Journal of Monetary Economics 136 (2023) 76–90

Contents lists available at ScienceDirect

Journal of Monetary Economics

journal homepage: www.elsevier.com/locate/jmoneco

Learning and the capital age premium☆

Kai Li[a,*], Chi-Yang Tsou[b], Chenjie Xu[c]

[a] Peking University HSBC Business School and PHBS Sargent Institute of Quantitative Economics and Finance, Xili University Town, Nanshan District, Shenzhen, 518055 China
[b] University of Manchester
[c] Shanghai University of Finance and Economics

ARTICLE INFO

Article history:
Received 16 August 2019
Revised 9 February 2023
Accepted 10 February 2023
Available online 11 February 2023

JEL classification:
E2
E3
G12

Keywords:
Learning
Capital age
Cross-section of expected returns
Capital misallocation
Cash flow duration

ABSTRACT

Imperfect information and learning are introduced into a production-based asset pricing model. Our model features slow learning about firms' exposure to aggregate productivity shocks over time. In contrast to a full information case, our framework provides a unified explanation for the stylized empirical features of the cross-section of stocks that differ in capital age: old capital firms (1) have higher capital allocation efficiency; (2) are more exposed to aggregate productivity shocks and, hence, earn higher expected returns, which we refer to as capital age premium; and (3) have shorter cash-flow duration, when compared with young capital firms.

1. Introduction

Parameter uncertainty is ubiquitous in economics and finance. In this paper, we study the role of learning in a general equilibrium asset pricing model on cross-sections of stock returns. In a large class of models that link production and investment to the cross-sections of expected returns, market participants can typically observe firm-specific productivity directly and distinguish its systematic component from the idiosyncratic component. This assumption is referred to as the full information paradigm. While this paradigm offers a natural starting place and an important analytic benchmark, our paper adopts an alternative but more realistic specification: that in which individual firms' managers have imperfect information about their productivity. Specifically, we explore cases in which firms with newly installed capital (i.e., young capital vintages) have limited information about their exposure to aggregate productivity shocks, but receive noisy signals from which

☆ We are grateful for helpful comments from Urban Jermann (the editor), Andrea Eisfeldt (the associate editor), an anonymous referee, as well as Hengjie Ai, Utpal Bhattacharya, Mariano Massimiliano (Max) Croce, Anthony M. Diercks (EFA discussant), Pierre Collin-Dufresne, Jack Favilukis (AFA discussant), Naveen Gondhi (MFA discussant), Po-Hsuan Hsu, Yan Ji, Xiaoji Lin, Dimitris Papanikolaou, Jincheng Tong, Fan Yang and Chu Zhang. We also thank the participants at AFA, EFA, and MFA, the 2018 Econometric Society European Winter Meeting, the 12th NCTU International Finance Conference, the 2019 Econometric Society Australian Meeting, the 4th SDU Finance Workshop, the 2019 Taiwan Economics Research, Hong Kong University of Science and Technology, and the University of Technology Sydney. We acknowledge the General Research Fund of the Research Grants Council of Hong Kong (Project Number: 16504618) for financial support.

* Corresponding author.
E-mail address: kaili@phbs.pku.edu.cn (K. Li).

https://doi.org/10.1016/j.jmoneco.2023.02.001
0304-3932/

they learn over time. Our paper shows that this single deviation from the full information case coherently explains a wide set of empirical facts with respect to the links between capital age, resource allocation efficiency, the timing of cash flows and, most importantly, expected returns in the cross-section of stocks.

Information about the separation between the systematic and idiosyncratic components of firm-level productivity is likely to be limited. Indeed, for newly installed capital, there is not enough historical data for firm managers to estimate its firm-specific exposure to aggregate productivity. Thus, the full information assumption may overestimate the amount of information that agents have. In this paper, we propose a novel general equilibrium production-based model with imperfect information. In the model, individual firms' managers have imperfect information about their firms' exposure with respect to aggregate productivity growth and, in turn, face a signal extraction problem. In particular, young capital vintage firms are assumed to have less precise signals about their firm-specific exposure than those of old vintage firms.[1,2]

This assumption is directly motivated and strongly supported by our empirical evidence in Section 2. Specifically, among U.S. publicly listed firms contained in the Compustat database, the dispersion of the marginal product of capital (MPK), which has been interpreted as a form of capital misallocation by Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), monotonically declines with firms' capital age.[3] We interpret this evidence by using a learning channel, and it is consistent with our model prediction of a negative correlation between capital age and capital misallocation. The intuition is that younger firms have less precise information about their firm-specific productivity to help them facilitate resource allocation, which leads to higher capital misallocations.

In our setting, the endogenous responses of firms' investments and payouts to news about future productivity can coherently explain the relationship among firm-level capital age, the cross-section of stock returns, and cash-flow duration. First and most importantly, the model generates a positive capital age and expected return relation in the cross-section. Second, it also simultaneously produces a negative link between capital age and the duration of cash flow. In contrast, in the full information case, both the expected returns and cash flow duration are entirely flat with capital age.

The intuition for these key model implications is as follows. Compared with young capital firms, old firms have more precise information about their exposure to aggregate productivity shocks. Therefore, they are more capable of taking advantage of aggregate technological progress. High information precision acts like a risk exposure amplification mechanism, making old capital firms more sensitive to aggregate productivity shocks, implying a higher average return. In other words, heterogeneity in information translates into heterogeneity in risk exposure across different capital vintages. For similar reasons, young capital firms have lower resource reallocation efficiency, are less able to take advantage of aggregate technology growth, and pay lower payouts than old capital firms. Therefore, young capital firms possess longer cash flow duration, as their future cash flows obtain more weight. Old capital firms, in contrast, have shorter cash flow duration.

These model implications are strongly supported by empirical evidence. To investigate the empirical link between capital age, the duration of cash flows, and expected returns in the cross-section, firms are sorted into quintile portfolios based on their capital ages within Fama-French 30 industries. As reported in Table 2, the return spread of a long-short old-minus-young (OMY) strategy, which is referred to as the "capital age premium", is statistically significant with a t-value of 2.91, and its Sharpe ratio is 0.44. The positive capital age-expected return relation is consistent with the first model implication in that old capital firms' more precise information translates into higher risk exposure, leading to a higher average return compared to that of young capital firms. Next, our paper documents a negative correlation between capital age and cash flow duration in Table 6, in which the cash flow duration is defined by Dechow et al. (2004), consistent with the model's second implication that explores the relation between the timing of cash flow and capital age.

Guided by our theory, we also provide extensive empirical evidence that directly supports the key learning mechanism. First, the productivity of new vintages of capital are shown to be less sensitive to aggregate productivity shocks than that of older vintages, consistent with our model implication. Second, young capital firms are shown to have a lower learning rate about their exposure compared to that of old capital firms, consistent with the assumption about heterogeneous information across capital vintages. Finally, normalized payouts for young capital firms are shown to have lower exposure to both long-run and short-run aggregate productivity shocks than old capital firms, which strongly supports our model.

We also examine empirical evidence differentiating our explanation from other alternative economic channels, including value premium, market power,[4] life cycles, and embodied technology shocks documented in the literature. The empirical evidence suggests that these explanations cannot explain the capital age premium. Due to space limitations, this empirical evidence is provided in Appendix F.4.

In our quantitative analysis, we show that our model, when calibrated to match both conventional macroeconomic quantity dynamics and asset pricing moments, generates a significant capital age premium, as well as a negative capital age and

[1] In our paper, we focus on a firm's capital age rather than its founding year age. In Appendix Appendix F.4.3, we provide empirical evidence that capital age carries different predictability of cross-section of stock returns from that of founding year age.

[2] In Section B.2 of the Appendix, we explicitly introduce costly learning into our static model to provide a micro-foundation of a positive relation between capital age and information precision. We thank Andrea Eisfeldt (the associate editor) and an anonymous referee for suggesting this fruitful direction.

[3] Feng (2019) documents a consistent negative relation between MPK dispersion and firm age by using firm-level panel data from China, Columbia, and Chile.

[4] Thanks to Andrea Eisfeldt (the associate editor) and the referee's great suggestion, we extend our endogenous learning model by allowing firms with different degrees of market power in Section B.2 of the Appendix. The extended model implies a non-monotonic relationship between market power and expected returns and justifies that market power alone is not likely to derive the cross-sectional return predictability of capital age.

cash flow duration relation. As the data confirm, firms with an older capital age have a higher average return and a shorter cash flow duration. Quantitatively, our model reproduces the joint empirical relationships among capital age, expected returns, and cash flow duration in the data quite well.

Literature review

Our paper builds on the literature that studies the implication of learning with respect to asset market valuations. Pastor and Veronesi (2009) provide a comprehensive review of learning models in finance. David (1997), Veronesi (2000), and Ai (2010) explore how learning and imperfect information affect both asset valuations and the risk premium on the aggregate equity market. Croce et al. (2014) and Collin-Dufresne et al. (2016) study the role that learning plays with respect to long-run cash-flow risks on the term structure of equity returns and equity risk premium, respectively. In particular, our paper is closely related to, yet distinguished from, Pástor and Veronesi (2003), which proposes a learning channel about firms' mean idiosyncratic profitability in a partial equilibrium setting, and implies a life-cycle pattern of firms' market-to-book ratio that is consistent with data. In contrast to this paper, our paper uses a general equilibrium model to study learning about firms' exposure with respect to aggregate productivity growth and quantifies its cross-sectional implications with respect to expected returns.

