UAB
Universitat Autònoma de Barcelona

Essays in Empirical Asset Pricing

Dijun Liu

Essays in Empirical Asset Pricing

Dijun Liu

Supervisor: Professor Abhay Abhyankar

IDEA Program
Universitat Autònoma de Barcelona

This dissertation is submitted for the degree of
Doctor of Philosophy in Economics

May 2023

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the acknowledgements.

Dijun Liu
May 2023

Acknowledgements

I would like to express my heartfelt gratitude to all the people who have supported me throughout my Ph.D. journey in the past six years in Barcelona, which has been a challenging yet rewarding experience. I had my best years here, which I will cherish forever.

Firstly, I am deeply grateful to my supervisor, Abhay Abhyankar, for his guidance, patience, and encouragement throughout my research. Thanks to his support, I can focus on the topic that interests me the most. As a seasoned professor with extensive knowledge, he has been instrumental in guiding me through the vast world of finance, helping me navigate where to start, and motivating me toward the academic frontiers. His keen insights and broad perspective have been precious, as he often offers useful suggestions and instructions that have helped to shape my ideas and enhance my analytical skills. Besides, he has provided me with unwavering support whenever I need his help, like applying for an exchange.

I am also grateful to my tutor, Professor Michael Creel, and Professor Jordi Caballé, who are also my referees. They always attended my annual evaluation and provided constructive feedback. The faculty in the IDEA program has been exceptional. Professor Inés Macho-Stadler has not only helped me to revise my website but has also shared numerous job market positions with me. Àngels López and Mercè Vicente have demonstrated incredible efficiency and have always been quick to prepare any documents I have needed. Moreover, all the professors who once taught me courses and attended my job market presentation and mock interviews are incredibly kind and patient, and they provide me with valuable feedback and suggestions. I learned a great deal from them, and I always enjoyed talking with them. I greatly appreciate our department and the Catalan government for the scholarship I received to make an undistracted and supportive study possible.

I am incredibly thankful to my parents, who have always been my rock and a constant source of motivation. Talking to them when I felt depressed or cried refreshes me. Their unconditional love and support have given me the courage to pursue my dreams. They never expect much from me, but their caring for my health and my happiness has always been enough to keep me going.

I would also like to express my gratitude to my friends, colleagues, and professors in Exeter: Bo, Hongfei, Ozlem, Huitong, Mridula, Yaning, Yudi, Shangdi, Christina, and others. Their support, encouragement, and companionship have been invaluable throughout my Ph.D. journey. I have had incredible years here in Barcelona due to the vibrant culture and beautiful surroundings. From the delicious food, architecture, and museums to the warm sunshine, natural parks, and mountains, every moment has been a source of great happiness and relief. Moreover, I have enjoyed engaging in cultural activities such as attending operas and concerts and visiting museums, which gave me a glimpse into Western Art's charm. I will always treasure the memories of my time here.

Last but not least, I give my special thanks to my supervisor back in my undergraduate studies, Yanyun Zhao. He was the one who encouraged me to study abroad and kickstarted my journey here. He told me that studying abroad was a difficult journey, but I would emerge from it like a phoenix rising from the ashes, and he encouraged me to contribute to society and the world. His passing has left me with profound sadness but has also prepared me to cope with the loss of my uncle and grandmother in the following years. While the pain of their loss still lingers, I have learned to cherish the people I love and hold them dear in my heart forever.

Abstract

This dissertation consists of three independent chapters that explore market anomalies and employ the present value decomposition method to analyze stock returns. The first chapter investigates the underperformance of the value strategy in recent decades and examines the impact of intangible capital on value investing. The second chapter, motivated by the close relationship between the book-to-market ratio and estimated equity duration, explores whether stocks with longer duration display higher sensitivity to the discount rate news and whether the short duration premium serves as a substitute for the value anomaly. In the third chapter, we apply the present value decomposition methodology to the behavioral factors to explore the primary driver of these portfolio returns, shedding light on the sources of their explanatory power and systematic risk.

Keywords: Characteristics, Anomalies, Factor model, Risk premia, Equity duration, Value premium, Short-sale constraints, Present value decomposition, Mispricing factor, Behavioral factor, Good beta bad beta.

JEL classification: G1, G10, G12, G14, G17.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Market anomalies are an important area of research in empirical finance due to the challenge they present to asset pricing models. The discovery and exploration of anomalies helps us understand the economic forces that affect risk in the stock market and inspires new asset pricing models that seek to explain why anomalies arise. The existence of anomalies has led to trading strategies implemented by traders and fund managers to achieve abnormal risk-adjusted returns. This dissertation seeks to contribute to the literature on market anomalies and factor models by focusing mainly on the value anomaly and on factors in recently proposed behavioral asset pricing models.

Chapter 2 studies the evolution of the value premium over time and across firms. We first revisit the value premium using a recent sample period and compare it to that over an earlier sample period used in previous work. Second, using three different methods for measuring intangible capital, we estimate an intangible adjusted value-growth premium and compare it to that using the conventional definition of book value using the sample period from 1976 to 2019. We supplement this with two sub-sample analyses that focus on the period before and after 1999–the year in which the number of listed stocks was at its peak over our sample period. In addition, we also study the difference between the value premium for high technology firms versus traditional industry firms. Third, we dig deeper into the decline in the returns of the value strategy. We investigate how, firms that populate extreme decile portfolios sorted on different definitions of the book-to-market ratio differ along on several characteristics including leverage, investment, and return on investment among others. Next, we study the value premium when the book-to-market ratio is computed using only the value of intangible capital for book value. Then, we turn to the High-Minus-Low (HML) portfolios created using the procedure followed for it in the Fama-French three-factor model. We use different book-to-market ratios, to study if they differ in their cumulative returns and their ability to span the mean-variance portfolios of factor models. Finally, we explore the role of the relative valuation on the returns to the HML factors.

Our main results are as follows. First, we confirm, as reported in previous research, that returns to the value strategy are lower from 1980 to 2019 compared to that from 1963 to 2012. To mitigate the concern of problems caused by the overlapping sample periods, we also report the returns for the period from 1976 to 1998 and from 1999 to 2019 when comparing different versions of the book-to-market ratio. We find that the returns are much smaller and not significantly different from zero for the more recent sample period.

Second, we find that accounting for intangible capital increases the returns to the value strategy, but the intangible adjusted value premium is also smaller post-1999, similar to that when using the traditional definition of the book value. When firms are classified into the high tech sector or traditional sector, we find that the effects of including intangible capital are prominent in both sectors.

Third, we find that the firms in the long leg sorted on intangible adjusted book-to-market ratio are less profitable, have lower levels of operating capital, and have higher leverage than firms in the short leg. This pattern is contrary to that when sorted on the traditional book-to-market ratio: the firms in the short leg sorted on the traditional book-to-market ratio are less profitable, have lower levels of operating capital, and have higher leverage than firms in the long leg. We confirm the positive relationship between the intangible capital over market value ratio (IM) and future stock returns, though the long-short decile portfolio returns are not significant or only marginally significant. We also find that the HML factors, except those constructed using IM, all experience large and persistent drawdowns, and deliver different spanning results depending on the contexts; the relative valuation is closely related to the returns of HML factors suffering from persistent drawdowns.

Chapter 3 studies the negative relationship between the book-to-market ratio and equity duration. Intuitively, a stock with a higher duration generates a larger portion of its cash flow over a longer horizon and as such is more sensitive to changes in the discount rate. We first study whether stocks with high duration are more sensitive to discount rate news obtained using the present value decomposition approach. Gormsen and Lazarus (2023) find that value, profitability, investment, low-risk, and payout factors have duration-based explanations. We also study if the duration premium explains the value anomaly by exploring the duration premium using sub-sample univariate portfolio analysis and comparing its power to predict future stock returns with different book-to-market ratios using Fama-MacBeth regression. Finally, we explore the effects of short-sale constraints and market capitalization on the short duration premium.

Our main results are as follows. First, we find that the cash flow news is the primary driver across portfolios composed of stocks with long or short duration. The variance of cash flow news over the variance of return news is always larger than that of discount rate news over return news.

Second, we find that the duration premium is relatively consistent compared with the value premium. In univariate portfolio analysis, the value-weighted return to a long-short portfolio is about -0.10 for sample periods from 1965 to 2020, from 1965 to 1998, and from 1999 to 2020. Using Fama-Macbeth regressions we find that duration subsumes the ability of the traditional book-to-market ratio to forecast future stock returns but not that of the intangible adjusted book-to-market ratio.

Finally, we find that the short duration premium is mainly concentrated in small and short-sale constrained stocks.

Chapter 4 applies the present value decomposition methodology to the behaviorally motivated factors in recent extension of the Fama-French three-factor model as well as to anomalies. Our motivation is to understand whether behavioral factors and traditional factors differ in their cash flow and discount rate decompositions. In contrast to econometrics, we explore whether these new factors reflect fundamental drivers of stock returns (i.e. cash flow and discount rate news) differently. We obtain the return decomposition using the standard Vector autoregression (VAR) methodology of Campbell and Shiller (1988) and the subsequent literature using individual stocks. We then obtain portfolio-level cash flow and discount rate components by suitable aggregation into anomaly, factor, and mean-variance efficient levels. We use these aggregate news components to analyze the driver of returns by comparing the ratio of the variance of each news to the variance of return news. Second, we do regressions using the news components to explore which part gives the explanatory power of these factors. Third, we leverage the beta decomposition to analyze the systematic risk associated with these factors.

Our main results are as follows. First, we find that the cash flow news is the main driver for all the underlying anomalies used to construct these factors, the factors themselves, and their mean-variance efficient frontier.

Second, we find that when regressing anomalies' return news on factors' news components, the adjusted R squared is much larger for cash flow news than discount rate news. Also, the magnitude of adjusted R squared using only the cash flow news as an independent variable is similar to that using the factors' return news as an independent variable. It, therefore, implies that the explanatory power of these factors stems from their cash flow news.

Third, we find that the cross-sectional variation of market betas also closely relates to the variation of beta components associated with cash flow news. Therefore, we highlight the importance of the fundamental cash flow of these factors. From this perspective, there is no essential difference between behavioral and rational models. This supports the argument in Kozak, Nagel, and Santosh (2018) that these factors are just a combination of stocks that may capture the main principal components through different repackaging of stock returns.

Overall, this dissertation examines the value strategy and its recent developments with intangible capital, investigates the short duration premium, and uses the present value decomposition to analyze behavioral factors. It contributes to the existing literature on market anomalies and factor models by investigating several important anomalies and exploring their latest advancements. The present value decomposition approach is extensively employed throughout the dissertation to examine the contribution of cash flow news and discount rate news to the unexpected return news of stocks with different characteristics. The findings of this study have significant implications for both academics and practitioners, serving as a foundation for further research in this area.

There are several avenues for extending this work. Firstly, it would be beneficial to utilize international data for comparative analysis. Additionally, incorporating the new decomposition method proposed by Cho, Kremens, Lee, and Polk (2022) or others could enhance the robustness of our analysis or offer novel perspectives. Furthermore, expanding the scope of anomalies and factors that can be examined through present value decomposition would be valuable. For example, it can shed light on the reduction of anomalies. Apart from statistical models such as LASSO, PCA, similarity analysis, and new Bayesian models, which address the exploding number of anomalies in the factor zoo, some papers begin from a theoretical perspective or incorporate a number of anomalies to create their factors. The reason we use present value decomposition is that the returns are ultimately determined by either cash flow or discount rate. Numerous characteristics, such as ROE, investment rate, historical anomalies, O-score, etc., convey either cash flow news or discount rate news, or both. By applying the present value decomposition to more anomalies and factors, we will be able to conduct the analysis at the cash flow news level or discount rate news level. This classification can help identify anomalies primarily driven by cash flow news or discount rate news, enabling effective categorization or reduction. Moreover, exploring the similarities among the news components of these anomalies can provide insights. For instance, if the cash flow news of certain anomalies is highly correlated, one of these anomalies may suffice to represent the variation in returns and the underlying economic rationale. Thus, variance decomposition can help the reduction of anomalies and facilitate comparisons of factor models.

Chapter 2

Does Intangible Capital Affect Returns to the Value-Growth Strategy?

2.1 Introduction

This chapter investigates the performance of the value-growth premium across firms and over time. The value-growth premium refers to the difference in average returns between stocks with high book-to-market ratios (value stocks) and those with low book-to-market ratios (growth stocks). This premium is an anomaly, as it generates positive returns when adjusted for risk using the Capital Asset Pricing Model (CAPM) or empirical asset pricing models. The value premium is a prominent strategy employed by traders and investors that seek to exploit these returns, which is studied thoroughly in this chapter. First, we revisit the return on this strategy and find that its profitability has been lower recently. Second, we note that the traditional BM ratio is based on the book value of tangible capital. However, in recent years, firms that have invested in intangible capital have dominated the largest firms in the stock market. Thus, we examine if the value-growth premium accounting for intangible capital performs differently from sorts using the conventional definition of a firm's book value. We employ three distinct methods for estimating intangible capital and compare their respective performances. Additionally, we delve deeper into the decline in the profitability of the value strategy by analyzing the firms' characteristics in the long leg and short leg sorted on different book-to-market ratios, using the book-to-market ratio where the numerator is only the value of intangible capital, constructing the High-Minus-Low (HML) factors, and exploring the role of the relative valuation on the returns to the HML factors.

The book-to-market (BM) ratio, which is the ratio of a firm's equity book value to its equity market value, is crucial in both the theory and practice of finance. The academic literature on asset pricing has extensively documented a positive relationship between a firm's BM ratio and future stock returns. This finding has led to the development of the value premium, which is the excess return that can be earned by employing the value strategy of holding long positions in value stocks and short positions in growth stocks. To form the value anomaly or long-short return, the conventional approach is to sort stocks on their BM and group them equally into ten portfolios. After calculating the average returns for each portfolio, the long-short returns are the difference in returns between the portfolio composed of stocks with the highest BM and the portfolio with the lowest BM. This anomaly earns positive returns in excess of what is predicted by the CAPM and other empirical asset pricing models. Based on this anomaly, Fama and French (1993) propose the HML factor, or value factor, in the famous three-factor model, which has been a workhorse in finance since 1993 as an extension of the Sharpe-Lintner-Mossion CAPM. The construction of HML is different from long-short returns, as the former eliminates the size effect. Each month, stocks listed on the NYSE, American Stock Exchange (AMEX), and NASDAQ are classified as big or small based on their market capitalization using the median NYSE market capitalization as the breakpoint. They are then independently divided into three groups based on their BM, using the 30th and 70th percentiles of BM among NYSE stocks. The HML factor is the value-weighted average return of a small and a big value portfolio minus the value-weighted average return of a small and a big growth portfolio.

In the realm of corporate finance, the modern theory of firm investment has demonstrated that investment is a function of marginal q, which denotes the increase in a firm's value if an additional unit of capital is invested. Tobin's Q is defined as the market value of firm capital over the replacement cost of a firm's capital. In empirical studies, the book value is often used as a proxy for the replacement cost. Thus, the BM ratio has emerged as a widely used proxy for the inverse of Tobin's Q. In Peters and Taylor (2017), the authors suggest a simple, new Tobin's Q proxy that accounts for intangible capital. They show that it is a superior proxy for both physical and intangible investment opportunities and that Tobin's Q also explains intangible investment better than tangible investment. Kogan and Papanikolaou (2014) utilizes the BM ratio as one proxy for growth opportunities.

In practice, the BM ratio is employed in value-growth investment strategies by many mutual and hedge funds, including Dimensional Fund Advisors, which manages 454 billion dollars in firm-wide assets.

However, the strategy has been losing its power in recent decades as growth stocks have outperformed value stocks on average by delivering higher returns. The HML factor experiences long and persistent drawdowns. We first confirm this phenomenon by replicating all the tables in Chapter 10 of Bali, Engle, and Murray (2016) and then extending the sample to a later period from 1980 to 2019 using the traditionally defined BM ratio. We obtain very similar results for the sample period from 1964 to 2012 and thus ensure that all the tricky procedures are followed correctly. The portfolio analysis in the later sample shows that the value premium is insignificant for the value-weighted long-short returns, indicating the deterioration of the value strategy.

Several reasons may have contributed to the decline of the value strategy, including the stock market becoming more efficient, lower transaction costs eliminating opportunities for free lunch, quicker information dissemination due to technological advancements, deficient construction methodology, the discovery effect, and so on. Böll, Thimme, and Uhrig-Homburg (2022) shows that optionability plays a role in anomaly returns. For stocks with traded options, they exhibit a much smaller long-short value premium than those without traded options. The difference is significant even after adjusting size and liquidity. Hasler (2021) explores the construction of the original HML portfolio, identifying six seemingly innocuous decisions that can affect the portfolio's performance. Soebhag, Van Vliet, and Verwijmeren (2022) further studies eleven construction choices when constructing several factors. They argue that factor returns are not only a function of their sorting characteristic but also a function of their construction choices. McLean and Pontiff (2016) and Dong, Liu, Lu, Sun, and Yan (2022) document that the discovery of an anomaly affects the investors' trading behavior by learning from academic publications. Among all the possible explanations, one concerns the misvaluation of the book value of equity. The most commonly applied approach is from Fama and French (1993), which we will document in detail in Section 2.3.2.

The underlying notion of misvaluation is that the conventional book value of equity primarily assesses the worth of tangible capital yet disregards intangible capital. However, since the last decade of the twentieth century, there has been an increase in listed companies such as Amazon, Google, and others that possess more intangible assets, including copyrights, human resources, electronic databases, and technologies, as opposed to tangible assets like buildings, machines, and equipments that are the primary assets of manufacturing sector firms. However, under existing accounting standards, the conservatism principle prohibits internally generated intangible capital from being recorded on the balance sheets. This principle necessitates that company accounts be prepared with caution and high verification levels. Due to the non-physical and non-financial nature of intangible assets, their valuation presents a significant challenge. Although patents, brands, trademarks, and copyrights are currently listed as long-term assets on the balance sheet, these items are primarily acquired from outside firms, and the amount reflects the cost of purchasing those externally produced items. Internally generated intangible assets, on the other hand, are absent from financial reports. Only minor direct costs incurred in the development of intangible assets, such as legal fees, are capitalized, while the remainder is expensed. As a result, expenses associated with creating internally generated assets, such as research and development, advertising, training, and others, are recognized as costs, reducing the firms' assets. Several models have been proposed that strive to integrate intangible capital in addition to traditional physical capital (see Corrado, Hulten, and Sichel (2009), Eisfeldt and Papanikolaou (2013), and Koh, Santaeulàlia-Llopis, and Zheng (2020) for instance).

Therefore, some papers argue that the misvaluation of BM causes the deterioration of the value strategy and that we should estimate and include the intangible capital when calculating the book value of equity. Park (2022) proposes revising the calculation of book value by including intangible assets in the calculation of book value of equity and finds that an intangible adjusted BM ratio still predicts stock returns. Ewens, Peters, and Wang (2020) and Eisfeldt, Kim, and Papanikolaou (2022) have similar observations. Vincenz (2023) confirms the positive effects of incorporating intangible capital not only for the U.S. stock market but globally in Europe, Japan, and Asia-Pacific as well. Gulen, Li, Peters, and Zekhnini (2020) underscore the significant role of off-balance-sheet intangible capital and intangible investment in factor models. Kazemi (2022) constructs a portfolio double-sorted on two key firm characteristics, the book-to-market ratio (including intangible capital) and the difference between the intangible and tangible investment rates. This portfolio generates large excess returns that existing models cannot explain. However, Rizova and Saito (2021) disfavor the adjustment because estimating intangible assets involves much noise, and the estimated intangible assets provide little additional information about future cash flow and profitability.

We compare the traditional BM and three versions of the intangible adjusted BM ratio ($BM^{Eisfeldt}$, BM^{Peters}, and BM^{Ewens}) by applying different methods to estimate the intangible capital. We perform two sub-sample analyses to investigate the performance of intangible adjustment in both time-series and cross-sectional aspects. In the first sub-sample analysis, we partition the sample period by the year 1999 due to the fact that the number of listed firms peaked at the end of 1997 and rapidly decreased after that, and publicly listed firms are typically larger and older. Our results indicate that the intangible adjusted BM produces higher long-short portfolio returns than the traditional BM. Nonetheless, even the intangible adjusted value premium is considerably smaller. The cross-sectional sub-sample analysis demonstrates that the improvement resulting from incorporating intangible capital is noticeable even in the traditional sector with less reliance on intangible capital.

