K}_t$, so that $\bar{v}(s_t) = \bar{V}(S_t)/(\bar{X}\mathbf{K}_t)$ and $\bar{k}_t = \bar{K}(S_t)/(\bar{X}\mathbf{K}_t)$. Since entry costs are homogeneous of degree one, we can write $\bar{H}(\bar{K}_t, \mathbf{K}_t) = \bar{X}\mathbf{K}_t \bar{H}(\bar{k}_t)$. The entrant's Bellman equation is

$$\bar{v}(s_t) = \max_{\bar{k}_t} -A_t \left[\bar{k}_t + \bar{H}(\bar{k}_t) \right] + (1 - \kappa_D) E \left[M_{t+1} v(k'_{t+1}, Z_{t+1}; s_{t+1}) \, \mathbf{G}_{t,t+1} e^{\mu_X + \sigma_X \varepsilon_{t+1}} \right], \tag{C18}$$

where the entrant's normalized capital in the following period is

$$k'_{t+1} = \frac{\bar{K}_t}{X_{t+1}\mathbf{K}_{t+1}} = \frac{\bar{k}_t}{\mathbf{G}_{t,t+1} e^{\mu_X + \sigma_X \varepsilon_{t+1}}}.$$

We next write the aggregate resource constraint (25) as integrals with respect to the summary measure m. The measure $\Gamma(K, Z, X)$ in equation (25) is defined on the state space of (K, Z, X), $\mathbb{K} \times \mathbb{Z} \times \mathbb{X}$. It induces a measure, $\tilde{\Gamma}(k, Z, X)$ on the state space of (k, Z, X) via the following relationship. For any subset $\boldsymbol{k} \times \boldsymbol{Z} \times \boldsymbol{X}$ of the product space of (k, Z, X), $\boldsymbol{k} \times \boldsymbol{Z} \times \boldsymbol{X} \subseteq \mathbb{k} \times \mathbb{Z} \times \mathbb{X}$,

$$\int_{\boldsymbol{k} \times \boldsymbol{Z} \times \boldsymbol{X}} \tilde{\Gamma}(k, Z, X) d(k, Z, X) = \int_{\mathbb{K} \times \mathbb{Z} \times \mathbb{X}} I_{\{\frac{K}{X\mathbf{K}} \in \boldsymbol{k}, Z \in \boldsymbol{Z}, X \in \boldsymbol{X}\}} \Gamma(K, Z, X) d(K, Z, X).$$

Given $\tilde{\Gamma}$, the summary measure m is defined as a measure on $\mathbb{k} \times \mathbb{Z}$, for any $\boldsymbol{k} \times \boldsymbol{Z} \subseteq \mathbb{k} \times \mathbb{Z}$,

$$\int_{\boldsymbol{k}\times\boldsymbol{Z}} m(k,Z)d(k,Z) = \int_{k\in\boldsymbol{k},Z\in\boldsymbol{Z},X\in\mathbb{X}} X\tilde{\Gamma}(k,Z,X)d(k,Z,X). \tag{C19}$$

The measure $m(k,Z)$ is the total permanent productivity X of firms at (k,Z). It is defined on a lower-dimensional space than $\tilde{\Gamma}(k,Z,X)$, but retains all the information in $\Gamma(K,Z,X)$ needed for aggregation because policy functions are homogeneous with respect to productivity.

To express the aggregate resource constraint in terms of m instead of $\Gamma(K,Z,X)$, we first use the definition of $\tilde{\Gamma}$ and divide (25) by the aggregate capital index $\mathbf{K}$. This gives

$$\begin{aligned} c(s) + \int \Big(i(k,Z;s) + H(i(k,Z;s),k) + Fk\Big) X\tilde{\Gamma}(k,Z,X)d(k,Z,X) \\ + \bar{\phi}A\Big(\bar{k}(s) + \bar{H}(\bar{k}(s))\Big)\bar{X} \\ = A\Big[\int u(k,Z;s)^{\nu}k^{\nu}Z^{1-\nu}X\tilde{\Gamma}(k,Z,X)d(k,Z,X)\Big], \end{aligned} \tag{C20}$$