Our paper is related to prior studies that highlight uncertainty and learning about different firm-level parameters, such as mean productivity (Pástor and Veronesi, 2003; 2009), R&D cost (Berk et al., 2004), and mean cash flow (Andrei et al., 2019). Specifically, our paper focuses on uncertainty about exposure to macroeconomic risk, similar to the approach in Ai et al. (2018) and Kim et al. (2021). In Appendix A, we provide a simple neoclassical framework to compare different learning mechanisms in the literature, focusing on asset pricing implications. We show that under this framework, (1) without fixed cost, learning about idiosyncratic productivity levels, as in Pástor and Veronesi (2003) and Andrei et al. (2019), won't affect risk premium; (2) with the presence of fixed costs, learning about idiosyncratic productivity levels or fixed costs, as in Pástor and Veronesi (2009) and Berk et al. (2004), will generate a leverage effect that affects risk premium; and (3) under some mild conditions, such a model setup with fixed costs is isomorphic to the alternative setup adopted in our paper (i.e., managers directly learn about the beta of the total payoff). Our paper studies the asset pricing implications of learning under a general equilibrium setup. To maintain tractability, we choose a setup without fixed costs and focus on the uncertainty about risk exposure directly.

Our paper is also closely related to the literature that emphasizes asset pricing implications of new technology frontiers and technological innovation, including Pástor and Veronesi (2009), Kogan et al. (2017), Gârleanu et al. (2012), Kogan et al. (2020), Garleanu et al. (2012), and Lin et al. (2019). The important difference between these papers and ours is that they explicitly allow for two technological shocks. Shocks of the first type are assumed to be technology "neutral" or "disembodied" because they affect the productivity of the entire amount of capital, irrespective of its vintage. Shocks of the second type correspond to (infrequent) arrivals of major technological or organizational innovations (e.g., automobiles, the Internet). These shocks do not affect the economy on impact, but only after firms have invested in new vintages of the capital stock that "embody" the technological improvements. This second type of shock creates displacement risks and generates a creative destruction effect. However, our paper only considers the first type of productivity growth and investigates the learning by doing effect within the existing technology.

The theoretical approach in our paper is connected to the investment-based asset pricing literature by endogenizing investment and linking it to the cross-section of expected returns.[5] Zhang (2005) provides an investment-based explanation for the value premium. Chan et al. (2001) and Lin (2012) focus on the relationship between R&D investment and stock returns. Eisfeldt and Papanikolaou (2013) develops a model of organizational capital and expected returns.

Moreover, our paper is related to the production-based asset pricing literature, for which Kogan and Papanikolaou (2012) provide an excellent literature survey. Compared with prior studies in this literature, our model departs from existing models in two significant aspects. First, although we address the equity premium puzzle in ways that are similar to those in other papers in the literature, our model generates heterogeneity in asset risk premia through different exposures to aggregate risk between old and young capital firms. Second, we nest a tractable vintage capital model into a general equilibrium in which individual firms have imperfect information about their productivity and must learn their productivity over time; in contrast, most previous papers assume perfect information in the model economy.

Our paper is related to a growing field of research that explores the impact of learning on capital misallocation. Feng (2019) finds that misallocation decreases with firm age and provides a firm's life-cycle learning mechanism to interpret the empirical finding. David et al. (2016) examines and quantifies the effects of resource misallocation across firms through information frictions. David et al. (2018) develops a theory to link misallocation to systematic investment risks, but without a learning channel. Our paper also identifies a consistent negative capital age and MPK dispersion relation. However, different from their papers, we use this empirical pattern as supporting evidence for our key model assumption of heterogeneous information across firms with different capital ages, and further study its asset pricing implications in the cross-section.

Our paper also contributes to the literature that studies the duration premium in the cross-section. Recent papers that explore the cross-section of stock returns, including Gonçalves (2019), Gormsen and Lazarus (2021) and Chen and Li (2018), show that long cash flow duration firms tend to earn a significantly lower average return than short duration firms, which

[5] There is a vast literature on investment-based asset pricing models, but we do not attempt to summarize it here. A partial list includes Cochrane (1991), Jermann (1998), Li et al. (2009), Tuzel (2010), Lin (2012), Belo and Lin (2012), Eisfeldt and Papanikolaou (2013), Belo et al. (2014a), Belo et al. (2014b), İmrohoroğlu and Tüzel (2014), Favilukis and Lin (2015), Belo et al. (2017), and Belo et al. (2018), among others.

Table 1
Misallocations on Portfolios Sorted by Capital Age.
This table reports time-series averages of capital misallocations within five capital age quintile portfolios, in which capital age is measured in quarters. The capital misallocation is computed through a two-step procedure. First, we compute the cross-sectional dispersion of marginal product of capital (MPK) relative to industry peers within a narrowly defined industry (either Fama-French 30 industries or SIC 2-digit industry code). Second, we take the average of the dispersion measure across industries. Misallocation 1 measures MPK by the ratio of operating income before depreciation (*oidbp*) to a one-year-lag net plant, property, and equipment (*ppent*) as in Chen and Song (2013), while Misallocation 2 measures MPK as sales (*sale*) over a one-year-lag net plant, property, and equipment (*ppent*) as in David et al. (2018). The sample period is from December 1978 to July 2016 and excludes utility, financial, and R&D intensive industries. A detailed definition of the variables can be found in the Appendix G.6.

	Y	2	3	4	O
Capital Age	9.71	15.04	19.86	24.66	35.95
Misallocation 1					
SIC 2-digit	1.02	0.88	0.81	0.80	0.77
FF30	1.09	0.93	0.87	0.85	0.79
Misallocation 2					
SIC 2-digit	1.12	0.91	0.84	0.79	0.82
FF30	1.13	0.93	0.87	0.90	0.93
Number of Firms	480	469	471	467	458

is referred to as the short duration premium. Our model predicts that firms with an older capital age have a higher average return and a shorter cash flow duration, and thus proposes a learning mechanism that can potentially rationalize the duration premium.

The rest of our paper is organized as follows. Section 2 summarizes our motivating empirical facts on the relationship between capital age, capital misallocations, and expected returns. Section 3 describes a general equilibrium model with production and learning and analyze its quantitative asset pricing implications. Section 4 provides a quantitative analysis of our model. Section 5 presents some additional evidence to support our model, and Section 6 concludes.

2. Empirical facts

In this section, we present several empirical facts that motivate our interest in studying the link between imperfect information, learning, and the cross-section of expected returns sorted by capital age. Details of the data construction can be found in Appendix G.

First, we investigate the empirical link between capital age and capital misallocation.[6] Following Salvanes and Tveteras (2004), Ai et al. (2013), and Lin et al. (2019), the capital age of an U.S. publicly listed firm is measured as a weighted average of the age of each capital vintage. Then, firms are sorted into quintile portfolios based on their capital ages within Fama-French 30 industries. Table 1 reports the cross-sectional dispersion of the marginal product of capital (MPK, hereafter) within each quintile portfolio as a measure of capital misallocation, following Hsieh and Klenow (2009). Within each quintile portfolio of firms, we first calculate the MPK dispersion within narrowly defined industries, either at the Fama-French 30 industry level or at a more refined SIC two-digit level, and then average the dispersion across industries.

Table 1 shows a salient inverse relationship between capital age and capital misallocation. That is, portfolios with higher capital age present lower levels of capital misallocation, ranging from 1.13 in the young capital age quintile to 0.77 in the old capital age quintile. Such a downward sloping pattern of misallocation across portfolios sorted by capital age is robust not only to different industry classifications but also to different measures of MPK dispersion, as used in Chen and Song (2013) or David et al. (2018), respectively.[7]

The negative relationship between capital age and capital misallocation supports our key model ingredient – that young capital firms possess less information about their exposure to common productivity shocks than old capital firms. Consistent with David et al. (2016), less information leads to lower resource reallocation efficiency. Such a model ingredient predicts a negative capital age and capital misallocation relation, as summarized in Lemma 2 of Section 3, which is strongly supported by the evidence in Table 1.[8] In addition to this indirect evidence, we present results in Table A5 of the Appendix that old capital firms feature lower levels of dispersion in analysts' forecasts and higher levels of analysts' coverage. This evidence further supports the positive relation between capital age and information precision.

Next, we present the evidence on the cross-section of stock returns based on portfolios sorted by capital age. Table 2 reports average annualized excess returns, t-statistics, and Sharpe ratios of the five portfolios sorted by capital age. The

[6] "Misallocation" is somewhat of a misnomer in our environment, as firms act optimally, given the information at hand. That said, we follow the literature and use the term to refer broadly to deviations from marginal product equalization.

[7] The average MPK dispersion over the entire sample in our calculation is broadly consistent with that in Chen and Song (2013) and David et al. (2018), although the latter papers did not calculate the dispersions across capital-age-sorted portfolios.

[8] In our model, old capital firms are assumed to have full information about their exposure; therefore, Lemma 2 suggests that the group of old capital firms should not have capital misallocation. However, the evidence shows that this firm group (i.e., portfolio O) still displays a positive MPK dispersion. This may be attributable to other factors, such as adjustment costs, financial constraints, taxes, and regulations that are not included in our model.

