

# Learning and the Anatomy of the Profitability Premium

Chi-Yang Tsou\*

February 17, 2021

[\[Latest Version\]](#)

## Abstract

I introduce imperfect information and learning for unobservable long-run productivity into a dynamic asset pricing model and provide an explanation for the profitability premium. Firms with high profitability have greater information precision and face greater exposure to updated long-run productivity shocks through the learning mechanism. Deviating from the existing models without learning, my framework provides a unified explanation for a wide set of empirical facts: firms with high cash-based operating profitability (1) have higher information precision and capital allocation efficiency; (2) are more exposed to aggregate productivity shocks and, hence, earn higher expected returns; and (3) exhibit shorter cash-flow duration.

**JEL Codes:** E2, E3, G11, G12, M41

**Keywords:** Operating profitability; Learning; Cross-section of expected returns; Capital misallocation; Investment; Tobin's  $q$ ; Cash flow duration

---

\*Chi-Yang Tsou (bentsou@ust.hk), Department of Finance, School of Business and Management, Hong Kong University of Science and Technology. I am grateful for helpful comments from Cheng Chen, Mikhail Bhatia, Wei Cui, Yao Deng, Xiang Fang, Po-Hsuan Hsu, Shiyang Huang, Dun Jia, Dunhong Jin, Amit Kumar, Alan Kwan, Kai Li, Jianan Liu, Tingjun Liu, Xin Liu, Yang Liu, Hunghua Pan, Vesa Pursiainen, Tatsuro Senga, Thomas Schmid, Guoman She, Jingcheng Tong, Yanzhi Wang, Tak-Yuen Wong, Jhih-Chian Wu, Hong Xiang, Wentao Yao, Lichen Zhang, and seminar and conference participants at the 14th NCTU International Finance Conference, National Chengchi University, SFM 2020, Macroeconometric Modelling Workshop 2020, Hong Kong University of Science and Technology, National Taiwan University, Shanghai University of Finance and Economics, University of Hong Kong, Academia Sinica, National Tsing Hua University, National Chiao Tung University, and National Central University for their helpful comments. I am responsible for any remaining errors or omissions.

# 1 Introduction

One of the key determinants of a firm’s fundamental asset valuation, investment, and hiring decisions depends on the manager’s assessments of the firm’s average productivity. In this paper, I investigate the role of managers’ rational learning on firms’ quantity dynamics and the cross-section of stock returns, both empirically and theoretically. To make the case for rational learning, I first document a set of empirical findings that link cash-based operating profitability (which presents better information quality than other types of profitability measures; see [Ball, Gerakos, Linnainmaa, and Nikolaev \(2016\)](#)) to information precision, resource allocation efficiency, investment-q regression, the timing of cash flows, and, most essentially, the expected returns in the cross-section of stocks. By incorporating a learning mechanism, I show that a standard investment-based asset pricing model provides a unified theory that rationalizes all these empirical patterns.

It remains an empirical challenge to measure the quality of information that managers learn from firms’ past performance. To address this challenge, I demonstrate that cash-based operating profitability contains more precise information about firms’ true economic conditions as follows. First, firms with high cash-based operating profitability have higher analyst coverage and lower dispersion in earnings forecasts. Second, I present evidence on a negative relation between cash-based operating profitability and the dispersion of static marginal product of capital (MPK), which has been interpreted as a form of capital misallocation by [Hsieh and Klenow \(2009\)](#) and [Chen and Song \(2013\)](#). These stylized facts show that cash-based operating profitability is a valid measure of information precision. Greater information precision facilitates firm managers’ effort to learn their firm-specific productivity and, therefore, attain higher efficiency in capital allocations.

To quantify the impact of information precision on firms’ decisions and stock returns, I deviate from standard investment-based asset pricing models by incorporating imperfect information and rational learning, following [Abel \(2018\)](#) and [Andrei, Mann, and Moyon \(2019\)](#). I consider an economy with imperfect information and develop a dynamic model with a large cross-section of firms. Firms invest in physical capital to maximize the value of the firm for existing shareholders, where the cross-sectional heterogeneity is driven by firm-level differences in information precision. There are two sources of aggregate risk: aggregate short-run and long-run shocks. Both aggregate shocks carry a positive price of risk. Aggregate long-run productivity is unobservable and evolves stochastically over time. An individual firm’s manager faces a signal extraction problem by observing aggregate short-run productivity and a firm-specific signal that contains information about changes in long-run productivity plus noise. Firms with greater information precision have lower loadings on the noise component

and, therefore, receive more precise signals to learn aggregate productivity than those with lower information precision.

This learning mechanism produces two important patterns in asset returns. First, high information precision causes firms to update their beliefs faster, which endogenously produces more variations in these firms'  $q$ . Such learning-induced variation is informative about firms' investment policies when firm managers make investment decisions to maximize the value of the firm, taking as given a stochastic discount factor.

Second, the model generates a positive relation between cash-based operating profitability and future stock return, and the return spread is sizable, as in the data. The intuition for this result is as follows. The difference in expected returns across firms is due to differences in variations of  $q$ , which reflects different exposure to aggregate productivity shocks. A higher variation in  $q$  leads to higher risk premia, as these firms' returns comove aggregate productivity shocks.

Relatedly, firms with low information precision have lower efficiency in capital reallocation, are less able to utilize aggregate technology growth, and pay lower payouts. Hence, consistent with the data, these firms feature longer cash flow durations, as their future cash flows account for greater weight; in contrast, their counterparts have shorter cash flow durations.

In my quantitative analysis, I present that my model, when calibrated to match both conventional macroeconomic quantity dynamics and asset pricing moments, generates a significant cash-based operating profitability premium and a negative relation between cash-based operating profitability and cash flow duration. Consistent with the data, high-cash-based operating profitability firms have a higher average return and a shorter cash flow duration. Quantitatively, my model replicates the joint empirical relationships among the MPK, investment rate, expected returns, and cash flow duration in the data reasonably well.

To further understand and test the model's economic mechanism, I perform several relevant analyses using data. First, to investigate the empirical link between cash-based operating profitability and expected returns in the cross-section, I construct a cash-based operating profitability measure. Regarding return predictability, consistent with [Ball, Gerakos, Linnainmaa, and Nikolaev \(2016\)](#), I find large and significant predictive power for cash-based operating profitability: Firms with high cash-based operating profitability earn higher expected returns than firms with low cash-based operating profitability by a value of 4.55% per annum. The difference in the return spread is economically large, and this value is 2.5 standard errors from zero.

Second, I provide empirical evidence that directly supports the negative link between cash-

based operating profitability and cash flow duration, where the measure of cash flow duration follows [Dechow et al. \(2004\)](#), [Weber \(2018\)](#), [Gonçalves \(2019\)](#), and [Gormsen and Lazarus \(2019\)](#). Therefore, this evidence validates the implication of the model I noted previously that explores the negative relation between the timing of cash flows and cash-based operating profitability.

In addition, I find better performance of the investment-q regression among firms with higher cash-based operating profitability. The interpretation is that firms with more precise information update their beliefs about future cash flows and therefore endogenously amplify volatility in their valuations. In addition, I document that productivity among firms with lower cash-based operating profitability is less sensitive to aggregate productivity shocks than that of their counterparts.

To test the model’s economic mechanism, I consider a two-factor asset pricing model that includes aggregate short-run and long-run productivity shocks as the pricing factors, following [Bansal and Shaliastovich \(2013\)](#) and [Ai, Croce, Diercks, and Li \(2018\)](#). I empirically document that the normalized payout (i.e., a proxy for cash flow) of firms with lower cash-based operating profitability is less exposed to both long- and short-run aggregate productivity shocks. Second, I show that exposure to these two shocks is positively priced and reasonably captures the cross-sectional variation in stock returns across portfolios sorted by cash-based operating profitability. Moreover, consistent with the model, the good fit of the two-factor model is driven by the increasing exposure of portfolios sorted on cash-based operating profitability to these two shocks, especially in the high-quintile portfolio. Summarizing these findings, high cash-based operating profitability firms provide higher expected stock returns because they have higher loadings on both the short- and long-run productivity risks that are both positively priced.

To establish the robustness of the empirical findings, I perform two additional analyses. I show that the unconditional asset pricing factor model cannot explain the return spread sorted on cash-based operating profitability. My paper documents that the CAPM model and other standard asset pricing factor models (e.g., [Fama and French \(1996\)](#), [Fama and French \(1996\)](#), [Carhart \(1997\)](#), [Fama and French \(2015\)](#), and [Hou, Xue, and Zhang \(2015\)](#)) cannot explain the size of the return spread in the data. Such a positive return relation is not driven by other known predictors according to [Fama and MacBeth \(1973\)](#) regressions. I show that the return spread is exhibited by in both small and large firms, although the magnitude of the return spread is larger among small firms. In addition, independently double-sorted portfolios indicate that return predictability cannot be attributed to behavioral bias and financial distress. Finally, further analyses do not find support for other explanations related to existing systematic risks.

**Related Literature** My paper belongs to the literature that studies implications of learning for asset market valuations.<sup>1</sup> Following this literature, I simply assume that both the firm and the financial market learn symmetrically and simultaneously. In contrast, many papers study the implications when managers can access better information than other agents (e.g., Myers and Majluf (1984)), when managers extract information from their stock prices and then make investment decisions (e.g., David et al. (2016)), or when managers with attention constraints learn both firm-specific and aggregate information and make investment decisions (e.g., Gondhi (2020)). Closer to my paper, Andrei, Mann, and Moyen (2019) build an investment model that allows the firm to learn passively for stochastic technological innovations and study the relationship between corporate investment and Tobin’s  $q$ . Distinct from existing research, I assume that investment per se depends on the updated belief about long-run productivity through learning and show that accounting for the learning channel delivers important and nontrivial implications for the determinants of corporate investments and, more important, the cross-section of stock returns.

The theoretical approach in this paper is connected to the production-based asset pricing literature by endogenizing investment and linking it to the cross-section of expected returns.<sup>2</sup> This strand of model primarily focuses on aggregate shocks that originate from the real economy, including investment-specific technology (IST) shocks (e.g., Kogan and Papanikolaou (2013, 2014)), adjustment cost shocks (e.g., Belo et al. (2014)), and financial shocks (e.g., Belo et al. (2018), Ai et al. (2019)).<sup>3</sup> My paper differs in that I focus on the heterogeneity in information precision, which allows me to provide a novel learning mechanism to explain the empirical failure of the standard capital asset pricing model (CAPM), which is different from the existing investment-based literature that highlights the role of either investment shocks or adjustment shocks.

This paper is related to a growing field of research that explores the impact of learning on capital misallocation. Using firm-level panel data from China, Columbia, and Chile, Feng (2018) finds that misallocation decreases with firm age and provides a firm life-cycle learning mechanism to interpret the empirical finding. Furthermore, using U.S. firm-level data, I find a consistent negative relation between cash-based operating profitability and MPK dispersion

---

<sup>1</sup>Pastor and Veronesi (2009) provides a comprehensive review of learning models in finance. David (1997), Veronesi (2000), and Ai (2010) study how learning and imperfect information affect both asset valuations and the risk premium in the aggregate equity market, while my work studies the implication of state variable learning for the cross-section of expected returns.

<sup>2</sup>The theoretical foundation for Tobin’s  $q$  theory of investment refers to Keynes (1936), Brainard and Tobin (1968), Tobin (1969), Mussa (1977), Lindenberg and Ross (1981), Abel (1983), Abel and Eberly (1993), Abel and Eberly (1997), Erickson and Whited (2000), Abel (2018), and among many others.

<sup>3</sup>There is a rich literature on investment-based asset pricing models, but I do not attempt to summarize it here. A partial list includes Cochrane (1991), Jermann (1998), Tuzel (2010), Lin (2012), Belo and Lin (2012), Eisfeldt and Papanikolaou (2013), Belo et al. (2014), Belo et al. (2017), Belo et al. (2018), and Lin et al. (2019), among others.

and use this empirical pattern as supporting evidence for my model assumption of heterogeneous information across firms with different cash-based operating profitability, which allows me to incorporate the role of learning and study asset pricing implications. [David et al. \(2016\)](#) examines and quantifies the effects of resource misallocation across firms through information frictions. [David, Schmid, and Zeke \(2018\)](#) develops a theory to link misallocation to systematic investment risks, but without a learning channel. In contrast, I focus on asset pricing in the cross-section. In particular, I show that through a learning channel, firms with high cash-based operating profitability are subject to less capital misallocation and feature higher integration with aggregate fluctuations. Therefore, these firms face higher aggregate risk exposure and are thus riskier than firms with low cash-based operating profitability.

My empirical analysis builds on a vast literature to investigate the predictive power of accruals for the cross-section of stock returns. One line of research has traditionally interpreted the accrual anomaly as being driven by mispricing, according to [Sloan \(1996\)](#) which connects accruals to earnings persistence and mispricing from errors in expectations (e.g., [Xie \(2001\)](#), [Barth and Hutton \(2004\)](#), [Richardson et al. \(2005\)](#)). The other line of research explains that investors fail to distinguish accruals from cash flows and, therefore, overreact to past growth by underestimating the sustainability of accruals (e.g., [Thomas and Zhang \(2002\)](#), [Hirshleifer et al. \(2004\)](#), [Richardson et al. \(2005\)](#), [Dechow et al. \(2008\)](#)). In addition, more recent works link accruals to investment and growth by classifying them as investments in working capital (e.g., [Zhang \(2007\)](#) and [Wu et al. \(2010\)](#)). To the best of my knowledge, the closest related empirical work to my paper is [Ball et al. \(2016\)](#), which constructs a cash-based operating profitability measure by purging non-cash accruals from operating profitability. Such a predictor presents significant predictive power by subsuming the power of accruals or operating profitability. Although the empirical work focuses on a measure, cash-based operating profitability, as documented in [Ball et al. \(2016\)](#), the primary difference is that I provide an equilibrium model with a learning mechanism to generate asymmetric exposures to common productivity shocks among firms with different cash-based operating profitability. In addition, such a learning framework also captures the positive relationship among cash-based operating profitability, the explanatory power of investment-q regression, and the negative link with cash flow duration.

My paper is also situated in the growing literature on continuous-time asset pricing theory (e.g., [Cochrane \(1991\)](#), [Berk et al. \(1999\)](#), [Gomes et al. \(2003\)](#), [Menzly et al. \(2004\)](#), [Zhang \(2005\)](#), [Gomes et al. \(2009\)](#), [Pástor and Veronesi \(2009\)](#), [Santos and Veronesi \(2010\)](#), [Papanikolaou \(2011\)](#), [Gârleanu et al. \(2012\)](#), [Eisfeldt and Papanikolaou \(2013\)](#), [Pástor and Veronesi \(2012, 2013\)](#), [Kogan and Papanikolaou \(2013, 2014\)](#), [van Binsbergen \(2016\)](#)) and production-based asset pricing, for which [Kogan and Papanikolaou \(2012\)](#) provide an exhaus-

tive literature survey. Different from these papers, my paper explores the effect of learning on the long-run productivity and contributes to the literature that relates consumption or productivity risk to equity risk premiums in an amplified manner from the perspective of learning.

The rest of my paper is organized as follows. Section 2 shows evidence by reporting summary statistics and stylized empirical facts. I describe my model framework and analyze its quantitative asset pricing implications in Section 3. I discuss further empirical tests for the model and its testable implications in Section 4. In Section 5, I discuss alternative explanations for the relation between cash-based-operating profitability and return. I conclude this paper in Section 6. The Internet Appendix contains details of variable constructions, additional empirical evidence, and the model solution.

## 2 Stylized Empirical Facts

First, I report summary statistics in my data. Second, I present several stylized facts that motivate my interest in studying the link between imperfect information, learning, and the cross-section of expected returns sorted on cash-based operating profitability. In particular, I examine the empirical link between cash-based operating profitability and analysts' forecast dispersion, as well as coverage, and between cash-based operating profitability and capital misallocation. I am able to clearly point out the learning mechanism underlying the cash-based profitability premium.

### 2.1 Summary Statistics

I calculate a firm's operating profitability by following the computations in [Ball, Gerakos, Linnainmaa, and Nikolaev \(2015\)](#): sales (SALEQ) minus cost of goods sold (COGSQ) minus sales general, and administrative expenses (XSGAQ) (excluding research and development expenditures (XRDQ)).<sup>4</sup> Such a measure captures the performance of the firm's operations and is not affected by non-operating items, such as leverage and taxes. To evaluate the ability of the cash portion of operating profitability to predict stock returns, I remove non-cash components included in the computation of operating profitability and scale by lagged total assets to create the cash-based operating profitability measure. These noncash components include the changes in accounts receivable (RECTQ), inventory (INVTQ), prepaid expenses (XPP), deferred revenue (DRCQ plus DRLTQ), accounts payable (APQ). Accruals are mea-

---

<sup>4</sup>Variables used in my sample are based on quarterly Compustat, while those in [Ball, Gerakos, Linnainmaa, and Nikolaev \(2015\)](#) and [Ball, Gerakos, Linnainmaa, and Nikolaev \(2016\)](#) are based on annual Compustat.

sured as changes in noncash working capital minus depreciation expense (DPQ), in which noncash working capital is equal to the change in noncash current assets (ACTQ-CHQ) minus the change in current liability (LCTQ) less short-term debt (DLCQ) and tax payable (TXPQ), and I normalize accruals by lagged total assets.<sup>5</sup>

In Table 1, I report pooled summary statistics and correlation matrix between cash-based operating profitability and other characteristics for each quarter. Other variables include accruals (ACR/AT), market capitalization (ME), book-to-market ratio (B/M), investment rate (I/K), and return on equity (ROE).<sup>6</sup>

In Panel A of Table 1, I report the pooled mean, median, standard deviation (Std), 5<sup>th</sup> percentile (P5), 25<sup>th</sup> percentile (P25), median, 75<sup>th</sup> percentile (P75), and 95<sup>th</sup> percentile (P95) of all these variables. Observations denote the valid number of observations for each variable. I have a total of 469,741 firm-year-quarter observations with non-missing cash-based operating profitability (COP/AT). The average of cash-based operating profitability is 0.02, and the median of it is 0.03, suggesting that the distribution is slightly skewed to the left. The mean of quarterly firm-level accruals is 0.00 with a standard deviation of 0.19. The average firm’s book-to-market ratio, investment rate, and return on equities are 0.72, 0.16, and 0.00. The average size of the firms in my sample 2,431.15 when I measure the size in 2009 (million) dollar term. I also provide the industry-level summary statistics of cash-based operating profitability in Section A.5 in the Internet Appendix.

In Panel B, I present the correlation coefficients of all variables considered in Panel A. I find that cash-based operating profitability is generally not highly correlated with other variables, except that its correlation coefficients with accruals (ACR/AT), size (ME), book-to-market ratio (B/M), investment rate (I/K), and return on equity (ROE) are 0.04, -0.02, 0.00, and 0.00, respectively. These mild correlations suggest that cash-based operating profitability contains different information in my sample.

## 2.2 Univariate Portfolio Sorting: Returns and Firm Characteristics

To investigate the link between cash-based operating profitability and the cross-section of stock returns, I construct quintile portfolios sorted on firms’ cash-based operating profitability normalized by total assets (ATQ) in Panel A, property plant and equipment (PPENTQ) in Panel B, book equity (BE) in Panel C and sales (SALEQ) in Panel D by one quarter, and

<sup>5</sup>My measure of accruals is consistent Zhang (2007) and Wu, Zhang, and Zhang (2010).

<sup>6</sup>Detailed information regarding the construction of other variables refers to Section A.1 in the Internet Appendix.



report each portfolio’s post-formation average stock returns. I form portfolios at the beginning of every January, April, July, and October by sorting all sample firms with non-missing cash-based operating profitability into five groups from low to high within the corresponding 30 industries, according to [Fama and French \(1997\)](#). As a result, I have industry-specific breaking points for quintile portfolios for each March, June, September, and December. I then assign all firms with non-missing cash-based operating profitability in January, April, July, and October into quintile portfolios. Thus, the low (high) portfolio contains firms with the lowest (highest) cash-based operating profitability in each industry. After forming the five portfolios (from low to high), I calculate the value-weighted monthly returns on these portfolios over the next three months (i.e., quarterly rebalanced portfolios). To examine the profitability-return relation, I also form a high-minus-low portfolio that takes a long position in the high-profitability portfolio and a short position in the low-profitability portfolio and calculate the returns on this portfolio.

In Panel A to Panel D of [Table 2](#), the top row presents the *annualized* average excess stock returns in percentage ( $E[R]-R_f$ , in excess of the risk-free rate), *t*-statistics, standard deviations, and Sharpe ratios of the six portfolios I form. [Table 2](#) shows that a firm’s cash-based operating profitability forecast stock returns. Taking Panel A based on cash-based operating profitability normalized by total assets (which is my primary proxy of cash-based profitability) as an example, the low, 2, 3, 4, and high portfolios provide annualized excess returns of 5.50%, 7.12%, 7.70%, 7.83%, and 10.05%, respectively. More importantly, the high-minus-low portfolio provides 4.55% annualized excess return with a *t*-statistic of 2.50. In addition, the Sharpe ratios of the low, 2, 3, 4, and high portfolios are 0.30, 0.39, 0.47, 0.51, and 0.64, respectively, and that of the high-minus-low portfolio is 0.47, which is comparable to the Sharpe ratio in the aggregate equity premium. Similar patterns are observed in other panels. The finding that the return on the high-minus-low portfolio is economically large and statistically significant across all panels suggests a significant predictive ability of firm-level cash-based operating profitability for stock returns.

**[Place [Table 2](#) about here]**

Overall, [Table 2](#) provides empirical evidence for the explanatory power of firm-level cash-based operating profitability for subsequent stock returns. For the rest of my analyses, I will focus on the cash-based operating profitability normalized by lagged total assets.

I then report the average firm characteristics across quintile portfolios in Panel C of [Table 1](#). On average, my sample contains 3,076 firms. Firms are evenly distributed across five portfolios sorted on cash-based operating profitability, where the average number of firms in each portfolio ranges from 605 to 624. The cross-sectional variations in cash-based

operating profitability are large. The average cash-based operating profitability in the lowest portfolio is  $-0.05$ ; in contrast, the average cash-based operating profitability in the highest portfolio is  $0.09$ . Given that I remove the non-cash component from operating profitability when calculating cash-based operating profitability, the accruals presents a downward sloping pattern across quintile portfolios. The size, investment rate ( $I/K$ ), and return on equity (ROE) increase in portfolios sorted on cash-based operating profitability, which implies firms with high cash-based operating profitability, on average, are larger and profitable firms with more investment opportunities. Moreover, I observe a downward sloping pattern of book-to-market ratio, since firms with higher cash-based operating profitability have higher market valuations and thus exhibit lower book-to-market ratio than their counterparts. Finally, I note a negative link between cash-based operating profitability and cash flow duration. Such a negative correlation is robust to a different measures of cash flow duration, according to [Weber \(2018\)](#), [Gonçalves \(2019\)](#), and [Gormsen and Lazarus \(2019\)](#).

In summary, firms with high cash-based operating profitability have lower accruals, larger market capitalizations, lower book-to-market ratios, higher investment rates, higher profits, and shorter cash flow duration.

## 2.3 Imperfect Information and Capital Misallocation

First, to identify the imperfect information and learning mechanism, I investigate the joint link between cash-based operating profitability and analysts’ forecast dispersion, as well as coverage, on the one hand, and between cash-based operating profitability and capital misallocation, on the other.<sup>7</sup> To do so, I implement a standard procedure and sort firms into quintile portfolios based on these firms’ cash-based operating profitability within [Fama and French \(1997\)](#) 30 industries. In [Table 3](#), I report time-series averages of the cross-sectional medians of analysts’ forecast dispersion and coverage,<sup>8</sup> and the cross-sectional dispersion of the marginal product of capital (MPK, hereafter) within each quintile portfolio as a mea-

---

<sup>7</sup>“Misallocation” is somewhat of a misnomer in my environment, as firms are acting optimally given the information at hand. I follow the literature and use the term to refer broadly to deviations from marginal product equalization.

<sup>8</sup>Analysts coverage (Coverage) is calculated using the number of distinct analysts who made fiscal year one earnings forecast for the stock during the month  $[t-11, t]$  as used in [Hong et al. \(2000\)](#) and [Ali and Hirshleifer \(2019\)](#). If a firm is not covered by the Institutional Brokers’ Estimate System (IBES), I assign its stock to zero analyst coverage. Moreover, following [Garfinkel \(2009\)](#), [Diether, Malloy, and Scherbina \(2002\)](#), and [Banerjee \(2011\)](#), I obtain the analyst forecast data from the IBES database, as well. Based on individual analysts’ forecasts, I construct the analysts’ forecast dispersion by calculating the standard deviation of earnings forecasts scaled by the absolute value of the mean earnings forecast. If the mean earnings forecast is zero, according to [Diether, Malloy, and Scherbina \(2002\)](#), then I assign the stock to the highest dispersion category. In general, dropping those observations with a mean earnings forecast of zero does not significantly affect my analysis.

sure of capital misallocation, following [Hsieh and Klenow \(2009\)](#). Here, I provide a brief description of the evidence but will reserve details of the data construction to the Internet Appendix [A.2](#). Within each quintile portfolio of firms, I first calculate the MPK dispersion within narrowly defined industries, either at the [Fama and French \(1997\)](#) 30 industries level or at a more refined SIC two-digit level, and then average the dispersion across the industries.

[Place Table 3 about here]

In Panel A of Table 3, I report the average analysts' forecast dispersion and coverage across quintile portfolios. On average, firms in the high (low) portfolio have 30 (18) analysts actively tracking and publishing opinions on these firms' stock. Moreover, I show a pattern of falling slopes in analysts' forecast dispersion across quintile portfolios sorted on cash-based profitability. In particular, the range of analysts' forecast dispersion is from 2.84 in the low portfolio to 2.22 in the high portfolio. Turning to capital misallocation, I observe a salient inverse relationship between cash-based operating profitability and capital misallocation in Panel B. Firms with higher cash-based operating profitability present lower capital misallocation, ranging from 1.33 in the lowest quintile to 0.76 in the highest quintile. Such a downward sloping pattern of misallocation across five quintile portfolios are robust not only to different industry classifications but also to different measures of MPK (i.e., [Chen and Song \(2013\)](#) and [David, Schmid, and Zeke \(2018\)](#)).<sup>9</sup>

The upward sloping pattern of analysts' coverage and the downward sloping pattern of analysts' forecast dispersion and capital misallocation provide suggestive evidence to support my key model assumption that low-profitability firms contain less information about their exposure to common productivity shocks than high-profitability firms. As documented in an extensive literature, analysts are specialized professionals to collect information about stocks and disseminate it to the public. A greater number and less disagreement of analysts producing and disseminating research about a given stock should result in a higher quality of information flow to market participants.<sup>10</sup> On the other hand, consistent with [David et al. \(2016\)](#), less information leads to lower resource reallocation efficiency. My model predicts a negative cash-based operating profitability and capital misallocation relation, as summarized in Lemma 1 of Section 4, strongly supported by the evidence in Table 3.

---

<sup>9</sup>The average MPK dispersion over the whole sample in my calculation is broadly consistent with that in [Chen and Song \(2013\)](#) and [David, Schmid, and Zeke \(2018\)](#), although the latter paper did not calculate the dispersions across cash-based-operating-profitability-sorted portfolios.

<sup>10</sup>A partial list of papers using analyst coverage as a measure of the quality of the information environment includes [Brennan, Jegadeesh, and Swaminathan \(1993\)](#), [Hong, Lim, and Stein \(2000\)](#), [Hou and Moskowitz \(2005\)](#), [Chan and Hameed \(2006\)](#), [Duarte, Han, Harford, and Young \(2008\)](#), [Kumar \(2009\)](#), [Hong and Kacperczyk \(2010\)](#), [Kelly and Ljungqvist \(2012\)](#), and [Ali and Hirshleifer \(2019\)](#). [Ganguli and Yang \(2009\)](#) and [Angeletos and Werning \(2006\)](#) investigate other dimensions of how the supply of information affects financial markets.

### 3 The Model

In this section, I develop a continuous-time equilibrium model that features imperfect information to account for firm investment and to explain the role of learning in stock prices and expected stock returns. My specification of the learning mechanism is according to [Abel \(2018\)](#) and [Andrei, Mann, and Moyen \(2019\)](#).

#### 3.1 Firms and Technology

I consider an economy with a continuum of firms that produce a flow of output given by

$$Y_{it} = e^{a_{it}+x_t} K_{it}, \quad (1)$$

where  $Y_{it}$  is firm  $i$ 's output,  $a_{i,t}$  is the firm-specific component of productivity,  $x_t$  is the aggregate short-run productivity shock that affecting the output of all existing firms, and  $K_{it}$  is firm  $i$ 's capital. Without loss of generality, I abstract from describing the flexible labor decision. The output of firm  $i$  is affected by the aggregate short-run productivity shock, whose evolution is given by a mean-reverting process:

$$dx_t = \rho_x(\mu_t - x_t)dt + \sigma_x dZ_t^x, \quad (2)$$

where  $Z_t^x$  is a standard Brownian motion,  $\rho_x$  and  $\sigma_x$  are strictly positive, and  $\mu_t$  is the long-run mean of  $x_t$ . The long-term mean  $\mu_t$  also follows a mean-reverting process:

$$d\mu_t = \rho_\mu(\bar{\mu} - \mu_t)dt + \sigma_\mu dZ_t^\mu, \quad (3)$$

where  $\rho_\mu$  and  $\sigma_\mu$  are strictly positive, and  $Z_t^\mu$  is a standard Brownian process. Moreover, the output has a firm-specific component  $a_{it}$ , which evolves according to another mean-reverting process:

$$da_{it} = -\rho_a a_{it}dt + \sigma_a dZ_{it}, \quad (4)$$

where  $\rho_a$  and  $\sigma_a$  are strictly positive, and  $Z_{it}$  is a standard Brownian process. Firm-specific productivity shocks are idiosyncratic in the sense that they are independent of each other of the economy-wide productivity shock. Such the mean reversion property in firms' productivity precludes the exploding growth rate of aggregate output and guarantees an equilibrium nondegenerate cross-sectional distribution of firms' productivity and capital, in which the heterogeneity and persistence in firms' productivity are consistent with findings in the existing study (i.e., [Bartelsman and Doms \(2000\)](#)).