where $c(s) = \frac{C}{\mathbf{K}}$ is the normalized aggregate consumption, $\bar{\phi} = \kappa_D/(1-\kappa_D)$ is the pre-survival measure of new entrants, and $\bar{k}(s) = \frac{\bar{K}(S)}{\bar{X}\mathbf{K}}$ is the normalized policy function for initial entrants. The bracketed term on the right-hand side is the aggregator identity (D26), which equals one, so normalized aggregate output per unit of the capital index is simply A, consistent with (18). We can compute the integrals in the above equation by first integrating over X.[18] For example, the cost of aggregate investment in the above equation can be computed as

$$\begin{aligned} &\int \Big(i(k,Z;s) + H(i(k,Z;s),k) + Fk\Big)\Big[X\tilde{\Gamma}(k,Z,X)dX\Big]d(k,Z) \\ &= \int \Big(i(k,Z;s) + H(i(k,Z;s),k) + Fk\Big)m(k,Z)d(k,Z), \end{aligned}$$

using the definition of the summary measure in (C19). Using the same computation, the

[18] We set the new entrants' initial productivity $\bar{X}$ such that $E[X] = \int X_{i,t}di = 1$. It is discussed in Appendix C.1.

aggregate resource constraint (C20) can be written as

$$\begin{aligned} c(s) + \int \Big(i(k,Z;s) + H(i(k,Z;s),k) + Fk\Big) m(k,Z) d(k,Z) \\ + \bar{\phi} A \Big(\bar{k}(s) + \bar{H}(\bar{k}(s))\Big) \bar{X} \\ = A\Big[\int u(k,Z;s)^{\nu} k^{\nu} Z^{1-\nu} m(k,Z) d(k,Z)\Big] = A, \end{aligned}$$

which is the resource constraint (D32) used in the numerical solution.

We next derive the transition of m using density notation. Integrating permanent productivity out of $\tilde{\Gamma}$ gives, for each $(k,Z) \in \Bbbk \times \mathbb{Z}$,

$$m\left(k,Z\right) = \int X \tilde{\Gamma}(k,Z,X) dX.$$

For any $(k',Z') \in \Bbbk \times \mathbb{Z}$, a firm in that state next period is either an incumbent previously at (k,Z) or a new entrant previously at $(\bar{k},\bar{Z})$. We derive the two contributions to m' separately.

Consider a firm that goes from (k,Z) to (k',Z'). The law of motion of k in Equation (C17) implies that $\ln k' + \ln \mathbf{G}' + \mu_X + \sigma_X \varepsilon' = \ln g(k,Z;s)$, where $\mathbf{G}' = \mathbf{K}'/\mathbf{K}$ and, conditional on the realized θ', $g(k,Z;s)$ is defined as

$$g(k,Z;s) = \left(1 - \delta\left(u(k,Z;s),\theta'\right)\right) k + i(k,Z;s).$$

That is to say, (k',Z') comes from (k,Z) if and only if Z' comes from Z and $\varepsilon' = \frac{1}{\sigma_X}\left\{\ln g(k,Z;s) - \left[\ln k' + \ln \mathbf{G}' + \mu_X\right]\right\}$. Because ε' is a standard normal random variable, the conditional transition density of k' is $\frac{1}{\sigma_X k'} \phi\left(\varepsilon(k', g(k,Z;s))\right) P(Z'|Z)$, where ϕ is the density of the standard normal distribution, $P(Z'|Z)$ is the transition density or probability mass of Z', and $\varepsilon(k',g)$ is defined as

$$\varepsilon(k',g) = \frac{1}{\sigma_X}\left[\ln g - \ln k' - \ln \mathbf{G}' - \mu_X\right].$$

In addition, conditional on this transition, X' equals $X \times e^{\mu_X + \sigma_X \varepsilon(k', g(k,Z;s))}$. Because incumbent firms survive to the next period with probability $1 - \kappa_D$, their contribution to $m'(k',Z')$ is therefore weighted by $(1-\kappa_D)\frac{1}{\sigma_X k'} \phi\left(\varepsilon(k', g(k,Z;s))\right) P(Z'|Z) e^{\mu_X + \sigma_X \varepsilon(k', g(k,Z;s))}$.