Table 2
Univariate Portfolio Sorting on Capital Age.
This table shows asset pricing tests for five portfolios sorted on capital age relative to a firm's industry peers, for which we use the Fama-French 30 industry classifications and rebalance portfolios at the beginning of January, April, July, and October. The results use monthly data, for which the sample period is July 1979 to December 2016 and excludes utility, financial, and R&D intensive industries. We report average excess returns over the risk-free rate E[R]-R_f, standard deviations Std, and Sharpe ratios SR across portfolios. Standard errors are estimated by Newey-West correction with ***, **, and * indicating significance at the 1, 5, and 10% levels. We include t-statistics in parentheses and annualize the portfolio returns by multiplying by 12. All portfolio returns correspond to value-weighted returns by firm market capitalization.

Variables	Y	2	3	4	O	OMY
E[R] - R_f (%)	3.32	6.50**	8.34***	8.19***	9.11***	5.79***
[t]	0.95	2.24	3.44	3.48	3.76	2.91
Std (%)	21.16	18.43	15.39	14.69	15.2	13.27
SR	0.16	0.35	0.54	0.56	0.60	0.44

average equity excess return for firms with old capital age (Quintile O) is 5.79% higher on an annualized basis than that of young capital firms (Quintile Y). The return spread of a long-short old-minus-young (OMY) strategy is referred to as the capital age premium. The return difference is statistically significant with a t-value of 2.91. Its Sharpe ratio is 0.44, which is almost comparable to that of the aggregate stock market index (around 0.5).[9]

In sum, the learning mechanism proposed in our paper provides a potential coherent explanation for the empirical evidence. On the one hand, old firms contain more precise information about their exposure to common productivity shocks, and hence have lower levels of capital misallocation; on the other hand, old firms with more information take better advantage of aggregate technological progress, and thus experience greater exposure to aggregate shocks, which results in higher average returns. In the next section, we develop a production-based general equilibrium model with a learning mechanism to formalize our preceding intuition and to quantitatively account for the capital age premium.

3. Model setup

The key mechanism of our model is that firms learn about their exposure to aggregate productivity shocks over time. In this section, to illustrate the basic premise of our theory, we first describe the learning mechanism in a static framework. We then incorporate learning into a dynamic general equilibrium model with different capital vintages to formalize our intuition and study its implications for the cross-section of expected returns.

3.1. Aggregation with learning

3.1.1. Static problem

The static model setup is similar to that of Melitz (2003) and Hsieh and Klenow (2009). In addition, it is augmented with an endogenous learning mechanism. Such a mechanism predicts that old capital age firms optimally choose higher information precision, which therefore justifies our assumption with respect to a positive relation between capital age and information precision in our dynamic model. It is also consistent with the direct and indirect evidence on a positive link between firms' capital age and information precision, as shown in Section 2.

In this economy, a group of infinitesimal firm units produces intermediate inputs y_i. There is also a single final good Y produced by a representative producer in a perfectly competitive final output market. This final good producer combines intermediate inputs using a constant elasticity of substitution (CES) production function:

$$Y = \left[\int y_i^{\nu} di\right]^{\frac{1}{\nu}}, \tag{1}$$

in which the parameter ν controls for the elasticity of substitution between intermediate inputs. Firms use capital and labor to produce intermediate goods through the production function:

$$y_i = k_i^{\alpha}(A_i n_i)^{1-\alpha}. \tag{2}$$

The productivity of firm i takes the form that $A_i = e^{m+\beta_i \Delta a}$, in which Δa is a common shock that affects the productivity of all firms, β_i is the firm-specific exposure to the common shock Δa, and m is a constant. The common productivity shock follows a Gamma distribution $\Delta a \sim \mathcal{G}(k_a, \theta_a)$.[10] It is assumed that firm managers do not exactly know their exposure, β_i,

[9] The evidence on the capital age premium is consistent with the empirical finding in Lin et al. (2019). However, we sort portfolios within industries to control for industry heterogeneity, while Lin et al. (2019) does not. Our paper proposes a learning mechanism that emphasizes that firms are uncertain about their firm-specific exposure. Because firms in the same industry are presumed to share the same industry-specific exposure to aggregate productivity, we compare firms within the same industry.

[10] The results of this section hold with any non-negative distributions. We restrict the common shock to the positive region because we focus on the implications of firms learning about their exposure to technological progress. Our learning mechanism does not directly apply in terms of technological

and must therefore make production decisions based on their interference on β_i. To facilitate a closed-form solution, we assume that conditioning on the common shock Δa, the prior distribution of β_i follows $N(\mu, \frac{1}{\Delta a}\sigma^2)$.[11] Before making the production decision, each firm receives a noisy signal of its exposure:

$$s_i = \beta_i + \epsilon_i, \tag{3}$$

in which $\epsilon_i \sim i.i.d.N(0, \frac{1}{\Delta a}\tau^2)$, conditional on Δa. The parameter τ determines the level of noise with respect to firm signals. When $\tau = 0$, firms have perfect information about their exposure to common shocks. As τ increases, firms are less certain about their exposure to common shocks, and input choices are less efficient. In the extreme case of $\tau \to \infty$, signals are not informative at all.

Each firm unit chooses capital and labor inputs to maximize expected profit under its information set:

$$\max_{k_i, n_i} E_s[k_i^{\alpha}(A_i n_i)^{1-\alpha} p_i] - Rk_i - Wn_i, \tag{4}$$

in which R is the capital rent and W is the wage rate, and E_s denotes the signal s_i, which explicitly emphasizes that a firm takes its signal into consideration when making the production decision. p_i is the market price of the intermediate good j, which can be determined as the marginal product of intermediate input $\frac{\partial Y}{\partial y_i}$.

The aggregate production function of the firm group is defined as:

$$F(K, N) \equiv \left[\int \left(k_i^{\alpha}(A_i n_i)^{1-\alpha}\right)^{\nu} di\right]^{\frac{1}{\nu}} \quad \text{subject to: } \int k_i di = K \text{ and } \int n_i di = N, \tag{5}$$

in which for each i, the choices of k_i, n_i must be measurable with respect to firm i's information. That is, k_i and n_i can only be functions of the signal s_i. In Appendix B.1, we prove that the optimality of resource allocation implies that the aggregate production of the firm group can be written as $Y = K^{\alpha}(\mathbf{A}N)^{1-\alpha}$, in which:

$$\mathbf{A} = \left[\int E_s\left(A_i^{(1-\alpha)\nu}\right)^{\frac{1}{1-\nu}} di\right]^{\frac{1-\nu}{(1-\alpha)\nu}}. \tag{6}$$

For simplicity's sake, we impose the normalization condition $\mu = 1 - \frac{1}{2}(1-\alpha)\nu\sigma^2$ throughout the paper.[12] This normalization assumption implies that the exposure to the aggregate productivity Δa is 1 in the case of no information ($\tau = \infty$). The following lemma gives the functional form of the group level production function:

Lemma 1. *The aggregate production function of the firm group is given by:*

$$F(K, N) = K^{\alpha}(\mathbf{A}N)^{1-\alpha}, \tag{7}$$

in which log productivity is given by $\ln \mathbf{A} = m + \lambda(\tau^2)\Delta a$, *and* $\lambda(\tau^2)$ *is defined as:*

$$\lambda(\tau^2) = 1 + \frac{1}{2}(1-\alpha)\frac{\nu^2}{1-\nu}\frac{\sigma^4}{\sigma^2+\tau^2}. \tag{8}$$

In the no private information case, $\lim_{\tau^2\to\infty}\lambda(\tau^2) = 1$, *and in the full information case,*

$$\lambda^* = \lim_{\tau^2\to 0}\lambda(\tau^2) = 1 + \frac{1}{2}(1-\alpha)\frac{\nu^2}{1-\nu}\sigma^2. \tag{9}$$

Proof. See the Proof in Appendix B.1. □

The realized log MPK dispersion is given by the following lemma:

Lemma 2. *The realized log MPK dispersion (cross-sectional variance) follows:*

$$Var[\log(MPK_i) - \log(MPK)] = (1-\alpha)^2\nu^2\frac{\sigma^2\tau^2}{\sigma^2+\tau^2}\Delta a. \tag{10}$$

Proof. See the Proof in Appendix B.1. □

regress because the information about exposure to old technologies is already present. Meanwhile, our setup is flexible enough to account for total factor productivity (TFP) growth patterns. In the data, TFP growth is positive over 90% of the time. To generate the 10% negative TFP growth, we can set the constant parameter m to be a negative number so that when the common shock Δa is small enough, TFP growth is negative.

[11] Hsieh and Klenow (2009) assume the prior distribution of exposure β_i to be $N(\mu, \sigma^2)$. Their setup implies that beta dispersion is constant over the business cycle, while ours implies a counter-cyclical beta dispersion.

[12] This normalization condition is only for simplification purposes. The original economy can be recovered by simultaneously scaling up/down other parameters. For example, conditional on Δa, an economy with parameters and conditioning variable $(\mu, \sigma^2, \tau^2, \Delta a)$ can be represented by a normalized version $(1, \frac{\sigma^2}{\mu}, \frac{\tau^2}{\mu}, \mu\Delta a)$.