Firm  $i$ 's capital stock  $K_{it}$  depreciates at a common fixed-rate  $0 \leq \delta \leq 1$ , and firm  $i$  can increase its capital stock by undertaking investment  $I_{it}$ . Therefore, the stock of capital accumulates according to the law of motion:

$$dK_{it} = (I_{it} - \delta K_{it})dt. \quad (5)$$

Investment is irreversible and costly to adjust. The adjustments to the capital stock, measured in units of the investment, follow an increasing and convex function of the investment

$$\Psi(I_{it}, K_{it}) = \frac{\chi}{2} \left( \frac{I_{it}}{K_{it}} \right)^2 K_{it}, \quad (6)$$

where  $\chi$  is a strictly positive adjustment parameter so that the adjustment cost function is strictly convex.<sup>11</sup>

### 3.2 Learning about Long-run Productivity

Firm  $i$  observes the short-run productivity but cannot observe the long-run mean  $\mu_t$ . Instead, firm  $i$  forms expectations over its future stream of cash flows but cannot perfectly infer the process driving cash flows from past realizations because the unobservable long-term mean  $\mu_t$  evolves stochastically.

The firm learns about the long-term mean  $\mu_t$  from two sources. The first source is to collect the information from past realizations of short-run productivity shocks embed in the firm's output for the inference of  $\mu_t$ . On the other hand, the firm may receive a signal  $s_t$  that contains true information about the long-run productivity shock plus noise, which takes the following form in continuous time:

$$ds_{it} = dZ_t^\mu + \frac{1}{\sqrt{\Phi_i}} dZ_t^s, \quad (7)$$

where the signals  $ds_{it}$  are uncorrelated across  $n$  and independent of any other shocks in the economy. That is,  $dZ_t^x$ ,  $dZ_t^\mu$ ,  $dZ_{it}$  and  $dZ_t^s$  are independent standard Brownian motions. In particular, idiosyncratic shocks,  $dZ_{it}$ , and signal shocks,  $dZ_t^s$ , are independent to aggregate shocks (i.e.,  $dZ_t^x$  or  $dZ_t^\mu$ ), respectively. I assume that firms with different cash-based operating profitability have heterogenous loadings on signal shocks, as captured by the parameter  $\Phi_i$ . For this assumption, I present some supportive evidence documented in Section 2.2 and 2.3. Specifically,  $\Phi_i s$  are assumed to be positively proportional to firms' cash-based profitability,

---

<sup>11</sup>The similar adjustment cost specification to model firm investment decisions refer to [Abel and Eberly \(1993\)](#), [Abel and Eberly \(1997\)](#), [Erickson and Whited \(2000\)](#), [Zhang \(2005\)](#), [Liu, Whited, and Zhang \(2009\)](#), [Clementi and Palazzo \(2019\)](#), and among others.

are drawn from a uniform distribution on the interval  $[\Phi_{min}; \Phi_{max}]$  at time 0, and then remain unchanged.

$\mathcal{F}_{it}$  denotes the information set of the firm  $i$  at time  $t$ . Conditional on its information set, the firm forms beliefs about the unobservable long-term mean  $\mu_t$ . I refer to the posterior mean of  $\mu_t$ ,  $\hat{\mu}_{it} \equiv \mathbb{E}[\mu_t | \mathcal{F}_{it}]$ , and to the posterior variance of  $\mu_t$ .  $\zeta_{it} \equiv \mathbb{E}[(\mu_t - \hat{\mu}_{it})^2 | \mathcal{F}_{it}]$  (i.e., the mean squared error). The standard Kalman-Bucy filtering theory, according to [Liptser and Shiryaev \(2013\)](#), implies that the distribution of  $\mu_t$  conditional on  $\mathcal{F}_{it}$  is Gaussian with mean  $\hat{\mu}_{it}$  and variance  $\zeta_{it}$ .

The resulting posterior distribution of  $\mu_t$  is also normally distributed, as shown in [Proposition 1](#) and its corollary, which characterize firm  $i$ 's learning problem as follows.

**Proposition 1.** *The posterior mean,  $\hat{\mu}_{it}$ , evolves as*

$$d\hat{\mu}_{it} = \rho_\mu(\bar{\mu} - \hat{\mu}_{it})dt + \left(\frac{\rho_x \zeta_{it}}{\sigma_x}\right)d\hat{Z}_{it}^x + \sigma_\mu \sqrt{\frac{\Phi_i}{1 + \Phi_i}}d\hat{Z}_t^s, \quad (8)$$

where  $d\hat{Z}_{it}^x = dZ_t^x + \left(\frac{\rho_x}{\sigma_x}\right)(\mu_t - \hat{\mu}_{it})dt$  denotes the incremental component of the short-run productivity shocks, and  $\left(\sqrt{\frac{1+\Phi_i}{\Phi_i}}\right)d\hat{Z}_t^s \equiv dZ_t^\mu + \frac{1}{\sqrt{\Phi_i}}dZ_t^s$  is a scaled version of the signal in [equation \(7\)](#) as a standardized Brownian motion. The posterior variance  $\zeta_{it}$  satisfies the Riccati differential equation as follows:

$$\frac{d\zeta_{it}}{dt} = \frac{\sigma_\mu^2}{1 + \Phi_i} - 2\rho_\mu\zeta_{it} - \left(\frac{\rho_x \zeta_{it}}{\sigma_x}\right)^2. \quad (9)$$

See the Proof in [Section B.1](#) in the Internet Appendix.

Equation (8) shows that firm  $i$ 's belief about its long-run productivity  $\hat{\mu}_{it}$  is driven by the Brownian motion shocks  $d\hat{Z}_{it}^x$ , which reflect the innovations in the short-run productivity  $dZ_t^x$  plus the difference between the unobservable long-run productivity  $\mu_t$  and the updated productivity  $\hat{\mu}_{it}$ . Since the signals shocks are independent of other shocks in the economy (i.e.,  $dZ_t^x$  and  $dZ_t^a$ ), the innovations  $d\hat{Z}_t^s$  represent signal shocks to the true value of the aggregate long-run productivity. In particular, holding a higher information precision  $\Phi_i$ , firm  $i$  leans to learn from these signal shocks and shapes its belief about the long-run productivity. In addition, more precise information suggests a lower posterior variance and a lower weight learn from short-run productivity shocks. Both short-run productivity and signal shocks command a risk premium in equilibrium; However, since firms with different information precision have heterogeneous exposures to short-run and signal shocks, they exhibit different risk compensations in equilibrium.

From now on, I replace the steady-state posterior variance  $\zeta_i$  in the model economy to

keep the stationarity. In particular, such a replacement allows me to compute the stationary solution for the instantaneous variance of the posterior mean  $\hat{\mu}_{it}$ , as summarized in Corollary 1. The comparative statics with respect to firm  $i$ 's information precision  $\Phi_i$  is presented in the following intuitive implications:

**Corollary 1.** *The instantaneous variance of the posterior mean is given by:*

$$\text{Var}[(d\hat{\mu}_{it})^2] = \sigma_\mu^2 - 2\rho_\mu\zeta_i, \quad (10)$$

*which is strictly increasing in  $\Phi_i$  (i.e.,  $\partial\text{Var}[(d\hat{\mu}_{it})^2]/\partial\Phi_i > 0$ ).*

*See the Proof in Section B.2 in the Internet Appendix.*

First, firm  $i$  with a higher information precision (i.e., a higher  $\Phi_i$ ) features higher volatility and acts as a risk exposure amplification mechanism. Because the volatility of  $\hat{\mu}_{it}$  is a decreasing function of  $\Phi_i$ , firm  $i$ 's exposure to common productivity shocks increases with information precision. The upper bound of information precision is the full information case. In the extreme case as  $\Phi_i$  goes to infinity, firm  $i$  then has a complete picture about the long-run productivity  $\mu_t$  without uncertainty, and the posterior variance  $\zeta_i$  collapses to zero. At this moment, the posterior mean shares the same instantaneous volatility with the unobservable  $\mu_t$  and reaches the instantaneous volatility of the unobserved process  $\sigma_\mu^2$ . The intuition is that, when firms are uncertain about their long-run productivity shocks, more information allows them to take better advantage of aggregate technological progress, and therefore they feature an amplified exposure to aggregate shocks.

Second, better information leads to a higher level of productivity. The intuition is that firm  $i$  with a higher information precision is more aggressively updated and able to take better advantage of aggregate productivity shocks, and  $\hat{\mu}_{it}$ , therefore, attain a higher level of productivity. Taking these together, when firms cannot perfectly observe their exposure to long-run productivity shocks, information precision allows them to utilize the aggregate technological progress, and therefore they attain a higher level of productivity and amplifies the exposure to aggregate shocks.

More importantly, the following lemma shows that the realized log marginal product of capital dispersion is decreasing in  $\Phi_i$ :

**Lemma 1.** *The partial derivative of the realized  $\log(\text{MPK}_i)$  dispersion (cross-sectional variance) with respect to information precision  $\Phi_i$  is negative:*

$$\frac{\partial}{\partial\Phi_i} \text{Var}[\log(\text{MPK}_i)] < 0 \quad (11)$$

*See the Proof in Section B.3 in the Internet Appendix.*

This paper studies the implications of different information precision on the cross-section of firms. Therefore, I do comparative statics with respect to information precision  $\Phi_i$ . The interpretation of the model implications from the above lemmas is as follows. As firms with more precise information acquire better knowledge about their productivity, they can better allocate capital across each other. According to Lemma 1 I can show that the realized MPK dispersion across firms decreases with respect to information precision. Apparently, as the information precision increases, capital misallocation decreases. In Section 2.2, I document a negative empirical correlation between cash-based operating profitability and capital misallocation. This evidence is consistent with the theoretical prediction in Lemma 1 and supports my key model assumption that low cash-based operating profitability firms contain less information about their exposure to common productivity shocks than their counterparts.

### 3.2.1 Aggregation of Information

My model economy features filtered short-run productivity and signal shocks, both of which contain the aggregate information with heterogeneous loadings as shown in Proposition 1. In Section 3.2.1, I present the result of the aggregation of information. Given the aggregate information, investors can determine the individual firm-level stock valuation and expected return by using the pricing kernel, as presented in Section 3.3.1. The aggregation of information at time  $t$  summarized in the pricing kernel is shown in the following Lemma.

**Lemma 2.** *The filtered short-run productivity aggregate signal shock is given by*

$$d\hat{Z}_t^x = dZ_t^x + \left(\frac{\rho_x}{\sigma_x}\right)(\mu_t - \hat{\mu}_t)dt, \quad (12)$$

and

$$ds_t = \int ds_{it}di = \left(\sqrt{\frac{1+\Phi}{\Phi}}\right)d\hat{Z}_t^s. \quad (13)$$

*See the Proof in Section B.4 in the Internet Appendix.*

## 3.3 Asset Prices

In this section, I use the following steps to study the asset pricing implications of the model economy. First, I show the optimal investment and how stock valuations depend on short-productivity and updated long-run productivity. Second, I decompose firms' risk premia into the risk compensations to these two shocks, respectively, and study that the heterogeneity in firms' information precision translates into the cross-sectional difference in expected stock



returns with respect to short-productivity and signal shocks.

### 3.3.1 Stock Valuation and Optimal Investment

Let  $\pi_t$  denote the state price of density. For simplicity, I assume that the aggregate short-productivity and signal shocks (i.e.  $d\hat{Z}_t^x$  and  $d\hat{Z}_t^s$ ) have constant constant prices of risk,  $\lambda_x$  and  $\lambda_\mu$ , respectively, and the risk-free rate  $r_f$  is also constant. Then the stochastic discount factor (SDF) in this economy evolves stochastically according to:

$$\frac{d\pi_t}{\pi_t} = -r_f dt - \lambda_x d\hat{Z}_t^x - \lambda_\mu d\hat{Z}_t^s, \quad (14)$$

where the stochastic discount factor in equation (14) is motivated by the general equilibrium model with short- and long-run productivity shocks in Ai (2010), Ai, Croce, and Li (2013), Ai, Croce, Diercks, and Li (2018), and Li, Tsou, and Xu (2019). Both shocks endogenously affect the representative household's consumption stream, and hence they are priced in equilibrium. For the model simplicity and tractability, I restrict the interest rate and the prices of risk to be constant. Such a setting refers to the literature of investment/production-based model on cross-sectional return predictability (e.g., Zhang (2005), Zhang (2005), Belo and Lin (2012), Eisfeldt and Papanikolaou (2013), Kogan and Papanikolaou (2014), Lin, Palazzo, and Yang (2019), and among others.)

Firm  $i$ 's objective is to maximize the expected discounted summation of future cash flows, net of investments, and cost. That is, the firm's investment decisions are based on a trade-off between the market value of additional capital from its investment and the accompanying adjustment cost. The following proposition states the equilibrium optimal firm investment policy and the stock valuation:

**Proposition 2.** *The stock valuation value of firm  $i$  is defined as*

$$V(a_{it}, x_t, \hat{\mu}_{it}, K_{it}) = \underset{I_s}{\text{Max}} \mathbb{E}_t \left[ \int_t^\infty \frac{\pi_s}{\pi_t} \underbrace{(Y_{is} - I_{is} - \Psi(I_{is}, K_{is}))}_{D_{is}} ds \right], \quad (15)$$

where the optimal investment is given by

$$\frac{I_{it}}{K_{it}} = -\frac{1}{\chi} + \frac{1}{\chi} q(a_{it}, x_t, \hat{\mu}_{it}). \quad (16)$$

In the model, I assume the shadow value of capital, marginal  $q$ , is equal to average  $q$ ,

$$V(a_{it}, x_t, \hat{\mu}_{it}, K_{it}) = q(a_{it}, x_t, \hat{\mu}_{it}) K_{it}. \quad (17)$$

See the Proof in Section B.5 in the Internet Appendix.

In equation (15),  $D_{it}$  denotes firm  $i$  divided at time  $t$  and is regarded the cash flow rate of firm  $i$  generated from production. The time- $t$  stock valuation of firm  $i$ ,  $V(a_{it}, x_t, \hat{\mu}_{it}, K_{it})$ , is equal to the present value of firm  $i$  cash flows valued by using the stochastic discount factor. Holding the condition of constant returns to scale in both production and adjustment cost, as shown in Hayashi (1982), the firm's marginal  $q$  equals to its average  $q$ , and, therefore, I conjecture the firm's stock valuation as the composition of the marginal  $q$  scaling by capital stock  $K_{it}$ . The firm's marginal  $q$  is increasing in both the level of the firm-specific productivity  $a_{it}$ , the level of the aggregate short-run productivity  $x_t$ , and the level of the updated long-run productivity  $\hat{\mu}_{it}$ , respectively.

The optimal investment policy stems from the first-order condition requiring firm  $i$  to the investment until the marginal benefit of investment, as measured by its  $q$ , is equal to its marginal cost. Firm  $i$ 's investment rate (i.e.,  $I_{it}/K_{it}$ ) equals its  $q$  minus one normalized a constant when the investment irreversibility is binding. The investment depends positively on firm  $i$ 's  $q$  and negatively on its cost of capital adjustment.<sup>12</sup> Supposed that firm  $i$  with low productivity, said its marginal  $q$  below to one, is better off disinvesting by selling rather than expanding its production scale. As the result, the optimal investment rate is negative as disinvestment.

Finally, firms learn about the unobservable long-run productivity growth  $\mu_t$  and form their beliefs  $\hat{\mu}_{it}$  according to signals that contain true information about the aggregate long-run productivity. When firm  $i$  features higher information precision, its  $q$  becomes more sensitive to innovations from both short-run productivity and signal shocks. Both these effects amplify the sensitivity of firm  $i$ 's  $q(a_{it}, x_t, \hat{\mu}_{it})$ , as well as the co-movement of investment and  $q$ . Such the theoretical prediction is consistent with the empirical result when I present in Section 4.2.

### 3.3.2 Risk and Risk Premia

An application of Ito's Lemma on firm  $i$ 's stock value implies that the dynamics of firm  $i$ 's stock returns are presented in the following proposition.

**Proposition 3.** *Firm  $i$ 's realized stock returns at time  $t$  follow the process:*

$$\frac{D_{it}dt + dV_{it}}{V_{it}} = \mathbb{E}_t \left[ \frac{D_{it}dt + dV_{it}}{V_{it}} \right] + \beta_{a,it} dZ_{it}^a + \beta_{x,it} d\hat{Z}_{it}^x + \beta_{\mu,it} d\hat{Z}_{it}^\mu, \quad (18)$$

---

<sup>12</sup>The rearrangement of the equation (16) shows  $\frac{I_{it}}{K_{it}} = \frac{1}{\chi} [q(a_{it}, x_t, \hat{\mu}_{it}) - 1]$ .

in which firm  $i$ 's risk exposures to firm-specific, short-run productivity, and signal shocks are  $\beta_{a,it}$ ,  $\beta_{x,it}$ , and  $\beta_{\mu,it}$ , respectively. The functional form of firm  $i$ 's risk exposure with respect to these shocks refers to the Internet Appendix.

See the Proof in Section B.6 of the Internet Appendix.

In equation (18), I decompose firm  $i$ 's realized stock returns into the risk exposures to firm-specific productivity shocks,  $\beta_{a,i}$ , to short-run productivity shocks,  $\beta_{x,it}$ , and to signal shocks,  $\beta_{\mu,it}$ . The first term of firm  $i$ 's on the right-hand side of the equation (18) shows that firms in the economy face the different exposure  $\sigma_a$  to firm-specific shocks. The second term in equation (18) determines the firm  $i$ 's exposure to short-run productivity shocks. Third, the last term  $\beta_{\mu,it}$  captures firm  $i$ 's exposure to signal shocks, in which firms with more information precision are able to learn true information about aggregate long-run productivity from signal shocks. As a result, their realized returns present higher loadings on signal shocks.

In equilibrium, asset risk premia are determined by factor loadings times the price of risk. That is, risk premia are determined by the Euler equation that characterizes the covariance of a firm's returns with the SDF. Combining the equation characterizing stock returns in equation (18) with the SDF in equation (14) yields the conditional risk premium as the compensation to hold firm  $i$ 's stocks:

$$\begin{aligned} \mathbb{E}_t \left[ \frac{D_{it}dt + dV_{it}}{V_{it}} \right] - r_f dt &= -\text{Cov}_t \left[ \frac{D_{it}dt + dV_{it}}{V_{it}}, \frac{d\pi_t}{\pi_t} \right] \\ &= \beta_{x,it} \lambda_x dt + \beta_{\mu,it} \lambda_\mu dt. \end{aligned} \tag{19}$$

In equation (19), I show that firm  $i$ 's risk premia are determined by its exposure to short-run productivity and signal shocks, respectively. The first term captures the risk premium of short-run productivity shocks in an amplified manner when I introduce the learning mechanism, as shown in Section B.6 in the Internet Appendix. The risk premium attributing to signal shocks is in the second term of equation (19). As summarized in Proposition 2 and Proposition 3, upon a positive short-productivity or signal shock, stock valuation rises exactly when the marginal utility and thus the SDF is low. Put together, investors demand a positive compensation for their exposures to such a short-productivity or signal shock. Finally, the heterogeneity in expected returns in the cross-section is determined by heterogeneous exposures to both shocks,  $\beta_{x,it}$  and  $\beta_{\mu,it}$ , the properties of which is summarized by Corollary 2 as below.

**Corollary 2.** *Firm  $i$ 's expected return depends on  $\Phi_i$ , which is the information precision to*

learn for the unobservable long-run productivity shock.

$$\frac{\partial}{\partial \Phi_i} (\beta_{x,it} \lambda_x + \beta_{\mu,it} \lambda_\mu) > 0. \quad (20)$$

See the Proof in Section B.7 of the Internet Appendix.

In equation (20), I show that firm  $i$  with a higher  $\Phi_i$  exhibits a larger sensitivity than does firm  $j$  with a lower  $\Phi_j$  in expected stock returns. This underlying difference in  $\Phi_i$  plays an essential role in determining heterogeneous responses to aggregate shocks and in formalizing the cross-sectional difference in expected stock returns.

More importantly, the heterogeneous risk compensation for productivity risks, especially for long-run productivity shocks, is responsible for the cross-sectional difference in expected returns among firms with different information precision. As shown in Corollary 2, firm  $i$ 's risk premium with respect to the short-run productivity and signal shock positively depends on its information precision  $\Phi_i$ , especially for the latter. As a positive realization of a signal shock, stock valuations of high  $\Phi$  firms with high information precision rise more than do those of low counterparts. Heterogenous responses to signal shocks, which contain true information about aggregate long-run productivity risks, translate into a cross-sectional difference in expected stock returns. My model predicts that high-cashed-based-operating-profitability firms contain more precise information and, thus, carry higher expected returns. This prediction is strongly supported by a statistically significant high-minus-low return spread among portfolios sorted on cash-based operating profitability. I call this return to spread the cash-based operating profitability premium.

The cross-sectional asset pricing implication is captured in the following proposition.

**Proposition 4.** *Suppose that there are two firms that differ in information precision in the economy: one is a high information precision firm, while the other is a low information precision firm. According to equation (17), two firms' expected stock returns are denoted as*

$$\mathbb{E}_t \left[ \frac{D_{Ht} dt + dV_{Ht}}{V_{Ht}} \right] - r_f dt = \beta_{x,Ht} \lambda_x dt + \beta_{\mu,Ht} \lambda_\mu dt, \quad (21)$$

and

$$\mathbb{E}_t \left[ \frac{D_{Lt} dt + dV_{Lt}}{V_{Lt}} \right] - r_f dt = \beta_{x,Lt} \lambda_x dt + \beta_{\mu,Lt} \lambda_\mu dt, \quad (22)$$

respectively. The long-short portfolio of high versus low information precision firms' expected

stock returns denotes

$$\mathbb{E}_t \left[ \frac{D_{Ht}dt + dV_{Ht}}{V_{Ht}} - \frac{D_{Lt}dt + dV_{Lt}}{V_{Lt}} \right] = [\beta_{x,Ht} - \beta_{x,Lt}] \lambda_x dt + [\beta_{\mu,Ht} - \beta_{\mu,Lt}] \lambda_\mu dt. \quad (23)$$

See the Proof in Section B.8 of the Internet Appendix.

I make several observations for the long-short portfolio in equation (23) as follows. First,  $\beta_{x,it}$  and  $\beta_{\mu,it}$  are the risk exposures to short-productivity and signal shocks. When there is a negative realization of these shocks, especially for the signal shock, stock valuations for all firms fall, but the stock valuation of firm  $H$  with high  $\Phi_H$  (high information precision) drops more than does that of firm  $L$  (low information precision). Therefore, firm  $H$  is more sensitive to signal shocks. Given the positive price of risk with respect to productivity and signal shocks ( $\lambda_x > 0$  and  $\lambda_\mu > 0$ ), investors demand a positive premium to hold high information precision firms  $H$  over low information precision firms  $L$ . In sum, the cash-based operating profitability premium compensates investors to hold high information precision stocks in terms of higher exposure to short-run productivity and signal shocks.

### 3.4 Cash Flow Duration

The learning mechanism in my model economy also affects the timing of cash flow, which is measured by cash flow duration. In this subsection, I provide a framework to understand my model's implication on the timing of cash flows.

Firm  $i$ 's dividend claim or dividend strip is a security that pays  $\Psi_{is}$  at time  $s$  per unit of capital, in which  $\Psi_{is}$  is equal to firm  $i$  dividend payout  $D_{it}$  normalized by its capital  $K_{is}$ . Its normalized dividend claim at time  $t$  is denoted as:  $\Pi_{it}^s = \mathbb{E}_t \left[ \frac{\pi_s}{\pi_t} \Psi_{is} \right]$ . In the model, I define the normalized dividend payout as the marginal product of capital minus investment rate and an normalized adjustment cost (i.e.,  $\Psi_{it} = MPK_{it} - \frac{I_{it}}{K_{it}} - \Phi(\frac{I_{it}}{K_{it}})$ ).

I further define the term distribution of firm  $i$  value as:

$$\omega_{it}^s = \frac{\Pi_{it}^s}{q_t}. \quad (24)$$

That is, the fraction of maturity- $s$  dividend strip's value in total firm  $i$  valuation, and firm  $i$ 's current valuation is equal to the sum of the normalized dividend strip value at all maturities,  $q_{it} = \sum_{s=1}^{\infty} \Pi_{it}^s$ . Therefore,  $\sum_{j=s}^{\infty} \omega_{it}^s = 1$ .<sup>13</sup> Under such notations, firm  $i$ 's Macaulay duration

<sup>13</sup>For the illustrative purpose, I start from the discrete-time framework. The remaining analysis keeps focusing on the continuous-time framework.

$M_{it}$  can also be expressed as:

$$M_{it} = \sum_{s=1}^{\infty} s \omega_{it}^s. \quad (25)$$

Notably, I implement some transparent reformulation of the above equation (25) and then obtain an equivalent of firm  $i$ 's Macaulay duration  $M_{it}$ , which can be demonstrated in the following proposition.

**Proposition 5.**

$$M(a_{it}, x_t, \hat{\mu}_{it})q(a_{it}, x_t, \hat{\mu}_{it}) = \mathbb{E}_t \left[ \int_t^{\infty} s \frac{\pi_s}{\pi_t} \Psi_{is} ds \right], \quad (26)$$

for which  $\frac{\partial M_{it}}{\partial a_{it}} < 0$ ,  $\frac{\partial M_{it}}{\partial x_t} < 0$ , and  $\frac{\partial M_{it}}{\partial \hat{\mu}_{it}} < 0$ .

See the Proof in Section B.9 of the Internet Appendix.

In equation (26), I show negative sensitivities of firm  $i$ 's Macaulay duration with respect to the firm-specific productivity  $a_{it}$ , the short-run productivity  $x_t$ , and the long-run productivity  $\hat{\mu}_{it}$ . More importantly, the negative sensitivity (i.e.,  $\frac{\partial M_{it}}{\partial \hat{\mu}_{it}} < 0$ ) to the long-run productivity suggests that firms with higher information precisions feature shorter cash flow durations, according to the learning mechanism significantly that affects the timing of cash flows. Intuitively speaking, firms with high cash-based profitability have shorter cash flow duration than those with high cash-based profitability. As shown in the previous subsections, these firms possess higher information precision and capital reallocation efficiency, attain a higher level of productivity, pay higher cash flows from the beginning, and allocate greater value weight at the short maturity when compared with their counterparts; therefore, they give shorter cash flow duration, as I show in Proposition 5.

### 3.5 Calibration and Quantitative Model Predictions

In this subsection, I calibrate my model at the annual frequency and evaluate its ability to replicate key moments of both real quantities and asset prices at the aggregate level. More importantly, I investigate its performance in terms of quantitatively accounting for the cash-based operating profitability premium in the cross-section of expected stock returns. Real quantities refer to the aggregate investment rate and Tobin's  $q$ , while the aggregate asset price refers to the equity premium.

In Table 4, I present a group of calibrated parameter values in my model. I adopt the following calibration procedure to determine a set of sensible parameters. All parameters are grouped into four categories. I determine parameters in the first and second category by following the previous literature; These parameters are in line with those in Pástor and

Veronesi (2006) and Andrei, Mann, and Moyen (2019). I determine parameters in the first and third category by matching a set of first moments of quantities to their empirical counterparts, including the average values and the cross-sectional distribution of the marginal product of capital, the investment rate, and Tobin’s  $q$ . The proportional adjustment cost for investment  $\chi$  is set to match 0.04, the median of the aggregate investment rate in my data. I set the depreciation rate of capital to 10%, to be consistent with commonly used values for the depreciation rate. When I determine parameters in the fourth category, I do not follow an exact one-to-one mapping to the first moment of a specific item in the data; instead, I determine these particular parameters by jointly matching to identify moments in the data: the parameters of the pricing kernel (i.e.,  $\lambda_x = 0.15$  and  $\lambda_\mu = 0.40$ ) are picked to approximately match the first and second moment of excess returns on the market portfolio. I set the interest rate  $r_f$  to 3%, which is close to the historical average real risk-free rate according to Campbell and Cochrane (1999). Last but not least, I do not use any information about the cross-sectional variation in portfolio returns when I use it in my calibration procedure. Instead, I compare the cross-sectional portfolio returns between the data and my simulation to follow my model implication.

[Place Table 4 about here]

I also evaluate the quantitative performance of the model at the aggregate level. In Table 5, I show that my model is broadly consistent with the key empirical features of real quantities and asset prices. With respect to real quantities and asset prices, my model produces comparable results to the data in my sample.

[Place Table 5 about here]

Next, I study the cash-based operating profitability premium at the cross-sectional level. For the purpose of cross-sectional analysis, I make use of several data sources at the micro-level, including (1) firm-level balance sheet data from the Compustat quarterly files and (2) monthly stock returns from CRSP. In Section A.1, I provide additional details regarding my data sources and constructions. Specifically, I set the distribution of information precision  $\Phi^i$  between 0 and 3, and then simulate 5,000 firms. In Table 6, I report the average excess returns, marginal product of capital, investment rate, Tobin’s  $q$ , and cash flow duration across different  $\Phi^i$ , and then compare them with my data.

[Place Table 6 about here]

I document several cross-sectional implications in terms of average returns and firm characteristics in Table 6. First, my model can quantitatively replicate the pattern in the data

by generating the upward sloping marginal product of capital, investment rate, Tobin's  $q$  but the downward sloping cash flow duration across five quintile portfolios sorted by information precisions.<sup>14</sup> Second, I use Table 6 to show that my model can generate a cash-based operating profitability premium (i.e., the return spread in the high-minus-low portfolio) as sizable as 4.24%, which is comparable to the 4.55% I obtain from my data in Section 2.1. To generate the sizable return spread, I identify a key mechanism: high-based-based-operating profitability firms contain more precise information about aggregate long-run productivity shocks, so they face higher risk exposures to these shocks. Hence, investors demand higher expected returns to hold high-based-based-operating firms' stocks.