For a firm that entered in the previous period, initial capital is $\bar{K}$ and permanent productivity is $\bar{X}$. Thus $\ln k' + \ln \mathbf{G}' + \mu_X + \sigma_X \varepsilon' = \ln \bar{k}$, where $\bar{k} = \frac{\bar{K}}{\bar{X}\mathbf{K}}$. The entrant reaches (k',Z') if and only if $\varepsilon' = \varepsilon(k',\bar{k})$ and Z' comes from $\bar{Z}$. Its transition density is $\frac{1}{\sigma_X k'} \phi\left(\varepsilon(k',\bar{k})\right) P(Z'|\bar{Z})$. A fraction κ_D of incumbents exits each period, so the

stationary mass of surviving entrants is also κ_D. Their total permanent productivity at this location is $\kappa_D \frac{1}{\sigma_X k'} \phi\left(\varepsilon(k', \bar{k})\right) P(Z'|\bar{Z}) \bar{X} e^{\mu_X + \sigma_X \varepsilon(k', \bar{k})}$. Combining the incumbent and entrant contributions gives

$$\begin{aligned} m'(k', Z') =& (1 - \kappa_D) \int \frac{\phi\left(\varepsilon(k', g(k, Z; s))\right)}{\sigma_X k'} P(Z'|Z) e^{\mu_X + \sigma_X \varepsilon(k', g(k, Z; s))} m(k, Z) dk dZ \\ &+ \kappa_D \frac{\phi\left(\varepsilon(k', \bar{k})\right)}{\sigma_X k'} P(Z'|\bar{Z}) \bar{X} e^{\mu_X + \sigma_X \varepsilon(k', \bar{k})}. \end{aligned} \tag{C21}$$

Equation (C21) is written for the density with respect to dk', which accounts for the Jacobian $1/(\sigma_X k')$. In the numerical solution, we instead express the distribution in log coordinates and reuse the notation m. The change-of-variables Jacobians are then absorbed into the log-coordinate density, yielding the normal-mixture representation in Equation (D23).

A recursive competitive equilibrium consists of the following functions:

- A value function for the representative household's utility $\mathcal{U}(s)$ and the associated consumption policy function $c(s)$, where $\mathcal{U}(s) = \frac{U(S)}{\mathbf{K}(S)}$ and $c(s) = \frac{C(S)}{\mathbf{K}(S)}$.
- A value function for incumbent firms, $v(k, Z; s)$, and the associated policy functions, $i(k, Z; s)$ and $u(k, Z; s)$.
- A value function for new firms, $\bar{v}(s)$, and the associated policy function, $\bar{k}(s)$, where $\bar{v}(s) = \frac{\bar{V}(S)}{\bar{X}\mathbf{K}(S)}$ and $\bar{k}(s) = \frac{\bar{K}(S)}{\bar{X}\mathbf{K}(S)}$.
- An aggregate growth factor $\mathbf{G}(s'|s) = \frac{\mathbf{K}(S')}{\mathbf{K}(S)}$, which is an equilibrium object determined jointly with firms' investment and utilization policies through the definition of the aggregate capital index (14).
- A stochastic discount factor, $M(s'|s)$, which is given by

$$M(s'|s) = \beta \left(\mathbf{G}(s'|s) \frac{c\left(s'\right)}{c\left(s\right)} \right)^{-\frac{1}{\psi}} \left(\frac{\left(\mathbf{G}(s'|s)\, \mathcal{U}\left(s'\right)\right)^{1-\gamma}}{E\left[\left(\mathbf{G}(s'|s)\, \mathcal{U}\left(s'\right)\right)^{1-\gamma} \middle| s\right]} \right)^{\frac{1/\psi - \gamma}{1-\gamma}},$$

  where the normalized utility index solves

$$\mathcal{U}(s) = \left\{ (1-\beta) c(s)^{1-\frac{1}{\psi}} + \beta \left(E\left[\left(\mathbf{G}(s'|s)\, \mathcal{U}(s')\right)^{1-\gamma} \middle| s \right] \right)^{\frac{1-1/\psi}{1-\gamma}} \right\}^{\frac{1}{1-1/\psi}}.$$

  The growth factor $\mathbf{G}(s'|s)$ enters both components because utility and consumption

are normalized by $\mathbf{K}_t$ today and by $\mathbf{K}_{t+1}$ tomorrow. These are equations (D34) and (D33) in the numerical solution.