Our paper studies the implications of different levels of information precision for the cross-section of firms. Therefore, we implement a comparatively static analysis with respect to signal noise τ. We have three intuitive implications from Lemma 1 and 2. First, as firms acquire better information about their productivity, they can better allocate capital and labor across each other. Lemma 2 shows that the realized MPK dispersion across firms decreases with respect to information precision. As the signal noise decreases, capital misallocation decreases. When firms receive perfect information, $\tau = 0$, the deadweight loss is equal to zero. In Section 2, we document a negative empirical correlation between capital age and capital misallocation. This evidence is consistent with our theoretical prediction in Lemma 2. It supports our key model ingredient that young capital firms contain less information about their exposure to common productivity shocks than old capital firms.

Second, better information leads to higher levels of firm group productivity, as a result of lower levels of capital misallocation. The productivity gap of the firm group with signal noise τ with respect to the full information benchmark can be computed as:

$$\ln A(0) - \ln A(\tau^2) = \frac{1}{2}\frac{1}{(1-\alpha)(1-\nu)}Var[\log(MPK_i) - \log(MPK)].$$

In the static model setup, the productivity gap is proportional to the difference in the capital misallocation of these two economies.

Third, and most importantly, better allocation induced by higher information precision acts as a risk exposure amplification mechanism. To illustrate this point, we inspect the firm group productivity expression $lnA = m + \lambda(\tau^2)\Delta a$: a 1 unit increase in common productivity shock Δa corresponds to $\lambda(\tau^2)$ units increase in firm group productivity. Because $\lambda(\tau^2)$ is a decreasing function of τ, the firm group's exposure to common productivity shocks increases with information precision. The upper bound on the exposure is attained under the full information and denoted by λ^*. The intuition is that when firms are uncertain about their exposure to common productivity shocks, more information allows them to take better advantage of aggregate technological progress; therefore, they feature a higher exposure to aggregate shocks.

In Appendix B.2, we further introduce costly learning to endogenize firms' information precision decisions as a microfoundation that old capital firms optimally choose a higher information precision. We also extend the model to compare firms with different degrees of market power and show that the model generally does not imply a monotonic relationship between market power and expected returns.

3.1.2. Dynamic learning

This subsection extends our preceding setup to a dynamic setting to deliver a tractable general equilibrium framework with learning to account for the cross-section of firms with heterogeneity in capital age and their expected returns. Firm i's productivity follows the following stochastic growth process:

$$A_{i,t} = \exp\left(\sum_{u=0}^{t} m + \beta_{i,u}\Delta a_u\right), \tag{11}$$

in which $\{\Delta a_u\}_{u=0}^t$ is a sequence of non-negative common productivity shocks. For $u = 0, 1, \dots, t$, $\beta_{i,u}$ is the exposure of firm i's productivity with respect to the common shock Δa_u. It is also assumed that $\{\beta_{i,u}\}_{u=0}^t$ is i.i.d across firm i and over t, and has a prior distribution of $N(\mu, \frac{1}{\Delta a_u}\sigma^2)$ as in the static setup.

In the static setup, young capital firms are less productive than old capital firms. If firms only learn the current exposure $\beta_{i,t}$ in the dynamic setting, this productivity difference accumulates over time, and the economy cannot have a balanced growth path. To guarantee balanced growth, we allow for perpetual learning in our dynamic setting: firms receive new signals about the entire history of their exposure coefficients and constantly update their beliefs in every period t. Specifically, in each period t, firm i will receive a sequence of new signals $\{s_{i,u,t}\}_{u=0}^t$, and each element in this sequence, $s_{i,u,t}(u \le t)$, is a signal received by firm i's manager in period t about the exposure coefficient $\beta_{i,u}(u \le t)$. The signals are normally distributed:

$$s_{i,u,t} = \beta_{i,u} + \epsilon_{i,u,t}, \qquad \epsilon_{i,u,t} \sim N\left(0, \frac{1}{\Delta a_u}\tau_{i,u,t}^2\right), \tag{12}$$

in which $\tau_{i,u,t}$ controls the variance of signal $s_{i,u,t}$. We omit firm index i for brevity's sake and describe the information updating process of firm i as follows:

- In period 0, $lnA_0 = m + \beta_0\Delta a_0$. After observing the signal:

$$s_{0,0} = \beta_0 + \epsilon_{0,0}, \qquad \epsilon_{0,0} \sim N\left(0, \frac{1}{\Delta a_0}\tau_{0,0}^2\right),$$

the posterior distribution of β_0 is updated as:

$$\beta_0 \sim N\left(\frac{1}{\frac{1}{\sigma^2}+\frac{1}{\tau_{0,0}^2}}\left[\frac{\mu}{\sigma^2}+\frac{s_{0,0}}{\tau_{0,0}^2}\right], \frac{1}{\Delta a_0}\frac{1}{\frac{1}{\sigma^2}+\frac{1}{\tau_{0,0}^2}}\right).$$

- In period 1, $lnA_1 = 2m + \beta_0 \Delta a_0 + \beta_1 \Delta a_1$. The manager receives a signal, $s_{1,1}$, on β_1. Under perpetual learning, the manager receives a signal $s_{0,1}$ on β_0 as well. The signals follow:

$$s_{0,1} = \beta_0 + \epsilon_{0,1}, \qquad \epsilon_{0,1} \sim N\Big(0, \tfrac{1}{\Delta a_0}\tau_{0,1}^2\Big),$$
$$s_{1,1} = \beta_1 + \epsilon_{1,1}, \qquad \epsilon_{1,1} \sim N\Big(0, \tfrac{1}{\Delta a_1}\tau_{1,1}^2\Big).$$

After observing these signals, the posterior distributions of β_0 and β_1 are updated as:

$$\beta_0 \sim N\left(\frac{1}{\frac{1}{\sigma^2}+\frac{1}{\tau_{0,0}^2}+\frac{1}{\tau_{0,1}^2}}\left[\frac{\mu}{\sigma^2}+\frac{s_{0,0}}{\tau_{0,0}^2}+\frac{s_{0,1}}{\tau_{0,1}^2}\right], \frac{1}{\Delta a_0}\frac{1}{\frac{1}{\sigma^2}+\frac{1}{\tau_{0,0}^2}+\frac{1}{\tau_{0,1}^2}}\right),$$
$$\beta_1 \sim N\left(\frac{1}{\frac{1}{\sigma^2}+\frac{1}{\tau_{1,1}^2}}\left[\frac{\mu}{\sigma^2}+\frac{s_{1,1}}{\tau_{1,1}^2}\right], \frac{1}{\Delta a_1}\frac{1}{\frac{1}{\sigma^2}+\frac{1}{\tau_{1,1}^2}}\right).$$

- In period 2, $lnA_2 = 3m + \beta_0 \Delta a_0 + \beta_1 \Delta a_1 + \beta_2 \Delta a_2$. The manager receives signals $\{s_{0,2}, s_{1,2}, s_{2,2}\}$ on $\{\beta_0, \beta_1, \beta_2\}$, and updates beliefs on β_0, β_1, and β_2.
- In general, in period t, $\ln A_t = \sum_{u=0}^{t}(m + \beta_u \Delta a_u)$. The manager receives signals, $\{s_{u,t} | u \le t\}$, on the current and all past exposures $\{\beta_u | u \le t\}$:

$$s_{u,t} = \beta_u + \epsilon_{u,t}, \quad \epsilon_{u,t} \sim N\Big(0, \frac{1}{\Delta a_u}\tau_{u,t}^2\Big), \quad \text{for } u = 0, 1, \cdots, t.$$

The posterior distributions of $\{\beta_u | u \le t\}$ are updated as:

$$\beta_u \sim N\left(\frac{1}{\frac{1}{\sigma^2}+\sum_{v=u}^{t}\frac{1}{\tau_{u,v}^2}}\left[\frac{\mu}{\sigma^2}+\sum_{v=u}^{t}\frac{s_{u,v}}{\tau_{u,v}^2}\right], \frac{1}{\Delta a_u}\frac{1}{\frac{1}{\sigma^2}+\sum_{v=u}^{t}\frac{1}{\tau_{u,v}^2}}\right), \quad \text{for } u = 0, 1, \cdots, t.$$

In this setup, firms will constantly learn about their exposures and improve their productivity, ensuring a balanced economic growth path.

3.1.3. Capital vintages, information structure, and aggregation

The main purpose of this paper is to study the implications of learning on the cross-section of firms. We choose a setup that allows us to study the link between capital age and expected returns, and at the same time, avoid keeping track of the distribution of firms with heterogeneous information. In particular, we assume that firms can be divided into $\bar{n}$ generations with the generation index n. In our quantitative model, $\bar{n}$ is set to 5, corresponding to 5 portfolios sorted by capital age in the empirical section. We use $n = 1$ to denote the youngest generation, and use $n = \bar{n}$ to denote mature firms. Within each generation, there is a continuum of firms indexed with i that produce intermediate inputs, y_i. These outputs can be transformed into a group-level output Y_n using a CES production function in the same fashion as in Eq. (1).