To comprehend why the inclusion of intangible capital enhances the value strategy's performance, we undertake similar analyses as Eisfeldt, Kim, and Papanikolaou (2022) by comparing firms' characteristics between firms in the long leg and firms in the short leg sorted on BM and BM^{Ewens}. While Eisfeldt, Kim, and Papanikolaou (2022) argue that the intangible value factor sorts more effectively on productivity, profitability, financial soundness, and other valuation ratios such as price-to-earnings or price-to-sales, we employ a wider range of variables and provide more statistical testing of the difference in their sorting efficiency. Various sorting patterns emerge from our results, but the primary insight is that the intangible adjusted BM favors firms that are less profitable, have lower levels of operating capital, and have higher leverage.

To investigate whether the improvement effects are caused by the positive relation between intangible capital and future stock returns, as documented in Eisfeldt and Papanikolaou (2013), we utilize the ratio of the intangible capital alone over the market value (IM) to examine the performance of the long-short returns and run Fama-MacBeth regression. A contemporary paper by Gulen, Li, Peters, and Zekhnini (2020) also stresses the differences between intangible and tangible investments. Gulen, Li, Peters, and Zekhnini (2020) highlights the importance of separating intangible and tangible investments and incorporating intangibles into various factors such as investment and profitability, in addition to the value factor, to demonstrate the significance of intangible capital in factor models. While my paper shares the same perspective of using intangible capital alone, our focus is solely on the value factor, and we provide results using alternative intangible estimation methods. Jagannathan, Korajczyk, and Wang (2023) use intangible assets to adjust the measure of firms' return on equity when constructing the profitability factor.

The drawdown refers to how much the investment is down from the peak (usually in percentages) before it recovers back to the peak, and it is an important measure when evaluating trading performance. We construct HML factors based on different BM ratios to investigate the cumulative returns, drawdowns, and spanning tests on them. Consistent with the results from long-short decile portfolio returns, we show that the intangible adjusted HML factors provide higher cumulative returns than the conventional HML factor in the Fama-French five-factor model. However, even these intangible adjusted HML factors have also experienced large and persistent drawdowns in the recent decade, especially in the traditional industry. On the other hand, the IHML, which is constructed using IM, generates much higher cumulative returns and smaller drawdowns. The spanning tests of HML factors vary across time periods and industries. It is difficult to find a superior version of BM, as it changes across criteria and contexts.

What is the reason for the persistent drawdowns experienced by the value strategy, even after adjusting for intangible assets? Arnott, Harvey, Kalesnik, and Linnainmaa (2021) decompose the excess returns and attribute the underperformance of value stocks to the decreased relative valuation of value stocks over growth stocks, in addition to the failure of book value to capture the intangible assets. But they also use Bootstrap to argue that reports of value's death may be greatly exaggerated. Ang (2022) documents the worst value drawdown ever experienced from 2017 to 2022, attributing it to both a decreasing trend component and downturns in cyclical components. In this study, we observe the close relationship between HML returns and the relative valuation using all versions of BM, confirming the finding in Arnott, Harvey, Kalesnik, and Linnainmaa (2021).

We contribute to the studies relevant to value strategy, especially those looking into the effects of modifying the BM ratio with intangible assets. This is the first paper to compare the three most commonly employed estimations of intangible assets. Dividing the whole sample into the high tech sector and not the high tech sector, we examine the potential reasoning behind the intangible adjustments. We also provide a more detailed examination of the argument in Arnott, Harvey, Kalesnik, and Linnainmaa (2021). The overall observation is that modifying book value by adding estimated intangible assets improves the value strategy's performance. However, there is no clear conclusion on which intangible estimation method is best. Furthermore, even the intangible adjusted HML factors have experienced persistent and large drawdowns in recent decades. The cross-sectional sub-sample analyses tend to doubt the narrative of book value mismeasurement. The ratio of intangible capital over market value alone gives rise to HML factors that feature much higher cumulative returns and smaller drawdowns than all versions of intangible adjusted HML factors. While Arnott, Harvey, Kalesnik, and Linnainmaa (2021) only inspects the relation between excess returns and relative valuation using the traditional BM^{FF}, we present more evidence supporting their argument using all variants of ratios.

The remaining sections of this dissertation are structured as follows: Section 2 presents a comprehensive review of the relevant literature. Section 3 details the sources of data, definitions of variables, and methodologies applied in this study. Section 4 discusses the difference in replication and reports findings from the later period. Section 5 provides a comparison of four variations of the BM ratio. Section 6 examines two sub-samples: the time series analyses considering the change in the number of listed firms and the cross-section analyses dividing firms into the high tech sector and the traditional sector. Section 7 conducts further analyses by exploring firms' characteristics, separating the intangible capital from the tangible capital, constructing HML factors, and investigating the relative valuation. Finally, Section 8 presents the concluding remarks.

2.2 Literature review

This chapter is closely related to research on the book-to-market (BM) ratio and its implications in finance, particularly concerning the re-examination of the value investing strategy in recent decades.

As its name indicates, the BM ratio is the book value of equity divided by the market value of equity. However, its calculation is not straightforward, as there are multiple ways to define the book value of equity, such as using total assets minus total liabilities or the value of common shares' equity. In empirical finance, the most commonly implemented calculation of the BM ratio is that of Fama and French (1992). Two fields of academia leverage the BM ratio to make striking findings: corporate finance and asset pricing.

In corporate finance, Tobin's Q has been playing an important role in investment theory since Tobin (1969) developed an intuitive and celebrated theory of investment. Marginal Q, which represents the marginal increase in the firm's value from investing one more unit of capital, is hard to measure. In practice, the BM ratio is often used as a proxy for Q. There is a vast body of academic work related to it in an equilibrium framework both in corporate finance and macroeconomics. In recent work, Peters and Taylor (2017) show that Tobin's Q also explains intangible investments better than tangible investments. They suggest a simple, new Tobin's Q proxy that accounts for the intangible capital and show that it is a superior proxy for both physical and intangible investment opportunities. The new proxy adjusts the denominator, the book value of equity, by adding intangible assets. Kogan and Papanikolaou (2014) use the BM ratio as a proxy for growth opportunities.

The theory of asset pricing leverages BM to study the value premium and its role in asset pricing models. As the empirical results in this chapter are mainly related to this field, a more detailed and thorough literature review will be provided below.

The value investing strategy, which involves holding long positions in value stocks characterized by a high BM ratio and short positions in growth stocks characterized by a low BM ratio, first appeared as a challenge to the traditionally assumed and proved market efficiency hypothesis (Fama and MacBeth (1973)). While the efficient market hypothesis claims that excess returns are not available as all market information is reflected in the stock prices, Basu (1977) finds a negative relation between the investment performance of equity securities and their price-to-earnings (P/E) ratio and that the P/E ratio, due to exaggerated investor expectations, may be indicators of future investment performance.

Since then, a growing number of academic papers have indicated a similar pattern: investing in value stocks can earn a higher return than investing in growth stocks. Although the definition of value stocks varies, such as those with a low price-to-earnings ratio (Basu (1977), Jaffe, Keim, and Westerfield (1989)), a high book value of equity to price (Barr Rosenberg and Lanstein (1984)), or a low debt-to-equity ratio (Bhandari (1988)), it all reflects the fact that these firms have a high book value of equity relative to the market value.

The book value of a firm represents the difference between total assets and total liabilities, thus reflecting the complete value of a company's assets that shareholders would receive if the company were to undergo liquidation. While many characteristics of firms, like P/E ratio and leverage, among others, are claimed to have correlations with cross-sectional stock returns, Fama and French (1992) suggest that only two measures—size and BM—are sufficient to capture all variations in stock returns. Based on this work, they proposed the famous three-factor model in 1993 (Fama and French (1993)), which demonstrated a significantly improved performance compared to the traditional CAPM model. One of the factors, the High minus Low factor (HML) constructed based on the BM ratio, is frequently cited and employed as a benchmark to assess portfolio performance.

Nonetheless, the reason why the value strategy works remains controversial. Some scholars (e.g., Fama and French (1993), Chen and Zhang (1998), Lettau and Ludvigson (2001), Zhang (2005), for instance) attribute it to risk, while some (e.g., Lakonishok, Shleifer, and Vishny (1994), Porta, Lakonishok, Shleifer, and Vishny (1997), Ali, Hwang, and Trombley (2003) among others) explain it from the perspective of behavior. Even within these two primary categories, the specific interpretations differ.

Up until recently, academic studies favoring the value strategy have been flourishing. However, they face more challenges than before, as more evidence has emerged indicating the failure of this strategy. HML factor becomes redundant in the Fama-French five-factor model. Over the last two decades, the HML factor has experienced a large and long drawdown, leading to the argument that the value strategy is no longer effective.

The remedies for the value strategy involve further inspection of the underlying reasons for the strategy's performance or modification of the calculation of the BM ratio. Golubov and Konstantinidi (2019) defend the value strategy by decomposing the BM ratio into market-to-value and value-to-book components. They find that the former drives all the value strategy returns. They also examine four value premium theories based on this decomposition and question their validity. Jaffe, Jindra, Pedersen, and Voetmann (2020) also confirm the behavioral explanation for the value premium, also using a decomposition approved by Rhodes-Kropf, Robinson, and Viswanathan (2005) as Golubov and Konstantinidi (2019) do. Arnott, Harvey, Kalesnik, and Linnainmaa (2021) use Bootstrap to show that the probability of seeing large drawdowns in HML factor observed in recent decades is about 1 in 20—unusual but not enough to support structural impairment—and thus conclude that reports of value's death may be greatly exaggerated. They also attribute the underperformance to the relative valuation. Ang (2022) attributes the drawdowns of value strategy to both a decreasing trend component and downturns in cyclical components. Fama and French (2021) claim that the high volatility of monthly premia prevents rejection of the hypothesis that expected premia are the same in both halves of the sample from 1963 to 2019. They also state that the failure of the value strategy in recent decades may only be concluded if the coefficients in regression remain constant in different sample periods, which is a strong assumption. Campbell, Giglio, and Polk (2023) interpret the returns of the value strategy through the intertemporal CAPM shocks and argue that the strategy's booms and busts primarily result from returns to value stocks within industries.

On the other hand, Asness and Frazzini (2013) find that aligning price data using fewer lags, i.e., more timely price data rather than price data from last December, will forecast the true BM ratio at fiscal year-end. The value portfolios based on the most timely measures earn statistically significant alphas ranging between 305 and 378 basis points per year against a five-factor model containing the standard measure of value as well as market, size, momentum, and a short-term reversal factor. Hasler (2021) and Soebhag, Van Vliet, and Verwijmeren (2022) explore a wider range of choices when constructing the HML factor.

This chapter closely relates to another modification to the calculation of the BM ratio, which incorporates intangible assets. Internally generated intangible capital is not recorded in the balance sheets under the current accounting principle, but it is estimated to take up to about 50% of firms' capital stock (see Falato, Kadyrzhanova, and Sim (2013)). Lev, Radhakrishnan, and Zhang (2009) find that organization capital is associated with five years of future operating, stock return performance, and executive compensation. Eisfeldt and Papanikolaou (2013) document that firms with more intangible capital as measured by organization capital have average returns that are 4.6% higher due to the firms' higher risk from shareholders' perspective. Peters and Taylor (2017), as mentioned earlier, use intangible assets to adjust Tobin's Q and find inspiring results. Given the increasing share of intangible capital yet its absence in the standard measure of book value in the fiscal report, we expect new features when it is combined with the value strategy. Indeed, Park (2022) suggests that an intangible adjusted BM ratio gives higher excess returns than the old one both in the 1976-2017 period and the 1997-2017 period. Eisfeldt, Kim, and Papanikolaou (2022) show that the intangible adjusted HML factor prices standard test assets with lower pricing errors and outperforms the traditional HML factor. Ewens, Peters, and Wang (2020) use market prices to estimate parameters needed in calculating intangible capital and find their version of the intangible adjusted HML factor performs better than other measurements of intangibles. Gulen, Li, Peters, and Zekhnini (2020) demonstrate that adding the off-balance-sheet intangibles to form factors enhances the explanatory power of the Fama-French three- and five-factor models and the q-factor model substantially. The adjusted value factor is no longer redundant in the Fama-French five-factor model. They especially highlight the need to separately consider on- and off-balance sheet assets as well as tangible and intangible investments. Jagannathan, Korajczyk, and Wang (2023) use intangible assets to adjust the measure of firms' return on equity while constructing the profitability factor. However, Rizova and Saito (2021) disfavors the adjustment as they argue that the estimation contains much noise and the estimated intangible assets provide little additional information about future cash flow and profitability, thus being unable to identify differences in expected stock returns.

Using three different versions of internally generated intangible capital, we adjusted the BM ratio and confirmed that including intangible capital improves the performance of the value strategy. However, there is no clear consensus on which measure of intangibles is superior. Additionally, we provide new evidence of sorting efficiency based on both traditional BM and intangible adjusted BM using a broad range of firms' characteristics. We also show a strong correlation between HML returns and relative valuation across all BM ratios and intangible capital over market value (IM) ratios.

2.3 Data and methodology

2.3.1 Data source

We use data from the Center for Research in Security Prices (CRSP) for stocks' prices, returns, delisting returns, delisting reasons, and outstanding shares; data from Compustat for firms' book value of equity, goodwill, and other financial statistics; from French's data library for five factors, five industry categories, and the risk-free rate; and from the Bureau of Labor Statistics website for the consumer price index. The depreciation parameters depreciation and accumulation to estimate intangible assets come from Li (2012) and Ewens, Peters, and Wang (2020). Information on firms' founding years is from Jay Ritter's website.

2.3.2 Variables calculation

How we construct the BM ratio is explained in detail in Bali, Engle, and Murray (2016). Book value is the total parent stockholders' equity (SEQ) adjusted by tax effects and the book value of preferred stocks. To be explicit, we take the sum of the parent stockholders' equity (SEQ), deferred taxes (TXDB), and investment tax credit (ITCB) and then subtract the book value of preferred stocks. All the variables mentioned are recorded in CPSP from 1961. Market capitalization is the absolute value of the product of alternate price (ALTPRC) and the number of shares outstanding (SHROUT) divided by 1000 to make it in $ millions. The size of the stocks is just the logarithm of market capitalization. These variables come from the monthly stock file in CRSP. The market value of equity is calculated in the same way, except that we only use data on the last trading day in December of a given year. The tricky part is aligning annual records of book equity with the market value of equity. The BM for stock i from June of year y to May of year y+1 is the book value ending in year y-1 over the market value in December of year t-1. We do this to ensure all the statistics we use when forming our portfolios are available, as firms usually fail to report the required data in time.

In the accounting standards, the conservatism principle prevents the internally generated intangible capital from being recorded in the balance sheets. For example, the R&D expenses will be viewed as expenses rather than capital. Still, some expenses will turn into technological accumulation and work as part of the firm's total capital. Many scholars have argued that the missing records of intangible assets lead to the bias of the book value, and a popular approach, the perpetual inventory method, is applied to estimate the internally generated intangible assets (see Li (2012), Eisfeldt and Papanikolaou (2013), Li and Hall (2016), Peters and Taylor (2017), Park (2022), Eisfeldt, Kim, and Papanikolaou (2022), Ewens, Peters, and Wang (2020) for example). However, it differs from one to another when the parameters' choice and the selection of initial values are considered. In this chapter, three ways of estimating internally generated intangible assets are used to compare with BM from Bali, Engle, and Murray (2016) which is formalized by Fama (hereafter, BM and BM^{FF} refer to the same ratio). They come from the method to estimate main or baseline internally generated intangible assets from Peters and Taylor (2017), Ewens, Peters, and Wang (2020), and Eisfeldt, Kim, and Papanikolaou (2022), respectively. In their paper, the authors explore various alternative approaches for handling situations where goodwill should be included, research and development expenses should be separately adjusted, and other commonly used parameters should be considered. They argue that their main findings remain robust despite these variations. However, due to the paper's length, only the methods used to construct their primary measure of internally generated intangible assets are discussed here. Though the calculation details are not recorded in this chapter, we offer an overview of the differences in calculating internally generated intangible assets. Peters and Taylor (2017) and Ewens, Peters, and Wang (2020) use the sampling method to account for the initial value of intangible assets and accumulate organization (O_{it}) and knowledge capital (K_{it}) separately while using a different set of parameters. Eisfeldt, Kim, and Papanikolaou (2022) on the other hand, initialize internally generated intangible assets as $SG\&A/0.3$ where $SG\&A$ is the first observation for selling and general administrative expenses when the firm first appears in Compustat. They do not consider research and development expenses separately. we call the intangible adjusted book value of equity as $be^{Eisfeldt}$, be^{Peters} and be^{Ewens} respectively.

Beta is from the most well-known capital asset pricing model (CAPM): $E[R_{i,t}] = R_{f,t} + \beta_i(E[R_{m,t}] - R_{f,t})$. In practice, we just run the following regression on the last trading day of every month to get the monthly estimated beta for each stock:

$$r_{i,t} = \alpha_i + \beta_i MKT_t + \varepsilon_{i,t}$$

where $r_{i,t}$ is the excess return for stock i during period t, MKT_t is the Fama-French market excess return (market factor). Notice that in this chapter, previous one-year daily data, with a minimum of 200 days of non-missing values, are required to do the estimation.

The following equations summarize all the variables mentioned except beta.

$$BE = BE^{FF} = SEQ + TXDB + ITCB - BVPS$$

$$Int_{it}^{Eisfeldt} = 0.8 \times Int_{i,t-1}^{Eisfeldt} + SG\&A$$

$$Int_{it}^{Peters} = [(1 - \delta_{Li,2012})G_{i,t-1} + R\&D] + [0.8 \times O_{i,t-1} + 0.3 \times SG\&A_{xrddeducted}]$$

$$Int_{it}^{Ewens} = [(1 - \delta_{Ewens})G_{i,t-1} + R\&D] + [0.8 \times O_{i,t-1} + \gamma_{Ewens} \times SG\&A_{xrddeducted}]$$

$$BE^{Eisfeldt} = BE^{FF} + Int^{Eisfeldt} - GDWL$$

$$BE^{Peters} = BE + Int^{Peters}$$

$$BE^{Ewens} = BE + Int^{Ewens}$$

$$ME = \frac{|ALTPRC_{Dec.} \times SHROUT_{Dec.}|}{1000}$$

$$BM = BM^{FF} = \frac{BE}{ME}; BM^{Eisfeldt} = \frac{BE^{Eisfeldt}}{ME}; BM^{Peters} = \frac{BE^{Peters}}{ME}; BM^{Ewens} = \frac{BE^{Ewens}}{ME}$$

$$IM^{Eisfeldt} = \frac{Int^{Eisfeldt}}{ME}; IM^{Peters} = \frac{Int^{Peters}}{ME}; IM^{Ewens} = \frac{Int^{Ewens}}{ME}$$

$$size = log(\frac{|ALTPRC \times SHROUT|}{1000})$$

$$r_{i,t} = \alpha_i + \beta_i MKT_t + \varepsilon_{i,t}$$

2.3.3 Methodology

Portfolios analysis

Portfolio analysis is a standard technique usually leveraged in empirical asset pricing to check future return predictability. It has the advantages of nonparametric regression and diversification away from the idiosyncratic risk. It does not require any assumptions in parameter distribution, regression form, etc. Taking averages of stock returns within each portfolio enables us to ignore the idiosyncratic risk which is the risk associated with a specific stock. However, it also has the downside of rapidly running out of degrees of freedom. To explain what it means, assume there are 10000 stocks. If we use univariate analysis and divide stocks into ten portfolios, there will be 1000 stocks in each portfolio. However, suppose we now care about two characteristics and do the bivariate portfolio analysis, dividing stocks into ten groups with each characteristic. In that case, it results in $10^2 = 100$ portfolios with only 100 stocks in each. The number of stocks in one portfolio decreases dramatically from 1000 to 100, which gives rise to the problem of freedom, and the conclusion obtained using much fewer stocks is unreliable. Things are even worse when we are interested in three or more characteristics.

How do we do the analysis? In univariate analysis, for each period, stocks are sorted by one of the characteristics from low to high, say, in this chapter, the BM ratio, and then stocks are divided into different groups by the sorting order to form portfolios. In independent bivariate analysis, stocks are ranked independently by two characteristics and then divided into groups.

