Essays on Trading Strategies, Corporate Activities, and Firm Performance

Siqi Liu

ICMA Centre

HENLEY BUSINESS SCHOOL

UNIVERSITY OF READING

A dissertation submitted in the requirements for the degree of

Doctor of Philosophy

April 2021

Abstract

This thesis focuses on three research questions in the areas of empirical asset pricing and corporate finance. I introduce the overview in the first chapter and conclude in the final chapter.

In the second chapter, we investigate the impact of the beta's statistical significance on the performance of the betting against beta (BAB) portfolios in the U.S. and major international markets. After dropping stocks with statistically insignificant betas, we find that a betting against statistically significant beta strategy reduces the monthly alphas of BAB portfolios by 20% – 50%, depending on beta estimation methods. If we replace the value of statistically insignificant beta by zero, a refined BAB strategy can generate a higher alpha than the original BAB strategy.

In the third chapter, we find a negative relationship between abnormal investment and future stock performance in the U.S. market. This negative relation is mainly driven by firm under-investment, not over-investment. Our explanations can be that market investors may not react promptly to the fundamental information contained in under-investment about a firm's future profitability, asset growth, and financial distress probability. Alternatively, the negative relation between under-investment and future stock returns is more pronounced for firms with lower investor monitoring and higher agency costs.

In the fourth chapter, we identify a positive link between peer firms' investment and focal firm's value of cash holding in the U.S. market. This effect is likely to result from the positive externalities brought by peer investment which are reflected in young and growing industries with ample investment opportunities that are shared by the focal firms. We find little evidence to support either the precautionary hypothesis or the learning hypothesis. Our further analyses show that firms increase their level of cash holdings while peer investment increases. Meanwhile, firms are less willing to use cash for dividend payments while they are more willing to use it for capital expenditure and R&D investment.

Acknowledgements

First and foremost, I would like to deeply thank my mentor Dr Yeqin Zeng. It has been a great honor to be his Ph.D. student, and I really appreciate his immensely unreserved guidance and support throughout my Ph.D. pursuit. He has taught me how to post a research question, formulate the structure, address the issues, write academic papers, and become an academic. Without his guidance, the quality of the research would not be promised. Moreover, I would like to thank my another esteemed supervisor Professor Michael P. Clements for his continuous support and tutelage. I am especially grateful for the warm advice he has offered during the Covid-19 period. Furthermore, I would like to thank Dr Chao Yin for his long-standing support and insightful comments and suggestions into this research. Particularly, I would like to express my sincere gratitude to Professor Mark Tippett for his unwavering support and belief in me which broadens my academic horizon. I would like to greatly thank Dr Lisa Schopohl for being my annual examiner in the last few years and supporting me.

My families have been great comfort and encouragements for my dissertation work. I would like to offer special thanks to my families who have always been there for me every step along the way. I am grateful to each of my Ph.D. friends for giving warm encouragements and valuable company.

I sincerely would like to thank ICMA Centre for funding my Ph.D. studies and supporting my daily work with various perspectives.

More importantly, I would like to thank my examiners, Dr Lisa Schopohl and

Dr Junhong Yang, for giving their precious time in reviewing and commenting the thesis.

Declaration

I confirm that this thesis is my own work and the use of all material from other sources has been properly and fully acknowledged.

Siqi Liu

April 2021

Contents

Chapter 1

Introduction

This thesis studies three research questions in the mix of empirical asset pricing and corporate finance.

Prior literature documents that the standard capital asset pricing model (CAPM) beta does not describe the cross-sectional average stock returns (Eugene and French, 1992; Lewellen and Nagel, 2006). High-beta stocks realize negative abnormal returns and low-beta stocks realize positive abnormal returns, which we refer to as the low-beta anomaly, has long received much attention from academics and practitioners (Jensen et al., 1972; Baker et al., 2011). To exploit such mis-valuation from agents who face severe leverage and margin constraints, Frazzini and Pedersen (2014) construct a betting against beta portfolio that longs low-beta stocks and shorts high-beta stocks, which can yield significant positive risk-adjusted returns.

As we all know, academic researchers usually rely on the statistical significance of regression coefficients to support or reject their hypotheses. Harvey et al. (2016) indicates that the statistical power of many empirical anomalies is overlooked in the finance and economics literature. They advocate that financial economists should focus on re-evaluating the statistical significance of hundreds of factors discovered in the previous literature. Beta measures a stocks' systematic risk, is estimated as the time-series correlation between a stock's returns and the market portfolio's returns. If a beta estimate is statistically insignificant, shall financial economists and portfolio managers take the estimate as zero or its raw value?

We apply the levels of statistical significance of beta estimates on BAB portfolios to measure the impact on BAB alphas. We find that, when betting against statistically significant betas, BAB alphas can be reduced by around 20%–50% based on different beta estimations. If we replace the insignificant beta as zero, a refined betting against beta trading strategy can generate higher alphas than original BAB strategy. We seek to shed new light on the potential impact of ignoring statistical significance in empirical asset pricing studies.

Our paper builds on a literature of low-beta anomaly that longs low-beta stocks and shorts high-beta stocks, which are able to generate significant positive risk-adjusted returns. Liu et al. (2018) overturn the prior argument that beta is the founding reason resulted in a beta anomaly and reveals a combination of negative alpha-IVOL relation among overpriced stocks and positive IVOL-beta correlation generates the beta anomaly[1]. Thus, controlling for IVOL or deleting overpriced high-IVOL stocks makes the beta anomaly disappear. Humphery-Jenner and Suchard (2013) investigate the existence of low-volatility anomaly outside the U.S. market and the relationship between low-volatility stocks, operating performance and stock earnings surprise via examining international stock returns. They find the low-volatility anomaly exists in both developed markets and emerging markets, which directly induces stronger future operating performance. Bali et al. (2011, 2017) investigate the relation between the beta anomaly and demand for lottery-like stocks which push up their prices with more volatile movements, thus decreases future expected abnormal returns relatively. However, this anomaly can disappear when controlling for lottery-like investors' demand which is identified as the maximum daily return over the previous month. Asness et al. (2014) construct BAB factors and decompose those into the industry-neutral bets and pure industry bets

[1]IVOL refers to a stock i's idiosyncratic volatility measured by following Ang et al. (2006).

to determine its primary performance driving component in the low-risk investing framework and they argue the low-risk strategies without industry bets outstandingly become the winner. Apart from this, Auer and Schuhmacher (2015) use stocks in the Dow Jones Industrial Average (DJIA) to test whether high-liquid portfolios have low-beta anomaly and how this arbitrage can be exploited if it exists for unconstrained investors. Exploiting the low-beta anomaly via both pure and combined BAB trading portfolios respectively captures substantive positive abnormal returns and provides sufficient diversification and low risk, which is unable to be explained by standard asset-pricing factors. They successfully challenge the conclusion made by Li et al. (2014) and Novy-Marx (2014) that low-beta anomaly only exists in the small capitalization markets instead of large capitalization markets. Furthermore, Baker et al. (2011) find an above-average rate of returns for low-volatility and low-beta stock combinations along with small setbacks from 1968 to 2008. They find genuinely low-risk stock portfolios have persistent out-performance compared to the high-risk portfolios. And beta is a stronger driver than volatility to the heart of the anomaly in large stocks and less effective in small stocks. This research contributes to minimizing the academic publication bias in finance research for publishing on a top-tier journal and encourage researchers to rigorously examine the existing factors instead of endlessly discovering new factors.

In addition to BAB factors, I focus on examining the negative investment-return relation in the Chapter 3. A large literature documents a negative relation between firm-level capital investment and future stock returns. For example, Titman et al. (2004) examine the investment growth anomaly in portfolio levels and suggest it may result from over-investment that investors under-react to the management's empire building behaviour of increased capital expenditures. This negative relation is stronger for firms that have great investment discretion and is significant only during the time periods with less hostile takeovers. Harvey et al. (2004) demonstrate that actively monitored debt creates value for shareholders of firms that face potentially extreme agency costs associated with misaligned managerial incentives and over-investment problem. Xing (2008) finds the negative relationship between investment growth and stock returns in both cross-sectional and time-series analyses. The rational Q theory explains the negative investment-return relation by suggesting that firms tend to invest more when the cost of capital (expected return) is lower, which induces a higher net present value of new investments. Polk and Sapienza (2009) find a positive relation between abnormal investment and discretionary accruals and that firms with high abnormal investment subsequently have low stock returns. They consider that if firms are misallocating resources due to market mis-valuation, then abnormal investment should predict risk-adjusted returns. Hou et al. (2015) propose a q-factor model including the market, size, investment, and profitability factors which largely explain the cross section of average stock returns. They consider that high costs of capital imply low net present value of new projects and low investment given discount rates.

Thus, the issue of whether abnormal investment, the difference between actual and predicted investment can explain the negative correlation, is investigated in the U.S. equity market. We use the investment expectation model from Richardson (2006) to measure the abnormal investment, and divide it into under- and over-investment, respectively, to test which one truly drives the negative impact.

The results show that market investors have delayed reactions to those information carried by firm under-investment about future fundamentals, such as future profitability, asset growth and the likelihood of financial distress risk, but not to systematic distress risk. Moreover, the negative relation between under-investment and future stock returns is more pronounced for firms with lower investor monitoring and higher agency costs. The results are consistently robust to alternative regression models and alternative investment models. Thus, when both investment and abnormal investment are considered simultaneously, future stock returns tend to be more closely associated with abnormal investment. Most importantly, it is the firm under-investment instead of over-investment that primarily drives the negative relation. This research suggests that future stock returns can be increased by decreasing the possibility of firm under-investment.

In addition, I also study peer effect of corporate investment on the firms' marginal value of cash holdings in the Chapter 4. Industry competitions become extremely intensive over past decades due to advanced technological innovations and product upgrades. Previous literature mainly examines that corporate capital structures can be affected in a parallel way by their peer firms' corporate decisions. Foucault and Fresard (2014) and Bustamante and Frésard (2017) find that firm investment responds positively to their product market peer firms' investment. Adhikari and Agrawal (2018) and Grennan (2019) conclude that firms' payouts to equity investors are positively correlated to their industry peer firms' payout decisions. Chen et al. (2019) argue that managers consider the level of their peer firms' cash holdings when deciding their own cash holdings. However, whether and how this effect of peer investment on firms' value of cash continues to work has received very little attention. Given increasing peer pressure everywhere, I examine how investors' valuation on firms' cash holdings changes if average investment of their product market peers increases in the U.S. equity market.

Empirically, I run the baseline regression model developed by Faulkender and Wang (2006a) and obtain a positive relation between peer firm average investment and the firms' value of cash holdings. To mitigate the endogeneity concerns, I use two instrumental variables (IVs) implemented in the two-stage least squares analyses. The IV results show highly statistically significant coefficient estimates so that both IVs are well-selected. The positive relation also remains robust in both the high-dimensional fixed effects and placebo tests.

Conducting several sub-sample analysis, we find this effect is likely to result from the positive externalities brought by peer investment from growing, young and competitive industries which contain sufficient investment options and are expected to fulfill capacity expansion. Since increasing aggregate investment of the whole industry could lower the price of key production factors, the generated externalities can benefit the firms that didn't make the investment (i.e., externalities of peer investment). In this type of industries, peer investment could increase the value of cash holding of the focal firms through the positive externality channel. We find little evidence to support two alternative hypotheses: cash holdings become more valuable because firms need to use it as a precautionary measure for competition escalation, and investors are learning the positive signal about the investment opportunities. Our further analyses confirm that firms on average increase their level of cash holdings, responding to their peer investment. In particular, firms are less willing to use cash for dividend payments while they are more willing to use it for capital and R&D investment. This research contributes to the growing literature of peer effect by highlighting a broader implication that is one set of peers' decisions may provide information for a different set of decisions of focal firms. We also demonstrate the value of cash as the reflection of pursuing future growth opportunities, which complements the current literature of regarding the increased value of cash holdings as a result of preserving cash for "war chest".

The reminder of the thesis is organized as follows. Chapter 2 presents the first paper which is titled as "Betting against significant beta". Chapter 3 presents the second paper which is named as "Abnormal Investment and Firm Performance". Chapter 4 documents the third paper which is titled as "Investment of product market peers and the value of cash holdings". Last, I conclude all those main findings comprehensively in the Chapter 5.

Chapter 2

Betting against significant beta

2.1. Introduction

In a perfect capital market, classical finance models and theories can often be far from enough to capture the full spectrum of rational behaviors in reality (Merton, 1987). This has provoked a strong interest of researchers in studying the refinement from a complex sphere of financial economics models. According to the capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965), stocks with larger market betas are expected to generate higher excess returns. However, Jensen et al. (1972) find that the realized abnormal returns of a portfolio with high-beta stocks are negative, whereas the realized abnormal returns of a portfolio with low-beta stocks are positive. Frazzini and Pedersen (2014) formally define a betting against beta (BAB) trading strategy, taking a long position on low-beta stocks and a short position on high-beta stocks, which can generate positive risk-adjusted returns. Frazzini and Pedersen (2014) show that many investors in the financial market are subject to funding constraints or are lawfully restricted by certain leverage and margin requirements, so that they over-invest on risky stocks with high betas to achieve their desired risk and expected return trade-off. Therefore, high-beta stocks are relatively over-valued and generate worse future returns than low-beta stocks.

Academic researchers usually rely on the statistical significance of regression coefficients to support or reject their null hypotheses. Statistical significance refers to the claim that a result from data generated by testing or experimentation is not likely to occur randomly or by chance, but is instead likely to be attributable to a specific cause. Harvey et al. (2016) indicate that the statistical significance of many empirical anomalies is overlooked in the finance and economics literature. They advocate that financial economists should focus on re-evaluating the statistical significance of hundreds of factors discovered in the previous literature. Beta, which measures a stock's systematic risk, is estimated as the time-series correlation between the stock returns and the stock market returns. If a beta estimate is statistically insignificant, should financial economists and portfolio managers take the estimate as zero or its raw value? Like many previous asset pricing studies, Frazzini and Pedersen (2014) construct beta-sorted portfolios based on stocks' raw beta estimates, ignoring the statistical significance of betas. In this chapter, we adopt Frazzini and Pedersen's (2014) BAB trading strategy as our empirical framework to study the impact of a beta estimate's statistical significance on empirical asset pricing and portfolio management. The BAB trading strategy is an arbitrage portfolio constructed upon the ranks of the beta estimates, which provides us with an ideal empirical setting to mitigate the potential confounding effect of multiple coefficient estimates that might be associated with both the exposure of interest and the excess returns in other more complicated trading strategies.

We employ two beta estimation methods in our empirical analyses. First, we follow Frazzini and Pedersen (2014) and estimate their ex-ante formula beta, using a sample of US public firms during 1926–2017. Second, we adopt ordinary least squares (OLS) regressions and estimate stock betas by regressing stock returns on the excess returns of the market portfolio. Based on t-statistics, we calculate the statistical significance of beta estimates and separate stocks with statistically significant betas from those with statistically insignificant betas. We define stocks with statistically significant (insignificant) betas at the 10%, 5%, or 1% levels if their corresponding absolute values of t-statistics are greater (lower) than the threshold of 1.65, 1.96, or 2.58, respectively. Next, we replicate Frazzini and Pedersen's (2014) main findings by sorting stocks into decile portfolios based on their beta estimates, without considering the statistical significance. Our replication results are consistent with Frazzini and Pedersen (2014) that the alphas of beta-sorted decile portfolios decrease monotonically from the lowest-beta portfolio to the highest-beta portfolio. The alpha of the BAB trading strategy, taking a long position on stocks with below the sample median beta and a short position on stocks with above the sample median beta, is positive and statistically significant.

In our sample, we find that 9%–40% of stocks have statistically insignificant formula beta at the 1% level, depending on the stock return frequencies and estimation windows. Comparing to stocks with statistically significant betas, stocks with statistically insignificant betas are more likely to be lottery-like stocks, smaller in firm size and market value, less likely to be value stocks, and with higher idiosyncratic volatilities. After dropping stocks with statistically insignificant betas at the 10%, 5% or 1% levels, we construct the betting against significant beta portfolios. Although the betting against significant beta portfolios still generate positive alphas, both the magnitude and the statistical significance of these alphas are lower than the corresponding BAB portfolios. The decreases in the alphas of the BAB portfolios range from 20% to 50%, depending on beta estimation methods, the statistical significance of beta estimates, and alpha estimation methods. We then construct a refined BAB strategy by replacing statistically insignificant betas with zeros.

The refined BAB trading strategy is based on updated stock betas instead of raw beta estimates. We find that the refined BAB strategy generates a higher alpha than the original BAB strategy. Liu et al. (2018) argue that since idiosyncratic volatility (IVOL) is positively related to betas and negatively related to alphas, it is IVOL rather than betas that drives the beta anomaly. According to Liu et al. (2018), the beta anomaly no longer exists after controlling for IVOL. We replicate Liu et al.'s (2018) findings and replace statistically insignificant betas by zeros. We find that based on the updated stock betas, the beta anomaly still exists after controlling for IVOL. Our main findings are robust to different beta estimation methods, various stock return frequencies, beta estimation windows, and factor pricing models. Our main results also remain robust across several major international equity markets: the UK, Canada, Belgium, Hong Kong, Singapore, Switzerland, and Netherlands.

Our paper contributes to the recent strand of studies on the beta anomaly. For example, Liu et al. (2018) show that the beta anomaly is much weaker after controlling for IVOL or deleting overpriced high-IVOL stocks. Bali et al. (2017) investigate the relation between the beta anomaly and demand for lottery-like stocks and find that the beta anomaly can be partially explained by the stocks' lottery features. To the best of our knowledge, we are the first to examine the potential impact of the beta's statistical significance on the beta anomaly. Our paper sheds light on the issues of ignoring portfolio sorting criteria's statistical significance in empirical asset pricing studies. We not only test the impact of the beta's statistical significance on the beta anomaly, but also provide investors with a refined BAB strategy that generates a better performance.

Our paper is organized as follow. Section 3.2 describes our sample, variable definition, and empirical methodologies. Section 2.3 reports our main results, and section 2.4 provides the robustness test results. Section 2.5 concludes.

2.2. Sample, variables, and summary statistics

2.2.1. Sample data

Our US sample covers observations with available daily stock return data in the Centre for Research in Security Prices (CRSP) from January 1926 to December

2017. Our US sample is composed of all common stocks actively traded on the New York Stock Exchange (NYSE), the American Stock Exchange (AMEX), and the National Association of Securities Dealers Automated Quotations (NASDAQ). Our international sample covers seven financial markets: the UK, Canada, Netherlands, Belgium, Switzerland, Hong Kong, and Singapore. We collect the data on the daily stock prices of these seven markets from the Compustat Global database. Due to data availability, the sample period for Canada is from January 1984 to December 2017, while the sample periods for the other six financial markets are from January 1985 to December 2017. We drop the observations with missing values of daily close prices. Our US sample includes 23,479 unique stocks, and our international sample includes 5,558 (the UK), 7,068 (Canada), 1,021 (Singapore), 409 (Belgium), 393 (Netherlands), 570 (Switzerland), and 381 (Hong Kong) unique stocks.

We collect the data on US stock risk factors from Wharton Research Data Services (WRDS), including Fama and French's (1993) three factors (excess market return: *MKTRF*; small minus big: *SMB*; and high minus low: *HML*), Carhart's (1997) momentum factor (*UMD*), Pástor and Stambaugh's (2003) liquidity factor (*LIQ*), and Fama and French's (2015) investment and profitability factors (*CMA* and *RMW*). For our seven international markets, the data on the corresponding MSCI local market indexes are collected from Datastream. We also collect the data on risk factors for our international sample, including *MKTRF*, *SMB*, *HML*, *CMA*, and *RMW*, from Kenneth French's website.[1]

2.2.2. Beta measures and their statistical significance

To estimate an individual stock's beta, we first follow the methodology developed by Frazzini and Pedersen (2014). Specifically, we estimate the one-year rolling standard deviations of stock returns (σ_i) and market returns (σ_m) using daily log

[1]The data can be downloaded from http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Developed.

returns. In the estimation of standard deviations, we require that a stock has non-missing daily returns for more than 120 trading days over a one-year rolling window. We then estimate the five-year rolling correlations (ρ_i) between stock i and the corresponding market index using log returns over three trading days. In the estimation of the correlations, we require that a stock has non-missing stock returns for more than 750 trading days over a five-year rolling window, in order to mitigate the non-synchronous trading concern. Following Frazzini and Pedersen (2014), the estimation windows of correlations are longer than the estimation windows of standard deviations, because the correlation between an individual stock and the corresponding market index changes relatively more slowly than the individual stock's standard deviation does. Our ex-ante time-series beta for stock i (β_i^{TS}) is estimated by the ratio of stock volatility (σ_i) to market volatility (σ_m), multiplied by the correlation (ρ_i) between stock i's returns and the corresponding market index's returns:

$$\beta_i^{TS} = \rho_i \frac{\sigma_i}{\sigma_m} \tag{2.1}$$

Following Frazzini and Pedersen (2014), we shrink the time-series beta (β_i^{TS}) toward its cross-sectional average (β_i^{XS}=1) and apply the weight (w_i) of 0.6 for all periods across all stocks[2], in order to mitigate the outlier effect. The beta estimated by this method is defined as $\beta^{Formula}$ in our paper hereafter:

$$\beta_i^{Formula} = w_i \beta_i^{TS} + (1 - w_i) * \beta_i^{XS} \tag{2.2}$$

[2]I follow Frazzini and Pedersen (2014) and Vasicek (1973) in which Vasicek (1973) constructs the Bayesian shrinkage factor as w_i=$1 - \sigma^2_{i,TS}/(\sigma^2_{i,TS,} + \sigma^2_{XS})$, where $\sigma^2_{i,TS}$ is the variance of the time-series estimated beta for stock i and σ^2_{XS} is the variance of cross-sectional estimated betas. Either when the time-series estimate of beta has a lower variance or when a large dispersion of cross-sectional beta estimates exist, the weight put on the time-series estimate of beta for stock i is higher. The weight across all US stocks has a mean value of 0.61. That's why we choose it as the shrinkage factor.

As a robustness check, we also calculate β_i^{TS} using the standard deviations and the correlations estimated by monthly returns over a five-year rolling window. We then define $\beta_i^{Formula,5YM}$ using Equation (2.2).

We define the statistical significance of $\beta_i^{Formula}$ according to Walpole et al. (2002):

$$t_{\beta^{Formula}} = \frac{\rho}{\sqrt{\frac{1-\rho^2}{N-2}}} \tag{2.3}$$

where N is the number of observations used to estimate the correlation coefficient ρ. If the absolute value of the t-statistic of correlation coefficient ρ is less than the threshold of 1.65, 1.96, or 2.58, then the corresponding $\beta_i^{Formula}$ is statistically insignificant at the 10%, 5%, or 1% levels, respectively.

An alternative method to estimate a stock beta and its statistical significance is the regression analysis. We adopt the capital asset pricing model (CAPM) and estimate β_i^{CAPM} in the following regression equation:

$$R_i - R_f = \alpha_i + \beta_i^{CAPM}(R_m - R_f) + \epsilon_i \tag{2.4}$$

where R_i is stock return, R_f is risk free rate, and R_m is market index return. The statistical significance of β_i^{CAPM} is based on the t-statistics:

$$t_{\beta^{CAPM}} = \frac{\hat{\beta} - 0}{SE(\hat{\beta})} \tag{2.5}$$

where $SE(\hat{\beta})$ is the standard error of the estimator in the CAPM regression. If the absolute value of $t_{\beta^{CAPM}}$ is less than the threshold of 1.65, 1.96, or 2.58, then the corresponding β_i^{CAPM} is statistically insignificant at the 10%, 5%, or 1% levels, respectively.

We estimate β_i^{CAPM} in Equation (2.4) using a one-year rolling window with daily returns and a five-year rolling window with both daily and monthly returns. To be consistent with $\beta_i^{Formula}$, we also shrink β_i toward its cross-sectional average

(β_i^{XS}=1):

$$\beta_i^{OLS} = w_i \beta_i^{CAPM} + (1 - w_i) * \beta_i^{XS} \tag{2.6}$$

Depending on the return frequency and the length of the estimation window, we define $\beta^{OLS,1YD}$ (one-year rolling window with daily returns), $\beta^{OLS,5YD}$ (five-year rolling window with daily returns), and $\beta^{OLS,5YM}$ (five-year rolling window with monthly returns).

2.2.3. Beta-sorted portfolios and summary statistics

We follow Frazzini and Pedersen (2014) to form the original ten beta-sorted portfolios. At the beginning of every calendar month, we sort all US stocks into ten decile portfolios, on the basis of their estimated betas at the end of the previous month. We employ five beta estimates: $\beta^{Formula}$ (Frazzini and Pedersen's (2014) formula beta), $\beta^{Formula,5YM}$ (Frazzini and Pedersen's (2014) formula beta estimated by monthly returns over a five-year rolling window), $\beta^{OLS,1YD}$, $\beta^{OLS,5YD}$, and $\beta^{OLS,5YM}$. All ranked stocks are assigned to one of the ten decile portfolios. NYSE stock beta breakpoints are adopted in constructing beta-sorted decile portfolios. Next, we examine the distribution of stocks with statistically insignificant betas across these ten decile portfolios. Panel A of Table 2.1 reports the time-series average number of stocks and the time-series average number of stocks with statistically insignificant betas at the 10%, 5%, and 1% levels, for each of the ten decile portfolios. The last column of Panel A reports the percentage of stocks with statistically insignificant betas in the corresponding portfolio.

As shown in Panel A, the numbers of stocks in the ten beta-sorted portfolios are not evenly distributed because stocks are assigned into these portfolios based on NYSE breakpoints. The stocks with statistically insignificant betas mostly concentrate in the portfolios with low beta stocks. A relatively lower proportion of stocks with statistically insignificant betas are in decile portfolios

2–10, while most stocks with statistically insignificant betas are in decile portfolio 1. For stocks sorted by $\beta^{Formula}$, the stocks with statistically insignificant betas at the 10%, 5%, and 1% levels, account for 5.56% ($= 167/3,002$), 6.76% ($= 203/3,002$), and 9.29% ($= 279/3,002$) of all sample stocks, respectively. For the other four beta estimates, we observe more stocks with statistically insignificant betas. At the 1% level, 31.40% ($= 945/3,010$), 39.50% ($= 1,159/2,934$), 16.37% ($= 371/2,267$), and 31.99% ($= 963/3,010$) of stocks have statistically insignificant betas for $\beta^{Formula,5YM}$, $\beta^{OLS,1YD}$, $\beta^{OLS,5YD}$, and $\beta^{OLS,5YM}$, respectively. For various beta estimation methods, we observe a similar distribution of stocks with statistically insignificant betas across ten beta-sorted portfolios. We also observe more stocks with statistically insignificant betas when we expand the beta estimation time window from one year to five years and change the return data frequency from daily to monthly. Our results suggest that betas and their statistical significance depend on the estimation method and return frequency.

Panel B of Table 2.1 reports the time-series average of estimated stock betas in each of the ten beta-sorted portfolios. For each beta estimate, we first calculate the mean value of stock betas in a portfolio at a given month, and then calculate the time-series average of the means. We also report the time-series average of the means of stock betas which are statistically insignificant at the 10%, 5%, and 1% levels if the absolute value of t-statistic of correlation coefficient ρ or β_i^{CAPM} estimate is less than the common thresholds of 1.65, 1.96, or 2.58. Within each beta-sorted portfolio, the mean of statistically insignificant betas is less than the mean of all stock betas in the portfolio. The means of statistically insignificant betas slightly increase from the 10% significance level to the 1% significance level.

Panel C of Table 2.1 reports the time-series average of US firm characteristics in each of the ten decile portfolios sorted by $\beta^{Formula}$. For each portfolio, we decompose stocks into those with statistically significant betas and those with statistically insignificant betas at the 1% level if the absolute value of t-statistic of correlation coefficient ρ is less than the threshold of 2.58. The firm characteristics are lottery demand (*Max*), market capitalization (*Mcap*), book-to-market ratio (*Btm*), logarithm of total assets (*Asset*), and idiosyncratic volatility (*Ivol*). *Max* is the average of a stock's five highest daily returns over a month. Bali et al. (2011) find that a stock's maximum daily returns over a month are negatively associated with its future returns. They also find that the negative relation between idiosyncratic volatility and returns documented by Ang et al. (2006) is often reversed after controlling for maximum daily returns. Given that large maximum daily stock returns are like lottery payoffs, Bali et al. (2011) conjecture that the negative relation between maximum daily returns and future returns is due to investors' preference for lottery stocks. *Ivol* is the standard deviation of the residuals from a regression of excess stock returns on Fama and French's (1993) three factors within a one-month period. *Mcap* and *Asset* are in millions. We calculate the mean value of firm characteristics for each portfolio at a given month, and then calculate the time-series average of the means. Panel C shows that stocks with statistically significant betas are less likely to be lottery-like stocks, larger in firm size and market value, more likely to be value stocks, and with a lower idiosyncratic volatility than stocks with statistically insignificant betas.

2.3. Main results

2.3.1. Betting against beta vs. Betting against significant beta

In order to examine how stocks with statistically insignificant betas may affect the BAB trading strategy, we first follow Frazzini and Pedersen (2014) and Horenstein (2017) to construct BAB portfolios. At the beginning of each calendar month, we assign US stocks into low-beta and high-beta portfolios, according to the medians of stock betas estimated in the previous month. The low-beta and high-beta portfolios are rebalanced every month. The weights of stocks in these two portfolios are proportional to the ranked betas, that is, a stock with a higher (lower) beta obtains a larger weight in the high-beta (low-beta) portfolio. Specifically, stock i's rank (zl_i) in the low-beta portfolio is equal to $rank(\beta_i)$ at portfolio formation, and its weight (wl_i) in the low-beta portfolio is equal to $\frac{(nl-zl_i+1)}{\sum_{i=1}^{i=nl} zl_i}$, where nl is the number of stocks in the low-beta portfolio. Similarly, stock i's rank (zh_i) in the high-beta portfolio is equal to $rank(\beta_i)$, and its weight (wh_i) in the high-beta portfolio is equal to $\frac{zh_i}{\sum_{i=1}^{i=nh} zh_i}$, where nh is the number of stocks in the high-beta portfolio. The sums of stock weights in the low-beta and high-beta portfolios are equal to one. The weighted stock betas for low- and high-beta portfolios are calculated, respectively as $\beta^L = \sum_{i=1}^{i=nl} wl_i\beta_i$ and $\beta^H = \sum_{i=1}^{i=nh} wh_i\beta_i$. The returns of the low- and high-beta portfolios are constructed accordingly as $r^L = \sum_{i=1}^{i=nl} r_i wl_i$ and $r^H = \sum_{i=1}^{i=nh} r_i wh_i$. Frazzini and Pedersen (2014) define the BAB portfolio as a self-financing zero-beta portfolio that takes a long position on low-beta stocks and a short position on high-beta stocks. The returns of the BAB portfolio are calculated as:

$$r_{t+1}^{BAB} = \frac{1}{\beta_t^L}(r_{t+1}^L - r^f) - \frac{1}{\beta_t^H}(r_{t+1}^H - r^f) \tag{2.7}$$

Next, we delete stocks with statistically insignificant betas and reconstruct the ten beta-sorted portfolios and the BAB portfolio using stocks with statistically significant betas. We choose 10%, 5%, and 1% statistical significance levels with respect to the thresholds of the absolute values of t-statistics being 1.65, 1.96 and 2.58, respectively. In Table 2.2, we adopt Frazzini and Pedersen's (2014) $\beta^{Formula}$ as our beta estimate and compare the performance of BAB portfolios to the performance of betting against significant beta portfolios. For ten beta-sorted portfolios, we follow Frazzini and Pedersen (2014) and calculate their equal-weighted monthly returns. Our portfolio performance measures are monthly returns in excess of the risk free rate, alphas of CAPM, alphas of Fama and French (1993) (FF) three-factor model, alphas of FF three-factor model augmented by Carhart's (1997) momentum factor and Pástor and Stambaugh's (2003) liquidity factor (FF-Carhart-PS) five-factor model, and alphas of Fama and French's (2015) (FF) five-factor model.

Table 2.2 shows that the abnormal returns of the well-documented BAB strategy hold over our sample period of 1926–2017. Consistent with Frazzini and Pedersen (2014), all the performance measures drop from decile portfolio 1 to decile portfolio 10. In addition, all the performance measures of the BAB portfolios are positive and statistically significant. For example, the excess return (0.74%) and CAPM alpha (0.62%) of the BAB portfolio are statistically significant at the original level.[3] After we drop the stocks with statistically insignificant betas at the 10%, 5%, and 1% levels if the absolute value of t-statistic of correlation coefficient ρ or β_i^{CAPM} estimate is less than the common thresholds of 1.65, 1.96, or 2.58, both the performance of the BAB portfolios and the statistical significance of the performance decrease consistently. The monthly excess return of the BAB portfolio drops from 0.74% (original) to 0.64% (10% Sig.), to 0.62% (5% Sig.), and to 0.59% (1% Sig.). The difference in the annualized excess returns between the BAB portfolio and the betting against 1% significant beta portfolio is 1.8%. The CAPM alphas of the BAB portfolios drop from 0.62% (original) to 0.53% (10% Sig.), to 0.51% (5% Sig.), and to 0.49% (1% Sig.). The difference in the CAPM alphas between the BAB portfolio and the betting against 1% significant beta portfolio is 1.6%. We observe a similar pattern of portfolio performance for the FF three-factor alphas, FF-Carhart-PS five-factor alphas, and FF five-factor alphas. The FF five-factor alpha of the BAB portfolio is 3.5% and statistically significant at the 1% level, while the FF five-factor alpha of the betting against 1% significant beta portfolio is 0.10% and statistically insignificant. Our findings suggest that the positive abnormal returns of the BAB portfolio is weaker when we drop stocks with statistically insignificant betas.