## 4 Testable Model Implications

In this section, I examine several key testable implications to support information and learning explanation. First, I document the negative relation between cash-based operating profitability and cash flow duration. Second, I also investigate the improvement of the investment- $q$  regression when firms contain more precise information. Second, firms with high cash-based profitability contain more precise information and, therefore, these firms' productivity growth exhibit higher co-movements with the aggregate productivity growth. Also, their payouts feature higher sensitivities to aggregate productivity shocks. Third, I then implement the generalized method of moments (GMM) test to show that the two-factor model, including the short- and long-run productivity shock, is positively priced in the cross-section of cash-based-operating-profitability-sorted portfolios. Together with the finding that high-cash-based-operating-profitability firms' stock returns are more exposed to both the short- and long-run productivity shock.

### 4.1 Cash Flow Duration in the Cross-Section

One implication of the learning mechanism in my model is that high cash-based operating profitability firms feature higher information precision and capital reallocation efficiency, pay higher cash flows from the beginning, and have higher value weight at the short maturity when compared with their counterparts; therefore, they give shorter cash flow duration based on the definition of cash flow duration in equation (26). In this subsection, I directly test this model prediction by showing the supporting evidence on the negative relation between cash-based operating profitability and cash flow duration.

---

<sup>14</sup>The cash flow duration in my quantitative exercise focuses on the measure of cash flow duration according to Dechow et al. (2004) and Weber (2018).



In Panel C of Table 1, I show the cash flow duration of each portfolio sorted on cash-based operating profitability. Appendix A.3 provides detailed construction of the cash flow duration.<sup>15</sup> As shown in Panel C of Table 1, firms with lower cash-based profitability feature longer cash flow duration than their counterparts. Cash flow duration decreases in portfolios sorted on cash-based operating profitability across three different measures of cash flow duration in the literature. Such the finding is consistent with the testable implication generated by my model. Intuitively speaking, less profitable firms are lower in average productivity and, therefore, possess lower dividend payouts. As time passes, they obtain better information and improve their efficiency in terms of resource allocations; meanwhile, their productivity and dividend payouts increase. Therefore, these firms have low cash flows at the short end and high cash flows at the far end.

## 4.2 Investment and Tobin’s q

Motivated by the evidence, as documented in Andrei, Mann, and Moyen (2019), that firms with more precise information are exhibiting a greater comovement between investment and Tobin’s q, I provide additional empirical facts to explore the learning mechanism by investigating the joint link between investment rate and Tobin’s q. To do so, I then implement a standard procedure and sort firms into quintile portfolios based on these firms’ cash-based operating profitability within Fama and French (1997) 30 industries. To measure the performance of the investment-q regression for firms with higher cash-based profitability, I estimate the sensitivity of Tobin’s q (i.e.,  $\beta_{q,n}$ ) with respect to investment rate by different portfolios sorted on cash-based operating profitability ( $n = L, 2, \dots, H$ ) by using the following regression:

$$\frac{I_{i,n,t}}{K_{i,n,t-1}} = \beta_{0,t} + \beta_{0,i} + \beta_{q,n}q_{i,n,t} + Controls_{i,n,t} + \varepsilon_{i,n,t}, \text{ for firms in portfolio } n \quad (27)$$

in which  $\frac{I_{i,n,t}}{K_{i,n,t-1}}$  denotes firm  $i$ ’s investment rate in portfolio  $n$  at time  $t$ ;  $\beta_{0,i}$  and  $\beta_{0,t}$  controls for the firm and time fixed effect, respectively; and  $q_{i,n,t}$  is firm  $i$ ’s q in portfolio  $n$ .  $Controls_{i,n,t}$  includes a list of control variables for firms’ fundamentals, including size, book-to-market ratio, and cash-based operating profitability.

[Place Table 7 about here]

---

<sup>15</sup>In my empirical calculation, the average of the first measure of cash flow duration (i.e., duration 1) over the whole sample is largely comparable with the mean cash flow duration in Weber (2018), although the latter paper did not calculate the duration measure across portfolios sorted on cash-based operating profitability. Data on the second measure of cash flow duration (i.e., duration 2) are from Andrei Gonçalves’s website (<https://andreigoncalves.com/>). I thank Niels Joachim Gormsen, and Eben Lazarus for sharing the third measure of cash flow duration (i.e., duration 3).

As reported in Table 7, I find that the standard investment-q panel regression fares better among firms in the high group sorted on cash-based operating profitability. The estimated coefficient  $\hat{\beta}_{q,n}$  rises from 0.1 to 0.7, and the  $R^2$  value from the regression roughly doubles from 0.58 to 0.81 when I move from the lowest to the highest portfolios sorted on cash-based operating profitability. This is consistent with the learning channel as shown in Section 2.3, and also highlights that the relation between cash-based operating profitability and information validates the learning mechanism that high-profitability firms contain more information about their exposure to long-run productivity shocks than low-profitability firms.

### 4.3 Firms' Exposure to Aggregate Shocks

#### 4.3.1 Productivity

To identify the contemporaneous exposure to aggregate productivity growth, I conduct a two-step empirical procedure. First, I estimate the firm-level productivity of public traded companies on the U.S. stock exchange, following Ai, Croce, and Li (2013) and Li, Tsou, and Xu (2019). Estimation details refer to the Internet Appendix A.4. Second, I estimate the exposure of firms' productivity growth with respect to aggregate productivity growth by different cash-based operating profitability groups ( $n = L, 2, \dots, H$ ) by using the following regression:

$$\Delta \ln A_{i,n,t} = \xi_{0,i} + \xi_{A,n} \Delta \ln A_t + Controls_{i,n,t} + \varepsilon_{i,n,t}, \text{ for firms in portfolio } n \quad (28)$$

in which  $\Delta A_{i,n,t}$  denotes firm  $i$ 's productivity growth in portfolio  $n$  at time  $t$ ,  $\xi_{0,i}$  controls for the firm-specific fixed effect, and  $\Delta \ln A_t$  is the growth rate of aggregate productivity as measured by the U.S. Bureau of Labor Statistics (BLS).  $Controls_{i,n,t}$  includes a list of control variables for firms' fundamentals, including size, book-to-market ratio, investment rate, and cash-based operating profitability.

[Place Table 8 about here]

The key coefficient of my interest, the coefficient  $\xi_{A,n}$  captures a cash-based operating profitability group  $j$ 's specific sensitivity to the aggregate productivity growth. I report my main findings in Table 8. The estimated coefficients (i.e.,  $\hat{\xi}_{A,n}$ ) on the aggregate productivity growth present an upward sloping pattern from the low to the high group based on cash-based operating profitability. These results are consistent with the learning mechanism and also confirm the intuition that that high-profitability firms have more precise information and, therefore, face higher exposures to the aggregate productivity growth.

### 4.3.2 Payout Sensitivities

One key premise of the learning mechanism in my model is that high cash-based operating profitability firms with more precise information take advantage of technology growth; therefore, their payouts are more exposed to aggregate productivity shocks. In this subsection, I directly test this model prediction. I show the supporting evidence of the positive relation between cash-based operating profitability and firms' payout exposure to aggregate short-run and long-run productivity shocks.

I proceed as follows. First, I measure firms' payout at the quarterly frequency by using their operating income before depreciation (XINTQ), net of interest expenses (TXTQ), income taxes (OIBDPQ), and common stock dividends (DVY-DVPQ), following [Croce, Marchuk, and Schlag \(2018\)](#).

In the second step, I estimate exposures by regressing firm  $i$ 's payout normalized by its lagged sales ratio on the short- and long-run productivity shocks, and other control variables as follows:

$$N_{i,t} = \beta_{0,i} + \beta_{sr}\varepsilon_t^x + \beta_{lr}\varepsilon_t^\mu + \rho N_{i,t-1} + \beta_\mu\mu_{t-1} + Controls_{i,t-1} + resid_t, \quad (29)$$

in which I follow [Ai, Croce, Diercks, and Li \(2018\)](#) to construct short-run and long-run shocks (i.e.,  $\varepsilon_t^x$  and  $\varepsilon_t^\mu$ , respectively). Specifically, I project TFP growth on predictors proposed by [Bansal and Shaliastovich \(2013\)](#) plus the integrated volatility of stock market returns to identify the long-run growth component and disentangle it from short-run TFP shocks.

In the model, a linear approximation of the equilibrium dividend processes suggests the dependence of payout on both contemporaneously short- and long- productivity shocks and predetermined variables. For the sake of parsimony, I use the lagged values of the payout ratio to capture the role of the endogenous state variables (i.e., capital shocks), so I may avoid additional measurement errors. Under the null of the model, this is an innocuous assumption.

I report my main findings in [Table 9](#). The estimated coefficients (i.e.,  $\hat{\beta}_{sr}$  and  $\hat{\beta}_{lr}$ ) present an upward sloping pattern from the low to high cash-based-operating-profitability-sorted portfolios. Such the result is consistent with my model setting, and also highlights the fact that the firms' payout in the highest quintile portfolio faces significantly higher exposure to short-run and long-run productivity shocks than do those in the lowest quintile.

[Place [Table 9](#) about here]

In summary, payout exposures present an upward sloping pattern with respect to both

short- and long-run productivity shocks, which is perfectly consistent with my model implication.

#### 4.4 Market Price and Risk Exposure

In this section, I first test the price of risk with respect to the short- and long-run productivity shock, which is positive as suggested in equation (14), and then examine cash-based-operating-profitability-sorted portfolios' exposure to the short- and long-run productivity shock, respectively.

My model implies a two-factor model in which the short-run risk is the first factor and the long-run risk is the second factor. I test the price of these two factors using the procedure detailed in [Cochrane \(2005\)](#) (pages 256-257). I first specify the stochastic discount factor (SDF) as:

$$\text{SDF}_t = 1 - \lambda_x \times \varepsilon_t^x - \lambda_\mu \times \varepsilon_t^\mu, \quad (30)$$

which specifies that investors' marginal utility is driven by two aggregate shocks:  $\varepsilon_x$  is the short-run productivity risk, and  $\varepsilon_\mu$  is the long-run productivity risk. I aim to test  $\lambda_x$  ( $\lambda_\mu$ ), which is sensitivity to  $\varepsilon_x$  ( $\varepsilon_\mu$ ) is proportional to the price of short-run (long-run) productivity shock  $\lambda_x$  ( $\lambda_\mu$ ) in equation (14). As a comparison of the two-factor model, I also consider a one-factor model and test the price of risk. The specification of the stochastic discount factor denotes:

$$\text{SDF}_t = 1 - \lambda \times R_{MKT} (\Delta \ln A_t), \quad (31)$$

where  $R_{MKT}$  is the market factor in the standard capital asset pricing model (CAPM), and  $\Delta \ln A_t$  is the growth rate of total factor productivity.

To test the price of risk, I use cash-based-operating-profitability-sorted portfolios (as presented in [Table 2](#)) as the test assets,<sup>16</sup> and implement the generalized method of moments (GMM) estimation using the following moment conditions:<sup>17</sup>

$$E[R_i^e] = -\text{Cov}(\text{SDF}, R_i^e), \quad (32)$$

which is the empirical equivalent to equation (19) of my model, but with the conditional moments replaced by their unconditional counterparts. I essentially assess the ability of  $\varepsilon_t^x$  and  $\varepsilon_t^\mu$  to price test assets on the basis of residuals of the Euler equation.

In addition, I follow the literature (e.g., [Papanikolaou \(2011\)](#), [Eisfeldt and Papanikolaou](#)

---

<sup>16</sup>This choice of test assets follows [Papanikolaou \(2011\)](#), [Eisfeldt and Papanikolaou \(2013\)](#), [Kogan and Papanikolaou \(2014\)](#), [Belo et al. \(2017\)](#), and [Lin et al. \(2019\)](#).

<sup>17</sup>Detailed information regarding moment conditions refers to [Table 10](#).

(2013), and Kogan and Papanikolaou (2014)) to estimate two statistics for cross-sectional fitting using the sum of squared errors (SSQE), mean absolute percent errors (MAPE), as well as the  $J$ -statistic of the overidentifying restrictions of the model. An insignificant  $J$ -statistic suggests that the null hypothesis of the pricing errors being zero is not rejected.

In Panel A of Table 10, I present the results of the one-factor (CAPM and TFP) and the two-factor SDF model. The price of short- and long-run risk is 0.57 and 0.81, respectively, which are both significantly positive. In terms of asset pricing errors, the SSQE of CAPM (TFP) is 6.02 (6.03). After I introduce the long-run productivity shock to my model, the SSQE reduces to 1.13. Moreover, the MAPE of CAPM (TFP) is 4.12 (4.25), and the MAPE of the two-factor model is 1.96. Last, the  $J$ -test is marginally significant to reject the CAPM model, while the  $J$ -test is statistically insignificant in the rest models. All these results indicate that the two-factor model is sufficient to capture the cross-sectional variations in the cash-based-operating-profitability-sorted portfolios.

[Place Table 10 about here]

In Panel B of Table 10, I present the cash-based-operating-profitability-sorted portfolios' risk exposures (betas) to the short- and long-run productivity shock ( $\varepsilon_t^x$  and  $\varepsilon_t^\mu$ ). Exposures to the short-run shock display a flat pattern across quintile sorted portfolios. In contrast, I find that the high-profitability portfolio has a significantly positive loading; in addition, I observe an increasing pattern in the betas from the low portfolio to the high portfolio. These portfolios present an upward-sloping pattern of covariances with the empirical measure of the long-run productivity shock. All these results thus support my model in that high-cash-based-operating-profitability firms predict higher stock returns because they have positive betas on the short- and long-run productivity shock that are positively priced.

## 5 Other Explanations for the Cash-based Operating Profitability Premium

In this section, I first provide additional empirical evidence for the relation between cash-based operating profitability and future stock returns. To do so, I perform several factor regressions to show that such a positive-return relation is literally unaffected by known return factors for other systematic risks. Moreover, I implement Fama and MacBeth (1973) regressions to examine if the positive relation between cash-based operating profitability and stock returns is mitigated by other firm characteristics. Second, I implement double sorting on size and cash-based operating profitability to confirm that the cash-based operating

profitability premium is not driven by the size effect. In addition to aggregate short-run and long-run productivity shocks as discussed in the previous section, I acknowledge that the positive cash-based-profitability-return relation could also be attributed to other explanations, including behavioral reasons, financial distress, and relevant risks documented in the literature. I discuss all these possible explanations in this section.

## 5.1 Factor Regressions and Risk Exposures

In this subsection, I investigate the extent to which the variation in the average returns of the profitability-sorted portfolios can be explained by exposures to standard risk factors, including the market factor in the CAPM model, the three factors in Fama and French (1993) (FF3), the four factors in Carhart (1997) (FF4), the five factors in Fama and French (2015) (FF5), and the four factors in Hou, Xue, and Zhang (2015) (HXZ).<sup>18</sup> To adjust for risk exposure, I perform time-series regressions of profitability-sorted portfolios' excess returns on the market factor (MKT) as the CAPM model in Panel A, on Fama-French three factors (MKT, the size factor-SMB, and the value factor-HML) in Panel B, on Carhart four factors (MKT, SMB, HML, and the momentum factor-UMD) in Panel C, on the Fama and French (2015) five factors (MKT, SMB, HML, the profitability factor-RMW, and the investment factor-CMA) in Panel D, and on the Hou, Xue, and Zhang (2015) q-factors (MKT, SMB, the investment factor-I/A, and the profitability factor-ROE) in Panel E, respectively. Such time-series regressions enable us to estimate the betas (i.e., risk exposures) of each portfolio's excess return on various risk factors and to estimate each portfolio's risk-adjusted return (i.e., alphas). These betas, together with annualized alphas (in %), are reported in Table 11.

[Place Table 11 about here]

First, as I show in Table 11, the positive profitability-return relation cannot be explained by the market factor in the CAPM model. The market betas are flat across quintile portfolios sorted on cash-based operating profitability, suggesting that high-profitability firms do not face higher market risk exposure. In addition, the intercept (i.e., alpha or risk-adjusted return) of the high-minus-low portfolio is 5.55% with a  $t$ -statistic of 3.20, which is both statistically and economically significant. Second, in Panels B to E, the risk-adjusted returns of the cash-based-operating-profitability-sorted high-minus-low portfolio remain large and significant, ranging from 4.48% for the HXZ model in Panel E to 7.67% for the FF3 model in Panel B, with  $t$ -statistics well above 3. Lastly, the high-minus-low portfolio carries signifi-

---

<sup>18</sup>The Fama and French factors are downloaded from Kenneth French's data library ([http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html)), and the HXZ factors are downloaded Global-q data library (<http://global-q.org/index.html>)

cant loadings on most risk factors except the investment factor. In summary, results from factor regressions in Table 11 suggest that the cross-sectional return spread across portfolios sorted on cash-based operating profitability cannot be eliminated by existing risk factors. Hence, common risk exposure cannot explain the positive profitability-return relation that I document.

I then examine how the time-series pattern of the risk-adjusted returns with respect to the Fama-French five-factor model and HXZ q-factor model, respectively, affects the return on the high-minus-low portfolio, which is my proxy for the cash-based operating profitability premium. Figure 1 plots the cumulative, risk-adjusted returns of the high-minus-low portfolio from July of 1980 to June of 2018. The positive profitability-return relation that I find appears to be a fairly persistent pattern across most years and increasing during economic downturns (denoted by economic recessions in shaded areas).

[Place Figure 1 about here]

## 5.2 Fama-MacBeth Regressions

I further investigate the predictive ability of cash-based operating profitability for stock returns using Fama-MacBeth cross-sectional regressions (Fama and MacBeth (1973)). This analysis allows me to control for an extensive list of firm characteristics that predict stock returns and to further examine whether the positive profitability-return relation is driven by other known predictors at the firm level.<sup>19</sup> I conduct cross-sectional regressions for each month from June of 1980 to July of 2018. In each month, monthly returns of individual stock returns (annualized by multiplying 12) are regressed on the cash-based operating profitability and different sets of control variables that are known 6 months prior to portfolio formation, except market capitalization, and industry fixed effects.<sup>20</sup> Control variables include the natural logarithm of market capitalization (Size), the natural logarithm of book-to-market ratio (B/M), investment rate (I/K), return on equity (ROE), and industry dummies based on Fama and French (1997) 30 industry classifications. All independent variables are normalized to a zero mean and one standard deviation after winsorization at the 1st and 99th percentiles to reduce the impact of outliers.

[Place Table 12 about here]

---

<sup>19</sup>Fama-MacBeth cross-sectional regressions provide a reasonable cross-check for the portfolio tests, as it is difficult to include multiple sorting variables with unique information about future stock returns by using a portfolio approach.

<sup>20</sup>Market capitalization is known at the end of every March, June, September, and December.

In Table 12, I report the average slopes (i.e., coefficient estimates) from monthly regressions, and the corresponding t-statistics are the average slopes divided by their time-series standard errors.<sup>21</sup> I annualize the slopes and standard errors. My results support the return predictive ability of cash-based operating profitability. In Specification 1, cash-based operating profitability significantly and positively predicts future stock returns with a slope coefficient of 3.00 with a t-statistic of 4.05. This finding implies that a one-standard-deviation increase in cash-based-operating profitability leads to a significant increase of 3.00% in the annualized stock return. As the difference in average cash-based operating profitability between two extreme quintile portfolios is approximate to 1.05 standard deviations, the slope of 3.00 in Specification 1 implies a return spread of 3.15% per annum, which approximately accounts for 75% of the annualized high-minus-low portfolio return of 4.55% as I report in Panel A of Table 2.<sup>22</sup>

From Specification 2 to 5, cash-based operating profitability positively predicts stock returns with a statistically significant slope when I further control for variables known to predict stock returns in the cross-section, including size, book-to-market ratio, investment rate, and return on equities (ROE). Additionally, Specification 4 highlights that the predictive ability of cash-based operating profitability is not subsumed by ROE. When I put all control variables together in Specification 6, the slope of the coefficient on cash-based operating profitability is still significantly positive. Overall, Table 12 suggests that the positive profitability-return relation cannot be attributed to other known predictors in the literature and confirms that cash-based operating profitability has a unique return predictive power.

### 5.3 Double Sorting on Size

To alleviate the concern that the return predictability I document is driven by firm size, I conduct two-way independent sorts for cash-based operating profitability and size. At the beginning of January, April, July, and October, I assign firms into big (B) and small (S) groups based on their market capitalization relative to industry peers and group firms into quintile portfolios (from low to high) based on their cash-based operating profitability relative to industry peers. I then track the value-weighted returns on each portfolio over the next three months. I report the annualized portfolio returns by multiplying 12 (and *t*-statistics in parentheses) in Table 13.

[Place Table 13 about here]

---

<sup>21</sup>My standard errors are based on Newey and West (1987).

<sup>22</sup>The Fama-MacBeth regressions weigh each observation equally and thus place substantial weight on small firms. However, my finding for the cash-based profitability premium from sorted portfolios is mainly based on value-weighted portfolios.



If the size is responsible for the cash-based operating profitability effect, then I would expect the return spread to concentrate on the small or big group. However, as shown in Table 13, high-profitability firms still outperform low-profitability ones in stock returns for both big and small groups. Moreover, the returns on the high-minus-low portfolios are both economically and statistically significant among big and small firms. Consequently, a positive profitability-return relation is not completely driven by the size effect.

## 5.4 Behavioral Explanations

### 5.4.1 Institutional Ownership

In this subsection, this paper aims to disentangle the heterogeneous information precision to update the aggregate long-run productivity in the main analysis from the alternative short-sale constraint explanation by using institutional ownership. The idea is that institutional ownership is widely used as a proxy for short-sale constraints in the literature, according to Nagel (2005) and Asquith et al. (2005). It has been studied in the short-sale constraint literature that stocks with higher institutional ownership are less subject to short-sale constraints as institutional investors are important suppliers of shares to borrow for short selling. Supposed that short-sale constraint is the driving force that explains the heterogeneity in the information precision story, one would expect the information precision effect to be stronger among stocks with lower institutional ownership but to dampen or even disappeared among stocks with higher institutional ownership.

To test this explanation, I calculate the institutional investors' ownership before examining whether the profitability-return relation varies across different types of institutional ownership. If the explanation based on institutional ownership holds, I expect that the profitability-driven return predictability will be absorbed by institutional investors' ownership. In particular, I collect quarterly institutional ownership data at the end of each quarter from the Thomson Reuters Institutional Holdings (13F) database, which keeps track of the 13F filings of professional money managers. An institutional investment manager with investment discretion over 100 million or more is required to file Form 13F with the SEC within 45 days at the end of a calendar quarter on the number of shares they hold of stocks. Institutional investors include investment advisers, banks, insurance companies, broker-dealers, pension funds, etc. A firm's institutional ownership is calculated as the ratio of the summation of its shares held by institutional investors to total shares outstanding in a given quarter. With the alternative explanation by the short-sale constraint, I would expect the information precision effect to get weaker with higher institutional ownership since institutional ownership is associated with less binding short-sale constraints.

I form double-sorted portfolios based on firms’ cash-based profitability and institutional ownership. In particular, I independently sort firms into two portfolios based on their institutional ownership, and into five portfolios based on their cash-based profitability at the beginning of every January, April, July, and October, all relative to industry peers. I then calculate the value-weighted returns on each portfolio over the next three months. I present the average returns of my double-sorted (2 by 5) portfolios in Panel A of Table 14. I include  $t$ -statistics in parentheses and annualize the portfolio returns by multiplying them by 12. In the high-intuitional-ownership group, the return spread sorted on cash-based operating profitability amounts to 3.07% and is significant at the 1% level. On the other hand, in the low-institutional-ownership group, the return spread based on two extreme portfolios sorted on cash-based operating profitability amounts to 6.96% that is significant at the 5% level with a  $t$ -statistic of 2.09. As the profitability-related return predictability is not eliminated when I control for institutional ownership, the cash-based operating profitability premium cannot be attributed to the difference in short-sale constraint proxied by institutional investors’ different shareholdings.

[Place Table 14 about here]

#### 5.4.2 Under-reaction to Earnings Report

It is well documented in the literature that investors may under-react to market news due to limited attention or lags in information diffusion.<sup>23</sup> It is possible that firms with high cash-based operating profitability are subject to great pressure from analysts and activist investors and are thus more likely to decorate their profitability in the next period. If high-profitability firms’ window dressing is under-reacted (underestimated) by investors who are myopic to pick stocks and rely on firms’ temporary performance, their stock prices may increase in the future and will result in the profitability-return relation that I find. To further examine this possibility, I implement the two-way portfolio sorting based on firms’ current and future cash-based operating profitability. At the beginning of every January, April, July, and October, I first sort stocks into five portfolios (from low “L” to high “H”) based on firms’ cash-based operating profitability (i.e., current profitability). Then, the firms in the highest quintile portfolios are further sorted into two portfolios based on their cash-based operating

---

<sup>23</sup>Prior studies suggest that investors tend to underreact to new information (e.g., Bernard and Thomas (1990)), especially for new but complex information (e.g., You and Zhang (2009)). For example, in the innovation literature, the evidence suggests that investors tend to over discount the cash flow prospects of R&D-intensive or patenting firms owing to high uncertainty and complexity associated with innovations or fail to take into account the benefits of innovation due to limited attention, which results in underpricing of innovation (see, e.g., Hall (1993); Lev and Sougiannis (1996); Aboody and Lev (1998, 2000); Chan et al. (2001); and Hirshleifer et al. (2013, 2017)).

profitability in the next quarter (i.e., future profitability):<sup>24</sup> the HL portfolio includes firms with future cash-based operating profitability below the median of the high group and the HH group includes those with future cash-based operating profitability above the median of the high group. If the under-reaction explanation is correct, then the profitability-return relation should only exist in the HL group but not in the HH group.

I then calculate the value-weighted portfolio returns of these two portfolios for the following 12 months and report their time-series average returns in the second and third columns of Panel B of Table 14. I include *t*-statistics in parentheses. I also report the average portfolio return in the lowest quintile portfolio (L), and also present the return difference between the HL and L groups and the return difference between the HH and L groups. My empirical results show that although the former (HL-L) is significantly positive on average (4.18% with a *t*-statistic of 2.35), the latter (HH-L) is also significantly positive on average (5.01% with a *t*-statistic of 2.51). In other words, even high-cash-based-operating-profitability firms that do not improve their cash-based operating profitability in the future still provide significantly higher returns than low-cash-based-operating-profitability firms. Hence, the under-reaction explanation is less likely to explain the cross-sectional variation in stock returns due to cash-based operating profitability.

### 5.4.3 Retail Investors' Behavioral Bias

Different from institutional investors who are more rational and have more complete information, retail investors may be more subject to behavioral bias (e.g., Daniel et al. (1998), Barberis et al. (1998), and Hong and Stein (1999), among others.). For example, retail investors may be sensitive and overreact to news about some local news (Ivković and Weisbenner (2005) and Seasholes and Zhu (2010)) and buy (sell) all their stocks at a huge premium (deep discounts). If such overreaction explains the cash-based operating profitability premium, then I expect the positive profitability-return relation to existing among stocks that experience significant jumps (drops) in the shares of retail investors. To examine this explanation, I first define the shares of retail investors in percentage as one minus the shares owned by institutional investors in percentage at the end of each quarter. At the beginning of every January, April, July, and October, I first sort all stocks with cash-based operating profitability into three portfolios by 30-40-30 based on the changes in retail investors on the report date of quarterly earnings (RDQ). The high (low) group includes stocks that experience the strongest increase (decrease) in retail investors' shares. Then, within each group, I further sort stocks into quintile portfolios based on firms' cash-based operating profitability within a particular industry. Moreover, within each portfolio of changes in retail investors'

---

<sup>24</sup>I present the transition matrix in Section A.6 of the Internet Appendix.

shares, I also form a high-minus-low portfolio (H-L) that takes a long position in the high quintile portfolio and a short position in the low quintile portfolio. As a result, I form total of 18 portfolios.

In Panel C of Table 14, I report the annualized monthly averages of value-weighted returns on all portfolios. I include  $t$ -statistics in parentheses. In addition, I report the mean (value-weighted) and the median of each group of the changes in retail investors' shares. I first find that, within the middle tercile (Group 2), the return spread (4.02% with a  $t$ -statistic of 2.27) is significant and comparable to that in the univariate portfolio sorting. In addition, the change in retail investors' shares is close to zero in the middle tercile (the mean and median are 0.29% and 0.19%, respectively). On the other hand, within the group of the lowest or highest changes (Group 1 or 3) in retail investors' shares, the return spread (i.e., the returns on the H-L portfolio) is *insignificant*. These results suggest that the profitability-return relation is orthogonal to the ownership of retail investors that are more subject to overreaction bias.

## 5.5 Financial Distress

Prior literature documents a negative relationship between the firm's financial distress and future stock returns in the cross-section. In addition, the cash-based operating profitability and financial distress are negatively correlated. Thus, part of the link between the firm's cash-based operating profitability and future stock returns may reflect the negative correlation between financial distress and future stock returns. In this subsection, I extend the previous analysis to investigate the joint link between cash-based profitability, financial distress, and future stock returns in the cross-section.