The sole distinction across firms in different generations is different levels of information precision. We make the critical assumption that older generation firms have more information about their exposure coefficients with respect to aggregate productivity. This assumption is also consistent with suggestive evidence in Section 2 that captures a negative relation between capital age and capital misallocation, together with Lemma 2. In the extreme case, mature firms (i.e., generation $\bar{n}$ firms) are assumed to know the exact value of $\{\beta_{i,u}\}_{u=0}^{t}$; that is, mature firms have complete information about their exposures.

The signal volatility at the micro-level $\{\{\tau_{u,t}\}_{u=0}^{t}\}_{t=0}^{\infty}$ is an infinite-dimensional object, and is not directly observable in the data. In Appendix B.3, we show that the sequence of $\{\{\tau_{u,t}\}_{u=0}^{t}\}_{t=0}^{\infty}$ can be specified as functions of parameters λ_n and ρ_n such that the group-level productivity growth follows an AR(1) process, as in Lemma 3:

Lemma 3. *The aggregate output of firm group n is:*

$$Y_{n,t} = K_{n,t}^{\alpha}(\mathbf{A_{n,t}} N_{n,t})^{1-\alpha}. \tag{13}$$

The productivity of mature firm group ($n = \bar{n}$) is:

$$\mathbf{A_{\bar{n},t}} = \exp\left[\sum_{u=0}^{t} m + \lambda^* \Delta a_u\right]. \tag{14}$$

If the sequence of noise parameters $\{\{\tau_{u,t}\}_{u=0}^{t}\}_{t=0}^{\infty}$ is specified as functions of λ_n and ρ_n for $n = 1, 2, \ldots, \bar{n}$ as equations (Appendix B.32) and dummyTXdummy-((Appendix B.33) in Appendix B.3, then the ratio between the productivity of young firms ($n < \bar{n}$) and that of mature firms ($n = \bar{n}$), $\chi_{n,t}$, is stationary and follows the following AR(1) process:

$$\chi_{n,t+1} = \ln \mathbf{A_{\bar{n},t+1}} - \ln \mathbf{A_{n,t+1}} = \rho_n \chi_{n,t} + (\lambda^* - \lambda_n)\Delta a_{t+1}. \tag{15}$$

In addition, the law of motion of the productivity for mature firms ($n = \bar{n}$) follows:

$$\ln \mathbf{A_{\bar{n},t+1}} - \ln \mathbf{A_{\bar{n},t}} = m + \lambda^* \Delta a_{t+1}, \tag{16}$$

and, with $\chi_{n,t}$ as the state variable, the law of motion for the productivity of young firms ($n < \bar{n}$) follows:

$$\ln \mathbf{A_{n,t+1}} - \ln \mathbf{A_{n,t}} = m + (1-\rho_n)\chi_{n,t} + \lambda_n \Delta a_{t+1}. \tag{17}$$

Proof. See the Proof in Appendix B.3. □

Eqs. (15)–(17) fully specify the aggregate productivity of young and mature firms. λ_n, which characterizes young firms' contemporaneous exposure to common shocks, decreases with the signal's noise (τ^2). This is consistent with Lemma 1, as firms with less information precision are less sensitive to aggregate productivity shocks. $1-\rho_n$ can be interpreted as the learning rate about productivity. ρ_n is increasing with the variances of the signals. Intuitively, because younger firms' information is less precise, the productivity gap between young and mature firms will persist for a longer time.

As detailed in Appendix F.1, the empirical evidence shows that $\lambda_{n+1} > \lambda_n$ and $\rho_n > \rho_{n+1}$; that is, younger firms have lower contemporaneous exposure to common productivity shocks and feature a lower learning rate. The evidence strongly supports our assumption that young firms have less information with respect to their firm-specific exposure to common productivity shocks than mature firms.

In the dynamic setup, generation n firms' log MPK dispersion, a capital misallocation measure, is directly linked to the productivity gap $\chi_{n,t}$, as summarized by Lemma 4. This highlights the fact that, in our model, information-induced misallocation is responsible for the productivity difference between different capital vintages.

Lemma 4. *In the dynamic setup, the realized log MPK dispersion (cross-sectional variance) in generation n follows:*

$$Var[\log(MPK_i) - \log(MPK)] = 2(1-\alpha)(1-\nu)\chi_{n,t}.$$

Proof. See the Proof in Appendix B.3. □

3.2. The full model

3.2.1. Preferences

Time is discrete and infinite, and is indexed by t. In this economy, there is a representative agent with Kreps and Porteus (1978) preferences, as in Epstein and Zin (1989):

$$V_t = \{(1-\beta)C_t^{1-\frac{1}{\psi}} + \beta(E_t[V_{t+1}^{1-\gamma}])^{\frac{1-1/\psi}{1-\gamma}}\}^{\frac{1}{1-1/\psi}} \tag{18}$$

in which C_t denotes the aggregate consumption at time t. For the sake of model parsimony, the utility function does not involve the dis-utility from labor; hence, the labor supply is inelastic.

3.2.2. Output producers

The stochastic process for the common productivity growth is specified as:

$$\begin{aligned} \Delta a_{t+1} &= \mu + x_t + e^{\sigma_a}\varepsilon_{a,t+1}, \\ x_{t+1} &= \rho_x x_t + e^{\sigma_x}\varepsilon_{x,t+1}, \\ \begin{bmatrix}\varepsilon_{a,t+1}\\ \varepsilon_{x,t+1}\end{bmatrix} &\sim i.i.d.N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\right). \end{aligned} \tag{19}$$

This specification follows Croce (2014) and captures long-run productivity risks.[13] In particular, the common productivity growth has two components: a short-run productivity shock $\varepsilon_{a,t+1}$ and a long-run shock $\varepsilon_{x,t+1}$. Short-run shocks affect contemporaneous output directly but do not affect future productivity growth. Long-run shocks do not affect the current output but carry news about future productivity growth rates. The log standard deviations of both shocks, σ_a and σ_x, are set to be constant over time.

There are $\bar{n}$ generations of firms based on their information precision. We denote $\mathbf{A}_t$ and $\mathbf{K}_t$ as the vectors of a firm's generation-wide productivity and capital stocks, respectively; that is, $\mathbf{A}_t = \{\mathbf{A_{n,t}}\}_{n=1}^{\bar{n}}$ and $\mathbf{K}_t = \{K_{n,t}\}_{n=1}^{\bar{n}}$. The aggregate production can be specified as the solution to the following optimal resource allocation problem:

$$\begin{aligned} F(\mathbf{A}_t, \mathbf{K}_t) = \max&_{N_{1,t},N_{2,t},\cdots,N_{\bar{n},t}} \textstyle\sum_{n=1}^{\bar{n}} K_{n,t}^{\alpha}(\mathbf{A_{n,t}} N_{n,t})^{1-\alpha}, \\ &\text{subject to } \textstyle\sum_{n=1}^{\bar{n}} N_{n,t} = 1. \end{aligned} \tag{20}$$

Despite featuring substantial heterogeneity across firms, the production side of our model can be summarized by a representative firm with the production function $Y_t = F(\mathbf{A}_t, \mathbf{K}_t)$, for which the law of motion for productivity is characterized by Eqs. (15)–(17), and the dynamics of capital stocks for firm generations are given by Eqs. (21)–(23), which will be detailed in the next subsection.

[13] In the full model, we specify a standard Gaussian long-run risk process for aggregate productivity growth, rather than the process that includes a non-negative Gamma process plus a negative constant term m used in our static/dynamic learning model. We choose to use a more standard productivity process to avoid the concern that non-negative shocks' specifications drive the return patterns. Our untabulated results show little difference in quantitative asset pricing implications with respect to these two alternative distributional assumptions about productivity growth shock processes.

3.2.3. Firm dynamics

New firms are created by investment. Upon creation, they belong to the youngest generation ($n = 1$). In each period, firms will exit with probability δ. The surviving firms of generation n ($n < \bar{n}$) become generation $n + 1$ with probability ϕ. For the sake of model parsimony, we assume that all firms are subject to the same exit rate and transition rate. Under this assumption, the capital dynamics of the youngest generation follow:

$$K_{1,t+1} = (1-\delta)(1-\phi)K_{1,t} + I_t, \tag{21}$$

in which I_t is investment. Eq. (21) shows that the next period's young capitals come from the existing young capitals that haven't moved to the next generation and new investment. The capital dynamics for the middle generations ($1 < n < \bar{n}$) follow:

$$K_{n,t+1} = (1-\delta)(1-\phi)K_{n,t} + (1-\delta)\phi K_{n-1,t}. \tag{22}$$

Eq. (22) shows that the inflow to the middle generation's capital only comes from the previous generation because old capital can not be directly created through investment. Lastly, the capital dynamics for the mature generation follow:

$$K_{\bar{n},t+1} = (1-\delta)K_{\bar{n},t} + (1-\delta)\phi K_{\bar{n}-1,t}, \tag{23}$$

Notably, there is no investment adjustment cost in our model; in contrast, a high adjustment cost is needed to generate a reasonable equity premium in standard RBC models. However, with the vintage capital and learning mechanism, our model can achieve a reasonable equity premium even without adjustment costs, which we will explore in Section 4.