To assign stocks into groups, we first need to decide how many groups, or portfolios, we would like to generate. Usually, it is 10 for univariate analysis and 5×5 groups in bivariate analysis. Next, we need to calculate the breakpoints, i.e., the percentiles of the characteristics, and distribute the stocks. In independent bivariate analysis, for example, we calculate the 20th, 40th, 60th, and 80th percentiles of the two characteristics denoted by X_1 and X_2 using the full sample or the sub-sample consisting of only NYSE stocks. A stock i is divided into the first group of X_1 and the second group of X_2, group $P_{1,2}$ if its X_1 is smaller or equal to the 20th percentile of X_1 and its X_2 is between the 20th and 40th percentiles (20th and 40th percentiles included) of X_2. Following this procedure until we get all the 5×5 groups.

Within each group, the average value of the one-month-ahead excess return is calculated using either the equal-weighted or the value-weighted method. After obtaining the time-series excess returns for every portfolio, the time-series means are taken, and then we can compare the one-month-ahead excess returns across different portfolios.

Fama-MacBeth Regression

Fama and MacBeth (1973) propose another procedure, Fama-MacBeth regression, to examine the relation between future excess returns and stocks' characteristics. Unlike portfolio analysis, which limits the number of variables we are interested in, Fama-MacBeth allows controlling for more sets of variables. We first run cross-sectional regression at each period t:

$r_{i,t+1} = \alpha_{0,t} + \alpha_{1,t}X1_{i,t} + \alpha_{2,t}X2_{i,t} + \alpha_{3,t}X3_{i,t} + \ldots + \epsilon_{i,t}$

Then, given the time series of estimated coefficients, adjusted R^2, and number of observations, we regress each of them on 1, with standard errors adjusted following Newey and West (1987) to get the final estimated coefficients.

Notice that the Fama-MacBeth coefficients can be explained as long-short portfolio returns. To see it, assume that we have only one independent variable; in each t, the following regression is run: $r_{i,t} = \alpha_{0,t} + \alpha_{1,t}X_{i,t-1} + \epsilon_{i,t}$, then we have

$$\hat{\alpha}_{1,t} = \frac{\sum_i (X_{i,t-1} - \bar{X}_{t-1})(r_{i,t} - \bar{r}_t)}{\sum_i (X_{i,t-1} - \bar{X}_{t-1})^2} = \frac{\sum_i (X_{i,t-1} - \bar{X}_{t-1}) r_{i,t}}{\sum_i \left(X_{i,t-1} - \bar{X}_{t-1}\right) X_{i,t-1}}$$

We can write $\hat{\alpha}_{1,t} = \sum_i w^{\alpha}_{i,t} r_{i,t}$, where

$$w^{\alpha}_{i,t} = \frac{\left(X_{i,t-1} - \bar{X}_{t-1}\right)}{\sum_i \left(X_{i,t-1} - \bar{X}_{t-1}\right) X_{i,t-1}}$$

It is obvious that $\sum_i w^{\alpha}_{i,t} = 0$, which implies that $\hat{\alpha}_{1,t}$ represents a zero-investment portfolio return by being long in stocks with above-average characteristic $X_{i,t-1}$ and short in stocks with below average characteristic $X_{i,t-1}$.

Spanning test

The spanning test, also known as the mean-variance efficiency test, is a time-series regression series regression proposed by Gibbons, Ross, and Shanken (1989) and Huberman and Kandel (1987) and extended by Kan and Zhou (2012), Gungor and Luger (2015) etc. The regression is as follows:

$$r_t = \alpha + \beta r_{k,t} + \varepsilon_t$$

We are interested in the relation between the mean-variance frontier spanned by the original K benchmark factors and the frontier spanned by K assets plus the tested asset on the left-hand side of the regression. Suppose the regression produces a significantly positive intercept. In that case, adding the tested asset to the original assets could expand the mean-variance frontier and thus improve the highest Sharpe ratio the investor could achieve. See more details in Gibbons, Ross, and Shanken (1989), Kan and Zhou (2012) and Gungor and Luger (2016).

2.4 Replication and extension using a later sample period

The appendix 5.1 presents tables we replicated from Chapter 10: the value premium in Bali, Engle, and Murray (2016) for the sample period from 1963 to 2012. These numbers are quite similar to those in the book. We document first the differences and reasons for them, then extend the analyses using a later sample period from 1980 to 2019.

First, in the summary table A.1, the maximum of BM is 18.64, which exhibits a significant discrepancy in the magnitude with 32.92 in the book. The reason is that we take a different approach to dealing with multiple stocks issued by the same firm. While most firms only issue one class share, some firms may choose to issue more classes even if we have confined the stocks to common stocks. Besides, firms are also allowed to be listed on more than one exchange. Bali, Engle, and Murray (2016) calculate BM at the stock level while what we do, following another most often used method, is to sum all the market capitalization of stocks belonging to the same firm, make it the market capitalization of the stock with the largest market value, and calculate BM at the firm level, leading to smaller values for this ratio. This approach, from my point of view, is more reasonable as the book value of equity is at the firm level, and certain stocks for the same firm present quite a small market capitalization, which causes some BM to have extreme values (while the 95 percentile of BM is around 2.3, BM at the stock level presents extreme BM more than 100). Once the stock level is applied, table A.2 delivers almost identical numbers as Bali, Engle, and Murray (2016). Another reason for the huge difference in the maximum statistics is that it takes an average of only forty-nine or fifty numbers and is therefore susceptible to extreme values. Why is that? Because BM is filled to a monthly frequency using yearly accounting information, as implied by its calculation.

Nevertheless, the divergence should not affect other results significantly in principle; otherwise, it will at least imply the non-robustness of the value premium. The similarity in numbers or patterns between my replication results and the results in Bali, Engle, and Murray (2016) confirms this argument.

Second, Bali, Engle, and Murray (2016) take an additional approach, using the CUSIP to merge Compustat and CRSP data besides using CRSP/Compustat link table, while we only leverage the link table as CUSIP changes over time and is not credible. This explains the smaller value of the average number of stocks in table A.2 compared with 3409 in the book.

Third, as explained above, treating BM at the firm level has another side effect: only one stock with the largest market capitalization among other stocks belonging to the same firm is kept, giving rise to the vanishment of returns for those stocks deleted.

Fourth, negligible differences also exist for β and market capitalization. So, when we do bivariate portfolio sorting using these additional variables, with each group consisting of only a small portion of stocks in the market (around the whole number of stocks divided by 25), the divergence accumulates. As we can see from the comparison between my replication results and those in Bali, Engle, and Murray (2016), the univariate sorting shows more similarity than the bivariate sorting does.

We regard the arguments above as reasonable, and the replication results are acceptable. Based on that, we then do the same analyses using the later sample period, i.e., from 1980 to 2019, to check if the stock market still exhibits a value premium and how its magnitude has changed. Table 2.1 to table 2.8 display these results. The summary table, table 2.1 shows that the mean BM ratio in the later sample is 0.81, a little bit smaller than 0.90 for the period 1964 to 2012. While most statistics for BM are smaller, its maximal value and kurtosis become larger, indicating the distribution of BM ratio in the later sample period is more subtle to extreme values. If we look at book value and market value separately and compare them with those in table A.1, we could figure that the decline in most statistics of BM ratio is caused by the disproportional variation of these two variables: book value, and market capitalization. The book value's mean, 95th percentile, and maximum increase around one-fifth, while market capitalization's statistics almost doubled except for its skewness and kurtosis. The cross-sectional correlation of the BM ratio with β and size keeps analogous, and the time series persistence of the BM ratio declines only a little, as we can see by comparing table 2.2, table 2.3 with table A.3 and table A.4.

After an overview of relevant statistics about the BM ratio, we care more about the value premium associated with it. Table 2.4 displays the univariate portfolio analysis of the BM ratio and future stock returns. Panel A presents the results using BM decile breakpoints calculated using all stocks in the three exchange markets. In contrast, only a subset of stocks, i.e. those listed in the NYSE stock market, is used to calculate the decile breakpoints in panel B. Again, compared with tables generated by the sample period from 1963 to 2012, the BM ratio's magnitude only declined a little for each portfolio, but market capitalization doubled. β for each portfolio shows minor changes compared with replication results. The negative relation between BM and size and between BM and β obtained in table 2.2 accords with this table. From portfolios 1 to 10, as the BM ratio increases, the market capitalization and β decrease. Within each group, the percentage of stocks listed in the NYSE market decreases regardless of which method is applied to calculate breakpoints. For example, in portfolio 1, the percentage of stock numbers listed on the NYSE for the sample period 1963–2012 is 28.55% and 31.29% separately, while it is now 17.58% and 21.36% for the later sample period. This indicates that the number of stocks issued in another two markets, AMEX and NASDAQ, increased.

The general positive relationship between the BM ratio and future stock returns still holds for equal-weighted portfolios, from an excess return of 0.01 for portfolio 1 to 1.3 for portfolio 10 in panel A and from 0.2 to 1.2 in panel B. Even if we use CAPM to adjust the excess returns, this positive relationship is still true, as we can see from the increase in CAPM α for the equal-weighted portfolio. However, value-weighted portfolios, which are the focus as they are more indicative of the returns an investor can achieve by following the portfolio strategy, tell a different story. Though in

Bali, Engle, and Murray (2016), the value-weighted portfolio delivers a substantial reduction in the average excess return and CAPM alpha for portfolios, especially the long-short portfolio, the positive relation between BM and future stock returns is strong and statistically significant. However, in the later sample period, we can see from the table that excess return and CAPM α fail to exhibit a monotonic increase. The long-short portfolio (10-1 column in the table) is not only smaller but also not significant anymore. This suggests a weaker performance of the value strategy in recent times. Bivariate sorting results basically agree with this indication.

Table 2.8 gives Fama-MacBeth regression results. When only the BM ratio or its log form is used as the independent variable, its associated coefficients are 0.37 and 0.40, respectively, with R^2 near 0 and 0.01 for the sample period from 1980 to 2019. They are smaller compared with 0.48 and 0.44, with R^2 near 0.01 and 0.01 in the replication table for the sample period from 1964 to 2012. When β and size are also included as independent variables, the coefficients associated with the BM ratio or its log form are also significantly positive. The predictive power of BM for future stock returns still exists in general for the universe of all stocks, but most coefficients are smaller compared with the earlier sample. The small R^2 is reasonable, as stock-level returns contain too much noise.

Overall, the portfolio analysis and Fama-MacBeth regression suggest the value premium is weaker in the later sample period.

2.5 Comparison between four versions of BM ratio

What causes the value premium to crumble? One possible explanation is that the current accounting principle prevents us from recording and calculating book value precisely as internally generated intangible assets like human capital and research and development costs which usually bring potential future earnings and thus shall be considered to be part of firms' assets, are actually recorded as costs or expenses which in contrast, decreases firms' assets and hence book value of equity. The accounting principle of conservatism requires firms to prepare their accounting reports with caution and high degrees of verification. Therefore, for research and development expenses that may bring large revenues in the future, firms cannot confirm the gains associated due to their uncertainty until the gains are fully realized. This principle has its own pros and cons which are not the concern of this chapter; its effects on the mismeasurement of the BM ratio become evident in recent decades. While traditional firms mainly rely on plant, property and equipment as their main assets, more and more firms like Amazon, Alibaba, Facebook, Apple, Dell, and other technology or e-commerce firms nowadays take intangible assets as their crucial assets. Therefore, some researchers propose to estimate internally generated intangible assets and adjust the book value accordingly. In this section, we consider three alternative measures of intangible assets, calculate the corresponding BM ratio, and compare them with the original BM ratio defined by Fama and French (1993). The construction of these four versions of the BM ratio has been explained in detail in the calculation of the variables. We focus on the comparison between them and check if one version outperforms another in several aspects, such as if its long-short portfolio premium beats others if it predicts future stock returns better.

Table 2.9 displays the summary statistics from 1976 to 2019. By construction, as we add missing internally generated intangible assets to the book value of equity, $BM^{Eisfeldt}$, BM^{Peters} and BM^{Ewebs} are all large than the originally defined BM BM^{FF}. Meanwhile, they are also more volatile and have larger kurtosis. The average monthly count of stocks is also larger. When BM^{FF} is non-positive and deleted from analyses, the other three versions of the BM ratio could be positive and kept in the sample. The extremely large average monthly maximum of $BM^{Eisfeldt}$, 144.44, is caused by the reason explained. When the stock's book value of equity is negative, its market value falls to the bottom as the market price decreases a lot. While this stock with negative BM^{FF} will not be considered anymore when forming portfolios and will be deleted from the sample when we use BM^{FF}, it gets a chance to exhibit a large BM ratio once its estimated intangible assets value is large enough to generate a large 'book' value of equity and therefore is kept for further analyses. The table shows that the average monthly statistics of $BM^{Eisfeldt}$ are the largest, with most percentiles nearly doubled than BM^{FF}. However, it is the most volatile and vulnerable to more extreme values. Though table 2.9 presents a basic overview of the average statistics of BM ratios, it hides the time series evolution of the ratio. Figure 2.1 to figure 2.4 fill this gap. Consistent with table 2.9, all ratios are positively skewed. Though sometimes volatile, the mean of the cross-sectional BM ratio is smaller in the later sample period. $BM^{Eisfeldt}$ has more stocks with a BM ratio larger than 3.

The correlation between these four versions of the BM ratio is presented in table 2.10. The Spearman rank correlation is shown in the above diagonal entries, while Pearson product-moment correlation is below-diagonal entries. As we see from the table, these four versions of the BM ratio are highly positively correlated, making it hard to distinguish the effect of one and the other when they are included as independent variables simultaneously. Spearman rank correlation is overall larger than Pearson product-moment correlation, but it is not too large to conclude a monotonic but non-linear relation between the variables. Bm^{Ewens} is most closely related to BM^{Peters}, which is not surprising as the only change in these two versions is the variation of parameters we use when estimating intangible assets. The Spearman correlation between BM^{Peters} and BM^{Ewens} is 0.99, indicating the results of portfolio sorting would be nearly identical for these two variables. Table 2.11 demonstrate the time persistence, i.e., the time series correlation for different BM ratios. $Bm^{Eisfeldt}$

is the most persistent one. Its ratio exhibits an average correlation to the ratio one year later ($\tau = 12$) of 0.816, larger than 0.762 for BM^{FF}, 0.769 for BM^{Peters} and 0.761 for BM^{Ewens}. For longer lags of months, the persistence of $BM^{Eisfeldt}$ is also stronger than others. At five years lag, the correlation for $BM^{Eisfeldt}$, 0.490, remains sizable. This phenomenon suggests two directions of reasoning. The fact that all versions of the BM ratio are rather persistent indicates the value premium is due to the BM ratio representing risk factor sensitivities. The logic behind it is that if the value premium is caused by mispricing, the market should correct it, and thus the time-series correlation shall either be small or even negative. Another deduction from the more persistent behavior of $BM^{Eisfeldt}$ than others is that we could potentially benefit from using $BM^{Eisfeldt}$ as the definition of BM ratio if mispricing tells the story of value premium as the correction for $BM^{Eisfeldt}$ is slower than others.

Table 2.12 shows the univariate sorting portfolio analyses. Consistent with the previous discussion, the value-weighted excess return for BM^{Peters} and BM^{Ewens} is close to each other for every portfolio. Using the sample period from 1976 to 2019, only the long-short portfolio constructed using $BM^{Eisfeldt}$ is significantly positive. The univariate portfolio analysis for BM^{FF} corresponds to our extension results using the sample period from 1980 to 2019. The adjustment for the book value of equity using estimation of intangible assets proposed by Peters and Taylor (2017) and Ewens, Peters, and Wang (2020) seems not to perform better, at least for the whole sample period we look into.

Table 2.13 displays the results of the Fama-MacBeth regression. The coefficient for $\log BM^{Ewens}$ is 0.39 with a t statistic of 6.12, the largest and the most significant among 0.27 for $\log BM^{FF}$, 0.26 for $\log bm^{Eisfeldt}$, 0.38 for $\log BM^{Peters}$. When all four BM ratios are included to be independent variables, it is again $\log BM^{Ewens}$ that delivers the most significant coefficients. We do not care too much about R^2 here because stock-level regression contains too much noise and is susceptible to extreme values. As Park (2022) argues, the intangible adjusted BM ratio outperforms the BM ratio since the coefficients associated with the intangible adjusted BM ratio are larger and more significant. Table 2.13 shows us BM^{Peters} and BM^{Ewens} outperform BM^{FF} (0.38, 0.39 v.s. 0.27) while the coefficient for $\log BM^{Eisfeldt}$ is 0.26, similar to 0.27 for BM^{FF}. It is confounding with the results in table 2.12 as Fama-MacBeth regression favors BM^{Ewens} while univariate portfolio analysis favors $BM^{Eisfeldt}$. We organize sub-sample analyses in the later section to delve more into that.

To summarize, we have the following observations: (1) $BM^{Eisfeldt}$ makes the BM ratio the largest and most persistent among others, (2) The intangible adjusted BM delivers higher long-short returns than the traditional BM in the univariate sorting, though only the value premium obtained using $BM^{Eisfeldt}$ is significant from 1976 to 2019, (3) In Fama-MacBeth regression, all BM still predicts future returns, but the coefficients of intangible adjusted ones are larger (except $BM^{Eisfeldt}$) and associated with larger t statistics. In the following analyses, we also compare these four versions of BM in subsamples, and through the lens of HML factors.

2.6 Sub-sample analyses

To delve into the BM ratio more closely, we conduct sub-sample analyses. The first sub-sample analysis explores the time dimension as the sample is divided by period, either before or after 1999. The second sub-sample analysis explores the cross-sectional dimension as the sample is divided by industries, whether it is a high tech firm or not.

2.6.1 Pre- and post-1999

The motivation for dividing the sample period to before or after 1999 is based on the stock market fact that the number of listed firms peaked in 1997 and decreased rapidly afterward and that the publicly listed firms are larger and older on average. Kahle and Stulz (2017) and Davydiuk, Glover, and Szymanski (2020) document in detail the empirical evolution of public firms. While their interest is to portray the features of the stock market and explore the mechanism behind it, we try to check if the value strategy exhibits a structural change.

Figure 2.5 displays the time series of listed firms in the NYSE, NASDAQ, and AMEX exchanges. The first two sharp jumps are due to the inclusion of AMEX stocks in July 1962 and NASDAQ in December 1972. After that, the listed number of firms peaks and downs but with the tendency of increasing, it reaches the peak in November 1997 with 7376 firms. Since then, the market has seen a dramatic drop in the number of listed firms, which has steadied at around 3600 firms in the recent decade. On the other hand, the average market capitalization has increased significantly and displayed more volatility since November 1997. Comparing the average market capitalization with the average value after deleting the largest 20 firms, we can see the gap between the two lines is expanding, indicating the largest 20 firms' market value is still growing.

Table 2.14 shows the summary statistics of the BM ratios and market value of equity from 1976 to 2019 divided by the year 1999. We use 1999 instead of 1998 because the peak appears at the end of 1997, and the one-year lag ensures the public availability of the book value of equity information. The mean of BM^{FF} is 0.95 in the pre-1999 sample period and 0.76 for the later sample period. While its percentiles decrease in general, BM^{FF} now is more skewed, and its kurtosis is larger. The smaller percentiles of the BM ratio are a common phenomenon across all versions.

While BM^{FF} and BM^{Peters} have larger kurtosis after 1999, $BM^{Eisfeldt}$ and BM^{Ewens} post 1999 exhibit smaller kurtosis. The market value of equity, which is the market capitalization in December, is around eight times larger in the post-1999 period than before. Combined with the overall decrease in the BM ratio, it implies that the book value of equity is also increasing, but not as rapidly as the market value. Table 2.15 and table 2.16 display the cross-sectional correlation between versions of the BM ratio and the time-series correlation within each ratio. The correlation decreases by a small magnitude in the later sample period.

The results of univariate portfolio analysis are presented in table 2.17. From panel A, we can see that the positive relation between the BM ratio and future stock returns holds in a general way but is not monotonic. The long-short portfolio delivers significant positive values, confirming the existence of the value premium. The intangible adjusted BM ratio produces a larger long-short return than BM^{FF} does. In panel B, things change dramatically. The positive relationship is much weaker in the later sample period. Though the return of portfolio one (with the lowest BM ratio) is usually the smallest, the highest return shows in portfolios 6, 6, 7, and 7, which are usually discarded when forming the HML factor. The long-short return is not significantly different from 0 anymore. Comparing A with B, we have several remarkable observations. First, the value-weighted excess returns for every portfolio all decrease. Take BM^{FF} for example, the excess return for portfolio 1 is 0.7 pre-1999 v.s. 0.49 post-1999, excess returns for other portfolios are 0.69 v.s. 0.54, 0.81 v.s. 0.53, 0.85 v.s. 0.52, 0.86 v.s. 0.52, 0.78 v.s. 0.55, 0.90 v.s. 0.49, 0.85 v.s. 0.49, 0.91 v.s. 0.55, 0.92 v.s. 0.51, respectively. Second, the excess returns for the portfolio are more significant in the earlier sample period. Third, the value premium delivered by the long-short portfolio is small and insignificant post-1999.