Another possible explanation for the profitability-return relation is that firms with low cash-based operating profitability could be close to financial distress (Griffin and Lemmon (2000), Vassalou and Xing (2004), Campbell et al. (2008), Da and Gao (2010), Liu and Wen (2017)); Existing studies document that financial distress is negatively related to subsequent returns.<sup>25</sup> If such the financial distress channel is responsible for the profitability effect, I would expect that there is no profitability-return relation within firms close to financial distress. To examine this explanation, I double sort firms' Z index, O index, distance to default, and failure probability into two portfolios (low and high) and firms' cash-based

---

<sup>25</sup>Potential explanations for the negative relation between distress risk and expected returns include firm's endogenous choice of leverage (George and Hwang (2010)), violations of shareholder priority in bankruptcy (Garlappi and Yan (2011)), and investors' preference for skewed and lottery-like payoffs (Conrad et al. (2014)), investor overconfidence (Gao et al. (2018)), and investors underreaction due to delayed prices (Liu and Wen (2017)).

operating profitability into quintile portfolios (low, 2, 3, 4, and high), all relative to their industry peers.<sup>26</sup> Moreover, within each governance portfolio, I also form a high-minus-low portfolio (H-L) that takes a long position in the high-cash-based-operating-profitability portfolio and a short position in the low counterpart portfolio. As a result, I form total of 12 portfolios.

[Place Table 15 about here]

### 5.5.1 O Index

Following [Ohlson \(1980\)](#) and [Griffin and Lemmon \(2000\)](#), I measure a firm's financial distress by using the O index. A high value of the O index suggests that the firm is subject to a dismal status financially. As reported in Panel B of Table 15, I show that the return spread from cash-based operating profitability is significantly positive in both less and more financially distress groups. The return spread is 3.56% (with a t-statistic above 2) in the high financial distress (i.e., high O index) group, and is 4.61% (with a t-statistic of 2.12) in the low financial distress (i.e., high Z index) group. Within each O index bins (within the row), firms with high cash-based operating profitability earn higher returns on average than firms with low cash-based operating profitability. Thus, the cash-based operating profitability contains some information about future stock returns that are not contained in the financial distress as measured by the O index.

### 5.5.2 Z Index

I consider the financial distress measure of the Z index, according to [Altman \(1968\)](#) and [Griffin and Lemmon \(2000\)](#). A lower value of the Z index implies that the firm is subject to financial distress to a greater extent. In Panel B of Table 15, I show that the return spread from cash-based operating profitability is significantly positive in both less and more financially distress groups. The return spread is 2.67% (with a t-statistic of 2.11) in the high financial distress (i.e., low Z index) group, and is 4.61% (with a t-statistic of 2.12) in the low financial distress (i.e., high Z index) group. Based on the finding that financially distressed firms' cash-based operating profitability still predict stock returns, I can rule out the measure of the Z index to explain the cash-based operating profitability premium.

---

<sup>26</sup>The detailed information of the O indexes, Z indexes, default probability, and failure probability refers to [Ohlson \(1980\)](#), [Altman \(1968\)](#), [Bharath and Shumway \(2008\)](#), and [Campbell, Hilscher, and Szilagyi \(2008\)](#), respectively.

### 5.5.3 Default Probability

To measure a firm’s financial distress based on the [Merton \(1974\)](#) model, I construct the firm-level measure of default probability. A high default probability suggests that the firm is likely to undergo financial distress. In Panel C of [Table 15](#), I show the results of double sorting that the return spread from cash-based operating profitability is significantly positive in both high and low default probability groups. The return spread is sizable and amounts to 9.06% at 1% significance level in the high default probability group, and is 2.95% (with a t-statistic of 2.21) in the low counterpart group. To briefly summarize findings in Panel C, the high quintile portfolio sorted cash-based operating profitability earns a higher return on average than the low quintile portfolio. As the result, I can exclude the possible explanation based on the measure of default probability.

### 5.5.4 Failure Probability

Turning to the two-way-sorted portfolios on cash-based operating profitability and failure probability, I follow [Campbell, Hilscher, and Szilagyi \(2008\)](#) to construct firm-level failure probability. As reported in Panel D of [Table 15](#), I present the returns of the 12 portfolios independently sorted on failure probability and cash-based operating profitability, respectively. Within both high and low failure probability groups, firms with high cash-based operating profitability significantly earn higher returns than firms with low cash-based operating profitability. Therefore, such a finding suggests that the profitability-return relation is not driven by a difference in failure probability.

Taken together, the return spread sorted on the cash-based operating profitability remains robust and suggests that the profitability-return relation is not driven by financial distress.

## 5.6 Other Risk-based Explanations

In this subsection, I also explore possible explanations based on systematic risks posited in prior studies. In particular, I consider five alternative channels that may drive variations in my cash-based-operating-profitability-sorted portfolios, including technology obsolescence ([Lin, Palazzo, and Yang \(2019\)](#)), value and profitability ([Fama and French \(2015\)](#), [Hou, Xue, and Zhang \(2015\)](#)), financial constraint ([Li \(2011\)](#), [Lins et al. \(2017\)](#)), economic and political uncertainty ([Brogaard and Detzel \(2015\)](#), [Bali, Brown, and Tang \(2017\)](#)), and adjustment costs ([Kim and Kung \(2016\)](#) and [Gu, Hackbarth, and Johnson \(2017\)](#)). I elaborate upon these alternative explanations as follows. First, high-profitability firms adopt more obsolete technology and invest in less advanced capital in production. The arrival of new technology

forces these firms to upgrade their capital, and, therefore, their cash flows are sensitive to the frontier technology shock. Second, high-profitability firms invest more because of more growth opportunities, or earning-based profitability measures, return on equity, contains similar information as cash-based operating profitability. Third, low-profitability firms may be subject to risk associated with financial constraints due to poor operating performance. In addition, these firms may be subject to risk associated with macroeconomic uncertainty (such as economic downturn or trade conflict) and political uncertainty (such as changes of the ruling party). Finally, high-profitability firms earn higher expected returns because it is costly for them to adjust their capital stock, especially during economic downturns.

To examine if the predictive ability of cash-based operating profitability can be attributed to other explanations of risk, I implement independent two-way sorting by assigning all sample firms by their values in a proxy for an alternative explanation (relative to their industry peers) into two groups (low and high) and by their cash-based operating profitability into quintile portfolios (low, 2, 3, 4, and high). In addition, within each portfolio sorted by the proxy for a risk-based explanation, I also form a high-minus-low portfolio (H-L) that takes a long position in the high-cash-based-operating-profitability portfolio and a short position in the low-cash-based-operating-profitability portfolio. As a result, I form total of 12 portfolios for each risk-based explanation. I then report the annualized monthly averages of value-weighted returns on all 12 portfolios in Table 16.

[Place Table 16 about here]

### 5.6.1 Technology obsolescence

I consider the capital age and investment rate to measure firm-level technology obsolescence, according to Lin, Palazzo, and Yang (2019) and Li, Tsou, and Xu (2019). A firm with an older capital age or a lower investment rate faces higher exposures to technology frontier shocks and, therefore, is riskier. In Panel A of Table 16, I show that the return spread from cash-based operating profitability (i.e., return on the high-minus-low portfolio) remains comparable to that in the univariate portfolio sort in both less and more capital age (investment rate) groups. The return spread is 7.22% (with a  $t$ -statistic of 3.20) in the young capital age group, and is 3.00% (with a  $t$ -statistic of 2.40) in the old capital age group. On the other hand, the return spread is 3.25% (with a  $t$ -statistic of 2.37) in the low investment rate group and is 5.81% (with a  $t$ -statistic of 3.18) in the high investment rate group. If technology obsolescence or investment channel is the main driving force for the cash-based operating profitability premium, then I should only observe significant return spreads in the old capital age and low investment rate group. In contrast, the return spreads are significant in the

young capital age and high investment rate groups. Therefore, the return spread sorted on the cash-based operating profitability cannot be explained by technology obsolescence.

### 5.6.2 Book-to-market and Return on Equity

Furthermore, I also form two sets of twelve portfolios two-way-sorted on cash-based operating profitability and book-to-market ratio and cash-based operating profitability and return on equity, respectively. To briefly summarize the main findings, Panel B of Table 16 presents the returns of the 12 portfolios sorted on book-to-market ratio and return on equity, respectively. Within both high and low book-to-market ratio (return on equity) group, the cash-based operating profitability spread remains significantly positive. These findings suggest that the profitability-return relation is not driven by known predictors in the literature, including book-to-market ratio and return on equity.

### 5.6.3 Financial Constraints

I consider the financial constraints measures of [Whited and Wu \(2006\)](#) and [Hadlock and Pierce \(2010\)](#).<sup>27</sup> A higher value of the SA or WW index suggests that the firm is subject to financial constraints to a greater extent. In Panel C of Table 16, I show that the return spread from cash-based operating profitability is significantly positive in both less and more financially constrained groups. When I use the SA index, the return spread is 3.55% (with a t-statistic of 2.41) in the low constraint group and is 8.82% (with a t-statistic of 3.73) in the high constraint group. When I use the WW index, the return spread is 3.74% (with a t-statistic of 2.47) in the low constraint group and is 11.25% (with a t-statistic of 4.82) in the high constraint group. The fact that financially unconstrained firms' cash-based operating profitability still predict stock returns suggests that financial constraints cannot explain the cash-based operating profitability premium.

### 5.6.4 Economic and Political Uncertainty

I follow [Bali, Brown, and Tang \(2017\)](#) to estimate firm-level exposure with respect to the macroeconomic uncertainty index based on [Jurado, Ludvigson, and Ng \(2015\)](#) and to the political uncertainty index based on [Bloom \(2009\)](#) by using rolling window regressions.<sup>28</sup> The

---

<sup>27</sup>Detailed information regarding the construction for the SA and WW indexes refers to [Farre-Mensa and Ljungqvist \(2016\)](#).

<sup>28</sup>For each stock with non-missing cash-based-operating profitability in each month in my sample, I estimate the uncertainty exposure from the monthly regressions of excess returns on the macroeconomic uncertainty index over a 60-month rolling window by controlling for empirical risk factors, including the market (MKT), size (SMB), value (HML), momentum (UMD), liquidity (LIQ), investment (I/A), and profitability (ROE).



results in the left and right parts of Panel D of Table 16 present the returns of the 12 portfolios sorted on macroeconomic uncertainty and political uncertainty, respectively. Within both high and low macroeconomic (political) uncertainty exposure group, the return spreads sorted on cash-based operating profitability are significantly positive. These findings suggest that the profitability-return relation is not driven by different exposures to macroeconomic (political) uncertainty.

### 5.6.5 Adjustment Costs

I follow the method of Kim and Kung (2016) and Gu, Hackbarth, and Johnson (2017) in measuring a firm’s asset redeployability and inflexibility, respectively.<sup>29</sup> If the adjustment costs in asset redeployability (inflexibility) indeed drive the cash-based operating profitability premium, I would expect that such premium would not exist in firms in the high asset redeployability (low inflexibility) group that is associated with lower adjustment costs. However, as shown in Panel D of Table 16, the return spread sorted on cash-based operating profitability is 6.36% with a t-statistic of 2.44 in the low asset redeployability group and 2.25% with a t-statistic of 2.01 in the high inflexibility group. On the other hand, the return spread sorted on cash-based operating profitability amounts to 3.28% at the 1% significance level in the high asset redeployability group. The return spread in the low inflexibility group is also significant at 1% and amounts to 9.43%. The fact that the profitability-return relation appears significantly positive in both high asset redeployability and low inflexibility groups suggests that the return predictability I document is unrelated to the systematic risk associated with adjustment costs.

Overall, I find that high-cash-based operating-profitability firms earn higher stock returns than the lower counterparts in all groups that stand for low exposures to systematic risks documented in the literature. All these results collectively point to a unique role of cash-based operating in return predictability on the cross-section.

## 6 Conclusion

The awareness of informational friction that distorts the allocation of resources and distresses productivity and output across heterogeneous firms has been documented in the macroeconomic literature as an important driver in the propagation and the amplification of shocks in the economy. This paper investigates the implications of cash-based operating profitability on the cross-section of stock returns. A long-short portfolio constructed from firms

---

<sup>29</sup>Detailed information regarding the construction for the asset redeployability index refers to Table 16.

with high versus low cash-based operating profitability relative to their industry peers generates an average excess return of around 4.54% per year. I also find that firms with high cash-based operating profitability are with more precise information about their exposure to common productivity shock for higher allocation efficiency and with better performance of the investment-q regression. Together with findings suggest that these firms' productivity growth exhibit higher co-movements with aggregate productivity growth.

To explain my empirical finding of a cash-based operating profitability premium, I consider an economy with imperfect information and develop a dynamic model economy with a large cross-section of firms. The long-run productivity is unobservable and evolves over time, while the learning takes place when firms can receive signals and update their beliefs about the long-run productivity. A high information precision causes firms to update their beliefs faster, which endogenously produce more variations in these firms' q, and generates higher exposure to aggregate productivity shocks. Therefore, these firms earn higher average excess returns as risk premia. Further empirical analyses provide supportive evidence to my model assumptions and implications. First, these firms' payout and realized returns exhibit higher co-movements with aggregate productivity shocks. Second, both two aggregate shock carries a positive price of risk.

In addition to the learning mechanism in my model, I document that the return spread cannot be explained by existing risk factors, including the Fama-French five-factor model (Fama and French (2015)) and the HXZ q-factor model (Hou, Xue, and Zhang (2015)). Fama and MacBeth (1973) regressions provide a valid cross-check for the positive relation between cash-based operating profitability and stock returns. Next, I examine other systematic risk explanations, including capital age, investment rate, value, return on equity, financial constraints, macroeconomic and political uncertainty, and asset redeployability and inflexibility. Moreover, I also consider other possibilities including institutional ownership, behavioral bias, and financial distress. None of these alternative explanations can fully explain the cash-based-operating-profitability-return relation. As a result, this study highlights that imperfect information and learning can have a significant impact on asset prices.

## References

- Abel, Andrew B, 1983, Optimal investment under uncertainty, *The American Economic Review* 73, 228–233.
- Abel, Andrew B, 2018, The effects of q and cash flow on investment in the presence of measurement error, *Journal of Financial Economics* 128, 363–377.
- Abel, Andrew B, and Janice C Eberly, 1993, A unified model of investment under uncertainty, *American Economic Review* .
- Abel, Andrew B, and Janice C Eberly, 1997, An exact solution for the investment and value of a firm facing uncertainty, adjustment costs, and irreversibility, *Journal of Economic Dynamics and Control* 21, 831–852.
- Aboody, David, and Baruch Lev, 1998, The value relevance of intangibles: The case of software capitalization, *Journal of Accounting research* 36, 161–191.
- Aboody, David, and Baruch Lev, 2000, Information asymmetry, r&d, and insider gains, *The Journal of Finance* 55, 2747–2766.
- Ai, Hengjie, 2010, Information quality and long-run risk: Asset pricing implications, *The Journal of Finance* 65, 1333–1367.
- Ai, Hengjie, Mariano Massimiliano Croce, and Kai Li, 2013, Toward a quantitative general equilibrium asset pricing model with intangible capital, *The Review of Financial Studies* 26, 491–530.
- Ai, Hengjie, Mariano Max Croce, Anthony M. Diercks, and Kai Li, 2018, News shocks and the production-based term structure of equity returns, *The Review of Financial Studies* hhy015.
- Ai, Hengjie, Jun E Li, Kai Li, and Christian Schlag, 2019, The collateralizability premium .
- Ali, Usman, and David Hirshleifer, 2019, Shared analyst coverage: Unifying momentum spillover effects, *Journal of Financial Economics* .
- Altman, Edward I, 1968, Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *The journal of finance* 23, 589–609.
- Andrei, Daniel, William Mann, and Nathalie Moyen, 2019, Why did the q theory of investment start working?, *Journal of Financial Economics* 133, 251–272.

- Angeletos, George-Marios, and Iván Werning, 2006, Crises and prices: Information aggregation, multiplicity, and volatility, *American Economic Review* 96, 1720–1736.
- Asquith, Paul, Michael B Mikhail, and Andrea S Au, 2005, Information content of equity analyst reports, *Journal of Financial Economics* 75, 245–282.
- Bali, Turan G, Stephen J Brown, and Yi Tang, 2017, Is economic uncertainty priced in the cross-section of stock returns?, *Journal of Financial Economics* 126, 471–489.
- Ball, Ray, Joseph Gerakos, Juhani T Linnainmaa, and Valeri Nikolaev, 2016, Accruals, cash flows, and operating profitability in the cross section of stock returns, *Journal of Financial Economics* 121, 28–45.
- Ball, Ray, Joseph Gerakos, Juhani T Linnainmaa, and Valeri V Nikolaev, 2015, Deflating profitability, *Journal of Financial Economics* 117, 225–248.
- Banerjee, Snehal, 2011, Learning from prices and the dispersion in beliefs, *The Review of Financial Studies* 24, 3025–3068.
- Bansal, Ravi, and Ivan Shaliastovich, 2013, A long-run risks explanation of predictability puzzles in bond and currency markets, *The Review of Financial Studies* 26, 1–33.
- Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307–343.
- Bartelsman, Eric J, and Mark Doms, 2000, Understanding productivity: Lessons from longitudinal microdata, *Journal of Economic Literature* 38, 569–594.
- Barth, Mary E, and Amy P Hutton, 2004, Analyst earnings forecast revisions and the pricing of accruals, *Review of Accounting Studies* 9, 59–96.
- Belo, Frederico, Jun Li, Xiaoji Lin, and Xiaofei Zhao, 2017, Labor-force heterogeneity and asset prices: The importance of skilled labor, *The Review of Financial Studies* 30, 3669–3709.
- Belo, Frederico, and Xiaoji Lin, 2012, The inventory growth spread, *The Review of Financial Studies* 25, 278–313.
- Belo, Frederico, Xiaoji Lin, and Santiago Bazdresch, 2014, Labor hiring, investment, and stock return predictability in the cross section, *Journal of Political Economy* 122, 129–177.

- Belo, Frederico, Xiaoji Lin, and Fan Yang, 2018, External Equity Financing Shocks, Financial Flows, and Asset Prices .
- Berk, Jonathan B, Richard C Green, and Vasant Naik, 1999, Optimal investment, growth options, and security returns, *The Journal of finance* 54, 1553–1607.
- Bernard, Victor L, and Jacob K Thomas, 1990, Evidence that stock prices do not fully reflect the implications of current earnings for future earnings, *Journal of Accounting and Economics* 13, 305–340.
- Bharath, Sreedhar T, and Tyler Shumway, 2008, Forecasting default with the merton distance to default model, *The Review of Financial Studies* 21, 1339–1369.
- Bloom, Nicholas, 2009, The impact of uncertainty shocks, *Econometrica* 77, 623–685.
- Blundell, Richard, and Stephen Bond, 2000, Gmm estimation with persistent panel data: an application to production functions, *Econometric reviews* 19, 321–340.
- Brainard, William C, and James Tobin, 1968, Pitfalls in financial model building, *The American Economic Review* 58, 99–122.
- Brennan, Michael J, Narasimhan Jegadeesh, and Bhaskaran Swaminathan, 1993, Investment analysis and the adjustment of stock prices to common information, *The Review of Financial Studies* 6, 799–824.
- Brogaard, Jonathan, and Andrew Detzel, 2015, The asset-pricing implications of government economic policy uncertainty, *Management Science* 61, 3–18.
- Campbell, John Y, and John H Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of political Economy* 107, 205–251.
- Campbell, John Y, Jens Hilscher, and Jan Szilagyi, 2008, In search of distress risk, *The Journal of Finance* 63, 2899–2939.
- Carhart, Mark M, 1997, On persistence in mutual fund performance, *The Journal of finance* 52, 57–82.
- Chan, Kalok, and Allaudeen Hameed, 2006, Stock price synchronicity and analyst coverage in emerging markets, *Journal of Financial Economics* 80, 115–147.
- Chan, Louis KC, Josef Lakonishok, and Theodore Sougiannis, 2001, The stock market valuation of research and development expenditures, *The Journal of Finance* 56, 2431–2456.

- Chen, Kaiji, and Zheng Song, 2013, Financial frictions on capital allocation: A transmission mechanism of tfp fluctuations, *Journal of Monetary Economics* 60, 683–703.
- Clementi, Gian Luca, and Bernardino Palazzo, 2019, Investment and the cross-section of equity returns, *The Journal of Finance* 74, 281–321.
- Cochrane, John H, 1991, Production-based asset pricing and the link between stock returns and economic fluctuations, *The Journal of Finance* 46, 209–237.
- Cochrane, John H, 2005, *Asset pricing: Revised edition* (Princeton university press).
- Conrad, Jennifer, Nishad Kapadia, and Yuhang Xing, 2014, Death and jackpot: Why do individual investors hold overpriced stocks?, *Journal of Financial Economics* 113, 455–475.
- Croce, Mariano Max Massimiliano, Tatyana Marchuk, and Christian Schlag, 2018, The leading premium, *Working Paper* .
- Da, Zhi, and Pengjie Gao, 2010, Clientele change, liquidity shock, and the return on financially distressed stocks, *Journal of Financial and Quantitative Analysis* 45, 27–48.
- Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under-and overreactions, *the Journal of Finance* 53, 1839–1885.
- David, Alexander, 1997, Fluctuating confidence in stock markets: Implications for returns and volatility, *Journal of Financial and Quantitative Analysis* 427–462.
- David, Joel M, Hugo A Hopenhayn, and Venky Venkateswaran, 2016, Information, misallocation, and aggregate productivity, *The Quarterly Journal of Economics* 131, 943–1005.
- David, Joel M, Lukas Schmid, and David Zeke, 2018, Risk-adjusted capital allocation and misallocation, *Working Paper* .
- Dechow, Patricia M, Scott A Richardson, and Richard G Sloan, 2008, The persistence and pricing of the cash component of earnings, *Journal of accounting research* 46, 537–566.
- Dechow, Patricia M, Richard G Sloan, and Mark T Soliman, 2004, Implied equity duration: A new measure of equity risk, *Review of Accounting Studies* 9, 197–228.
- Diether, Karl B, Christopher J Malloy, and Anna Scherbina, 2002, Differences of opinion and the cross section of stock returns, *The Journal of Finance* 57, 2113–2141.

- Duarte, Jefferson, Xi Han, Jarrad Harford, and Lance Young, 2008, Information asymmetry, information dissemination and the effect of regulation on the cost of capital, *Journal of Financial Economics* 87, 24–44.
- Eisfeldt, Andrea L, and Dimitris Papanikolaou, 2013, Organization capital and the cross-section of expected returns, *The Journal of Finance* 68, 1365–1406.
- Erickson, Timothy, and Toni M Whited, 2000, Measurement error and the relationship between investment and q, *Journal of political economy* 108, 1027–1057.
- Fama, Eugene F, and Kenneth R French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.
- Fama, Eugene F, and Kenneth R French, 1996, Multifactor explanations of asset pricing anomalies, *The Journal of Finance* 51, 55–84.
- Fama, Eugene F, and Kenneth R French, 1997, Industry costs of equity, *Journal of financial economics* 43, 153–193.
- Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1 – 22.
- Fama, Eugene F, and James D MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 607–636.
- Farre-Mensa, Joan, and Alexander Ljungqvist, 2016, Do measures of financial constraints measure financial constraints?, *The Review of Financial Studies* 29, 271–308.
- Feng, Ying, 2018, Firm life-cycle learning and misallocation, Technical report, V Working Paper.
- Ganguli, Jayant Vivek, and Liyan Yang, 2009, Complementarities, multiplicity, and supply information, *Journal of the European Economic Association* 7, 90–115.
- Gao, Pengjie, Christopher A Parsons, and Jianfeng Shen, 2018, Global relation between financial distress and equity returns, *The Review of Financial Studies* 31, 239–277.
- Garfinkel, Jon A, 2009, Measuring investors’ opinion divergence, *Journal of Accounting Research* 47, 1317–1348.
- Garlappi, Lorenzo, and Hong Yan, 2011, Financial distress and the cross-section of equity returns, *The journal of finance* 66, 789–822.

- Gârleanu, Nicolae, Leonid Kogan, and Stavros Panageas, 2012, Displacement risk and asset returns, *Journal of Financial Economics* 105, 491–510.
- George, Thomas J, and Chuan-Yang Hwang, 2010, A resolution of the distress risk and leverage puzzles in the cross section of stock returns, *Journal of Financial Economics* 96, 56–79.
- Gomes, Joao, Leonid Kogan, and Lu Zhang, 2003, Equilibrium cross section of returns, *Journal of Political Economy* 111.
- Gomes, Joao F, Leonid Kogan, and Motohiro Yogo, 2009, Durability of output and expected stock returns, *Journal of Political Economy* 117.
- Gonçalves, Andrei, 2019, The short duration premium, *Available at SSRN 3385579* .
- Gondhi, Naveen, 2020, Rational inattention, misallocation, and asset prices .
- Gormsen, Niels Joachim, and Eben Lazarus, 2019, Duration-driven returns, *Available at SSRN 3359027* .
- Griffin, John M, and Michael L Lemmon, 2000, Does book-to-market equity proxy for distress risk or overreaction, Technical report, Citeseer.
- Gu, Lifeng, Dirk Hackbarth, and Tim Johnson, 2017, Inflexibility and stock returns, *The Review of Financial Studies* 31, 278–321.
- Hadlock, Charles J, and Joshua R Pierce, 2010, New evidence on measuring financial constraints: Moving beyond the kz index, *The Review of Financial Studies* 23, 1909–1940.
- Hall, Bronwyn H, 1993, The stock market’s valuation of r&d investment during the 1980’s, *The American Economic Review* 83, 259–264.
- Hayashi, Fumio, 1982, Tobin’s marginal q and average q: A neoclassical interpretation, *Econometrica: Journal of the Econometric Society* 213–224.
- Hirshleifer, David, Kewei Hou, Siew Hong Teoh, and Yinglei Zhang, 2004, Do investors overvalue firms with bloated balance sheets?, *Journal of Accounting and Economics* 38, 297–331.
- Hirshleifer, David, Po-Hsuan Hsu, and Dongmei Li, 2013, Innovative efficiency and stock returns, *Journal of Financial Economics* 107, 632–654.
- Hirshleifer, David, Po-Hsuan Hsu, and Dongmei Li, 2017, Innovative originality, profitability, and stock returns, *The Review of Financial Studies* 31, 2553–2605.



- Hong, Harrison, and Marcin Kacperczyk, 2010, Competition and bias, *The Quarterly Journal of Economics* 125, 1683–1725.
- Hong, Harrison, Terence Lim, and Jeremy C Stein, 2000, Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, *The Journal of Finance* 55, 265–295.
- Hong, Harrison, and Jeremy C Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *The Journal of finance* 54, 2143–2184.
- Hou, Kewei, and Tobias J Moskowitz, 2005, Market frictions, price delay, and the cross-section of expected returns, *The Review of Financial Studies* 18, 981–1020.
- Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: An investment approach, *Review of Financial Studies* hhu068.
- Hsieh, Chang-Tai, and Peter J Klenow, 2009, Misallocation and manufacturing tfp in china and india, *The Quarterly journal of economics* 124, 1403–1448.
- Ivković, Zoran, and Scott Weisbenner, 2005, Local does as local is: Information content of the geography of individual investors' common stock investments, *The Journal of Finance* 60, 267–306.
- Jermann, Urban J, 1998, Asset pricing in production economies, *Journal of Monetary Economics* 41, 257–275.
- Jurado, Kyle, Sydney C Ludvigson, and Serena Ng, 2015, Measuring uncertainty, *American Economic Review* 105, 1177–1216.
- Kelly, Bryan, and Alexander Ljungqvist, 2012, Testing asymmetric-information asset pricing models, *The Review of Financial Studies* 25, 1366–1413.
- Keynes, John Maynard, 1936, The general theory of interest, employment and money.
- Kim, Hyunseob, and Howard Kung, 2016, The asset redeployability channel: How uncertainty affects corporate investment, *The Review of Financial Studies* 30, 245–280.
- Kogan, Leonid, and Dimitris Papanikolaou, 2012, Economic activity of firms and asset prices, *Annual Review of Financial Economics* 4, 1–24.
- Kogan, Leonid, and Dimitris Papanikolaou, 2013, Firm characteristics and stock returns: The role of investment-specific shocks, *The Review of Financial Studies* 26, 2718–2759.

- Kogan, Leonid, and Dimitris Papanikolaou, 2014, Growth opportunities, technology shocks, and asset prices, *The Journal of Finance* 69, 675–718.
- Kumar, Alok, 2009, Hard-to-value stocks, behavioral biases, and informed trading, *Journal of Financial and Quantitative Analysis* 1375–1401.
- Lev, Baruch, and Theodore Sougiannis, 1996, The capitalization, amortization, and value-relevance of r&d, *Journal of accounting and economics* 21, 107–138.
- Li, Dongmei, 2011, Financial constraints, r&d investment, and stock returns, *The Review of Financial Studies* 24, 2974–3007.
- Li, Kai, Chi-Yang Tsou, and Chenjie Xu, 2019, Learning and the capital age premium, *Available at SSRN 3225041* .
- Lin, Xiaoji, 2012, Endogenous technological progress and the cross-section of stock returns, *Journal of Financial Economics* 103, 411–427.
- Lin, Xiaoji, Bernardino Palazzo, and Fan Yang, 2019, The risks of old capital age: Asset pricing implications of technology adoption, *Journal of Monetary Economics* .
- Lindenberg, Eric B, and Stephen A Ross, 1981, Tobin’s q ratio and industrial organization, *Journal of business* 1–32.
- Lins, Karl V, Henri Servaes, and Ane Tamayo, 2017, Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis, *The Journal of Finance* 72, 1785–1824.
- Liptser, Robert, and Albert N Shiryaev, 2013, *Statistics of random Processes: I. general Theory*, volume 5 (Springer Science & Business Media).
- Liu, Laura Xiaolei, Toni M Whited, and Lu Zhang, 2009, Investment-based expected stock returns, *Journal of Political Economy* 117, 1105–1139.
- Liu, Xiaochun, and Quan Wen, 2017, Financial distress risk innovations and the distress risk-return relation, *Available at SSRN 2074674* .
- Menzly, Lior, Tano Santos, and Pietro Veronesi, 2004, Understanding predictability, *Journal of Political Economy* 112, 1–47.
- Merton, Robert C, 1974, On the pricing of corporate debt: The risk structure of interest rates, *The Journal of finance* 29, 449–470.