- A law of motion of the summary measure m, that is given by (C21)

such that the following conditions are met:

1. Given the equilibrium stochastic discount factor and law of motion of summary measure m, the value function $v(k, Z; s)$ and the policy functions, $i(k, Z; s)$ and $u(k, Z; s)$, solve the normalized maximization problem (C15), (C16), and (C17) for incumbent firms.
2. The value function $\bar{v}(s)$ and the policy function $\bar{k}(s)$ solve the profit maximization problem for new entrants (C18).
3. The equilibrium SDF is consistent with household consumption.
4. The law of motion of summary measure (C21) is consistent with firms' optimal policies.
5. Market clearing conditions (C20) hold.
6. The aggregate growth factor $\mathbf{G}(s'|s)$ is consistent with firms' policies, in the sense that the aggregator identity (D26) holds at every aggregate state.
7. Labor market clearing gives rise to the optimal labor allocation across firms.

D Numerical solutions

D.1 The solution method

We solve the normalized recursive equilibrium in Appendix C.3 globally. The model's asset pricing implications depend on the market prices of the two aggregate risks. With endogenous growth, aggregate shocks have permanent effects on the level of the economy through the capital index $\mathbf{K}_t$. Risk premiums and their variation over time are nonlinear functions of the aggregate state, particularly under recursive preferences with high risk aversion. We therefore compute all equilibrium objects as fully nonlinear functions on a discretized state space spanning the economy's ergodic range.

We evaluate expectations over firm-specific states by numerical quadrature, which provides accurate integration over the large conditional volatilities of firm-specific shocks in our calibration.

The summary measure m is an infinite-dimensional state variable. We approximate it with a parsimonious parametric family that permits efficient and accurate integration, and use equilibrium restrictions in each period to determine its parameters. This keeps the aggregate state low-dimensional and gives its transition in closed form. We can therefore compute the equilibrium fixed point without the intermediate simulation step used in Krusell and Smith (1998).

Throughout this section, we work with the normalized equilibrium objects of Lemma 1: firm-specific state variables are (k, Z) and $\mathbf{G}(s'|s)$ denotes the endogenous gross growth rate of the aggregate capital index, (B4), between today's aggregate state s and tomorrow's state s'. The summary measure m is defined on $(\ln k, \ln Z)$, and the normalization of $\bar{X}$ in Appendix C.1 guarantees that m integrates to one, so that we can treat it as a probability measure. We denote by $E_m[\cdot]$ the integral against m and by $\phi_\sigma(\cdot)$ the density of a normal distribution with mean zero and standard deviation σ.

The law of motion of m and the aggregate state vector We first show that the exact law of motion of m is a mixture of normal distributions. Conditional on survival and on the realization of (θ', s'), an incumbent firm at (k, Z) today moves to

$$\ln k' = \ln g(k, Z; s) - \ln \mathbf{G}(s'|s) - \mu_X - \sigma_X \varepsilon', \tag{D22}$$

where

$$g(k, Z; s) = (1 - \delta(u(k, Z; s), \theta'))\, k + i(k, Z; s)$$

is the end-of-period capital implied by the policy functions, and ε' is the innovation to the firm's permanent productivity. Because m weights firms by $X' = Xe^{\mu_X + \sigma_X \varepsilon'}$, this weighting tilts the transition density of $\ln k'$, which, after completing the square, shifts its mean by $-\sigma_X^2$ and scales the total mass by $e^{\mu_X + \frac{1}{2}\sigma_X^2}$:

$$\begin{aligned} m'(\ln k', Z') = & e^{\mu_X + \frac{1}{2}\sigma_X^2} \Bigg\{ (1 - \kappa_D) \sum_Z P(Z'|Z) \int \phi_{\sigma_X} \Big(\ln k' - \ln g(k, Z; s) \\ & + \ln \mathbf{G}(s'|s) + \mu_X + \sigma_X^2 \Big) m(\ln k, Z)\, d \ln k \\ & + \kappa_D \bar{X} P(Z'|\bar{Z})\, \phi_{\sigma_X} \left(\ln k' - \ln \bar{k}(s) + \ln \mathbf{G}(s'|s) + \mu_X + \sigma_X^2 \right) \Bigg\}, \end{aligned} \tag{D23}$$

where the second line accounts for new entrants, who enter with normalized capital $\bar{k}(s)$ and initial productivity $(\bar{X}, \bar{Z})$. Integrating (D23) over $(\ln k', Z')$ and using the normalization of $\bar{X}$ in Appendix C.1 confirms that m' integrates to one. The entrants' share of the total mass is

$$\lambda_e \equiv \kappa_D \bar{X} e^{\mu_X + \frac{1}{2}\sigma_X^2}. \tag{D24}$$

Equation (D23) expresses m' as a mixture of normal distributions in $\ln k$, with common variance σ_X^2. This motivates approximating the distribution of $(\ln k, \ln Z)$ under m with a bivariate normal density:

$$(\ln k, \ln Z) \sim N\left(\begin{bmatrix} \mu_k \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_k^2 & 0 \\ 0 & \tau_Z^2 \end{bmatrix} \right). \tag{D25}$$

Because the marginal distribution of $\ln Z$ is stationary, we fix its mean at zero and its variance at the unconditional value, $\tau_Z^2 = \frac{\sigma_Z^2}{1-\rho_Z^2}$, and we set the correlation between $\ln k$ and $\ln Z$ to zero.[19] The cross-sectional mean μ_k is the only unrestricted moment of this family and is an endogenous aggregate state variable. The dispersion τ_k is determined in each period by the exact aggregation identity derived below, so it does not add a state variable. The normalized economy therefore has three aggregate state variables,

$$s = (A, \theta, \mu_k).$$

The approximation summarizes the infinite-dimensional firm distribution with the single endogenous state μ_k. This allows us to solve globally for equilibrium objects on a finite-dimensional grid.

The cross-sectional projection of utilization and the normalization identity The definition of the aggregate capital index, (14), divided by $\mathbf{K}_t^\nu$, implies the identity

$$\int Z^{1-\nu} \left(u(k, Z; s)\, k\right)^\nu m(\ln k, \ln Z)\, d\ln k\, d\ln Z = 1, \qquad \forall s. \tag{D26}$$

This restriction must hold at every date in equilibrium. To impose it in closed form, we project the log utilization policy on the firm-level state variables. For each aggregate state s, we approximate

$$\ln u(k, Z; s) \approx a_u(s) + b_k(s) \ln k + b_Z(s) \ln Z, \tag{D27}$$

[19] Strictly, because entrants are born at $\bar{Z}$, the cross-sectional moments of $\ln Z$ under m differ slightly from the unconditional moments of the AR(1) process, and a small correlation between $\ln k$ and $\ln Z$ arises in equilibrium. Both corrections are constants scaled by the entrant share λ_e, which is of the order of $\kappa_D \bar{X}$, and are quantitatively negligible.

where the coefficients (a_u, b_k, b_Z) are obtained from a weighted least squares regression of $\ln u$ on $(1, \ln k, \ln Z)$ with the density m as the weight, so that the projection is most accurate where firms concentrate. In our estimated economy this projection fits the utilization policy closely, with an m-weighted R^2 above 0.99 at every aggregate state. Defining the composite elasticities

$$B_K(s) = \nu\,(1 + b_k(s))\,, \qquad B_Z(s) = 1 - \nu + \nu b_Z(s), \tag{D28}$$

the integrand in (D26) is lognormal under (D25), and the identity reduces to

$$\nu a_u(s) + B_K(s)\mu_k + \frac{1}{2}B_K(s)^2\tau_k^2 + \frac{1}{2}B_Z(s)^2\tau_Z^2 = 0, \tag{D29}$$

We solve this equation for τ_k^2 at each aggregate state. Exact aggregation determines the second moment of the firm distribution, with no separate law of motion to approximate. Approximation errors in cross-sectional dispersion therefore do not accumulate over time.