Table 2.18 says that if we use BM^{FF} as the BM ratio, its predictive power for future stock returns fades. Before 1999, the coefficient with $\log BM^{FF}$ is 0.41 with a t value of 5.02. However, after 1999, the coefficient is much smaller and insignificant (0.12 with a t value of 1.23 associated). The reduction of the BM ratio coefficients in both the magnitude and t values associated is also present for intangible adjusted ratios. Yet the intangible adjusted BM ratios still have certain predictive power for future returns at the stock level post-1999.

Overall, what we can take from this sub-sample analysis is that: (1) the BM ratio has decreased in recent decades as the book value is not increasing as rapidly as the market value does, (2) the value-weighted excess returns for each portfolio formed by sorting BM ratio are all smaller post-1999 than those pre-1999, (3) the positive relationship between BM ratio and future stock returns is much weaker after 1999, even when intangible assets are included to adjust the book value of equity.

2.6.2 High tech industry and traditional industry

Suppose the mismeasurement of the book value of equity causes the death of the value strategy. In that case, it is reasonable to deduce that the adjustment of intangible assets will restore the problem for technology-intensive or brain-intensive firms and spend more portions of total assets on research and development. For the firms belonging to the more traditional industries like manufacturing, mines, construction, and alike, even though they may also rely more on intangible assets nowadays, the improvement of value strategy brought by adjusting intangible assets is expected to be much less significant compared with technology or brain intensive industries. Therefore, this sub-sample analysis is a cross-sectional division induced by industry and examines the reasoning of intangible adjustments.

The first question is how to divide firms into high tech and traditional industries. Based on the five industries classified by Fama, Ewens, Peters, and Wang (2020) revise it by moving *sic* >= 8000 & *sic* <= 8099 that are originally classified into health care, medical equipment and drugs into industry5 referring to other–mines, construction, construction materials, etc., and moving some "high tech" TV/radio providers into the consumer sector. We use the five industries as Ewens, Peters, and Wang (2020). Intuitively, among those five industry classes, we expect the high tech industry involving 'computer programming and data processing, computer-integrated systems design and others', and 'health industry involving healthcare, medical equipment and drugs' to be technology and brain intensive. Figure 2.6 and 2.7 support this thought. Compustat item 'ppegt' records the value of plant, property and equipment which are major physical assets for traditional industries. Figure 2.6 plots the average of how much a firm's physical assets (proxied by Compustat item 'ppegt') takes up in total assets (Compustat item 'at'). It is clear from the picture that the 'Manuf' and 'Cnsmr' industries have a larger portion of assets as their physical assets, and this portion is either stable or even increasing in the past years. Industries 'HiTec', 'Hlth', and 'Other', on the other hand, not only have a smaller portion of physical assets but also see this portion decreasing over time. Figure 2.7 displays the average of how much expenses on research and development account for total assets. The portion for industry 'Cnsmr', 'Manuf' and 'Other' is less than 2.5% while it is increasing for 'HiTec' and 'Hlth' sectors. Combining the two plots, it is reasonable to include the 'HiTec' and 'Hlth' sectors in the high tech industry and others in the traditional industry. Figure 2.8 plots the average portion of intangible assets (both externally obtained and internally generated estimated) over the intangible adjusted total assets for four versions. As we can see, the portion of intangible assets for the high tech industry is higher than the traditional industry. The highest ratio in the

FF version for the high tech is the minimum ratio in other versions for the high tech industry. Besides, two more phenomena are notable here. First, in the first subplot with title FF, the portions of (mainly externally obtained) intangible assets over total assets in both sectors are close to each other until around the year 2000, when the high tech sector firms start to have far more intangible assets. Second, the intangible assets are forming a greater part of total assets, except we use the version of intangible assets estimated by Eisfeldt, which does not accumulate knowledge and organization capital separately but directly accumulates Compustat item *sg&a* which in most times also includes expenses on research.

After classifying firms into two categories, we do similar analyses and check how the adjustment of intangibles differs and how the value premium within the two sectors differs from each other. The basic summary statistics are shown in table 2.19. Across panels, the average number of firms belonging to the traditional sector is more than double that of firms in the high tech sector. While the average and maximal market value of equity for both sectors is similar in magnitude, the 5th, 25th, 50th, 75th, and 95th percentiles of ME for high tech firms are all smaller than those in panel B, implying the existence of a small number of giant high tech conglomerates taking up the majority of market value. Whichever version of the BM ratio we use, that ratio is larger in the traditional sector than in the high tech sector. Within each panel, for the high tech sector, the intangible adjusted BM ratios are in general twice or more than twice as large as BM^{FF} while for the traditional sector, the intangible adjusted ratios are in general 1.5 times or more than 1.5 times as large as BM^{FF}. Table 2.20 shows that no significant differences in the correlation between versions of BM ratio are exhibited comparing panel A and panel B. Table 2.21 shows that the traditional sector features a more persistent BM ratio.

Table 2.22 displays the univariate portfolio analysis results. We can observe that some exceptions exist for the positive relationship between ratios and future stock returns. Will it be caused by fewer stocks? Given that the traditional sector comprises more firms than the high tech sector does, yet the positive relation is challenged by more objections and that the long-short return is smaller in the traditional sector, we tend to drop this argument. The excess returns for every portfolio in the high tech sector are larger than those in the traditional sector. The long-short returns are positive though sometimes not significant in the traditional sector. What are the outcomes of intangibles adjustment? For the high tech sector, the returns of the long-short portfolio are similar in magnitude (0.42 v.s. 0.42 v.s 0.44 v.s. 0.41) across versions of the BM ratio but they are more significant (1.9 v.s. 2.44 v.s. 2.33 v.s. 0.41) when intangibles are used to adjust for the book value of equity. For the traditional sector, $BM^{Eisfeldt}$ delivers larger and more significant long-short portfolio returns while BM^{Peters} and BM^{Ewens} do not perform better than BM^{FF}. Table 2.23 shows the Fama-MacBeth regression results after controlling for β and size. As all coefficients associated with the BM ratio are significantly positive, it confirms the existence of value premium at least for the whole sample period from 1976 to 2019. For both sectors, BM^{Peters} and BM^{Ewens} rather $BM^{Eisfeldt}$ have larger and more significant coefficients than BM^{FF} do. When four versions of ratios are included in the regression, it becomes hard to explain the results. Within the high tech sector, logBM^{Peters} and logBM^{Ewens} subsume logBM^{FF} and log$BM^{Eisfeldt}$ but for the two significant coefficients, their sign is opposite, which may be caused by their high similarity (0.96 Pearson-product correlation). On the other hand, the coefficients of all versions of the BM ratio are not significant within the traditional sector.

From the analyses of the sub-sample for high tech, and traditional sectors respectively, we have the following arguments: (1) the BM ratios are higher, more volatile, and more persistent in the traditional sector than in the high tech sector, (2) the high tech sector exhibits higher value-weighted returns for every portfolio formed by sorting BM ratios and the traditional sector can hardly create a significantly positive long-short premium, (3) there is no generally better intangible adjusted version as $BM^{Eisfeldt}$ outperforms others in univariate sorting, but only for the traditional sector, BM^{Ewens} and BM^{Peters} work similarly in Fama-MacBeth regression with coefficients larger and more significant than $BM^{Eisfeldt}$.

2.7 Further analyses

In this section, we aim to investigate the reasons behind the superior performance of the intangible adjusted value strategy, delve into the HML factors constructed and explore why the strategy experiences large and persistent drawdowns. Section 2.7.1 provides the comparison of firms' characteristics in the long leg and short leg sorted on BM^{ff} and BM^{Ewens}. Section 2.7.2 explores the value premium brought by intangible capital by separating intangible capital from tangible capital. Section 2.7.3 looks at the HML factors constructed. Section 2.7.4 investigates the relationship between HML returns and the relative valuation of growth stocks over value stocks.

2.7.1 Firms' characteristics

We compare the average values of firms' characteristics in the long leg and short legs sorted by BM^{ff} and BM^{Ewens}. The data come from Compustat, CRSP, IBES, FRED, Refinitiv, French Kenneth's library, and Robert F. Stambaugh's website. We acknowledge that the construction of many variables closely follows the codes by Chen and Zimmermann (2022) and is subject to changes to adapt to our purpose.

Table 2.24 describes the variables used. A 98% winsorization is applied to all ratio variables each year to mitigate the problem caused by extreme values. Table 2.25 is similar to table 12 in Eisfeldt, Kim, and Papanikolaou (2022). Still, we provide more variables covering not only firms' accounting characteristics but also other indicators representing systematic risk, distress risk, mispricing scores, etc. For each variable in the first column, we first calculate the cross-sectional average each year within long or short legs sorted by BM^{ff} or BM^{Ewens}. Then the table displays the time-series average of each variable in the long leg (the second and seventh columns), short leg (the third and eighth columns), the ratio of the difference between the long and the short leg over the mean in the short leg (the fourth and ninth columns), the difference between the two legs (the fifth and tenth columns), the p-value testing if the difference is significantly different from zero (the sixth and eleventh columns).

We observe several patterns detailed below.

- Both versions of the BM ratio can deliver similar sorting effects on the variable: for the long-term EPS forecast (fgr5yr), the difference between the long and short leg is -12.6223 for BM^{ff}, -11.0667 for BM^{Ewens}, and both are significantly different from zero. For the residual momentum of the past 11 months (RMom11), the difference between the long and the short leg is 0.0342 and 0.0550, respectively. Still, both are not significantly different from 0, meaning that both ratios do not efficiently sort the residual momentum of the past 11 months.

- The ratio of difference over mean in the short leg changes a lot through the sorting direction is the same: for mean estimated earnings to price (sfe), sfe is larger in short legs for both BM ratios. However, when using BM^{Ewens}, the mean sfe in the short leg is much larger (-0.0485 v.s. -0.1727), and the mean sfe in the long leg is much smaller (-0.3922 v.s. -0.2017).

- The sorting direction of variables is the same, but the significance level changes: for the industry-adjusted change in capital investment (ChInvIA), it is smaller in the long leg compared to the short leg for both BM ratios, but the difference between the long and the short leg when using BM^{ff} is not significant, while the difference when using BM^{Ewens} is significant. For earnings predictability (ErnPred), the situation reverses.

- The sorting direction of variables changes: for operating leverage (OPLenerage), the stocks in the long leg of BM^{ff} exhibit smaller leverage than those in the short leg do, but the stocks in the long of BM^{Ewens} have larger leverage than those in the short leg do. The change in sorting direction also applies to roa, roe, roi, op, prof, eps, BookLeverage, ChNNCOA, ChNWC, NOA, and others.

The main insight comes from the fourth pattern. We see that the intangible adjusted BM tends to favor firms that are less profitable, have lower levels of operating capital, and have higher leverage.

2.7.2 The role of intangible assets

In previous sections, we observe that adjusting the book value of equity with estimated intangible assets contributes to a higher value premium and the resurgence of the value strategy across time and cross sections. However, it leads to the question of whether the improvement is totally caused by the positive relation between intangible capital and future stock returns as documented in Eisfeldt and Papanikolaou (2013). To investigate this prospect, we substitute the BM ratio with the intangible capital-to-market value of equity (IM) ratio and analyze the value premium based on IM.

Table 2.26 displays the value-weighted average returns of ten decile portfolios, and the long-short portfolio. The long-short returns formed using $IM^{Eisfeldt}$ is 0.12 with a t value of 1.45, even smaller than the long-short returns formed by BM^{FF}. The IM^{Peters} and IM^{Ewens} deliver long-short returns of 0.15 and 0.16, respectively. They are marginally significant with t values of 1.64 and 1.69. The $IM^{Eisfeldt}$ ratio is similar to the organization-to-book value (O/K) ratio in Eisfeldt and Papanikolaou (2013) except that we use the market value as the denominator and do not rank firms with industries. Eisfeldt and Papanikolaou (2013) shows a strong positive relation between O/K: in the five portfolios sorted on O/K, firms with more organization capital have average returns that are 4.63% higher than firms with less organization. In Gulen, Li, Peters, and Zekhnini (2020), the authors report a long-short return of 0.46% for the ten portfolios sorted on IM from July 1998 to June 2022, but this premium is also not significant.

Table 2.27 displays the Fama-MacBeth regression results using the independent variables shown in the first column. The significantly positive coefficients (0.19 for $IM^{Eisfeldt}$, 0.21 for IM^{Peters} and 0.22 for IM^{Ewens}) are consistent the results in Eisfeldt and Papanikolaou (2013) that there is a positive relation between organization capital and future stock returns. Though these coefficients are smaller than 0.27, it does not imply fewer effects of intangible capital on one-month-ahead excess returns. Combined with the standard deviation of these ratios (0.885, 5.3, 1.756, and 1.327) in summary statistics that are not listed in this chapter, they indicate that a one standard deviation difference in BM or IM is associated with an increase in expected stock returns of 0.236%, 1.01%, 0.35%, and 0.295%, respectively. When all versions of ratios are contained in the specification (column (5)), the inclusion of BM^{FF} does not make all other coefficients redundant.

In summary, there is a positive relationship between *IM* and future stock returns, though long-short decile portfolio returns are not significant or are only marginally significant. If we use the long-short return as the criterion to compare ratios, then IM^{Ewens} is the best for this specific sample period.

2.7.3 HML factors

We focus on long-short portfolio returns and Fama-MacBeth regression results to study the value premium in previous sections. The HML factor is also an important value factor broadly used in asset pricing models. Now we explore the HML factor constructed using various BM ratios to provide new insights. To be specific, we draw the cumulative returns and drawdowns and report the spanning tests of these factors.

Figure 2.9 and Figure 2.10 draw the cumulative returns and drawdowns of HML^{FF} from 1976 to 2019, respectively. Consistent with lower long-short returns in recent decades, we see a huge drop in cumulative returns around 2000 and a persistent decline in the recent decade. The drawdown around 2000 is near -0.5, meaning that an investment in this HML factor loses nearly half of the total wealth. More than that, the HML factor experienced a large and long drawdown over the last two decades.

Figure 2.11 and figure 2.12 draw the cumulative returns and drawdowns of four HML factors constructed using BM^{FF} and three intangible adjusted BM from 1976 to 2019. The HML factor formed by $BM^{Eisfeldt}$ generates higher cumulative returns than others at any time. The cumulative returns of HML^{Peters} and HML^{Ewens} are similar to each other; both perform better than HML^{FF}. Yet even the intangible adjusted HML factors failed to avoid losses in the recent decade, as we see their cumulative returns decline in general since 2014 or so. It is in line with our previous results in the time-series subsample analysis that all value premia are smaller and weaker in the later sample period. Figure 2.12 displays the losses more clearly. The intangible adjusted HML factors usually present smaller drawdowns than traditionally defined HML^{FF}. In more recent decades, $HML^{Eisfeldt}$'s drawdown is usually the smallest one, but all of them experience persistent and even increasing drawdown.

Table 2.28 presents the spanning tests of the traditional HML factor and intangible adjusted HML factors for the whole sample period. The dependent variables in columns 1 to 4 are HML^{FF}, $HML^{Eisfeldt}$, HML^{Peters} and HML^{Ewens}, respectively. Numbers in parentheses are p-values associated. As explained in the methodology part, a significant positive intercept implies that adding the tested assets to the original assets can expand the mean-variance frontier and thus improve the highest Sharpe ratio the investor can obtain. Regretfully, none of the intercepts (-0.1, 0.04, 0.04, and 0.06) are significant. The non-significant intercept for the HML factor is expected, as Fama and French (2015) argue that the HML factor becomes redundant in the five-factor model and that its average return is captured by its exposure to RMW and CMA. However, the intangible adjusted BM ratios also fail to deliver a more serviceable HML factor, which means that it is infeasible to increase the power of four factors (market factor, SMB, RMW, and CMA) to explain the average returns by adding the HML factor. But the redundancy of HML factors does not mean the value premium disappears. The value premium emphasizes the positive relationship between the BM ratio and stock returns. As long as the positive relationship still holds, investors can expect benefits from applying the value strategy.

When it comes to the comparison among versions of BM, $BM^{Eisfeldt}$ brings the highest cumulative return and lower drawdowns most of the time. However, none of the HML factors can expand the mean-variance frontier. Though the intangible adjustments do generate higher cumulative returns, they still experience persistent drawdowns in the recent decade.

Will the spanning tests depend on the time period we use, as the intangible capital plays an important role in the later sample period? Table 2.29 exhibits the results for two time-series subsamples. To our surprise, it is the HML factor constructed by BM^{FF} that can expand the efficient frontier spanned by the other four factors in both sub-sample periods. The intercept is 0.27 (p=0.032) and -0.34 (0.042), respectively. However, among the intangible adjusted HML factors, BM^{Peters} and BM^{Ewens} can deliver value factors that have significant and positive intercepts, but only before 19999. Another observation is that the intercepts are all positive pre-1999, but negative afterward. Comparing pre- and post-1999 panels, we see that RMW, CMA, and market factors exhibit positive relation with the HML factors in the later sample period, but some of them are either negatively related to or have no connection to the HML factors in the earlier sample period.

For the division of high tech sector and traditional sector, figure 2.13, 2.14, 2.15, and 2.16 plot the cumulative returns and drawdowns. Though the intangible adjusted HML factors generate much higher cumulative returns in both sectors, they sometimes generate larger drawdowns. HML^{Ewens} and $HML^{Eisfeldt}$ reach a drawdown of more than -0.5 around 1993 among high tech companies. In the recent decade, all HML factors have experienced persistent drawdowns, especially in the traditional sector. HML^{Peters} outperforms within the high tech sector while $HML^{Eisfeldt}$ outperforms within the traditional sector. While the distinct improvement brought by intangible adjustment started around 1986 in the traditional sector, it started ten years later for HML^{Peters} in the high tech sector and started only around 2001 for the other two. The spanning tests are shown in table 2.30. The intercept in column (3) within the high tech sector is 0.29 with a p-value of 0.058. None of the other intercepts are significantly different from 0. So only the HML constructed using BM^{Peters} spans the mean-variance efficient within the high tech sector. While the high tech sector favors BM^{Peters} in the sense of higher cumulative returns, the traditional sector favors $BM^{Eisfeldt}$.

We now move to the HML factors constructed using IM ratios. The cumulative returns and drawdowns are displayed separately in Figure 2.17, 2.18. First, notice that the cumulative returns of IHML factors are even higher than intangible adjusted HML factors (2.23, 12.12, 14.12, 15.74 v.s. 2.23, 8.77, 5.72, 5.43 in figure2.11). Second, the IHML factors do not exhibit large drawdowns in recent decades as HML^{FF} and intangible adjusted HML factors do. The results of the spanning test are shown in table 2.31. The intercept for the dependent variable $IHML^{Peters}$ and $IHML^{Ewens}$ are significant, which means that they can expand the original mean-variance efficient frontier spanned by the other four factors. Besides, the explanatory power of the other four factors also declines for the IHML factors as the R^2 is smaller.

To summarize, the HML factors constructed using BM^{FF} and intangible adjusted BM exhibit similar features as the long-short hedge returns. Their cumulative returns are larger by including intangible capital, yet they have still experienced persistent drawdowns in recent decades, indicating the deterioration of the value strategy. The comparison among versions of BM is more difficult, as different criteria or contexts favor different BM. The IM, on the other hand, provides HML factors that deliver higher cumulative returns and fewer drawdowns, though remember that their long-short returns are not significant.

2.7.4 Relative valuation

The univariate portfolio analysis in Section 2.6.2 and the performance of HML factors within the high tech sector and the traditional sector in Section 2.7.3 pose a challenge to the narrative attributing the deterioration of the value strategy to book value mismeasurement. Arnott, Harvey, Kalesnik, and Linnainmaa (2021) propose another explanation for the value's underperformance: the valuations of value stocks relative to growth stocks have decreased. They argue that the relative valuation for the value factor is at its lowest level in recent decades, i.e., the value factor is much cheaper in the market, and the revaluation component is the largest contributor to the value stocks' underperformance. In this section, we examine this demonstration.