- Mussa, Michael, 1977, External and internal adjustment costs and the theory of aggregate and firm investment, *Economica* 44, 163–178.
- Myers, Stewart C., and Nicholas S. Majluf, 1984, Corporate financing and investment decisions when firms have information that investors do not have, *Journal of Financial Economics* 13, 187 – 221.
- Nagel, Stefan, 2005, Short sales, institutional investors and the cross-section of stock returns, *Journal of financial economics* 78, 277–309.
- Newey, WK, and KD West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.
- Ohlson, James A, 1980, Financial ratios and the probabilistic prediction of bankruptcy, *Journal of accounting research* 109–131.
- Papanikolaou, Dimitris, 2011, Investment shocks and asset prices, *Journal of Political Economy* 119, 639–685.
- Pástor, L’uboš, and Pietro Veronesi, 2006, Was there a nasdaq bubble in the late 1990s?, *Journal of Financial Economics* 81, 61–100.
- Pastor, Lubos, and Pietro Veronesi, 2009, Learning in financial markets, *Annu. Rev. Financ. Econ.* 1, 361–381.
- Pástor, L’uboš, and Pietro Veronesi, 2009, Technological revolutions and stock prices, *The American Economic Review* 99, 1451–1483.
- Pástor, L’uboš, and Pietro Veronesi, 2012, Uncertainty about government policy and stock prices, *The Journal of Finance* 67, 1219–1264.
- Pástor, L’uboš, and Pietro Veronesi, 2013, Political uncertainty and risk premia, *Journal of Financial Economics* 110, 520–545.
- Richardson, Scott A, Richard G Sloan, Mark T Soliman, and Irem Tuna, 2005, Accrual reliability, earnings persistence and stock prices, *Journal of accounting and economics* 39, 437–485.
- Santos, Tano, and Pietro Veronesi, 2010, Habit formation, the cross section of stock returns and the cash-flow risk puzzle, *Journal of Financial Economics* 98, 385–413.
- Seasholes, Mark S, and Ning Zhu, 2010, Individual investors and local bias, *The Journal of Finance* 65, 1987–2010.

- Sloan, Richard G, 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings?, *Accounting review* 289–315.
- Thomas, Jacob K, and Huai Zhang, 2002, Inventory changes and future returns, *Review of Accounting Studies* 7, 163–187.
- Tobin, James, 1969, A general equilibrium approach to monetary theory, *Journal of money, credit and banking* 1, 15–29.
- Tuzel, Selale, 2010, Corporate real estate holdings and the cross-section of stock returns, *The Review of Financial Studies* 23, 2268–2302.
- van Bingbergen, Jules, 2016, Good-specific habit formation and the cross-section of expected returns, *The Journal of Finance* .
- Vassalou, Maria, and Yuhang Xing, 2004, Default risk in equity returns, *The journal of finance* 59, 831–868.
- Veronesi, Pietro, 2000, How does information quality affect stock returns?, *The Journal of Finance* 55, 807–837.
- Weber, Michael, 2018, Cash flow duration and the term structure of equity returns, *Journal of Financial Economics* 128, 486–503.
- Whited, Toni M, and Guojun Wu, 2006, Financial constraints risk, *The Review of Financial Studies* 19, 531–559.
- Wu, Jin, Lu Zhang, and X Frank Zhang, 2010, The q-theory approach to understanding the accrual anomaly, *Journal of Accounting Research* 48, 177–223.
- Xie, Hong, 2001, The mispricing of abnormal accruals, *The accounting review* 76, 357–373.
- You, Haifeng, and Xiao-jun Zhang, 2009, Financial reporting complexity and investor under-reaction to 10-k information, *Review of Accounting studies* 14, 559–586.
- Zhang, Lu, 2005, The value premium, *The Journal of Finance* 60, 67–103.
- Zhang, X Frank, 2007, Accruals, investment, and the accrual anomaly, *The Accounting Review* 82, 1333–1363.

Figure 1: **Cumulative Abnormal Returns of the High-minus-Low Portfolio**

Cumulative abnormal returns are computed for the risk-adjusted returns (based on Fama-French five-factor model and HXZ q-factor model) on the high-minus-low portfolio sorted by cash-based-operating profitability. Fama-French five-factor model and HXZ q-factor model are defined in Table 11. I plot the time-series of the cumulative abnormal returns from an initial investment of one dollar. The shaded bands are labeled as recession periods, according to NBER recession dates. The sample period is July 1980-June 2018.

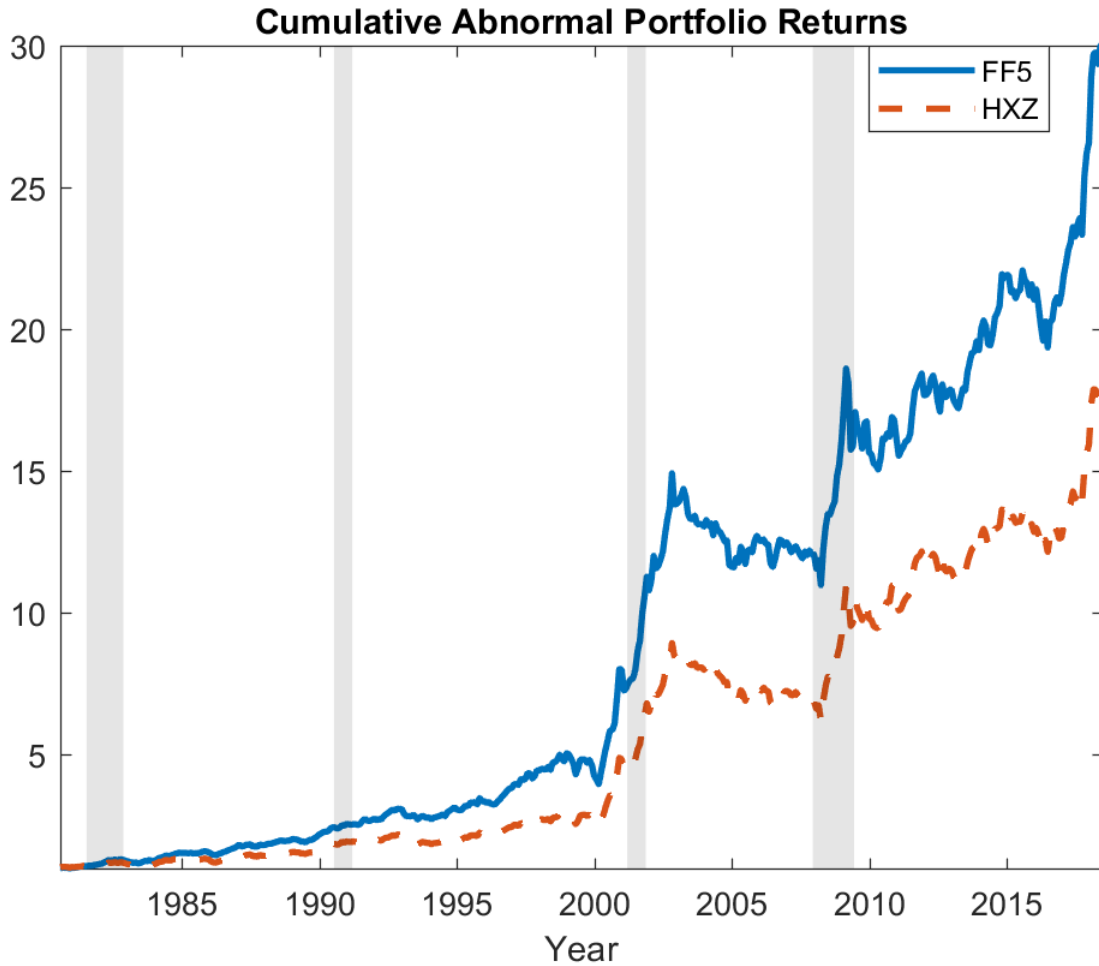


Table 1: Summary Statistics, Correlations, and Firm Characteristics

This table presents summary statistics in Panel A and a correlation matrix in Panel B for the firm-year-quarter sample. Cash-based operating profitability (COP/AT) is by removing non-cash components included in the computation of operating profitability and scaling by lagged total assets. The measure of accruals (ACR/AT) is defined as changes in non-cash working capital minus depreciation expense, in which non-cash working capital is equal to the change in non-cash current assets minus the change in current liability less short-term debt and tax payable, and then scaled by lagged total assets. ME is market capitalization deflated by CPI (measured in 2009 millions USD) at the end of March, June, September, and December. B/M is the ratio of book equity to market capitalization. Investment rate (I/K) is gross investment divided by lagged property, plant, and equipment. Return on equity (ROE) is defined as income before extraordinary items divided by the previous quarter book value of equity. In Panel A, I report the pooled mean, median, standard deviation (Std), 5<sup>th</sup> percentile (P5), 25<sup>th</sup> percentile (P25), 75<sup>th</sup> percentile (P75), and 95<sup>th</sup> percentile (P95). Observations denote the valid number of observations for each variable. The sample period is 1979 - 2017 at a quarterly frequency and excludes utility and financial industries. Panel C reports the time-series average of the cross-sectional medians of firm characteristics for five portfolios sorted on cash-based operating profitability. Duration 1 to 3 refer to [Weber \(2018\)](#), [Gonçalves \(2019\)](#), and [Gormsen and Lazarus \(2019\)](#), respectively. The sample period is from 1979 to 2017 at a quarterly frequency. The detailed definition of the variables refers to Section [A.1](#)

	COP/AT	ACR/AT	ME	B/M	I/K	ROE
<b>Panel A: Summary Statistics</b>						
Mean	0.02	0.00	2,431.15	0.72	0.16	0.00
Std	0.14	0.19	14,001.64	0.86	12.45	16.98
P5	-0.11	-0.14	6.70	0.10	-0.07	-0.26
P25	-0.01	-0.04	38.20	0.30	0.01	-0.02
Median	0.03	-0.01	166.09	0.53	0.03	0.02
P75	0.06	0.01	821.65	0.90	0.08	0.04
P95	0.13	0.11	8,257.59	1.90	0.31	0.10
Observations	469,741	356547	467,332	469,741	336,758	469,716
<b>Panel B: Correlation</b>						
COP/AT	1					
ACR/AT	-0.35	1				
ME	0.04	0.00	1			
B/M	-0.02	-0.04	-0.07	1		
I/K	0.00	0.00	0.00	0.00	1	
ROE	0.00	0.00	0.00	0.00	0.00	1
<b>Panel C: Firm Characteristics</b>						
	L	2	3	4	H	
COP/AT	-0.05	0.00	0.03	0.05	0.09	
ACR/AT	0.00	-0.01	-0.01	-0.02	-0.02	
Log ME	4.51	5.42	6.02	6.26	6.11	
B/M	0.55	0.67	0.63	0.55	0.44	
I/K	0.03	0.03	0.03	0.03	0.05	
ROE	-0.03	0.01	0.02	0.03	0.04	
Duration 1	20.79	20.55	19.25	18.50	18.55	
Duration 2	78.98	64.68	55.26	50.57	50.31	
Duration 3	-8.78	-10.87	-13.58	-16.92	-18.64	
Number	624	615	616	616	605	

Table 2: **Univariate Portfolio Sorting**

This table shows average excess returns for five portfolios sorted on cash-based operating profitability normalized by total assets (ATQ) in Panel A, by property, plant, and equipment (PPENTQ) in Panel B, by book equity (BE) in Panel C, and by sales (SALEQ) in Panel D by one-quarter relative to their industry peers, for which I use the [Fama and French \(1997\)](#) 30 industry classifications, and rebalance portfolios at the beginning of every January, April, July, and October. The sample starts from July 1980 to June 2018 and excludes financial and utility industries. I report average excess returns over the risk-free rate ( $E[R]-R_f$ ),  $t$ -statistics, standard deviations (Std), and Sharpe ratios (SR) across five portfolios in each panel. Portfolio returns are value-weighted by firms' market capitalization and are multiplied by 12 to make the magnitude comparable to annualized returns.  $t$ -statistics based on standard errors using the Newey-West correction are reported in parentheses.

	<b>L</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>H</b>	<b>H-L</b>
<b>Panel A: Total Assets</b>						
E[R]- $R_f$ (%)	5.50	7.12	7.70	7.83	10.05	4.55
[ $t$ ]	1.64	2.35	2.94	2.94	3.72	2.50
Std (%)	18.44	18.27	16.34	15.48	15.79	9.66
SR	0.30	0.39	0.47	0.51	0.64	0.47
<b>Panel B: Property, Plant, and Equipment</b>						
E[R]- $R_f$ (%)	5.25	8.30	7.71	8.38	9.54	4.29
[ $t$ ]	1.49	2.98	3.06	3.14	3.42	2.49
Std (%)	19.04	17.32	15.63	15.02	17.43	9.61
SR	0.28	0.48	0.49	0.56	0.55	0.45
<b>Panel C: Book Equity</b>						
E[R]- $R_f$ (%)	6.15	6.27	7.28	8.02	9.91	3.76
[ $t$ ]	1.85	2.10	2.69	3.01	3.72	2.20
Std (%)	18.44	18.16	16.28	15.89	15.58	9.35
SR	0.33	0.35	0.45	0.50	0.64	0.40
<b>Panel D: Sales</b>						
E[R]- $R_f$ (%)	5.84	7.77	7.63	8.50	9.15	3.31
[ $t$ ]	1.67	3.01	2.89	3.35	3.24	2.23
Std (%)	19.07	16.67	15.57	15.50	16.30	9.49
SR	0.31	0.47	0.49	0.55	0.56	0.35

Table 3: **Imperfect Information and Misallocation**

This table reports time-series averages of the cross-sectional medians of analysts coverage, analysts' forecast dispersion, and time-series averages of capital misallocations in five portfolios sorted on cash-based profitability, relative to their industry peers, for which I use the [Fama and French \(1997\)](#) 30 industry classifications, and rebalance portfolios at the beginning of every January, April, July, and October. Analysts coverage (Coverage) is calculated using the number of distinct analysts who made fiscal year one earnings forecast for the stock during the month  $[t-11, t]$  as used in [Hong, Lim, and Stein \(2000\)](#) and [Ali and Hirshleifer \(2019\)](#). If a firm is not covered by the IBES database, I assign its stock to zero analyst coverage. Analysts' forecast dispersion (Dispersion) is calculated as the standard deviation of earnings forecasts scaled by the absolute value of the mean earnings forecast. If the mean earnings forecast is zero, according to [Diether, Malloy, and Scherbina \(2002\)](#), then I assign the stock to the highest dispersion category. Misallocation 1 is calculated using the MPK measure in [Chen and Song \(2013\)](#), and Misallocation 2 is calculated using the MPK measure in [David, Schmid, and Zeke \(2018\)](#). I present each misallocation using SIC 2-digit and [Fama and French \(1997\)](#) 30 industry classifications. The sample period is from 1979 to 2017 at a quarterly frequency and excludes utility and financial industries from the analysis. The detailed definition of the measure of capital misallocation refers to the Internet Appendix [A.2](#).

	<b>L</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>H</b>
<b>Panel A: Analyst Coverage</b>					
Coverage	18.23	20.56	24.00	28.02	30.17
Dispersion	2.84	2.70	2.51	2.34	2.22
<b>Panel B: Capital Misallocation</b>					
Misallocation 1					
<i>SIC 2-digit</i>	1.14	0.96	0.90	0.81	0.76
<i>FF30</i>	1.19	0.98	0.92	0.83	0.78
Misallocation 2					
<i>SIC 2-digit</i>	1.31	1.01	0.94	0.88	0.81
<i>FF30</i>	1.33	1.03	0.97	0.92	0.88



Table 4: **Calibrated Parameters**

<b>The Description of Parameter</b>	<b>Symbol</b>	<b>Value</b>
<b>Technology</b>		
Mean-reversion parameter of firm-specific shock	$\rho_a$	0.10
Volatility of firm-specific shock	$\sigma_a$	0.01
Mean-reversion parameter of short-run productivity shock	$\rho_x$	0.30
Volatility of short-run productivity shock	$\sigma_x$	0.12
Mean-reversion parameter of long-run productivity shock	$\rho_\mu$	0.01
Volatility of long-run productivity shock	$\sigma_\mu$	0.25
Mean of long-run productivity shock	$\bar{\mu}$	0.25
<b>Production</b>		
Depreciation rate of capital	$\delta$	0.10
<b>Investment</b>		
Proportional adjustment cost for investment	$\chi$	16
<b>Stochastic Discount Factor (SDF)</b>		
Risk-free rate	$r_f$	0.03
Price of risk of short-run productivity shock	$\lambda_x$	0.15
Price of risk of long-run productivity shock	$\lambda_\mu$	0.40

Table 5: **Unconditional Aggregate Moments**

<b>Moments</b>	<b>Data</b>	<b>Benchmark</b>
$E[R_m]-R_f$ (%)	5.71	8.91
$\text{Std}[R_m]$ (%)	17.61	28.78
SR	0.32	0.31
I/K	0.04	0.04
Tobin's q	1.43	1.63

Table 6: **Portfolios, Firm Characteristics, and Model Comparison**

This table reports time-series averages of the cross-sectional averages of firm characteristics across five portfolios sorted on cash-based-operating profitability. Panel A is based on quintile portfolios as I present in Panel C of Table 1; Panel B reports quintile portfolios based on my model economy as the benchmark. The returns  $E[R]-R_f$  are annualized.

<b>Moments</b>	<b>L</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>H</b>	<b>H-L</b>
<b>Panel A: Data</b>						
$E[R]-R_f$ (%)	5.50	7.12	7.70	7.83	10.05	4.55
MPK	1.28	1.29	1.34	1.35	1.67	
I/K	0.03	0.03	0.03	0.03	0.05	
Tobin's q	1.29	1.48	1.34	1.51	1.81	
Duration	20.79	20.55	19.25	18.50	18.55	
<b>Panel B: Model</b>						
$E[R]-R_f$ (%)	5.44	8.01	9.05	9.40	9.68	4.24
MPK	1.31	1.32	1.40	1.45	1.57	
I/K	0.02	0.02	0.02	0.03	0.05	
Tobin's q	1.28	1.31	1.40	1.50	1.76	
Duration	21.94	20.96	20.84	20.68	19.81	

Table 7: **Investment and q Regressions**

This table reports the results of investment regression among quintile portfolios sorted on a firm’s cash-based profitability. All estimates are based on the following regression:

$$\frac{I_{i,n,t}}{K_{i,n,t-1}} = \beta_{0,t} + \beta_{0,i} + \beta_{q,n}q_{i,n,t} + Controls_{i,n,t} + \varepsilon_{i,n,t}, \text{ for firms in portfolio } n$$

in which  $Controls_{i,n,t}$  are control variables for firms’ fundamentals, including size, book-to-market ratio, investment rate, and cash-based profitability. All independent variables are normalized to a zero mean and a one standard deviation after winsorization at the 1st and 99th percentiles to reduce the impact of outliers.  $t$ -statistics based on standard errors that are clustered at the firm level are reported. The sample period is from 1979 to 2017 at quarterly frequency.

	<b>L</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>H</b>
Tobin’s q	0.01	0.05	0.06	0.06	0.07
[t]	0.53	5.20	4.84	6.27	3.17
Log ME	0.16	0.08	0.02	0.01	-0.02
[t]	12.17	9.65	1.34	0.95	-0.72
Log B/M	-0.00	-0.00	-0.01	-0.01	-0.00
[t]	-0.28	-0.35	-1.39	-1.66	-0.28
COP/AT	-0.14	-0.03	0.01	0.01	0.09
[t]	-14.28	-4.80	1.81	1.77	3.66
Observations	63,874	66,445	69,705	69,495	65,604
R-squared	0.58	0.67	0.70	0.71	0.81
Time FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes

Table 8: **Exposure to Aggregate Productivity Shocks**

This table reports the aggregate productivity exposures among quintile portfolios sorted on a firm’s cash-based profitability. All estimates are based on the following regression:

$$\Delta \ln A_{i,n,t} = \xi_{0i} + \xi_{A,n} \Delta \ln A_t + Controls_{i,n,t} + \varepsilon_{i,n,t}, \text{ for firms in portfolio } n$$

in which  $Controls_{i,n,t}$  are control variables for firms’ fundamentals, including size, book-to-market ratio, investment rate, and cash-based profitability. Panel A and B differ in that they use two alternative estimation methods in the first stage to estimate  $\Delta \ln A_{i,n,t}$ . Productivity in Panel A is based on the fixed effect procedure, whereas productivity in Panel B is based on the dynamic error component method of [Blundell and Bond \(2000\)](#). These estimation methods are described in the Internet Appendix [A.4](#), following [Ai, Croce, and Li \(2013\)](#) and [Li, Tsou, and Xu \(2019\)](#). All independent variables are normalized to a zero mean and a one standard deviation after winsorization at the 1st and 99th percentiles to reduce the impact of outliers.  $t$ -statistics based on standard errors that are clustered at the firm level are reported. The sample period is from 1979 to 2017.

	<b>L</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>H</b>	<b>L</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>H</b>
	<b>Firm-level TFP 1</b>					<b>Firm-level TFP 2</b>				
$\Delta \ln A_t$	0.43	0.36	0.35	0.42	0.78	0.35	0.38	0.39	0.38	0.88
[t]	4.51	3.81	2.80	2.23	2.21	4.99	5.49	4.43	2.37	2.52
Log ME	-2.03	-1.52	-2.43	-2.35	-0.27	-1.80	-1.03	-1.59	-1.10	0.75
[t]	-5.26	-3.59	-4.66	-3.01	-0.25	-6.05	-3.32	-3.93	-1.65	0.73
Log BM	-2.21	-2.26	-2.80	-3.46	-2.48	-2.00	-2.05	-2.36	-2.88	-2.50
[t]	-8.34	-7.88	-8.26	-6.75	-3.33	-9.32	-9.56	-9.11	-6.89	-3.44
I/K	-1.36	-1.09	-1.72	-1.23	-1.12	-0.37	-0.31	-0.60	-0.14	0.21
[t]	-7.27	-6.76	-7.09	-3.30	-1.40	-2.46	-2.43	-3.28	-0.44	0.26
CP/AT	1.27	0.97	1.30	1.91	0.74	1.26	0.73	0.87	1.67	-0.49
[t]	6.08	5.53	5.61	4.25	0.77	7.40	6.21	5.15	4.42	-0.57
Observations	25,739	24,641	24,494	22,199	18,714	25,866	24,875	24,838	22,696	19,098
R-squared	0.41	0.44	0.42	0.44	0.37	0.47	0.46	0.45	0.43	0.38
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 9: **Payout Sensitivities to Productivity Shocks**

This table shows payout sensitivities to short-run and long-run productivity shocks by cash-based operating profitability sorted quintile portfolios. All estimates are based on the following panel regression:

$$N_{i,t} = \beta_{0,i} + \beta_{sr}\varepsilon_t^x + \beta_{lr}\varepsilon_t^\mu + \rho N_{i,t-1} + \beta_\mu \mu_{t-1} + Controls_{i,t-1} + resid_t,$$

in which  $N_{i,t}$  denotes firm  $i$ 's payout (income-to-sales) ratio,  $\varepsilon_t^x$  and  $\varepsilon_t^\mu$  denotes short- and long-run shocks, respectively. Controls variables for a firm's fundamentals include size and book-to-market ratio. I further control for the predetermined value of the long-run component,  $\mu_{t-1}$ , and the lagged payout ratio  $N_{i,t-1}$ . Standard errors are adjusted for heteroscedasticity and clustered at the firm level.

<b>Payout Sensitivities</b>					
	<b>L</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>H</b>
$\varepsilon_{x,t}$	-2.16	0.64	0.59	0.85	0.64
[t]	-1.11	2.89	10.22	6.70	0.94
$\varepsilon_{\mu,t}$	-2.74	0.27	0.23	0.22	1.77
[t]	4.62	9.80	7.86	4.97	5.67
$N_{i,t-1}$	0.55	0.30	0.43	0.47	0.45
[t]	132.43	116.67	134.81	134.28	104.25
$x_{t-1}$	-7.49	-0.79	0.68	1.02	4.04
[t]	-3.54	-3.38	11.06	7.67	5.56
lagged log ME	26.01	1.57	2.46	2.33	3.22
[t]	5.69	2.65	14.29	6.04	1.58
lagged B/M	8.11	1.21	0.89	0.85	4.23
[t]	2.92	3.87	10.35	4.52	4.00
Observations	33,719	40,568	43,080	44,598	44,267
Firm FE	Yes	Yes	Yes	Yes	Yes

Table 10: Estimating the Market Price of Risk

In Panel A, I present GMM estimates of the parameters of the stochastic discount factor  $SDF_t = 1 - \lambda_x \times \varepsilon_t^x - \lambda_\mu \times \varepsilon_t^\mu$ , using the quintile portfolios sorted on cash-based operating profitability.  $\varepsilon_a$  and  $\varepsilon_x$  denotes the short- and long-run productivity shock. I do the normalization such that  $E[m] = 1$  (See, e.g., Cochrane (2005)).  $R_{MKT}$ ,  $\Delta \ln A_t$ ,  $\varepsilon_t^x$ , and  $\varepsilon_t^\mu$  are normalized to zero mean and unit standard deviation. I report  $t$ -statistics and computed errors using the Newey-West procedure adjusted for three lags. As a measure of fit, I report the sum of squared errors (SSQE), mean absolute pricing errors (MAPE), and the  $J$ -statistic of the overidentifying restrictions of the model. In Panel B, I present testing portfolios' risk exposures ( $\beta_x$  and  $\beta_\mu$ ) to the short- and long-run productivity shock, respectively.

<b>Panel A: Price of Risk</b>			
	<b>CAPM Model</b>	<b>TFP Model</b>	<b>Two-factor Model</b>
$\lambda_x$	0.28	0.79	0.57
[t]	2.97	1.88	3.41
$\lambda_\mu$			0.81
[t]			2.80
SSQE (%)	6.02	6.03	1.13
MAPE (%)	4.12	4.25	1.96
J-Test	9.49	7.56	2.47
p	0.09	0.18	0.65

<b>Panel B: Portfolio Risk Exposures</b>						
	<b>L</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>H</b>	<b>H-L</b>
$\beta_x$	1.01	0.67	0.59	1.09	1.09	0.11
[t]	1.16	0.86	0.83	1.30	1.21	0.24
$\beta_\mu$	2.51	3.01	3.36	3.84	4.08	1.77
[t]	2.12	2.55	2.82	2.80	2.64	2.65

Table 11: **Asset Pricing Factor Tests**

This table shows asset pricing factor tests for five portfolios sorted on cash-based operating profitability relative to their industry peers, for which I use the [Fama and French \(1997\)](#) 30 industry classifications and rebalance portfolios at the beginning of every January, April, July, and October. The results reflect monthly data, for which the sample starts from July 1980 to June 2018 and excludes financial and utility industries. To adjust for risk exposure, I perform time-series regressions of cash-based-profitability-sorted portfolios' excess returns on the market factor (MKT) as the CAPM model in Panel A, on the [Fama and French \(1996\)](#) three factors (MKT, the size factor-SMB, and the value factor-HML) in Panel B, on the [Fama and French \(1996\)](#) three factors plus [Carhart \(1997\)](#) factor (MKT, the size factor-SMB, the value factor-HML, and the momentum factor-UMD), on the [Fama and French \(2015\)](#) five factors (MKT, the size factor-SMB, the value factor-HML, the profitability factor-RMW, and the investment factor-CMA) in Panel D, and on the [Hou, Xue, and Zhang \(2015\)](#) q-factors (MKT, SMB, the investment factor-I/A, and the profitability factor-ROE) in Panel E, respectively. Data on the Fama-French five-factor and Carhart factors are from Kenneth French's website. Data on the I/A and ROE factor are provided by Kewei Hou, Chen Xue, and Lu Zhang. These betas, together with alphas, are annualized by multiplying 12. Standard errors are estimated by using the Newey-West correction, and the corresponding  $t$ -statistics are reported in parentheses.