Closed forms for the growth rate and the law of motion of μ_k The same projection delivers the endogenous growth rate of the economy in closed form. Applying the definition (14) at $t+1$, dividing by $\mathbf{K}_{t+1}^{\nu}$, and substituting $K' = g(k, Z; s)X\mathbf{K}$ for incumbents and $\bar{K} = \bar{k}(s)\bar{X}\mathbf{K}$ for entrants, the growth factor solves

$$\begin{aligned}\mathbf{G}(s'|s)^{B'_K} = {} & e^{\nu a'_u + \frac{1}{2}(B'_Z)^2\sigma_Z^2}\, E_\varepsilon\left[e^{(1-B'_K)(\mu_X + \sigma_X\varepsilon)}\right] \\ & \times\left\{(1-\kappa_D)E_m\left[g(k,Z;s)^{B'_K} e^{B'_Z\rho_Z \ln Z}\right] + \kappa_D \bar{X}\,\bar{k}(s)^{B'_K} e^{B'_Z\rho_Z \ln \bar{Z}}\right\},\end{aligned} \tag{D30}$$

where (a'_u, B'_K, B'_Z) denote the projection coefficients evaluated at the future state s', the factor $e^{\frac{1}{2}(B'_Z)^2\sigma_Z^2}$ together with $e^{B'_Z\rho_Z \ln Z}$ is the conditional lognormal moment $E[Z'^{B'_Z}|Z]$, and the expectation over ε integrates out the permanent productivity innovation. The growth factor depends on the entire cross-section of investment and utilization decisions through g, on entry through $\bar{k}(s)$, and on the realized depreciation shock θ' through the survival of installed capital.

Taking the first moment of the mixture (D23) gives the law of motion of the endogenous aggregate state variable μ_k in closed form:

$$\mu'_k(s, s') = (1-\lambda_e)E_m\left[\ln g(k, Z; s)\right] + \lambda_e \ln \bar{k}(s) - \ln \mathbf{G}(s'|s) - \mu_X - \sigma_X^2. \tag{D31}$$

Both (D30) and (D31) are exact conditional on the normal representation of m and the projection (D27); no simulation is required to compute the transition of the aggregate state.

The household block Under the normalization of Lemma 1, aggregate consumption satisfies $C_t = c(s_t)\mathbf{K}_t$. Dividing the aggregate resource constraint (25) by $\mathbf{K}_t$ and using the identity (D26), which sets normalized aggregate output equal to A,

$$c(s) = A - E_m\left[i(k, Z; s) + H\left(i(k, Z; s), k\right) + Fk\right] - A\bar{\phi}\bar{X}\left[\bar{k}(s) + \bar{H}\left(\bar{k}(s)\right)\right]. \tag{D32}$$

Similarly, the household's utility index inherits the homogeneity of the economy, $U_t = \mathcal{U}(s_t)\mathbf{K}_t$, and the recursion (1) becomes

$$\mathcal{U}(s) = \left\{(1-\beta)c(s)^{1-\frac{1}{\psi}} + \beta\left(E\left[\left(\mathbf{G}(s'|s)\,\mathcal{U}(s')\right)^{1-\gamma}\middle|s\right]\right)^{\frac{1-1/\psi}{1-\gamma}}\right\}^{\frac{1}{1-1/\psi}}, \tag{D33}$$

where the growth factor $\mathbf{G}(s'|s)$ scales the continuation utility because utility is normalized by $\mathbf{K}_t$ today and by $\mathbf{K}_{t+1}$ tomorrow. The stochastic discount factor (3) is then

$$M(s'|s) = \beta\left(\mathbf{G}(s'|s)\frac{c(s')}{c(s)}\right)^{-\frac{1}{\psi}}\left(\frac{\left(\mathbf{G}(s'|s)\,\mathcal{U}(s')\right)^{1-\gamma}}{E\left[\left(\mathbf{G}(s'|s)\,\mathcal{U}(s')\right)^{1-\gamma}|s\right]}\right)^{\frac{1/\psi-\gamma}{1-\gamma}}. \tag{D34}$$