To complete our model, we also have the market clearing condition:

$$C_t + I_t = Y_t. \tag{24}$$

3.2.4. Equilibrium conditions

Our full benchmark model adopts the Dixit-Stiglitz aggregate production function among intermediate inputs with imperfect substitution; however, instead of using the monopolistic competition setup as in Hsieh and Klenow (2009), we assume that the intermediate good producers are perfectly competitive. This assumption allows us to focus on imperfect information as the only source of resource reallocation inefficiency in our dynamic setup.[14] Under this assumption, all the intermediate firms take price p_i when solving their maximization problems. Moreover, as shown in our quantitative analysis (Section 4.2), this additional ingredient of monopolistic power does not significantly impact asset pricing, although it does lead to further inefficiency in capital accumulation in the dynamic setup.

Because of the perfect competition setup, the standard welfare theorems apply in our economy; therefore, equilibrium prices and quantities can be constructed from the solution of a planner's problem. Detailed derivations of the equilibrium conditions are provided in Appendix C. Given the equilibrium quantities, the cum-dividend price of mature firms, $p_{K_{\bar{n}},t}$, satisfies:

$$p_{K_{\bar{n}},t} = \alpha A_{\bar{n},t}^{1-\alpha}\left(\frac{K_{\bar{n},t}}{N_{\bar{n},t}}\right)^{\alpha-1} + (1-\delta)E_t[\Lambda_{t,t+1}p_{K_{\bar{n}},t+1}]. \tag{25}$$

in which $\Lambda_{t,t+1}$ is the stochastic discount factor. Also, the cum-dividend price of adolescent firms, $p_{K_n,t}$ ($n < \bar{n}$), satisfies:

$$p_{K_n,t} = \alpha A_{n,t}^{1-\alpha}\left(\frac{K_{n,t}}{N_{n,t}}\right)^{\alpha-1} + (1-\delta)\{(1-\phi)E_t[\Lambda_{t,t+1}p_{K_n,t+1}] + \phi E_t[\Lambda_{t,t+1}p_{K_{n+1},t+1}]\}. \tag{26}$$

Eq. (25) implies that the cum-dividends marginal value of mature firms equals the current period marginal product of capital, $A_{\bar{n},t}^{1-\alpha}\left(\frac{K_{\bar{n},t}}{N_{\bar{n},t}}\right)^{\alpha-1}$, plus the expected continuation value of future payoffs, $p_{K_{\bar{n}},t+1}$, adjusted for the survival probability $1-\delta$. According to Eq. (26), the value of adolescent firms is determined by the marginal product of its capital in the current period, $\alpha A_{n,t}^{1-\alpha}\left(\frac{K_{n,t}}{N_{n,t}}\right)^{\alpha-1}$, plus the continuation value. Conditional on surviving to the next period, the continuation value consists of two parts: with probability ϕ, they become the next generation and pay $p_{K_{n+1},t+1}$ going forward; with probability $1-\phi$, they remain in the same generation and pay the continuation value of $p_{K_n,t+1}$.

The optimal investment should satisfy the Euler equation:

$$1 = E_t[\Lambda_{t,t+1}p_{K_1,t+1}]. \tag{27}$$

Eq. (27) equates the marginal cost of investment on the left-hand side, which equals one due to zero adjustment costs, to the marginal benefit of investment on the right-hand side.

[14] This perfect competitive market setup promises that, under the current imperfect information structure, the allocation is Pareto optimal, such that in the dynamic model, we may solve the problem with a social planner. However, when compared to the perfect information economy, there is a deadweight loss. Hsieh and Klenow (2009) assume monopolistic competition, which does not distort the allocation of capital across firms in their static model, but will lead to inefficiency in capital accumulation in our dynamic setup.

Table 3
Calibration Parameters.
This table reports a summary of parameters for our quarterly calibrations. The benchmark capital model features limited information and learning. We set the discount rate (β), risk aversion (γ), and intertemporal elasticity of substitution (ψ) as in standard long-run risk models. The capital share (α) and capital depreciation rate (δ_k) are consistent with the standard real business cycles literature. We calibrate μ and σ_a to match the mean and the volatility, respectively, of output growth in the U.S. economy in our sample period, 1930–2016. We then set relative volatility at $\exp(\sigma_x - \sigma_a) = 0.13$ and the autocorrelation of long-run risk at $\rho_x = 0.946$, following Croce (2014). We choose transition rate (ϕ) to match the steady-state distribution of capital age and capital share, and choose the learning parameter λ and ρ to match their empirical counterparts as detailed in Appendix F.1.

Preference parameters		
Risk aversion	γ	10
Intertemporal elasticity of substitution	ψ	2
Discount factor	β	0.99
Technology parameters		
Capital share	α	0.3
Depreciation rate of capital	δ	0.03
Learning parameters		
Capital stage transition rate	ϕ	0.125
Productivity exposure of mature generation	λ^*	2.07
Cointegration speed	ρ_s	0.966
Common productivity parameters		
Average growth rate	μ	0.001
Volatility of short-run risk	$\exp(\sigma_a)$	0.014
Relative volatility of long-run risk	$\exp(\sigma_x)/\exp(\sigma_a)$	0.13
Auto-correlation of expected growth	ρ_x	0.946

3.3. Asset returns

With these equilibrium conditions, we can compute the asset returns for each firm group. The ex-dividend price of $K_{n,t}$, denoted as $q_{K_n,t}$, should satisfy:

$$q_{K_n,t} = E_t[\Lambda_{t,t+1} p_{K_n,t+1}], \text{ for } n = 1, 2, \ldots, \bar{n}. \tag{28}$$

The return of capital takes the form:

$$R_{K_n,t+1} = \frac{\alpha A_{n,t+1}^{1-\alpha}\left(\frac{K_{n,t+1}}{N_{n,t+1}}\right)^{\alpha-1} + (1-\delta)\left[(1-\phi)q_{K_n,t+1} + \phi q_{K_{n+1},t+1}\right]}{q_{K_n,t}}, \text{ for } n < \bar{n},$$

$$R_{K_{\bar{n}},t+1} = \frac{\alpha A_{\bar{n},t+1}^{1-\alpha}\left(\frac{K_{\bar{n},t+1}}{N_{\bar{n},t+1}}\right)^{\alpha-1} + (1-\delta)q_{K_{\bar{n}},t+1}}{q_{K_{\bar{n}},t}}, \text{ for } n = \bar{n}.$$

The key mechanism that generates the return spread between old versus young capital is the difference in the marginal product of capital's exposure to the common productivity shock. As discussed previously, mature firms with more information can take better advantage of aggregate technological progress, and their productivity features a higher exposure to aggregate shocks. Therefore, old firms' marginal products of capital are more sensitive to common productivity shocks and thus earn higher expected returns.

The market return can be computed as a weighted average of the returns on different capital vintages in this economy, $R_{m,t+1} = \sum_{n=1}^{\bar{n}} \frac{q_{K_n,t}K_{n,t}}{\sum_{n=1}^{\bar{n}} q_{K_n,t}K_{n,t}} R_{K_n,t+1}$.

4. Quantitative model implications

In this section, we calibrate our model at a quarterly frequency and evaluate its ability to replicate key moments of macroeconomic quantities and asset prices at the aggregate level. More importantly, we investigate its performance in quantitatively accounting for capital age premium in the cross-section. In addition, we also show that our model offers a joint framework to understand the cash flow duration and the duration premium in the cross-section.

4.1. Calibration

Table 3 presents calibration parameters at the quarterly frequency. In our model, there are three parameters intimately related to our key learning mechanism: rate of transition to the next capital age group (ϕ), mature firms' exposure to contemporaneous aggregate shocks (λ^*), and the persistence of the cointegration residual (ρ). As shown in Table 4, the transition rate ϕ is set to 0.125 to broadly match the steady-state distribution of capital age and capital share. The other two learning parameters λ^* and ρ are identified by their empirical counterparts (i.e., the exposure to aggregate productivity shocks and the auto-correlation of the productivity gap). Appendix F.1 provides the detailed empirical evidence and calibration procedure.

Table 4
Capital Age and Capital Share.
This table reports the average capital age (measured in years) and capital share of each capital-age firm group in the data and the benchmark model. Capital share is defined as the time series average of group *ppent* share ($\frac{ppent_n}{\sum_{n=1}^{5} ppent_n}$). The benchmark capital model features limited information and learning, but does not have adjustment costs. A detailed calculation of the model counter-part is described in Appendix D.1.

Panel A: Capital Age					
	Y	2	3	4	O
Data	3.44	5.39	6.59	8.48	14.2
Model	1.63	3.26	4.89	6.52	14.2
Panel B: Capital Share					
Data	0.12	0.19	0.25	0.26	0.19
Model	0.21	0.17	0.13	0.11	0.38

Table 5
Aggregate Moments.
This table reports macro quantities and asset returns in the model and data. The benchmark capital model features limited information and learning and is calibrated as in Table 3. RBC is the real business cycle model with convex adjustment costs. We retain the same parameter except that we keep five capital age groups but eliminate the learning channel (by imposing $\lambda_n = 1$ and $\rho_n = 0$). We also add the adjustment cost of investment to generate equity premium. We model the adjustment cost following Jermann (1998):
$G(I, K) = K\left[\alpha_0 + \frac{\alpha_1}{1-1/\xi}\left(\frac{I}{K}\right)^{1-1/\xi}\right]$.
$\{\alpha_0, \alpha_1\}$ are set such that in a steady state, $G = I$ and $G_I = 1$. We set the adjustment cost parameter $\xi = 2.2$ to obtain the same equity premium as in our benchmark model. Column Adj. Cost shows the result of the benchmark model with convex adjustment costs. Column M.C. shows the result of the benchmark model with monopolistic competition. We set $\nu = 0.8$, which corresponds to a markup of 1.25, following De Loecker et al. (2020). Panel A reports the moments of output, consumption, and investment. Panel B reports the equity premium, the risk-free rate, and the capital age spread. $E(r_m - r_f)$ is the levered equity premium. $E(r_5 - r_1)$ is the levered spread between capital age group 5 and group 1. We assume a constant financial leverage ratio of 2. All the moments are annualized.