	<b>L</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>H</b>	<b>H-L</b>		<b>L</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>H</b>	<b>H-L</b>
<b>Panel A: CAPM</b>													
$\alpha_{CAPM}$	-3.51	-2.12	-0.70	-0.15	2.04	5.55							
[t]	-2.74	-2.31	-1.05	-0.20	2.24	3.20							
MKT	1.12	1.15	1.04	0.99	1.00	-0.12							
[t]	37.36	41.36	53.66	50.33	45.12	-3.12							
<b>Panel B: FF3</b>													
$\alpha_{FF3}$	-4.31	-3.14	-0.99	0.53	3.36	7.67	$\alpha_{FF4}$	-3.52	-2.43	-0.52	0.85	3.20	6.73
[t]	-3.58	-3.02	-1.44	0.76	4.14	4.51	[t]	-2.96	-2.39	-0.80	1.19	3.91	4.04
MKT	1.08	1.13	1.03	0.98	0.96	-0.12	MKT	1.06	1.12	1.02	0.97	0.96	-0.10
[t]	43.65	38.22	54.58	61.44	59.52	-3.97	[t]	43.41	46.05	51.20	70.01	57.62	-3.64
SMB	0.32	0.25	0.12	-0.07	-0.07	-0.39	SMB	0.33	0.26	0.12	-0.07	-0.07	-0.40
[t]	7.12	6.40	4.66	-2.63	-2.71	-6.84	[t]	7.25	6.24	4.60	-2.66	-2.72	-6.83
HML	-0.06	0.04	-0.02	-0.09	-0.23	-0.17	HML	-0.09	0.01	-0.04	-0.11	-0.22	-0.13
[t]	-1.34	1.11	-0.86	-4.14	-6.84	-2.49	[t]	-1.96	0.29	-1.81	-4.53	-6.86	-1.89
							UMD	-0.08	-0.07	-0.05	-0.03	0.02	0.10
							[t]	-2.93	-3.41	-3.45	-1.15	0.92	2.99
<b>Panel D: FF5</b>							<b>Panel E: HXZ</b>						
$\alpha_{FF5}$	-2.90	-2.23	-0.97	0.46	2.20	5.10	$\alpha_{HXZ}$	-1.89	-1.08	-0.29	1.12	2.59	4.48
[t]	-2.66	-2.10	-1.40	0.69	2.68	3.40	[t]	-1.65	-1.04	-0.37	1.48	2.54	2.76
MKT	1.06	1.12	1.03	0.98	0.99	-0.07	MKT	1.05	1.12	1.03	0.98	0.99	-0.07
[t]	39.25	41.86	64.52	78.74	67.89	-2.46	[t]	39.60	53.33	51.32	62.56	41.44	-2.02
SMB	0.30	0.24	0.12	-0.07	-0.06	-0.36	SMB	0.25	0.19	0.09	-0.10	-0.08	-0.33
[t]	5.96	7.59	4.33	-2.92	-2.38	-5.77	[t]	5.35	8.20	3.32	-4.50	-2.53	-5.44
HML	-0.03	0.03	-0.04	-0.09	-0.29	-0.25	HML	-0.09	0.09	-0.02	-0.12	-0.22	-0.13
[t]	-0.68	0.64	-1.56	-2.06	-9.02	-3.67	[t]	-1.48	1.80	-0.43	-3.04	-4.08	-1.50
RMW	-0.12	-0.09	0.03	-0.01	0.15	0.26	ROE	-0.15	-0.17	-0.01	-0.02	0.13	0.28
[t]	-1.83	-2.13	0.55	-0.34	3.97	2.86	[t]	-2.96	-5.54	-0.46	-0.52	3.89	4.10
CMA	0.00	0.07	0.06	-0.01	0.08	0.08							
[t]	0.02	1.22	1.19	-0.18	1.51	0.70							

Table 12: **Fama-MacBeth Regressions**

This table reports Fama-MacBeth regressions of individual stock excess returns on their cash-based operating profitability and other firm characteristics. I conduct cross-sectional regressions for each month. In each month, monthly returns of individual stock returns (annualized by multiplying them by 12) are regressed on cash-based profitability, different sets of control variables that are known for 6 months prior to portfolio formation, except the market capitalization (Size), and industry fixed effects. Control variables include the natural logarithm of market capitalization (Size), the natural logarithm of book-to-market ratio (B/M), investment rate (I/K), return on equity (ROE), and industry dummies based on [Fama and French \(1997\)](#) 30 industry classifications. All independent variables are normalized to a zero mean and one standard deviation after winsorization at the 1st and 99th percentiles to reduce the impact of outliers. t-statistics based on standard errors estimated using the Newey-West correction are reported. The sample period is July 1980 to June 2018.

	(1)	(2)	(3)	(4)	(5)	(6)
COP/AT	3.00	3.98	2.96	2.55	2.74	2.72
[t]	4.05	7.33	4.25	5.81	4.52	7.92
Log ME		-4.38				-3.43
[t]		-3.73				-3.45
Log B/M			5.51			3.32
[t]			6.93			4.43
ROE				1.65		3.10
[t]				1.60		3.43
I/K					-3.56	-2.63
[t]					-5.49	-5.56
Observations	1,404,098	1,396,461	1,404,082	1,404,023	1,007,172	1,001,909
R-squared	0.04	0.04	0.04	0.04	0.05	0.06
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes



Table 13: **Double Sorting - Size**

This table reports average excess stock returns of ten portfolios independently sorted on five portfolios based on cash-based operating profitability and two portfolios based on size, all relative to their industry peers based on the [Fama and French \(1997\)](#) 30 industry classifications. At the beginning of January, April, July, and October, I assign firms into big (B) and small (S) groups based on their market capitalization relative to industry peers and group firms into quintile portfolios (from low to high) based on their cash-based operating profitability relative to industry peers. I then track the value-weighted returns on each portfolio for a quarter. I report the annualized portfolio returns by multiplying them by 12 (and report  $t$ -statistics in parentheses).

	<b>L</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>H</b>	<b>H-L</b>
<b>S</b>	2.01	8.50	10.95	13.29	11.73	9.72
[t]	0.58	2.87	3.88	5.16	4.94	4.88
<b>B</b>	5.95	7.12	7.66	7.83	9.93	3.98
[t]	2.10	2.82	3.31	3.01	3.87	2.85

Table 14: **Double Sorting - Behavioral**

This table reports average excess stock returns of portfolios sorted on cash-based-operating profitability and the alternative measure of a behavioral explanation relative to their industry peers based on the [Fama and French \(1997\)](#) 30 industry classifications. In Panel A, I report average excess stock returns of two by five portfolios independently double sorted on cash-based operating profitability and institutional ownership. I collect quarterly institutional ownership data at the end of each quarter from the Thomson Reuters Institutional Holdings (13F) database, which keeps track of the 13F filings of professional money managers. An institutional investment manager with investment discretion over 100 million or more is required to file Form 13F with the SEC within 45 days at the end of a calendar quarter on the number of shares they hold of stocks. A firm's institutional ownership is calculated as the ratio of the summation of its shares held by an institutional investors to total shares outstanding in a given quarter. The sorting on cash-based operating profitability is reported across columns L to H, and the sorting on the measure of institutional ownership is reported across row L and H. The column H-L stands for the high-minus-low portfolio (across columns) within portfolios sorted on institutional ownership. In Panel B, at the beginning of every January, April, July, and October, I first sort stocks into five portfolios based on firms' cash-based operating profitability. Then, firms in the high quintile portfolio are further sorted into two portfolios based on their cash-based operating profitability in the next quarter, and I report their portfolio returns denoted as HL and HH, respectively. I also report the average portfolio return in the low quintile portfolio sorted on cash-based operating profitability at the beginning of every January, April, July, and October. Lastly, I calculate the return difference between HL and L portfolios (HL-L), and the return difference between HH and L portfolios (HH-L). In the left Panel C, I report average excess stock returns of three by five portfolios dependently double sorted on changes in the fraction of firm shares outstanding owned by retail investors and then on cash-based operating profitability on the report date of quarterly earnings. The sorting on cash-based operating profitability is reported across columns L to H, and the sorting on changes in the fraction of firm shares outstanding owned by retail investors is reported across row L, 2, and H. The column H-L stands for the high-minus-low portfolio (across columns) within each portfolio sorted by the fraction of firm shares outstanding owned by retail investors. In the right Panel C, I report the time-series average of the cross-sectional mean and median of changes in the fraction of firm shares outstanding owned by retail investors across tercile portfolios. All portfolio returns correspond to value-weighted returns by firm market capitalization. I report the annualized portfolio returns by multiplying them by 12 (and  $t$ -statistics in parentheses). The sample period is July 1980 to June 2018.

<b>Panel A: Institutional Ownership</b>						
	<b>L</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>H</b>	<b>H-L</b>
L	2.76	7.01	7.13	6.19	9.72	6.96
[t]	0.60	2.15	2.21	1.58	3.02	2.09
H	7.58	8.21	7.75	9.49	10.65	3.07
[t]	2.11	2.31	2.42	2.81	3.03	2.32

<b>Panel B: Underreaction</b>					
	<b>HL</b>	<b>HH</b>	<b>L</b>	<b>HL-L</b>	<b>HH-L</b>
E[R]-R <sub>f</sub> (%)	9.69	10.51	5.50	4.18	5.01
[t]	3.82	3.43	1.64	2.35	2.51

<b>Panel C: Overreaction</b>								
	<b>L</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>H</b>	<b>H-L</b>	<b>Mean</b>	<b>Median</b>
L	8.26	8.21	10.81	9.31	10.25	1.99	-2.35	-2.31
[t]	3.04	2.69	4.12	2.70	2.39	0.77		
2	6.38	9.62	6.44	9.00	10.45	4.08	0.05	0.04
[t]	2.16	2.92	2.05	2.60	3.35	2.96		
H	5.43	10.13	10.81	8.60	9.11	3.68	2.39	2.64
[t]	1.43	2.82	3.21	2.63	2.15	1.57		

Table 15: **Double Sorting - Financial Distress**

This table reports average excess stock returns of two by five portfolios independently sorted on cash-based operating profitability and the measure of financial distress, referring to the O index in Panel A, Z index in Panel B, distance to default in Panel C, and failure probability in Panel D, all relative to their industry peers based on the [Fama and French \(1997\)](#) 30 industry classifications. The detailed information of the O index, Z index, default probability, and failure probability refers to [Ohlson \(1980\)](#), [Altman \(1968\)](#), [Bharath and Shumway \(2008\)](#), and [Campbell, Hilscher, and Szilagyi \(2008\)](#), respectively. At the beginning of every January, April, July, and October, I assign firms into low (L) and high (H) groups based on their financial distress measures. I then track the value-weighted returns on each portfolio over the next three months. I report the annualized portfolio returns by multiplying them by 12 (and report t-statistics in parentheses).

	<b>L</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>H</b>	<b>H-L</b>	<b>L</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>H</b>	<b>H-L</b>
	<b>Panel A: O Index</b>						<b>Panel B: Z Index</b>					
<b>L</b>	5.56	7.48	7.42	7.63	9.99	4.43	7.06	9.27	8.8	7.89	9.74	2.67
[t]	1.55	2.31	2.55	2.12	2.87	2.04	1.92	2.91	3.19	2.99	3.23	2.11
<b>H</b>	6.25	6.67	8.61	11.38	9.81	3.56	5.61	4.52	6.63	8.08	10.22	4.61
[t]	1.48	1.64	2.70	4.28	3.54	2.10	1.50	1.38	2.20	1.99	2.91	2.12
	<b>Panel C: Distance to Default</b>						<b>Panel D: Failure Probability</b>					
<b>L</b>	7.04	6.23	7.72	7.68	9.99	2.95	7.41	6.48	7.14	7.88	9.85	2.45
[t]	2.51	2.57	3.39	2.88	3.84	2.21	2.85	2.74	3.09	2.84	3.63	1.88
<b>H</b>	1.85	8.65	7.16	10.71	10.91	9.06	2.50	7.47	8.96	6.76	8.39	5.89
[t]	0.59	2.68	2.64	4.63	4.79	4.91	0.74	2.43	3.54	3.01	3.41	3.74

Table 16: **Double Sorting - Other Risks**

This table reports average excess stock returns of independent two-way sorting by assigning all sample firms' values in a proxy for an alternative explanation, including technology obsolescence (capital age and investment rate) in Panel A, book-to-market ratio, and return on equities in Panel B, financial constraints (SA and WW index) in Panel C, the exposure of economic uncertainty (macroeconomic uncertainty and political uncertainty index) in Panel D, and adjustment cost (asset redeployability and inflexibility) in Panel E, into two groups (low and high) and by assigning them by their cash-based profitability into quintile portfolios (low, 2, 3, 4, and high). Asset redeployability is constructed in a three-step procedure. First, I compute the asset-level redeployability as the proportion of industries that use a given asset. Second, the industry-level redeployability index is the value-weighted average of each asset's redeployability. Finally, I obtain the firm-level measure of asset redeployability by the value-weighted average of industry-level redeployability indices across business segments in which the firm operates. The asset redeployability data is available from Howard Kung's website. In addition, within each portfolio sorted by the proxy for a risk-based explanation, I also form a high-minus-low portfolio (H-L) that takes a long position in the high-cash-based-profitability portfolio and a short position in the low-cash-based-profitability portfolio. As a result, I form total of 12 portfolios for each risk-based explanation. I then report the monthly averages of value-weighted returns on all 12 portfolios. I annualize the portfolio returns by multiplying them by 12 and include t-statistics in parentheses.

	<b>L</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>H</b>	<b>H-L</b>	<b>L</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>H</b>	<b>H-L</b>
<b>Panel A: Technology Obsolete</b>												
	<b>Capital Age</b>						<b>I/K</b>					
L	3.49	7.84	6.85	7.83	10.72	7.22	7.17	7.27	7.66	7.95	10.52	3.35
[t]	1.00	2.99	2.64	2.90	3.75	3.20	2.61	2.70	3.07	3.02	3.90	2.37
H	6.70	7.03	8.21	7.48	9.69	3.00	2.95	6.90	7.74	7.62	8.76	5.81
[t]	2.44	2.68	3.76	2.97	3.91	2.40	0.85	2.87	3.48	2.84	3.60	3.18
<b>Panel B: Value and Profitability</b>												
	<b>B/M</b>						<b>ROE</b>					
L	4.90	5.26	7.48	7.65	9.86	4.96	1.99	6.70	8.31	8.90	11.08	9.09
[t]	1.70	2.00	3.58	2.89	3.77	2.45	0.52	1.91	2.75	2.99	3.43	2.39
H	6.00	9.30	8.08	7.43	12.50	6.50	7.55	7.06	7.52	7.92	9.95	2.40
[t]	1.73	3.78	2.82	2.76	5.26	2.24	3.12	3.37	3.55	3.05	3.85	2.03
<b>Panel C: Financial Constraint</b>												
	<b>SA</b>						<b>WW</b>					
L	6.44	7.22	7.67	7.75	9.98	3.55	6.14	7.25	7.67	7.79	9.88	3.74
[t]	2.29	2.86	3.32	3.00	3.81	2.41	2.14	2.83	3.33	3.03	3.77	2.47
H	0.25	5.04	6.59	10.91	9.06	8.82	-0.30	5.01	9.18	10.76	10.95	11.25
[t]	0.07	1.59	2.37	4.34	3.78	3.73	-0.09	1.71	3.50	4.29	5.04	4.82
<b>Panel D: Uncertainty</b>												
	<b>Macroeconomic Uncertainty</b>						<b>Political Uncertainty</b>					
L	4.24	5.71	7.46	7.46	9.91	5.67	5.90	6.55	7.15	7.48	9.36	3.46
[t]	1.32	2.00	2.93	3.20	4.11	2.98	1.86	2.16	2.75	2.54	3.23	1.98
H	7.78	8.44	8.10	8.45	10.82	3.04	5.74	8.46	8.34	8.04	11.72	5.98
[t]	2.77	2.80	3.10	2.36	3.14	1.99	1.85	3.37	3.31	2.78	3.89	2.13
<b>Panel E: Adjustment Cost</b>												
	<b>Redeployability</b>						<b>Inflexibility</b>					
L	3.40	7.79	8.14	8.17	9.76	6.36	2.87	7.33	8.26	7.31	12.30	9.43
[t]	0.84	2.51	3.12	2.62	3.40	2.44	0.83	2.59	3.33	2.52	4.56	4.91
H	7.11	7.97	8.45	8.03	10.39	3.28	6.92	7.65	7.46	8.11	9.17	2.25
[t]	2.2	2.59	2.86	2.65	3.46	3.04	2.65	3.22	3.48	3.46	3.46	2.01

# Internet Appendix for “Learning and the Anatomy of the Profitability Premium”\*

Chi-Yang Tsou

November 1, 2020

## Contents

<b>A Supplementary Analyses</b>	<b>IA. 1</b>
A.1 Data Sources . . . . .	IA. 1
A.2 Capital Misallocation . . . . .	IA. 2
A.3 Cash flow duration . . . . .	IA. 3
A.4 Firm-level Productivity Estimation Details . . . . .	IA. 5
A.5 Summary Statistics across Industries . . . . .	IA. 7
A.6 Transition Matrix . . . . .	IA. 7
A.7 Investment and Tobin’s q: Cumulant Estimator . . . . .	IA. 7
<b>B Mathematical Details of the Model</b>	<b>IA. 11</b>
B.1 Proof of Proposition 1 . . . . .	IA. 11
B.2 Proof of Corollary 1 . . . . .	IA. 13
B.3 Proof of Lemma 1 . . . . .	IA. 13
B.4 Proof of Lemma 2 . . . . .	IA. 16
B.5 Proof of Proposition 2 . . . . .	IA. 16

---

\*Citation format: Chi-Yang Tsou (2020) Internet Appendix to “Learning and the Anatomy of the Profitability Premium.” Any queries can be directed to the author of the article.

B.6 Proof of Proposition 3 . . . . .	IA. 19
B.7 Proof of Corollary 2 . . . . .	IA. 19
B.8 Proof of Proposition 4 . . . . .	IA. 21
B.9 Proof of Proposition 5 . . . . .	IA. 21

# A Supplementary Analyses

## A.1 Data Sources

My sample consists of firms that lie in the intersection of quarterly Compustat and CRSP (Center for Research in Security Prices). I obtain accounting data from Compustat and stock returns data from CRSP. My sample firms include those with domestic common shares (SHRCD = 10 and 11) trading on NYSE, AMEX, and NASDAQ, excluding utility firms with SIC 4-digit (Standard Industrial Classification) codes between 4900 and 4949 and finance firms with SIC codes between 6000 and 6999 (finance, insurance, trusts, and real estate sectors). Delisting returns are taken from CRSP; if a delisting return is missing and the delisting is performance-related, I impute a return of -30%, according to [Shumway \(1997\)](#). According to [Fama and French \(1993\)](#), I further drop closed-end funds, trusts, American depository receipts, real estate investment trusts, and units of beneficial interest. To mitigate backfilling bias, firms in my sample must be listed on Compustat for two years before including them in my sample. I match the firms on CRSP against quarterly Compustat, and lag quarterly accounting information by six months. For example, if a firm's first fiscal quarter ends in March, I assume that this information is public by the end of the following September. Macroeconomic data refers to the Bureau of Economic Analysis (BEA) maintained by the United States Department of Commerce, except the total factor productivity that is obtained from the U.S. Bureau of Labor Statistics (BLS).

I calculate a firm's operating profitability by following the computations in [Ball, Gerakos, Linnainmaa, and Nikolaev \(2015\)](#): sales (SALEQ) minus cost of goods sold (COGSQ) minus sales general, and administrative expenses (XSGAQ) (excluding research and development expenditures (XRDQ)).<sup>1</sup> Such a measure captures the performance of the firm's operations and is not affected by non-operating items, such as leverage and taxes. To evaluate the ability of the cash portion of operating profitability to predict stock returns, I remove the accrual components included in the computation of operating profitability to create the cash-based operating profitability measure. These components include the changes in accounts receivable (RECTQ), inventory (INVTQ), prepaid expenses (XPP), deferred revenue (DRCQ plus DRLTQ), accounts payable (APQ). Accruals are calculated in accordance with [Wu, Zhang, and Zhang \(2010\)](#).

I measure a firm's capital and investment following [Lin, Palazzo, and Yang \(2019\)](#) and [Li, Tsou, and Xu \(2019\)](#). I first denote an initial measure of firm-level capital stock ( $K_{i,0}$ ) for firm

---

<sup>1</sup>Variables constructed in my sample are based on quarterly Compustat, while those in [Ball, Gerakos, Linnainmaa, and Nikolaev \(2015\)](#) and [Ball, Gerakos, Linnainmaa, and Nikolaev \(2016\)](#) are based on annual Compustat.

$i$  using net property plant and equipment (PPENTQ) as the initial measure of firm  $i$ 's capital. After obtaining the initial capital by calculating the ratio of accumulated depreciation and amortization (DPACTQ) over current depreciation and amortization (DPQ), I recursively construct a measure of firm-level capital stock as follows:

$$K_{i,t+1} = K_{i,t} + I_{i,t}^N, \quad (\text{A.1})$$

in which  $I_{i,t}^N$  denotes firm  $i$  net investment between period  $t$  and  $t + 1$ . Net investment is defined as the difference in net property plant and equipment between two consecutive quarters. I denote firm  $i$ 's gross investment as

$$I_{i,t}^N = \delta_j K_{i,t} + I_{i,t}^G, \quad (\text{A.2})$$

in which  $\delta_j$  refers to the depreciation rate of industry  $j$  calculated by using industry-level depreciation data from BEA. All the quantities are expressed in 2009 dollars using the seasonally adjusted implicit price deflator for non-residential fixed investment.

Market capitalization is calculated by using data from CRSP and it is equal to the number of shares outstanding (SHROUT) multiplied by the share price (PRC). When size is reported to levels, I express it in 2009 dollars using the seasonally adjusted implicit price deflator for non-residential fixed investment. Quarter book value of equity is constructed following [Hou, Xue, and Zhang \(2015\)](#), and it is equal to shareholder's equity (SEQQ) plus deferred taxes and investment tax credit (TXDITCQ, if available) minus the book value of preferred stock (PSTKRQ). If shareholder's equity is not available, I use common equity (CEQQ) plus carrying value of the preferred stock (PSTKQ). If common equity is not available, I measure shareholder's equity as the difference between total assets (ATQ) and total liabilities (LTQ). The book-to-market ratio is the book value of equity divided by the market capitalization (PRCCQ times CSHOQ) at the end of the fiscal quarter. When I construct firm-level capital stock and gross investment observations, firm  $i$ 's investment rate is measured as gross investment  $\delta_j K_{i,t} + I_{i,t}^N$  divided by the beginning of the period of capital stock  $K_{i,t}$ . Return on equity (ROE) is defined as income before extraordinary items (IBQ) divided by the previous quarter book value of equity.

## A.2 Capital Misallocation

Following [Chen and Song \(2013\)](#), I measure the marginal product of capital by the ratio of operating income before depreciation (OIBDPQ) to one-year-lag net property, plant, and equipment (PPENTQ). For robustness, I follow [David et al. \(2018\)](#) to construct an alternative



measure of marginal product of capital by replacing the operating income before depreciation (OIBDPQ) with sales (SALEQ).<sup>2</sup> All the quantities are expressed in 2009 constant dollars using the seasonally adjusted implicit price deflator for non-residential fixed investment. Following Hsieh and Klenow (2009), I extend from manufacturing to all sectors, except utility, financial, high R&D industries, and compute the cross-sectional standard deviation as the dispersion measure within narrowly defined industries, as classified by the 2-digit SIC (Standard Industry Classification) industries, or broadly defined industries, as classified by the Fama-French 30 industries. Specifically, for firm  $i$  in industry  $j$ , I compute:

$$\log \left( \frac{MPK_{i,j}}{MPK_j} \right),$$

in which  $MPK_j$  is the cross-sectional average of  $MPK$  measured at the industry-level. I construct the measure of misallocation as follows. First, I compute the standard deviation of  $\log \left( \frac{MPK_{i,j}}{MPK_j} \right)$  at the industry-level within each portfolio sorted on capital age, in which the number of observations within each narrowly or broadly defined industry must be larger than 10 to avoid biased standard deviations driven by a few extreme values.<sup>3</sup> Second, I take the cross-sectional average of standard deviations across industries within each portfolio. Finally, as shown in Table 3, I report time-series averages of the cross-sectional dispersions of MPK in five portfolios sorted on cash-based operating profitability.

### A.3 Cash flow duration

I construct firm  $i$ 's cash flow duration to reflect the timing of cash flows, according to the model proposed by Lettau and Wachter (2007). Duration is the equity implied cash flow duration. Dechow, Sloan, and Soliman (2004) proposes the measure of cash flow duration and documents a negative relationship between cash flow durations and stock returns; in addition, Weber (2018) recently studies asset pricing implications, including exposure to existing risk factors, time variations in the slope, and the effect of short-sale constraints.

Duration resembles the traditional Maculay duration for bonds, which reflects the weighted average time to maturity of cash flows. The ratio of discounted cash flows to price determines

---

<sup>2</sup>Using sales (SALEQ) to proxy a firm's output alleviates any missing data concerns, given that the coverage of sales (SALEQ) is higher than that of operating income before depreciation (OIBDPQ) in Compustat.

<sup>3</sup>The industry coverage attrition issue is more salient for narrowly defined industries. Specifically, after I impose this restriction, the valid number of industries according to SIC 2-digit industry classifications drops from 75 to 42 for both  $MPK$  measures; in contrast, the valid number of industries according to Fama-French 30 industry classifications drops from 28 to 24 for the  $MPK$  measure following Chen and Song (2013) and also drops to 26 for the  $MPK$  measure following David, Schmid, and Zeke (2018).

the weights:

$$Dur_{i,t} = \frac{\sum_{s=1}^T s \times CF_{i,t+s}/(1+r)^s}{P_{i,t}}, \quad (\text{A.3})$$

in which I denote  $Dur_{i,t}$  as firm  $i$ 's duration at the end of fiscal year  $t$ ,  $CF_{i,t+s}$  is the cash flow at time  $t+s$ ,  $P_{i,t}$  is the current stock price, and  $r$  is the expected return on equity. Following [Dechow, Sloan, and Soliman \(2004\)](#), I assume the constant expected return on equity across both stocks and time. Relaxing such an assumption for firm-specific discount rates generates larger cross-sectional variations in the duration measure since firms with high cash flow duration tend to be growth firms with lower stock returns. For simplicity's sake, I focus on the measure of cash flow duration by imposing the constant expected return.<sup>4</sup>

In contrast to fixed income securities, such as bonds, stocks cannot have a well-defined finite maturity,  $t+T$ , and cash flows are not known in advance. Therefore, I split the duration formula into a finite detailed forecasting period and an infinite terminal value. I also assume that the later component is paid out as the level perpetuity for simplicity. Such the assumption allows us to rewrite the equation [\(A.3\)](#) as follows:

$$Dur_{i,t} = \frac{\sum_{s=1}^T s \times CF_{i,t+s}/(1+r)^s}{P_{i,t}} + \left( T + \frac{1+r}{r} \right) \times \frac{P_{i,t} - \sum_{s=1}^T CF_{i,t+s}/(1+r)^s}{P_{i,t}}. \quad (\text{A.4})$$

Moreover, I impose a clean surplus accounting, based on an accounting identify, and forecast cash flows via forecasting return on equity ( $ROE$ ),  $E_{i,t+s}/BV_{i,t+s-1}$ , and growth in book equity,  $(B_{i,t+s} - B_{i,t+s-1})/BV_{i,t+s-1}$ :

$$\begin{aligned} CF_{i,t+s} &= E_{i,t+s} - (B_{i,t+s} - B_{i,t+s-1}) \\ &= B_{i,t+s-1} \times \left[ \frac{E_{i,t+s}}{B_{i,t+s-1}} - \frac{B_{i,t+s} - B_{i,t+s-1}}{B_{i,t+s-1}} \right]. \end{aligned} \quad (\text{A.5})$$

Following [Dechow, Sloan, and Soliman \(2004\)](#), I model returns on equity and growth in equity as autoregressive process based on recent findings in financial accounting literature. In [Weber \(2018\)](#), the author estimates autoregressive parameters by using the merged CRSP-Compustat universe and assumes the mean reversion of  $ROE$  to the average cost of equity. I also follow the estimated autoregressive parameters in [Weber \(2018\)](#) by assuming that the growth in book equity is mean reverting to the long-run average growth rate in the economy with a coefficient of mean reversion equal to the average historical mean reversion in sales growth. The persistence of AR(1) model is 0.41 for  $ROE$  and 0.24 for  $BV$ , respectively. I assign the discount rate  $r$  to a value 0.12, which is equal to a steady-state average cost of

---

<sup>4</sup>According to [Weber \(2018\)](#), the variation over time in the return on equity  $r$  does not alter the cross-sectional ranking, which alleviate the concern for the cross-sectional implications.

equity of 0.12. Finally, I assign the average long-run nominal growth rate to a value 0.06, and use a detailed forecasting period of 15 years.

## A.4 Firm-level Productivity Estimation Details

**Firm-level productivity estimation** Data and firm-level productivity estimation are constructed as follows. I consider publicly traded companies on U.S stock exchanges listed in both the annual Compustat and the CRSP (Center for Research in Security Prices) databases for the period 1950-2017. In what follows, I report the annual Compustat items in parentheses and defined industry at the level of two-digit SIC codes. The output, or value added, of firm  $i$  in industry  $j$  at time  $t$ ,  $y_{i,j,t}$ , is calculated as sales (SALE) minus the cost of goods sold (COGS) and is deflated by the aggregate gross domestic product (GDP) deflator from the U.S. National Income and Product Accounts (NIPA). I measure the capital stock of the firm,  $k_{i,j,t}$ , as the total book value of assets (AT) minus current assets (ACT). This allows us to exclude cash and other liquid assets that may not be appropriate components of physical capital. I use the number of employees in a firm (EMP) to proxy for its labor inputs,  $n_{i,j,t}$ , because data for total hours worked are not available.

I assume that the production function at the firm level is Cobb-Douglas and allow the parameters of the production function to be industry-specific:

$$y_{i,j,t} = A_{i,j,t} k_{i,j,t}^{\alpha_{1,j}} n_{i,j,t}^{\alpha_{2,j}},$$

where  $A_{i,j,t}$  is the firm-specific productivity level at time  $t$ . This is consistent with our original specification because the observed physical capital stock,  $k_{i,j,t}$ , corresponds to the mass of production units owned by the firm.

I estimate the industry-specific capital share,  $\alpha_{1,j}$ , and labor share,  $\alpha_{2,j}$ , using the dynamic error component model adopted in [Blundell and Bond \(2000\)](#) to correct for endogeneity. Given the industry-level estimates for  $\widehat{\alpha}_{1,j}$  and  $\widehat{\alpha}_{2,j}$ , the estimated log productivity of firm  $i$  is computed as follows:

$$\ln \widehat{A}_{i,j,t} = \ln y_{i,j,t} - \widehat{\alpha}_{1,j} \cdot \ln k_{i,j,t} - \widehat{\alpha}_{2,j} \cdot \ln n_{i,j,t}.$$

I allow for  $\widehat{\alpha}_{1,j} + \widehat{\alpha}_{2,j} \neq 1$ , but my results hold also when I impose constant returns to scale in the estimation, that is,  $\widehat{\alpha}_{1,j} + \widehat{\alpha}_{2,j} = 1$ .

I use the multi-factor productivity index for the private non-farm business sector from the BLS as the measure of aggregate productivity.