[3]We replicate the returns of the BAB portfolio over the period of January 1926–March 2012, exactly the same as Frazzini and Pedersen (2014). Untabulated results show that the performance of the BAB portfolio is comparable with Frazzini and Pedersen's (2014) original results.

2.3.2. Betting against beta vs. Betting against significant beta: Alternative beta estimates

Beta estimates and their statistical significance vary with respect to beta estimation methods and return data frequency. In this section, we examine whether our findings in Section 2.3.1 hold for alternative beta estimates. We sort stocks into ten beta-sorted portfolios and construct BAB portfolios using $\beta^{Formula,5YM}$, $\beta^{OLS,1YD}$, $\beta^{OLS,5YD}$, and $\beta^{OLS,5YM}$. For brevity, we choose the FF-Carhart-PS five-factor alpha and FF five-factor alpha as our performance measures. We also only compare the performance of original beta-sorted portfolios to the performance of 1%-significant-beta-sorted portfolios. Table 2.3 reports the results.

For our four beta estimates, the low decile beta-sorted portfolios still outperform the high decile beta-sorted portfolios. However, the declining pattern of decile portfolio performance is not persistent for decile portfolios with only 1% significant beta stocks. The performance of original BAB portfolios remains positive and statistically significant. But after dropping stocks with 1% insignificant betas if the absolute value of t-statistic of correlation coefficient ρ or β_i^{CAPM} estimate is less than the threshold of 2.58, the performance of betting against significant beta portfolios are all statistically insignificant, except for the FF-Carhart-PS five-factor alphas of $\beta^{OLS,1YD}$ and $\beta^{OLS,5YD}$ estimates. Thus, the positive abnormal returns of the BAB portfolios do not remain robust when we drop stocks with statistically insignificant betas and sort stocks by alternative beta estimates.

2.3.3. A refined betting against beta trading strategy

Given that researchers usually treat statistically insignificant coefficients as zeros, in this section we propose a refined BAB strategy and compare it with the original BAB strategy. Instead of dropping stocks with statistically insignificant betas from our portfolios, we first replace statistically insignificant betas by zeros, and then construct beta-sorted portfolios and BAB portfolios based on the updated stock betas.

Table 2.4 presents the monthly alphas of the original and refined beta-sorted portfolios estimated by the Fama and French (2015) five-factor model. The last column shows the monthly alphas of the original and refined BAB portfolios. After replacing statistically insignificant betas by zeros, we observe that the FF five-factor alphas of ten beta-sorted portfolios consistently decrease from decile portfolio 1 to decile portfolio 10. More importantly, the FF five-factor alphas of the refined BAB portfolios are larger than the corresponding original BAB portfolios, except for $\beta^{OLS,1YD}$. For example, sorting stocks by Frazzini and Pedersen's (2014) beta ($\beta^{Formula}$), the FF five-factor alpha of the original BAB portfolio is 0.35%, while the FF five-factor alpha of the refined BAB portfolio is 0.42%. The annualized difference in the performance between the original and refined BAB strategy is equal to 0.84% $(= (0.42\% - 0.35\%) * 12)$. The improvement is even more economically significant for stocks sorted by $\beta^{Formula,5YM}$, $\beta^{OLS,5YD}$, and $\beta^{OLS,5YM}$.

2.4. Robustness check and further discussions

2.4.1. Beta and idiosyncratic volatility

Liu et al. (2018) argue that since there exists a cross-sectionally positive relation between idiosyncratic volatility and betas and a negative relation between idiosyncratic volatility and alphas only among overpriced stocks, we usually observe a beta anomaly for overpriced stocks. They also argue that idiosyncratic volatility is one of the underlying factors leading to the beta anomaly. They find that the negative relation between idiosyncratic volatility and alphas is robust after controlling for betas, but there is little evidence of a beta anomaly when controlling for idiosyncratic volatility.

At first, we replicate Liu et al.'s (2018)'s result in our sample. At the beginning of each calendar month, we sort US stocks based on Liu et al.'s (2018) beta and Ang et al.'s (2006) idiosyncratic volatility ($IVOL$), estimated at the end of the previous month, into 50 (5 by 10) value-weighted portfolios. Specifically, following Liu et al. (2018), we first regress a stock's monthly return on the contemporaneous market return and one-month lagged market return, over a five-year rolling window for each month from January 1926 to December 2017. The stock's beta is the shrinked summation of the two regression coefficients (Vasicek, 1973; Dimson, 1979), as described in Equation (2.2). We follow Ang et al. (2006) to estimate $IVOL$, which is defined as the standard deviation of the residuals from regressing the daily stock returns on Fama and French (1993) three factors over the previous month. At the beginning of every month, we sort stocks into ten decile portfolios based on the beta estimated at the end of the previous month. Then for each beta-sorted decile portfolio, we further sort stocks into five quintile portfolios based on the $IVOL$ estimated at the end of the previous month. For each of these 50 portfolios, we estimate its Fama and French (1993) three-factor alpha based on monthly portfolio returns.

Panel A of Table 2.5 reports the results. The column labeled "Long 1, Short 10" presents the FF three-factor model alpha of a portfolio that takes a long position on stocks with lowest betas (portfolio 1) and a short position on stocks with the highest betas (portfolio 10), within the corresponding $IVOL$ quintile. The row labeled "Long 1, Short 5" reports the FF three-factor model alpha of a portfolio that takes a long position on stocks with the lowest IVOL (portfolio 1) and a short position on stocks with the highest IVOL (portfolio 5), within the corresponding beta decile. The last column reports the averages of the FF three-factor model alphas of the ten beta-sorted portfolios, within the corresponding $IVOL$ quintile. The last row reports the averages of the FF three-factor model alphas of the five IVOL-sorted portfolios, within the corresponding *IVOL* quintile. The alphas of all five “Long 1, Short 10” portfolios are statistically insignificant at the 10% level. Moreover, the average alpha of five“Long 1, Short 10” portfolios across *IVOL* quintiles is 0.18% and statistically insignificant at the 10 % level. Our result is consistent with the findings in Table 5 of Liu et al. (2018) that there is little evidence of a beta anomaly when controlling for stock idiosyncratic volatility.

Next, we replace stock betas which are statistically insignificant at the 1% level by zeros, and repeat our analyses in Panel A of Table 2.5. Panel B of Table 2.5 shows that two alphas of all five “Long 1, Short 10” portfolios are statistically significant. In the second row, the alpha of the “Long 1, Short 10” portfolio is 0.19% with a t-statistic of 1.67. In the fifth row, the alpha of the “Long 1, Short 10” portfolio is 0.91% ($= 0.51\% - (-0.40\%)$) with a t-statistic of 2.12. Most of all, the average alpha of five “Long 1, Short 10” portfolios across *IVOL* quintiles is 0.25% and statistically significant at the 5% level. Our results suggest that after replacing statistically insignificant betas by zeros, the beta anomaly still exists when controlling for stock idiosyncratic volatility.

2.4.2. International evidence

So far, our empirical tests are all based on US stocks. Frazzini and Pedersen (2014) test the abnormal returns of BAB portfolios across 19 MSCI developed countries. The excess returns of the BAB portfolios are positive and statistically significant for 10 out of 19 financial markets. In addition, the alphas of Fama and French (1993) three-factor model augmented by Carhart’s (1997) momentum factor is positive and statistically significant for 6 out 19 financial markets. Frazzini and Pedersen (2014) argue that the small number of stock observations intensifies the difficulty of rejecting the null hypothesis of zero abnormal returns in the international financial markets. In this section, we choose the six countries with positive and statistically significant four-factor alphas: Canada, Netherlands, Belgium, Switzerland, Hong Kong, and Singapore. We also add the UK in our international sample.

Table 2.6 reports the Fama and French (2015) five-factor alphas of original BAB portfolios and betting against significant beta portfolios. We construct two types of betting against significant beta portfolios. First, we drop stocks with statistically insignificant betas at the 1% level. Second, we replace betas that are statistically insignificant at the 1% level by zeros, and then sort stocks with the updated betas. For our five beta estimates, Table 2.6 shows that most of the original BAB portfolios generate positive and statistically significant alphas. Using Frazzini and Pedersen's (2014) beta estimate $\beta^{Formula}$ as an example, our results are consistent with Frazzini and Pedersen (2014) that the UK is the only financial market in which the original BAB portfolio does not generate a statistically significant beta at the 10% level. When we construct betting against significant beta portfolios using only stocks with statistically significant betas, most alphas of the betting against significant beta portfolios decrease in terms of both value and statistical significance. However, when we replace statistically insignificant betas by zeros and construct BAB portfolios using the updated stock betas, most alphas of the BAB portfolios increase in terms of both value and statistical significance. Using the UK market as an example, four alphas out of the five betting against significant beta portfolios are positive and statistically significant at the 10% level, except for $\beta^{OLS,1YD}$. Our results indicate that our empirical evidence based on the US sample can be extended to the international financial markets. After replacing statistically insignificant betas by zeros, the BAB portfolios can generate higher positive abnormal returns in the international financial markets.

2.4.3. Betting against beta with higher statistical significance

Harvey et al. (2016) argue that the conventional statistical significance cutoffs (1.65, 1.96, or 2.58) are inadequate to establish significance in current asset pricing studies. They propose a higher statistical significance threshold of t-statistics at the 3.00 level. In our untabulated results, we replicate our main tests using 3.00 as the statistical significance threshold. Comparing to the results reported in Tables 2.2 and 2.3, the alphas of the BAB portfolios decrease in terms of value and statistical significance after dropping stocks with betas' t-statistics less than 3.00. Comparing to the results reported in Table 2.4, the alphas of the BAB portfolios increase in terms of value and statistical significance after replacing betas with t-statistics less than 3.00 by zeros. Our findings further demonstrate the importance of beta significance on the performance of BAB portfolios.

2.5. Conclusion

In this study, we examine the impact of the beta's statistical significance on the performance of the BAB trading strategy. We find that the performance of the BAB trading strategy is weakened after dropping stocks with statistically insignificant betas. However, when replacing statistically insignificant betas by zeros, our refined BAB trading strategy can generate a higher alpha than the corresponding BAB trading strategy. In empirical asset pricing studies, a standard procedure is to sort stocks by a regression coefficient without considering its statistical significance. The findings in our paper highlight the concern on this standard procedure. Our paper also has an important implication for practitioners on their portfolio management when they trade based on a well-documented stock return anomaly. Future research could comprehensively investigate the sorting variables' statistical significance in the previously documented stock return anomalies.

Table 2.1. Summary statistics of ten beta-sorted portfolios

Panel A. Time-series average numbers of stocks. This panel reports the time-series average of the numbers of stocks and the numbers of stocks with statistically insignificant betas in ten beta-sorted decile portfolios. At the beginning of each calendar month, US stocks are ranked in an ascending order. The stocks are assigned in ten decile portfolios, based on their estimated betas at the end of the previous month. NYSE stock beta breakpoints are adopted in constructing beta-sorted portfolios. We estimate stock betas using five methods: $\beta^{Formula}$, $\beta^{Formula,5YM}$, $\beta^{OLS,1YD}$, $\beta^{OLS,5YD}$, and $\beta^{OLS,5YM}$. The detailed estimation methods are discussed in Section 2.2.2. Our sample period is from January 1926 to December 2017. Portfolio 1 (10) includes stocks with the lowest (highest) betas. For portfolios 1–10, we report the time-series average numbers of stocks and the time-series average numbers of stocks with statistically insignificant betas at the 10%, 5%, and 1% levels, for ten beta-sorted portfolios. In the last two columns, we report the time-series average of the total number of stocks and the percentage of stocks with statistically insignificant betas in the corresponding beta-sorted portfolios.

Portfolio	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Total**	**Pct.**
$\boldsymbol{\beta^{Formula}}$												
Original	588	296	259	240	222	221	224	238	274	438	3,002	
10% Insig.	141	7	3	2	1	1	1	1	2	7	167	5.56%
5% Insig.	167	11	5	3	2	2	2	2	3	8	203	6.76%
1% Insig.	214	21	10	6	4	3	3	3	4	10	279	9.29%
$\boldsymbol{\beta^{Formula,5YM}}$												
Original	442	293	255	237	229	231	240	260	302	522	3,010	
10% Insig.	280	90	50	31	22	17	14	12	10	10	535	17.77%
5% Insig.	305	113	68	44	33	26	23	20	18	18	667	22.16%
1% Insig.	341	152	102	73	58	49	45	42	41	44	945	31.40%
$\boldsymbol{\beta^{OLS,1YD}}$												
Original	744	296	255	233	220	213	211	213	224	326	2,934	
10% Insig.	551	97	51	32	21	15	11	8	7	7	799	27.23%
5% Insig.	595	120	67	43	30	22	17	13	11	10	927	31.60%
1% Insig.	653	160	98	67	49	37	29	24	21	20	1,159	39.50%
$\boldsymbol{\beta^{OLS,5YD}}$												
Original	575	237	207	189	175	169	164	165	171	215	2,267	
10% Insig.	208	11	5	2	1	1	1	0	1	0	230	10.15%
5% Insig.	242	17	8	4	2	1	1	0	1	1	277	12.22%
1% Insig.	304	31	16	8	4	3	2	1	2	0	371	16.37%
$\boldsymbol{\beta^{OLS,5YM}}$												
Original	457	293	255	236	228	229	241	259	303	510	3,010	
10% Insig.	293	91	50	31	22	17	14	11	10	10	548	18.21%
5% Insig.	319	113	68	44	33	27	23	20	19	18	683	22.69%
1% Insig.	357	152	102	73	58	49	45	42	42	43	963	31.99%

Panel B. Time-series average of stock betas. This panel reports the time-series average of estimated stock betas in ten beta-sorted portfolios. At the beginning of each calendar month, US stocks are ranked in an ascending order. The stocks are assigned in ten decile portfolios, based on their estimated betas at the end of the previous month. NYSE stock beta breakpoints are adopted in constructing beta-sorted portfolios. We estimate stock betas using five methods: $\beta^{Formula}$, $\beta^{Formula,5YM}$, $\beta^{OLS,1YD}$, $\beta^{OLS,5YD}$, and $\beta^{OLS,5YM}$. The detailed estimation methods are discussed in Section 2.2.2. Our sample period is from January 1926 to December 2017. Portfolio 1 (10) includes stocks with the lowest (highest) betas. For portfolios 1–10, we report the time-series average numbers of stocks and the time-series average numbers of stocks with statistically insignificant betas at the 10%, 5%, and 1% levels, for ten beta-sorted portfolios. We calculate the mean of stock betas in a portfolio at a given month, and report the time-series average of the means.

Portfolio	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
$\boldsymbol{\beta^{Formula}}$										
Original	0.62	0.80	0.90	0.99	1.07	1.15	1.25	1.37	1.53	2.04
10% Insig.	0.44	0.70	0.79	0.84	0.87	0.94	1.03	1.11	1.25	1.93
5% Insig.	0.45	0.71	0.80	0.86	0.89	0.95	1.04	1.11	1.26	1.93
1% Insig.	0.48	0.73	0.81	0.87	0.91	0.98	1.05	1.12	1.26	1.97
$\boldsymbol{\beta^{Formula,5YM}}$										
Original	0.59	0.80	0.91	1.00	1.08	1.16	1.25	1.36	1.51	1.97
10% Insig.	0.50	0.75	0.83	0.92	0.97	1.03	1.11	1.18	1.26	1.64
5% Insig.	0.52	0.76	0.85	0.93	0.98	1.06	1.14	1.22	1.33	1.62
1% Insig.	0.54	0.77	0.87	0.95	1.02	1.08	1.17	1.26	1.39	1.67
$\boldsymbol{\beta^{OLS,1YD}}$										
Original	0.49	0.68	0.77	0.84	0.92	1.00	1.09	1.19	1.32	1.64
10% Insig.	0.44	0.63	0.70	0.75	0.80	0.86	0.91	0.97	1.05	1.38
5% Insig.	0.45	0.64	0.71	0.77	0.83	0.90	0.95	1.01	1.10	1.42
1% Insig.	0.46	0.65	0.73	0.80	0.86	0.93	1.00	1.07	1.17	1.46
$\boldsymbol{\beta^{OLS,5YD}}$										
Original	0.56	0.70	0.79	0.86	0.93	1.00	1.08	1.16	1.28	1.50
10% Insig.	0.45	0.64	0.70	0.75	0.80	0.84	0.91	0.96	1.08	1.25
5% Insig.	0.46	0.66	0.72	0.77	0.81	0.85	0.91	0.96	1.06	1.25
1% Insig.	0.47	0.68	0.74	0.80	0.84	0.87	0.94	0.96	1.05	1.23
$\boldsymbol{\beta^{OLS,5YM}}$										
Original	0.59	0.81	0.91	1.00	1.08	1.16	1.25	1.36	1.51	1.97
10% Insig.	0.50	0.75	0.84	0.92	0.98	1.03	1.10	1.19	1.26	1.64
5% Insig.	0.51	0.76	0.85	0.93	0.99	1.05	1.14	1.22	1.32	1.64
1% Insig.	0.53	0.77	0.87	0.95	1.02	1.08	1.16	1.27	1.38	1.67

Panel C. Time-series average of firm-level characteristic variables. This panel reports the time-series average of US firm characteristics in ten beta-sorted portfolios. For each portfolio, we separate its stocks into those with statistically significant betas and those with statistically insignificant betas at the 1% level. At the beginning of each calendar month, US stocks are ranked in an ascending order. The stocks are assigned in ten decile portfolios, based on their estimated betas at the end of the previous month. NYSE stock beta breakpoints are adopted in constructing beta-sorted portfolios. We estimate stock betas using $\beta^{Formula}$ which is defined in Section 2.2.2. Our sample period is from January 1926 to December 2017. Portfolio 1 (10) includes stocks with the lowest (highest) betas. The five firm characteristics are lottery demand (*Max*), market capitalization (*Mcap*), book-to-market ratio (*Btm*), natural logarithm of total assets (*Asset*), and idiosyncratic volatility (*Ivol*). *Mcap* and *Asset* are in millions. We calculate the mean of firm characteristics in a portfolio at a given month, and then calculate the time-series average of the means.

Portfolio	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1% Significant										
Max	0.03	0.03	0.03	0.03	0.03	0.04	0.04	0.040	0.04	0.05
Mcap	904	1,327	1,356	1,455	1,559	1,575	1,573	1,439	1,233	780
Btm	0.93	0.83	0.80	0.78	0.77	0.76	0.75	0.74	0.72	0.66
Asset	5.21	5.52	5.54	5.58	5.66	5.65	5.61	5.48	5.29	4.84
Ivol	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.03	0.04
1% Insignificant										
Max	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.06	0.07
Mcap	107	114	132	170	176	139	91	63	44	27
Btm	1.01	0.95	0.97	0.95	1.01	1.08	0.95	0.97	0.92	0.94
Asset	3.52	3.78	4.05	4.20	3.97	3.97	3.95	3.66	3.37	3.00
Ivol	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.04	0.05	0.05

Table 2.2. Betting against beta vs. betting against significant beta: $\beta^{Formula}$

This table presents the returns of beta-sorted portfolios and those with 10% significant beta stocks, 5% significant beta stocks, and 1% significant beta stocks. At the beginning of each calendar month, US stocks are ranked in an ascending order. The stocks are assigned in ten decile portfolios, based on their estimated betas at the end of the previous month. NYSE stock beta breakpoints are adopted in constructing beta-sorted portfolios. We estimate stock betas using five methods: $\beta^{Formula}$, $\beta^{Formula,5YM}$, $\beta^{OLS,1YD}$, $\beta^{OLS,5YD}$, and $\beta^{OLS,5YM}$. The detailed estimation methods are discussed in Section 2.2.2. Our sample period is from January 1926 to December 2017. Portfolio 1 (10) includes stocks with the lowest (highest) betas. The construction of the betting against beta (BAB) portfolio is discussed in Section 2.3.1. We report the portfolios' market excess returns, capital asset pricing model (CAPM) alpha, Fama and French (1993) three-factor model (FF three-factor) alpha, alpha of Fama and French (1993) three-factor model augmented by Carhart (1997) momentum factor and Pástor and Stambaugh (2003) liquidity factor (FF-Carhart-PS five-factor), and alpha of Fama and French (2015) five-factor model (FF five-factor). t-statistics are reported in brackets and 5% statistical significance is indicated in **bold**.

	Portfolio	1	2	3	4	5	6	7	8	9	10	BAB
Excess return	Original	**0.91%**	**0.98%**	**0.99%**	**1.00%**	**0.96%**	**1.03%**	**1.02%**	**0.99%**	**0.99%**	**1.01%**	**0.74%**
		[6.47]	[5.59]	[5.38]	[4.73]	[4.26]	[4.21]	[3.78]	[3.53]	[3.28]	[3.06]	[7.35]
	10% Sig.	**0.86%**	**0.96%**	**0.97%**	**0.98%**	**0.97%**	**1.04%**	**1.01%**	**0.98%**	**1.00%**	**0.99%**	**0.64%**
		[6.22]	[5.52]	[5.21]	[4.59]	[4.25]	[4.25]	[3.74]	[3.50]	[3.27]	[2.99]	[6.72]
	5% Sig.	**0.85%**	**0.96%**	**0.96%**	**0.97%**	**0.98%**	**1.04%**	**1.00%**	**0.98%**	**0.99%**	**1.00%**	**0.62%**
		[6.22]	[5.54]	[5.15]	[4.54]	[4.26]	[4.24]	[3.71]	[3.49]	[3.27]	[3.00]	[6.62]
	1% Sig.	**0.82%**	**0.95%**	**0.95%**	**0.96%**	**0.97%**	**1.04%**	**1.01%**	**0.97%**	**1.00%**	**1.00%**	**0.59%**
		[6.11]	[5.51]	[5.13]	[4.52]	[4.21]	[4.22]	[3.75]	[3.44]	[3.28]	[2.99]	[6.31]
CAPM alpha	Original	**0.48%**	**0.40%**	**0.36%**	**0.27%**	0.17%	0.17%	0.09%	0.02%	-0.04%	-0.05%	**0.62%**

Continued on next page

	Portfolio	1	2	3	4	5	6	7	8	9	10	BAB
		[5.59]	[4.47]	[4.36]	[2.87]	[1.84]	[1.73]	[0.74]	[0.13]	[-0.28]	[-0.25]	[6.44]
	10% Sig.	**0.42%**	**0.38%**	**0.32%**	**0.24%**	0.17%	0.18%	0.07%	0.00%	-0.05%	-0.08%	**0.53%**
		[5.38]	[4.33]	[3.98]	[2.52]	[1.76]	[1.76]	[0.60]	[0.02]	[-0.36]	[-0.41]	[5.79]
	5% Sig.	**0.42%**	**0.37%**	**0.31%**	**0.23%**	0.17%	0.18%	0.06%	0.00%	-0.05%	-0.07%	**0.51%**
		[5.40]	[4.41]	[3.86]	[2.42]	[1.79]	[1.74]	[0.53]	[0.00]	[-0.35]	[-0.40]	[5.69]
	1% Sig.	**0.39%**	**0.37%**	**0.30%**	**0.22%**	0.15%	0.17%	0.07%	-0.01%	-0.05%	-0.08%	**0.49%**
		[5.20]	[4.40]	[3.83]	[2.38]	[1.66]	[1.69]	[0.61]	[-0.12]	[-0.33]	[-0.41]	[5.41]
FF three-factor	Original	**0.41%**	**0.31%**	**0.26%**	**0.16%**	0.04%	0.03%	-0.09%	**-0.16%**	**-0.23%**	-0.24%	**0.60%**
alpha		[6.17]	[5.09]	[5.21]	[2.76]	[0.90]	[0.64]	[-1.47]	[-2.48]	[-2.91]	[-1.91]	[6.25]
	10% Sig.	**0.35%**	**0.29%**	**0.23%**	**0.13%**	0.04%	0.04%	-0.10%	**-0.16%**	**-0.24%**	**-0.25%**	**0.51%**
		[5.71]	[4.94]	[4.53]	[2.25]	[0.86]	[0.81]	[-1.61]	[-2.59]	[-2.95]	[-2.03]	[5.57]
	5% Sig.	**0.35%**	**0.29%**	**0.21%**	**0.12%**	0.05%	0.04%	-0.11%	**-0.17%**	**-0.24%**	**-0.25%**	**0.49%**
		[5.71]	[5.00]	[4.27]	[2.08]	[0.90]	[0.77]	[-1.75]	[-2.61]	[-2.93]	[-2.02]	[5.47]
	1% Sig.	**0.33%**	**0.28%**	**0.21%**	0.11%	0.03%	0.03%	-0.10%	**-0.18%**	**-0.23%**	**-0.25%**	**0.47%**
		[5.47]	[4.90]	[4.14]	[1.94]	[0.62]	[0.66]	[-1.57]	[-2.85]	[-2.90]	[-2.03]	[5.21]
FF-Carhart-PS	Original	**0.35%**	**0.29%**	**0.29%**	**0.18%**	**0.13%**	**0.14%**	0.05%	0.08%	0.09%	0.06%	**0.47%**
five-factor alpha		[4.54]	[4.06]	[4.28]	[2.70]	[1.96]	[2.15]	[0.65]	[0.99]	[0.94]	[0.42]	[3.81]
	10% Sig.	**0.25%**	**0.26%**	**0.25%**	**0.14%**	**0.13%**	**0.15%**	0.04%	0.07%	0.08%	0.02%	**0.32%**
		[3.55]	[3.68]	[3.83]	[2.16]	[2.08]	[2.18]	[0.59]	[0.87]	[0.84]	[0.17]	[2.82]
	5% Sig.	**0.24%**	**0.25%**	**0.24%**	**0.13%**	**0.13%**	**0.14%**	0.03%	0.07%	0.08%	0.03%	**0.30%**
		[3.36]	[3.57]	[3.70]	[2.07]	[2.05]	[2.13]	[0.41]	[0.90]	[0.82]	[0.19]	[2.63]
	1% Sig.	**0.22%**	**0.24%**	**0.22%**	**0.13%**	**0.13%**	**0.13%**	0.03%	0.07%	0.08%	0.02%	**0.28%**
		[3.14]	[3.51]	[3.43]	[2.08]	[1.96]	[2.04]	[0.42]	[0.85]	[0.81]	[0.17]	[2.39]
FF five-factor	Original	**0.27%**	**0.15%**	**0.15%**	0.03%	-0.01%	0.00%	-0.11%	-0.06%	-0.04%	0.06%	**0.35%**
alpha		[3.53]	[2.23]	[2.35]	[0.44]	[-0.21]	[0.02]	[-1.51]	[-0.71]	[-0.35]	[0.46]	[2.94]

Continued on next page

Table 2.2 – Continued from previous page

Portfolio	1	2	3	4	5	6	7	8	9	10	BAB
10% Sig.	0.13%	0.12%	0.12%	0.00%	0.00%	0.00%	-0.12%	-0.08%	-0.05%	0.03%	0.16%
	[1.93]	[1.73]	[1.81]	[-0.07]	[-0.07]	[0.03]	[-1.65]	[-0.86]	[-0.46]	[0.22]	[1.56]
5% Sig.	0.11%	0.10%	0.10%	-0.01%	-0.01%	0.00%	-0.14%	-0.07%	-0.05%	0.04%	0.13%
	[1.64]	[1.47]	[1.56]	[-0.17]	[-0.14]	[-0.02]	[-1.81]	[-0.82]	[-0.47]	[0.26]	[1.28]
1% Sig.	0.09%	0.08%	0.07%	-0.02%	-0.02%	-0.01%	-0.14%	-0.08%	-0.05%	0.03%	0.10%
	[1.34]	[1.24]	[1.15]	[-0.34]	[-0.33]	[-0.20]	[-1.84]	[-0.87]	[-0.47]	[0.24]	[0.93]

Table 2.3. Betting against beta vs. betting against significant beta: Alternative beta estimates

This table presents the returns of beta-sorted portfolios and those with only 1% significant beta stocks. At the beginning of each calendar month, US stocks are ranked in an ascending order. The stocks are assigned in ten decile portfolios, based on their estimated betas at the end of the previous month. NYSE stock beta breakpoints are adopted in constructing beta-sorted portfolios. We estimate stock betas using four methods: $\beta^{Formula,5YM}$, $\beta^{OLS,1YD}$, $\beta^{OLS,5YD}$, and $\beta^{OLS,5YM}$. The detailed estimation methods are discussed in Section 2.2.2. Our sample period is from January 1926 to December 2017. Portfolio 1 (10) includes stocks with the lowest (highest) betas. The construction of the betting against beta (BAB) portfolio is discussed in Section 2.3.1. We report the portfolios' alpha of Fama and French (1993) three-factor model augmented by Carhart (1997) momentum factor and Pástor and Stambaugh (2003) liquidity factor (FF-Carhart-PS five-factor), and alpha of Fama and French (2015) five-factor model (FF five-factor). t-statistics are reported in brackets and 5% statistical significance is indicated in **bold**.

	Portfolio	1	2	3	4	5	6	7	8	9	10	BAB
$\beta^{Formula,5YM}$												
FF-Carhart-PS	Original	**0.15%**	**0.21%**	**0.22%**	**0.21%**	**0.23%**	**0.16%**	0.04%	0.08%	0.15%	0.25%	**0.23%**
five-factor alpha		[2.05]	[3.39]	[3.55]	[3.45]	[3.61]	[2.39]	[0.55]	[1.10]	[1.65]	[1.92]	[2.36]
	1% Sig.	-0.09%	0.13%	0.07%	0.04%	0.00%	0.02%	-0.09%	-0.10%	-0.02%	0.08%	-0.03%
		[-0.82]	[1.42]	[1.04]	[0.63]	[-0.03]	[0.23]	[-1.19]	[-1.34]	[-0.27]	[0.63]	[-0.35]
FF five-factor	Original	0.13%	0.10%	0.09%	0.08%	0.07%	-0.01%	-0.12%	-0.08%	0.03%	0.21%	**0.22%**
alpha		[1.86]	[1.65]	[1.51]	[1.26]	[1.15]	[-0.10]	[-1.66]	[-0.97]	[0.27]	[1.53]	[2.18]
	1% Sig.	**-0.31%**	-0.08%	-0.13%	**-0.18%**	**-0.21%**	**-0.19%**	**-0.31%**	**-0.29%**	-0.16%	0.10%	-0.13%
		[-2.69]	[-0.87]	[-1.86]	[-3.20]	[-3.38]	[-2.82]	[-4.21]	[-3.69]	[-1.73]	[0.78]	[-1.80]
$\beta^{OLS,1YD}$												

Continued on next page

Table 2.3 - Continued from previous page

	Portfolio	1	2	3	4	5	6	7	8	9	10	BAB
FF-Carhart-PS	Original	**0.43%**	**0.26%**	**0.25%**	0.13%	0.08%	0.10%	0.00%	0.03%	-0.10%	**-0.27%**	**1.04%**
five-factor alpha		[4.36]	[3.38]	[3.49]	[1.89]	[1.10]	[1.43]	[0.03]	[0.41]	[-1.15]	[-2.10]	[4.84]
	1% Sig.	0.17%	0.15%	**0.19%**	0.13%	0.06%	0.05%	0.01%	-0.01%	-0.09%	**-0.32%**	**0.39%**
		[0.96]	[1.90]	[2.49]	[1.70]	[0.87]	[0.67]	[0.10]	[-0.19]	[-1.06]	[-2.60]	[2.98]
FF five-factor	Original	**0.35%**	0.10%	0.10%	-0.06%	-0.12%	-0.07%	**-0.17%**	-0.15%	**-0.25%**	**-0.27%**	**0.89%**
alpha		[3.42]	[1.32]	[1.33]	[-0.77]	[-1.64]	[-0.94]	[-2.42]	[-1.75]	[-2.53]	[-2.02]	[4.17]
	1% Sig.	0.17%	-0.01%	-0.05%	**-0.14%**	**-0.19%**	**-0.23%**	**-0.28%**	**-0.30%**	**-0.31%**	**-0.33%**	0.14%
		[0.92]	[-0.10]	[-0.71]	[-2.13]	[-2.97]	[-3.39]	[-3.85]	[-3.66]	[-3.25]	[-2.55]	[1.21]
$\boldsymbol{\beta^{OLS,5YD}}$												
FF-Carhart-PS	Original	**0.46%**	**0.33%**	**0.25%**	**0.21%**	**0.15%**	**0.19%**	0.08%	0.04%	0.00%	-0.08%	**0.77%**
five-factor alpha		[5.22]	[4.19]	[3.54]	[2.87]	[2.09]	[2.78]	[1.12]	[0.55]	[-0.01]	[-0.64]	[5.26]
	1% Sig.	**0.27%**	**0.24%**	**0.22%**	**0.16%**	0.11%	**0.13%**	0.04%	0.01%	-0.02%	-0.10%	**0.38%**
		[3.50]	[3.05]	[3.01]	[2.19]	[1.65]	[2.01]	[0.55]	[0.16]	[-0.29]	[-0.75]	[2.92]
FF five-factor	Original	**0.38%**	**0.19%**	0.13%	0.08%	0.00%	0.03%	-0.08%	-0.15%	-0.17%	-0.19%	**0.71%**
alpha		[4.32]	[2.46]	[1.87]	[1.04]	[0.05]	[0.38]	[-0.95]	[-1.62]	[-1.69]	[-1.34]	[4.84]
	1% Sig.	0.13%	0.05%	0.05%	-0.01%	-0.06%	-0.05%	-0.13%	**-0.18%**	**-0.19%**	-0.21%	0.22%
		[1.69]	[0.66]	[0.66]	[-0.12]	[-0.86]	[-0.66]	[-1.67]	[-2.02]	[-1.96]	[-1.45]	[1.85]
$\boldsymbol{\beta^{OLS,5YM}}$												
FF-Carhart-PS	Original	0.13%	**0.19%**	**0.22%**	**0.24%**	**0.23%**	0.09%	0.09%	0.09%	0.16%	**0.27%**	**0.21%**
five-factor alpha		[1.82]	[3.13]	[3.59]	[3.78]	[3.60]	[1.41]	[1.28]	[1.23]	[1.79]	[2.04]	[2.13]
	1% Sig.	-0.12%	0.13%	0.06%	0.06%	0.04%	-0.04%	-0.04%	-0.10%	0.00%	0.11%	-0.03%
		[-0.97]	[1.04]	[0.81]	[0.97]	[0.58]	[-0.55]	[-0.55]	[-1.41]	[-0.02]	[0.84]	[-0.41]

Continued on next page

Table 2.3 - Continued from previous page

	Portfolio	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**BAB**
FF five-factor	Original	0.12%	0.08%	0.09%	0.10%	0.09%	-0.07%	-0.08%	-0.06%	0.04%	0.22%	**0.20%**
alpha		[1.67]	[1.43]	[1.56]	[1.64]	[1.32]	[-1.04]	[-1.12]	[-0.77]	[0.41]	[1.60]	[2.04]
	1% Sig.	-0.21%	-0.06%	-0.11%	**-0.13%**	**-0.15%**	**-0.24%**	**-0.27%**	**-0.32%**	-0.18%	0.07%	-0.14%
		[-1.67]	[-0.50]	[-1.58]	[-2.44]	[-2.50]	[-3.75]	[-3.79]	[-4.20]	[-1.94]	[0.53]	[-1.82]

Table 2.4. A refined betting against beta strategy: Replacing statistically insignificant betas as zeros

This table presents the performance of original BAB trading strategies and the corresponding refined BAB trading strategies. At the beginning of each calendar month, US stocks are ranked in an ascending order. The stocks are assigned in ten decile portfolios, based on their estimated betas at the end of the previous month. NYSE stock beta breakpoints are adopted in constructing beta-sorted portfolios. We estimate stock betas using five methods: $\beta^{Formula}$, $\beta^{Formula,5YM}$, $\beta^{OLS,1YD}$, $\beta^{OLS,5YD}$, and $\beta^{OLS,5YM}$. The detailed estimation methods are discussed in Section 2.2.2. Our sample period is from January 1926 to December 2017. Portfolio 1 (10) includes stocks with the lowest (highest) betas. The construction of the BAB portfolio is discussed in Section 2.3.1. In refined BAB trading strategies, we replace betas that are statistically insignificant at the 1% level by zeros, and reconstruct ten beta-sorted portfolios and the BAB portfolios. For brevity, we only report the alpha of Fama and French (2015) five-factor model (FF five-factor). t-statistics are reported in brackets and 5% statistical significance is indicated in **bold**.