To construct the relative valuation, stocks are sorted into value (growth) portfolios each month if their IM or BM ratios are higher than the 70 percentile (lower than the 30 percentile). Then, we calculate the sum of intangible capital or (intangible adjusted) book value over the sum of market capitalization for the two portfolios. The relative valuation is just the ratio of the growth portfolio over that of the value portfolio. It is smaller than one by construction and measures the relative market price for unit book value or intangible capital of the value portfolio to the growth portfolio. For example, the ratio of 0.25 implies that the price of unit book value or intangible capital of the value portfolio is only 1/4 of the price of the growth portfolio.

Figure 2.19 plots the cumulative returns (left axis) and relative valuation (right axis) of HML factors specified in each subplot constructed using BM^{FF}, $BM^{Eisfeldt}$, BM^{Peters} and BM^{Ewens} from July 1976 to November 2019, while figure 2.20 portrays those using IM ratios. The same pattern is observed across BM or intangible capital over market value specifications. In the short run, the relative valuation moves along with the cumulative returns. In the long run, they diverge in figure 2.19. It is consistent with the results from Arnott, Harvey, Kalesnik, and Linnainmaa (2021). In addition, we examine the relationship not only for BM^{FF} but also for all other ratio variants (intangible adjusted, using only intangible capital).

Figure 2.19 explains why even HML factors constructed using the intangible adjusted BM ratio still experience large drawdowns in recent decades, why the long-short premium post-1999 is smaller, and why the positive relation between ratios and one-month ahead excess returns is weaker after 1999. The relative valuation is smaller compared with the early period, and the relative valuation is almost at its lowest level in recent decades and continues to fall. It means that growth stocks are priced much more expensively than value stocks compared to the early days. We see the relative valuation fall from around 0.25 to around 0.1 in the first subplot. Those numbers translate to the fact that growth stocks, which were about four times more expensive than value stocks, are now nearly ten times more expensive than value stocks. Adding estimated intangible capital to the book value of equity only mitigates the pricing gap but does not work strongly enough to avoid the large drawdowns in recent decades. The R^2 shown in the picture is the adjusted R^2 from the regression of annual returns in July (in log form) on the difference in log relative valuation between the current year and the previous year. The change in relative valuation explains approximately half of the variance in excess returns.

IHML factors, on the other hand, have exhibited stable relative valuation in recent decades. For IHML constructed using $Int^{Eisfeldt}/M$, the relative ratio is around 0.07 in recent decades, smaller than 0.1 from the earlier sample period. It is consistent with the worse performance and larger drawdowns of $IHML^{Eisfeldt}$ as shown in figure 2.18. The relative valuation for the other IHML factors has been around 0.1 in the last ten years, which is at its highest level. Associated with that, the cumulative returns keep increasing and deliver small drawdowns.

In summary, we see no additional benefits to using intangible capital to adjust book value. Instead, the ratio of IM on its own can give rise to a factor that features much higher cumulative returns and smaller drawdowns in recent decades. The performance of $IHML^{Peters}$ and $IHML^{Ewens}$ is better than $IHML^{Eisfeldt}$. The relative valuation explains the persistent and large drawdowns of HML^{FF} and other intangible adjusted HML factors. Given that growth stocks are priced more expensively than value stocks, if the mean reversion of pricing occurs in the future, the value strategy is expected to gain positive returns in the future.

2.8 Conclusion

We confirm the recent deterioration of the value strategy and the positive impact of incorporating intangible capital into the calculation of the book value of equity. However, even the intangible adjusted long-short decile portfolio returns are smaller in recent decades. There is no superior version of intangible adjusted BM, as it varies across criteria and contexts.

Motivated by the sudden drop of listed firms around 1998, we conduct a time-series subsample analysis to investigate if there is a structural change. Our results show that the average BM ratio after 1999 is smaller than the earlier sample period, indicating that the book value of equity increases more slowly than the market value. While the number of listed firms falls for around 15 years and keeps steady with about 3800 firms from 2013, the average market capitalization amplifies dramatically at the same time. Though the inclusion of estimated intangible assets increases the long-short returns, the positive relationship between the BM ratio and future stock returns is much weaker, and the average value-weighted long-short returns are not significantly different from 0 after 1999.

To investigate if the conservative accounting biases in book value contribute to the death of the value strategy, we classify firms into the high tech sector which is more vulnerable to this bias, and the traditional industry which is less vulnerable to the bias. While the statistics of the BM ratio and the univariate portfolio returns exhibit diversity, the improvement brought by adjusting intangibles in the performance of the value strategy is significant in both sectors, casting doubt on the story of book value mismeasurement.

In further analyses, we show that intangible adjusted BM tends to favor firms that are less profitable, have lower levels of operating capital, and have higher leverage. Separating the intangible and tangible capital, we use the book-to-market ratio where the numerator is only the value of intangible capital to show that there is also a positive relationship between intangible capital and future stock returns. However, the long-short returns are either not larger than the traditional value premium or are marginally significant.

The HML factor is the important value factor constructed using BM, which plays an important role in asset pricing models and is broadly used by the private sector. So we also delve into the cumulative returns, drawdowns, and spanning tests on HML factors constructed using various BM ratios. We find that consistent with lower long-short returns in the later sample, the HML factors constructed have experienced a large and persistent drawdown in recent decades. On the other hand, the IHML factors which are constructed using the IM ratios deliver much larger cumulative returns and smaller drawdowns. The relative valuation shows that the growth stocks are priced more expensive than value stocks, if the mean reversion of pricing occurs in the future, the value strategy is expected to gain positive returns again.

2.9 Figures and tables

2.9.1 Figures

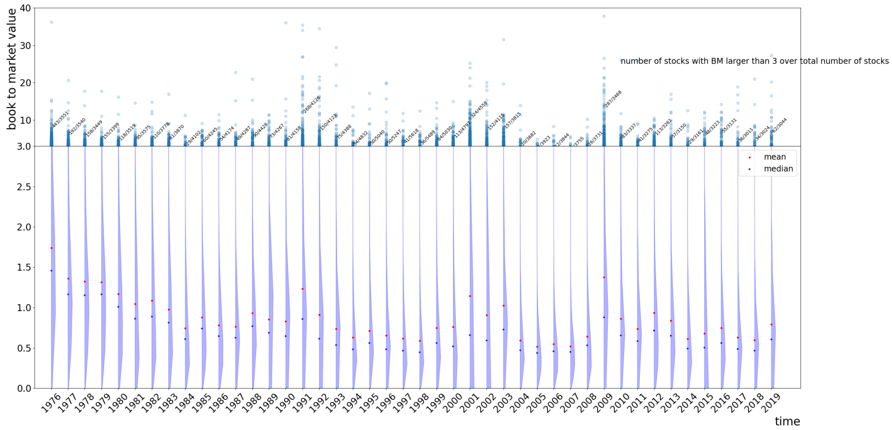

Fig. 2.1: The time series of the cross-sectional distribution of BM^{FF} from 1976 to 2019

This figure plots the annual 'distribution' time series of BM^{FF} from 1976 to 2019. Each year only data in June are used to draw the picture. The lower part is the distribution up to the value of 3, with red points indicating the means and blue points indicating the medians. The upper part draws the scatter with labels indicating the number of stocks with $BM > 3$ over the number of stocks in June.

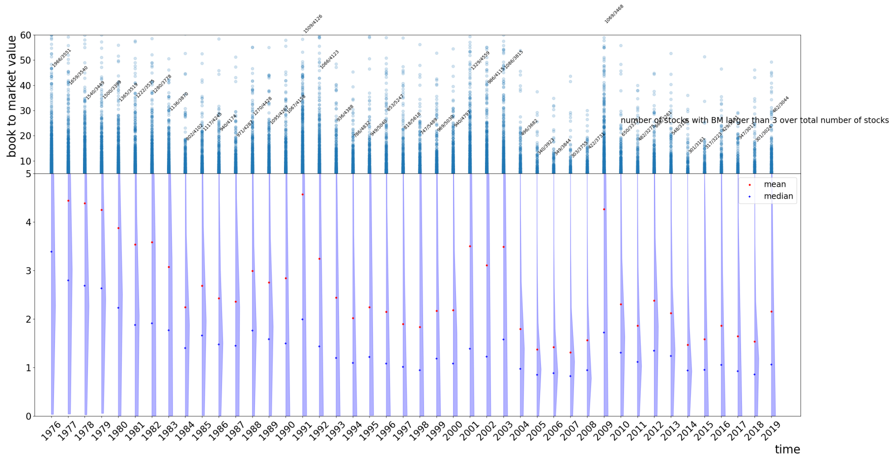

Fig. 2.2: The time series of the cross-sectional distribution of $BM^{Eisfeldt}$ from 1976 to 2019

This figure plots the time series of annual 'distribution' of $BM^{Eisfeldt}$ from 1976 to 2019. Each year only data in June are used to draw the picture. The lower part is the distribution up to the value of 5 with red points indicating the means and blue points indicating the medians. The upper part draws the scatter with labels associated indicating the number of stocks with $BM^{Eisfeldt} > 3$ over the number of stocks in June.

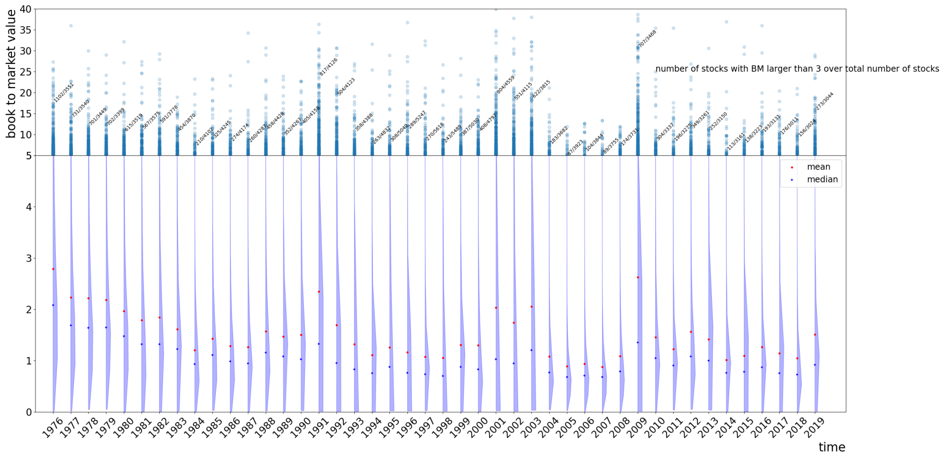

Fig. 2.3: The time series of cross-sectional distribution of BM^{Peters} from 1976 to 2019
This figure plots the time series of annual 'distribution' of BM^{Peters} from 1976 to 2019. Each year only data in June are used to draw the picture. The lower part is the distribution up to the value of 5, with red points indicating the means and blue points indicating the medians. The upper part draws the scatter with labels associated indicating the number of stocks with $BM^{Peters} > 3$ over the number of stocks in June.

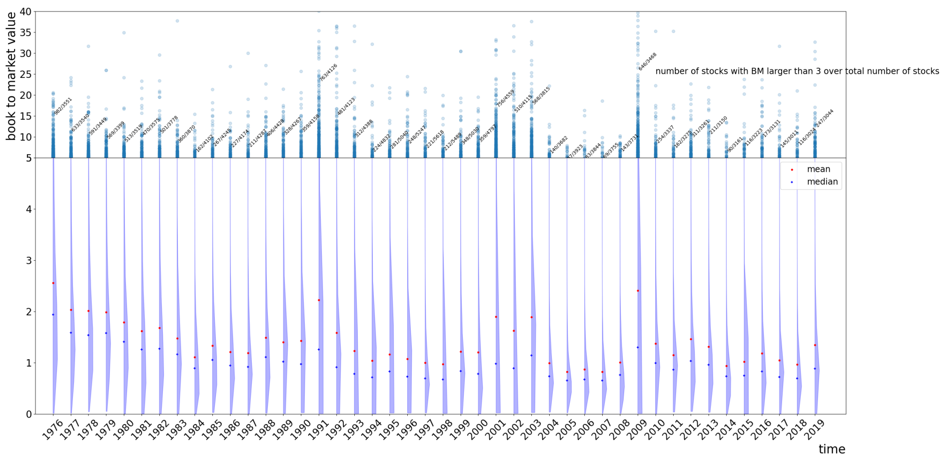

Fig. 2.4: The time series of cross-sectional distribution of BM^{Ewens} from 1976 to 2019
This figure plots the annual 'distribution' time series of BM^{Ewens} from 1976 to 2019. Each year only data in June are used to draw the picture. The lower part is the distribution up to the value of 5, with red points indicating the means and blue points indicating the medians. The upper part draws the scatter with labels associated indicating the number of stocks with $BM^{Ewens} > 3$ over the number of stocks in June.

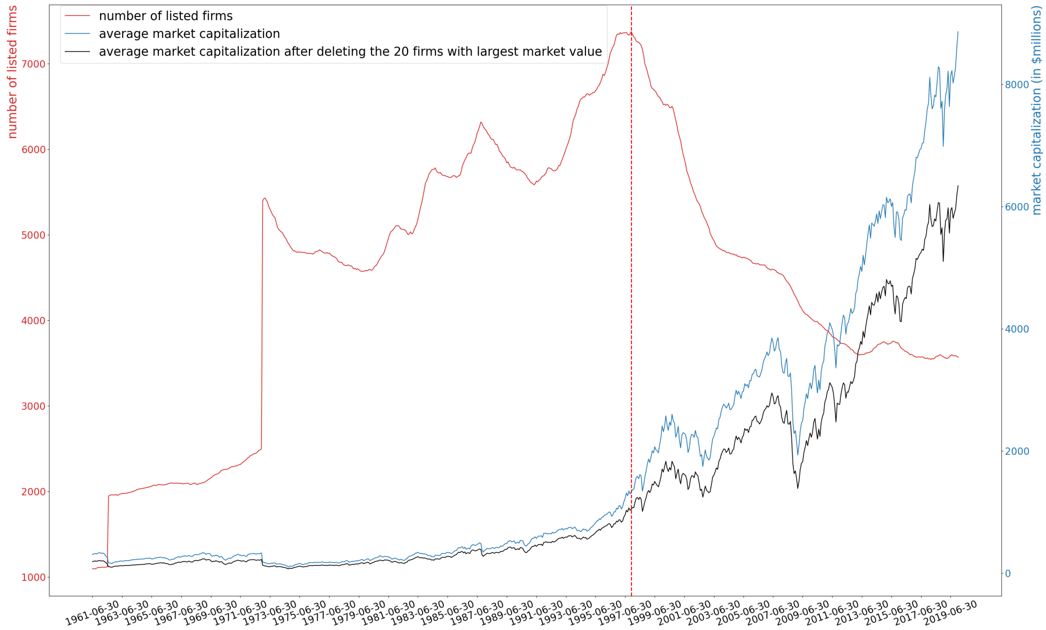

Fig. 2.5: The number of firms listed and average market capitalization from 1961 to 2019
This figure plots the time series of the number of listed firms in NYSE/AMEX/Nasdaq (red line), the average market capitalization of the stocks (blue line), and the average market capitalization of the stocks after deleting the largest 20 firms in each month (black).

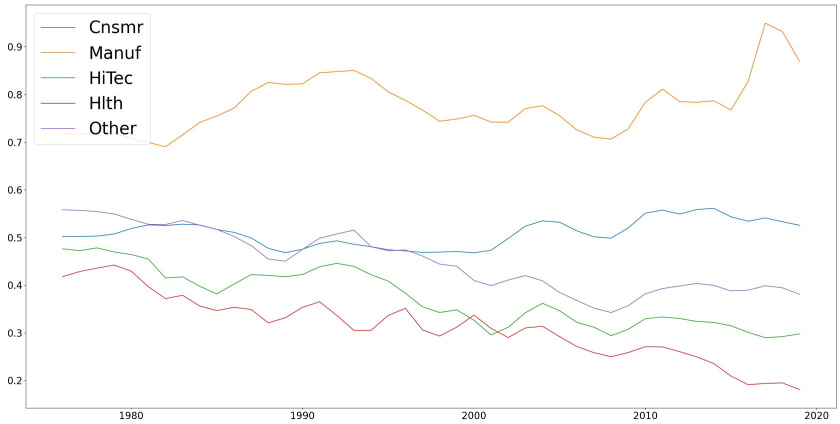

Fig. 2.6: The average of firms' percentages of property, plant, and equipment over total assets
This figure plots the time series of the average firms' percentages of PP&E (property, plant, and equipment) over total assets for five general sectors classified by Fama and French from 1976 to 2019. Each year, we calculate the value of PP&E over total assets for each firm and then take the averages using all the firms belonging to the same sector.

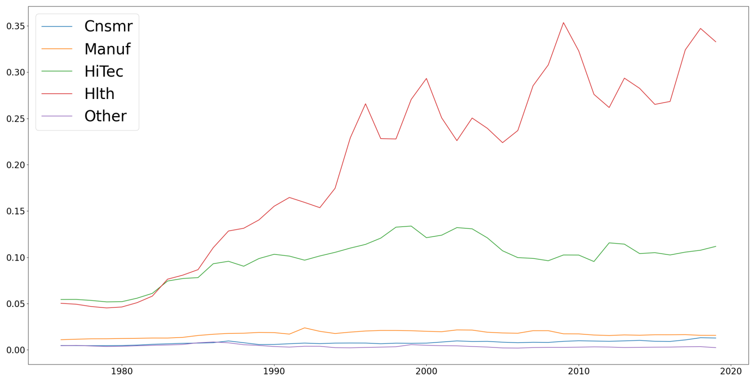

Fig. 2.7: The average of firms' percentages of expenditure on research and development over total assets
This figure plots the time series of the average firms' percentages of expenditure on research and development over total assets for five general sectors classified by Fama and French from 1976 to 2019. Each year, we calculate the value of the expense on research and development over total assets for each firm and then take the averages using all the firms in the same sector.

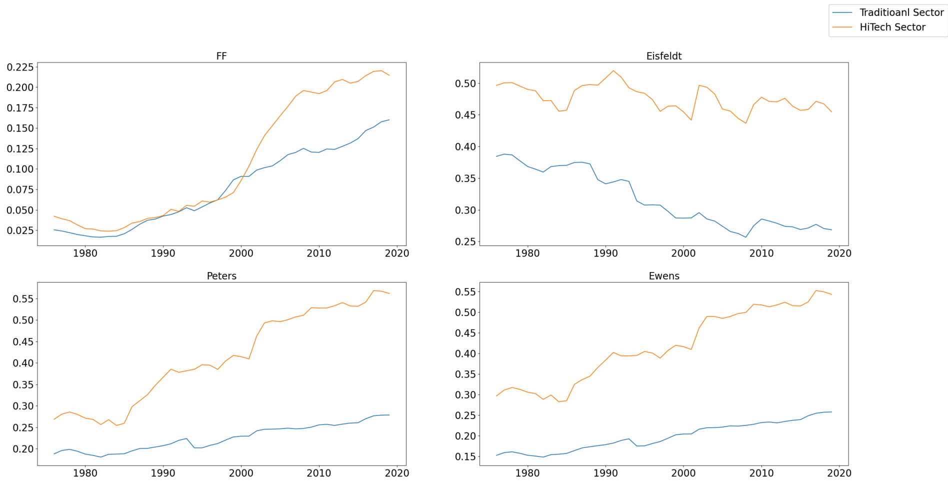

Fig. 2.8: The average of firms' percentages of intangible assets over intangible adjusted total assets

This figure plots the time series for the average of firms' percentages of the intangible assets (both externally obtained assets and estimated internally generated assets) over the intangible adjusted total assets for the traditional sector (blue line) and the high tech sector (orange line) defined in this chapter from 1976 to 2918. The titles of sub-plots indicate the version of estimated intangible assets.

Fig. 2.9: The cumulative returns of HML factor from 1963 to 2019
This figure plots the cumulative returns of the HML factor constructed following Fama and French (1993) from July 1963 to November 2019.