**Endogeneity and the dynamic error component model.** I follow [Blundell and Bond \(2000\)](#) and write the firm-level production function as follows:

$$\begin{aligned}\ln y_{i,t} &= z_i + w_t + \alpha_1 \ln k_{i,t} + \alpha_2 \ln n_{i,t} + v_{i,t} + u_{i,t} \\ v_{i,t} &= \rho v_{i,t-1} + e_{i,t},\end{aligned}\tag{A.6}$$

where  $z_i$ ,  $w_t$ , and  $v_{i,t}$  indicate a firm fixed effect, a time-specific intercept, and a possibly autoregressive productivity shock, respectively. The residuals from the regression are denoted by  $u_{i,t}$  and  $e_{i,t}$  and are assumed to be white noise processes. The model has the following dynamic representation:

$$\begin{aligned}\Delta \ln y_{i,j,t} &= \rho \Delta \ln y_{i,j,t-1} + \alpha_{1,j} \Delta \ln k_{i,j,t} - \rho \alpha_{1,j} \Delta \ln k_{i,j,t-1} + \alpha_{2,j} \Delta \ln n_{i,j,t} - \rho \alpha_{2,j} \Delta \ln n_{i,j,t-1} \\ &\quad + (\Delta w_t - \rho w_{t-1}) + \Delta \kappa_{i,t},\end{aligned}\tag{A.7}$$

where  $\kappa_{i,t} = e_{i,t} + u_{i,t} - \rho u_{i,t-1}$ . Let  $x_{i,j,t} = \{\ln(k_{i,j,t}), \ln(n_{i,j,t}), \ln(y_{i,j,t})\}$ . Assuming that  $E[x_{i,j,t-l} e_{i,t}] = E[x_{i,j,t-l} u_{i,t}] = 0$  for  $l > 0$  yields the following moment conditions:

$$\begin{aligned}E[x_{i,i,t-l} \Delta \kappa_{i,t}] &= 0 \text{ for } l \geq 3 \\ E[x_{i,j,t-l} \Delta \kappa_{i,t}] &= 0 \text{ for } l \geq 3.\end{aligned}\tag{A.8}$$

that are used to conduct a consistent GMM estimation of equation (A.7). Given the estimates  $\hat{\alpha}_{1,j}$  and  $\hat{\alpha}_{2,j}$ , log productivity of firm  $i$  is computed as

$$\ln \hat{a}_{i,j,t} = \ln y_{i,j,t} - \hat{\alpha}_{1,j} \ln k_{i,j,t} - \hat{\alpha}_{2,j} \ln n_{i,j,t},\tag{A.9}$$

where  $\hat{a}_{i,j,t}$  is the productivity for firm  $i$  in industry  $j$ .

**Endogeneity and fixed effects.** An alternative way to estimate the production function avoiding endogeneity issues is to work with the following regression:

$$\ln y_{i,j,t} = v_j + z_{i,j} + w_{j,t} + \alpha_{1,j} \ln k_{i,j,t} + \alpha_{2,j} \ln n_{i,j,t} + u_{i,j,t}.\tag{A.10}$$

The parameter  $v_j$ ,  $z_{i,j}$ , and  $w_{j,t}$  indicate an industry dummy, a firm fixed effect, and an industry-specific time dummy, respectively. The residual from the regression is denoted by  $u_{i,j,t}$ . Given our point estimate of  $\hat{\alpha}_{1,j}$  and  $\hat{\alpha}_{2,j}$ , I can use equation (A.9) to estimate  $\hat{a}_{i,j,t}$ . Given this estimation of firms' productivity, I obtain the alternative estimation of firms' productivity.

## A.5 Summary Statistics across Industries

In Table A1, I report the summary statistics of cash-based profitability among firms in each industry according to the Fama and French (1997) 30 (FF30) industry classifications. There are comparatively large cross-industry variations in cash-based profitability. Specifically, the standard deviation of cash-based profitability ranges from 0.07 for the Textiles industry to 0.32 for the Apparel industry. Therefore, to make sure our results are not driven by any particular industry, I control for industry effects in our further analyses.

[Place Table A1 about here]

## A.6 Transition Matrix

Whether firms' cash-based profitability is persistent or not is important for our analysis of the profitability-return relation. To examine such persistence, I present the transition across quintiles over time. I present this analysis in Table A2. The left panel of Panel A2 shows the transition of firms' cash-based profitability from quarter  $t$  to quarter  $t+1$ , while the right panel shows the transition of firms' cash-based profitability from quarter  $t$  to quarter  $t+5$ . For firms in the top or bottom quintile of the distribution of cash-based profitability, the probability of staying in the same quintile in the next quarter (five quarters later) is above 65% (44%). The persistent cash-based profitability is intuitive because firms cannot easily adjust their information precision about their productivity. More importantly, such persistence has important asset pricing implications: if there is any profitability-return relation, then it should be attributed to long-lasting fundamental issues rather than transitory effects such as overreaction (underreaction) or mispricing.

[Place Table A2 about here]

## A.7 Investment and Tobin's q: Cumulant Estimator

In our model, marginal q is identical to average q. However, empirical evidence documented in the literature suggests that average q is a poor proxy for marginal q due to several reasons. One of these reasons is measurement error in the firms's capital stock. As the result, I empirically account for measurement error following Erickson and Whited (2000, 2002, 2012). The panel regression results are robust when I implement the estimators in . The results are reported in Table A3. Notably, the correlation between investment and Tobin's q are increasing from the low to high portfolio sorted on cash-based profitability.

[Place Table A3 about here]

Table A1: Cash-based Profitability across Fama-French 30 Industries

This table reports summary statistics of the industry-year observations of non-missing cash-based profitability across industries, including the pooled mean (Mean), standard deviation (Std), 5th percentile (P5), 25th percentile (P25), median (P50), 75th percentile (P75), and 95th percentile (P95). Obs denotes the average number of firms with non-missing cash-based profitability in each industry. Industries are based on Fama-French 30 industry classifications (FF30), excluding financial and utility industries. The sample period is 1979-2017.

FF30	Name	Obs	Mean	Std	P5	P25	Median	P75	P95
1	Food	3,613	-0.03	0.11	-0.18	-0.04	0.00	0.02	0.06
2	Liquor	492	-0.01	0.07	-0.14	-0.03	0.00	0.02	0.06
3	Tobacco	148	-0.03	0.15	-0.28	-0.04	0.02	0.04	0.10
4	Recreation	3,444	-0.04	0.27	-0.21	-0.07	-0.01	0.02	0.07
5	Printing	2,183	-0.01	0.08	-0.16	-0.03	0.01	0.03	0.07
6	Consumer Goods	3,102	-0.03	0.11	-0.18	-0.04	-0.01	0.02	0.07
7	Apparel	2,102	-0.06	0.32	-0.24	-0.09	-0.03	0.01	0.08
8	Healthcare	15,278	-0.03	0.21	-0.19	-0.06	-0.01	0.02	0.07
9	Chemicals	3,047	-0.01	0.09	-0.15	-0.03	0.01	0.03	0.06
10	Textiles	1,061	-0.02	0.07	-0.14	-0.03	0.00	0.02	0.05
11	Construction	5,375	-0.03	0.12	-0.16	-0.04	-0.01	0.02	0.05
12	Steel	2,230	-0.01	0.08	-0.10	-0.02	0.00	0.02	0.05
13	Machinery	6,096	-0.03	0.16	-0.16	-0.04	0.00	0.02	0.05
14	Electrical	4,355	-0.03	0.19	-0.19	-0.05	-0.01	0.02	0.07
15	Automobiles	2,257	-0.02	0.11	-0.16	-0.03	0.00	0.02	0.06
16	Carry	1,091	-0.02	0.11	-0.17	-0.03	0.00	0.02	0.05
17	Mines	1,585	-0.04	0.18	-0.19	-0.07	-0.02	0.01	0.06
18	Coal	278	0.00	0.05	-0.09	-0.02	0.01	0.02	0.04
19	Oil	6,954	-0.02	0.26	-0.14	-0.03	0.00	0.03	0.06
21	Communication	3,617	-0.02	0.15	-0.16	-0.03	0.01	0.03	0.06
22	Service	19,595	-0.03	0.18	-0.19	-0.05	-0.01	0.03	0.08
23	Business Equipment	17,938	-0.02	0.12	-0.16	-0.04	0.00	0.03	0.08
24	Business Supplies	2,431	0.00	0.08	-0.11	-0.01	0.01	0.03	0.06
25	Transportation	3,242	0.00	0.12	-0.10	-0.01	0.01	0.03	0.06
26	Wholesale	7,028	-0.05	0.29	-0.24	-0.07	-0.02	0.01	0.05
27	Retail	8,456	-0.03	0.22	-0.16	-0.05	-0.01	0.02	0.07
28	Meals	3,462	0.00	0.11	-0.10	-0.01	0.02	0.03	0.06
30	Other	896	-0.07	0.84	-0.28	-0.08	-0.02	0.02	0.09

Table A2: **Transition Matrix: Persistence of Cash-based Profitability**

This table presents transition frequency (%) across cash-based profitability quintiles from quarter  $t$  to  $t+1$  in the left panel (column 1 to column 6) and from quarter  $t$  to  $t+5$  in the right panel (column 7 to column 12). The cash-based profitability quintiles in each quarter are sorted in the same way as Tables ?? and 11. The sample period is 1979-2017.

	<b>Panel A: One-Quarter</b>					<b>Panel B: Five-Quarter</b>				
	L(t+1)	2(t+1)	3(t+1)	4(t+1)	H(t+1)	L(t+5)	2(t+5)	3(t+5)	4(t+5)	H(t+5)
L(t)	65.52	19.28	6.81	4.55	4.84	42.25	21.07	13.77	10.87	12.04
2(t)	17.74	48.46	22.22	7.61	3.82	17.28	33.93	24.83	15.24	8.71
3(t)	6.63	21.81	44.99	21.13	5.43	8.93	25.25	31.68	23.31	10.84
4(t)	4.52	8.48	21.17	47.89	17.94	7.03	15.07	25.38	32.56	19.96
H(t)	4.39	4.70	6.87	19.48	64.56	7.50	9.73	13.69	24.98	44.10

Table A3: **Investment and q Regressions: Cumulant Estimator**

This table reports the results of investment regression using cumulant estimator among quintile portfolios sorted on a firm's cash-based profitability. All estimates are based on the following regression:

$$\frac{I_{i,n,t}}{K_{i,n,t-1}} = \beta_{0,t} + \beta_{0,i} + \beta_{q,n}q_{i,n,t} + Controls_{i,n,t} + \varepsilon_{i,n,t}, \text{ for firms in portfolio } n$$

in which  $Controls_{i,n,t}$  are control variables for firms' fundamentals, including size, book-to-market ratio, investment rate, and cash-based profitability. All independent variables are normalized to a zero mean and a one standard deviation after winsorization at the 1st and 99th percentiles to reduce the impact of outliers.  $t$ -statistics based on standard errors that are clustered at the firm level are reported. The sample period is from 1979 to 2017 at quarterly frequency.

	<b>L</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>H</b>
Tobin's q	-0.12	-0.19	0.01	0.03	0.80
[t]	-5.45	-2.36	0.34	1.86	2.41
Log ME	0.03	-0.01	-0.04	-0.05	-0.08
[t]	5.91	-2.27	-12.70	-17.6	-6.75
Log B/M	-0.15	-0.22	-0.08	-0.05	0.49
[t]	-8.53	-4.00	-4.24	-3.04	2.15
COP/AT	-0.21	-0.10	-0.00	0.03	0.08
[t]	-23.87	-9.23	-0.23	8.74	2.03
Observations	63,874	66,445	69,705	69,495	65,604
$\rho^2$	0.01	0.04	0.06	0.09	0.12
$\tau^2$	0.68	0.35	0.41	0.60	0.62
J-Test	117.58	18.10	82.28	145.56	5.98



## B Mathematical Details of the Model

### B.1 Proof of Proposition 1

This proposition is a simple application of the Kalman-Bucy filter. The observable variables are the short-run productivity dynamics in equation (4) and the signal process in equation (9). Given that  $\mu_t$  is unobservable, I specifically rewrite the dynamics of observable processes as

$$d\mathbf{S}_t = (\mathbf{A}_0 + \mathbf{A}_1\mu_t)dt + \mathbf{B}_1dZ_t^\mu + \mathbf{B}_2d\mathbf{Z}_t, \quad (\text{B.1})$$

where

$$\mathbf{A}_0 = \begin{pmatrix} -\rho_x \\ 0 \end{pmatrix}, \quad \mathbf{A}_1 = \begin{pmatrix} \rho_x \\ 0 \end{pmatrix}, \quad \mathbf{B}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \mathbf{B}_2 = \begin{pmatrix} \sigma_x & 0 \\ 0 & \frac{1}{\sqrt{\Phi_i}} \end{pmatrix}, \quad (\text{B.2})$$

and of the unobservable variable  $\mu_t$  as

$$d\mu_t = [a_0 + a_1\mu_t]dt + b_1dZ_t^\mu + \mathbf{b}_2d\mathbf{Z}_t, \quad (\text{B.3})$$

where

$$a_0 = \rho_\mu\bar{\mu}, \quad a_1 = -\rho_\mu, \quad b_1 = \sigma_\mu, \quad \mathbf{b}'_2 = \begin{pmatrix} 0 & 0 \end{pmatrix}. \quad (\text{B.4})$$

For the brevity of notation, I temporarily ignore the subscript  $i$ . The posterior mean and the posterior variance evolve according to

$$d\hat{\mu}(a_0 + a_1\hat{\mu})dt + [(b \circ B) + \zeta\mathbf{A}'_1](B \circ B)^{-1} \times [d\mathbf{S}_t - (\mathbf{A}_0 + \mathbf{A}_1\hat{\mu})dt], \quad (\text{B.5})$$

$$\frac{d\zeta}{dt} = 2a_1\zeta + (b \circ b) - [(b \circ B) + \zeta\mathbf{A}'_1](B \circ B)^{-1}[(b \circ B) + \zeta\mathbf{A}'_1]', \quad (\text{B.6})$$

where

$$(b \circ b) = b_1b'_1 + \mathbf{b}_2\mathbf{b}'_2 = \sigma_\mu^2, \quad (\text{B.7})$$

$$(B \circ B) = \mathbf{B}_1\mathbf{B}'_1 + \mathbf{B}_2\mathbf{B}'_2 = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \frac{\Phi+1}{\Phi} \end{pmatrix}, \quad (\text{B.8})$$

$$(b \circ B) = b_1\mathbf{B}'_1 + \mathbf{b}_2\mathbf{B}'_2 = \begin{pmatrix} 0 & \sigma_\mu \end{pmatrix}, \quad (\text{B.9})$$

$$[(b \circ B) + \zeta\mathbf{A}'_1](B \circ B)^{-1} = \begin{pmatrix} \frac{\lambda\zeta}{\sigma_x^2} & \frac{\sigma_\mu\Phi}{1+\Phi} \end{pmatrix}. \quad (\text{B.10})$$

Therefore, I obtain the posterior mean

$$\begin{aligned}
d\hat{\mu}_{it} &= \rho_\mu(\bar{\mu} - \hat{\mu}_{it})dt + \left( \frac{\lambda\zeta_{it}}{\sigma_x^2} \quad \frac{\sigma_\mu\Phi_i}{1+\Phi_i} \right) \begin{pmatrix} dx_t - \rho_x(\hat{\mu}_{it} - x_t)dt \\ \sqrt{\frac{1+\Phi_i}{\Phi_i}}d\hat{Z}_t^s \end{pmatrix} \\
&= \rho_\mu(\bar{\mu} - \hat{\mu}_{it})dt + \left( \frac{\lambda\zeta_{it}}{\sigma_x^2} \quad \frac{\sigma_\mu\Phi_i}{1+\Phi_i} \right) \begin{pmatrix} \sigma_x dZ_t^x + \rho_x(\mu_t - \hat{\mu}_{it})dt \\ \sqrt{\frac{1+\Phi_i}{\Phi_i}}d\hat{Z}_t^s \end{pmatrix} \\
&= \rho_\mu(\bar{\mu} - \hat{\mu}_{it})dt + \left( \frac{\rho_x\zeta_{it}}{\sigma_x} \right) d\hat{Z}_{it}^x + \sigma_\mu \sqrt{\frac{\Phi_i}{1+\Phi_i}} d\hat{Z}_t^s,
\end{aligned} \tag{B.11}$$

where

$$d\hat{Z}_t^x = dZ_t^x + \left( \frac{\rho_x}{\sigma_x} \right) (\mu_t - \hat{\mu}_{it})dt. \tag{B.12}$$

The signal process can be rewritten as

$$ds_{it} = dZ_t^\mu + \frac{1}{\sqrt{\Phi_i}}dZ_t^s = \sqrt{\frac{1+\Phi_i}{\Phi_i}}d\hat{Z}_t^s, \tag{B.13}$$

where  $\hat{Z}_t^s$  is a standard Brownian motion orthogonal to  $\hat{Z}_t^x$  and has heterogeneous loadings  $\frac{1+\Phi_i}{\Phi_i}$  across firms.

According to equation (B.6), the dynamics of the posterior variance is summarized by the Riccati differential equation as follows:

$$\frac{d\zeta_{it}}{dt} = \frac{\sigma_\mu^2}{1+\Phi_i} - 2\rho_\mu\zeta_{it} - \left( \frac{\rho_x\zeta_{it}}{\sigma_x} \right)^2. \tag{B.14}$$

The stationary solution to the learning problem is available when I impose the assumption that the posterior variance has reached a steady state as time elapsed. Such the assumption is common in the literature, according to [Scheinkman and Xiong \(2003\)](#) and [Andrei, Mann, and Moyen \(2019\)](#), and fits well in the infinite horizon problem in my model economy. The steady-state value for  $\zeta_{it}$  is denoted as  $\zeta_i$  by imposing the Riccati differential equation in equation (B.14) equal to zero:

$$\frac{\sigma_\mu^2}{1+\Phi_i} - 2\rho_\mu\zeta_i - \left( \frac{\rho_x\zeta_i}{\sigma_x} \right)^2 = 0. \tag{B.15}$$

This yields a quadratic equation, and I take the only one positive root as the stationary

posterior variance:

$$\begin{aligned}\zeta_i &= \frac{1}{2\left(\frac{\rho_x}{\sigma_x}\right)^2} \left[ -2\rho_x + \sqrt{4\rho_\mu^2 + 4\left(\frac{\rho_x}{\sigma_x}\right)^2 \left(\frac{\sigma_\mu^2}{1+\Phi_i}\right)} \right] \\ &= \left(\frac{\sigma_x}{\rho_x}\right)^2 \left[ \sqrt{\rho_\mu^2 + \left(\frac{1}{1+\Phi_i}\right) \left(\frac{\rho_x \sigma_\mu}{\sigma_x}\right)^2} - \rho_\mu \right].\end{aligned}\quad (\text{B.16})$$

## B.2 Proof of Corollary 1

Plugging the stationary posterior variance  $\zeta_i$  into equation (B.11) and replacing  $\zeta_{it}$  yields

$$d\hat{\mu}_{it} = \rho_\mu(\bar{\mu} - \hat{\mu}_{it})dt + \left(\frac{\sigma_x}{\rho_x}\right) \left[ \sqrt{\rho_\mu^2 + \left(\frac{1}{1+\Phi_i}\right) \left(\frac{\rho_x \sigma_\mu}{\sigma_x}\right)^2} - \rho_\mu \right] d\hat{Z}_{it}^x + \sigma_\mu \sqrt{\frac{\Phi_i}{1+\Phi_i}} d\hat{Z}_t^s, \quad (\text{B.17})$$

and the instantaneous variance of the posterior mean  $\hat{\mu}_{it}$  is

$$\begin{aligned}\text{Var}[d\hat{\mu}_{it}] &= \left(\frac{\sigma_x}{\rho_x}\right)^2 \left[ \sqrt{\rho_\mu^2 + \left(\frac{1}{1+\Phi_i}\right) \left(\frac{\rho_x \sigma_\mu}{\sigma_x}\right)^2} - \rho_\mu \right]^2 + \sigma_\mu^2 \left(\frac{\Phi_i}{1+\Phi_i}\right) \\ &= \sigma_\mu^2 - \left(\frac{2\rho_\mu \sigma_x^2}{\rho_x^2}\right) \left[ \sqrt{\rho_\mu^2 + \left(\frac{1}{1+\Phi_i}\right) \left(\frac{\rho_x \sigma_\mu}{\sigma_x}\right)^2} - \rho_\mu \right] \\ &= \sigma_\mu^2 - 2\rho_\mu \zeta_i\end{aligned}\quad (\text{B.18})$$

Trivially, the variance is increasing in  $\Phi_i$  by taking the partial derivative as follows:

$$\begin{aligned}\frac{\partial \text{Var}[d\hat{\mu}_{it}]}{\partial \Phi_i} &= -2\rho_\mu \frac{\partial \zeta_i}{\partial \Phi_i} \\ &= \frac{\rho_\mu \sigma_x \sigma_\mu^2}{(1+\Phi_i)^2 \sqrt{\rho_\mu^2 \sigma_x^2 + \frac{\rho_x^2 \sigma_\mu^2}{1+\Phi_i}}} > 0.\end{aligned}\quad (\text{B.19})$$

## B.3 Proof of Lemma 1

According to equation (3), firm  $i$ 's realized marginal product of capital (MPK) denotes:

$$\text{MPK}_i = \frac{\partial Y_i}{\partial K_i} = e^{a_{it} + x_t}. \quad (\text{B.20})$$

Therefore, firm  $i$ 's the logarithm of MPK is given by

$$\log \text{MPK}_i = a_{it} + x_t, \quad (\text{B.21})$$

and the variance captures the dispersion of the logarithm of MPK on the cross-section:

$$\text{Var}[\log(\text{MPK}_i)] = \text{Var}[a_{it}] + \text{Var}[x_t]. \quad (\text{B.22})$$

To compute the unconditional variance, I discretize the dynamics with respect to the firm-specific, short-run, and filtered long-run productivity:

$$a_{it+\Delta t} = e^{-\rho_a \Delta t} a_{it} + \sigma_a \sqrt{\frac{1 - e^{-2\rho_a \Delta t}}{2\rho_a}} \varepsilon_{it+\Delta t}^a, \quad (\text{B.23})$$

$$x_{t+\Delta t} = e^{-\rho_x \Delta t} x_t + \hat{\mu}_{it}(1 - e^{-\rho_x \Delta t}) + \sigma_x \sqrt{\frac{1 - e^{-2\rho_x \Delta t}}{2\rho_x}} \varepsilon_{t+\Delta t}^x, \quad (\text{B.24})$$

$$\hat{\mu}_{it+\Delta t} = e^{-\rho_\mu \Delta t} \hat{\mu}_{it} + \bar{\mu}(1 - e^{-\rho_\mu \Delta t}) + \sigma_\mu \sqrt{\frac{1 - e^{-2\rho_\mu \Delta t}}{2\rho_\mu}} \left( \Omega_i \varepsilon_{it+\Delta t}^x + \sigma_\mu \sqrt{\frac{\Phi_i}{1 + \Phi_i}} \varepsilon_{t+\Delta t}^\mu \right), \quad (\text{B.25})$$

where  $\varepsilon_{it}^a$ ,  $\varepsilon_t^x$ , and  $\varepsilon_t^\mu$  are i.i.d. standard normal variables<sup>5</sup>, and  $\Omega_i$  is given by

$$\Omega_i = \left( \frac{\sigma_x}{\rho_x} \right) \left[ \sqrt{\rho_\mu^2 + \left( \frac{1}{1 + \Phi_i} \right) \left( \frac{\rho_x \sigma_\mu}{\sigma_x} \right)^2} - \rho_\mu \right]. \quad (\text{B.26})$$

If  $\Delta t$  equal to 1, then the above discretized mean-reverting dynamics exactly follow AR(1) processes. A stationary AR(1) process implies the mean and variance are time-invariant. Therefore, the variance of the firm-specific productivity is

$$\text{Var}[a_{it+1}] = e^{-2\rho_a} \text{Var}[a_{it}] + \sigma_a^2 \left( \frac{1 - e^{-2\rho_a}}{2\rho_a} \right) \text{Var}[\varepsilon_{it+1}^a]. \quad (\text{B.27})$$

Given that  $\text{Var}[a_{it+1}]$  equals to  $\text{Var}[a_{it}]$ , rearrangement the above equation yields

$$\text{Var}[a_{it}] = \frac{\sigma_a^2}{2\rho_a}. \quad (\text{B.28})$$

Since the variance of the short-run productivity depends on that of the filtered long-run productivity, I first compute the variance of the filtered long-run productivity as follows:

$$\text{Var}[\hat{\mu}_{it+1}] = e^{-2\rho_\mu} \text{Var}[\hat{\mu}_{it}] + \sigma_\mu^2 \left( \frac{1 - e^{-2\rho_\mu}}{2\rho_\mu} \right) \text{Var} \left[ \Omega_i \varepsilon_{t+1}^x + \sigma_\mu \sqrt{\frac{\Phi_i}{1 + \Phi_i}} \varepsilon_{t+1}^\mu \right], \quad (\text{B.29})$$

---

<sup>5</sup>  $\varepsilon_{it}^a, \varepsilon_t^x, \varepsilon_t^\mu \stackrel{iid}{\sim} N(0, 1)$

where  $\text{Var}\left[\Omega_i \varepsilon_{t+1}^x + \sigma_\mu \sqrt{\frac{\Phi_i}{1+\Phi_i}} \varepsilon_{t+1}^\mu\right]$  is similar with the result in equation (B.18). As a result, I obtain

$$\begin{aligned}\text{Var}[\hat{\mu}_{it}] &= \left(\frac{1}{2\rho_\mu}\right)(\sigma_\mu^2 - 2\rho_\mu \zeta_i) \\ &= \left(\frac{\sigma_\mu^2}{2\rho_\mu}\right) - \zeta_i.\end{aligned}\tag{B.30}$$

The unconditional variance of  $x_t$  is available when  $t$  and  $\Delta t$  go to infinity:

$$\begin{aligned}\text{Var}[x_{t+\Delta t}] &= e^{-2\rho_x \Delta t} \text{Var}[x_t] + (1 - e^{-\rho_x \Delta t})^2 \text{Var}[\hat{\mu}_{it}] \\ &\quad + 2(1 - e^{-\rho_x \Delta t}) \sigma_x \sqrt{\frac{1 - e^{-2\rho_x \Delta t}}{2\rho_x}} \text{Cov}(\hat{\mu}_{it}, \varepsilon_{t+\Delta t}^x) \\ &\quad + \sigma_x^2 \left(\frac{1 - e^{-2\rho_x \Delta t}}{2\rho_x}\right) \text{Var}[\varepsilon_{t+\Delta t}^x].\end{aligned}\tag{B.31}$$

$$\begin{aligned}\text{Var}[x_t] &= \lim_{\Delta t \rightarrow \infty} \frac{(1 - e^{-\rho_x \Delta t})^2}{(1 - e^{-2\rho_x \Delta t})} \text{Var}[\hat{\mu}_{it}] + \lim_{\Delta t \rightarrow \infty} \left(\frac{\sigma_x^2}{1 - e^{-2\rho_x \Delta t}}\right) \left(\frac{1 - e^{-2\rho_x \Delta t}}{2\rho_x}\right) \\ &\quad + 2\sigma_x \lim_{\Delta t \rightarrow \infty} \left(\frac{1 - e^{-\rho_x \Delta t}}{1 - e^{-2\rho_x \Delta t}}\right) \left(\sqrt{\frac{1 - e^{-2\rho_x \Delta t}}{2\rho_x}}\right) \sigma_\mu \left(\sqrt{\frac{1 - e^{-2\rho_\mu \Delta t}}{2\rho_\mu}}\right) \Omega_i \\ &= \frac{\sigma_\mu^2}{2\rho_\mu} - \zeta_i + \frac{\sigma_x^2}{2\rho_x} + \frac{\sigma_\mu \rho_x \zeta_i}{\sqrt{4\rho_x \rho_\mu}} \\ &= \frac{\sigma_\mu^2}{2\rho_\mu} + \frac{\sigma_x^2}{2\rho_x} + \left(-1 + \sigma_\mu \sqrt{\frac{\rho_x}{\rho_\mu}}\right) \zeta_i,\end{aligned}\tag{B.32}$$

where I replace  $\Omega_i$  in the second equality according to equation (B.16) and (B.42):

$$\Omega_i = \left(\frac{\rho_x}{\sigma_x}\right) \zeta_i.\tag{B.33}$$

The partial derivative of  $\text{Var}[x_t]$  with respect to  $\Phi_i$  is:

$$\frac{\partial}{\partial \Phi_i} \text{Var}[x_t] = \left(-1 + \sigma_\mu \sqrt{\frac{\rho_x}{\rho_\mu}}\right) \frac{\partial \zeta_i}{\partial \Phi_i} < 0,\tag{B.34}$$

when the sufficient condition is held:

$$\sigma_\mu^2 \rho_x > \rho_\mu.\tag{B.35}$$

Finally, I can then compute:

$$\frac{\partial}{\partial \Phi_i} \text{Var}[\log(\text{MPK}_i)] = \frac{\partial}{\partial \Phi_i} \text{Var}[x_t] = \left( -1 + \sigma_\mu \sqrt{\frac{\rho_x}{\rho_\mu}} \right) \frac{\partial \zeta_i}{\partial \Phi_i} < 0. \quad (\text{B.36})$$

## B.4 Proof of Lemma 2

Aggregating filtered short-run productivity shocks across all firms yields

$$\begin{aligned} d\hat{Z}_t^x &= \int d\hat{Z}_{it}^x di \\ &= dZ_t^x + \left( \frac{\rho_x}{\sigma_x} \right) (\mu_t - \hat{\mu}_t) dt. \end{aligned} \quad (\text{B.37})$$

An application of the law of large numbers implies that  $\Phi_i$  converges to its cross-sectional mean  $\Phi$ . The aggregate signal shocks at time  $t$  is given by

$$\begin{aligned} ds_t &= \int ds_{it} di = \int \left( dZ^\mu + \frac{1}{\sqrt{\Phi_i}} dZ_t^s \right) di \\ &= dZ_t^\mu + \int \frac{1}{\sqrt{\Phi_i}} dZ_t^s di \\ &= dZ_t^\mu + \int \left( \frac{1}{\sqrt{\Phi_i}} \right) di dZ_t^s \\ &\rightarrow dZ_t^\mu + \mathbb{E}_i \left( \frac{1}{\sqrt{\Phi_i}} \right) dZ_t^s \\ &= dZ_t^\mu + \frac{1}{\sqrt{\Phi}} dZ_t^s \\ &= \left( \frac{1 + \Phi}{\sqrt{\Phi}} \right) d\hat{Z}_t^s, \end{aligned} \quad (\text{B.38})$$

where the integration and differentiation operator, without loss of generality, are interchangeable, and  $\mathbb{E}_i[\cdot] = \int di$  is the operator of cross-sectional expectation.<sup>6</sup>

## B.5 Proof of Proposition 2

Firm  $i$ 's optimization problem is to maximize its expected summation of future dividends (i.e., cash-flows) discounted by the SDF, after subtracting its investment and adjustment cost, subject to the law of motion of capital accumulation in equation (7) and the adjustment cost

---

<sup>6</sup>The exemplary of interchange operators is as follows:

$$d\omega_t = \int (d\omega_{it}) di = d \left( \int \omega_{it} di \right),$$

where  $d\omega_t$  is the aggregation of the dynamics  $d\omega_{it}$ .

in equation (8). Consider the risk-neutral probability measure  $\mathcal{Q}$ , implicitly defined given my specification for the SDF. Under this measure, the firm's valuation equals:

$$\begin{aligned} V(a_{it}, x_t, \hat{\mu}_{it}, K_{it}) &= \text{Max}_{I_s} \mathbb{E}_t \left[ \int_t^\infty \frac{\pi_s}{\pi_t} (Y_{is} - I_{is} - \Psi(I_{is}, K_{is})) ds \right] \\ &= \text{Max}_{I_s} \mathbb{E}_t^{\mathcal{Q}} \left[ \int_t^\infty e^{-r_f(s-t)} (Y_{is} - I_{is} - \Psi(I_{is}, K_{is})) ds \right], \end{aligned} \quad (\text{B.39})$$

where  $\mathbb{E}_t^{\mathcal{Q}}$  denotes expectations under the risk-neutral measure  $\mathcal{Q}$ , and

$$\frac{d\mathcal{Q}}{d\mathcal{P}} = \exp \left( -\lambda_x \hat{Z}_t^x - \lambda_\mu \hat{Z}_t^\mu - \frac{1}{2} \lambda_x^2 t - \frac{1}{2} \lambda_\mu^2 t \right), \quad (\text{B.40})$$

where  $\mathcal{P}$  denotes the physical probability measure. The Hamilton–Jacobi–Bellman (HJB) equation associated with the above optimization problem is

$$r_f V_{it} = \text{Max}_{I_{it}} e^{a_{it} + x_t} K - I_{it} - \Psi(I_{it}, K_{it}) + \mathcal{D}[V_{it}], \quad (\text{B.41})$$

where  $\mathcal{D}$  is the differential operator. I guess that firm  $i$  value can be written as

$$V(a_{it}, x_t, \hat{\mu}_{it}, K_{it}) = q(a_{it}, x_t, \hat{\mu}_{it}) K_{it}. \quad (\text{B.42})$$

The first order condition from the Hamilton–Jacobi–Bellman equation (B.39) yields the level of optimal investment:

$$\begin{aligned} 0 &= -1 - \chi \left( \frac{I_{it}}{K_{it}} \right) + q(a_{it}, x_t, \hat{\mu}_{it}) \\ \frac{I_{it}}{K_{it}} &= -\frac{1}{\chi} + \frac{1}{\chi} q(a_{it}, x_t, \hat{\mu}_{it}). \end{aligned} \quad (\text{B.43})$$

The second equality of the above equation implies a one-to-one mapping relation between the optimal investment rate and Tobin's  $q$ .