Portfolio	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**BAB**
$\beta^{Formula}$											
Original	**0.27%**	**0.15%**	**0.15%**	0.03%	-0.01%	0.00%	-0.11%	-0.06%	-0.04%	0.06%	**0.35%**
	[3.53]	[2.23]	[2.35]	[0.44]	[-0.21]	[0.02]	[-1.51]	[-0.71]	[-0.35]	[0.46]	[2.94]
Refined	**0.27%**	**0.15%**	**0.15%**	0.02%	-0.02%	-0.01%	-0.13%	-0.09%	-0.07%	0.01%	**0.42%**
	[3.42]	[2.11]	[2.34]	[0.34]	[-0.32]	[-0.13]	[-1.66]	[-0.96]	[-0.64]	[0.06]	[3.08]
$\beta^{Formula,5YM}$											
Original	0.13%	0.10%	0.09%	0.08%	0.07%	-0.01%	-0.12%	-0.08%	0.03%	0.21%	**0.22%**
	[1.86]	[1.65]	[1.51]	[1.26]	[1.15]	[-0.10]	[-1.66]	[-0.97]	[0.27]	[1.53]	[2.18]
Refined	0.13%	0.09%	0.09%	0.08%	0.08%	-0.01%	-0.13%	-0.09%	0.01%	0.17%	**0.42%**

Continued on next page

Table 2.4 - continued from previous page

Portfolio	1	2	3	4	5	6	7	8	9	10	BAB
	[1.86]	[1.51]	[1.47]	[1.27]	[1.19]	[-0.10]	[-1.69]	[-1.10]	[0.15]	[1.22]	[2.17]
$\beta^{OLS,1YD}$											
Original	**0.35%**	0.10%	0.10%	-0.06%	-0.12%	-0.07%	-0.17%	-0.15%	**-0.25%**	**-0.27%**	**0.89%**
	[3.42]	[1.32]	[1.33]	[-0.77]	[-1.64]	[-0.94]	[-2.42]	[-1.75]	[-2.53]	[-2.02]	[4.17]
Refined	**0.35%**	0.11%	0.10%	-0.05%	-0.12%	-0.07%	**-0.18%**	-0.15%	**-0.24%**	-0.26%	**0.66%**
	[3.46]	[1.40]	[1.36]	[-0.74]	[-1.67]	[-0.99]	[-2.47]	[-1.75]	[-2.49]	[-1.95]	[2.63]
$\beta^{OLS,5YD}$											
Original	**0.38%**	**0.19%**	0.13%	0.08%	0.00%	0.03%	-0.08%	-0.15%	-0.17%	-0.19%	**0.71%**
	[4.32]	[2.46]	[1.87]	[1.04]	[0.05]	[0.38]	[-0.95]	[-1.62]	[-1.69]	[-1.34]	[4.84]
Refined	**0.38%**	**0.19%**	**0.14%**	0.08%	0.01%	0.03%	-0.08%	-0.15%	-0.18%	-0.21%	**0.94%**
	[4.35]	[2.36]	[1.97]	[1.02]	[0.08]	[0.42]	[-0.97]	[-1.64]	[-1.81]	[-1.45]	[5.30]
$\beta^{OLS,5YM}$											
Original	0.12%	0.08%	0.09%	0.10%	0.09%	-0.07%	-0.08%	-0.06%	0.04%	0.22%	**0.20%**
	[1.67]	[1.43]	[1.56]	[1.64]	[1.32]	[-1.04]	[-1.12]	[-0.77]	[0.41]	[1.60]	[2.04]
Refined	0.12%	0.08%	0.09%	0.10%	0.09%	-0.07%	-0.08%	-0.07%	0.02%	0.19%	**0.38%**
	[1.66]	[1.31]	[1.49]	[1.68]	[1.31]	[-1.05]	[-1.11]	[-0.86]	[0.22]	[1.32]	[2.00]

Table 2.5. Betting against significant beta and stock idiosyncratic volatility

Panel A. Replicating Liu et al.'s (2018)'s result. This panel presents the replication of Liu et al.'s (2018)'s Table 5 in our sample. At the beginning of each calendar month, we double sort stocks based on Liu et al.'s (2018) beta and Ang et al.'s (2006) idiosyncratic volatility (*IVOL*), estimated at the end of the previous month, into 50 (5 by 10) value-weighted portfolios. Following Liu et al. (2018), we first regress a stock's monthly return on the contemporaneous market return and one-month lagged market return. The stock's beta is the shrinked summation of the two regression coefficients (Vasicek, 1973; Dimson, 1979). For each of the 50 portfolios, we estimated its Fama and French (1993) (FF) three-factor alpha based on monthly returns. The column labeled "Long 1, Short 10" presents the FF three-factor model alpha of a portfolio that takes a long position on stocks with lowest betas (portfolio 1) and a short position on stocks with the highest betas (portfolio 10). The row labeled "Long 1, Short 5" reports the FF three-factor model alpha of a portfolio that takes a long position on stocks with the lowest *IVOL* (portfolio 1) and a short position on stocks with the highest *IVOL* (portfolio 5). The last column reports the averages across the ten beta-sorted portfolios and the last row reports the averages across the five *IVOL*-sorted portfolios. t-statistics are reported in brackets and 5% statistical significance is indicated in **bold**.

	Beta decile											
***IVOL* quintile**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Long 1, Short 10**	**Average**
1	0.02%	0.04%	0.11%	-0.01%	-0.03%	0.10%	-0.10%	-0.09%	-0.03%	-0.01%	0.03%	-0.02%
	[0.24]	[0.62]	[1.80]	[-0.17]	[-0.51]	[1.44]	[-1.39]	[-1.10]	[-0.45]	[-0.23]	[0.34]	[-0.54]
2	0.12%	-0.01%	-0.06%	-0.01%	0.08%	-0.12%	-0.17%	**-0.28%**	-0.03%	-0.02%	0.14%	-0.05%
	[1.48]	[-0.08]	[-0.70]	[-0.18]	[1.00]	[-1.37]	[-1.82]	[-3.17]	[-0.38]	[-0.23]	[1.17]	[-1.31]
3	-0.07%	-0.06%	0.14%	-0.02%	-0.08%	0.01%	-0.05%	-0.18%	**-0.26%**	-0.20%	0.12%	-0.05%
	[-0.69]	[-0.60]	[1.48]	[-0.15]	[-0.71]	[0.08]	[-0.42]	[-1.37]	[-1.96]	[-1.78]	[0.81]	[-1.00]
4	-0.25%	0.25%	-0.10%	0.17%	0.20%	0.22%	-0.10%	0.22%	-0.23%	-0.18%	-0.07%	0.07%
	[-1.52]	[1.71]	[-0.69]	[1.11]	[1.37]	[1.38]	[-0.66]	[1.22]	[-1.34]	[-1.19]	[-0.32]	[0.78]
5	0.41%	0.20%	**0.59%**	**0.64%**	0.43%	0.21%	0.29%	-0.28%	-0.21%	-0.26%	0.67%	0.29%
	[1.31]	[0.70]	[2.04]	[2.23]	[1.44]	[0.67]	[0.93]	[-0.85]	[-0.61]	[-0.84]	[1.57]	[1.64]
Long1, Short 5	-0.40%	-0.16%	-0.47%	**-0.65%**	-0.46%	-0.11%	-0.40%	0.19%	0.17%	0.24%	-0.64%	-0.31%
	[-1.27]	[-0.56]	[-1.61]	[-2.24]	[-1.58]	[-0.37]	[-1.29]	[0.6]	[0.52]	[0.8]	[-1.46]	[-1.73]
Average	0.04%	0.08%	0.14%	**0.15%**	0.12%	0.08%	-0.02%	-0.12%	-0.15%	-0.13%	0.18%	
	[0.52]	[1.09]	[1.83]	[2.10]	[1.64]	[1.05]	[-0.30]	[-1.33]	[-1.55]	[-1.54]	[1.50]	

Panel B. Replicating Liu et al.'s (2018)'s result: Replacing statistically insignificant betas by zeros. This panel presents the replication of Liu et al.'s (2018)'s Table 5 by replacing statistically insignificant betas by zeros in our sample. At the beginning of each calendar month, we double sort stocks based on Liu et al.'s (2018) beta and Ang et al.'s (2006) idiosyncratic volatility (*IVOL*), estimated at the end of the previous month, into 50 (5 by 10) value-weighted portfolios. Following Liu et al. (2018), we first regress a stock's monthly return on the contemporaneous market return and one-month lagged market return. The stock's beta is the shrinked summation of the two regression coefficients (Vasicek, 1973; Dimson, 1979). Stock betas that are statistically insignificant at the 1% level are replaced by zeros. For each of the 50 portfolios, we estimated its Fama and French (1993) (FF) three-factor alpha based on monthly returns. The column labeled "Long 1, Short 10" presents the FF three-factor model alpha of a portfolio that takes a long position on stocks with lowest betas (portfolio 1) and a short position on stocks with the highest betas (portfolio 10). The row labeled "Long 1, Short 5" reports the FF three-factor model alpha of a portfolio that takes a long position on stocks with the lowest *IVOL* (portfolio 1) and a short position on stocks with the highest *IVOL* (portfolio 5). The last column reports the averages across the ten beta-sorted portfolios and the last row reports the averages across the five *IVOL*-sorted portfolios. t-statistics are reported in brackets and 5% statistical significance is indicated in **bold**.

	Beta decile											
***IVOL* quintile**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Long1, Short 10**	**Average**
1	0.04%	0.03%	0.01%	0.00%	0.02%	0.00%	-0.07%	-0.05%	0.03%	0.00%	0.04%	-0.02%
	[0.56]	[0.52]	[0.21]	[0.06]	[0.31]	[-0.03]	[-0.94]	[-0.64]	[0.40]	[-0.07]	[0.47]	[-0.64]
2	0.10%	0.06%	-0.04%	0.05%	-0.12%	-0.11%	**-0.21%**	**-0.18%**	-0.08%	-0.10%	0.19%	-0.05%
	[1.23]	[0.76]	[-0.48]	[0.63]	[-1.52]	[-1.41]	[-2.37]	[-1.97]	[-0.90]	[-1.11]	[1.67]	[-1.47]
3	-0.13%	-0.06%	0.08%	0.12%	0.05%	0.05%	0.00%	-0.04%	-0.17%	-0.16%	0.03%	0.00%
	[-1.23]	[-0.53]	[0.75]	[1.19]	[0.46]	[0.44]	[0.03]	[-0.34]	[-1.41]	[-1.43]	[0.2]	[-0.08]
4	-0.17%	0.10%	-0.18%	-0.08%	0.23%	0.09%	**0.35%**	-0.14%	-0.20%	-0.24%	0.07%	0.01%
	[-1.00]	[0.64]	[-1.11]	[-0.50]	[1.63]	[0.60]	[2.13]	[-0.88]	[-1.10]	[-1.61]	[0.31]	[0.13]
5	0.51%	0.05%	0.55%	**0.65%**	0.34%	**0.91%**	0.04%	0.00%	-0.42%	-0.40%	**0.91%**	0.30%
	[1.58]	[0.18]	[1.88]	[2.27]	[1.16]	[3.06]	[0.13]	[-0.01]	[-1.34]	[-1.32]	[2.12]	[1.70]
Long1, Short 5	-0.47%	-0.02%	-0.53%	**-0.65%**	-0.32%	**-0.91%**	-0.11%	-0.04%	0.45%	0.39%	**-0.87%**	-0.33%
	[-1.48]	[-0.07]	[-1.83]	[-2.23]	[-1.08]	[-3.23]	[-0.36]	[-0.14]	[1.45]	[1.31]	[-1.98]	[-1.81]
Average	0.07%	0.04%	0.08%	**0.15%**	0.10%	**0.19%**	0.02%	-0.08%	-0.17%	**-0.18%**	**0.25%**	
	[0.81]	[0.46]	[1.06]	[2.01]	[1.33]	[2.54]	[0.26]	[-0.97]	[-1.83]	[-2.17]	[2.16]	

Table 2.6. International evidence

This table presents the Fama and French (2015) five-factor alphas of original BAB portfolios, betting against significant beta portfolios and the corresponding refined BAB portfolios. Our sample covers seven international stock markets: United Kingdom (UK), Canada (CA), Netherlands (NL), Belgium (BE), Switzerland (CH), Hong Kong (HK), and Singapore (SG). We estimate stock betas using five methods: $\beta^{Formula}$, $\beta^{Formula,5YM}$, $\beta^{OLS,1YD}$, $\beta^{OLS,5YD}$, and $\beta^{OLS,5YM}$. The detailed estimation methods are discussed in Section 2.2.2. At the beginning of each calendar month, the stocks of a financial market are assigned in either low- or high-beta portfolios, based on their estimated betas at the end of the previous month. We then construct the BAB portfolio based on the discussions in Section 2.3.1. When constructing betting against significant beta portfolios, we first drop stocks with statistically insignificant betas at the 1% level. For refined BAB portfolios, we replace betas that are statistically insignificant at the 1% level by zeros, and reconstruct ten beta-sorted portfolios and the BAB portfolios. t-statistics are reported in brackets and 5% statistical significance is indicated in **bold**.

Portfolio	**UK**	**CA**	**NL**	**BE**	**CH**	**HK**	**SG**
$\boldsymbol{\beta^{Formula}}$							
Original	-0.03%	**2.66%**	**0.92%**	**2.76%**	**0.75%**	**0.66%**	0.49%
	[-0.09]	[4.54]	[2.62]	[2.56]	[3.64]	[2.32]	[1.65]
1% Sig.	-0.11%	-0.30%	0.41%	**0.75%**	0.34%	**0.61%**	0.01%
	[-0.60]	[-1.41]	[1.46]	[3.58]	[1.93]	[2.45]	[0.02]
Refined	**0.87%**	**6.38%**	**1.24%**	**7.49%**	**1.16%**	**0.90%**	**1.99%**
	[1.98]	[5.81]	[3.49]	[3.81]	[4.66]	[2.88]	[5.73]
$\boldsymbol{\beta^{Formula,5YM}}$							
Original	**0.82%**	**3.57%**	**1.42%**	-0.68%	**1.98%**	**1.09%**	**2.08%**
	[3.25]	[4.31]	[3.51]	[-0.75]	[5.51]	[3.54]	[5.87]
1% Sig.	**0.28%**	-0.13%	0.10%	0.18%	0.29%	0.28%	-0.14%
	[2.46]	[-0.86]	[0.42]	[0.82]	[1.84]	[1.39]	[-0.86]
Refined	**1.02%**	**4.89%**	**1.17%**	0.30%	**1.62%**	**2.11%**	**3.18%**
	[3.75]	[8.63]	[3.16]	[1.24]	[6.85]	[4.48]	[7.50]
$\boldsymbol{\beta^{OLS,1YD}}$							
Original	0.47%	**8.28%**	0.14%	**0.92%**	**1.09%**	**1.23%**	**2.66%**
	[0.75]	[5.98]	[0.18]	[2.84]	[3.25]	[4.18]	[4.22]
1% Sig.	0.08%	0.23%	0.02%	**0.80%**	**0.43%**	**0.53%**	0.23%
	[0.44]	[1.09]	[0.03]	[4.28]	[2.31]	[2.43]	[1.29]
Refined	0.55%	**5.51%**	0.48%	**0.64%**	**1.45%**	**1.72%**	**3.67%**

Continued on next page

Table 2.6 - continued from previous page

Portfolio	UK	CA	NL	BE	CH	HK	SG
	[1.10]	[11.71]	[1.92]	[2.29]	[4.55]	[4.35]	[6.23]
$\beta^{OLS,5YD}$							
Original	-0.29%	**2.89%**	**1.09%**	**1.14%**	**0.79%**	**1.20%**	**0.69%**
	[-1.36]	[6.46]	[3.01]	[2.59]	[2.95]	[4.36]	[2.41]
1% Sig.	-0.12%	-0.17%	0.34%	**0.66%**	0.44%	**0.96%**	0.00%
	[-0.68]	[-0.71]	[1.24]	[2.40]	[2.13]	[3.77]	[-0.02]
Refined	0.91%	**5.95%**	**1.35%**	0.56%	**0.97%**	**1.26%**	**1.39%**
	[1.78]	[7.38]	[3.64]	[1.69]	[3.14]	[4.52]	[4.23]
$\beta^{OLS,5YM}$							
Original	**0.70%**	0.92%	**1.55%**	**2.06%**	**1.54%**	**0.92%**	**2.42%**
	[3.79]	[1.93]	[3.89]	[1.36]	[4.83]	[3.16]	[6.55]
1% Sig.	**0.36%**	**-0.40%**	0.21%	0.38%	0.15%	0.35%	-0.12%
	[2.74]	[-2.15]	[0.84]	[1.75]	[0.74]	[1.67]	[-0.66]
Refined	**0.73%**	**1.83%**	**1.23%**	**4.76%**	**1.72%**	**1.80%**	**3.15%**
	[3.53]	[6.36]	[3.12]	[3.78]	[5.94]	[3.80]	[8.00]

Chapter 3

Abnormal Investment and Firm Performance

3.1. Introduction

In an efficient capital market, firms with better future growth options usually have higher equity valuation. To exercise these growth options, firms with a higher market valuation should have a lower payout ratio and invest more on projects with positive net present value (NPV). However, Lee et al. (2016) document that there is a negative correlation between capital expenditures and industry Tobin's Q since the middle of the 1990s. Furthermore, previous studies document an investment growth anomaly that there is a negative relation between firm-level capital investment and future stock returns.[1] In the capital budgeting context, Hou et al. (2015) argue that given expected cash flows, lower costs of capital lead to higher NPV of new projects and higher firm investment, and higher costs of capital imply lower NPV of new projects and lower firm investment. Since lower costs of capital is also associated with lower expected stock returns, researchers observe a negative investment-return

[1] A large literature documenting the negative investment-return relation includes Gomes et al. (2003), Titman et al. (2004), Liu et al. (2009), Polk and Sapienza (2009), and Kogan and Papanikolaou (2013).

relation.

Since a firm's investment decisions are subject to managerial discretion and random systematic and idiosyncratic shocks, the actual firm investment may deviate from the "optimal investment" predicted by theoretical models. Unlike the previous literature on the relation between investment and stock returns, we focus on firm abnormal investment, which is the gap between actual and predicted investment levels. All information on changes in future firm cash flows, including firm investment decisions, will be instantaneously transferred into a firm's stock prices in an efficient market. Therefore, abnormal investment may reflect shocks to a firm's long-run growth opportunities and carry new information about the firm's fundamentals in the future. For instance, Chen et al. (2007) and Bakke and Whited (2010) show that managers use private information when making their investment decisions. If market investors fully incorporate such new information into stock prices contemporaneously, we should not observe an empirical association between abnormal investment and future stock returns. However, if market investors react to such new information and update their expectations on a firm's future growth with a delay, the current abnormal investment may exhibit certain predictability of future stock returns.

A firm's abnormal investment may also be a proxy for agency costs due to conflicts of interests. On the one hand, the managers of a firm with poor investment opportunities and high free cash flow have an incentive to over-invest for their own benefits, e.g. empire building, rather than for the benefits of shareholders (Jensen, 1986). Fairfield et al. (2003) and Titman et al. (2004) provide empirical evidence that over-investment may generate inefficiency and impair firms' stock performance. On the other hand, agency issues may also be associated with firm under-investment. Due to the conflict of interest between shareholders and bondholders, overhang debts prevent shareholders from capturing the benefits of positive NPV investment opportunities, giving rise to firm under-investment (Myers, 1977). The conflict of interest between managers and shareholders may also lead to firm under-investment. Hart (1983) and Bertrand and Mullainathan (2003) propose the "lazy manager" hypothesis that managers prefer a quiet life and choose not to spend effort on firm investment. Guerrieri and Kondor (2012) and Aghion et al. (2013) offer the "career concern" hypothesis that managers forgo positive NPV projects because the risk associated with new investment may cost them their jobs. Besides the delayed market reaction explanation, the empirical relation between abnormal investment and future stock returns may reflect the agency cost reduction in firm market value.

Using a large sample of U.S. public firms during 1974–2017, we adopt an accounting-based investment model proposed by Richardson (2006) to decompose firm investment into predicted and abnormal components. "Abnormal investment" is defined as the absolute value of the difference between actual and predicted investment, which measures the degree of a firm's investment deviating from its predicted level. We also define over-investment (under-investment) as the absolute value of the abnormal investment which is greater (less) than zero. Next, we sort firms into quintile portfolios at the end of June over our sample period, based on the ranks of most recent estimated abnormal investment, under-investment, and over-investment. The quintile portfolios are rebalanced every year. After adjusting for common systematic risk factors, we find that both abnormal investment and under-investment are negatively related to the performance of the quintile portfolios. However, we do not find any evidence that over-investment affects the performance of the quintile portfolios. A portfolio taking a long position on the firms with bottom quintile under-investment and a short position on the firms with top quintile under-investment generates a positive and statistically significant alpha after adjusting for Fama and French (1993) three factors or Fama and French (2015) five factors. The portfolio's annualized three-factor model alpha and five-factor model alpha are 7.80% and 4.56%, which are also economically significant.

We then employ the Fama and MacBeth (1973) regressions to examine the empirical association between abnormal investment and future stock returns, controlling for firm characteristics.[2] We find that abnormal investment is negatively correlated with future stock returns. When we include both investment and abnormal investment in the same regression, we find that the variation in abnormal investment retains the power of explaining future stock returns, while the coefficient of investment is statistically insignificant. Consistent with the portfolio analysis results, our multivariate regression shows a negative relation between under-investment and future stock returns. However, we cannot find a similar relation between over-investment and stock returns. Taken together, our results suggest that it is the under-investment that mainly drives the negative relation between abnormal investment and future stock returns.

We next examine the two potential mechanisms (discussed above) through which under-investment leads to a decrease in future stock returns: (1) *the market delayed reaction channel* and (2) *the agency cost channel*. With respect to the first mechanism, we first investigate whether under-investment conveys information about future profitability, asset growth, and financial distress. After controlling for firm characteristics, we find that under-investment is negatively associated with the change in earnings and the change in assets over the next year. With one standard deviation increase in under-investment, a firm's earnings growth rate over the next year will decrease 0.06%, which is about 60% of the sample mean of earnings growth rates. A one standard deviation increase in under-investment will also be associated with a 0.63% decrease in a firm's asset growth rate over the next year, which is 5.73% of the sample mean of asset growth rates. Using Shumway's (2001) bankruptcy prediction model, we find that firms with under-investment are more likely to experience future financial distress. With one standard deviation increase

[2] Fama and MacBeth (1973) regressions correct for the cross-sectional correlation among standard errors. In addition, all sample years have equal weights when estimating Fama and MacBeth (1973) regression coefficients, while the panel regression coefficients are biased toward the sample years with more observations.

in under-investment, the probability of financial distress will increase 0.30%, which is 5.77% of the sample mean of unconditional financial distress probabilities. These results confirm that under-investment contains information about firm fundamentals in the future. In an efficient market, investors should promptly incorporate the information carried by abnormal investment into stock prices. To show that the negative relation between under-investment and future stock returns is partly due to the delayed market reaction to under-investment, we employ an empirical test which is similar to the research design used by Abarbanell and Bernard (1992) and Shane and Brous (2001) in their examinations of the post-earnings announcement drift. We show that after controlling for the *future* change in earnings, the *future* change in assets, and the likelihood of *future* financial distress, the relation between under-investment and future stock returns is not statistically significant. About 47.06% of the negative association between under-investment and future stock returns is due to the association between under-investment and future firm fundamentals, supporting *the market delayed reaction channel*.

To explore the second mechanism, *the agency cost channel*, we investigate whether the negative relation between under-investment and future stock returns is more pronounced for firms with weaker external monitoring or higher agency costs. If under-investment is driven by potential agency issues, then market investors will adjust firm value according to under-investment related agency costs, leading to lower stock returns. We first classify firm–year observations with under-investment into two sub-samples using the annual industry medians of blockholder ownership, the ownership of a firm's blockholders who hold more than 5% of the firm's outstanding shares. Firms with blockholder ownership above the median are classified as those with stronger external monitoring and lower agency costs. We find that the negative relationship between under-investment and future stock returns is only statistically significant in the low blockholder ownership sub-sample. We next divide firm–year observations with under-investment into two sub-samples based on the two direct proxies of agency costs proposed by Ang et al. (2000): expense ratio and asset utilization ratio. Higher expense ratios are associated with less efficiency and higher agency costs, while higher asset utilization ratios are associated with greater efficiency and lower agency costs. We find that although the negative relation between under-investment and future stock returns is statistically significant in both partitions, the economic impact of under-investment on future stock returns is larger for firms with higher agency costs. Combined, these findings support *the agency cost channel* that agency conflicts may lead to firm under-investment and hurt firm value.[3]

Finally, we conduct a set of robustness tests to validate our main findings. First of all, we re-estimate the impact of abnormal investment, under-investment, and over-investment on future stock returns using a panel regression with the year and industry fixed effects. To mitigate the concern about econometric issues in estimating the investment model, we reconstruct our three abnormal investment proxy variables using a single panel regression between 1974 and 2017 and rolling panel regressions with five-year estimation windows. To mitigate any concern on the potential model misspecification in Richardson's (2006) framework, we estimate the predicted investment using the two alternative investment models developed by Harvey et al. (2004) and Titman et al. (2004). These robustness tests generally support our main findings that there is a negative relation between abnormal investment and future stock returns and that the negative relation is mainly driven by under-investment, not over-investment. In our supplementary tests, we examine whether the negative relation between under-investment and future stock returns is due to the firm-specific information carried by under-investment or the potential positive association between under-investment and the systematic financial distress

[3]If stock markets are efficient, agency costs associated with under-investment may lead to a contemporaneous change in stock prices and should not be associated with lower future stock returns. We acknowledge that in an efficient market, the agency cost channel would also require that investors underreact to the implications of agency costs for firm investment decisions.

risk factor. We do not find evidence supporting the systematic financial distress risk exposure explanation. We also investigate the impact of market recessions on the negative relation between abnormal investment and future stock returns. We find that the negative relation between abnormal investment and future stock returns is much weaker during market recession periods than non-recession periods, suggesting that, during market recession periods, market investors are more likely to react to the negative information carried by under-investment without a delay.

Our paper is closely related to Titman et al. (2004), which also investigates the association between abnormal capital investment and subsequent stock performance. Titman et al. (2004) find that firms with the most over-investment are likely to under-perform during the following five years. This empirical relation is stronger for firms with more cash flows or fewer debts. Our paper differs from Titman et al. (2004) in two dimensions. First, Titman et al. (2004) measure the abnormal capital investment as the deviation of a firm's capital expenditures from its average capital expenditures over the past three years, whereas our abnormal investment is estimated based on an accounting-based framework which controls for the cross-sectional and time-series variations of firms' growth opportunity, leverage, cash holding, age, size, stock returns, and historical investment. We further divided abnormal investment into under-investment and over-investment since market investors may react differently to the information conveyed by under-investment or over-investment. Second, Titman et al. (2004) find that firms with the least abnormal capital investment tend to out-perform firms with the highest abnormal capital investment in terms of stock returns. Our paper shows that after adjusting for the cross-sectional and time-series variations in firm characteristics, it is under-investment that drives the negative relation between abnormal investment and future stock returns, not over-investment.

The remainder of the paper is structured as follows. Section 3.2 presents our data source, investment model, and summary statistics of the key variables in our empirical analyses. Section 3.3 discusses our main empirical results. Section 3.4 provides robustness test results and further discussions. Finally, Section 4.6 concludes.

3.2. Sample, variables, and summary statistics

3.2.1. Data source and sample selection

Our sample starts with U.S. firm–year observations with available stock return data in the Centre for Research in Security Prices (CRSP) and accounting information in the Compustat Fundamentals Annual files. Following Richardson (2006), we delete firm–year observations without U.S. ordinary common shares, with a negative book value of equity, and with the absolute value of the free cash flow to total assets ratio being greater than one. We also exclude financial firms (SIC codes 6000–6999) from our sample because the investment decisions of financial firms may not convey the same information as those of non-financial firms. After applying these data cleaning filters, we arrive at our main sample of 122,180 firm–year observations over the fiscal year 1974–2017. For the common stocks of our sample firms, we collect their systematic risk factor return data, including the market ($MKTRF$), size (SMB), and book-to-market (HML) from Wharton Research Data Services (WRDS). To apply Fama and French's (2015) five-factor model, we collect the profitability (RMW) and investment factor (CMA) from Kenneth R. French's website. Finally, BAA and AAA rating bond yield data are collected from the Federal Reserve's H-15 report and stock market model $Beta$s are calculated by Eventus using the most recent 255 trading days' returns and CRSP value-weighted index returns as the proxy for market returns. Detailed definitions of all the variables and their data sources are described in Appendix A.

3.2.2. Measures of abnormal investment

Our objective in this paper is to study the empirical association between firm-level abnormal investment and stock returns. Abnormal investment is the deviation from the investment level which would be predicted by a firm-specific investment model. Following Richardson (2006) and Stoughton et al. (2017), we estimate the following accounting-based investment model and use the regression residuals as our proxy for the firm-level abnormal investment:

$$\begin{aligned} INew_{i,t} =& \alpha + \beta_1 \frac{V}{P}_{i,t-1} + \beta_2 Leverage_{i,t-1} + \beta_3 Cash_{i,t-1} + \beta_4 Age_{i,t-1} + \beta_5 Size_{i,t-1} \\ &+ \beta_6 Return_{i,t-1} + \beta_7 INew_{i,t-1} + \sum Industry_j + \sum Year_t + \epsilon_{i,t} \end{aligned} \tag{3.1}$$

where $INew_{i,t}$ is the new investment for firm i in year t, defined as the difference between $ITotal_{i,t}$ and $IMaintenance_{i,t}$. $ITotal_{i,t}$ is the annual total investment expenditure, including capital expenditure, R&D expense, acquisitions, sales of property, plant and equipment, divided by $Asset$. $IMaintenance_{i,t}$ is the required investment expenditure (depreciation and amortization) to maintain assets in place divided by $Asset$. The existing financial economics studies indicate that a firm's new investment depends on future growth opportunities, financial constraints, and other firm characteristics (Hubbard, 1998). Firm growth opportunities are measured by V/P_{t-1}, where V is the value of assets in place and P is the firm's market value (Ohlson, 1995; Feltham and Ohlson, 1996).[4] A firm's market value (P) is the sum of the value of assets in place (V) and the value of future growth opportunities, therefore V/P is negatively related to a firm's future growth opportunities. The financial constraints are measured by firm leverage ratios which are equal to the

[4]Following Ohlson (1995) and Richardson (2006), the value of assets in place is estimated as $(1 - \alpha r)BV + \alpha(1 + r)X - \alpha rd$, α=$\omega/1 + r - \omega$, $r = 12\%$, $\omega = 0.62$, BV is the Book Value of Common Equity (CEQ), X is Operating Income After Depreciation (OIADP), and d is annual Dividend (DVC).

total debts divided by total debts and book value of common equity ($Leverage_{t-1}$) and cash holdings ($Cash_{t-1}$). The other firm characteristics controlled in Equation (3.1) include the natural log of firm age (Age_{t-1}), the natural log of total assets ($Size_{t-1}$), cumulative stock returns over the previous year ($Return_{t-1}$), and the lag of new investment ($INew_{t-1}$). We also include the Fama–French 48 industry fixed effects ($\sum Industry_j$) to control for the variation of firm investment across industries and the year fixed effects ($\sum Year_t$) to control for the time-series variation of firm investment related to stock market trends and business cycles. To mitigate the influence of outliers, we follow Richardson (2006) and winsorize all financial variables in Equation (3.1) at the 1st and 99th percentiles.