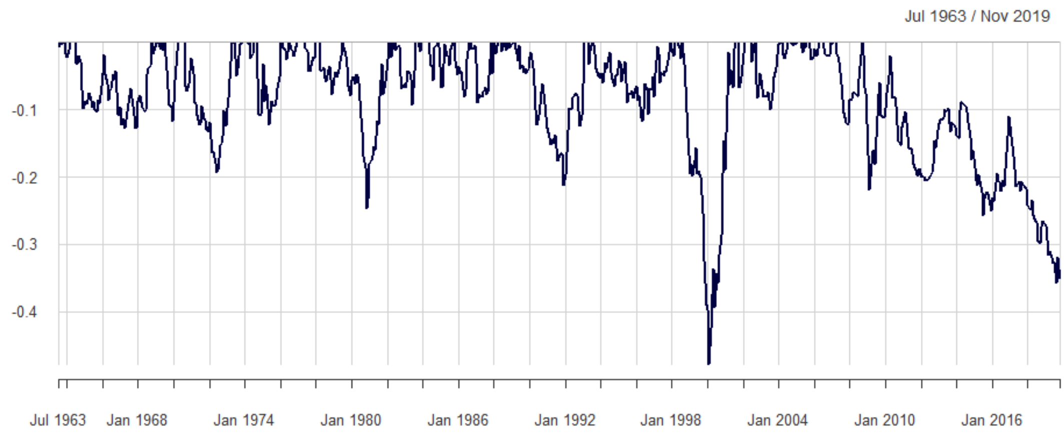

Fig. 2.10: The drawdowns of HML factor from 1963 to 2019
This figure plots the drawdowns of the HML factor constructed following Fama and French (1993) from July 1963 to November 2019.

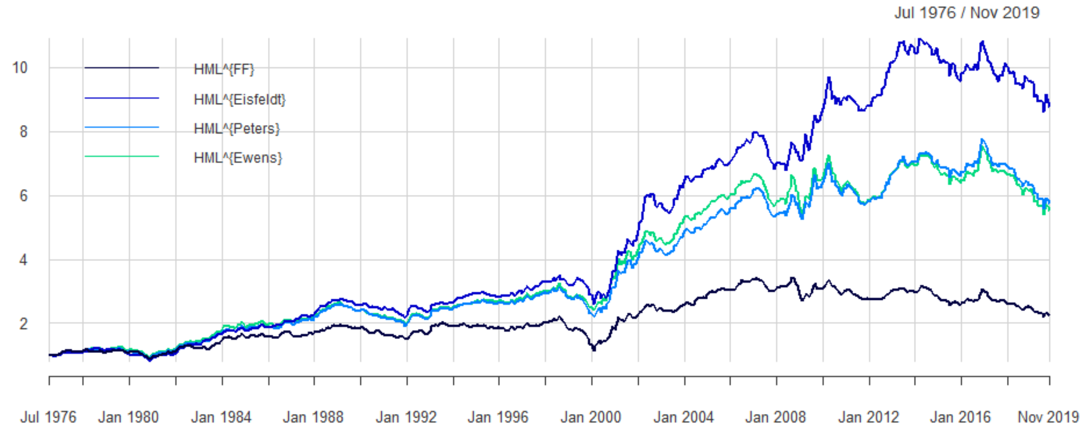

Fig. 2.11: The cumulative returns of four HML factors from 1976 to 2019

This figure draws the cumulative returns of HML factors constructed using BM^{FF} (dark line), $BM^{Eisfeldt}$ (blue line), BM^{Peters} (light blue line) and BM^{Ewens} (green line) from July 1976 to November 2019.

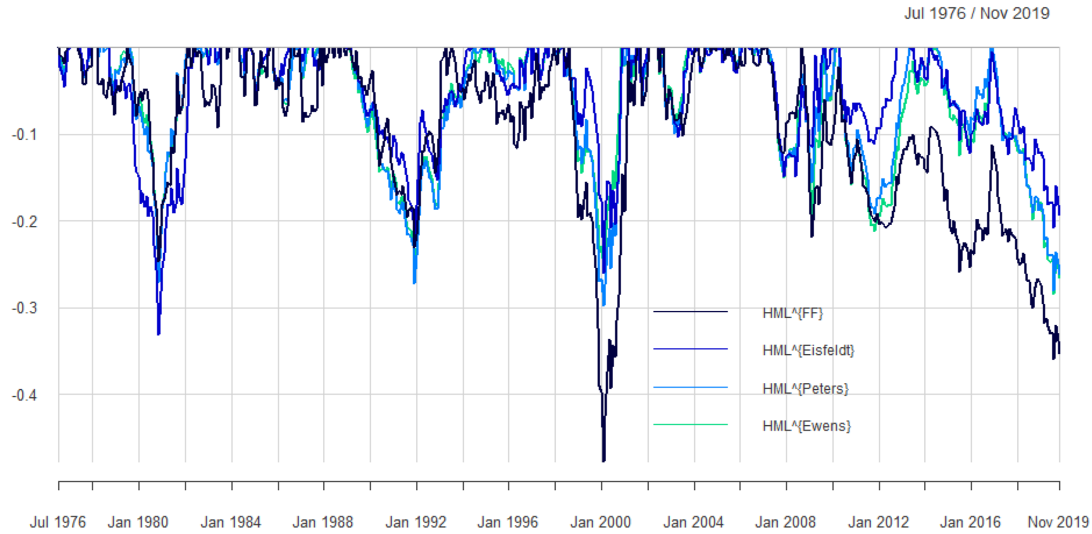

Fig. 2.12: The drawdowns of four HML factors from 1976 to 2019

This figure draws the drawdowns of HML factors constructed using BM^{FF} (dark line), $BM^{Eisfeldt}$ (blue line), BM^{Peters} (light blue line) and BM^{Ewens} (green line) from July 1976 to November 2019.

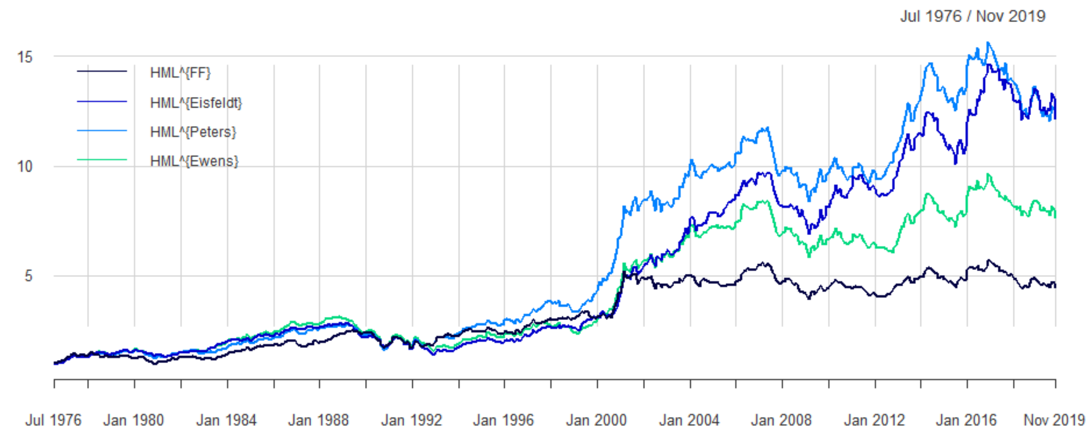

Fig. 2.13: The cumulative returns of four HML factors from 1976 to 2019 within the high tech sector

This figure draws the cumulative returns of HML factors constructed in the high tech sector using BM^{FF} (dark line), $BM^{Eisfeldt}$ (blue line), BM^{Peters} (light blue line) and BM^{Ewens} (green line) from July 1976 to November 2019.

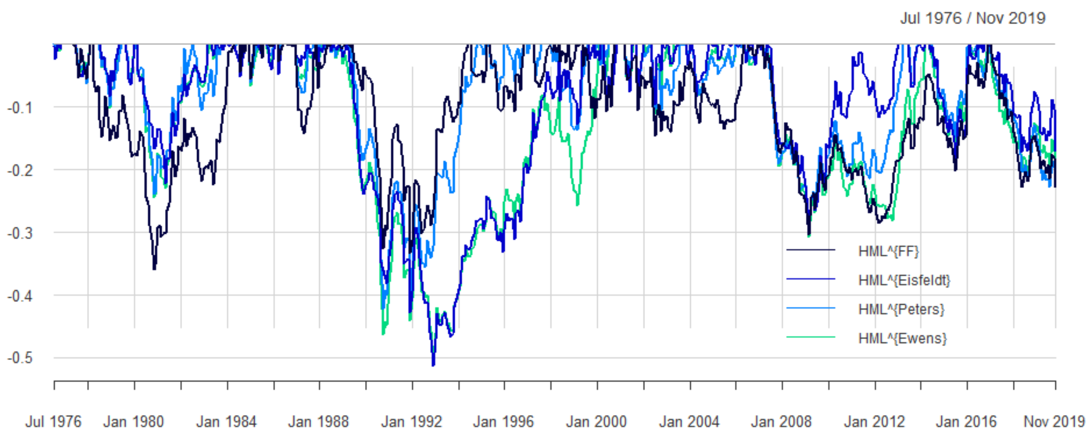

Fig. 2.14: The drawdowns of four HML factors from 1976 to 2019 within the high tech sector

This figure draws the drawdowns of HML factors constructed in the high tech sector using BM^{FF} (dark line), $BM^{Eisfeldt}$ (blue line), BM^{Peters} (light blue line) and BM^{Ewens} (green line) from July 1976 to November 2019.

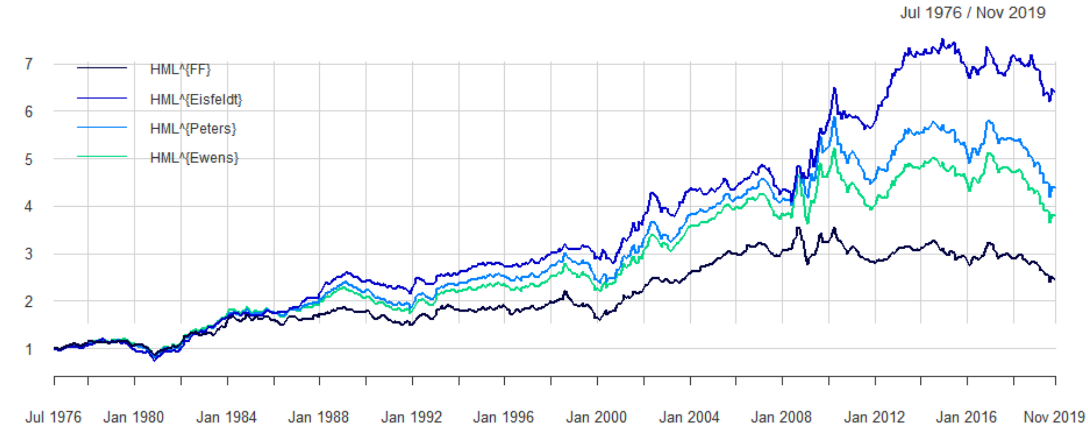

Fig. 2.15: The cumulative returns of four HML factors from 1976 to 2019 within the traditional sector

This figure draws the cumulative returns of HML factors constructed in the traditional sector using BM^{FF} (dark line), $BM^{Eisfeldt}$ (blue line), BM^{Peters} (light blue line) and BM^{Ewens} (green line) from July 1976 to November 2019.

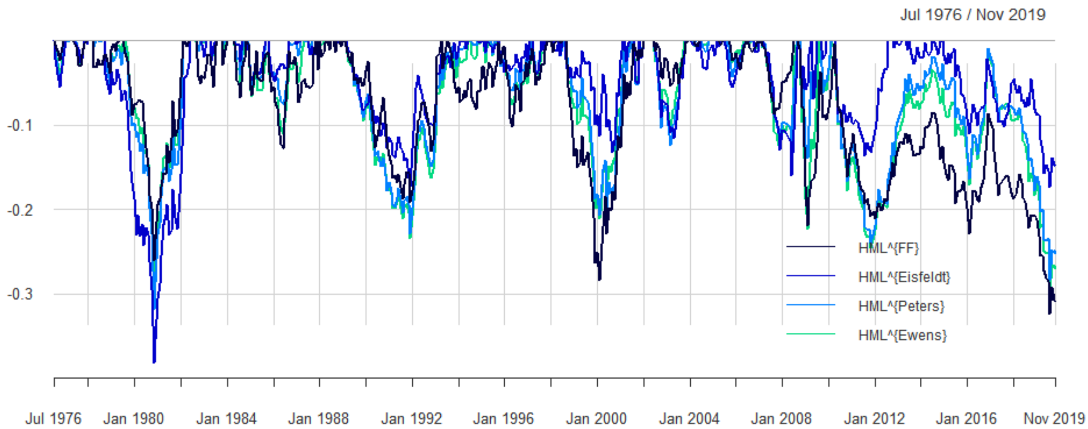

Fig. 2.16: The drawdowns of four HML factors from 1976 to 2019 within the traditional sector

This figure draws the drawdowns of HML factors constructed in the traditional sector using BM^{FF} (dark line), $BM^{Eisfeldt}$ (blue line), BM^{Peters} (light blue line) and BM^{Ewens} (green line) from July 1976 to November 2019.

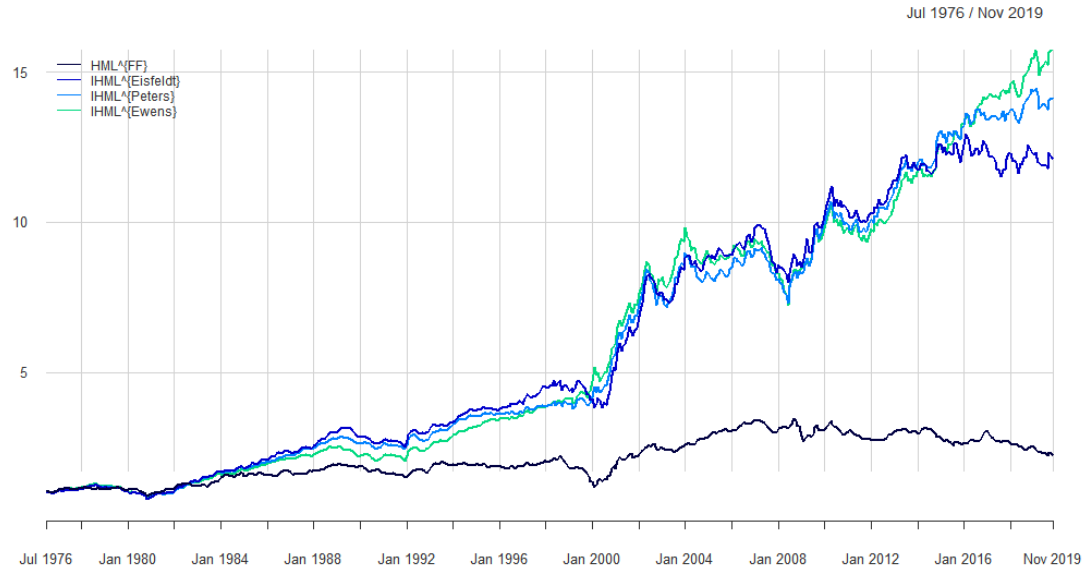

Fig. 2.17: The cumulative returns of HML, IHML factors from 1976 to 2019

This figure draws the cumulative returns of HML and IHML factors constructed using BM^{FF} (dark line), $im^{Eisfeldt}$ (blue line), im^{Peters} (light blue line) and im^{Ewens} (green line) from July 1976 to November 2019.

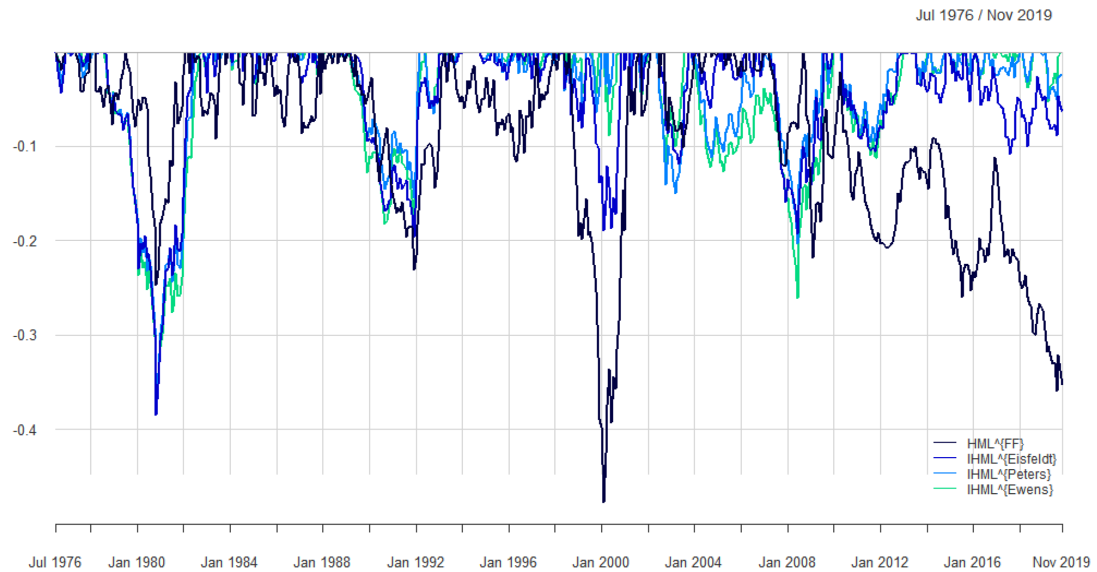

Fig. 2.18: The drawdowns of HML, IHML factors from 1976 to 2019

This figure draws the drawdowns of HML, IHML factors constructed using BM^{FF} (dark line), $im^{Eisfeldt}$ (blue line), im^{Peters} (light blue line) and im^{Ewens} (green line) from July 1976 to November 2019.

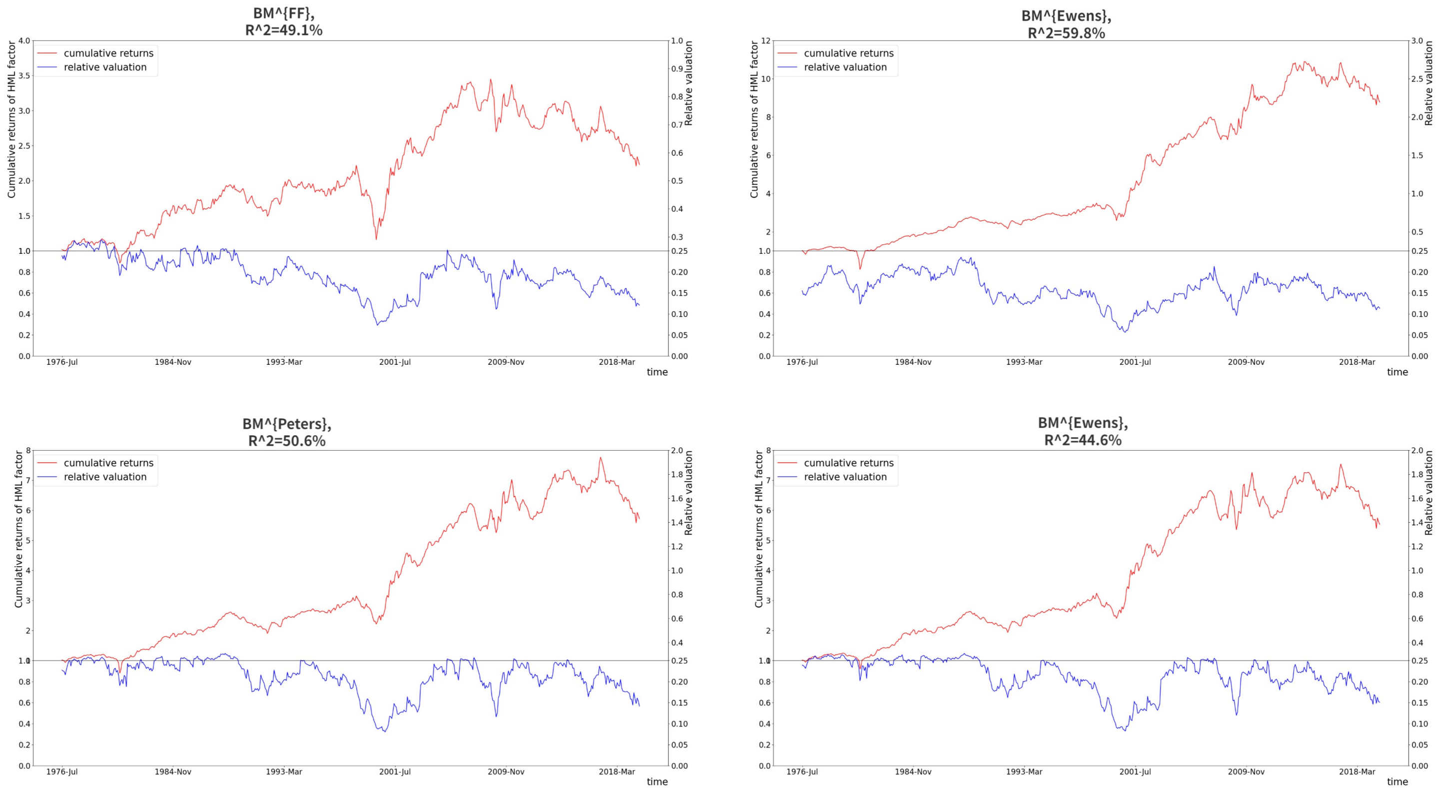

Fig. 2.19: The cumulative returns and relative valuations of four HML factors from 1976 to 2019

This figure draws the cumulative returns (left axis) and relative valuation (right axis) of HML factors specified in each subplot constructed using BM^{FF}, $BM^{Eisfeldt}$, BM^{Peters} and BM^{Ewens} from July 1976 to November 2019. The Rˆ2 is the adjusted R^2 from the regression of annual returns in July (in log form) on the difference of log relative valuation between the current year and the previous year.