For the notation brevity, I temporarily suppress subscript. Using conjectured value function equation (B.42) and the optimal investment rate in equation (B.43), I obtain the following partial differential equation for  $q$  according to the HJB equation (B.41):

$$\begin{aligned} 0 &= e^{a+x} - \left( -\frac{1}{\chi} + \frac{1}{\chi} q \right) - \frac{\chi}{2} \left( -\frac{1}{\chi} + \frac{1}{\chi} q \right)^2 + q \left[ -\frac{1}{\chi} + \frac{1}{\chi} q - (r_f + \delta) \right] \\ &\quad q_t - q_a \rho_a a + q_x \rho_x (\hat{\mu} - x) + q_{\hat{\mu}} \rho_{\hat{\mu}} (\bar{\mu} - \hat{\mu}) \\ &\quad + \frac{1}{2} q_{aa} \sigma_a^2 + \frac{1}{2} q_{xx} \sigma_x^2 + \frac{1}{2} q_{\hat{\mu}\hat{\mu}} (\sigma_\mu^2 - 2\rho_\mu \zeta) + q_{x\hat{\mu}} \sigma_x \left( \frac{\rho_x \zeta}{\sigma_x} \right). \end{aligned} \quad (\text{B.44})$$

Following Ai (2010), I conjecture that  $q$  takes the exponential form:

$$q(a, x, \hat{\mu}, t) \equiv e^{\eta_a a + \eta_x x + \eta_{\hat{\mu}} \hat{\mu} + \eta(t)}, \quad (\text{B.45})$$

where  $q_a(a, x, \hat{\mu}, t) = \eta_a q(a, x, \hat{\mu}, t)$ ,  $q_x(a, x, \hat{\mu}, t) = \eta_x q(a, x, \hat{\mu}, t)$ ,  $q_{\hat{\mu}}(a, x, \hat{\mu}, t) = \eta_{\hat{\mu}} q(a, x, \hat{\mu}, t)$ , and  $q_t(a, x, \hat{\mu}, t) = \eta'(t) q(a, x, \hat{\mu}, t)$ . Substituting these terms into equation (B.44) and dividing by  $q_a(a, x, \hat{\mu}, t)$ , I would get

$$\begin{aligned} 0 &= \left[ e^{a+x} q^{-1} + \frac{1}{2\chi} q^{-1} + \eta'(t) \right] + \left[ 1 - \frac{2}{\chi} - (r_f + \delta) \right] + \frac{1}{2\chi} q \\ &\quad - \eta_a \rho_a a + \eta_x \rho_x (\hat{\mu} - x) + \eta_{\hat{\mu}} \rho_{\hat{\mu}} (\bar{\mu} - \hat{\mu}) \\ &\quad + \frac{1}{2} \eta_a^2 \sigma_a^2 + \frac{1}{2} \eta_x^2 \sigma_x^2 + \frac{1}{2} \eta_{\hat{\mu}}^2 (\sigma_{\hat{\mu}}^2 - 2\rho_{\hat{\mu}} \zeta) + \eta_x \eta_{\hat{\mu}} \rho_x \zeta. \end{aligned} \quad (\text{B.46})$$

Using  $e^c - 1 \approx c$  to approximate and simplify, the above HJB equation becomes

$$\begin{aligned} 0 &= \left[ 1 + (1 - \eta_a) a + (1 - \eta_x) x + \eta_{\hat{\mu}} \hat{\mu} - \eta(t) \right] + \frac{1}{2\chi} \left[ 1 - \eta_a a - \eta_x x - \eta_{\hat{\mu}} \hat{\mu} - \eta(t) \right] \\ &\quad + \eta'(t) + \left[ 1 - \frac{2}{\chi} - (r_f + \delta) \right] + \frac{1}{2\chi} \left[ 1 + \eta_a a + \eta_x x + \eta_{\hat{\mu}} \hat{\mu} + \eta(t) \right] \\ &\quad - \eta_a \rho_a a + \eta_x \rho_x (\hat{\mu} - x) + \eta_{\hat{\mu}} \rho_{\hat{\mu}} (\bar{\mu} - \hat{\mu}) \\ &\quad + \frac{1}{2} \eta_a^2 \sigma_a^2 + \frac{1}{2} \eta_x^2 \sigma_x^2 + \frac{1}{2} \eta_{\hat{\mu}}^2 (\sigma_{\hat{\mu}}^2 - 2\rho_{\hat{\mu}} \zeta) + \eta_x \eta_{\hat{\mu}} \rho_x \zeta \\ &= \left[ 2 - \frac{1}{\chi} - (r_f + \delta) \right] - \eta(t) + \eta'(t) + \eta_{\hat{\mu}} \rho_{\hat{\mu}} \bar{\mu} + \frac{1}{2} \eta_a^2 \sigma_a^2 + \frac{1}{2} \eta_x^2 \sigma_x^2 + \frac{1}{2} \eta_{\hat{\mu}}^2 (\sigma_{\hat{\mu}}^2 - 2\rho_{\hat{\mu}} \zeta) + \eta_x \eta_{\hat{\mu}} \rho_x \zeta \\ &\quad + (1 - \eta_a - \eta_a \rho_a) a + (1 - \eta_x - \eta_x \rho_x) x + (1 - \eta_{\hat{\mu}} + \eta_x \rho_x - \eta_{\hat{\mu}} \rho_{\hat{\mu}}) \hat{\mu}. \end{aligned} \quad (\text{B.47})$$

Matching the coefficients with respect to  $a$ ,  $x$ ,  $\hat{\mu}$ , and  $t$  yields:

$$\eta_a = \frac{1}{1 + \rho_a}, \quad (\text{B.48})$$

$$\eta_x = \frac{1}{1 + \rho_x}, \quad (\text{B.49})$$

$$\eta_{\hat{\mu}} = \frac{1 + 2\rho_x}{(1 + \rho_x)(1 + \rho_{\hat{\mu}})}. \quad (\text{B.50})$$

Given the initial condition  $\eta(0) = 0$ , I determine the functional form of  $\eta(t)$  by the solving an ordinary differential equation (ODE) as follows:

$$\eta'(t) - \eta(t) + \bar{\eta} = 0, \quad (\text{B.51})$$



where  $\bar{\eta}$  is a time-unvarying constant and given by

$$\begin{aligned}\bar{\eta} &= 2 - \frac{1}{\chi} - (r_f + \delta) + \eta_{\hat{\mu}}\rho_{\mu}\bar{\mu} + \frac{1}{2}\eta_a^2\sigma_a^2 + \frac{1}{2}\eta_x^2\sigma_x^2 \\ &\quad + \frac{1}{2}\eta_{\hat{\mu}}^2(\sigma_{\mu}^2 - 2\rho_{\mu}\zeta) + \eta_x\eta_{\hat{\mu}}\rho_x\zeta,\end{aligned}\tag{B.52}$$

and the initial condition is  $\eta(0) = 0$ . Note that in order to solve for asset prices, I do not have to pin down the functional form of  $\eta(t)$ . Nevertheless, I need the functional form of  $\eta(t)$  to compute market equities and book-to-market ratios in the economy.

$\eta(t)$  follows a linear first order ODE and can be solved using the integrating factor method:

$$\eta(t) = \bar{\eta}(1 - e^{-t})\tag{B.53}$$

## B.6 Proof of Proposition 3

I define firm  $i$ 's realized stock returns by using Ito's Lemma on firm  $i$ 's stock value in equation (19). The firm's realized returns are the summation of dividend yields and capital gains in the first and second term of the first equality at the left hand side in the following equation:

$$\begin{aligned}\frac{D_{it}dt + dV_{it}}{V_{it}} &= \mathbb{E}_t \left[ \frac{D_{it}dt + dV_{it}}{V_{it}} \right] + \frac{\mathcal{D}[q(a_{it}, x_t, \hat{\mu}_{it})K_{it}]}{q(a_{it}, x_t, \hat{\mu}_{it})K_{it}} \\ &= \mathbb{E}_t \left[ \frac{D_{it}dt + dV_{it}}{V_{it}} \right] + \beta_{a,it}dZ_{it}^a + \beta_{x,it}d\hat{Z}_{it}^x + \beta_{\mu,it}d\hat{Z}_{it}^{\mu},\end{aligned}\tag{B.54}$$

where

$$\beta_{a,it} = \eta_a\sigma_a,\tag{B.55}$$

$$\beta_{x,it} = \eta_x\sigma_x + \eta_{\hat{\mu}}\left(\frac{\rho_x\zeta_i}{\sigma_x}\right),\tag{B.56}$$

$$\beta_{\mu,it} = \eta_{\hat{\mu}}\sigma_{\mu}\sqrt{\frac{\Phi_i}{1 + \Phi_i}}.\tag{B.57}$$

## B.7 Proof of Corollary 2

According to equation (21), firm  $i$ 's risk premium is summarized as two components with respect to the short- and signal shocks. The heterogeneity in expected returns depends on heterogeneity in information precision among firms. Therefore, I present the partial derivative

of firm's risk premium with respect to its information precision as follows:

$$\frac{\partial}{\partial \Phi_i} (\beta_{x,it} \lambda_x + \beta_{\mu,it} \lambda_\mu) = \frac{\partial \beta_{x,it}}{\partial \Phi_i} \lambda_x + \frac{\partial \beta_{\mu,it}}{\partial \Phi_i} \lambda_\mu, \quad (\text{B.58})$$

where  $\lambda_x$  and  $\lambda_\mu$  are constant according to equation (16). The comparative static analysis is going to focus on  $\frac{\partial \beta_{x,it}}{\partial \Phi_i}$  and  $\frac{\partial \beta_{\mu,it}}{\partial \Phi_i}$ , respectively. For the convenience of the following analysis, I apply the log-linear approximation of equation (B.16) and obtain,

$$\begin{aligned} \zeta_i &= \left(\frac{\sigma_x}{\rho_x}\right)^2 \left[ \sqrt{\rho_\mu^2 + \left(\frac{1}{1+\Phi_i}\right) \left(\frac{\rho_x \sigma_\mu}{\sigma_x}\right)^2} - \rho_\mu \right] \\ &= \rho_\mu \left(\frac{\sigma_x}{\rho_x}\right)^2 \left[ \sqrt{1 + \left(\frac{1}{1+\Phi_i}\right) \left(\frac{\rho_x \sigma_\mu}{\rho_\mu \sigma_x}\right)^2} - 1 \right] \\ &\approx -\frac{1}{2} \left(\frac{\sigma_\mu^2}{\rho_\mu}\right) \left(\frac{1}{1+\Phi_i}\right). \end{aligned} \quad (\text{B.59})$$

As a result, plugging equation (B.56) and (B.59) into the first component in the right hand side of equation (B.58) yields

$$\frac{\partial \beta_{x,it}}{\partial \Phi_i} \lambda_x = -\frac{1}{2} \left(\frac{\eta_{\hat{\mu}} \rho_x}{\sigma_x}\right) \left(\frac{\sigma_\mu^2}{\rho_\mu}\right) \left(\frac{1}{1+\Phi_i}\right)^2 \lambda_x. \quad (\text{B.60})$$

To analysis the sensitivity of risk premium attributing to signal shocks, I apply the similar trick of the log-linear approximation and obtain

$$\sqrt{\frac{\Phi_i}{1+\Phi_i}} \approx 1 - \frac{1}{2} \left(\frac{\Phi_i}{1+\Phi_i}\right), \quad (\text{B.61})$$

and

$$\frac{\partial}{\partial \Phi_i} \sqrt{\frac{\Phi_i}{1+\Phi_i}} \approx \frac{1}{2} \left(\frac{1}{1+\Phi_i}\right)^2. \quad (\text{B.62})$$

The second component in the right hand side of equation (B.58) can be rewritten as

$$\frac{\partial \beta_{\mu,it}}{\partial \Phi_i} \lambda_\mu = \frac{1}{2} \eta_{\hat{\mu}} \sigma_\mu \left(\frac{1}{1+\Phi_i}\right)^2 \lambda_\mu. \quad (\text{B.63})$$

Finally, the sensitivity of firm  $i$ 's risk premium to its information precision  $\Phi_i$  is characterized as follows:

$$\begin{aligned}
\frac{\partial}{\partial \Phi_i} (\beta_{x,it} \lambda_x + \beta_{\mu,it} \lambda_\mu) &= \frac{\partial \beta_{x,it}}{\partial \Phi_i} \lambda_x + \frac{\partial \beta_{\mu,it}}{\partial \Phi_i} \lambda_\mu \\
&= -\frac{1}{2} \left( \frac{\eta_{\hat{\mu}} \rho_x}{\sigma_x} \right) \left( \frac{\sigma_\mu^2}{\rho_\mu} \right) \left( \frac{1}{1 + \Phi_i} \right)^2 \lambda_x + \frac{1}{2} \eta_{\hat{\mu}} \sigma_\mu \left( \frac{1}{1 + \Phi_i} \right)^2 \lambda_\mu \\
&= \frac{1}{2} \eta_{\hat{\mu}} \sigma_\mu \left[ - \left( \frac{\rho_x \sigma_\mu}{\rho_\mu \sigma_x} \right) \lambda_x + \lambda_\mu \right] > 0,
\end{aligned} \tag{B.64}$$

where the positive partial derivative is positive in the above equation when the sufficient condition is held:

$$- \left( \frac{\rho_x \sigma_\mu}{\rho_\mu \sigma_x} \right) \lambda_x + \lambda_\mu > 0. \tag{B.65}$$

Alternatively, I can numerically verify that a firm's expected return is increasing in its information precision  $\Phi_i$ . Figure A1 illustrates the equilibrium expected return as a function of  $\Phi_i$ . As information precision increases (as I move along the x-axis), a firm are better to update its belief about unobservable long-run productivity, and, its cash flows become more procyclical. Therefore, its stock is riskier to earn a higher expected return.

[Place Figure A1 about here]

## B.8 Proof of Proposition 4

The long-short portfolio return in equation (24) is available when I substitute equation (20) with respect to high versus low information precision firm into the covariance term in equation (21), and then take a difference of expected returns.

## B.9 Proof of Proposition 5

According to equation (28), I calculate its Macaulay duration recursively as follows:

$$\begin{aligned}
M(a_{it}, x_t, \hat{\mu}_{it}) q(a_{it}, x_t, \hat{\mu}_{it}) &= \mathbb{E}_t \left[ \int_t^\infty s \frac{\pi_s}{\pi_t} \Psi_{is} ds \right] \\
&= \mathbb{E}_t^Q \left[ \int_t^\infty s e^{r_f(s-t)} \Psi_{is} ds \right],
\end{aligned} \tag{B.66}$$

in which  $\mathbb{E}_t^{\mathcal{Q}}$  denotes expectations under the risk-neutral measure  $\mathcal{Q}$ , and  $\mathcal{P}$  denotes the physical probability measure. To derive Macaulay duration, I consider a small fraction of time  $\Delta t$  such that:

$$\begin{aligned}
M_{it}q_{it} &= \mathbb{E}_t \left[ \int_t^{t+\Delta t} s \frac{\pi_s}{\pi_t} \Psi_{is} ds \right] + \mathbb{E}_t \left[ \int_{t+\Delta t}^{\infty} s \frac{\pi_s}{\pi_t} \Psi_{is} ds \right] \\
&= \mathbb{E}_t^{\mathcal{Q}} \left[ \int_t^{\infty} \Delta t e^{r_f(s-t)} \Psi_{is} ds \right] + \mathbb{E}_t^{\mathcal{Q}} \left[ \int_t^{\infty} (s - \Delta t) e^{r_f(s-t)} \Psi_{is} ds \right] \\
&= \mathbb{E}_t^{\mathcal{Q}} \left[ \int_t^{\infty} e^{r_f(s-t)} \Psi_{is} ds \right] \Delta t + \mathbb{E}_t^{\mathcal{Q}} \left[ \mathbb{E}_{t+\Delta t}^{\mathcal{Q}} \left[ \int_t^{\infty} (s - \Delta t) e^{r_f(s-t)} \Psi_{is} ds \right] \right] \\
&= q_{it} \Delta t + \mathbb{E}_t^{\mathcal{Q}} \left[ e^{-r_f \Delta t} M_{it+\Delta t} q_{it+\Delta t} \right] \tag{B.67}
\end{aligned}$$

For the notation brevity, I temporarily suppress subscript and take the limit as  $\Delta t \rightarrow 0$ . Applying Ito's lemma to expand the multiplication of  $M$  and  $q$ , the Hamilton–Jacobian–Bellman (HJB) equation associated with the above equation denotes:

$$\begin{aligned}
r_f M q dt &= q dt + \mathcal{D}[Mq] \\
r_f dt &= \frac{1}{M} dt + \mathbb{E}_t^{\mathcal{Q}} \left[ \frac{dM}{M} + \frac{dq}{q} + \frac{dM}{M} \frac{dq}{q} \right]. \tag{B.68}
\end{aligned}$$

As is shown in Section B.5, I conjecture that  $M$  takes the exponential form:

$$M = e^{\phi_a + \phi_x x + \phi_{\hat{\mu}} \hat{\mu} + \phi(t)}, \tag{B.69}$$

in which  $M_a = \phi_a M$ ,  $M_x = \phi_x M$ ,  $M_{\hat{\mu}} = \phi_{\hat{\mu}} M$ , and  $M_t = \phi'(t) M$ . The above HJB equation is simplified when I use the first order to approximation for  $1/M$  and substitute these terms and the functional form of  $q$  in equation (B.45) into equation (B.68):

$$\begin{aligned}
r_f &= 1 - \phi_a a - \phi_x x - \phi_{\hat{\mu}} \hat{\mu} - \phi_a \rho_a a + \phi_x \rho_x (\hat{\mu} - x) + \phi_{\hat{\mu}} \rho_{\hat{\mu}} (\bar{\mu} - \hat{\mu}) \\
&\quad + \frac{1}{2} \phi_a^2 \sigma_a^2 + \frac{1}{2} \phi_x^2 \sigma_x^2 + \frac{1}{2} \phi_{\hat{\mu}}^2 (\sigma_{\hat{\mu}}^2 - 2\rho_{\hat{\mu}} \zeta) + \phi_x \phi_{\hat{\mu}} \sigma_x \left( \frac{\rho_x \zeta}{\sigma_x} \right) \\
&\quad - \eta_a \rho_a a + \eta_x \rho_x (\hat{\mu} - x) + \eta_{\hat{\mu}} \rho_{\hat{\mu}} (\bar{\mu} - \hat{\mu}) + \frac{1}{2} \eta_a^2 \sigma_a^2 + \frac{1}{2} \eta_x^2 \sigma_x^2 + \frac{1}{2} \eta_{\hat{\mu}}^2 (\sigma_{\hat{\mu}}^2 - 2\rho_{\hat{\mu}} \zeta) \\
&\quad + \eta_x \eta_{\hat{\mu}} \sigma_x \left( \frac{\rho_x \zeta}{\sigma_x} \right) + \phi_a \eta_a \sigma_a^2 + \phi_x \eta_x \sigma_x \left( \frac{\rho_x \zeta}{\sigma_x} \right) + \phi_{\hat{\mu}} \eta_{\hat{\mu}} (\sigma_{\hat{\mu}}^2 - 2\rho_{\hat{\mu}} \zeta) + \phi_{\hat{\mu}} \eta_x \sigma_x \left( \frac{\rho_x \zeta}{\sigma_x} \right). \tag{B.70}
\end{aligned}$$

Matching the coefficients with respect to  $a$ ,  $x$ ,  $\hat{\mu}$ , and  $t$  determines:

$$\phi_a = \frac{-\eta_a \rho_a}{1 + \rho_a} < 0, \quad (\text{B.71})$$

$$\phi_x = \frac{-\eta_x \rho_x}{1 + \rho_x} < 0, \quad (\text{B.72})$$

$$\begin{aligned} \phi_{\hat{\mu}} &= \frac{\phi_x \rho_x - \eta_x \rho_x - \eta_{\hat{\mu}} \rho_{\hat{\mu}}}{1 + \rho_{\hat{\mu}}} \\ &= \frac{1}{1 + \rho_{\hat{\mu}}} \left[ -\frac{1 + 2\rho_x}{1 + \rho_{\hat{\mu}}} \rho_x \eta_x - \frac{1 + 2\rho_x}{(1 + \rho_x)(1 + \rho_{\hat{\mu}})} \right] < 0, \end{aligned} \quad (\text{B.73})$$

Since  $\eta_x$  is strictly positive according to equation (B.49), I can show that  $\phi_a$ ,  $\phi_x$ , and  $\phi_{\hat{\mu}}$  are strictly negative.

Given the initial condition  $\phi(0) = 0$ , I determine the functional form of  $\phi(t)$  by the solving an ordinary differential equation (ODE) as follows:

$$\phi'(t) - \phi(t) + \bar{\phi} = 0, \quad (\text{B.74})$$

where  $\bar{\phi}$  is a time-unvarying constant and characterized as follows:

$$\begin{aligned} \bar{\phi} &= 1 - r_f + \phi_{\hat{\mu}} \rho_{\hat{\mu}} \bar{\mu} + \frac{1}{2} \phi_a^2 \sigma_a^2 + \frac{1}{2} \phi_x^2 \sigma_x^2 + \frac{1}{2} \phi_{\hat{\mu}}^2 (\sigma_{\hat{\mu}}^2 - 2\rho_{\hat{\mu}} \zeta) + \phi_x \phi_{\hat{\mu}} \sigma_x \left( \frac{\rho_x \zeta}{\sigma_x} \right) \\ &\quad + \eta_{\hat{\mu}} \rho_{\hat{\mu}} \bar{\mu} + \frac{1}{2} \eta_a^2 \sigma_a^2 + \frac{1}{2} \eta_x^2 \sigma_x^2 + \frac{1}{2} \eta_{\hat{\mu}}^2 (\sigma_{\hat{\mu}}^2 - 2\rho_{\hat{\mu}} \zeta) + \eta_x \eta_{\hat{\mu}} \sigma_x \left( \frac{\rho_x \zeta}{\sigma_x} \right) + \phi_a \eta_a \sigma_a^2 + \phi_x \eta_{\hat{\mu}} \sigma_x \left( \frac{\rho_x \zeta}{\sigma_x} \right) \\ &\quad + \phi_{\hat{\mu}} \eta_{\hat{\mu}} (\sigma_{\hat{\mu}}^2 - 2\rho_{\hat{\mu}} \zeta) + \phi_{\hat{\mu}} \eta_x \sigma_x \left( \frac{\rho_x \zeta}{\sigma_x} \right), \end{aligned} \quad (\text{B.75})$$

and the initial condition is  $\phi(0) = 0$ . Note that in order to solve for asset prices, I do not have to pin down the functional form of  $\phi(t)$ . Nevertheless, I need the functional form of  $\phi(t)$  to compute market equities and book-to-market ratios in the economy.

$\phi(t)$  follows a linear first order ODE and can be solved using the integrating factor method:

$$\phi(t) = \bar{\phi}(1 - e^t) \quad (\text{B.76})$$

## References

- Ai, Hengjie, 2010, Information quality and long-run risk: Asset pricing implications, *The Journal of Finance* 65, 1333–1367.
- Andrei, Daniel, William Mann, and Nathalie Moyen, 2019, Why did the q theory of investment start working?, *Journal of Financial Economics* 133, 251–272.
- Ball, Ray, Joseph Gerakos, Juhani T Linnainmaa, and Valeri Nikolaev, 2016, Accruals, cash flows, and operating profitability in the cross section of stock returns, *Journal of Financial Economics* 121, 28–45.
- Ball, Ray, Joseph Gerakos, Juhani T Linnainmaa, and Valeri V Nikolaev, 2015, Deflating profitability, *Journal of Financial Economics* 117, 225–248.
- Blundell, Richard, and Stephen Bond, 2000, Gmm estimation with persistent panel data: an application to production functions, *Econometric reviews* 19, 321–340.
- Chen, Kaiji, and Zheng Song, 2013, Financial frictions on capital allocation: A transmission mechanism of tfp fluctuations, *Journal of Monetary Economics* 60, 683–703.
- David, Joel M, Lukas Schmid, and David Zeke, 2018, Risk-adjusted capital allocation and misallocation, *Working Paper* .
- Dechow, Patricia M, Richard G Sloan, and Mark T Soliman, 2004, Implied equity duration: A new measure of equity risk, *Review of Accounting Studies* 9, 197–228.
- Erickson, Timothy, and Toni M Whited, 2000, Measurement error and the relationship between investment and q, *Journal of political economy* 108, 1027–1057.
- Erickson, Timothy, and Toni M Whited, 2002, Two-step gmm estimation of the errors-in-variables model using high-order moments, *Econometric Theory* 776–799.
- Erickson, Timothy, and Toni M Whited, 2012, Treating measurement error in tobin’s q, *The Review of Financial Studies* 25, 1286–1329.

- Fama, Eugene F, and Kenneth R French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.
- Fama, Eugene F, and Kenneth R French, 1997, Industry costs of equity, *Journal of financial economics* 43, 153–193.
- Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: An investment approach, *Review of Financial Studies* hhu068.
- Hsieh, Chang-Tai, and Peter J Klenow, 2009, Misallocation and manufacturing tfp in china and india, *The Quarterly journal of economics* 124, 1403–1448.
- Lettau, Martin, and Jessica A Wachter, 2007, Why is long-horizon equity less risky? a duration-based explanation of the value premium, *The Journal of Finance* 62, 55–92.
- Li, Kai, Chi-Yang Tsou, and Chenjie Xu, 2019, Learning and the capital age premium, *Available at SSRN 3225041* .
- Lin, Xiaoji, Bernardino Palazzo, and Fan Yang, 2019, The risks of old capital age: Asset pricing implications of technology adoption, *Journal of Monetary Economics* .
- Scheinkman, Jose A, and Wei Xiong, 2003, Overconfidence and speculative bubbles, *Journal of Political Economy* 111, 1183–1220.
- Shumway, Tyler, 1997, The delisting bias in crsp data, *The Journal of Finance* 52, 327–340.
- Weber, Michael, 2018, Cash flow duration and the term structure of equity returns, *Journal of Financial Economics* 128, 486–503.
- Wu, Jin, Lu Zhang, and X Frank Zhang, 2010, The q-theory approach to understanding the accrual anomaly, *Journal of Accounting Research* 48, 177–223.

Figure A1: **Expected Return and Information Precision**

This figure plots the expected return as a function of information precision  $\Phi_i$ . Other parameter values are set according to Table 6.