The fitted value of the accounting-based investment model, $INew^*_{i,t}$, is taken as the predicted level of new investment for firm i at year t. Then we define firm i's abnormal investment ($AInvt_{i,t}$) as the absolute value of the deviation from the predicted investment: $AInvt_{i,t} = |INew_{i,t} - INew^*_{i,t}|$. $AInvt_{i,t}$ indicates the deviation of investment from its predicated value, without distinguishing between under- and over-investment. Our investment model in Equation (3.1) allows us to further differentiate between firm i's under- and over-investment. If $INew_{i,t} < INew^*_{i,t}$, then the under-investment proxy variable is defined as $Under_{i,t} = |INew_{i,t} - INew^*_{i,t}|$. If $INew_{i,t} > INew^*_{i,t}$, then the over-investment proxy variable is defined as $Over_{i,t} = |INew_{i,t} - INew^*_{i,t}|$. Since market investors may react differently to the information conveyed by under- or over-investment, it is important for us to differentiate the direction of abnormal investment in our empirical analyses. We measure the general abnormal investment, over-investment, and under-investment in absolute value, so that the estimated coefficients of these three proxies are comparable to each other in our empirical analyses.

To investigate the empirical association between the firm-level abnormal investment and future stock returns, we need to avoid the "look ahead bias" due to the use of future information in estimating the current abnormal investment. In other words, the information used to estimate abnormal investment should be available to market investors before stock returns are measured. For each year t between 1980 and 2017, we estimate a historical panel regression on a sample of firm–year observations between 1974 and year $t-1$. For example, a firm's abnormal investment in 1980 is estimated by running a panel regression based on firm–year observations between 1974 and 1979, and a firm's abnormal investment in 1981 is estimated by running a panel regression based on firm–year observations between 1974 and 1980, and so on. In our robustness tests, we estimate abnormal investment with two alternative regression methods. First, we follow Richardson (2006) and Stoughton et al. (2017) to estimate Equation (3.1) by a single panel regression between 1974 and 2017. Second, for abnormal investment in year t, we estimate Equation (3.1) by a five-year rolling window between year $t-4$ and year t.

3.2.3. Summary statistics

Table 4.1 presents the descriptive statistics of all the variables in our main empirical analyses. The number of observations, mean, standard deviation, 1st percentile, 25th percentile, median, 75th percentile, and 99th percentile are reported from left to right, in sequence for each variable. The mean and standard deviation of *INew* are 0.07 and 0.11, which are comparable to those (0.08 and 0.13) reported in Richardson (2006).

Panel A of Table 3.2 summarizes the regression coefficients estimated by the investment model. For each year t between 1980 and 2017, we run a panel regression of Equation (3.1) based on firm–year observations between 1974 and year $t-1$. We only report the time-series average of the coefficients estimated by thirty-eight historical panel regressions from 1980 to 2017. Year and Fama–French 48 industry fixed effects are controlled in these regressions. The t-values of regression coefficients are based on standard errors clustered by firm. The numbers of positive and negative coefficients at the 1% statistical significance level are reported in parentheses. The negative coefficients of V/P_{t-1} suggest that firms with better future growth opportunities make a higher investment. Since a lower leverage ratio and higher cash holdings indicate lower financial constraints, the negative coefficients of $Leverage_{t-1}$ and the positive coefficients of $Cash_{t-1}$ show that firms with lower financial constraints make a higher investment. The negative coefficients of Age_{t-1} suggest that firms in the later stage of their life cycle tend to invest less, while the positive coefficients of $Size_{t-1}$ suggests that larger firms tend to make a higher investment. $Return_{t-1}$ captures additional variations in investment expenditure that are not explained by growth opportunities and financial constraints but may temporarily affect firms' investment decisions. The positive coefficients of $Return_{t-1}$ suggest that firms with higher past stock performance tend to invest more. The positive coefficients of $INew_{t-1}$ suggest that new investment expenditure is increasing in prior investment activities. The signs of these coefficients are all consistent with Richardson (2006). The average R^2 of the thirty-eight historical panel regressions is 0.342, suggesting that the investment model can explain a large portion of the cross-sectional and time-series variations in firm-level investment.

Panel B of Table 3.2 presents the descriptive statistics of the predicted firm investment $INew^*$ and our abnormal investment proxy variables. We observe that about 59.0% (41.0%) of the firm–year observations in our sample have a lower (higher) investment than the predicted investment level. The means (standard deviations) of our three abnormal investment proxies, $AInvt$, $Under$, and $Over$, are 0.057 (0.064), 0.049 (0.046), and 0.070 (0.082), respectively. Since the mean of the predicted new investment $INew^*$ is 0.069, our three abnormal investment proxy variables are economically important.

3.3. Empirical results

This section presents our main empirical findings.

3.3.1. Quintile portfolio analysis

To examine the empirical association between firm-level abnormal investment and future stock returns, we begin by forming quintile portfolios based on firm-level abnormal investment and estimating the performance of these quintile portfolios. Following Fama and French's (1993) portfolio construction method, we sort all stocks into quintile portfolios based on one of their most recent estimated abnormal investment proxies *AInvt*, *Under*, and *Over*, at the end of June in each year of 1980–2017. Stocks with the lowest (highest) abnormal investment measures are allocated to portfolio 1 (5). Then we calculate the equally weighted monthly returns of these quintile portfolios over the next twelve-month holding period. To evaluate the performance of these quintile portfolios, we estimate their portfolio alphas using the Fama and French (FF) three-factor model (Fama and French, 1993) and five-factor model (Fama and French, 2015):

$$R_{p,t} - Rf_t = \alpha_p + \beta_1(MKTRF_t) + \beta_2 SMB_t + \beta_3 HML_t + \epsilon_{p,t} \tag{3.2}$$

$$R_{p,t} - Rf_t = \alpha + \beta_1(MKTRF_t) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 RMW_t + \beta_5 CMA_t + \epsilon_t \tag{3.3}$$

where $R_{p,t}$ denotes the portfolio p's return over month t, Rf_t denotes the risk-free return measured by the one-month Treasury bill rate over month t, $MKTRF_t$ denotes the excess return for the market portfolio[5] over month t, SMB_t denotes the return of a size factor mimicking portfolio over month t, HML_t denotes the return

[5]This is a portfolio consisting of all securities in the capital market where the proportion invested in each security corresponds to its relative market value.

of a value factor mimicking portfolio over month t, RMW_t denotes the return of a profitability factor mimicking portfolio over month t, and CMA_t denotes the return of an investment factor mimicking portfolio over month t.

Table 3.3 reports the alphas of quintile portfolios estimated by the two multi-factor models. The last column reports the corresponding factor model alpha of a portfolio that takes a long position on stocks in portfolio 1 and a short position on stocks in portfolio 5. For stocks sorted by the abnormal investment proxy $AInvt$, the quintile portfolio alpha decreases from quintile 1 portfolios to quintile 5 portfolios in terms of both statistical significance and value. The FF three-factor model alpha of the portfolio taking a long position on stocks in portfolio 1 and a short position on stocks in portfolio 5 is positive and statistically significant. The annualized FF three-factor alpha of the long–short portfolio is 3.60% ($= 0.30\% * 12$), which is economically significant. The FF five-factor model alpha of the long–short portfolio is positive and statistically insignificant.

For stocks sorted by the under-investment proxy *Under*, we also find that the quintile portfolio alphas estimated by both multi-factor models monotonically decrease from quintile 1 portfolios to quintile 5 portfolios in terms of both value and statistical significance. For example, The FF three-factor model alpha of portfolio 1 is 0.33% with a t-statistic of 4.22 and the FF three-factor model alpha of portfolio 5 is -0.32% with a t-statistic of -2.58. We then construct a long–short portfolio taking a long position on stocks in portfolio 1 and a short position on stocks in portfolio 5. Both multi-factor model estimated alphas of the long–short portfolio are positive and statistically significant at the 1% level. Using the FF three-factor model as an example, the alpha of portfolio 1 is 0.33% with a t-statistic of 4.22 and the alpha of portfolio 5 is -0.32% with a t-statistic of -2.58. The annualized FF three-factor model alpha and FF five-factor model alpha of the long–short portfolios are 7.80% ($= 0.65\% * 12$) and 4.56% ($= 0.38\% * 12$), which are economically significant. For stocks sorted by the over-investment proxy *Over*, we do not find a similar decreasing pattern for the quintile portfolio alphas. The alphas of the long–short portfolio are negative and statistically insignificant for both multi-factor models. These results suggest that the negative relation between abnormal investment and future stock returns is mainly explained by under-investment, not by over-investment.[6]

3.3.2. Abnormal investment and future stock returns

In this section, we investigate the empirical association between firm-level abnormal investment and future stock returns using the following multivariate regression:

$$\begin{aligned} BHR_{i,t+1} = &\alpha + \beta_1 Investment\ components_{i,t} + B * Control\ variables_{i,t} \\ &+ \sum Industry_j + \epsilon_{i,t} \end{aligned} \tag{3.4}$$

where $BHR_{i,t+1}$ is one-year buy-and-hold returns starting from the beginning of the fourth month after the end of firm i's fiscal year t. To make sure that all the information on our explanatory variables are available to the market investors when we measure stock returns, we follow the previous literature and forward the stock returns by three months. *Investment components* is one of the following three variables: $AInvt_t$, $Under_t$, and $Over_t$. Our control variables include market-to-book ratio (MTB_t), leverage ($Leverage_t$), cash holdings ($Cash_t$), firm size ($Size_t$), and lag one-year buy-and-hold returns (BHR_t). We also control for the industry fixed effects based on the Fama–French 48 industry classification. Since stock returns are the dependent variable in Equation (3.4) and have cross-sectional correlation, we adopt the Fama and MacBeth (1973) regression to estimate Equation (3.4). The Fama and MacBeth (1973) regression helps to correct for the cross-sectional correlation among standard errors. Furthermore, the panel regression coefficients may be affected by the years with more observations. This concern is also mitigated by the Fama and

[6]Our results are qualitatively the same if we use raw excess returns instead of the multi-factor model alpha.

MacBeth (1973) regression, in which all years are treated as equally important.

We present the Fama and MacBeth (1973) regression results in Table 3.4. In column (1), the coefficient of $AInvt_t$ is negative and statistically significant, suggesting that abnormal investment is still negatively associated with future stock returns after controlling for firm characteristics. Next, we separate firm–year observations into those with under-investment and those with over-investment. Columns (2) and (3) of Table 3.4 show that both under-investment and over-investment are negatively associated with future stock returns. However, in terms of the coefficient value and statistical significance level, the negative relation between under-investment and future stock returns is much stronger than the relation between over-investment and future stock returns, suggesting that the negative relation between abnormal investment and future stock returns is mostly driven by under-investment.

In columns (4)–(6), we include both abnormal investment components and $INew_t$ in the multivariate regressions. Column (4) shows that the coefficient of $INew_t$ is statistically insignificant while the coefficient of $AInvt_t$ remains negative and statistically significant at the 1% level. A one standard deviation increase in $AInvt$ is associated with a 0.86% ($= 0.064 * -0.135$) decrease in firm annual buy-and-hold stock returns, which accounts for 5.51% ($= 0.86\%/15.6\%$) of an average firm's annual buy-and-hold stock returns. As shown in column (5), after including $INew_t$ in the multivariate regressions, the coefficient of $Under_t$ remains negative and statistically significant at the 1% level. A one standard deviation increase in $Under$ results in a 2.06% ($= 0.046 * -0.447$) decrease in firm annual buy-and-hold stock returns, which is equivalent to 13.21% ($= 2.06\%/15.6\%$) of an average firm's annual buy-and-hold stock returns. Column (6) shows that the coefficient of $Over_t$ turns into positive but only statistically significant at the 10% level, after controlling for $INew_t$. After controlling for $INew_t$, we do not find the evidence shown in (Titman et al., 2004) that firms increasing capital investments achieve negative stock returns subsequently.

Overall, the multivariate regression results reported in Table 3.4 are consistent with those documented in our quintile portfolio analysis. Taken together, the results in Table 3.4 have the following three implications. First, when a firm's actual investment deviates from its predicted level, its future stock performance is weaker. Second, it is under-investment per se rather than over-investment that explains the negative relation between abnormal investment and future stock returns. Last, there is no consistent evidence supporting the investment puzzle documented in the previous studies that firms investing above the predicted investment level have a worse future stock performance.

3.3.3. Potential mechanisms

So far, we have decomposed firm-level investment into the predicted and abnormal components. Our findings show that the negative association between investment and future stock returns is mainly due to abnormal investment. In this section, we investigate the two potential channels through which abnormal investment has a negative impact on future stock returns.

Delayed market reaction to under-investment

In a standard project valuation model, managers should incorporate their private information about the firms' future profitability and distress risk into their investment decisions (Chen et al., 2007). Bakke and Whited (2010) also find that managers may incorporate private investor information when making investment decisions. Therefore, abnormal investment may provide the market with new information about the evolution of the firms' future fundamentals. When market imperfections prevent investors from processing the new information embedded in firms' abnormal investment, stock prices may not fully react to such forward-looking information. Then the contemporaneous stock prices cannot fully reflect the fundamental information conveyed by abnormal investment, leading to stock misvaluation (*the delayed market reaction channel*).

To test this channel, we first investigate whether abnormal investment captures the information relevant to three firm fundamentals: future profitability, future asset growth, and the likelihood of future financial distress. To test whether abnormal investment predicts future profitability or asset growth, we adopt the following two panel regressions:

$$\begin{aligned}\Delta Earnings_{i,t\ to\ t+1} =& \alpha + \beta_1 Abnormal\ investment_{i,t} + B * Control\ variables_{i,t} \\ &+ \sum Year_t + \sum Industry_j + \epsilon_{i,t}\end{aligned} \tag{3.5}$$

$$\begin{aligned}\Delta Assets_{i,t\ to\ t+1} =& \alpha + \beta_1 Abnormal\ investment_{i,t} + B * Control\ variables_{i,t} \\ &+ \sum Year_t + \sum Industry_j + \epsilon_{i,t}\end{aligned} \tag{3.6}$$

where $\Delta Earnings_{t\ to\ t+1}$ is equal to $(Earnings_{t+1} - Earnings_t)/Assets_t$, where $Earnings_t$ is equal to income before extraordinary items plus interest expense divided by total assets. $\Delta Assets_{t\ to\ t+1}$ is equal to $(Assets_{t+1} - Assets_t)/Assets_t$. *Abnormal investment* is one of the three abnormal investment proxies: $AInvt_t$, $Under_t$, and $Over_t$. The control variables in Equation (3.5) include the book-to-market ratio (BTM_t), total assets ($Size_t$), capital structure ($Leverage_t$), and current earnings ($Earnings_t$). The control variables in Equation (3.6) include BTM_t, $Size_t$, and $Leverage_t$. Year and Fama–French 48 industry fixed effects are also included in these two regressions.

On the one hand, Jensen (1986) indicates that managers with an empire building tendency have an incentive to over-invest and grow their firms beyond the optimal size. Arif and Lee (2014) show that firms with higher capital spending are more likely to experience a decrease in future earnings. On the other hand, managers who anticipate potential future financial constraints may forgo positive NPV projects, which negatively affect firms' future profitability. Table 3.5 reports the panel regression results. In columns (1)–(3), the dependent variable is the change in earnings over the next one-year horizon. The coefficient of *AInvt* in column (1) is -0.011 and statistically significant at the 1% level. A one standard deviation increase in *AInvt* results in a 0.07% ($= 0.064 * -0.011$) decrease in future annual earnings growth rate, which is about 70% ($= 0.07\%/0.10\%$) of an average firm's annual earnings growth rate in our sample. Column (2) shows that the coefficient of *Under* in column (2) is -0.013 and statistically significant at the 5% level. A one standard deviation increase in *Under* results in a 0.06% ($= 0.046 * -0.013$) decrease in future earnings growth rate, which is about 60% ($= 0.06\%/0.10\%$) of an average firm's annual earnings growth rate in our sample. We also find weak evidence that over-investment is negatively related to future profitability. The coefficient of *Over* in column (3) is -0.005 and statistically significant at the 5% level. These results suggest that abnormal investment, especially under-investment, negatively predicts future profitability, which may explain the negative relation between under-investment and future stock returns. In columns (4)–(6), the dependent variable is the change in total assets over the next one-year horizon. The coefficients of *AInvt* and *Over* are positive and statistically significant, while the coefficient of *Under* is negative and statistically significant. A one standard deviation increase in *Under* results in a 0.63% ($= 0.046 * -0.136$) decrease in future annual asset growth rate, which is about 5.73% ($= 0.63\%/11.0\%$) of an average firm's annual assets growth rate in our sample. These results indicate that it is the under-investment which contains the negative information about firm future asset growth. Such negative information also helps to explain the negative relation between under-investment and future stock returns.

Second, we direct our attention to whether abnormal investment contains the information of future financial distress. On the one hand, stockholders have an incentive to take riskier projects than bondholders do, since stockholders only have limited liability. Firms with a potential bankruptcy risk may choose to borrow money from debt holders and over-invest on risky projects. On the other hand, it is costly for firms with financial constraints to raise money from the external credit market. Such firms may choose to under-invest and forgo projects with positive net present value. We follow Shumway's (2001) bankruptcy prediction model to estimate the relation between abnormal investment and the probability of future financial distress. Specifically, we run the following logit regression of distress probability on our abnormal investment proxies:

$$\begin{aligned} Delist_{i,t\ to\ t+3} =& \alpha + \beta_1 Abnormal\ investment_{i,t} + B * Control\ variables_{i,t} \\ &+ \sum Year_t + \sum Industry_j + \epsilon_{i,t} \end{aligned} \tag{3.7}$$

where $Delist_{i,t\ to\ t+3}$ is equal to one if the firm i is delisted due to performance-related reasons in the next three years, and zero otherwise. The control variables are profitability which is equal to net income divided by total assets (*Profit*), leverage which is equal to the total debts divided by total debts and book value of common equity (*Leverage*), market value of equity to the total market values in year t (*MVE/Total MV*), abnormal returns in the prior fiscal year (*AR*), stock return volatility which is equal to the standard deviation of the residuals from the regression of monthly stock returns on the value-weighted market index return (*Volatility*), firm size (*Size*). We also control for year and Fama–French 48 industry fixed effects in Equation (3.7).

Table 3.6 presents the marginal effect results of the bankruptcy prediction model. In column (1), we only include the control variables. The coefficients of all the control variables are statistically significant and their signs are generally consistent with Shumway (2001). The coefficients of *AInvt* and *Under* are positive and statistically significant at the 1% level in columns (2)–(3), while the coefficient of *Over* is not statistically significant. The Pseudo R^2 and the area under the receiver operating characteristic (ROC) curve are larger in column (3) than those in column

(1), suggesting that adding under-investment in the bankruptcy prediction model increases the model's ability to identify financial distresses firms. A one standard deviation increase in *Under* is associated with a 0.30% (= 0.046 ∗ 0.065) increase in the probability of future financial distress. Given that the sample mean value of the unconditional probability of financial distress is 5.20%, such an increase in the probability of financial distress is equivalent to 5.77% (= 0.30%/5.20%) of the sample mean. These results support the view that abnormal investment carries information on the probability of future financial distress. More importantly, it is under-investment, not over-investment, that appears to have incremental value for predicting future financial distress.

We have shown that under-investment contains additional information about future changes in profitability, changes in asset growth, and the likelihood of financial distress. If stock prices incorporate the information carried by under-investment immediately, then we should not observe an empirical relation between under-investment and future stock returns. However, the negative relation between under-investment and future stock returns we have documented suggests that market investors may fail to fully react to such information. To test *the delayed market reaction channel*, we follow the empirical design of Caskey et al. (2012), and examine the following regression model[7]:

$$\begin{aligned} BHR_{i,t+1} = &\alpha + \beta_1 Under_{i,t} + B_1 * Control\ variables_{i,t} \\ &+ B_2 * Future\ fundamentals_{i,t} + \sum Industry_j + \epsilon_{i,t} \end{aligned} \tag{3.8}$$

where the dependent variable is next year's stock returns and the control variables are the ratio of book value to market value of equity ($BTM_{i,t}$), the natural log of market value of equity ($Ln(MVE)_{i,t}$), and market systematic risk measured by the beta of the standard market model ($Beta_{i,t}$). *Future fundamentals* include next year's

[7] Abarbanell and Bernard (1992) and Shane and Brous (2001) also use similar analyses to study the post-earnings announcement drift.

change in debt ($\Delta Debt_{t\ to\ t+1}$), next year's change in earnings ($\Delta Earnings_{t\ to\ t+1}$), next year's asset growth ($\Delta Asset_{t\ to\ t+1}$), and performance related delisting indicator variable over the next three years ($Delist_{t\ to\ t+3}$). Adding *Future fundamentals* one by one into Equation (3.8) will reduce the value and the statistical significance of the coefficient on $Under_{i,t}$, if the return predictability of under-investment is partially due to the market's delayed reaction to the information carried by under-investment.

Table 3.7 presents the Fama and MacBeth (1973) regression results of Equation (3.8). Without adding fundamentals, the coefficient of *Under* in column (1) is -0.238 and statistically significant at the 1% level. In columns (2)–(4), we add the three fundamental variables one by one in Equation (3.8). The coefficients of *Under* decreases in terms of both value and statistical significance. These results suggest that the negative relation between under-investment and future stock returns could be partially explained by the market's failure to efficiently incorporate the fundamental information carried by under-investment into stock prices. In column (5), we include all the three fundamental variables together in Equation (3.8), the coefficient of *Under* is -0.126 and statistically insignificant. About 47.06% ($= (0.238 - 0.126)/0.238$) of the negative association between under-investment and future stock returns is due to the future changes in firm fundamentals conveyed by under-investment.

In sum, the empirical findings in this section support *the delayed market reaction channel* that market investors do not fully incorporate the future fundamentals associated with under-investment into the contemporaneous stock prices, which results in a negative relation between under-investment and future stock returns.

Agency costs

In the previous section, we document the market inefficiency channel through which under-investment may lead to lower future stock returns. However, even if there is no delayed market reaction to the fundamental information conveyed by under-investment, we may still observe a negative relation between under-investment and future stock returns due to agency problems. If under-investment is due to the conflicts of interests between shareholders and bondholders or between managers and shareholders, then the agency costs related to under-investment have a negative impact on firm value and are associated with lower future stock returns (*the agency cost channel*).

Myers (1977) first discusses the debt overhang problem that the existence of debt may lead to an under-investment problem because a firm with outstanding debt has an incentive to forgo positive NPV investment opportunities if the benefits of the new projects accrue to bondholders instead of shareholders. Bergman and Callen (1991) also identify the possibility of opportunistic under-investment by firm managers in debt renegotiation. Bergman and Callen (1991) argue that if managers act strictly in the shareholders' interests, due to the conflicts of interests between bondholders and shareholders, managers may optimally use their discretion over firm investment decisions to force concessions from the firms' creditors by threatening to sap firm value through under-investment. Therefore the conflict of interests between shareholders and bondholders may lead to under-investment. In addition, firms may under-invest due to the conflict of interests between managers and shareholders. With asymmetric information and the lack of external monitoring, managers may prefer a "quiet life" (e.g., Hart, 1983; Bertrand and Mullainathan, 2003), since it is costly for them to make complicated investment decisions. Moreover, managers may be risk averse and intentionally choose not to invest in risky projects due to "career concerns". Instead of being lazy, managers may worry about losing their jobs if their new projects have unfavorable outcomes due to random factors (Aghion et al., 2013). Both "quiet life" and "career concerns" may explain why managers bypass positive NPV projects, leading to inferior future stock returns.

To test *the agency cost channel*, we adopt sub-sample analyses and divide firm–year observations with under-investment into two sub-samples based on the annual industry medians of *Blockholder Ownership*, *Expense Ratio*, and *Asset Utilization Ratio*. If the negative relation between under-investment and future stock returns is partly due to the agency costs, then such negative relation is likely to be more pronounced among firms subject to a poorer external monitoring environment. Edmans (2014) reviews the theoretical and empirical studies on blockholders and summarizes the "voice" and "exit" channels through which blockholders may engage in corporate governance. *Blockholder Ownership* is defined as the ownership of a firm's blockholders who hold more than 5% of the firm's outstanding shares. A higher *Blockholder Ownership* indicates better corporate governance quality and fewer agency costs. We examine whether the negative relation between under-investment and future stock returns can be explained by the cross-sectional differences in *Blockholder Ownership*. Columns (1) and (2) of Table 3.8 report the results of sub-sample analyses for firms with low and high *Blockholder Ownership*. The coefficient of $Under_t$ remains negative for both sub-samples, but is only statistically significant in the low *Blockholder Ownership* partition.

Next, we adopt two direct proxies for agency costs proposed by Ang et al. (2000): *Expense Ratio* and *Asset Utilization Ratio*.[8] *Expense Ratio* is defined as operating expenses scaled by total sales, which is a measure of how effectively a firm's managers control operating costs, including excessive perquisite consumption and other direct agency costs. *Expense Ratio* is positively related to agency costs. *Asset Utilization Ratio* is defined as total sales scaled by total assets, which is a measure of how effectively a firm's managers deploy its assets. In the contrary to *Expense Ratio*, *Asset Utilization Ratio* is negatively related to agency costs. Columns (3)–(4) and (5)–(6) of Table 3.8 report the results of sub-sample analyses for firms with low and high *Expense Ratio* and *Asset Utilization Ratio*, respectively. The coefficients of $Under_t$ remain negative and statistically significant for both sub-

[8] Other commonly used proxies for agency costs from the previous literature, such as managerial ownership and anti-takeover rights, are only available for firms included in the S&P 1500 index.

samples. However, the coefficients of $Under_t$ are larger in terms of the absolute value in high *Expense Ratio* and low *Asset Utilization Ratio* partitions, suggesting that the negative relation between under-investment and future stock returns is more pronounced in firms with high agency costs.

Taken as a whole, we also find evidence supporting *the agency cost channel* that the negative relation between under-investment and future stock returns is due to the agency costs associated with under-investment.[9]

3.4. Robustness tests and further discussions

In this section, we provide the results of robustness tests and further discussions on our results.

3.4.1. Abnormal investment and future stock returns: Alternative econometric estimation methods

Our main results reported in Table 3.4 rely on the abnormal investment proxies estimated by historical panel regressions. Also, the empirical relation between abnormal investment and future stock returns is estimated by Fama and MacBeth (1973) regressions. In this section, we check whether our main results are robust to alternative econometric estimation methods. Table 3.9 reports the results of these robustness tests.

In Panel A, B, and C, we report the robustness test results for *AInvt*, *Under*, and *Over*, respectively. In column (1), the abnormal investment proxies are estimated by the historical panel regressions between 1974 and year t. In column (1), we use a panel regression to examine the empirical relation between the ab-

[9]To formally test the statistical significance of the differences in coefficients between two subsamples, we compare the mean differences in the thirty-eight regression coefficients from the Fama and MacBeth (1973) regressions. We find that the mean differences are statistically significant at the 5% level for *Blockholder Ownership* and *Asset Utilization Ratio*.

normal investment proxies and future stock returns. In columns (2)–(3), we follow Richardson (2006) and Stoughton et al. (2017) to estimate the abnormal investment proxies by running a single panel regression of Equation (3.1) between 1974 and 2017. Then we examine the empirical relation between the abnormal investment proxies and future stock returns using a panel regression and a Fama and MacBeth (1973) regression in columns (2) and (3), respectively. In columns (4)–(5), we estimate the abnormal investment proxies by rolling panel regressions with five-year fixed windows. Specifically, for abnormal investment in year t, we estimate Equation (3.1) with a five-year rolling window from year $t-4$ to year t. Then we examine the empirical relation between the abnormal investment proxies and future stock returns using a panel regression and a Fama and MacBeth (1973) regression in columns (4) and (5), respectively. The coefficients of the control variables in Equation (3.1) are suppressed for brevity.

Panel A of Table 3.9 shows that the coefficients of *AInvt* remain negative and statistically significant. Panel B of Table 3.9 shows that the coefficients of *Under* are negative and statistically significant at the 1% level in all the respective columns. Panel C of Table 3.9 shows that the coefficients of *Over* are positive and statistically significant, except in column (3). The coefficients of under-investment are generally larger than those of over-investment in the corresponding columns. Overall, our main results remain robust to these alternative econometric estimation methods. We still find a negative relation between abnormal investment and future stock returns, which is mainly driven by under-investment.

3.4.2. Alternative measures of abnormal investment

In our empirical analyses, we measure the level of abnormal investment following an accounting-based investment model proposed by Richardson (2006). As a result, the inferences drawn from our empirical analyses are contingent on the reliability of the investment expectation model. In this section, we check whether our main results are robust to two alternative measures of abnormal investment which have been developed in the previous investment literature. First, we follow Harvey et al. (2004) and use industry median investment as the benchmark investment level. We measure *AInvt* as the absolute value of the difference between a firm's investment and its industry median investment. In addition, *Under* is the absolute value of the difference when a firm's investment is less than its industry median investment and *Over* is the absolute value of the difference when a firm's investment is greater than its industry median investment. Second, we follow Titman et al. (2004) and use a firm's average capital expenditure during the previous three years as its benchmark investment level. We measure *AInvt* as the absolute value of the difference between a firm's capital expenditure in year $t-1$ and its average capital expenditures during the previous three years. Similarly, *Under* is the absolute value of the difference when a firm's investment is less than its benchmark investment level and *Over* is the absolute value of the difference when a firm's investment is greater than its benchmark investment level.

Table 3.10 presents the Fama and MacBeth (1973) regression results of future stock returns on alternative abnormal investment proxy variables. For both alternative measures, the coefficients of *Under* are negative and statistically significant. The coefficients of *AInvt* and *Over* are negative and statistically significant for Harvey et al.'s (2004) measure, but statistically insignificant for Titman et al.'s (2004) measure.[10] These results suggest that the negative relation between underinvestment and future stock returns remains robust to these two alternative measures of abnormal investment.

[10] If we define *AInvt* as the raw difference, instead of the absolute value of the difference, the coefficient of *AInvt* is -0.250 and statistically significant at the 1% level for Harvey et al.'s (2004) measure, and is -0.013 and statistically insignificant for Titman et al.'s (2004) measure.

3.4.3. Systematic financial distress risk and the relation between under-investment and future stock returns

Section 3.3.3 shows that abnormal investment carries information on future firm fundamentals. We further provide evidence that the negative relation between under-investment and future stock returns is partly due to the fact that markets fail to fully react to the information on firm-specific financial distress conveyed by under-investment. Besides the delayed market reaction explanation, an alternative explanation of our findings is that firms with low under-investment have high exposure to systematic financial distress risk. According to this alternative explanation, the high abnormal returns of firms with low under-investment stem from the high risk-premium of systematic financial distress risk. To differentiate our market delayed reaction explanation and the alternative risk-premium-based explanation, we augment the Fama and French (2015) five-factor model by an additional systematic financial distress risk factor FDR:

$$\begin{aligned} R_{p,t} - Rf_t =& \alpha + \beta_1 MKTRF_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 RMW_t + \beta_5 CMA_t + \\ & \beta_6 FDR_t + \epsilon_{p,t} \end{aligned} \tag{3.9}$$

where $R_{p,t}$ denotes the portfolio p's return over month t, Rf_t denotes the risk-free return measured by the one-month Treasury bill rate over month t; $MKTRF_t$, SMB_t, HML_t, RMW_t, and CMA_t are returns of Fama and French (2015) five factors [11]; and FDR_t is the factor mimicking portfolio return for systematic financial distress risk in month t. FDR_t is estimated by a hedge portfolio that takes a long position on BAA

[11] RMW_t denotes the profitability factor and is measured by the return of a profitability factor mimicking portfolio over month t. CMA_t denotes the investment (CMA) factor and is the return of an investment factor mimicking portfolio over month t. Fama and French (2015) show that over the sample period of 1963–2013, adding the profitability and investment factors makes the value factor redundant since the time series of HML returns are completely explained by the other four factors.

rated bonds and a short position in AAA rated bonds. We collect the monthly bond yields from the Federal Reserve's H-15 reports and convert bond yields to returns using the log-linear approximation defined in Campbell et al. (1997).

At the end of every June over our sample period 1980–2017, we sort firms into five equally weighted portfolios based on their most recent *Under*. Portfolio 1 (5) include firms with the lowest (highest) under-investment. Table 3.11 presents the results of time-series regressions of portfolio excess returns on the six systematic risk factors. We do not find evidence that firms with low under-investment have high exposure to the systematic financial distress factor. The coefficients of *FDR* actually increase from 0.0035 for portfolio 1 to 0.0054 for portfolio 5, suggesting that firms with high under-investment have high exposure to the financial distress factor. Our empirical result does not support the alternative risk-premium-based explanation. Furthermore, the alphas of these five portfolios decrease monotonically from portfolio 1 to 5. The annualized return spread between portfolio 1 and 5 is 8.16%. Overall, our results show that the negative relation between under-investment and future stock returns remains robust after controlling for the systematic financial distress risk factor. Therefore, it is more likely that the negative relation between under-investment and future stock returns is due to market inefficiencies, such as delayed reaction to under-investment or potential agency costs.

3.4.4. The impact of market recessions on our results

Our sample covers two notorious stock market downturns: the burst of the internet bubble between 2000 and 2002 and the recent financial crisis between 2007 and 2009. Both stock returns and firm investment strategies are notably affected by these two negative market-level shocks. In untabulated tests, we examine the impact of these two market recessions on the empirical relation between abnormal investment and future stock returns. The results of Fama and MacBeth (1973) regressions show that the relation between under-investment and future stock returns is not statistically significant during these two periods, suggesting that stock markets are less likely to react to negative information with a delay during the recession periods. Outside of these two recession periods, the negative relation between under-investment and future stock returns remain statistically significant.