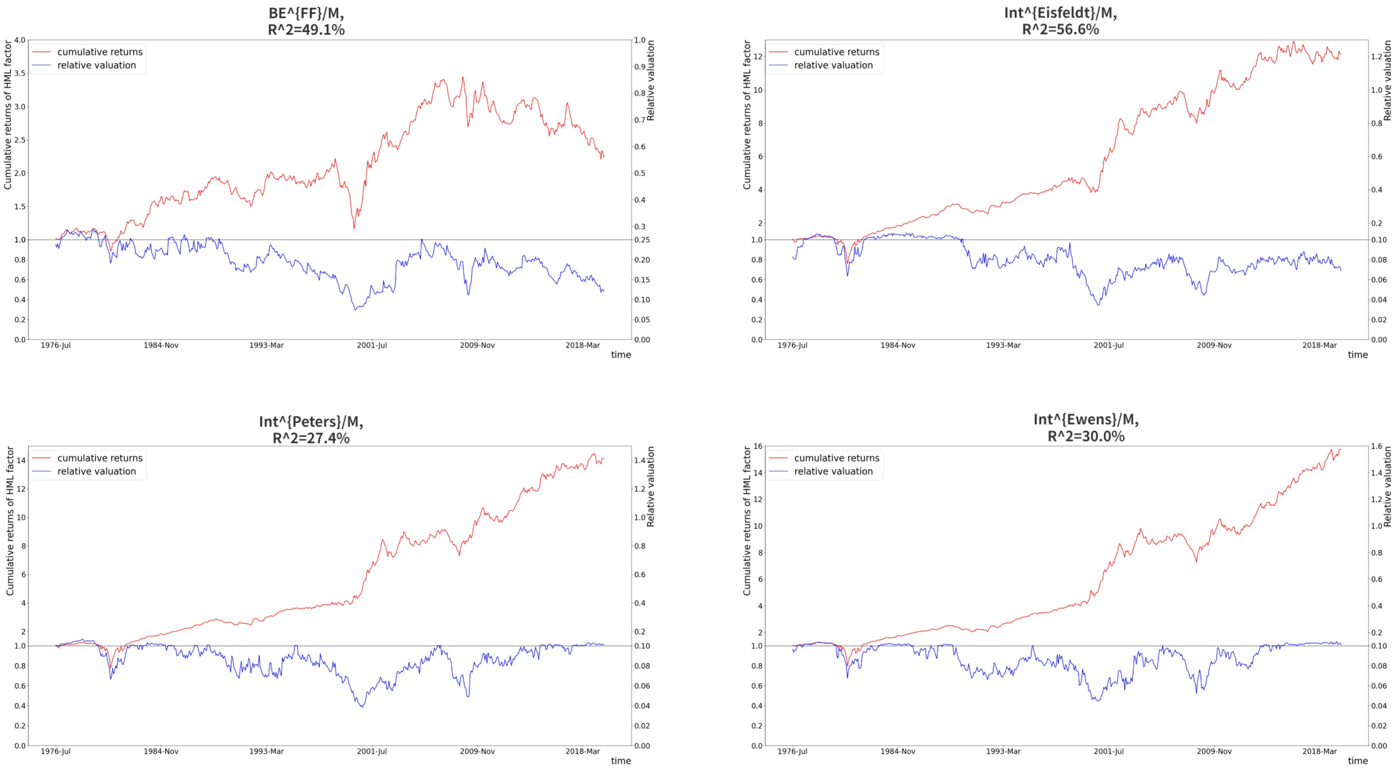

Fig. 2.20: The cumulative returns and relative valuations of HML, IHML factors from 1976 to 2019

This figure draws the cumulative returns (left axis) and relative valuation (right axis) of HML or IHML factors specified in each subplot constructed using BM^{FF}, $im^{Eisfeldt}$, im^{Peters} and im^{Ewens} from July 1976 to November 2019. The Rˆ2 is the adjusted R^2 from the regression of annual returns in July (in log form) on the difference of log relative valuation between the current year and the previous year.

2.9.2 Tables

Table 2.1: Summary Statistics
This table presents the summary statistics for variables relevant to the book-to-market ratio calculated using the CRSP, Compustat. The sample period is from 1980 to 2019. Each month, the mean (mean), standard deviation (std), skewness (skew), excess kurtosis (kurtosis), minimum (min), fifth percentile (5%), 25th percentile (25%), median (median), 75th percentile (75%), 95th percentile (95%), and maximum (max), number of stocks (count) of the cross-sectional distribution of each variable are calculated. The table presents the time-series means for each cross-sectional value. BM for months from June of year y through May of year y + 1 is calculated as the book value of common equity as of the end of the fiscal year ending in calendar year y-1 to the market value of common equity as of the end of December of year y-1. lnBM is the natural log of BM. BEadj and MEadj are the book value and market value adjusted to reflect the 2012 dollar using the consumer price index and recorded in millions of dollars. MktCap is the share price times the number of shares outstanding.

	mean	std	skew	kurt	min	5%	25%	median	75%	95%	max	count
BM	0.81	0.87	7.18	142.18	0.00	0.11	0.33	0.62	1.02	2.03	20.09	3884
lnBM	-0.62	0.93	-0.89	3.00	-7.24	-2.23	-1.12	-0.51	-0.01	0.67	2.83	3884
BE	1254	6039	15	363	0	6	37	136	534	4793	155398	3884
ME	2732	13085	15	313	1	12	66	271	1128	10186	329579	3884
Mktcap	2607	12414	14	304	1	10	59	256	1095	9851	307561	3883

Table 2.2: Correlations
This table presents the time-series averages of the annual cross-sectional Pearson product–moment (below-diagonal entries) and Spearman rank (above-diagonal entries) correlations between pairs BM, ln BM, β, and size. The sample period is from 1980 to 2019.

	BM	lnBM	β	size
BM		1.00	-0.25	-0.24
lnBM	0.83		-0.25	-0.24
β	-0.22	-0.24		0.40
size	-0.27	-0.21	0.38	

Table 2.3: Persistence
This table presents the results of persistence analyses of BM and ln BM. Each month t, the cross-sectional Pearson product–moment correlation between the month t and month $t+\tau$ values of the given variable is calculated. The table presents the time-series averages of the monthly cross-sectional correlations. The column labeled τ indicates the lag at which the persistence is measured. The sample period is from 1980 to 2019.

τ	BM	lnBM
12	0.755	0.793
24	0.607	0.668
36	0.507	0.586
48	0.437	0.526
60	0.387	0.477
120	0.305	0.375

Table 2.4: Univariate Portfolio Analysis
This table presents the results of univariate portfolio analyses of the relation between the book-to-market ratio and future stock returns. The sample period is from 1980 to 2019. Panel A displays the average values within each portfolio formed by sorting BM using all stocks in the sample while panel B displays the same statistics within portfolios formed by sorting BM using NYSE-listed stocks. The Characteristics section of each panel shows the average values of BM, lnBM, MktCap, and β, the percentage of stocks that are listed on the New York Stock Exchange, and the number of stocks for each decile portfolio. The EW portfolios (VW portfolios) section in each panel shows the average equal-weighted (value-weighted) one-month-ahead excess return and CAPM alpha (in percent per month) for each of the 10 decile portfolios as well as for the long-short zero-cost portfolio that is long the 10th decile portfolio and short the first decile portfolio (column 10 – 1). Newey and West (1987) t-statistics, adjusted using six lags, testing the null hypothesis that the average portfolio excess return or CAPM alpha is equal to zero, are shown in parentheses.

Panel A: NYSE/AMEX/NASDAQ Breakpoints												
	Value	1	2	3	4	5	6	7	8	9	10	10-1
Characteristics	BM	0.11	0.23	0.34	0.44	0.56	0.69	0.84	1.02	1.31	2.52	
	lnBM	-2.46	-1.51	-1.13	-0.85	-0.61	-0.40	-0.21	-0.01	0.23	0.78	
	Mktcap	4426	4633	3999	3014	2569	2038	1886	1676	1365	636	
	β	1.01	0.98	0.92	0.88	0.84	0.79	0.74	0.68	0.61	0.56	
	%NYSE	17.58%	26.01%	31.83%	35.11%	36.85%	37.80%	37.13%	36.40%	33.25%	25.22%	
	n	386	386	386	386	386	386	386	386	386	386	
EW portfolios	Excess return	0.01	0.37	0.57	0.74	0.84	0.93	0.92	0.96	1.06	1.30	1.30
		(0.01)	(1.11)	(1.97)	(2.65)	(3.0)	(3.46)	(3.5)	(3.6)	(3.93)	(3.53)	(4.96)
	CAPM α	-0.92	-0.50	-0.24	-0.04	0.09	0.23	0.26	0.33	0.45	0.66	1.58
		(-4.18)	(-2.95)	(-1.58)	(-0.26)	(0.62)	(1.54)	(1.73)	(2.1)	(2.65)	(2.54)	(5.69)
VW portfolios	Excess return	0.57	0.56	0.69	0.67	0.67	0.70	0.73	0.62	0.80	0.92	0.34
		(2.0)	(2.47)	(3.37)	(3.08)	(3.17)	(3.48)	(3.43)	(2.73)	(4.04)	(3.35)	(1.27)
	CAPM α	-0.19	-0.13	0.04	-0.01	0.02	0.09	0.15	0.03	0.23	0.26	0.45
		(-1.29)	(-1.46)	(0.65)	(-0.17)	(0.27)	(0.92)	(1.21)	(0.2)	(1.57)	(1.42)	(1.53)
Panel B: NYSE Breakpoints												
	Value	1	2	3	4	5	6	7	8	9	10	10-1
Characteristics	BM	0.15	0.30	0.41	0.51	0.61	0.72	0.85	1.00	1.22	2.22	
	lnBM	-2.09	-1.25	-0.94	-0.72	-0.53	-0.36	-0.20	-0.03	0.17	0.68	
	Mktcap	4677	4225	3864	2888	2410	2215	1897	1748	1433	689	
	β	1.00	0.93	0.89	0.85	0.82	0.78	0.73	0.68	0.63	0.57	
	%NYSE	21.36%	30.06%	34.51%	37.20%	38.10%	38.53%	38.43%	37.44%	34.99%	26.63%	
	n	616	412	360	331	323	319	318	328	355	480	
EW portfolios	Excess return	0.20	0.59	0.70	0.80	0.91	0.90	0.90	0.93	1.05	1.20	1.20
		(0.58)	(2.0)	(2.54)	(2.86)	(3.34)	(3.36)	(3.52)	(3.66)	(3.9)	(3.46)	(4.44)
	CAPM α	-0.70	-0.22	-0.07	0.04	0.19	0.22	0.26	0.31	0.42	0.57	1.26
		(-3.65)	(-1.46)	(-0.48)	(0.29)	(1.27)	(1.44)	(1.77)	(2.04)	(2.58)	(2.36)	(5.23)
VW portfolios	Excess return	0.61	0.74	0.69	0.63	0.69	0.67	0.75	0.66	0.72	0.87	0.25
		(2.5)	(3.57)	(3.4)	(2.91)	(3.32)	(3.32)	(3.58)	(3.14)	(3.61)	(3.4)	(1.09)
	CAPM α	-0.10	0.08	0.04	-0.04	0.06	0.07	0.19	0.09	0.16	0.22	0.32
		(-0.98)	(1.2)	(0.52)	(-0.5)	(0.6)	(0.75)	(1.46)	(0.62)	(1.15)	(1.3)	(1.29)

Table 2.5: Bivariate Dependent-Sort Portfolio Analysis
This table presents the results of bivariate dependent-sort portfolio analyses of the relation between BM and future stock returns after controlling for the effect of each of β and MktCap (control variables) from 1980 to 2019. Each month, all stocks in the CRSP sample are sorted into five groups based on an ascending sort of one of the control variables. Within each control variable group, all stocks are sorted into five portfolios based on an ascending sort of BM. The quintile breakpoints used to create the portfolios are calculated using all stocks in the CRSP sample. Panel A presents the average return and CAPM alpha (in percent per month) of the long-short zero-cost portfolios that are long the fifth BM quintile portfolio and short the first BM quintile portfolio in each quintile, as well as for the average quintile, of the control variable. Panel B presents the average return and CAPM alpha for the average control variable quintile portfolio within each BM quintile, as well as for the difference between the fifth and first BM quintiles. Results for equal-weighted (Weights = EW) and value-weighted (Weights = VW) portfolios are shown. Newey and West (1987) t-statistics using six lags, testing the null hypothesis that the average return or alpha is equal to zero, are shown in parentheses.

Panel A: BM Difference Portfolios								
Control	Weights	Value	Control1	Control2	Control3	Control4	Control5	ControlAvg
β	EW	Return	1.04	0.81	0.64	0.75	0.92	0.83
			(5.17)	(4.43)	(4.06)	(3.99)	(3.41)	(5.02)
		CAPM α	1.16	0.89	0.74	0.83	0.97	0.92
			(5.69)	(4.89)	(4.51)	(4.2)	(3.55)	(5.39)
	VW	Return	0.49	0.43	0.08	0.15	-0.03	0.22
			(2.03)	(2.3)	(0.48)	(0.78)	(-0.11)	(1.47)
		CAPM α	0.44	0.51	0.12	0.21	0.01	0.26
			(1.77)	(2.77)	(0.68)	(1.0)	(0.04)	(1.62)
Mktcap	EW	Return	1.03	1.30	0.95	0.57	0.17	0.80
			(3.65)	(5.0)	(3.64)	(2.14)	(0.77)	(3.61)
		CAPM α	1.28	1.53	1.22	0.86	0.43	1.07
			(4.73)	(5.77)	(4.53)	(3.05)	(1.8)	(4.63)
	VW	Return	1.14	1.19	0.89	0.54	0.03	0.76
			(4.18)	(4.47)	(3.38)	(2.06)	(0.15)	(3.44)
		CAPM α	1.37	1.43	1.16	0.83	0.16	0.99
			(5.03)	(5.19)	(4.29)	(2.98)	(0.7)	(4.24)
Panel B: Avarage Control Variable Portfolios								
Control	Weights	Value	BM1	BM2	BM3	BM4	BM5	BM5-1
β	EW	Return	0.30	0.69	0.82	0.94	1.13	0.83
			(0.89)	(2.45)	(3.09)	(3.51)	(3.38)	(5.02)
		CAPM α	-0.52	-0.06	0.10	0.24	0.40	0.92
			(-2.79)	(-0.39)	(0.73)	(1.75)	(1.96)	(5.39)
	VW	Return	0.60	0.63	0.69	0.69	0.83	0.22
			(2.81)	(3.18)	(3.61)	(3.66)	(3.72)	(1.47)
		CAPM α	-0.02	0.02	0.10	0.14	0.23	0.26
			(-0.33)	(0.35)	(1.59)	(1.67)	(1.85)	(1.62)
Mktcap	EW	Return	0.24	0.72	0.91	0.96	1.05	0.80
			(0.66)	(2.41)	(3.33)	(3.64)	(3.57)	(3.61)
		CAPM α	-0.66	-0.07	0.19	0.31	0.40	1.07
			(-3.16)	(-0.42)	(1.33)	(2.04)	(2.18)	(4.63)
	VW	Return	0.14	0.61	0.78	0.88	0.90	0.76
			(0.39)	(2.16)	(2.98)	(3.48)	(3.15)	(3.44)
		CAPM α	-0.74	-0.16	0.08	0.24	0.25	0.99
			(-3.83)	(-1.14)	(0.62)	(1.75)	(1.43)	(4.24)

Table 2.6: Bivariate Independent-Sort Portfolio Analysis–Control for β
This table presents the results of bivariate independent-sort portfolio analyses of the relation between BM and future stock returns after controlling for the effect of β from 1980 to 2019. Each month, all stocks in the CRSP sample are sorted into five groups based on an ascending sort of β. All stocks are independently sorted into five groups based on an ascending sort of BM. The quintile breakpoints used to create the groups are calculated using all stocks in the CRSP sample. The intersections of the β and BM groups are used to form 25 portfolios. The table presents the average one-month-ahead excess return (in percent per month) for each of the 25 portfolios as well as for the average β quintile portfolio within each quintile of BM and the average BM quintile within each β quintile. Also shown are the average return and CAPM alpha of a long-short zero-cost portfolio that is long the fifth BM (β quintile portfolio and short the first BM (β quintile portfolio in each β(BM) quintile. Newey and West (1987) t-statistics using six lags, testing the null hypothesis that the average return or alpha is equal to zero, are shown in parentheses. Panel A presents results for equal-weighted portfolios. Panel B presents results for value-weighted portfolios.

Panel A: Equal-Weighted Portfolios

	β1	β2	β3	β4	β5	βAvg	β5-1	β5-1CAPMα
BM1	0.21	0.34	0.47	0.23	0.05	0.26	-0.16	-0.69
							(-0.69)	(-2.76)
BM2	0.66	0.68	0.72	0.71	0.58	0.67	-0.08	-0.65
							(-0.38)	(-3.02)
BM3	1.07	0.83	0.86	0.91	0.76	0.89	-0.31	-0.95
							(-1.43)	(-4.06)
BM4	1.02	0.92	0.93	0.87	0.82	0.91	-0.20	-0.85
							(-0.73)	(-3.33)
BM5	1.30	1.25	1.18	1.06	0.87	1.13	-0.43	-1.05
							(-1.48)	(-3.8)
BMAvg	0.85	0.80	0.83	0.76	0.61		-0.24	-0.84
							(-1.15)	(-4.06)
BM5-1	1.10	0.91	0.70	0.83	0.83	0.87		
	(4.79)	(4.46)	(3.81)	(4.18)	(2.85)	(4.95)		
BM 5-1	1.27	1.01	0.79	0.92	0.92	0.98		
CAPM	(5.95)	(4.89)	(4.18)	(4.32)	(3.11)	(5.45)		

Panel B: Value-Weighted Portfolios

	β1	β2	β3	β4	β5	βAvg	β5-1	β5-1CAPMα
BM1	0.46	0.48	0.71	0.57	0.53	0.55	0.07	-0.57
							(0.2)	(-1.82)
BM2	0.47	0.69	0.62	0.84	0.55	0.64	0.08	-0.44
							(0.31)	(-1.66)
BM3	0.73	0.70	0.70	0.59	0.70	0.68	-0.03	-0.60
							(-0.12)	(-2.25)
BM4	0.84	0.69	0.71	0.67	0.64	0.71	-0.20	-0.84
							(-0.73)	(-3.86)
BM5	0.93	0.86	0.85	0.91	0.48	0.80	-0.45	-1.15
							(-1.45)	(-3.63)
BMAvg	0.69	0.68	0.72	0.71	0.58		-0.11	-0.72
							(-0.44)	(-3.31)
BM5-1	0.47	0.38	0.14	0.35	-0.05	0.25		
	(1.79)	(2.06)	(0.8)	(1.7)	(-0.18)	(1.69)		
BM 5-1	0.50	0.46	0.18	0.38	-0.08	0.29		
CAPM	(1.87)	(2.59)	(1.04)	(1.78)	(-0.29)	(1.87)		

Table 2.7: Bivariate Independent-Sort Portfolio Analysis–Control for Mktcap
This table presents the results of bivariate independent-sort portfolio analyses of the relation between BM and future stock returns after controlling for the effect of Mktcap from 1980 to 2019. Each month, all stocks in the CRSP sample are sorted into five groups based on an ascending sort of Mktcap. All stocks are independently sorted into five groups based on an ascending sort of BM. The quintile breakpoints used to create the groups are calculated using all stocks in the CRSP sample. The intersections of the Mktcap and BM groups are used to form 25 portfolios. The table presents the average one-month-ahead excess return (in percent per month) for each of the 25 portfolios as well as for the average Mktcap quintile portfolio within each quintile of BM and the average BM quintile within each Mktcap quintile. Also shown are the average return and CAPM alpha of a long-short zero-cost portfolio that is long the fifth BM (Mktcap) quintile portfolio and short the first BM (Mktcap) quintile portfolio in each Mktcap (BM) quintile. Newey and West (1987) t-statistics using six lags, testing the null hypothesis that the average return or alpha is equal to zero, are shown in parentheses. Panel A presents results for equal-weighted portfolios. Panel B presents results for value-weighted portfolios.