		Data	Benchmark	RBC	Adj. Cost	M.C.
Parameter Learning			Yes	No	Yes	Yes
Adjustment Cost			No	Yes	Yes	No
Monopolistic Competition			No	No	No	Yes
Panel A: Aggregate Quantities						
Average output growth	$E(\Delta y)$	2.00	2.00	2.00	2.00	2.00
Volatility of output growth	$\sigma(\Delta y)$	3.49	3.49	3.49	3.49	3.49
Volatility of consumption growth	$\sigma(\Delta c)$	2.53	2.78	3.19	3.17	2.90
Volatility of investment	$\sigma(\Delta i)$	16.40	6.16	4.92	4.83	6.56
Autocorrelation of output	$AC_1(\Delta y)$	0.45	0.54	0.46	0.53	0.54
Autocorrelation of consumption	$AC_1(\Delta c)$	0.49	0.67	0.48	0.56	0.64
Corr of consumption and investment	$corr(\Delta c, \Delta i)$	0.39	0.83	0.81	0.87	0.82
Panel B: Asset Prices						
Equity premium	$E(r_m - r_f)$	5.70	4.03	4.03	5.72	4.07
Risk-free rate	r_f	0.89	0.89	0.89	0.89	0.73
Volatility of equity return	$\sigma(r_m)$	17.61	3.11	4.49	4.73	3.11
Volatility of risk-free rate	$\sigma(r_f)$	0.97	1.17	1.35	1.36	1.19
Capital age premium	$E(r_5 - r_1)$	5.79	5.89	0	5.91	5.95

4.2. Aggregate moments

We now turn to the quantitative performance of the model at the aggregate level. We solve and simulate our model at a quarterly frequency and aggregate the model-generated data to compute annual moments.[15] Table 5 reports the key moments of macroeconomic quantities (top panel) and those of asset returns (bottom panel) respectively. The model-implied moments are broadly consistent with the key empirical features of macroeconomic quantities and asset prices.

In addition to our benchmark calibration, we also calibrate an RBC model with adjustment costs for comparison's sake. In terms of aggregate moments on macro quantities (top panel), our model has similar implications to those standard RBC models. In particular, our calibration features low volatility of consumption growth (2.78%) and matches the mean, volatility, and autocorrelations of output and consumption growth with the data reasonably well. Due to the absence of adjustment costs, the volatility of an investment in our benchmark calibration (6.16%) is reasonably high when compared with that of the RBC model.

[15] The model is solved by using a second-order local approximation around the steady-state in the `Dynare` package.

Table 6
Capital Age Premium and Cash Flow Duration.
This table reports excess returns and cash flow duration in the benchmark model and data. Panel A reports the average excess return (in annualized percentage term) of each capital age group. Panel B reports the cash flow duration (measured in years). Duration 1 denotes the cash flow duration in Weber (2018), and Duration 2 denotes the cash flow duration in Gonçalves (2019). The benchmark model features limited information and learning, but does not have adjustment costs.

Panel A: Excess Return						
	Y	2	3	4	O	OMY
Data	3.32	6.50	8.34	8.19	9.11	5.79
Model	0.27	2.57	4.28	5.49	6.16	5.89
Panel B: Cash Flow Duration						
Duration 1	21.13	20.09	19.64	19.66	19.49	1.64
Duration 2	49.64	43.25	40.34	38.15	37.22	12.42
Model	32.25	31.83	31.66	31.56	31.50	0.75

Turning our attention to asset pricing moments (bottom panel), our model produces a low risk-free rate (0.89%) and a high equity premium (4.03%) with a leverage ratio of 2, comparable to key empirical moments for aggregate stock markets. It is noteworthy that our model can generate a high equity premium even without adjustment costs. Because new capital does not fully enjoy productivity growth, the consumption-smoothing effect of investment is mitigated. To achieve a comparable equity premium, the RBC model needs high adjustment costs, which leads to a lower investment volatility (4.92%). Our benchmark model also produces a sizable capital age premium as in the data. In contrast, because there is no informational difference between old capital firms and young capital firms in the RBC model, it fails to generate any capital age spread.

To better understand the implications of the learning mechanism in our model, we further plot and analyze impulse response functions of our benchmark model and compare them with those of the RBC model in Section D.4 of the Appendix.

We also extend our benchmark model to add adjustment costs or monopolistic competition. The corresponding moments are reported in columns "Adj. Cost" and "M.C.," respectively.

As shown in column "Adj. Cost", adding adjustment costs to the benchmark model significantly increases the market equity premium to 5.72% while the capital age spread remains unchanged. Compared to the RBC calibration, the benchmark model with adjustment costs can achieve a significantly higher equity premium (5.72% vs. 4.03%) while keeping investment volatility at a similar magnitude (4.83% vs. 4.92%). This suggests that, compared with adjustment costs, the learning mechanism has a similar effect on equity premium, but a lower impact on investment volatility.

Column "M.C." reports the results of the extended model with monopolistic competition. Also Appendix E provides the equilibrium conditions of the extended model. The quantitative performance of our model with monopolistic competition is very similar to the benchmark calibration in terms of macro quantities and asset price moments. It is noteworthy that, because market power creates additional inefficiency in intertemporal capital accumulation, the economy features a lower level of output, consumption, investment, and capital stock, all of which are quantified in Table A2 of the Appendix.

4.3. Cross-sectional implications

Panel A of Table 6 reports the average excess return of each capital age group in the data and model. Our model can generate a capital age spread as large as 5.89%, which is comparable to the data. The key mechanism for generating the return spread is as follows: mature firms have better information about their exposure to common shocks, demonstrating better resource allocation and also resulting in an amplification effect that makes their marginal product of capital more exposed to common productivity shocks.

Panel B of Table 6 shows the cash flow duration of each capital age group in the data and model. To demonstrate the robustness of the cash flow duration and capital age pattern, we report two alternative measures of cash flow duration that are used in Weber (2018) and Gonçalves (2019), respectively.[16] In the data, young capital firms have higher cash flow duration than those of old capital firms. This pattern is robust for both cash flow duration measures. Our model generates the same monotonic pattern. Intuitively, young firms have lower levels of average productivity, so they also have lower dividend payouts. As they grow into mature firms, they acquire better information and demonstrate better allocation of resources; hence, their productivity and dividend payouts increase. Therefore, young firms have low cash flows at the short end, and high cash flows at the far end. In contrast, the cash flow for mature firms is evenly distributed.

[16] In Appendix G.4, we provide a detailed construction of the cash flow duration following Weber (2018), which we label as "Duration 1" in our table. Meanwhile, we thank Andrei Gonçalves for sharing his cash flow duration measure with us, which we label as "Duration 2."

5. Empirical analysis

In this section, we test several implications of our model to support the learning mechanism. Our model predicts that: (1) the productivity of old capital firms is more sensitive to aggregate productivity shocks than that of young capital firms, (2) old capital firms feature a higher learning rate in terms of their exposure to aggregate shocks than young capital firms, and (3) the exposure of firms' payouts with respect to both long-run and short-run productivity shocks increases with firms' capital age, all of which are strongly supported in the data.

5.1. Firms' Exposure to aggregate shocks

Because better information will amplify risk exposure through learning, our model predicts that the exposure increases with capital age groups. To identify the contemporaneous exposure, we regress firm-level productivity growth on the aggregate productivity growth across different capital age groups by controlling for a list of variables that represent firms' fundamentals. The variable of interest is the coefficient on aggregate productivity growth, which captures the capital age group-specific exposure with respect to aggregate productivity shocks. As shown in Panel A of Table A3 in the Appendix, productivity exposure increases with capital age, which strongly supports our learning mechanism.

5.2. Estimation of the learning rate

Our model predicts that mature groups learn faster due to their information advantage. Therefore, given $1-\rho_n$ as the learning rate, ρ_n should decrease from young to mature groups. In our model, the learning rate parameter ρ_n in a group is the persistence of the co-integration residual. We identify these parameters through the autocorrelation of the productivity gap, which is measured by the log productivity differences between the oldest capital age group and a young group. The estimated autocorrelations of capital age groups 1–4 are reported in Panel B of Table A3 in the Appendix. There is a decreasing pattern from the young group to the old group, which is consistent with our model prediction.