The previous psychology and economics studies suggest that investors' sensitivity to news is most pronounced when they are going through hard times. In the psychology literature, Smith and Ellsworth (1985) find that people's emotions, such as anxiety, hope, and sadness, are associated with a greater sense of uncertainty. Tiedens and Linton (2001) show that the reliance on heuristic versus systematic processing varies with emotions. Consistent with these findings, the behavioral economics literature suggests that investor behavior differs in times of anxiety and fear versus periods of prosperity and tranquillity (Akerlof and Shiller, 2010; Garcia, 2013). For example, Hirshleifer and Shumway (2003) and Cortés et al. (2016) show that stock returns and credit approval rates are affected by weather. Edmans et al. (2007) also show that stock returns are affected by the outcomes of major sporting events. One potential explanation of our empirical findings is that market investors tend to pay more attention to the negative firm-specific information during the market downturns so that stock prices are more likely to fully reflect the negative firm-specific information during the recession periods.

3.4.5. Additional robustness test results

In this section, we further discuss our robustness test results which are not tabulated in this chapter. First, Fama and French (2008) point out that microcap stocks comprise 60% of the stocks in the U.S. market, but on average only account for 3% of the market capitalization. Microcap stocks also tend to disproportionately inhabit stock return anomalies because the cross-sectional dispersion of anomaly variables tends to be the highest among them. To test whether our main results are driven by small stocks, we either drop stocks with the lowest 20% market capitalization or those with stock prices less than $5 from our sample. Our baseline regression results remain robust.

Second, acquisitions (particularly stock-financed deals) predict poor subsequent stock returns, which could have important implications for stock return prediction models. Since our measure of investment includes acquisitions, it is possible that the periods of measured under-investment immediately follow major deals. To mitigate this concern, we exclude acquisition expenses from our investment measures and find that our baseline regression results remain qualitatively the same.

Third, the abnormal investment variable in Titman et al. (2004) takes negative values for under-investment and positive values for over-investment. In contrast, our abnormal investment variable ($AInvt$) is in its absolute value term, suggesting that large positive values of abnormal investment can reflect either substantial under- or over-investment. We use variables $Under$ and $Over$ to distinguish between under- and over-investment in our empirical analyses. Following Titman et al. (2004), we define Raw_AInvt as the difference between our actual and predicted investment model without taking absolute value. We replace $AInvt$ by Raw_AInvt in columns (1) and (4) of Table 3.4. The coefficient of Raw_AInvt in column (1) is -0.004 and statistically insignificant, suggesting that future stock returns are not negatively related to firm investment. The coefficient of Raw_AInvt in column (4) is 0.338 and statistically significant at the 1% level, suggesting that firms with over-investment tend to outperform those with under-investment.

Fourth, Cooper et al. (2008) document a negative relation between asset growth and subsequent stock returns. Since asset growth captures common return effects across the components of a firm's total investment or financing activities, it can predict cross-sectional stock returns (Cooper et al., 2008). In Columns (4)–(6) of Table 3.4, we have controlled for the proxies of firm total investment ($INew$) and financing activities (*Leverage*). To alleviate the concern that our abnormal investment measures simply capture the asset growth effect, we directly add asset growth as a control variable in these three regression specifications. Following Cooper et al. (2008), asset growth in year t is defined as the percentage change in total assets from fiscal year $t-2$ to fiscal year $t-1$. All the coefficients of our abnormal investment variables remain robust, and the coefficients of asset growth are negative and statistically significant.

3.5. Conclusions

In a standard firm growth model, corporate earnings which are not paid out as dividends will be invested in positive NPV projects. The book value of firm equities will increase accordingly. In a dynamic financial market, a firm's actual investment may deviate from its model predicted investment level due to random economic shocks or managerial discretion. If the deviation is due to random economic shocks, then the gap between the actual and model predicted investment conveys information about the firm's future fundamentals. The return predictability of abnormal investment may be explained by possible market inefficiencies, such as a market delayed reaction to the fundamental information contained in abnormal investment. If the deviation is due to managerial discretion, then abnormal investment may impact future stock returns through the costs of agency problems.

We employ Richardson's (2006) investment model to decompose firm investment into two components: abnormal investment and model predicted investment. We find that when both investment and abnormal investment are considered simultaneously, future stock returns tend to be more closely associated with abnormal investment, rather than investment. More importantly, we find that the negative relation between abnormal investment and future stock returns is mainly driven by under-investment instead of over-investment. We provide weak evidence on a positive relation between over-investment and future stock returns. We then investigate two mechanisms through which under-investment is negatively associated with future stock returns. With respect to *the delayed market reaction channel*, we show that under-investment conveys fundamental information about firms' future profitability, asset growth, and the likelihood of financial distress. The negative relation between under-investment and future stock returns can be partially explained by the market investors' delayed reaction to the fundamental information in under-investment. We then show that the negative relation between under-investment and future stock returns is more pronounced for firms with less external monitoring and higher agency costs, which supports the *the agency cost channel*. Combined, these results support the notion that market inefficiency along with agency costs associated with under-investment helps to explain the negative empirical relation between under-investment and future stock returns. The earlier investment studies show that high (low) corporate investment predicts low (high) future returns (e.g., Titman et al., 2004). Our paper contributes to the previous literature by showing that after adjusting for firm growth opportunities along with other characteristics in our investment model and stock return prediction model, under-investment, instead of over-investment, is negatively associated with future stock returns.

Appendix A

Table A1. Variable definitions

This table provides variable definitions and corresponding data sources. Compustat refers to the Capital IQ from Standard & Poor's database, CRSP refers to the Centre for Research in Security Prices, FF refers to Kenneth French's data library, WRDS refers to the Fama French & Liquidity Factors database on Wharton Research Data Services, and 13F refers to the Thomson Reuters 13F Database.

Variable	Definition	Source
Asset	Total assets (millions).	Compustat
ITotal	Annual total investment expenditure divided by *Asset*: [Capital expenditure (CAPX) + R&D Expenditure (XRD) + Acquisitions (AQC) − Sale of Property, Plant and Equipment (SPPE)]/*Asset* (Richardson, 2006).	Compustat
IMain	Annual required investment expenditure to maintain assets in place divided by *Asset*: Depreciation and Amortization (DPC)/*Asset* (Richardson, 2006).	Compustat
INew	Investment expenditure on new projects divided by *Asset*: *ITotal* − *IMain* (Richardson, 2006).	Compustat
AInvt	Abnormal investment proxy variable: $\lvert INew - \widehat{INew} \rvert$, where $\widehat{INew_t}$ is estimated by a historical panel regression over the period 1974 to year $t-1$.	CRSP & Compustat
Under	Under-investment proxy variable: $\lvert AInvt \rvert$ if $AInvt$<0.	CRSP & Compustat
Over	Over-investment proxy variable: $\lvert AInvt \rvert$ if $AInvt$>0.	CRSP & Compustat
MVE	Market value of equity (millions): Common Outstanding Shares (CSHO) * Stock Price (PRCC_F).	Compustat
V/P	Growth opportunity: Assets in place/*MVE*, where the assets in place are estimated as $(1-\alpha r)BV + \alpha(1+r)X - \alpha r d$, α=$\omega/1+r-\omega$, $r = 12\%$, $\omega = 0.62$, BV is the Book Value of Common Equity (CEQ), X is Operating Income After Depreciation (OIADP), and d is annual Dividend (DVC) (Ohlson, 1995; Richardson, 2006).	Compustat
Leverage	Leverage ratio: [Short-term Debt (DLC) + Long-term Debt (DLTT)]/ [DLC + DLTT + CEQ] (Richardson, 2006).	Compustat
Cash	Cash holdings: Cash and Short-term Investment (CHE) divided by *Asset* (Richardson, 2006).	Compustat

Continued on next page

Table A1 - continued from previous page

Variable	Definition	Source
Age	Firm age: the natural log of (1+ the number of years the firm has been listed on CRSP as of the start of year) (Richardson, 2006).	CRSP
Size	Firm size: the natural log of *Asset* (Richardson, 2006).	Compustat
Return	The percentage change in firm market value over the previous year: MV_t/MV_{t-1} -1 (Richardson, 2006).	CRSP
R_p	The monthly return on quintile portfolio p by abnormal investment proxies.	CRSP & Compustat
MKTRF	The monthly excess return on the market portfolio (Fama and French, 1993).	FF
SMB	The monthly average return on the three small portfolios minus the average return on the three big portfolios (Fama and French, 1993).	FF
HML	The monthly average return on the two value portfolios minus the average return on the two growth portfolios (Fama and French, 1993).	FF
RMW	The monthly average return on the two robust operating profitability portfolios minus the average return on the two weak operating profitability portfolios (Fama and French, 2015).	FF
CMA	The monthly average return on the two conservative investment portfolios minus the average return on the two aggressive investment portfolios (Fama and French, 2015).	FF
FDR	Systematic financial distress risk: monthly return on a hedge portfolio with a long position in BAA rated bonds and a short position on AAA rated bonds. We follow Campbell et al. (1997) to convert yields to returns.	Federal Reserve H-15 reports
BHR	One-year buy-and-hold returns starting from the fourth month after a fiscal year end and we require at least 6 available monthly returns.	CRSP & Compustat
MTB	Market-to-Book ratio: $[MV$ + DLC + DLTT $]/Asset$ (Stoughton et al., 2017).	Compustat
Earnings	Firm earnings: [Income Before Extraordinary Items (IB) + Interest Expense(XINT)] $/Asset$.	Compustat
BVE	Book value of equity (millions): CEQ + Preferred Treasury Stock (TSTKP) − Preferred Dividends In Arrears (DVPA).	Compustat
BTM	Book-to-Market ratio: BVE/MVE.	Compustat
Profit	Profitability: Net Income (NI)/*Asset*.	Compustat

Continued on next page

Table A1 - continued from previous page

Variable	Definition	Source
MVE/Total MV	The natural log of market value of equity to the total market value.	Compustat
AR	Abnormal returns: a firm's buy-and-hold return during the fiscal year subtracted by the value-weighted market index.	CRSP
Volatility	Volatility: the standard deviation of the residuals from the regression of monthly stock returns on the value-weighted market index return.	CRSP
Delist	An indicator variable for performance-rated delisting which equals one if a firm delists within three years of the fourth month after the fiscal year end with a CRSP delisting code 500 or between 520 and 584, and zero otherwise. (Shumway, 2001).	CRSP & Compustat
Δ*Earnings*	The change in earnings: $(Earnings_{t+1} - Earnings_t)/Asset_t$.	Compustat
Δ*Asset*	The change in assets: $(Assets_{t+1} - Assets_t)/Asset_t$.	Compustat
Beta	Beta of a standard market model, using the most recent 255 trading days' returns and CRSP value-weighted index returns as the proxy for market returns.	CRSP & Eventus
Blockholder Ownership	The percentage ownership of blockholders who hold more than 5% of a firm's outstanding shares.	13F
Expense Ratio	Operating expenses divided by total sales. Operating expenses are defined as total expenses less cost of goods sold, interest expense, and managerial compensation (Ang et al., 2000).	Compustat
Asset Utilization Ratio	Total sales divided by total assets (Ang et al., 2000).	Compustat

Table 3.1. Summary statistics

This table presents the summary statistics of the variables used in our main empirical analysis. For the variables included in our investment model Equation (3.1), the sample consists of 122,180 firm–year observations over the period 1974–2017. For the rest of the variables, the sample period is 1980–2017. The number of observations, mean, standard deviation, 1st percentile, 25th percentile, median, 75th percentile, and 99th percentile are reported from left to right, in sequence for each variable. See Appendix A for variable definitions.

Variables	**Obs.**	**Mean**	**S.D.**	**p1**	**p25**	**Median**	**p75**	**p99**
Asset	122,180	2028.2	13000	2.258	33.780	137.9	714.4	33756.2
Itotal	122,180	0.119	0.112	-0.029	0.043	0.088	0.160	0.580
Imain	122,180	0.048	0.034	0.000	0.026	0.04	0.059	0.204
INew	122,180	0.070	0.110	-0.154	0.002	0.042	0.109	0.523
MVE	122,180	1416.8	4589.4	1.313	23.726	114.0	641.0	34164.1
V/P	122,180	0.788	0.706	-0.456	0.342	0.619	1.041	3.727
Leverage	122,180	0.312	0.248	0.000	0.076	0.297	0.497	0.917
Cash	122,180	0.155	0.193	0.000	0.024	0.075	0.212	0.872
Age	122,180	2.427	0.896	0.693	1.792	2.485	3.091	4.290
Size	122,180	5.099	2.136	0.815	3.653	5.075	6.711	10.425
Return	122,180	0.241	0.809	-0.798	-0.182	0.085	0.421	4.549
MTB	108,135	1.473	1.341	0.273	0.720	1.020	1.663	8.391
BHR	108,135	0.156	0.609	-0.825	-0.198	0.069	0.367	2.904
Earnings	95,356	0.014	0.192	-0.769	0.009	0.059	0.092	0.238
BE	95,356	855.6	4360.8	0.954	20.382	86.932	397.1	13519.1
BTM	95,356	0.728	0.874	0.050	0.329	0.562	0.907	3.228
Profit	122,072	0.006	0.200	-0.742	-0.001	0.042	0.080	0.248
AR	122,072	0.058	0.748	-0.875	-0.293	-0.048	0.227	2.649
Volatility	122,072	0.121	0.091	0.027	0.068	0.100	0.149	0.445
Delist	59,706	0.052	0.222	0.000	0.000	0.000	0.000	1.000
Beta	59,706	0.784	0.629	-0.568	0.349	0.734	1.162	2.481
Δ*Earnings*	59,706	0.001	0.029	-0.037	0.000	0.000	0.000	0.045
Δ*Debt*	59,706	0.041	0.325	-0.354	-0.032	0.000	0.047	1.001
Δ*Asset*	59,706	0.110	0.481	-0.438	-0.038	0.046	0.154	1.648
FDR	455	1.539	0.668	0.741	1.073	1.384	1.776	4.447

Table 3.2. Analysis of investment expenditure

Panel A. Coefficients of investment prediction regressions. This panel summarizes the regression coefficients estimated by the investment model Equation (3.1). The dependent variable is new investment expenditure ($INew_t$). The independent variables are growth opportunity (V/P_{t-1}), leverage ($Leverage_{t-1}$), cash holdings ($Cash_{t-1}$), firm age (Age_{t-1}), firm size ($Size_{t-1}$), past stock performance ($Return_{t-1}$), and the lag of new investment expenditure ($INew_{t-1}$). We estimate the investment model using thirty-eight historical panel regressions. For each year t between 1980 and 2017, we run a panel regression using firm–year observations between 1974 and year $t-1$. We only report the time-series average of the coefficients estimated by these historical panel regressions. Year and Fama–French 48 industry fixed effects are controlled in all regressions. The coefficients of the year and industry fixed effects are suppressed for brevity. See Appendix A for variable definitions. The t-values of regression coefficients are based on standard errors clustered by firm. The numbers of positive and negative coefficients at the 1% statistical significance level are reported in parentheses.

Dependent variable: $INew_t$	
V/P_{t-1}	-0.010
	(Negative 38, Positive 0)
$Leverage_{t-1}$	-0.033
	(Negative 38, Positive 0)
$Cash_{t-1}$	0.078
	(Negative 0, Positive 38)
Age_{t-1}	-0.004
	(Negative 38, Positive 0)
$Size_{t-1}$	0.003
	(Negative 0, Positive 38)
$Return_{t-1}$	0.012
	(Negative 0, Positive 38)
$INew_{t-1}$	0.409
	(Negative 0, Positive 38)
Constant	0.037
	(Negative 0, Positive 38)
Average Observations	70,598
Average adj. R^2	0.342
Year fixed effects	Yes
Industry fixed effects	Yes
Number of historical panels	38

Panel B. Descriptive statistics of abnormal investment. This panel presents the descriptive statistics of the abnormal investment variables estimated by the investment model Equation (3.1). The main sample covers 108,273 firm–year observations over the period 1980–2017. The number of observations, mean, standard deviation, 1st percentile, 25th percentile, median, 75th percentile, and 99th percentile are reported from left to right, in sequence for each variable. See Appendix A for variable definitions.

Variables	**Obs.**	**Mean**	**S.D.**	**p1**	**p25**	**Median**	**p75**	**p99**
$INew^*$	108,273	0.069	0.064	-0.039	0.026	0.055	0.099	0.285
$INew$ - $INew^*$	108,273	0.000	0.086	-0.195	-0.043	-0.011	0.027	0.311
$AInvt$	108,273	0.057	0.064	0.001	0.017	0.037	0.073	0.319
$Under$	63,932	0.049	0.046	0.001	0.018	0.036	0.065	0.221
$Over$	44,341	0.070	0.082	0.001	0.015	0.039	0.092	0.384

Table 3.3. Quintile portfolio analysis

This table presents the results of the quintile portfolio analysis of the relation between firm abnormal investment and stock returns. We sort stocks into quintile portfolios based on the most recently estimated *AInvt*, *Under*, and *Over* at the end of June for each year from 1980–2017. Portfolio 1 (5) includes stocks with the lowest (highest) abnormal investment measures. For each quintile portfolio, we calculate its equally weighted monthly returns. Then we estimate its Fama and French (1993) (FF) three-factor model alpha and Fama and French (2015) five-factor model alpha based on monthly portfolio returns. The last column reports the corresponding factor model alpha of a portfolio that takes a long position on stocks in portfolio 1 and a short position on stocks in portfolio 5. The t-statistics are reported in brackets. $***$, $**$, and $*$ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Quintile portfolios					
Factor models	**1 (Lowest)**	**2**	**3**	**4**	**5 (Highest)**	**Long 1, Short 5**
			Stocks sorted by *AInvt*			
FF three-factor alpha	0.37%***	0.28%***	0.30%***	0.14%	0.07%	0.30%**
	[4.76]	[3.57]	[3.66]	[1.62]	[0.60]	[2.21]
FF five-factor alpha	0.32%***	0.24%***	0.31%***	0.20%**	0.25%**	0.07%
	[3.97]	[2.98]	[3.70]	[2.23]	[2.21]	[0.47]
			Stocks sorted by *Under*			
FF three-factor alpha	0.33%***	0.19%**	0.23%***	0.07%	-0.32%***	0.65%***
	[4.22]	[2.44]	[2.73]	[0.81]	[-2.58]	[4.56]
FF five-factor alpha	0.28%***	0.16%**	0.24%***	0.13%	-0.11%	0.38%***
	[3.42]	[1.97]	[2.80]	[1.32]	[-0.86]	[2.44]
			Stocks sorted by *Over*			
FF three-factor alpha	0.38%***	0.41%***	0.40%***	0.43%***	0.49%***	-0.11%
	[4.29]	[4.05]	[4.04]	[4.15]	[3.84]	[-0.72]
FF five-factor alpha	0.34%***	0.33%***	0.44%***	0.53%***	0.59%***	-0.25%
	[3.78]	[3.18]	[4.24]	[4.88]	[4.48]	[-1.47]

Table 3.4. Investment, abnormal investment, and stock performance

This table presents the Fama and MacBeth (1973) regression results of firm future stock returns on abnormal investment proxy variables. The dependent variable is BHR_{t+1}, one-year buy-and-hold returns starting from the fourth month after the end of fiscal year t. The independent variables of interest are the abnormal investment proxies, estimated by the investment expenditure Equation (3.1): $AInvt_t$, $Under_t$, and $Over_t$. See Appendix A for variable definitions. The t-statistics of Fama–MacBeth regression coefficients are reported in brackets. The coefficients of the Fama–French 48 industry fixed effects are suppressed for brevity in the respective columns. $***$, $**$, and $*$ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
$AInvt_t$	-0.165***			-0.135**		
	[-3.62]			[-2.65]		
$Under_t$		-0.249***			-0.447***	
		[-3.20]			[-4.54]	
$Over_t$			-0.094*			0.352*
			[-1.73]			[1.73]
$INew_t$				-0.053	-0.400***	-0.399**
				[-1.25]	[-2.85]	[-2.43]
MTB_t	-0.032***	-0.034***	-0.034***	-0.031***	-0.031***	-0.029***
	[-5.02]	[-5.27]	[-5.05]	[-4.90]	[-4.96]	[-4.06]
$Leverage_t$	-0.031	-0.031	-0.025	-0.032	-0.041	-0.040
	[-1.19]	[-1.10]	[-0.83]	[-1.21]	[-1.45]	[-1.42]
$Cash_t$	0.048	0.055	0.037	0.046	0.074**	0.088*
	[1.24]	[1.61]	[0.63]	[1.27]	[2.39]	[1.89]
$Size_t$	-0.002	-0.002	-0.003	-0.001	-0.000	-0.002
	[-0.34]	[-0.49]	[-0.65]	[-0.30]	[-0.04]	[-0.43]
BHR_t	-0.033	-0.031	-0.040	-0.032	-0.028	-0.045
	[-1.24]	[-1.26]	[-1.38]	[-1.22]	[-1.10]	[-1.54]
Constant	0.243***	0.143**	0.197**	0.244***	0.135**	0.189**
	[3.43]	[2.58]	[2.51]	[3.40]	[2.37]	[2.40]
Observations	108,135	63,856	44,279	108,135	63,856	44,279
Average adj. R^2	0.107	0.121	0.145	0.109	0.126	0.149
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of groups	38	38	38	38	38	38

Table 3.5. Information in abnormal investment about future earnings and asset growth

This table presents the panel regression results of the change in earnings and change in assets on abnormal investment proxy variables. Our sample consists of 95,356 firm–year observations with available data for the analysis during 1980–2017. The dependent variable in columns (1)–(3) is the change in firm earnings over a one-year horizon normalized by total assets: $\Delta Earnings_{t\ to\ t+1}$=$(Earnings_{t+1}-Earnings_t)/Assets_t$. The dependent variable in columns (4)–(6) is the change in firm total assets over a one-year horizon normalized by total assets: $\Delta Assets_{t\ to\ t+1}$=$(Assets_{t+1}-Assets)_t/Assets_t$. The independent variables of interest are firm abnormal investment proxies: $AInvt_t$, $Under_t$, and $Over_t$. See Appendix A for variable definitions. Year and Fama–French 48 industry fixed effects are controlled for in all regressions. The t-statistics reported in brackets are clustered by firm. $***$, $**$, and $*$ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Δ**Earnings**$_{t\ to\ t+1}$			Δ**Assets**$_{t\ to\ t+1}$		
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
$AInvt_t$	-0.011***			0.163***		
	[-3.77]			[4.93]		
$Under_t$		-0.013**			-0.136**	
		[-2.30]			[-2.36]	
$Over_t$			-0.005**			0.208***
			[-1.97]			[5.11]
BTM_t	-0.002***	-0.003***	-0.001***	-0.111***	-0.103***	-0.116***
	[-9.12]	[-7.10]	[-5.71]	[-23.73]	[-23.80]	[-12.63]
$Size_t$	0.000***	0.000**	0.000**	-0.013***	-0.012***	-0.015***
	[3.57]	[2.50]	[2.46]	[-10.65]	[-7.23]	[-9.10]
$Leverage_t$	0.002***	0.000	0.004***	-0.096***	-0.100***	-0.109***
	[3.38]	[0.42]	[3.13]	[-11.00]	[-9.59]	[-7.72]
$Earnings_t$	-0.046***	-0.056***	-0.034***			
	[-9.74]	[-6.84]	[-12.06]			
Constant	0.004***	0.005**	0.002***	0.319***	0.295***	0.370***
	[3.12]	[2.18]	[2.98]	[14.80]	[11.38]	[13.12]
Observations	95,356	56,190	39,166	95,356	56,190	39,166
Average adj. R^2	0.071	0.087	0.053	0.0405	0.0363	0.0474
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Table 3.6. Abnormal investment and future financial distress

This table reports the logit regression results (marginal effect reported) of future financial distress on abnormal investment proxy variables. Our regression design follows Shumway's (2001) bankruptcy prediction model. The dependent variable is $Delist_{t\ to\ t+3}$, an indicator variable that equals one if a firm is delisted in the next three years due to performance reasons, and zero otherwise. The independent variables of interest are firm abnormal investment proxies: $AInvt_t$, $Under_t$, and $Over_t$. See Appendix A for variable definitions. Year and Fama–French 48 industry fixed effects are controlled for in all regressions. The z-values reported in brackets are clustered by firm. $***$, $**$, and $*$ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Variables	**(1)**	**(2)**	**(3)**	**(4)**
$AInvt_t$		0.019***		
		[2.66]		
$Under_t$			0.065***	
			[4.95]	
$Over_t$				0.002
				[0.24]
$Profit_t$	-0.022***	-0.024***	-0.023***	-0.028***
	[-11.40]	[-9.42]	[-6.88]	[-6.76]
$Leverage_t$	0.052***	0.053***	0.061***	0.047***
	[18.62]	[16.45]	[14.54]	[9.54]
$MVE/Total\ MV_t$	-0.007***	-0.008***	-0.009***	-0.009***
	[-15.54]	[-15.27]	[-11.96]	[-9.60]
AR_t	-0.030***	-0.032***	-0.035***	-0.029***
	[-16.29]	[-15.78]	[-13.22]	[-9.06]
$Volatility_t$	0.029***	0.031***	0.034***	0.026***
	[5.75]	[5.02]	[3.84]	[3.62]
$Size_t$	-0.008***	-0.007***	-0.007***	-0.006***
	[-12.50]	[-9.56]	[-8.25]	[-5.41]
Observations	122,072	94,962	55,987	36,378
Pseudo R^2	0.229	0.229	0.238	0.219
Area under ROC curve	0.874	0.874	0.877	0.869
Year fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes

Table 3.7. Under-investment and stock returns: Controlling for fundamentals

This table presents the Fama and MacBeth (1973) regression results of annual buy-and-hold returns on under-investment and firm future fundamentals. The dependent variable is buy-and-hold stock return BHR_t, 12-month buy-and-hold returns starting from the fourth month after the fiscal year t end. The independent variable of interest is $Under_t$. Firm fundamentals include the change in earnings ($\Delta Earnings_{t\ to\ t+1}$), the change in assets ($\Delta Assets_{t\ to\ t+1}$), performance related delist indicator variable ($Delist_{t\ to\ t+3}$). We follow Caskey et al. (2012) and use the following three control variables: book-to-market (BTM_t), stock beta ($Beta_t$), and firm size ($Size_t$). See Appendix A for variable definitions. The t-statistics of Fama–MacBeth regression coefficients are reported in brackets. $***$, $**$, and $*$ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
$Under_t$	-0.238***	-0.187**	-0.134*	-0.132*	-0.126
	[-2.97]	[-2.41]	[-1.73]	[-1.72]	[-1.64]
$\Delta Earnings_{t\ to\ t+1}$		4.022***			3.410***
		[4.78]			[4.36]
$\Delta Assets_{t\ to\ t+1}$			0.180***		0.165***
			[9.86]		[9.21]
$Delist_{t\ to\ t+3}$				-0.421***	-0.322***
				[-20.34]	[-13.51]
BTM_t	0.060***	0.072***	0.089***	0.072***	0.094***
	[4.93]	[5.25]	[6.35]	[5.87]	[6.72]
$Beta_t$	0.041*	0.044*	0.033	0.038	0.032
	[1.74]	[1.82]	[1.37]	[1.65]	[1.35]
$Size_t$	-0.002	-0.007	-0.005	-0.010*	-0.008
	[-0.34]	[-1.24]	[-0.85]	[-1.82]	[-1.38]
Constant	0.110	0.147	0.065	0.174	0.090
	[1.04]	[1.49]	[0.65]	[1.64]	[0.89]
Observations	59,768	53,403	53,403	59,768	53,403
Average adj. R^2	0.103	0.118	0.126	0.130	0.142
Industry fixed effects	Yes	Yes	Yes	Yes	Yes
Number of groups	35	35	35	35	35

Table 3.8. Abnormal investment and stock returns: Agency costs

This table reports the cross-sectional relation between under-investment, agency costs, and annual buy-and-hold returns. In the Fama and MacBeth (1973) regressions, the dependent variable is buy-and-hold stock return BHR_{t+1}, 12-month buy-and-hold returns starting from the fourth month after the fiscal year t end, and the independent variable of interest is $Under_t$. We divide firm–year observations with under-investment into two sub-samples based on the annual industry medians of *Blockholder Ownership*, *Expense Ratio*, and *Asset Utilization Ratio*. The high (low) sub-samples include firm–year observations with above(below and equal to) median corresponding variables. The control variables are the same as those reported in Table 3.4. See Appendix A for variable definitions. Fama–French 48 industry fixed effects are controlled for in all regressions. The t-statistics of Fama–MacBeth regression coefficients are reported in brackets. $***$, $**$, and $*$ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Blockholder Ownership$_t$		***Expense Ratio$_t$***		***Asset Utilization Ratio$_t$***	
	Low	**High**	**Low**	**High**	**Low**	**High**
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
$Under_t$	-0.455***	-0.078	-0.272**	-0.404***	-0.472***	-0.263**
	[-3.18]	[-0.44]	[-2.07]	[-2.93]	[-4.74]	[-2.24]
$INew_t$	-0.343**	0.093	-0.237**	-0.303*	-0.279***	-0.217
	[-2.57]	[0.32]	[-2.27]	[-1.97]	[-2.84]	[-1.66]
MTB_t	-0.029***	-0.019**	-0.019***	-0.034*	-0.029***	-0.034***
	[-4.12]	[-2.37]	[-3.53]	[-1.95]	[-4.71]	[-4.83]
$Leverage_t$	0.012	-0.007	0.001	-0.076	-0.031	-0.001
	[0.34]	[-0.19]	[0.05]	[-1.35]	[-1.10]	[-0.02]
$Cash_t$	0.095**	0.012	0.064*	0.084***	0.110***	0.115**
	[2.50]	[0.23]	[1.69]	[2.87]	[4.50]	[2.68]
$Size_t$	-0.006	-0.000	-0.004	0.018	0.006	-0.005
	[-1.23]	[-0.03]	[-0.95]	[0.87]	[1.20]	[-1.04]
BHR_t	-0.053*	-0.046	-0.062**	0.015	-0.040	-0.016
	[-1.87]	[-1.36]	[-2.09]	[0.26]	[-1.30]	[-0.47]
Constant	0.258***	0.240**	0.190***	0.052	0.206**	0.169**
	[2.80]	[2.04]	[2.91]	[0.23]	[2.51]	[2.09]
Observations	22,562	20,566	33,542	29,913	32,470	31,317
Average adj. R^2	0.180	0.194	0.164	0.157	0.163	0.154
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of groups	38	38	38	38	38	38

Table 3.9. Abnormal investment and future stock returns: Alternative econometric estimation methods

This table presents the robustness test results of future stock return on three abnormal investment proxies between 1980 and 2017. The dependent variable is BHR_{t+1}, 12-month buy-and-hold returns starting from the fourth month after the fiscal year t end. The independent variables of interest are $AInvt_t$, $Under_t$, and $Over_t$ in Panels A, B, and C, respectively. In column (1), the abnormal investment proxies are estimated by Equation (3.1) with the historical panels starting from 1974. In columns (3)–(4), the abnormal investment proxies are estimated by Equation (3.1) with the whole panel period of 1974–2017. In columns (5)–(6), the abnormal investment proxies are estimated by Equation (3.1) with five-year rolling windows between year $t-4$ and year t. The other independent variables are the same as those reported in Table 3.4. See Appendix A for variable definitions. In columns (1)–(2) and (4), we use panel regressions to estimate the relation between abnormal investment and future stock returns. In columns (3) and (5), we use a Fama and MacBeth (1973) regression to estimate the relation between abnormal investment and future stock returns. In columns (1), (2), and (4), we control for year and Fama–French 48 industry fixed effects and cluster standard errors by year and industry (Petersen, 2009). The t-statistics of panel and Fama–MacBeth (FM) regression coefficients are reported in brackets. $***$, $**$, and $*$ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Abnormal investment

	Historical panel	***Whole panel***		***Rolling panel***	
	Panel	***Panel***	***FM***	***Panel***	***FM***
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
$AInvt_t$	-0.147***	-0.163***	-0.143***	-0.146***	-0.133**
	[-4.03]	[-4.46]	[-2.81]	[-3.95]	[-2.67]
Control variables	Yes	Yes	Yes	Yes	Yes
Observations	108,087	108,135	108,135	108,135	108,135
Average adj./Adj. R^2	0.119	0.119	0.109	0.119	0.109
Year fixed effects	Yes	Yes	No	Yes	No
Industry fixed effects	Yes	Yes	Yes	Yes	Yes
Number of groups			38		38

	Panel B. Under-investment				
	Historical panel	***Whole panel***		***Rolling panel***	
	Panel	***Panel***	***FM***	***Panel***	***FM***
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
$Under_t$	-0.451***	-0.420***	-0.421***	-0.299***	-0.404***
	[-6.65]	[-6.18]	[-4.05]	[-4.37]	[-3.77]
Control variables	Yes	Yes	Yes	Yes	Yes
Observations	63,825	62,830	62,830	62,578	62,578
Average adj./Adj. R^2	0.107	0.117	0.128	0.118	0.125
Year fixed effects	Yes	Yes	No	Yes	No
Industry fixed effects	Yes	Yes	Yes	Yes	Yes
Number of groups			38		38

	Panel C. Over-investment				
	Historical panel	***Whole panel***		***Rolling panel***	
	Panel	***Panel***	***FM***	***Panel***	***FM***
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
$Over_t$	0.296***	0.220**	0.339	0.185**	0.282**
	[3.28]	[2.41]	[1.66]	[2.12]	[2.09]
Control variables	Yes	Yes	Yes	Yes	Yes
Observations	44,262	45,305	45,305	45,557	45,557
Average adj./Adj. R^2	0.140	0.125	0.148	0.122	0.143
Year fixed effects	Yes	Yes	No	Yes	No
Industry fixed effects	Yes	Yes	Yes	Yes	Yes
Number of groups			38		38