The growth factor enters both consumption growth and continuation utility in the pricing kernel. Firms' investment and utilization choices affect this factor and hence the market prices of aggregate risk.

The firm block Given the pricing kernel and the transition of the aggregate state, we solve the firm's optimality conditions on a grid of the firm-level state variables. The normalized capital transition, Equation (B3), implies that tomorrow's normalized capital is

$$\hat{k}' = \frac{g(k, Z; s)}{\mathbf{G}(s'|s)\,e^{\mu_X + \sigma_X \varepsilon'}}. \tag{D35}$$

Both first-order conditions depend on the shadow price of capital $\tilde{q}$, defined recursively in (B10). Substituting the date-$t+1$ conditions (B7) and (B9) into the envelope condition (B10) eliminates the continuation expectation. This gives $\tilde{q}$ in closed form as a function of date-$t+1$ policies, without numerically differentiating the value function. Writing $\iota(k, Z; s) \equiv i(k, Z; s)/k$ for the investment rate and using $H(I, K) = h\left(I/K\right)K$, we obtain

$$\begin{aligned}\tilde{q}(k', Z'; s') = {} & \alpha\nu A' Z'^{1-\nu} u'^{\nu} k'^{\nu-1} - \left(\delta_1 + \delta_u f(u')\right)\frac{\alpha\nu}{\delta_u} A' Z'^{1-\nu} u'^{\nu-1-\lambda} k'^{\nu-1} \\ & - h_K\left(\iota'\right) + (1-\delta_0)\left[1 + h'\left(\iota'\right)\right],\end{aligned} \tag{D36}$$

where $u' = u(k', Z'; s')$ and $\iota' = \iota(k', Z'; s')$ are evaluated by interpolation, and $h_K(\iota) = h(\iota) - \iota h'(\iota)$ as in Section 4. The first term of (D36) is the marginal product of capital; the second term prices the marginal depreciation induced by utilization, using the date-$t+1$ utilization condition; the third term is the adjustment-cost saving; and the fourth term is the value of undepreciated capital, using the date-$t+1$ investment condition.

Given $\tilde{q}$, the firm's optimality conditions on the grid read

$$1 + h'\left(\iota(k, Z; s)\right) = (1 - \kappa_D) E\left[M(s'|s)\, \tilde{q}(\hat{k}', Z'; s') \,\middle|\, k, Z, s\right], \tag{D37}$$

$$\alpha\nu A\, Z^{1-\nu} u(k, Z; s)^{\nu-1-\lambda} k^{\nu-1} = \delta_u (1 - \kappa_D) E\left[M(s'|s)\, \theta' \tilde{q}(\hat{k}', Z'; s') \,\middle|\, k, Z, s\right], \tag{D38}$$

which are the numerical counterparts of Propositions 1 and 2. The conditional expectations integrate over the transition of the exogenous aggregate shocks (A', θ'), the transitory productivity transition $P(Z'|Z)$, and the permanent productivity innovation ε' by Gaussian quadrature. Both conditions can be inverted for the updated policies in closed form. The entrant's optimality condition takes the same form,

$$A\Big(1 + \bar{H}'\left(\bar{k}(s)\right)\Big) = (1 - \kappa_D) E\left[M(s'|s)\, \tilde{q}\left(\frac{\bar{k}(s)}{\mathbf{G}(s'|s) e^{\mu_X + \sigma_X \varepsilon'}}, Z'; s'\right) \Big| s\right], \tag{D39}$$

with Z' drawn from $P(\cdot|\bar{Z})$.