5.3. Productivity shocks and payouts

As the learning mechanism in our model implies that old capital firms with more precise information take better advantage of technology growth, these firms' payouts should be more exposed to aggregate productivity shocks. We estimate exposures by regressing firms' payout ratios on short- and long-run productivity shocks and other control variables. As presented in Table A4 of the Appendix, there is an upward-sloping pattern on coefficients for both short- and long-run productivity shocks from young to old capital age portfolios. Such an increasing pattern provides direct evidence that supports the learning mechanism in our model.

6. Conclusion

In this paper, we study the role of learning on the cross-section of stock returns through the lens of a general equilibrium asset pricing model. There is a large class of models that links production and investment to the cross-section of expected returns, which assumes that market participants directly observe firm-specific productivity and can also distinguish their systematic components from their idiosyncratic components. In this paper, an alternative but more realistic imperfect information structure is proposed: that of individual firms having imperfect information about their productivity and facing a signal extraction problem. In particular, young capital vintage is assumed to receive less precise signals about their exposure when compared with that of old capital vintage.

We demonstrate that this learning channel is highly useful for studying the joint link between capital age, reallocation efficiency, the timing of cash flows, and, most importantly, expected returns in the cross-section. Our model framework provides an unified explanation of a wide set of empirical facts in the cross-section: (1) a negative relation between capital age and capital misallocation, (2) a negative link between capital age and cash flow duration, and, most importantly, (3) a positive capital age and expected return relation, which is referred to as the capital age premium. In contrast, a standard asset pricing model with full information generates flat relations between capital age and these aforementioned important economic variables. Our findings suggest that imperfect information and learning can significantly impact asset prices.

Data availability

Data will be made available on request.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jmoneco.2023.02.001

References

Ai, H., 2010. Information quality and long-run risk: asset pricing implications. J. Finance 65 (4), 1333–1367.
Ai, H., Croce, M.M., Diercks, A.M., Li, K., 2018. News shocks and the production-based term structure of equity returns. Rev. Financ. Stud. hhy015. doi:10.1093/rfs/hhy015.
Ai, H., Croce, M.M., Li, K., 2013. Toward a quantitative general equilibrium asset pricing model with intangible capital. Rev. Financ. Stud. 26 (2), 491–530.
Andrei, D., Mann, W., Moyen, N., 2019. Why did the q theory of investment start working? J. Financ. Econ. 133 (2), 251–272.
Belo, F., Li, J., Lin, X., Zhao, X., 2017. Labor-force heterogeneity and asset prices: the importance of skilled labor. Rev. Financ. Stud. 30 (10), 3669–3709.
Belo, F., Lin, X., 2012. The inventory growth spread. Rev. Financ. Stud. 25 (1), 278–313.
Belo, F., Lin, X., Bazdresch, S., 2014. Labor hiring, investment, and stock return predictability in the cross section. J. Polit. Econ. 122 (1), 129–177.
Belo, F., Lin, X., Vitorino, M.A., 2014. Brand capital and firm value. Rev. Econ. Dyn. 17 (1), 150–169.
Belo, F., Lin, X., Yang, F., 2018. External equity financing shocks, financial flows, and asset prices. Rev. Financ. Stud..
Berk, J.B., Green, R.C., Naik, V., 2004. Valuation and return dynamics of new ventures. Rev. Financ. Stud. 17 (1), 1–35.
Chan, L.K.C., Lakonishok, J., Sougiannis, T., 2001. The stock market valuation of research. J. Finance 56 (6), 2431– 2456. doi:10.1111/0022-1082.00411.
Chen, K., Song, Z., 2013. Financial frictions on capital allocation: atransmission mechanism of tfp fluctuations. J. Monet. Econ. 60 (6), 683–703.
Chen, S., Li, T., 2018. A unified duration-based explanation of the value, profitability, and investment anomalies. Profitab. Invest. Anomal. (November 26, 2018).
Cochrane, J.H., 1991. Production-Based asset pricing and the link between stock returns and economic fluctuations. J. Finance 46 (1), 209–237. doi:10.2307/2328694.
Collin-Dufresne, P., Johannes, M., Lochstoer, L.A., 2016. Parameter learning in general equilibrium: the asset pricing implications. Am. Econ. Rev. 106 (3). 664–98
Croce, M.M., 2014. Long-run productivity risk: a new hope for production-based asset pricing? J. Monet. Econ. 66, 13–31.
Croce, M.M., Lettau, M., Ludvigson, S.C., 2014. Investor information, long-run risk, and the term structure of equity. Rev. Financ. Stud. 28 (3), 706–742.
David, A., 1997. Fluctuating confidence in stock markets: implications for returns and volatility. J. Financ. Quant. Anal. 32 (4), 427–462.
David, J.M., Hopenhayn, H.A., Venkateswaran, V., 2016. Information, misallocation, and aggregate productivity. Q. J. Econ. 131 (2), 943–1005.
David, J.M., Schmid, L., Zeke, D., 2018. Risk-adjusted capital allocation and misallocation. Working Paper.
De Loecker, J., Eeckhout, J., Unger, G., 2020. The rise of market power and the macroeconomic implications. Q. J. Econ. 135 (2), 561–644.
Dechow, P.M., Sloan, R.G., Soliman, M.T., 2004. Implied equity duration: a new measure of equity risk. Rev. Account. Stud. 9 (2–3), 197–228.
Eisfeldt, A.L., Papanikolaou, D., 2013. Organization capital and the cross-section of expected returns. J. Finance 68 (4), 1365–1406. doi:10.1111/jofi.12034.
Epstein, L.G., Zin, S.E., 1989. Substitution, risk aversion, and the temporal behavior of consumption. Econometrica 57 (4), 937–969.
Favilukis, J., Lin, X., 2015. Wage rigidity: a quantitative solution to several asset pricing puzzles. Rev. Financ. Stud. 29 (1), 148–192.
Feng, Y., 2019. Firm life-cycle learning and misallocation. Working Paper.
Gârleanu, N., Kogan, L., Panageas, S., 2012. Displacement risk and asset returns. J. Financ. Econ. 105 (3), 491–510.
Garleanu, N., Panageas, S., Yu, J., 2012. Technological growth and asset pricing. J. Finance 67 (4), 1265–1292.
Gonçalves, A., 2019. The short duration premium. Available at SSRN 3385579.
Gormsen, N.J., Lazarus, E., 2021. Duration-driven returns. Available at SSRN 3359027.
Hsieh, C.-T., Klenow, P.J., 2009. Misallocation and manufacturing tfp in china and india. Q. J. Econ. 124 (4), 1403–1448.
İmrohoroğlu, A., Tüzel, Ş., 2014. Firm-level productivity, risk, and return. Manage. Sci. 60 (8), 2073–2090.
Jermann, U.J., 1998. Asset pricing in production economies. J. Monet. Econ. 41 (2), 257–275.
Kim, Y., Kuehn, L.-A., Li, K., 2021. Learning about the consumption risk exposure of firms. Available at SSRN 3897353.
Kogan, L., Papanikolaou, D., 2012. Economic activity of firms and asset prices. Annual Rev. Financ. Econ. 4, 1–24.
Kogan, L., Papanikolaou, D., Seru, A., Stoffman, N., 2017. Technological innovation, resource allocation, and growth. Q. J. Econ. 132 (2), 665–712.
Kogan, L., Papanikolaou, D., Stoffman, N., 2020. Left behind: creative destruction, inequality, and the stock market. J. Polit. Econ. 128 (3), 855–906.
Kreps, D.M., Porteus, E.L., 1978. Temporal resolution of uncertainty and dynamic choice theory. Econometrica 185–200.
Li, E.X., Livdan, D., Zhang, L., 2009. Anomalies. Rev. Financ. Stud. 22 (11), 4301–4334.
Lin, X., 2012. Endogenous technological progress and the cross-section of stock returns. J. Financ. Econ. 103 (2), 411–427. doi:10.1016/j.jfineco.2011.08.013.
Lin, X., Palazzo, B., Yang, F., 2019. The risks of old capital age: asset pricing implications of technology adoption. J. Monet. Econ..
Melitz, M.J., 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. Econometrica 71 (6), 1695–1725.
Pástor, L., Veronesi, P., 2003. Stock valuation and learning about profitability. J. Finance 58 (5), 1749–1789.
Pastor, L., Veronesi, P., 2009. Learning in financial markets. Annu. Rev. Financ. Econ. 1 (1), 361–381.
Pástor, L., Veronesi, P., 2009. Technological revolutions and stock prices. Am. Econ. Rev. 99 (4). 1451–83
Restuccia, D., Rogerson, R., 2008. Policy distortions and aggregate productivity with heterogeneous establishments. Rev. Econ. Dyn. 11 (4), 707–720.
Salvanes, K.G., Tveteras, R., 2004. Plant exit, vintage capital and the business cycle. J. Ind. Econ. 52 (2), 255–276.
Tuzel, S., 2010. Corporate real estate holdings and the cross-section of stock returns. Rev. Financ. Stud. 23 (6), 2268–2302.
Veronesi, P., 2000. How does information quality affect stock returns? J. Finance 55 (2), 807–837.
Weber, M., 2018. Cash flow duration and the term structure of equity returns. J. Financ. Econ. 128 (3), 486–503.
Zhang, L., 2005. The value premium. J. Finance 60 (1), 67–103. doi:10.2307/3694834. http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.2005.00725.x/full