Table 3.10. Alternative measures of abnormal investment

This table presents the Fama and MacBeth (1973) regression results of firm future stock returns on abnormal investment proxy variables. The dependent variable is BHR_{t+1}, 12-month buy-and-hold returns starting from the fourth month after the fiscal year t end. The independent variables of interest are the abnormal investment estimated by investment models developed in Harvey et al. (2004) and Titman et al. (2004). In columns (1)–(3), the abnormal investment is defined as the absolute value of the difference between a firm's capital investment expenditure and its industry median investment level. In columns (4)–(6), the abnormal investment is defined as the absolute value of the difference between a firm's capital investment expenditure at year $t-1$ and its average capital investment expenditure in the past three years. See Appendix A for variable definitions. Fama–French 48 industry fixed effects are controlled for in all regressions. The t-statistics of Fama–MacBeth regression coefficients are reported in brackets. $***$, $**$, and $*$ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Harvey et al. (2004)			***Titman et al. (2004)***		
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
$AInvt_t$	-0.250***			-0.013		
	[-2.85]			[-1.49]		
$Under_t$		-0.466**			-0.023*	
		[-2.27]			[-1.80]	
$Over_t$			-0.281*			-0.014
			[-1.95]			[-1.41]
$INew_t$	-0.059	0.003	-0.113**	-0.086**	-0.055	-0.011
	[-1.46]	[0.06]	[-2.63]	[-2.28]	[-1.55]	[-0.10]
MTB_t	-0.031***	-0.038***	-0.027***	-0.029***	-0.028***	-0.031***
	[-4.52]	[-4.89]	[-4.23]	[-3.86]	[-3.44]	[-4.12]
$Leverage_t$	-0.034	-0.009	-0.047*	-0.033	-0.031	-0.019
	[-1.31]	[-0.30]	[-1.81]	[-1.26]	[-1.13]	[-0.57]
$Cash_t$	0.040	0.050	0.063	0.021	0.039	0.029
	[1.13]	[1.49]	[1.31]	[0.63]	[1.03]	[0.53]
$Size_t$	-0.001	-0.004	-0.000	-0.002	-0.002	-0.008
	[-0.29]	[-0.78]	[-0.04]	[-0.47]	[-0.39]	[-1.10]
BHR_t	-0.032	-0.042	-0.030	-0.043	-0.052*	-0.010
	[-1.24]	[-1.61]	[-1.06]	[-1.59]	[-1.82]	[-0.27]
Constant	0.284***	0.364***	0.245***	0.216***	0.247**	0.163*
	[4.38]	[3.76]	[2.72]	[2.94]	[2.44]	[1.83]
Observations	108,402	54,943	54,602	81,566	46,528	35,038
Average adj. R^2	0.109	0.128	0.139	0.116	0.129	0.161
Number of groups	38	38	38	38	38	38
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Table 3.11. Under-investment and systematic distress risk

This table presents the time-series regressions of portfolio returns controlling for six risk factors. At each end of June over the period 1980–2017, firms are divided into five equally weighted portfolios based on *Under*. The dependent variable is the monthly excess returns of the corresponding under-investment quintile portfolio in year t. The independent variables are the returns of factor mimicking portfolios: market ($MKTRF$), size (SMB), book-to-market (HML), profitability (RMW), investment (CMA), and financial distress risk (FDR). The financial distress risk factor return is the return of a mimicking portfolio that longs BAA bonds and shorts AAA bonds. We download monthly bond yield data from the Federal Reserve's H-15 report and convert bond yields to returns using the log-linear approximation defined in Campbell et al. (1997). The t-statistics are presented in the brackets. $***$, $**$, and $*$ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Under **portfolio**	**Alpha**	***MKTRF***	***SMB***	***HML***	***RMW***	***CMA***	***FDR***	**Obs.**	**Adj.** R^2
1 (Lowest)	-0.0026	1.0033	0.7385	0.2733	0.0010	0.0002	0.0035	455	0.9069
	[-1.33]	[51.01]	[25.05]	[7.41]	[2.59]	[0.38]	[3.03]		
2	-0.0033	0.9792	0.6860	0.2933	0.0007	-0.0002	0.0031	455	0.9054
	[-1.72]	[50.82]	[23.75]	[8.12]	[1.83]	[-0.31]	[2.79]		
3	-0.0039	1.0024	0.7689	0.2968	-0.0001	-0.0006	0.0041	455	0.9031
	[-1.88]	[47.92]	[24.52]	[7.57]	[-0.25]	[-1.11]	[3.35]		
4	-0.0047	1.0129	0.7917	0.1473	-0.0011	0.0000	0.0039	455	0.8953
	[-2.08]	[44.13]	[23.01]	[3.43]	[-2.54]	[-0.05]	[2.89]		
5 (Highest)	-0.0094	1.0049	0.8537	0.1131	-0.0041	-0.0012	0.0054	455	0.8634
	[-3.21]	[33.90]	[19.21]	[2.04]	[-7.19]	[-1.39]	[3.13]		

Chapter 4

Investment of product market peers and the value of cash holdings

4.1. Introduction

In an efficient market, all relevant information should be considered when managers make corporate decisions. Recent studies find that product market peers' information and decisions, such as their stock price, dividend policy, capital structure, and investment, complement the information set of the focal firms, and thus influence their decision-making. The benefit of holding cash equates the firms' ability of financing expected investment opportunities in future. Thus, a firm's cash policy is typically assumed to be determined as a function of the joint distribution of investment opportunities and cash flows over time (Duchin, 2010). Bikhchandani et al. (1998) argue that learning by observing the behavior of others has long implications and impact in economics and business strategies. Fresard (2010) consider that a firm's cash reserves will decrease its product market rivals' future market shares when rivals are financially constrained and when the competition is intensive. Bustamante and Frésard (2017) find that a firm's investment responds positively to its peer firms' investment in the same industry. Grennan (2019) argues that corporate dividend payments have peer effects when raising the dividend payments instead of lowering it in the same product market. Chen et al. (2019) find that managers consider their peer firms' average cash holdings when deciding on their own cash holdings. Bernard et al. (2020a) indicate that peer firms' information can facilitate own firms' investment and product decisions. Fairhurst and Nam (2020) suggest that peer effects in capital structure choice exists to those firms that are under weak external corporate governance. While these studies provide comprehensive evidence on the information-driven corporate policy synchronicity, the discussion about the impact of peer firms' decisions on the focal firms' valuation of cash holdings is less well-understood.

Holding cash on the balance sheet could both add benefits and incur costs to a firm. On the one hand, cash holdings are used as an efficient preemptive weapon for weeding out rivals' entry or finance competitive strategies to gain future market shares (Fresard, 2010; Benoit, 1984). External financing is costly while internally generated cash flows can be volatile. Therefore, preserving highly liquid assets, such as cash holdings, is necessary to ensure timely funding of necessary spending (Denis and Sibilkov, 2010). On the other hand, the risk-free nature of cash implies a minimum return. By holding the asset in the form of cash, firms essentially give up the investment opportunities that could have been value-enhancing to the shareholders. More importantly, the highly liquid nature of cash indicates that its use is mostly at the manager's discretion. Higher cash holding could easily lead to exacerbation of agency cost and detrimental effect on the shareholders' wealth (Myers and Rajan (1998); Dittmar and Mahrt-Smith (2007))

As Bustamante and Frésard (2020) show that corporate managers are unable to perfectly observe the value of future investment opportunities, the firms might follow their peer firms to invest in those similar new projects with a positive NPV. Thus, one expects the firms' value of cash holdings to be positively associated with the investment of their peer firms because of preemptive motives in front of better future investment opportunities. We define peer investment of a firm i as the average investment (i.e., Capital Expenditure scaled by lagged Property, Plant and Equipment) of all its non-local product market peers.

The increasing investments made by the product market peers lead to strategic response from firms' managers. With increasing uncertainty, the role of cash reserve as a "buffer" for spending will become more valuable. Therefore, we can expect the value of corporate cash holding to increase as a result of higher peer firms' investment. Meanwhile, managers value their reputation in the labor market. When their compensation is based on the performance relative to their product market peers, they are incentivized to imitate peers' decisions, even if this means ignoring important private information obtained internally (Scharfstein and Stein, 1990). If peer firms' increasing investment triggers the managerial reputation related to inefficient use of cash, the firms' value of cash holding will be lower.

Our analysis for value of cash holding is based on the framework developed by Faulkender and Wang (2006b). Using an US sample of 49,544 firm-year observation between 1996 and 2017, we find that for a given firm, an increase in peer investment would lead to a higher value of cash holding. This effect is both statistical and economically significant and a one standard deviation increase of peer investment would lead to \$0.14 higher value for \$1 of cash. In comparison, the same one standard deviation change in firms' own investment will lead to about \$0.15 higher value of cash holding.

We applied several methods to ensure that our results are not driven by potential endogeneity issues. First, following Bustamante and Frésard (2020) , we use the average investment made by peer firms' local neighbor firms that are unrelated to the focal firm as the instrumental variable for peer investment to conduct a two-stage least square regression. This instrumental variable satisfies the relevance condition because due to local knowledge externalities, the investments made by firms that in the same local area are correlated. It also satisfies the exclusion condition because we exclude any peer firms' local neighbors that might directly or indirectly relate to the decision made by the focal firms. Second, to further control for potential omitted time-invariant factors that might influence both peer investment and firms' value of cash holding, we conducted additional analyses by controlling for firm fixed effect and industry year dummies. Lastly, to make sure that the effect can only be attributed to the investment made by a firm's product market peer instead of any random firms that invest in the same time, we conducted a placebo test by randomly selecting a set of firms as pseudo-firms in the regression. The application of these approaches doesn't change our main findings that peer investment would increase the firms' value of cash holdings.

We also conducted a battery of robustness tests to further validate our findings. For example, our definition of peers is based on the TNIC classification developed by Hoberg and Phillips (2015) while some of the previous studies rely on the SIC code to define peers, to make sure that our analyses are not sensitive to the definition of peers, we also conducted the tests for the 3-digit SIC defined peers. We also tested our results against alternative definitions of unexpected change in cash holdings, and a different definition of cash holdings. Our results remain robust to all these tests.

While our baseline findings confirm the dominating benefit effect of holding cash in response to peer investment, the channel through which the effect manifests itself is still unclear. One of the potential explanations of our findings is peer learning. In reality, decisions made by other agents are frequently viewed as a form of endorsement and we frequently change our own decisions based on others' decisions. For example, we would assume that the restaurant with more consumers to be a better one (Bikhchandani et al., 1998). Similarly, the action taken by other firms might also be used by a focal firm to guide their own decision-making. If a firm and its shareholders believe that the peer firms are in possession of private information unknown to market, they would assume that the investment behavior made by the peer firms is a confirmation of better investment opportunities. Therefore, when peer firms increase the investment, focal firms' value of cash holdings would increase. Previous studies argue that firms that are relatively smaller, producing more similar products, and with worse information environment are more likely to be learners of the information (Bustamante and Frésard, 2020; Foucault and Fresard, 2014). Therefore, if the learning effect is the main driver of our findings, we should observe this effect to be stronger for these firms. Our empirical tests don't support this hypothesis as we don't find that the effect of peer investment on the firms' value of cash holding is more prominent for these firms.

The relation between peer investment and the firms' value of cash holding can also be a result of the industry structure dynamic. First, the investment expansion of a peer firm may lead to positive externalities which can spill over to the whole industry. For example, a mobile phone producer expands their production, the increasing demand and innovation will drive the cost of chipset down, therefore, benefit the rest of producers that didn't expand. Capital investment associated with employee training and technological innovation could also be beneficial to the whole industry. Since these activities are broadly beneficial to the shareholders, the value of cash holding would improve consequently. Second, peer investment also has the potential to escalate the competition. When peer firms boost their investment, their production costs would drop which give them an edge in the potential price war. Feeling threatened by peer investment, the precautionary demand for cash will rise so the firms' value of cash will increase accordingly.

These two mechanisms will have different implications for industries in different life stages. For a growing industry with plenty of investment opportunities, the potential for capacity expansion is high. Therefore, the positive externality that comes from peer investment could be dominant. In comparison, for a mature industry with fewer investment opportunities, the aggregate market size of the whole industry is relatively stable. The escalation of competition led by peer investment is more likely to be a concern and a reason for a higher value of the shareholding. To test these two hypotheses, we divide our sample into the growing industries (with high Tobin's Q and low average firm age) and the mature industries (with low Tobin's Q and high average firm age). We find that the positive effect of peer investment on the firms' value of cash holdings is predominantly concentrated in the growing industries. These results imply that the positive externality led by peer investment is more likely to be the driver of our main findings.

Our further analyses also provide evidence for this claim. When breaking down the use of cash into dividend payments, capital investment, R&D spending, and advertisement expense, we find that peer investment would lead to a reduction of dividend payments and an increase in both capital and R&D investment. Meanwhile, we find no evidence that the use of cash for advertisement spending increases. These findings are consistent with the picture that firms utilize the positive externality brought by peer investment so as to increase the value of cash. And firms are on average unlikely to expand the advertisement spending to defend their market shares.

Our contribution to the literature is twofold. First, it advances our understanding of firms' responses to peer firms' new information. Most of the extant studies on peer effects underscore the consequent corporate decision changes resulting from the new information contained either in the peer firms' stock price (Faulkender and Wang, 2006b) or in peer firms' like-for-like corporate decisions, such as capital structure, dividend policy, and investment (Leary and Roberts, 2014; Grennan, 2019; Bustamante and Frésard, 2020). In comparison, by highlighting the effect of peer firms' investment decisions on the valuation of corporate cash holdings, we provide novel evidence that peer firms' policies may have an implication broader than previously believed: one set of peers' decisions may provide information for a different set of decisions of focal firms. Second, this study also supplements research related to the value of cash holding. The benefits of preserving cash have been widely recognized by previous studies. However, the majority of these papers would regard cash holdings as funds prepared for the "rainy day", for example, when firms are under financial constraints (Denis and Sibilkov, 2010), highly leveraged (Faulkender and Wang, 2006b), or when the capital market frictions amplify the difficulty for external financing (Bates et al., 2018; Drobetz et al., 2010). By documenting the observational-learning driven positive relationship between peer investment and the firms' value of cash holdings, we highlight the value of cash as the reflection of pursuing future growth opportunities.

This paper is closely related to the research by Bustamante and Frésard (2020), which identifies the positive relationship between firms' investment and peers' investment. Our study is different in two main aspects. First, instead of focusing on the interplays of investment decisions between different firms, our study investigates the effect of peer investment on the firms' value of cash holdings. While the two issues are related, the value of cash holdings has its own unique cost-benefit drivers and, therefore, is worth of taking independent investigation. Second, similar to Bustamante and Frésard (2020), this study relates to the information value of peers' investment. However, while their goal is mainly to answer whether firms learn the information contained in peer investment, our study builds on their conclusion and tries to explain how the information influences the market's perceived value of cash holdings.

This remainder of the paper proceeds as follows. Section 4.2 illustrates the research desgin and sample selection process. Section 4.3 demonstrates the main results and robustness tests results. Section 4.4 discusses the main economic drivers of our main findings. Section 4.5 concludes.

4.2. Research design and sample selection

4.2.1. Define investment of product market peers

In our paper, for each firm, the peer firms are defined as other firms that produce similar products. The similarity of the products is based on the Test-based Network Industry Classification (TNIC) developed by Hoberg and Phillips (2015). By utilizing the information provided in the product description sessions of 10K fillings, a textual analysis was conducted to quantify the level of similarity of products between each pair of firms. Intuitively, when two firms using more common words in their product descriptions, the level of similarity of their products are likely to be higher. A minimum level of similarity threshold is then applied to match the fraction of TNIC peer firms in all firm-pairs with the fraction that could have been calculated from a 3-digit SIC code. In each year t, for each firm i, all firms with a product description similarity above the threshold are classified as firm i's product market peers (Hoberg and Phillips, 2015).

Following Bustamante and Frésard (2020), for each firm i, we exclude all the local peer firms in calculating the peer investment. Specifically, if firm i's headquarter is located in a metropolitan statistical area A in year t, then all the other firms from A will be excluded when calculating the peer investment in that year. This is mainly to avoid the concern of superfluous relation caused by common local factors, instead of common product market. By excluding the local product market peers, both the number of peer firms identified and the similarity between a firm i and its peers are likely to reduce so that our estimations will become more conservative. To sum up, the peer investment of a firm i is calculated as the average investment (i.e., Capital Expenditure scaled by lagged Property, Plant and Equipment) of all its non-local product market peers.

4.2.2. The model for the value of cash holdings

Similar to an extensive list of previous studies, this paper applies the framework propelled by Faulkender and Wang (2006b) to quantify the marginal value of cash holding. In this approach, an OLS model is estimated to identify the effect of the increase in cash on the market value of the firm. To capture the influence of peer investment on the firms' value of cash holding, we augmented the model by adding an interaction term between peer investment and the change in cash holding. Our model can be described as the following equation:

$$
\begin{aligned}
r_{i,t} - R^B_{i,t} = \alpha + \beta_1 * (p\_Capx_{i,t} * \frac{\Delta C_{i,t}}{MVE_{i,t-1}}) + \beta_2 * \frac{\Delta C_{i,t}}{MVE_{i,t-1}} + \beta_3 * p\_Capx_{i,t} + \gamma' Firm_{i,t} \\
+ \lambda' Peer_{i,t} + \delta' \mu_j + \phi' \nu_t + \epsilon_{i,t}
\end{aligned}
\tag{4.1}
$$

where the dependent variable $r_{i,t} - R^B_{i,t}$ is the excess return of firm i in year t in excess to the return of one of Fama-French 25 portfolios, consisting of firms with similar size and book-to-market ratios. Since the excess return can be approximately interpreted as the percentage change in the market value of equity (MVE) in year t ($\Delta MVE_{i,t}/MVE_{i,t-1}$) and the change of cash ($\Delta C_{i,t}$) in the model is scaled by the similar factor ($MVE_{i,t-1}$); the coefficient β_2 can be interpreted as the effect of \$1 change in the cash holding on the dollar value of the firm i's firm value and is defined as the marginal value of cash holdings. We extended the model by including the interaction term of peer investment and the change in the cash ($p\_Capx_{i,t} * \frac{\Delta C_{i,t}}{MVE_{i,t-1}}$), and the coefficient β_1 identifies the effect of peer investment on the value of cash holding. Similar to the previous study (Leary and Roberts, 2014; Bustamante and Frésard, 2020), we measure the peer investment at time t to limit the amount of time for firms to respond to peers in our main analyses while discussing the qualitatively similar results using the lagged peer investment in Section 4.3.3.

A list of firm and peer characteristics is introduced to mitigate the concern of omitted variables and ensure robust estimation results. These variables include the change in earnings ($\Delta E_{i,t}$), the change in net assets ($NA_{i,t}$) , the change in R&D expense ($\Delta R\&D_{i,t}$), interest expense, dividends, lagged cash ($C_{i,t-1}$), and leverage ($L_{i,t}$). We also add two additional interaction terms, $C_{i,t-1} * \Delta C_{i,t}$ and $L_{i,t} * \Delta C_{i,t}$ to capture the effect of cash holdings and leverage on the marginal value of cash holdings. We summarize these characteristics of the focal firm i as the "Firm-specific factors ($Firm$)" and the average value of peer characteristics as "Peer firm averages ($Peer$)" and include both sets of variables in our regression. Moreover, we add the investment made by firm i and its interaction with $\Delta C_{i,t}$ to ensure that the effect of peer investment on the value of cash holding identified is not purely mirroring the effect of firm i's own investment on it. Lastly, we include a set of industry and year dummies (μ and ν) in the regression to ensure that our results are not driven by the time-invariant industry characteristics and the cross-sectional common factors. Detailed definitions of all variables can be found in Appendix A.

4.2.3. Data, sample selection, and summary statistics

The data used in our study is based on U.S. public listed firms and is mainly from three data sources: the stock return data is from the Centre for Security Price (CRSP), the accounting information is from Compustat annual database, and the peer identification is from Horberg and Phillips data library. In addition, we downloaded Fama French benchmark portfolio return data from Kenneth French's data library and consumer price index-related data from Federal Reserve St. Louis.

Following a standard procedure as in the previous studies, we excluded financial and utility firms (with SIC codes 6000-6999 and 4900-4999) from our sample due to their special characteristics and regulatory requirements (Faulkender and Wang, 2006b). To minimize the potential bias caused by data error, we also followed previous studies and dropped a very small number of observations with negative net assets, negative market value of equity, or negative dividends. After applying all the data filters and excluding observations with missing values, our final sample covers 49,544 firm-year observations over the fiscal year 1996–2017. To ensure that our estimation results are amplified by inflation, we converted all accounting variables to their real value in 2017 dollars by using the consumer index data.

The summary statistics of our main variables have been presented in Table 4.1. Even though compared to the previous studies, our sample period is slightly different, the statistics of our variables, are overall in line with those that have been reported previously. For example, the mean of firm investment ($Capx$) and peer investment(p_Capx) are about 0.318 and 0.325 while in Bustamante and Frésard (2020) are about 0.36 and 0.38. The statistics of our sample seem slightly smaller. This is partially due to the fact that the investment variable is constructed by normalizing the Capital Expenditure at time t by the Property, Plant and Equipment at time $t-1$. The lag structure between the numerator and denominator implies an embedded inflation return, which has been removed when we converted the data to the real value in 2017. When taking this into consideration, the mean value of peer investment could be even more close to those reported in Bustamante and Frésard (2020)'s paper. The mean and median of our accounting variables are quite similar to the ones reported by Faulkender and Wang (2006b).

4.3. Main results and robustness tests

4.3.1. Baseline regression results

In this section, we discuss our results of estimating the baseline regression of Equation 4.1, which is presented in Table 4.2. In column 2, we estimated the model with peer investment and its interaction of peer investment and the change in cash. The coefficient of the interaction term is positive and statistically significant at 1% level, supporting the hypothesis that peer investment is positively associated with the marginal value of cash holding. The significance and magnitude of this effect wouldn't diminish after we further controlled for the firms' own investment ($Capx$) and its interaction with the change in cash ($Capx$*ΔC) in column 3, the peer characteristics in column 4, and both sets of controls altogether in column 5. This effect is also economically important: by applying the most conservative results of column 5, a one standard deviation in peer investment (0.153) will lead to a 13.8% (=0.902∗0.153) higher value for $1 extra cash for the firm. It is worth noting that the effect of the value of cash increasing associated peers' investment is not significantly weaker compared to the effect of firms' own (0.148=0.444∗0.335). This is consistent with the findings of Bustamante and Frésard (2020) and Foucault and Fresard (2014) which supports the role of peer firms in providing information for focal firms.

4.3.2. Identification and endogeneity

Theoretically, the marginal value of cash holding of a firm is unlikely to be a determinant of its peer firms' investment decision. Therefore, the issue of reverse causality is not a major concern in our research setting. However, it is likely that peer firms' investment and focal firms' value of cash have some unobserved common drivers. For example, a shock to the credit supply or government policy may jointly include the investments of peer firms and the value of cash holdings of focal firms, while these factors are not easy to be quantified or controlled. Although the industry and year fixed effects in our main specification could help to mitigate this issue, the concern of omitted factors wouldn't be fully addressed without further identification attempts. We adopted several strategies, namely a two stage least square approach (2SLS) using the instrumental variables, a robustness test with the highest dimensional fixed effects, and a placebo test, to ensure that our results are not driven by the unobserved confounding factors.

Instrumental variables and 2SLS

Our first identification strategy is the 2SLS analyses based on the instrumental variables (IVs) that are related to peer investments but exogenous to the focal firms' value of cash holding. We use two different IVs in this setting: the first one is the average idiosyncratic return of firm i's product market peers and the other one is the average investment of peer firms unrelated to local neighbors firms. The extant literature, for example, Titman et al. (2004) has identified a close association between stock return and firms' investment. Therefore, the relevance condition of stock return is naturally satisfied. However, the stock return impounded a wide range of information of which some are related to market or industry-specific factors that could impede the endogeneity of using it directly as an IV. To address this concern, we use an augmented market model for the stock return to remove any market or industry-specific component embedded in the stock returns. Specifically, we run the following regression:

$$r_{i,t} = \alpha + \beta_{i,t}^{M}(rm_t - rf_t) + \beta_{i,t}^{IND}(p\_r_{i,t} - rf_t) + \eta_{i,t} \tag{4.2}$$

where $r_{i,t}$ is the daily stock return of firm i in the industry at year t, rm_t and rf_t are the daily return of CRSP all firm value-weighted total return index and risk-free rate, and $p\_r_{i,t}$ is the average return of firms within the same industry (classified by 3-digit SIC code), excluding the focal firm itself. The model is estimated for each firm year observation, and thus the idiosyncratic returns ($Iret$) are obtained as:

$$Iret_{i,t} \equiv \hat{\eta}_{i,t} = r_{i,t} - \hat{r}_{i,t} \tag{4.3}$$

where $Iret_{i,t}$ is the idiosyncratic returns of firm i at year t and $\hat{r}_{i,t}$ is the fitted value of daily stock return of firm i in the industry at year t.

We then calculate the average idiosyncratic returns of all product market peer firms and use it as the IV for our estimation. Since the information contained in the idiosyncratic returns is orthogonal to market and industry-wide common factors, by construction, the idiosyncratic returns are exogenous to peer firms' information, such as their value of cash holdings. Therefore, the exclusion condition of the IV has been met. Our approach of using peer firms idiosyncratic return as an IV follows Leary and Roberts (2014) in their evaluation of the peer effect on firms' capital structure decisions, and a similar approach has also been used in other papers such as Grennan (2019); Adhikari and Agrawal (2018); Fairhurst and Nam (2020).

The second IV that we used for peer investment is the investment made by the peer firms local neighbors that produces a completely unrelated product. This approach has firstly been applied by Bustamante and Frésard (2020). Intuitively, the investment made by a Michigan sugar company and a Michigan car marker is likely to be correlated due to the local economic and social commonalities. But there is little reason to argue the Michigan sugar company's investment would have any implication for a California car maker's value of cash holding. Therefore, to investigate how the value of the focal firms (the California car-maker) cash holdings would be influenced by the product market peers (the Michigan car maker), we could use the investment of unrelated local neighbor firms (the Michigan sugar company) as the IV.

Although the intuition of using unrelated local firms' investment as the IV is straightforward, the real challenge is to identify the neighbor firms with businesses unrelated to the peer firms. To do so, we first use the zip code of firms' headquarters to identify the metropolitan statistical area that each firm is located. If firm A's product market peer B is located in M, then we can identify all other firms that are also domiciled in M. Next, we start to exclude firms that are likely to be running a business that relates to firm B's operation. The first set of firms we exclude is the firms that might have to produce similar or related products. To do so, we exclude all local firms that are in the same Fama-French 12 industries. Then we exclude that firms that might potentially be firm B's customers or suppliers. We use two sources to identify firms that potentially share a supply chain. The first source is the vertical firm relatedness data provided by Frésard et al. (2015). In this approach, firms' business descriptions in 10K fillings have been analyzed to identify the potential customer-supplier relation. The second source is the Compustat Segment Data from which we identify the customers for each firm. This is also commonly used in previous studies, such as Bakke and Gu (2017) and Li and Tang (2016). After we identify all local firms that are potentially firm B's customers or suppliers, we excluded them in our IV constructions. The last set of firms that we excluded in calculating the IV are the firms that share the indirect link with firm A. It is likely that firm B's local neighbors might also be firm A's product market peers, and if that is the case, the IV will be associated with the characteristics of focal firm A. To address this issue, all firm A's peers are also excluded. By applying these filters, the neighbor firms of firm B left are those unrelated to firm A and their investment is then averaged and used as the IV for firm B' investment. We apply this process for each of firm A's product market peers and then calculate the average of unrelated local neighbors' investment and use it as our second instrumental variable.

The estimation results of our 2SLS analyses are presented in columns (1)–(4) of Table 4.3. Columns 1 and 3 present the first stage estimation results. We can see that the coefficient of peer average idiosyncratic return (p_Iret) is negative and statistically related to the peer investment, consistent with the findings in the previous studies that the investment is negatively related to stock return (Titman et al., 2004). The coefficient of second IV (Nb_inv) is positively and statistically significantly related to the peer investment, consistent with our argument that the local firms are likely to be subject to the common local shocks. The second stage estimation results are reported in columns 2 and 4. We can see that the interaction between instrumented peer investment ($p\_Capx\_hat * \Delta Cash$) is still positive and statically significantly related to the excess return, while the baseline regression result is only slightly larger in magnitude. These results confirm our baseline findings that the increasing peer investment could lead to a higher marginal value of cash holding.

Higher-dimensional of fixed effects

The second identification attempt that we make is to introduce the higher-dimensional fixed effects in the model. As discussed at the beginning of Section 4.3.2, the main identification challenge in our study is to mitigate the potential bias caused by the potential unobserved omitted factors. By including the time-invariant fixed effects and its interaction with the time dummy in our model, we will be able to absorb the effect of omitted variables to a large extent and ensure the robustness of our findings. Columns (5)–(7) of Table 4.3 present the result of this robustness test. Column 5 shows the result with no fixed effects. Column 6 shows the result of our baseline estimation with industry and year fixed effect, and column 7 displays the result with firm fixed effects and (industry $*$ year) fixed effects. Since the effect of peers' investment is essentially an industry average, by construction, it is highly correlated with the interactions between industry dummies and year dummies. We can see that, consistent with our expectation, the size of the estimated coefficient of $p\_Capx * \Delta Cash$ drops from 0.902 to 0.513, while it is still positive and statistically significant at 1% level.

Placebo Test

The third approach that we used to address the endogeneity concern is a placebo test. If the observed positive effect between peer investment and the marginal value of cash holdings is resulting from some unobserved market-wide common factors, then a similar positive association might exist between firms that are not product market peers. To test it, for each firm year observation, we replace its real peer investment with a pseudo peer investment of a firm that is randomly drawn from the whole sample firms by following Grennan (2019) and Gao and Zhang (2019). Then, we estimate our baseline model by using the pseudo peer investment and pseudo controls instead of the real peer investment and real controls:

$$
\begin{aligned}
r_{i,t} - R^B_{i,t} = \alpha + \beta_1' * (pseudop\_Capx_{i,t} * \frac{\Delta C_{i,t}}{MVE_{i,t-1}}) + \beta_2 * \frac{\Delta C_{i,t}}{MVE_{i,t-1}} \\
+ \beta_3 * pseudop\_Capx_{i,t} + \gamma' Firm_{i,t} + \lambda' pseudopPeer_{i,t} + \delta' \mu_j + \phi' \nu_t + \epsilon_{i,t}
\end{aligned}
\tag{4.4}
$$

We repeated this process 1,000 times and the distribution of the estimated effect of pseudo peer investment on the value of cash holdings (β_1') are reported in Figure 4.1. We can see that the estimated average coefficient is -0.074 that is far away from our estimation result by using the investment of real product market peers (0.902). These results support our arguments that the positive relation between peer investment and the value of cash holdings is unlikely to be driven by the unobserved latent common factors.

4.3.3. Other robustness tests

We further conducted a battery of other robustness tests to ensure the robustness of our findings. Our main analysis is based on the TNIC industry classification of Hoberg and Phillips (2015, 2016), which is relatively new compared to the traditional industry classification based on the SIC code. To compare the results obtained by using two different approaches, we re-estimated the model by calculating the peer investment based on the 3-digit SIC code. The average peer investment and control are therefore the average of all firms in the same 3-digit SIC industry, excluding the focal firm itself. Our estimated result is reported in column (1) of Table 4.4. In the table, we can see that the coefficient of $p\_Capx * \Delta Cash$ is still positive and statistically significantly related to the value of cash holding. The magnitude of the coefficient is slightly smaller, consistent with the argument that the TNIC industry classification could better capture the characteristics of a similar product market peer compared to the SIC code (Hoberg and Phillips, 2015, 2016).

In an efficient capital market, new information will be incorporated into stock prices quickly. If the change of cash is expected, it shouldn't be reflected in the stock prices at the beginning of the year (Faulkender and Wang, 2006b). Our research setting so far contains an implicit assumption that the expected change is zero so that the change in cash ($\Delta Cash$) is entirely unexpected. Following Faulkender and Wang (2006b), we adopted three alternative measures of the expected change in cash to validate our main results. The first measure, the average change in cash of the benchmark portfolio, assumes that firms with similar size and book-to-market ratio would have a similar level of cash holding. Therefore, when firms' cash holdings deviate from the average level, firms are expected to adjust the cash holdings towards the average, and the level of deviation is expected to change in cash holdings. The second and third measures, motivated by Almeida et al. (2004), adopt a different approach and posit that the expected change of cash are predicted by source and use of cash, and can be proxied by the fitted value of estimating the following models:

$$\begin{aligned}\Delta C_{i,t} = \beta_0 + \beta_1 * \mathit{Cash\ flow}_{i,t-1} + \beta_2 * Q_{i,t-1} + \beta_3 * \mathit{Size}_{i,t-1} \\ + \mathit{Industry\ fixed\ effects}_i + \epsilon_{i,t}\end{aligned} \tag{4.5}$$

where $\mathit{Cash\ flow}_{i,t-1}$ is cash flow of firm i in year $t-1$ that is the ratio of earnings before extraordinary items and depreciation subtract dividends to total assets, $Q_{i,t-1}$ is Tobin's Q of firm i in year $t-1$ and measured by the market-to-book ratio, $Size_{i,t-1}$ is the natural log of assets.

$$\Delta C_{i,t} = \beta_0 + \beta_1 * \mathit{Cash\ flow}_{i,t-1} + \beta_2 * Q_{i,t-1} + \beta_3 * \mathit{Size}_{i,t-1} + \beta_4 * \mathit{Capital\ expenditure}_{i,t-1}$$
$$+ \beta_5 * \mathit{Acquisitions}_{i,t-1} + \beta_6 * \Delta \mathit{Net\ working\ capital}_{i,t} + \beta_7 * \Delta \mathit{Short\ term\ debt}_{i,t}$$
$$+ \mathit{Industry\ fixed\ effects}_i + \epsilon_{i,t} \tag{4.6}$$

We create three alternative measures of the unexpected change in cash by subtracting the expected component of cash change from the overall change in cash and estimate the baseline model based on these alternative measures. These results are reported in columns 2–4 of Table 4.4, which are similar to the ones in our baseline regression, showing that our main conclusion is a robust alternative approach of defining unexpected change in cash.