Finally, once the policies have converged, the normalized firm value function is computed by iterating the Bellman equation (B5),

$$v(k, Z; s) = d(k, Z; s) + (1 - \kappa_D) E\left[M(s'|s)\, \mathbf{G}(s'|s)\, e^{\mu_X + \sigma_X \varepsilon'}\, v(\hat{k}', Z'; s') \,\middle|\, k, Z, s\right], \tag{D40}$$

with $d(k, Z; s) = \pi(k, Z; s) - i(k, Z; s) - H\left(i(k, Z; s), k\right) - Fk$, holding the converged policies fixed. Firm-level returns and the value and profitability portfolio sorts reported in Section 5 are constructed from v and the associated one-period return $(1 - \kappa_D)\, \mathbf{G}(s'|s)\, e^{\mu_X + \sigma_X \varepsilon'} v(\hat{k}', Z'; s') / \left[v(k, Z; s) - d(k, Z; s)\right]$.

D.2 The algorithm

We discretize the exogenous shocks using the method of Rouwenhorst (1995), which Kopecky and Suen (2010) show is accurate for highly persistent processes. We use 7 states for $\ln Z$ and 3 states each for $\ln A$ and $\ln \theta$. The endogenous aggregate state μ_k has 7 grid points spanning $\pm 50\%$ in levels around its deterministic steady state value. The firm-level capital grid has 20 points, with bounds that cover the ergodic distribution of k in simulations.

We integrate the permanent productivity innovation ε using a 10-point Gauss–Legendre rule truncated at ± 3 standard deviations, with weights proportional to the normal density. Cross-sectional integrals against m, including those in (D29)–(D32), use an 11×11 tensor-product Gauss–Legendre rule on a ± 3 standard deviation box around $(\mu_k, 0)$. The weights are proportional to the bivariate normal density and sum to one. We use shape-preserving cubic interpolation for policy functions in $\ln k$ and μ_k, and cubic splines when evaluating policies at the cross-sectional quadrature nodes.

We compute equilibrium by fixed-point iteration, starting from the deterministic steady state of the normalized system. This initial solution satisfies the same equations with both aggregate shocks fixed at their unconditional means and a constant growth factor. Each iteration uses the current guesses of the policy functions $\left(\iota, u, \bar{k}\right)$ and projection coefficients (a_u, b_k, b_Z) in the following steps:

1. At every aggregate state s, solve the normalization identity (D29) for τ_k^2 and construct the quadrature nodes and weights for integration against m.

2. For every pair (s, s'), compute the growth factor $\mathbf{G}(s'|s)$ from (D30) and the predicted mean $\mu_k'(s, s')$ from (D31). The projection coefficients at s' are evaluated by interpolating in the μ_k dimension at the predicted mean; because the predicted mean itself depends on these coefficients, we evaluate them at the previous iteration's prediction, and this simultaneity is resolved as part of the outer fixed point.

3. At every s, compute aggregate consumption $c(s)$ from (D32).

4. Solve the recursion (D33) for $\mathcal{U}(s)$ by an inner fixed-point iteration, and construct the pricing kernel $M(s'|s)$ from (D34).

5. At every firm-level grid point (k, Z, s), evaluate the conditional expectations on the right-hand sides of (D37) and (D38) using the closed form (D36), and invert the two conditions for the updated investment and utilization policies. Update the entrant policy $\bar{k}(s)$ from (D39) in the same way.

6. Refit the projection coefficients (a_u, b_k, b_Z) by m-weighted least squares using the updated utilization policy.

7. Compute the sup-norm distance between the updated and current iterates of all policy functions and projection coefficients. If the distance exceeds the tolerance of 10^{-8}, update the guess with a damped step and return to step 1.

Once the policies have converged, we solve (D40) by iteration to obtain the firm value function v.

We compute the estimation moments from a simulated panel of firms. Aggregate shocks follow their Markov chains, and μ_k follows (D31). We recover the levels of aggregate quantities by cumulating the realized growth factors $\mathbf{G}$ and construct firm-level variables from the normalized policies. The resulting moments are compared with the empirical evidence in Section 2.