Duchin (2010) argue that the traditionally used measure of cash holding measurement, the Compustat item CHE (Cash and Short term investment), contains not only a risk-free and liquid component of "cash and cash equivalents", but also a "short-term investment" component that is sometimes risky and highly illiquid. If the change in peer investment is driven by the change in the level of risk aversion of investors, it is likely that the return of a company's stock and its holding of risky assets would exhibit similar movements, leading to a positive relation between the peer investment and value of cash holdings. To ensure that our results are not contaminated by the risky investment contained in the measurement of cash holding, we use an alternative measure that only contains the real cash and cash equivalent component to re-estimate our model. Column 5 of Table 4.4 shows that our main estimation results remain unchanged by using this measurement, and rejecting the hypothesis that our findings are driven by the risky assets contained in cash holdings.

4.4. Economics drivers of our findings

Our findings so far show that an increase in peer firms' investment is associated with the higher value of firms' cash holdings. Yet it is still unclear how such a positive relationship exists. We investigate two potential mechanisms that could drive our findings.

4.4.1. Observational learning channel

Decision-makers who do not possess full information about their investment opportunity may use the observed actions of others as a signal for information. Bikhchandani et al. (1998) provide an example to illustrate this type of actions. We can assume that there are two identical firms that produce the same product: A and B. Each of them holds private information regarding the growth opportunity unknown to the other. While the information collected by firm A couldn't be observed directly by firm B, this information could be inferred by observing the actual investment decision made by firm A. When firm A increases the level of investment, firm B can predict that firm A has collected a positive signal and vice versa. Bernard et al. (2020a) and Bernard et al. (2020b) highlight the information value of the peers' investment decision and provide evidence for such learning.

The observational learning from product market peers is not unique to managers. The shareholders can also infer the private signal from the product market peers' decision. When learning information from the financial market, their acquisition of investment signal also follows a similar trajectory and views the peer investment as a signal for a better investment opportunity, therefore, reward the cash holdings with higher market value.

We designed several tests to investigate whether our findings are driven by the observational learning channel. First, compared with the investment decision made by firms that produce very different products, the decision made by firms produce similar product could contain more relevant information value (Bustamante and Frésard, 2020). If the higher value of cash associated with peer investment is indeed driven by observational learning, we should observe that such an effect would be more prominent when a firm's products are more similar to the products of its peers. We use the textual analysis-based product market similarity measure developed by Hoberg and Phillips (2015) as our measurement to divide our sample into a high similarity group (that firms with product similarity higher than the annual industry median value) and a low similarity group (that firms with similarity lower than the annual industry median). Then we estimated the baseline model for each sub-sample. Our results show that the effect of peer investment on the value of cash are indifferent in both sub-samples, against the prediction of the observational learning channel.

Secondly, the observational learning effect should be stronger with relatively smaller firms. Intuitively, smaller firms are more incentivized to learn from their product market peers (Bustamante and Frésard, 2020). We created a measurement of relative size: the ratio of firm size to the average size of peer firms. By construction, when the ratio is smaller, firms are smaller compared to the peer average size. They should have a higher incentive to learn from their industry peer firms. On the other hand, if the ratio is higher than 1, then firms are relatively larger than the average peer firms. They should have less incentive to learn from peers. We estimated our baseline model for each sub-sample of firms, respectively, and our results show there is not a significant difference between the coefficients of $p\_Capx * \Delta Cash$, showing that the effect of peer investment on the value of cash holding does not vary according to firm size. This is also against the prediction of the observational channel.

Thirdly, we conducted further analyses to check whether the effect of peer investment on the value of cash varies with the stock price informativeness. Foucault and Fresard (2014) find that firms can learn information contained in peer firms' stock prices to make their own investment decisions. Such learning will be increasing when the firms' stock price informativeness is low and the informativeness of peer firms' average stock prices is high. Since firms with better stock price informativeness are also likely to be the ones with better information quality (Gelb and Zarowin, 2002), their investment decisions could also be more informative. If the effect of peer investment on the value of cash is due to observational learning, we should expect such effect to be more prominent when the firms' stock price informativeness is lower and the peer firms' average stock price informativeness is higher. We use the number of analysts who cover a firm's accounting release as the measure of stock price informativeness. To test this hypothesis, we split our sample into the low sub-sample and high sub-sample of firm stock price informativeness and peer firms' stock price informativeness, respectively, and compare each stock price informativeness measure with its annual industry medians. Then, we estimated the effect of peer investment on the value of cash for each sub-sample. We find that, for both sub-samples of firms' stock price informativeness, the coefficients of $p\_Capx * \Delta Cash$ are close and our further test shows that the coefficient difference is not statistically significant. For both sub-samples of peer firms' stock price informativeness, the coefficients of $p\_Capx * \Delta Cash$ are indifferent and our further test shows that the difference is not statistically significant. These results again are not consistent with the prediction of the observational learning story.

4.4.2. Strategic response channel (market size effects vs. market share effects)

Peer investment could also influence the value of cash holding of a firm due to the channel of strategic reaction. When peer firms decide to increase their investment, the focal firms need to respond by changing their own financial policies.

Consequently, the value of cash holding will change. The optimal decision made by the firm in reaction to peer investment can vary when they are operating in different industries. For example, in a growing industry where firms are relatively young and with ample investment opportunities, increasing aggregate investment of the whole industry could lower the price of key production factors, therefore benefit the firm that didn't make the investment (i.e., externalities of peer investment). Therefore, in this type of industries, peer investment could increase the value of cash holding of the focal firms through the positive externality channel. On the other hand, for mature industries that are relatively well established with relatively few investment opportunities, the market capacity is relatively stable. When peer firms increase the investment, the potential product price will become cheaper therefore focal firms' product market shares may shrink. This will also lead to an increase in the value of cash holdings since these firms may need to hoard cash to prepare for potential competition in the future.

To further investigate whether or how the industry dynamics associated strategic response would shape the effect of peer investment on the value of cash holding, we conduct further tests by dividing our sample into growing firms and mature firms before we estimate the model for each type of firm. We use two proxies for such division. The first one is to use industry average Tobins' Q, and industries with higher Q are usually the ones with better investment opportunities. The second one is the average age of firms within an industry, and industries that consist of younger firms are more likely to enjoy better investment potential. We define industries with above-average Tobin's Q and below-average age as a growing industry while other firms as mature industries and investigate the effect of peer investment on the value of cash holding in each industry.

Our results reported in Table 4.6 show that the positive effect of peer investment on the value of cash holding is only significant in the high growth industries which are characterized by high industry Q and young average age. These results show that the effect of peer investment on the value of cash holdings is likely to be resulting from the positive externalities brought by peer firms' investment that is reflected in the growing industries and shared by the focal firms.

4.5. Peer investment and cash policy

We also investigate the effect of peer investment on firms' cash policy.

4.5.1. Peer investment and corporate cash holdings

For a given firm, the marginal value of cash holding is negatively related to the level of cash holding. Therefore, the positive relation between peer investment and the marginal value of cash holding could simply be a reflection of the negative relation between peer investment and the level of cash holding. If peer firms' investment causes the focal firms to reduce their cash holdings, then we might also observe a negative relation between peer investment and the value of cash holding. To investigate this hypothesis, we follow Bates et al. (2009) and conduct further analyses by regressing the level of cash holdings of a firm on the peer investment and a set of the control variables.

Our regression results are reported in Table 4.7. In both our contemporaneous specification (column 1) and our lagged specification (column 2), we can see that peer investment is positive and statistically significant associated with the level of cash holdings of a firm. The effect is also economically significant. For instance, a one standard deviation of peer investment (0.153) would lead to 0.26% (= 0.017∗0.153) higher in a firm's level of cash holdings.

4.5.2. Peer investment and the use of cash

To understand why peer investment could lead to an increase in the level of cash holdings, we conduct further tests to find out how did firms decide to use their cash, in response to the increasing in peer investment. Follow Harford et al. (2008), firms may spend their cash on dividend payments, capital investment, R&D, and marketing. We, therefore, could look at how peer investment influences the use of cash for each of these purposes. Specifically, we regress the level of spending for dividends, capital expenditure, research and development, and advertisement, on the peer investment and its interaction with peer investment, respectively. The coefficient of the interaction term indicates when peer investment changes, how a firm would change their use of cash for each of these purposes.

Our results are reported in Table 4.8. We can see that the coefficient of the interaction of peer investment and level of cash holdings is negative in the dividend payment regression, positive in the capital investment and R&D regressions, and insignificant in the advertisement regression. These results show us that when peer firms increase the investment, the focal firms would use less cash for dividend payments while using more cash for capital and R&D investment. Since these activities are broadly value-enhancing, it is not surprising that investors would have a positive view to the focal firms, therefore, provide their cash holdings with a higher valuation.

4.6. Conclusions

We find a positive relation between peer investment and the marginal value of cash holdings held by firms in the cross section. We find this relation is more pronounced among firms in the high growth industries with high Tobin's Q and young average age, suggesting this positive effect is resulting from the positive externalities brought by peer firms' investment. Such externalities can be shared to those firms that didn't invest in the same product markets. Moreover, we find little supportive evidence that cash holdings can be more valuable when firms use it as a precautionary measure for intensive competition, and investors are learning the positive signal about the investment opportunities from peer investment. In a robustness check, our results remain positive and statistically significant in alternative industry classification, model specifications of changes in cash, and measure of cash holdings. Furthermore, we find firms will increase their level of cash holdings in response to the increasing peer investment. In addition, firms are less likely to use cash for dividend payments, and more likely to spend it for capital and R&D investment. We conclude that peer investment is of great importance associated with firms' value of cash.

Appendix A

Table A1. Variable definitions

This table provides variable definitions and corresponding data sources. CRSP refers to the Centre for Research in Security Prices, Compustat refers to the Capital IQ from Standard & Poor's database, ISS refers to the Institutional Shareholder Services (formerly RiskMetrics), and FF refers to the Kenneth French's data library.

Variable	Definition	Source
$r_t - R_t^B$	Excess stock returns for a firm i with the benchmark portfolios defined as Fama-French 25 portfolios formed on size and book-to-market factor returns. (Faulkender and Wang, 2006b)	CRSP, Compustat and FF
AT_t	Total assets (millions).	Compustat
MVE_t	Market value of equity (millions): number of shares outstanding (CSHPRI) * Stock Price (PRCC_F). (Faulkender and Wang, 2006b)	Compustat
$Capx_t$	Capital expenditure (CAPEX) divided by lagged Property, plant and equipment (PPENT).	Compustat
$Cash_t$	Cash and short-term investment(CHE) divided by MVE. (Faulkender and Wang, 2006b)	Compustat
$\Delta Cash_t$	Change in cash holdings from fiscal year $t-1$ to year t, divided by MVE at the beginning of fiscal year t. (Faulkender and Wang, 2006b)	Compustat
$\Delta Earnings_t$	Change in firm earnings from fiscal year $t-1$ to year t, divided by MVE at the beginning of fiscal year t. Earnings = income before extraordinary items (IB) + interest expense(XINT) + deferred tax credits (TXDI) + investment tax credits (ITCI). (Faulkender and Wang, 2006b)	Compustat
$\Delta Net\ assets_t$	Change in net assets from fiscal year $t-1$ to year t, divided by MVE at the beginning of fiscal year t. Net assets= AT - cash holdings(CHE). (Faulkender and Wang, 2006b)	Compustat
$\Delta R\&D_t$	Change in research and development expenditure (XRD) from fiscal year $t-1$ to year t, divided by MVE at the beginning of fiscal year t. (Faulkender and Wang, 2006b)	Compustat
$\Delta Interest\ expenses_t$	Change in interest expenses(XINT) from fiscal year $t-1$ to year t, divided by MVE at the beginning of fiscal year t. (Faulkender and Wang, 2006b)	Compustat

Continued on next page

Table A1 - continued from previous page

Variable	Definition	Source
$\Delta Dividends_t$	Change in total common share dividends (DVC) from fiscal year $t-1$ to year t, divided by MVE at the beginning of fiscal year t. (Faulkender and Wang, 2006b)	Compustat
$Leverage_t$	Leverage ratio= [Short-term Debt (DLC) + Long-term Debt (DLTT)]/ [DLC + DLTT + MVE]. (Faulkender and Wang, 2006b)	Compustat
$Net\ financing_t$	Net financing = equity issuance (SSTK) - repurchases (PRSTKC) + debt issuance(DLTIS) - debt redemption (DLTR), divided by MVE at the beginning of fiscal year t. (Faulkender and Wang, 2006b)	Compustat
Firm-Specific Factors	These are firm i's characteristic variables, including $\Delta Cash_t$, $\Delta Earnings_t$, $\Delta Net\ assets_t$, $\Delta R\&D_t$, $\Delta Interest\ expenses_t$, $\Delta Dividends_t$, $Cash_{t-1}$, $Leverage_t$, $Net\ financing_t$, $Cash_{t-1} * \Delta Cash_t$, and $Leverage_t * \Delta Cash_t$.	
p_Capx_t	Equally-weighted average peer firms' investment of a firm i is constructed by following Hoberg and Phillips (2015) TNIC industry classifications for each industry and fiscal year.	
Peer Firm Averages	These are equally-weighted average peer firms' characteristic variables of a firm i, including $p\_\Delta Cash_t$, $p\_\Delta Earnings_t$, $p\_\Delta Net\ assets_t$, $p\_\Delta R\&D_t$, $p\_\Delta Interest\ expenses_t$, $p\_\Delta Dividends_t$, $p\_Leverage_t$, p_Cash_{t-1}, $p\_Net\ financing_t$, $p\_(Cash_{t-1} * \Delta Cash_t)$, and $p\_(Leverage_t * \Delta Cash_t)$. The above variables are constructed by following Hoberg and Phillips (2015) TNIC industry classifications for every industry and fiscal year.	
p_Iret_t	Average peer firms' idiosyncratic stock returns of a firm i are defined as the difference between average realized and expected returns of product market peers. (Leary and Roberts, 2014)	Compustat
Nb_inv_t	Average investment of neighbor firms of non-local product market peers. (Bustamante and Frésard, 2017)	Compustat
$G\text{-}index_t$	Corporate governance index is composed of twenty-four provisions on investor rights and takeover protections applied to a firm i. (Gompers et al., 2003)	ISS
$E\text{-}index_t$	Entrenchment index is composed of the six most important provisions in *G-index*. (Bebchuk et al., 2008)	ISS
$Numest_t$	The number of analysts who cover a firm i's accounting releases.	Compustat
$Size_t$	Firm size: the natural log of AT. (Kim et al., 2014)	Compustat

Continued on next page

Table A1 - continued from previous page

Variable	Definition	Source
$Prfmrg_t$	Profit margin = Operating income (OIBDP)/Total sales (SALE). (Fairfield and Yohn, 2001)	Compustat
Tobin's Q_t	Tobin's Q is the market-to-book ratio, *MTB*. (Hennessy et al., 2007)	Compustat
$Opert\_exp_t$	Operating expense ratio: Total Operting Expenss (XOPR)/Total sales (SALE). (Ang et al., 2000)	Compustat
NWC_t	Net working capital = [Working capital (WCAP) − $Cash$]/AT. (Bates et al., 2009)	Compustat
$R\&D\_Sale_t$	The ratio of R&D expenses to total sales = XRD/SALE. (Bates et al., 2009)	Compustat
$Acquisitions_t$	Acquisition expenditures (AQC), normalized by AT. (Bates et al., 2009)	Compustat
div_dum_t	Dividend dummy equals to one if a firm i pays positive common dividend, and zero otherwise. (Bates et al., 2009)	Compustat
$Sigma_t$	The average of the standard deviations of cash flows over ten years for firms with the same three-digit SIC codes. (Bates et al., 2009)	Compustat
Cash flow$_t$	The ratio of earnings before extraordinary items and depreciation subtract dividends to total assets.	Compustat

Table 4.1. Summary statistics

This table presents the summary statistics of the variables used in our main empirical analysis. For the variables included in our baseline regression model, the sample consists of 49,544 firm–year observations over the period 1996-2017. The prefix $p\_$ denotes peer firms' equal-weighted average characteristics within each industry and fiscal year for a firm i in year t by following Hoberg and Phillips (2015). The number of observations, mean, standard deviation, 1st percentile, 25th percentile, median, 75th percentile, and 99th percentile are reported from left to right, in sequence for each variable. See Appendix A for variable definitions.

Variables	Obs.	Mean	S.D.	p1	p25	Median	p75	p99
$r_t - R_t^B$	49,544	0.022	0.617	-0.997	-0.335	-0.075	0.218	3.003
MVE_t	49,544	4514.369	15000	4.877	106.924	509.758	2173.207	120000
$Capx_t$	49,544	0.318	0.335	0.005	0.122	0.216	0.385	2.079
$Cash_{t-1}$	49,544	0.186	0.241	0.000	0.037	0.104	0.236	1.507
$\Delta Cash_t$	49,554	0.005	0.136	-0.525	-0.032	0.001	0.036	0.621
$\Delta Earnings_t$	49,554	0.021	0.231	-0.754	-0.030	0.004	0.038	1.331
$\Delta Net\ assets_t$	49,554	0.014	0.384	-1.746	-0.059	0.011	0.092	1.677
$\Delta R\&D_t$	49,554	-0.001	0.025	-0.144	0.000	0.000	0.001	0.084
$\Delta Interest\ expenses_t$	49,554	0.001	0.018	-0.083	-0.001	0.000	0.002	0.092
$\Delta Dividends_t$	49,554	0.000	0.009	-0.061	0.000	0.000	0.000	0.047
$Leverage_t$	49,544	0.252	0.290	0.000	0.007	0.151	0.392	1.203
$Net\ financing_t$	49,544	0.038	0.209	-0.567	-0.032	0.000	0.051	1.130
p_Capx_t	49,544	0.325	0.153	0.058	0.207	0.307	0.426	0.781
p_Cash_{t-1}	49,544	0.184	0.122	0.008	0.094	0.159	0.251	0.618
$p\_\Delta Cash_t$	49,544	0.005	0.044	-0.131	-0.016	0.003	0.024	0.172

Continued on next page

Table 4.1 - continued from previous page

Variables	Obs.	Mean	S.D.	p1	p25	Median	p75	p99
$p\_\Delta Earnings_t$	49,544	0.019	0.086	-0.251	-0.016	0.009	0.040	0.365
$p\_\Delta Net\ assets_t$	49,544	0.021	0.157	-0.588	-0.031	0.021	0.082	0.535
$p\_\Delta R\&D_t$	49,544	-0.001	0.007	-0.034	-0.001	0.000	0.001	0.016
$p\_\Delta Interest\ expenses_t$	49,544	0.001	0.006	-0.023	-0.001	0.001	0.003	0.026
$p\_\Delta Dividends_t$	49,544	0.000	0.003	-0.012	0.000	0.000	0.001	0.011
$p\_Leverage_t$	49,544	0.244	0.187	0.000	0.096	0.184	0.353	0.861
$p\_Net\ financing_t$	49,544	0.040	0.084	-0.183	-0.010	0.024	0.084	0.323
p_Iret_t	49,434	-0.005	0.023	-0.080	-0.016	-0.006	0.005	0.080
Nb_inv_t	49,261	0.462	0.160	0.192	0.343	0.425	0.554	0.903
Peer similarity	24,768	0.031	0.021	0.002	0.017	0.025	0.038	0.111
$Numest_t$	17,067	7.386	6.801	1.000	2.167	5.083	10.333	30.583
p_Numest_t	17,406	8.546	3.472	1.000	6.450	8.228	10.373	30.583
$Tobin's\ Q_t$	23,280	1.962	1.403	0.546	1.114	1.501	2.250	8.545
Industry age	24,765	2.807	0.417	1.991	2.516	2.773	3.041	3.978
Cash flow	44,914	0.018	0.190	-0.957	0.011	0.066	0.107	0.273
NWC	44,914	0.065	0.174	0.404	0.041	0.048	0.170	0.538
R&D_Sale	44,914	0.236	1.170	0.000	0.000	0.003	0.077	10.122
Acquisitions	44,914	0.023	0.058	-0.003	0.000	0.000	0.011	0.333
div_dum	44,914	0.349	0.477	0.000	0.000	0.000	1.000	1.000
Sigma	44,914	0.068	0.045	0.012	0.033	0.057	0.085	0.215

Table 4.2. Baseline regression model

This table presents panel regression results estimated by the baseline regression model (4.1). The dependent variable is excess stock return, defined as the difference between stock return and Fama-French 25 portfolio returns formed on size and book-to-market factors: $r_t - R_t^B$. The independent variable of our interest is the interaction of average peer investment and firms' own value of cash holdings: $p\_Capx_t * \Delta Cash_t$. Firm-specific factors denote a firm i's characteristic variables in year t. Peer firm averages denote equally-weighted average characteristic variables of a firm i's product market peers and are constructed by following Hoberg and Phillips (2015) TNIC industry classifications. Column (1) estimates firms' value of cash holdings. Column (2) estimates the interaction of average peer investment and firms' value of cash holdings, controlling for firm-specific factors only. Column (3) estimates the same interaction by controlling for firm-specific factors and the interaction of firm investment and firms' value of cash. Column (4) tests the same interaction by controlling for both firm-specific and peer firm average factors only. Column (5) tests the same interaction by controlling for firm-specific factors, peer firm averages, and the interaction of firm investment and firms' value of cash. ΔX_t is compact notation for the 1-year change in the variable X: $X_t - X_{t-1}$. Year and industry fixed effects are controlled in all regressions. Industries are defined by three-digit SIC code. See Appendix A for variable definitions.

Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
$p\_Capx_t * \Delta Cash_t$		1.155***	0.909***	1.145***	0.902***
		[5.33]	[4.02]	[5.29]	[4.00]
p_Capx_t		-0.055**	-0.060**	-0.126***	-0.131***
		[-2.38]	[-2.57]	[-4.98]	[-5.13]
$Capx_t * \Delta Cash_t$			0.449***		0.444***
			[3.78]		[3.76]
$Capx_t$			0.010		0.012
			[1.01]		[1.14]
Firm-specific factors					

Continued on next page

Table 4.2 - continued from previous page

Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
$\Delta Cash_t$	2.111***	1.704***	1.634***	1.661***	1.591***
	[36.39]	[17.67]	[16.71]	[17.26]	[16.31]
$\Delta Earnings_t$	0.535***	0.535***	0.535***	0.520***	0.520***
	[28.52]	[28.46]	[28.48]	[27.68]	[27.70]
$\Delta Net\ assets_t$	0.218***	0.218***	0.216***	0.211***	0.209***
	[17.95]	[17.92]	[17.69]	[17.29]	[17.04]
$\Delta R\&D_t$	1.000***	0.994***	0.974***	0.971***	0.949***
	[5.97]	[5.94]	[5.79]	[5.82]	[5.66]
$\Delta Interest\ expenses_t$	-1.959***	-1.961***	-1.956***	-1.826***	-1.820***
	[-8.37]	[-8.38]	[-8.36]	[-7.81]	[-7.79]
$\Delta Dividends_t$	1.337***	1.324***	1.331***	1.291***	1.299***
	[4.37]	[4.33]	[4.35]	[4.25]	[4.27]
$Cash_{t-1}$	0.443***	0.450***	0.452***	0.449***	0.451***
	[23.11]	[23.44]	[23.54]	[22.99]	[23.10]
$Leverage_t$	-0.475***	-0.477***	-0.474***	-0.462***	-0.460***
	[-43.55]	[-43.53]	[-42.84]	[-41.55]	[-40.90]
$Net\ financing_t$	-0.026	-0.026	-0.032	-0.031	-0.036
	[-1.06]	[-1.06]	[-1.27]	[-1.23]	[-1.44]
$Cash_{t-1} * \Delta Cash_t$	-0.935***	-0.870***	-0.831***	-0.838***	-0.799***
	[-10.32]	[-9.55]	[-9.10]	[-9.21]	[-8.76]
$Leverage_t * \Delta Cash_t$	-1.666***	-1.513***	-1.477***	-1.493***	-1.457***

Continued on next page

Table 4.2 - continued from previous page

Variables	(1)	(2)	(3)	(4)	(5)
	[-16.07]	[-14.23]	[-13.90]	[-14.11]	[-13.78]
Peer firm averages					
$p\_\Delta Cash_t$				1.786***	1.780***
				[9.72]	[9.69]
$p\_\Delta Earnings_t$				0.233***	0.233***
				[6.31]	[6.33]
$p\_\Delta Net\ assets_t$				0.117***	0.118***
				[4.87]	[4.90]
$p\_\Delta R\&D_t$				-0.192	-0.198
				[-0.38]	[-0.40]
$p\_\Delta Interest\ expenses_t$				-1.428***	-1.420***
				[-2.82]	[-2.80]
$p\_\Delta Dividends_t$				-0.705	-0.734
				[-0.78]	[-0.82]
$p\_Leverage_t$				-0.062***	-0.061***
				[-2.85]	[-2.83]
p_Cash_{t-1}				0.033	0.031
				[0.96]	[0.91]
$p\_Net\ financing_t$				0.033	0.029
				[0.69]	[0.60]
$p\_(Cash_{t-1} * \Delta Cash_t)$				-1.634***	-1.607***

Continued on next page

Table 4.2 - continued from previous page

Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
				[-2.89]	[-2.84]
$p\_(Leverage_t * \Delta Cash_t)$				-1.759***	-1.754***
				[-4.75]	[-4.74]
Constant	-0.004	0.010	0.009	0.024	0.023
	[-0.11]	[0.28]	[0.24]	[0.66]	[0.63]
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	49,546	49,546	49,546	49,544	49,544
Adj. R^2	0.230	0.231	0.231	0.238	0.238

Table 4.3. Peer investment and value of cash: Mitigating endogenous concerns

This table presents two-stage least squares analyses and panel regression results estimated by the baseline regression model (4.1). The dependent variable is excess stock return, defined as the difference between stock return and Fama-French 25 portfolio returns formed on size and book-to-market factors: $r_t - R_t^B$. The independent variable of our interest is $p\_Capx\_hat_{t+1} * \Delta Cash_t$. Columns (1)-(2) are two-stage least squares analyses with an instrument of non-local average peer investment of neighbor firms (Nb_inv). Columns (3)-(4) are two-stage least squares analyses with an instrument of average peer idiosyncratic returns (p_Iret). Columns (5)-(7) are OLS panel regressions by controlling for no fixed effects, (Year $*$ Industry) fixed effects, and (Firm and Year $*$ Industry) triple fixed effects, respectively. Other firm-specific factors denote a firm i's characteristic variables in year t except the variable of $\Delta Cash_t$. Peer firm averages denote equally-weighted average characteristic variables of a firm i's product market peers and are constructed by following Hoberg and Phillips (2015) TNIC industry classifications. ΔX_t is compact notation for the 1-year change in the variable X: $X_t - X_{t-1}$. Year and industry fixed effects are controlled for in all regressions except for Column (5). Industries are defined by three-digit SIC code. The t-statistics are robust to white heteroskedasticity and reported in parentheses. See Appendix A for variable definitions. The t-statistics are presented in the brackets. $***$, $**$, and $*$ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Instrumental Variables				**High-dimensional FE**		
	1st-stage	2nd-stage	1st-stage	2nd-stage		OLS	
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
$p\_Capx\_hat_{t+1} * \Delta Cash_t$		1.146***		1.196***			
		[4.11]		[4.38]			

Continued on next page

Table 4.3 - continued from previous page

	Instrumental Variables				**Panel**		
	1st-stage	2nd-stage	1st-stage	2nd-stage	OLS		
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
$p\_Capx\_hat_{t+1}$		-5.561***		-0.783***			
		[-13.93]		[-4.62]			
$p\_Capx_t * \Delta Cash_t$					0.859***	0.902***	0.513***
					[3.76]	[4.00]	[3.09]
p_Capx_t					-0.144***	-0.131***	-0.008
					[-6.92]	[-5.13]	[-0.26]
$Capx_t * \Delta Cash_t$	0.005	0.425***	0.003	0.390***	0.435***	0.444***	0.470***
	[0.32]	[3.52]	[0.23]	[3.24]	[3.64]	[3.76]	[5.63]
$Capx_t$	0.038***	0.222***	0.038***	0.037***	0.009	0.012	0.010

Continued on next page

Table 4.3 - continued from previous page

	Instrumental Variables				Panel		
	1st-stage	2nd-stage	1st-stage	2nd-stage		OLS	
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
	[23.90]	[11.85]	[23.46]	[3.03]	[0.91]	[1.14]	[1.10]
p_Iret_t	-0.328***						
	[-10.72]						
Nb_inv_t			0.220***				
			[25.02]				
$\Delta Cash_t$	-0.007	1.443***	-0.006	1.473***	1.643***	1.591***	1.608***
	[-0.87]	[12.93]	[-0.74]	[13.37]	[16.63]	[16.31]	[22.72]
Constant	0.309***	1.705***	0.142***	0.229***	0.025***	0.023	0.017
	[33.39]	[13.15]	[12.78]	[3.64]	[2.62]	[0.63]	[0.97]

Continued on next page

Table 4.3 - continued from previous page

	Instrumental Variables				Panel		
	1st-stage	2nd-stage	1st-stage	2nd-stage	OLS		
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Other firm-specific factors	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Peer firm average	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	No	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	No	Yes	Yes
Firm fixed effects	No	No	No	No	No	No	Yes
Shea's partial R^2	0.0051	-	0.0225	-	-	-	-
First stage F-stat	114.815	-	626.099	-	-	-	-
Observations	49,434	49,434	49,261	49,261	49,544	49,544	47,670
Adj. R^2	0.587	0.242	0.595	0.239	0.225	0.238	0.322

Table 4.4. Robustness tests

This table presents robustness test results of the baseline regression model (4.1) by using alternative industry classification and cash holdings measures. The dependent variable is excess stock return, defined as the difference between stock return and Fama-French 25 portfolio returns formed on size and book-to-market factors: $r_t - R_t^B$. The independent variable of our interest is the interaction of average peer firms' investment and firms' own value of cash holdings: $p\_Capx_t * \Delta Cash\_alt_t$. Other firm-specific factors denote a firm i's characteristic variables except the variable of $\Delta Cash_t$. Peer Firm Averages denote equally-weighted average characteristic variables of a firm i's product market peers and are constructed by following 3-digit SIC industry classifications. Column (1) is our robustness check using the alternative industry measure of 3-digit SIC classifications. Column (2)-(4) are the robustness check using the alternative cash holdings measures. Column (5) is the robustness check using the measure of cash only, not cash and short-term investments used in Column (1)-(4). ΔX_t is compact notation for the 1-year change in the variable X: $X_t - X_{t-1}$. Year and industry fixed effects are controlled for in all regressions. Industries are defined by three-digit SIC code. The t-statistics are robust to white heteroskedasticity and reported in parentheses. See Appendix A for variable definitions. $***$, $**$, and $*$ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	SIC3 industry	**Alternative cash holdings**			**Cash only**
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
$p\_Capx_t * \Delta Cash_t$	0.587***	0.891***	0.812***	0.661**	0.768***
	[2.66]	[3.89]	[3.32]	[2.55]	[2.89]
p_Capx_t	-0.124***	-0.122***	-0.122***	-0.118***	-0.132***
	[-3.79]	[-4.74]	[-4.52]	[-4.28]	[-5.12]
$Capx_t * \Delta Cash_t$	0.436***	0.315**	0.421***	0.542***	0.531***
	[4.33]	[2.56]	[3.23]	[3.76]	[3.71]
$Capx_t$	0.007	0.016	0.028**	0.036***	0.020*
	[0.74]	[1.57]	[2.55]	[3.10]	[1.89]

Continued on next page

Table 4.4 - continued from previous page

	SIC3 industry	Alternative cash holdings			Cash only
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
$\Delta Cash_t$	1.626***	1.638***	1.749***	1.687***	1.538***
	[18.94]	[16.65]	[17.03]	[15.50]	[13.81]
Constant	0.055*	0.054	0.156***	0.156***	0.011
	[1.83]	[1.48]	[4.13]	[3.97]	[0.32]
Other firm-specific factors	Yes	Yes	Yes	Yes	Yes
Peer firm averages	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	66,839	49,465	46,318	42,343	49,544
Adj. R^2	0.230	0.233	0.240	0.230	0.220

Table 4.5. Peer investment and value of cash: Observational learning

This table presents the panel regression results estimated by the baseline regression model (4.1) for the observational learning channel. The dependent variable is defined as the difference between stock return and Fama-French 25 portfolio returns formed on size and book-to-market factors: $r_t - R^B_t$. The independent variable of our interest is the interaction of average peer investment and firms' value of cash holdings: $p\_Capx_t * \Delta Cash_t$. Columns (1)-(2) estimate the interaction of average peer investment and firms' value of cash holdings between a firm i's size and average peer firms' size. Columns (3)-(4) estimate the same interaction in the low and high similar peer firms groups. Columns (5)-(6) estimate the same interaction in the low and high level of stock price informativeness groups. Columns (7)-(8) estimate the same interaction in the low and high level of average peer stock price informativeness groups. Other firm-specific factors denote a firm i's characteristic variables in year t except the variable of $\Delta Cash_t$. Peer firm averages denote equally-weighted average characteristic variables of a firm i's product market peers and are constructed by following Hoberg and Phillips (2015) TNIC industry classifications. ΔX_t is compact notation for the 1-year change in the variable X: $X_t - X_{t-1}$. Industry and year fixed effects are controlled for in all regressions. Industries are defined by three-digit SIC code. The t-statistics are robust to white heteroskedasticity and reported in parentheses. See Appendix A for variable definitions. $***$, $**$, and $*$ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Size		**Peer similarity**		**Firm informativeness**		**Peer informativeness**	
	Firm	Peer	Low	High	Low	High	Low	High
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
$p\_Capx_t * \Delta Cash_t$	0.791***	0.640*	1.011***	0.590*	0.650*	0.708	0.765*	0.794*

Continued on next page

Table 4.5 - continued from previous page

	Size		Peer similarity		Firm informativeness		Peer informativeness	
	Firm	Peer	Low	High	Low	High	Low	High
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
	[2.81]	[1.72]	[3.31]	[1.78]	[1.68]	[1.53]	[1.93]	[1.73]
p_Capx_t	-0.108***	-0.139***	-0.052	-0.226***	-0.140***	-0.212***	-0.148***	-0.205***
	[-2.98]	[-4.00]	[-1.60]	[-5.33]	[-3.14]	[-5.68]	[-3.71]	[-4.64]
$Capx_t * \Delta Cash_t$	0.257*	1.037***	0.309*	0.584***	0.229	0.736***	0.586***	0.269
	[1.79]	[5.10]	[1.72]	[3.72]	[1.22]	[3.20]	[2.94]	[1.24]
$Capx_t$	0.026**	-0.027	0.027*	0.003	0.003	-0.028*	-0.015	0.003
	[2.02]	[-1.61]	[1.81]	[0.22]	[0.20]	[-1.74]	[-0.95]	[0.16]
$\Delta Cash_t$	1.724***	1.313***	1.579***	1.646***	1.799***	1.903***	1.823***	1.740***
	[14.12]	[8.36]	[11.46]	[11.90]	[10.97]	[9.95]	[10.36]	[9.41]

Continued on next page

Table 4.5 - continued from previous page

	Size		Peer similarity		Firm informativeness		Peer informativeness	
	Firm	Peer	Low	High	Low	High	Low	High
Variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Constant	-0.112*	0.142***	-0.013	0.074	0.050	0.121**	0.041	0.114*
	[-1.73]	[3.36]	[-0.29]	[1.25]	[0.62]	[2.19]	[0.58]	[1.74]
Other firm-specific factors	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Peer firm averages	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27,359	22,185	24,776	24,768	17,067	18,057	17,406	17,718
Adj. R^2	0.245	0.240	0.227	0.256	0.253	0.237	0.254	0.234

Table 4.6. Peer investment and value of cash: Strategic response

This table presents panel regression analyses results estimated by the baseline regression model (4.1) for the strategic response channel. The dependent variable is excess stock return, defined as the difference between stock return and Fama-French 25 portfolio returns formed on size and book-to-market factors: $r_t - R_t^B$. The independent variable of our interest is the interaction of average peer investment and firms' value of cash holdings: $p\_Capx_t * \Delta Cash_t$. For every firm-year observation, our sample is divided into low and high sub-samples based on annual industry median values of Tobin's Q and Industry age, respectively. Other firm-specific factors denote a firm i's characteristic variables except the variable of $\Delta Cash_t$. Peer firm averages denote equally-weighted average characteristic variables of a firm i's product market peers and are constructed by following Hoberg and Phillips (2015) TNIC industry classifications. Columns (1)-(2) estimate the interaction of average peer investment and firms' value of cash holdings in the low Tobin's Q group versus the high Tobin's Q group. Columns (3)-(4) estimate the same interaction in the young industry group versus the old industry group. ΔX_t is compact notation for the 1-year change in the variable X, $X_t - X_{t-1}$. Year and industry fixed effects are controlled in all regressions. Industries are defined by three-digit SIC code. The t-statistics are robust to white heteroskedasticity and reported in parentheses. $***$, $**$, and $*$ denote statistical significance at the 1%, 5%, and 10% level, respectively.

	Tobin's Q		**Industry Age**	
	Low	High	Young	Old
Variables	**(1)**	**(2)**	**(3)**	**(4)**
$p\_Capx_t * \Delta Cash_t$	0.232	1.472***	0.631**	0.531
	[0.98]	[3.88]	[1.99]	[1.43]
p_Capx_t	-0.191***	-0.112***	-0.065	-0.118***
	[-6.19]	[-3.02]	[-1.64]	[-3.34]
$Capx_t * \Delta Cash_t$	0.061	0.169	0.328**	0.727***
	[0.47]	[0.95]	[2.32]	[3.41]
$Capx_t$	-0.077***	-0.016	0.009	0.017

Continued on next page

Table 4.6 - continued from previous page

	Tobin's Q		Industry Age	
	Low	High	Young	Old
Variables	**(1)**	**(2)**	**(3)**	**(4)**
	[-6.82]	[-1.19]	[0.71]	[0.92]
$\Delta Cash_t$	1.184***	2.355***	1.862***	1.326***
	[11.37]	[14.16]	[12.20]	[10.42]
Constant	-0.207***	0.110*	-0.063	0.085*
	[-5.10]	[1.95]	[-1.14]	[1.74]
Other firm-specific factors	Yes	Yes	Yes	Yes
Peer firm averages	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Observations	23,280	25,670	24,765	24,779
Adj. R^2	0.239	0.319	0.254	0.236

Table 4.7. Peer investment and value of cash: Cash holdings

This table presents panel regression analyses results of examining the effect of average peer investment on corporate cash holdings by following Bates et al. (2009). The dependent variable is a firm i's cash holdings: $Cash_t$. The independent variable of interest is the average peer investment: p_Capx. Firm-specific factors denote a firm i's characteristic variables in year t. Peer firm averages denote equally-weighted average characteristic variables of a firm i's product market peers and are constructed by following Hoberg and Phillips (2015) TNIC industry classifications. Column (1) estiamtes the impact of average peer investment on firms' cash holdings in year t: p_Capx_t. Column (2) estimates the impact of one-year lagged average peer investment on current firms' cash holdings: p_Capx_{t-1}. Year and industry fixed effects are controlled in all regressions. Industries are defined by three-digit SIC code. The t-statistics are robust to white heteroskedasticity and reported in parentheses. $***$, $**$, and $*$ denote statistical significance at the 1%, 5%, and 10% level, respectively.

Variables	**(1)**	**Variables**	**(2)**
p_Capx_t	0.017**	p_Capx_{t-1}	0.014*
	[2.27]		[1.71]
Firm-specific factors			
$Size_t$	-0.010***	$Size_{t-1}$	-0.011***
	[-20.65]		[-19.69]
$Cash\ flow_t$	-0.050***	$Cash\ flow_{t-1}$	-0.085***
	[-7.34]		[-10.10]
MTB_t	0.020***	MTB_{t-1}	0.019***
	[28.65]		[23.79]
NWC_t	-0.319***	NWC_{t-1}	-0.289***
	[-51.87]		[-40.81]

Continued on next page

Table 4.7 - continued from previous page

Variables	**(1)**	**Variables**	**(2)**
$R\&D\_Sale_t$	0.026***	$R\&D\_Sale_{t-1}$	0.023***
	[24.68]		[18.83]
$Acquisition_t$	-0.272***	$Acquisition_{t-1}$	-0.234***
	[-28.98]		[-22.60]
div_dum_t	-0.012***	div_dum_{t-1}	-0.013***
	[-7.76]		[-7.68]
$Sigma_t$	0.273***	$Sigma_{t-1}$	0.279***
	[5.64]		[5.25]
$Leverage_t$	-0.311***	$Leverage_{t-1}$	-0.288***
	[-60.76]		[-51.72]
Peer firm averages			
p_Size_t	-0.000	p_Size_{t-1}	0.001
	[-0.07]		[0.96]
$p\_Cash\_flow_t$	-0.118***	$p\_Cash\_flow_{t-1}$	-0.160***
	[-7.31]		[-8.46]
p_MTB_t	-0.005***	p_MTB_{t-1}	-0.006***
	[-3.05]		[-3.09]
p_NWC_t	-0.050***	p_NWC_{t-1}	-0.045***
	[-4.88]		[-3.77]
$p\_rd\_sale_t$	0.033***	$p\_rd\_sale_{t-1}$	0.027***

Continued on next page

Table 4.7 - continued from previous page

Variables	**(1)**	**Variables**	**(2)**
	[10.10]		[7.45]
$p\_Acquisition_t$	-0.025	$p\_Acquisition_{t-1}$	-0.063**
	[-0.93]		[-2.06]
$p\_div\_dum_t$	-0.025***	$p\_div\_dum_{t-1}$	-0.028***
	[-8.19]		[-7.95]
p_Sigma_t	-0.059	p_Sigma_{t-1}	-0.053
	[-1.24]		[-0.96]
$p\_Leverage_t$	-0.077***	$p\_Leverage_{t-1}$	-0.089***
	[-8.40]		[-8.28]
Constant	0.273***	Constant	0.271***
	[22.03]		[19.46]
Industry fixed effects	Yes	Industry fixed effects	Yes
Year fixed effects	Yes	Year fixed effects	Yes
Observations	44,914	Observations	37,402
Adj. R^2	0.602	Adj. R^2	0.587

Table 4.8. Peer investment and value of cash: Use of cash

This table presents panel regression analyses results of estimating the impact of average peer investment on the use of cash holdings by following Harford et al. (2008). The dependent variable is dividend payments, capital expenditure, R&D expenditure, and advertising expense, respectively. The independent variable of our interest is the interaction of average peer investment and firms' cash holdings: $p\_Capx_t * Cash_t$. Firm-specific factors denote a firm i's characteristic variables in year t. Peer firm averages denote equally-weighted average characteristic variables of a firm i's product market peers and are constructed by following Hoberg and Phillips (2015) TNIC industry classifications. Column (1) estimates the impact of average peer investment on firms' use of cash as dividend payments. Column (2) estimates the impact of average peer investment on firms' use of cash as capital expenditure. Column (3) estimates the impact of average peer investment on firms' use of cash as R&D expense. Column (4) estimates the impact of average peer investment on firms' use of cash as advertising expense. Year and industry fixed effects are controlled in all regressions. Industries are defined by three-digit SIC code. The t-statistics are robust to white heteroskedasticity and reported in parentheses. $***$, $**$, and $*$ denote statistical significance at the 1%, 5%, and 10% level, respectively.

	Dividend	**Capital expense**	**R&D**	**Advertising expense**
Variables	**(1)**	**(2)**	**(3)**	**(4)**
$p\_Capx_t * Cash_t$	-0.020***	0.005***	0.120***	-0.005
	[-6.07]	[2.88]	[6.84]	[-1.05]
p_Capx_t	-0.001	-0.001	-0.018***	0.009***
	[-1.61]	[-1.44]	[-4.90]	[5.85]
$Capx_t * Cash_t$	-0.000		-0.015*	0.003
	[-0.18]		[-1.76]	[1.62]
$Capx_t$	-0.004***		-0.005**	-0.000
	[-10.88]		[-2.20]	[-0.21]
Firm-specific factors				

Continued on next page

Table 4.8 - continued from previous page

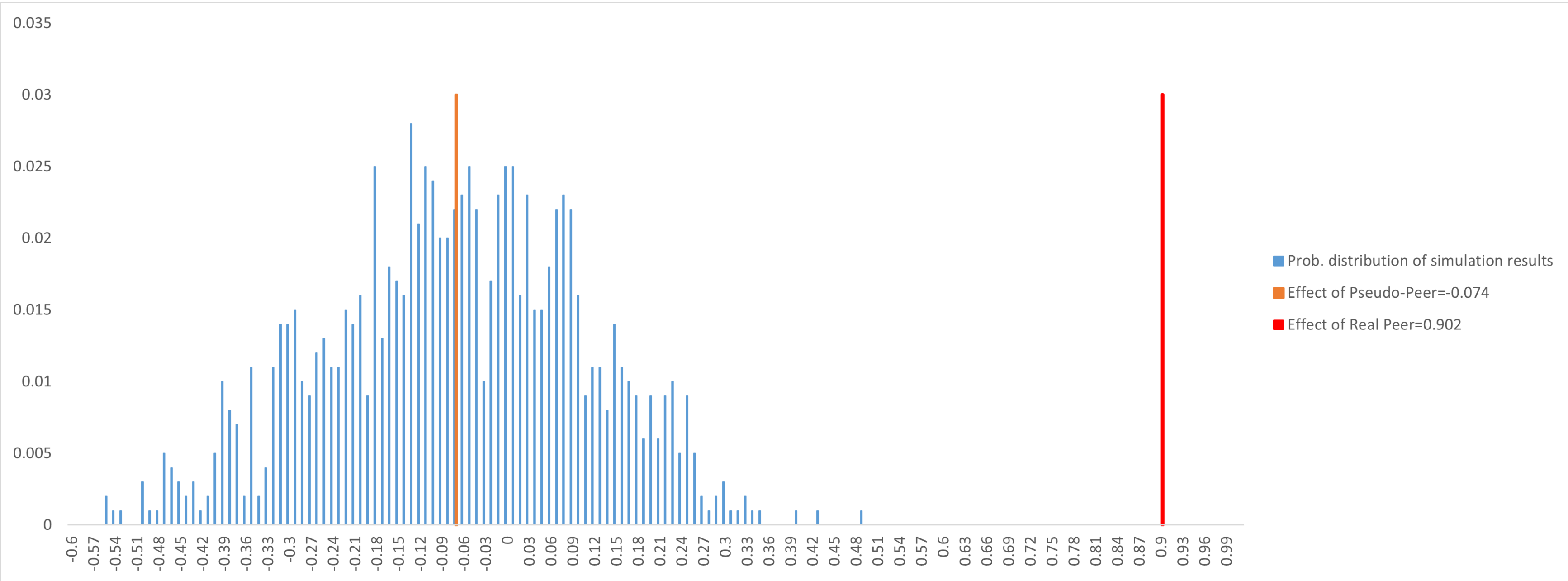
	Dividend	Capital expense	R&D	Advertising expense
Variables	**(1)**	**(2)**	**(3)**	**(4)**
MTB	0.002***	0.001***	0.009***	0.001***
	[20.42]	[15.24]	[19.14]	[6.37]
Cash flow$_t$	-0.000	-0.001***	-0.047***	0.001***
	[-1.16]	[-15.02]	[-42.81]	[5.41]
Ln_at	0.001***	-0.002***	-0.007***	-0.001***
	[25.91]	[-68.63]	[-29.90]	[-7.35]
Leverage	-0.007***	-0.001***	-0.005***	-0.002***
	[-32.61]	[-13.26]	[-5.33]	[-4.23]
Cash	0.008***	-0.004***	-0.019***	-0.001
	[5.54]	[-5.97]	[-2.88]	[-0.34]
Peer firm averages				
p_MTB	-0.000	0.000	-0.003***	0.000
	[-0.75]	[0.83]	[-4.19]	[0.69]
p_Cash flow	0.002***	0.000	-0.019***	0.005***
	[6.00]	[1.28]	[-9.31]	[7.50]
p_Ln_at	0.000***	0.000**	0.005***	0.000
	[3.25]	[2.32]	[15.05]	[1.58]
p_Leverage	-0.001*	-0.000	-0.011***	-0.002
	[-1.91]	[-1.04]	[-6.27]	[-1.63]

Continued on next page

Table 4.8 - continued from previous page

	Dividend	Capital expense	R&D	Advertising expense
Variables	**(1)**	**(2)**	**(3)**	**(4)**
p_Cash	-0.015***	-0.000	0.151***	0.003
	[-10.64]	[-0.29]	[27.75]	[1.45]
Constant	0.006***	0.015***	0.009*	-0.002
	[2.79]	[23.11]	[1.82]	[-0.98]
Year fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Observations	46,739	46,739	46,739	46,739
Adj. R^2	0.218	0.311	0.631	0.287

Figure 4.1. A Placebo test: This figure addresses the concern that our results may reflect unobserved correlations between industries. Specifically, we use randomly assigned peer investment groups into the baseline regression regression (4.1) and it proves these random peer investment groups do not have significant peer effect.



Chapter 5

Conclusions

This thesis explains three aspects of empirical finance studies. In the first chapter, using all the U.S. common equity data from 1926 to 2017, we examine the relation between stock beta's statistical significance and BAB portfolio alphas. We find that BAB portfolio alphas can be significantly decreased when dropping stocks with statistically insignificant betas. If we replace statistically insignificant betas as zeros and re-run the portfolio construction, a refined BAB trading strategy can generate a higher BAB alpha than original BAB strategy. The negative impact of betas' statistical significance on BAB portfolio alphas remains robust in several robustness tests and in major international equity markets. Our evidence reveals the importance of betas' statistical significance on BAB strategy. The implications of my research contribute to minimize the academic publication bias in finance research for publishing on top-tier journals and encourage academics to re-examine the existing factors instead of endlessly discovering new factors. My findings shed new light on the potential impact of ignoring statistical significance in empirical asset pricing studies. My research also has important impact for practitioners on their portfolio management when they trade based on a well-documented stock return anomaly, and policymakers can also learn from my work that there are always arbitrageurs who can exploit anomalous trading opportunities. One limitation of this work is that we do not know if this framework can be applied to other empirical asset pricing studies. I suggest that future research should explore and try if it can be used on any other anomaly factors.

Prior evidence that making investment decreases the firm's future stock returns leaves unknown answer about which part of and how investment results in such negative relation. Therefore, I study this negative investment-return relation and question whether abnormal investment can explain the negative relation between firm investment and future stock returns. Then, I split abnormal investment into under-investment and over-investment. In this case, the research question is whether a decline of firm under-investment or a decline of firm over-investment leads to an increase in its future stock returns. The evidence shows that it is firm under-investment that leads to this negative relation, rather than over-investment. One explanation is that firms that are expected to have higher stock returns are found to deviate less from predicted investment levels, which indicates they have less abnormal investment. These results suggest that both delayed market reaction and agency issues may lead to the anomalous return predictability of under-investment. We suggest that firms should control their levels of investment and avoid the deviation from predicted level of investment as it comes at a cost. The limitations for this study are few. Even though the essential channels of the causal link between investment and future stock returns are identified, we do not examine the potential impact brought by a transitional change from firm under-investment to the over-investment. We also do not take into account of monopolist effect on the negative relation. Another limitation is the proper use of effective measures for firms to lessen the under-investment but also not to over-invest. I suggest that future research could examine how to avoid making abnormal investment.

Many parallel peer effects research studies are conducted in recent literature, such as peer investment to firm investment (Bustamante and Frésard, 2017), and peer cash flows to firm cash flows (Chen et al., 2019). It attracts our attention and raises a further interest of whether peer effects can occur in a non-parallel way, so we study the effect of peer investment on the firms' value of cash holdings. We find that they are positively related because of the positive externalities brought by peer investment which are shared by focal firms in growing and young industries with many investment opportunities. And we do not find evidence to support alternative hypotheses: the precautionary demand hypothesis and observational learning hypothesis. The findings imply that the positive externalities led by peer investment are more likely to be the drivers of our results. We advise that firms should keep a close eye on peer firms' investment agenda as the externalities generated through it can spill over the entire industry which would affect investors' valuation on firms' own cash holdings. If you know the enemies and know yourself, you can fight a hundred battles with no fear of defeat. The limitations of this study also exist. Although we can identify such a positive and significant relationship, we do not know how long the impact of peer firms' investment would reflect on focal firms' value of cash. We do not account for other forms of investment except capital investment when measuring the peer investment, such as R&D investment and acquisitions. We also do not know if it exists in an international context. I suggest the future research could examine a broader impact of peer investment.

Bibliography

Abarbanell, J. S. and Bernard, V. L. (1992). Tests of analysts' overreaction/underreaction to earnings information as an explanation for anomalous stock price behavior. *Journal of Finance*, 47:1181–1207.

Adhikari, B. K. and Agrawal, A. (2018). Peer influence on payout policies. *Journal of Corporate Finance*, 48:615–637.

Aghion, P., Reenen, J. V., and Zingales, L. (2013). Innovation and institutional ownership. *The American Economic Review*, 103:277–304.

Akerlof, G. A. and Shiller, R. J. (2010). *Animal spirits: How human psychology drives the economy, and why it matters for global capitalism.* Princeton University Press.

Almeida, H., Campello, M., and Weisbach, M. S. (2004). The cash flow sensitivity of cash. *Journal of Finance*, 59(4):1777–1804.

Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X. (2006). The cross-section of volatility and expected returns. *Journal of Finance*, 61:259–299.

Ang, J. S., Cole, R. A., and Lin, J. W. (2000). Agency costs and ownership structure. *Journal of Finance*, 55:81–106.

Arif, S. and Lee, C. M. (2014). Aggregate investment and investor sentiment. *Review of Financial Studies*, 27:3241–3279.

Asness, C. S., Frazzini, A., and Pedersen, L. H. (2014). Low-risk investing without industry bets. *Financial Analysts Journal*, 70(4):24–41.

Auer, B. R. and Schuhmacher, F. (2015). Liquid betting against beta in dow jones industrial average stocks. *Financial Analysts Journal*, 71(6):30–43.

Baker, M., Bradley, B., and Wurgler, J. (2011). Benchmarks as limits to arbitrage: Understanding the low-volatility anomaly. *Financial Analysts Journal*, 67(1):40–54.

Bakke, T.-E. and Gu, T. (2017). Diversification and cash dynamics. *Journal of Financial Economics*, 123(3):580–601.

Bakke, T.-E. and Whited, T. M. (2010). Which firms follow the market? An analysis of corporate investment decisions. *Review of Financial Studies*, 23:1941–1980.

Bali, T. G., Brown, S. J., Murray, S., and Tang, Y. (2017). A lottery-demand-based explanation of the beta anomaly. *Journal of Financial and Quantitative Analysis*, 52(6):2369–2397.

Bali, T. G., Cakici, N., and Whitelaw, R. F. (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics*, 99:427–446.

Bates, T. W., Chang, C.-H., and Chi, J. D. (2018). Why has the value of cash increased over time? *Journal of Financial and Quantitative Analysis*, 53(2):749–787.

Bates, T. W., Kahle, K. M., and Stulz, R. M. (2009). Why do us firms hold so much more cash than they used to? *Journal of Finance*, 64(5):1985–2021.

Bebchuk, L., Cohen, A., and Ferrell, A. (2008). What matters in corporate governance? *Review of Financial Studies*, 22(2):783–827.

Benoit, J.-P. (1984). Financially constrained entry in a game with incomplete information. *The RAND Journal of Economics*, pages 490–499.

Bergman, Y. Z. and Callen, J. L. (1991). Opportunistic underinvestment in debt renegotiation and capital structure. *Journal of Financial Economics*, 29:137–171.

Bernard, D., Blackburne, T., and Thornock, J. (2020a). Information flows among rivals and corporate investment. *Journal of Financial Economics*, 136(3):760–779.

Bernard, D., Ge, W., Matsumoto, D., and Toynbee, S. (2020b). Implied tradeoffs of chief financial officer accounting expertise: Evidence from firm-manager matching. *Management Science*.

Bertrand, M. and Mullainathan, S. (2003). Enjoying the quiet life? Corporate governance and managerial preferences. *Journal of Political Economy*, 111:1043–1075.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1998). Learning from the behavior of others: Conformity, fads, and informational cascades. *Journal of economic perspectives*, 12(3):151–170.

Bustamante, M. C. and Frésard, L. (2017). Does firm investment respond to peers' investment? *Robert H. Smith School Research Paper No. RHS*, 2827803.

Bustamante, M. C. and Frésard, L. (2020). Does firm investment respond to peers' investment? *Management Science*.

Campbell, J. Y., Champbell, J. J., Campbell, J. W., Lo, A. W., Lo, A. W., and MacKinlay, A. C. (1997). *The econometrics of financial markets*. Princeton University Press.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, 52(1):57–82.

Caskey, J., Hughes, J., and Liu, J. (2012). Leverage, excess leverage, and future returns. *Review of Accounting Studies*, 17:443–471.

Chen, Q., Goldstein, I., and Jiang, W. (2007). Price informativeness and investment sensitivity to stock price. *Review of Financial Studies*, 20:619–650.

Chen, Y.-W., Chan, K., and Chang, Y. (2019). Peer effects on corporate cash holdings. *International Review of Economics & Finance*, 61:213–227.

Cooper, M. J., Gulen, H., and Schill, M. J. (2008). Asset growth and the cross-section of stock returns. *Journal of Finance*, 63:1609–1651.

Cortés, K., Duchin, R., and Sosyura, D. (2016). Clouded judgment: The role of sentiment in credit origination. *Journal of Financial Economics*, 121:392–413.

Denis, D. J. and Sibilkov, V. (2010). Financial constraints, investment, and the value of cash holdings. *Review of Financial Studies*, 23(1):247–269.

Dimson, E. (1979). Risk measurement when shares are subject to infrequent trading. *Journal of Financial Economics*, 7(2):197–226.

Dittmar, A. and Mahrt-Smith, J. (2007). Corporate governance and the value of cash holdings. *Journal of Financial Economics*, 83(3):599–634.

Drobetz, W., Grüninger, M. C., and Hirschvogl, S. (2010). Information asymmetry and the value of cash. *Journal of banking & finance*, 34(9):2168–2184.

Duchin, R. (2010). Cash holdings and corporate diversification. *Journal of Finance*, 65(3):955–992.

Edmans, A. (2014). Blockholders and corporate governance. *Annual Review of Financial Economics*, 6:23–50.

Edmans, A., Garcia, D., and Norli, Ø. (2007). Sports sentiment and stock returns. *Journal of Finance*, 62:1967–1998.

Eugene, F. and French, K. (1992). The cross-section of expected stock returns. *Journal of Finance*, 47(2):427–465.

Fairfield, P. M., Whisenant, J. S., and Yohn, T. L. (2003). Accrued earnings and growth: Implications for future profitability and market mispricing. *The Accounting Review*, 78:353–371.

Fairfield, P. M. and Yohn, T. L. (2001). Using asset turnover and profit margin to forecast changes in profitability. *Review of Accounting Studies*, 6(4):371–385.

Fairhurst, D. and Nam, Y. (2020). Corporate governance and financial peer effects. *Financial Management*, 49(1):235–263.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33:3–56.

Fama, E. F. and French, K. R. (2008). Dissecting anomalies. *Journal of Finance*, 63:1653–1678.

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

Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81:607–636.

Faulkender, M. and Wang, R. (2006a). Corporate financial policy and the value of cash. *Journal of Finance*, 61:1957–1990.

Faulkender, M. and Wang, R. (2006b). Corporate financial policy and the value of cash. *Journal of Finance*, 61(4):1957–1990.

Feltham, G. A. and Ohlson, J. A. (1996). Uncertainty resolution and the theory of depreciation measurement. *Journal of Accounting Research*, 34:209–234.

Foucault, T. and Fresard, L. (2014). Learning from peers' stock prices and corporate investment. *Journal of Financial Economics*, 111(3):554–577.

Frazzini, A. and Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics*, 111(1):1–25.

Fresard, L. (2010). Financial strength and product market behavior: The real effects of corporate cash holdings. *Journal of Finance*, 65(3):1097–1122.

Frésard, L., Hoberg, G., and Phillips, G. (2015). Innovation activities and the incentives for vertical acquisitions and integration. *Social Science Research Network, working paper.*

Gao, P. and Zhang, G. (2019). Accounting manipulation, peer pressure, and internal control. *The Accounting Review*, 94(1):127–151.

Garcia, D. (2013). Sentiment during recessions. *Journal of Finance*, 68:1267–1300.

Gelb, D. S. and Zarowin, P. (2002). Corporate disclosure policy and the informativeness of stock prices. *Review of accounting studies*, 7(1):33–52.

Gomes, J., Kogan, L., and Zhang, L. (2003). Equilibrium cross section of returns. *Journal of Political Economy*, 111:693–732.

Gompers, P., Ishii, J., and Metrick, A. (2003). Corporate governance and equity prices. *Quarterly Journal of Economics*, 118(1):107–156.

Grennan, J. (2019). Dividend payments as a response to peer influence. *Journal of Financial Economics*, 131(3):549–570.

Guerrieri, V. and Kondor, P. (2012). Fund managers, career concerns, and asset price volatility. *American Economic Review*, 102:1986–2017.

Harford, J., Mansi, S. A., and Maxwell, W. F. (2008). Corporate governance and firm cash holdings in the us. *Journal of financial economics*, 87(3):535–555.

Hart, O. D. (1983). The market mechanism as an incentive scheme. *The Bell Journal of Economics*, 14:366–382.

Harvey, C. R., Lins, K. V., and Roper, A. H. (2004). The effect of capital structure when expected agency costs are extreme. *Journal of Financial Economics*, 74:3–30.

Harvey, C. R., Liu, Y., and Zhu, H. (2016). ... and the cross-section of expected returns. *Review of Financial Studies*, 29(1):5–68.

Hennessy, C. A., Levy, A., and Whited, T. M. (2007). Testing Q theory with financing frictions. *Journal of Financial Economics*, 83(3):691–717.

Hirshleifer, D. and Shumway, T. (2003). Good day sunshine: Stock returns and the weather. *Journal of Finance*, 58:1009–1032.

Hoberg, G. and Phillips, G. (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, 124(5):1423–1465.

Hoberg, G. and Phillips, G. M. (2015). Industry choice and product language. *Available at SSRN 1858007*.

Horenstein, A. R. (2017). Betting against alpha. *Working Paper,*, University of Miami.

Hou, K., Xue, C., and Zhang, L. (2015). Digesting anomalies: An investment approach. *Review of Financial Studies*, 28:650–705.

Hubbard, R. G. (1998). Capital–market imperfections and investment. *Journal of Economic Literature*, 36:193–225.

Humphery-Jenner, M. and Suchard, J.-A. (2013). Foreign venture capitalists and the internationalization of entrepreneurial companies: Evidence from china. *Journal of International Business Studies*, 44(6):607–621.

Jensen, M. C. (1986). Agency costs of free cash flow, corporate finance, and takeovers. *The American Economic Review*, 76:323–329.

Jensen, M. C., Black, F., and Scholes, M. S. (1972). The capital asset pricing model: Some empirical tests.

Kim, K., Mauldin, E., and Patro, S. (2014). Outside directors and board advising and monitoring performance. *Journal of Applied Econometrics*, 57(2-3):110–131.

Kogan, L. and Papanikolaou, D. (2013). Firm characteristics and stock returns: The role of investment-specific shocks. *Review of Financial Studies*, 26:2718–2759.

Leary, M. T. and Roberts, M. R. (2014). Do peer firms affect corporate financial policy? *Journal of Finance*, 69(1):139–178.

Lee, D., Shin, H., and Stulz, R. M. (2016). Why does capital no longer flow more to the industries with the best growth opportunities? *Working Paper, Available at SSRN 2839832*.

Lewellen, J. and Nagel, S. (2006). The conditional capm does not explain asset-pricing anomalies. *Journal of financial economics*, 82(2):289–314.

Li, J. Y. and Tang, D. Y. (2016). The leverage externalities of credit default swaps. *Journal of Financial Economics*, 120(3):491–513.

Li, X., Sullivan, R. N., and Garcia-Feijóo, L. (2014). The limits to arbitrage and the low-volatility anomaly. *Financial Analysts Journal*, 70(1):52–63.

Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *Journal of Finance*, 20(4):587–615.

Liu, J., Stambaugh, R. F., and Yuan, Y. (2018). Absolving beta of volatility's effects. *Journal of Financial Economics*, 128(1):1–15.

Liu, L. X., Whited, T. M., and Zhang, L. (2009). Investment-based expected stock returns. *Journal of Political Economy*, 117:1105–1139.

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *Journal of Finance*, 42(3):483–510.

Myers, S. C. (1977). Determinants of corporate borrowing. *Journal of Financial Economics*, 5:147–175.

Myers, S. C. and Rajan, R. G. (1998). The paradox of liquidity. *The Quarterly Journal of Economics*, 113(3):733–771.

Novy-Marx, R. (2014). Understanding defensive equity. Technical report, National Bureau of Economic Research.

Ohlson, J. A. (1995). Earnings, book values, and dividends in equity valuation. *Contemporary accounting research*, 11:661–687.

Pástor, L. and Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, 111(3):642–685.

Petersen, M. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies*, 22:435–480.

Polk, C. and Sapienza, P. (2009). The stock market and corporate investment: A test of catering theory. *Review of Financial Studies*, 22:187–217.

Richardson, S. (2006). Over-investment of free cash flow. *Review of Accounting Studies*, 11:159–189.

Scharfstein, D. S. and Stein, J. C. (1990). Herd behavior and investment. *The American economic review*, pages 465–479.

Shane, P. and Brous, P. (2001). Investor and (value line) analyst underreaction to information about future earnings: The corrective role of non-earnings-surprise information. *Journal of Accounting Research*, 39:387–404.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442.

Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *Journal of Business*, 74:101–124.

Smith, C. A. and Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology*, 48:813.

Stoughton, N. M., Wong, K. P., and Yi, L. (2017). Investment efficiency and product market competition. *Journal of Financial and Quantitative Analysis*, 52:2611–2642.

Tiedens, L. Z. and Linton, S. (2001). Judgment under emotional certainty and uncertainty: The effects of specific emotions on information processing. *Journal of Personality and Social Psychology*, 81:973.

Titman, S., Wei, K. J., and Xie, F. (2004). Capital investments and stock returns. *Journal of Financial and Quantitative Analysis*, 39:677–700.

Vasicek, O. A. (1973). A note on using cross-sectional information in Bayesian estimation of security betas. *Journal of Finance*, 28(5):1233–1239.

Walpole, R., Myers, R., Myers, S., and Ye, K. (2002). *Probability and Statistics for Engineers and Scientists, 7th Edition*. Pearson Education.

Xing, Y. (2008). Interpreting the value effect through the q-theory: An empirical investigation. *The Review of Financial Studies*, 21(4):1767–1795.