Panel A: Equal-Weighted Portfolios								
	Mktcap 1	Mktcap 2	Mktcap 3	Mktcap 4	Mktcap 5	Mktcap Avg	Mktcap 5-1	Mktcap 5-1 CAPM α
BM1	0.76	-0.39	-0.07	0.34	0.60	0.25	-0.16	-0.10
							(-0.38)	(-0.24)
BM2	0.87	0.33	0.66	0.77	0.71	0.67	-0.16	-0.16
							(-0.49)	(-0.48)
BM3	1.34	0.76	0.85	0.86	0.79	0.92	-0.54	-0.55
							(-1.8)	(-1.86)
BM4	1.33	0.78	0.92	0.94	0.73	0.94	-0.60	-0.59
							(-2.31)	(-2.22)
BM5	1.63	0.89	0.89	0.84	0.84	1.02	-0.79	-0.81
							(-2.77)	(-2.75)
BMAvg	1.19	0.47	0.65	0.75	0.73		-0.45	(-0.44)
							(-1.53)	(-1.52)
BM5-1	0.87	1.28	0.96	0.50	0.24	0.77		
	(2.83)	(4.93)	(3.53)	(1.83)	(1.11)	(3.36)		
BM 5-1	1.19	1.53	1.21	0.75	0.48	1.03		
CAPM α	(4.15)	(5.81)	(4.32)	(2.53)	(2.06)	(4.38)		
Panel B: Valuel-Weighted Portfolios								

Table 2.7 continued from previous page

	Mktcap 1	Mktcap 2	Mktcap 3	Mktcap 4	Mktcap 5	Mktcap Avg	Mktcap 5-1	Mktcap 5-1 CAPM α
BM1	0.00	-0.30	-0.02	0.37	0.62	0.14	0.62	0.77
							(1.46)	(1.95)
BM2	0.22	0.36	0.66	0.77	0.67	0.54	0.46	0.55
							(1.46)	(1.74)
BM3	0.81	0.73	0.88	0.85	0.66	0.78	-0.15	-0.09
							(-0.51)	(-0.31)
BM4	0.81	0.80	0.92	0.93	0.63	0.82	-0.19	-0.16
							(-0.71)	(-0.61)
BM5	1.14	0.86	0.88	0.84	0.76	0.90	-0.38	-0.38
							(-1.43)	(-1.41)
BMAvg	0.59	0.49	0.66	0.75	0.67		0.07	0.14
							(0.26)	(0.5)
BM5-1	1.14	1.17	0.89	0.47	0.14	0.76		
	(3.8)	(4.42)	(3.27)	(1.73)	(0.66)	(3.33)		
BM 5-1	1.43	1.41	1.16	0.72	0.28	1.00		
CAPM α	(5.01)	(5.21)	(4.08)	(2.47)	(1.2)	(4.19)		

Table 2.8: Fama-MacBeth Regression Analysis
This table presents the results of Fama and MacBeth (1973) regression analyses of the relation between expected stock returns and book-to-market ratio using a sample period from 1980 to 2019. Each column in the table presents results for a different cross-sectional regression specification. The dependent variable in all specifications is the one-month-ahead excess stock return. The independent variables are indicated in the first column. Independent variables are winsorized at the 0.5% level on a monthly basis. The table presents average slope and intercept coefficients along with t-statistics (in parentheses), adjusted following Newey and West (1987) using six lags, testing the null hypothesis that the average coefficient is equal to zero. The rows labeled Adj.R^2 and n present the average adjusted R-squared and the number of data points, respectively, for the cross-sectional regressions.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BM	0.37	0.32	0.30	0.25				
	(4.2)	(3.84)	(3.12)	(3.02)				
lnBM					0.40	0.35	0.34	0.30
					(5.19)	(5.26)	(3.89)	(4.23)
β		-0.20		-0.11		-0.15		-0.04
		(-1.4)		(-0.53)		(-1.14)		(-0.24)
Size			-0.07	-0.06			-0.07	-0.07
			(-1.41)	(-0.98)			(-1.35)	(-1.08)
Intercept	0.45	0.66	0.84	0.93	0.98	1.08	1.27	1.28
	(1.52)	(2.68)	(1.64)	(2.01)	(3.41)	(4.14)	(2.69)	(2.9)
Adj. R^2	0.00	0.02	0.01	0.03	0.01	0.02	0.01	0.03
n	3885	3862	3885	3862	3885	3862	3885	3862

Table 2.9: Summary Statistics for Comparison between the Four Versions of Book-to-Market Ratio
This table presents summary statistics for four versions of book-to-market ratios (bm^{FF}, $bm^{Eisfeldt}$, bm^{Peters} and bm^{Ewens} and $logbm^{FF}$ (the natural log of bm^{FF}) and market value of equity (ME) from period June 1976 to December 2019. ME is adjusted to reflect the 2021 dollar using the consumer price index and recorded in millions of dollars.

	mean	std	skew	kurtosis	min	5%	25%	median	75%	95%	max	count
logbm^{FF}	-0.56	0.91	-0.89	3.01	-7.09	-2.14	-1.05	-0.45	0.04	0.71	2.85	3849
bm^{FF}	0.86	0.88	6.97	135.60	0.00	0.13	0.37	0.67	1.09	2.14	20.38	3849
$bm^{Eisfeldt}$	2.67	5.07	11.59	294.90	0.01	0.27	0.76	1.44	2.84	8.51	144.44	3932
bm^{Peters}	1.48	1.97	9.11	201.43	0.01	0.22	0.58	1.02	1.71	4.13	48.55	3934
bm^{Ewens}	1.37	1.65	7.25	125.08	0.01	0.21	0.56	0.97	1.61	3.75	36.95	3929
ME	2224	10402	15	355	1	10	53	222	941	8435	265186	4397

Table 2.10: Correlation between the Four Versions of Book-to-Market Ratio
This table presents the time-series averages of cross-sectional Pearson product-moment (below-diagonal entries) and Spearman rank (above-diagonal entries) correlations between the four versions of the book-to-market ratio for the sample period from June 1976 to December 2019.

	bm^{FF}	$bm^{Eisfeldt}$	bm^{Peters}	bm^{Ewens}
bm^{FF}	1.00	0.73	0.87	0.89
$bm^{Eisfeldt}$	0.66	1.00	0.91	0.90
bm^{Peters}	0.81	0.92	1.00	0.99
bm^{Ewens}	0.84	0.90	0.97	1.00

Table 2.11: Persistence of the Four Versions of Book-to-Market Ratio
This table presents the results of persistence analyses of four versions of the book-to-market ratio for the sample period from June 1976 to December 2019. Each month t, the cross-sectional Pearson product–moment correlation between the month t and month $t+\tau$ values of the given variable is calculated. The table presents the time-series averages of the monthly cross-sectional correlations. The column labeled τ indicates the lag at which the persistence is measured.

τ	bm^{FF}	$bm^{Eisfeldt}$	bm^{Peters}	bm^{Ewens}
12	0.762	0.816	0.769	0.761
24	0.616	0.692	0.626	0.613
36	0.517	0.610	0.532	0.516
48	0.447	0.546	0.464	0.447
60	0.395	0.490	0.405	0.387
120	0.301	0.366	0.287	0.269

Table 2.12: Univariate Portfolio Analysis for Comparison between the Four Versions of Book-to-Market Ratio
This table presents the averages of value-weighted one-month-ahead excess returns and corresponding Newey and West (1987) t-statistics using 6 lags within portfolios formed by sorting different book-to-market ratios. The first row indicates the sorting ratio used, and the first column indicates the portfolio with ratios from low to high, and also for the long-short portfolio (10-1).. The sample period is from June 1976 to December 2019.

	bm^{FF}		$bm^{Eisfeldt}$		bm^{Peters}		bm^{Ewens}	
	VW excess return	t VW	VW excess return	t VW	VW excess return	t VW	VW excess return	t VW
1	0.60	(2.6)	0.55	(2.34)	0.57	(2.42)	0.59	(2.53)
2	0.62	(2.96)	0.62	(2.94)	0.63	(3.01)	0.61	(2.9)
3	0.68	(3.4)	0.65	(3.34)	0.62	(3.08)	0.62	(3.07)
4	0.69	(3.36)	0.65	(3.53)	0.71	(3.69)	0.70	(3.68)
5	0.70	(3.45)	0.68	(3.52)	0.64	(3.37)	0.67	(3.51)
6	0.67	(3.38)	0.75	(3.91)	0.71	(3.59)	0.69	(3.48)
7	0.70	(3.41)	0.77	(3.71)	0.73	(3.63)	0.72	(3.5)
8	0.67	(3.26)	0.75	(3.44)	0.73	(3.46)	0.74	(3.57)
9	0.74	(3.67)	0.76	(3.44)	0.76	(3.38)	0.74	(3.31)
10	0.72	(3.27)	0.74	(3.27)	0.73	(3.2)	0.73	(3.17)
10-1	0.13	(1.4)	0.19	(2.15)	0.16	(1.8)	0.14	(1.55)

Table 2.13: Fama-MacBeth Regression Using Four Versions of Book-to-Market Ratio
This table presents the results of Fama and MacBeth (1973) regression analyses of the relation between expected stock returns and book-to-market ratio using the sample period from June 1976 to December 2019. Each column in the table presents results for a different cross-sectional regression specification. The dependent variable in all specifications is the one-month-ahead excess stock return. The independent variables are indicated in the first column. Independent variables are winsorized at the 0.5% level on a monthly basis. The table presents average slope and intercept coefficients along with t-statistics (in parentheses), adjusted following Newey and West (1987) using six lags, testing the null hypothesis that the average coefficient is equal to zero. The rows labeled Adj. R^2 and n present the average adjusted R-squared and the number of data points, respectively, for the cross-sectional regressions.

	(1)	(2)	(3)	(4)	(5)
logbm^{FF}	0.27				-0.26
	(3.98)				(-1.79)
log$bm^{Eisfeldt}$		0.26			-0.23
		(4.6)			(-1.93)
logbm^{Peters}			0.38		-0.04
			(5.97)		(-0.16)
logbm^{Ewens}				0.39	0.92
				(6.12)	(3.18)
β	0.00	-0.02	0.01	0.01	-0.02
	(0.02)	(-0.1)	(0.05)	(0.04)	(-0.1)
Size	-0.09	-0.06	-0.05	-0.05	-0.05
	(-1.49)	(-0.98)	(-0.92)	(-0.88)	(-0.96)
Intercept	1.35	0.98	1.04	1.06	1.07
	(3.27)	(2.31)	(2.52)	(2.56)	(2.86)
Adj. R^2	0.03	0.03	0.03	0.03	0.03
n	3851	3934	3936	3931	3785

Table 2.14: Summary Statistics for Pre-1999 and Post-1999 Periods
This table presents summary statistics for four versions of book-to-market ratios (bm^{FF}, $bm^{Eisfeldt}$, bm^{Peters} and bm^{Ewens} and $logbm^{FF}$ (the natural log of bm^{FF}) and market value of equity (ME). Panel A covers the sample period from June 1976 to December 1998 and panel B covers the sample period from January 1999 to December 2019. ME is adjusted to reflect the 2021 dollar using the consumer price index and recorded in millions of dollars.

Panel A: pre 1999												
	mean	std	skew	kurtosis	min	5%	25%	median	75%	95%	max	count
logbm^{FF}	-0.43	0.89	-0.95	2.92	-6.84	-2.00	-0.90	-0.31	0.17	0.79	2.87	4092
bm^{FF}	0.95	0.89	6.33	122.19	0.00	0.15	0.43	0.77	1.23	2.28	19.81	4092
$bm^{Eisfeldt}$	3.17	5.84	11.94	337.19	0.01	0.32	0.92	1.75	3.47	9.94	167.37	4180
bm^{Peters}	1.62	1.99	8.49	185.21	0.01	0.23	0.65	1.15	1.93	4.39	47.76	4165
bm^{Ewens}	1.50	1.71	7.33	140.70	0.01	0.23	0.62	1.10	1.81	3.99	39.42	4161
ME	538	2520	18	469	0	3	14	53	230	2160	74430	4608
Panel B: post 1999												
	mean	std	skew	kurtosis	min	5%	25%	median	75%	95%	max	count
logbm^{FF}	-0.70	0.93	-0.83	3.11	-7.36	-2.30	-1.21	-0.59	-0.10	0.62	2.83	3588
bm^{FF}	0.76	0.88	7.66	150.02	0.00	0.10	0.30	0.56	0.93	1.98	21.00	3588
$bm^{Eisfeldt}$	2.13	4.25	11.21	249.42	0.00	0.20	0.59	1.10	2.16	6.97	119.78	3666
bm^{Peters}	1.33	1.94	9.78	218.88	0.01	0.21	0.51	0.87	1.47	3.85	49.39	3686
bm^{Ewens}	1.23	1.59	7.16	108.28	0.00	0.20	0.48	0.83	1.39	3.48	34.29	3681
ME	4037	18878	13	232	1	16	95	403	1705	15184	470323	4169

Table 2.15: Correlation between Book-to-Market Ratios for Pre-1999 and Post-1999 Periods
This table presents the time-series averages of cross-sectional Pearson product-moment (below-diagonal entries) and Spearman rank (above-diagonal entries) correlations between the four versions of the book-to-market ratio for the sample period from June 1976 to December 1998 (Panel A) and from January 1999 to December 2019 (Panel B).

Panel A: pre 1999				
	bm^{FF}	$bm^{Eisfeldt}$	bm^{Peters}	bm^{Ewens}
bm^{FF}	1.00	0.75	0.89	0.90
$bm^{Eisfeldt}$	0.66	1.00	0.93	0.92
bm^{Peters}	0.82	0.94	1.00	0.99
bm^{Ewens}	0.84	0.90	0.98	1.00
Panel B: post 1999				
	bm^{FF}	$bm^{Eisfeldt}$	bm^{Peters}	bm^{Ewens}
bm^{FF}	1.00	0.69	0.83	0.87
$bm^{Eisfeldt}$	0.66	1.00	0.88	0.88
bm^{Peters}	0.80	0.91	1.00	0.98
bm^{Ewens}	0.84	0.90	0.97	1.00

Table 2.16: Persistence of Book-to-Market Ratios for Pre-1999 and Post-1999 Periods
This table presents the results of persistence analyses of four versions of the book-to-market ratio for the sample period from June 1976 to December 1998 (Panel A) and from January 1999 to December 2019 (Panel B). Each month t, the cross-sectional Pearson product–moment correlation between the month t and month $t+\tau$ values of the given variable is calculated. The table presents the time-series averages of the monthly cross-sectional correlations. The column labeled τ indicates the lag at which the persistence is measured.

Panel A: pre 1999				
τ	bm^{FF}	$bm^{Eisfeldt}$	bm^{Peters}	bm^{Ewens}
12	0.781	0.844	0.811	0.798
24	0.634	0.724	0.674	0.654
36	0.528	0.639	0.577	0.551
48	0.449	0.572	0.500	0.470
60	0.390	0.513	0.437	0.403
120	0.274	0.393	0.308	0.264
Panel B: post 1999				
τ	bm^{FF}	$bm^{Eisfeldt}$	bm^{Peters}	bm^{Ewens}
12	0.746	0.788	0.726	0.724
24	0.603	0.663	0.578	0.575
36	0.516	0.585	0.490	0.488
48	0.457	0.527	0.433	0.432
60	0.410	0.475	0.378	0.381
120	0.327	0.351	0.268	0.279

Table 2.17: Univariate Portfolio Analysis for Pre-1999 and Post-1999 Periods
This table presents the averages of value-weighted one-month-ahead excess returns and corresponding Newey and West (1987) t-statistics using 6 lags within portfolios formed by sorting different book-to-market ratios. The first row indicates the sorting ratio used, and the first column indicates the portfolio with ratios from low to high, and also for the long-short portfolio (10-1). The sample periods are from June 1976 to December 1998 (Panel A) and from January 1999 to December 2019 (Panel B) respectively.

Panel A: pre 1999								
	bm^{FF}		$bm^{Eisfeldt}$		bm^{Peters}		bm^{Ewens}	
	VW excess return	t VW	VW excess return	t VW	VW excess return	t VW	VW excess return	t VW
1	0.70	(2.14)	0.64	(1.94)	0.67	(2.03)	0.69	(2.1)
2	0.69	(2.32)	0.72	(2.41)	0.71	(2.28)	0.66	(2.19)
3	0.81	(2.85)	0.80	(2.91)	0.76	(2.71)	0.78	(2.78)
4	0.85	(2.95)	0.77	(3.15)	0.88	(3.31)	0.87	(3.33)
5	0.86	(3.1)	0.82	(3.19)	0.77	(2.96)	0.81	(3.05)
6	0.78	(2.93)	0.90	(3.39)	0.89	(3.36)	0.88	(3.36)
7	0.90	(3.23)	0.94	(3.26)	0.88	(3.33)	0.87	(3.13)
8	0.85	(3.18)	0.91	(2.92)	0.92	(3.12)	0.93	(3.29)
9	0.91	(3.3)	0.97	(3.04)	0.98	(3.07)	0.96	(3.05)
10	0.92	(2.97)	0.92	(2.78)	0.92	(2.79)	0.92	(2.75)
10-1	0.23	(1.73)	0.28	(2.2)	0.25	(2.14)	0.23	(1.87)
Panel B: post 1999								
	bm^{FF}		$bm^{Eisfeldt}$		bm^{Peters}		bm^{Ewens}	
	VW excess return	t VW	VW excess return	t VW	VW excess return	t VW	VW excess return	t VW
1	0.49	(1.52)	0.46	(1.36)	0.46	(1.38)	0.48	(1.46)
2	0.54	(1.83)	0.52	(1.72)	0.55	(1.94)	0.55	(1.88)
3	0.53	(1.88)	0.49	(1.77)	0.47	(1.64)	0.44	(1.54)
4	0.52	(1.77)	0.52	(1.87)	0.53	(1.88)	0.52	(1.86)
5	0.52	(1.76)	0.54	(1.81)	0.51	(1.79)	0.53	(1.89)
6	0.55	(1.86)	0.60	(2.08)	0.52	(1.74)	0.50	(1.63)
7	0.49	(1.61)	0.58	(1.92)	0.57	(1.83)	0.56	(1.81)
8	0.49	(1.52)	0.58	(1.88)	0.53	(1.72)	0.53	(1.74)
9	0.55	(1.85)	0.54	(1.75)	0.52	(1.64)	0.51	(1.58)
10	0.51	(1.61)	0.56	(1.77)	0.52	(1.67)	0.52	(1.66)
10-1	0.02	(0.18)	0.10	(0.82)	0.06	(0.47)	0.04	(0.35)

Table 2.18: Fama-MacBeth Regression for Pre-1999 and Post-1999 Periods
This table presents the results of Fama and MacBeth (1973) regression analyses of the relation between expected stock returns and book-to-market ratio using a sample period from June 1976 to December 1999 (Pre-1999) and from January 1999 to December 2019 (Post-1999). Each column in the table presents results for a different cross-sectional regression specification. The dependent variable in all specifications is the one-month-ahead excess stock return. The independent variables are indicated in the first column. Independent variables are winsorized at the 0.5% level on a monthly basis. The table presents average slope and intercept coefficients along with t-statistics (in parentheses), adjusted following Newey and West (1987) using six lags, testing the null hypothesis that the average coefficient is equal to zero. The rows labeled Adj. Adj. R^2 and n present the average adjusted R-squared and the number of data points, respectively, for the cross-sectional regressions.

	Pre 1999					Post 1999				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
logbm^{FF}	0.41				-0.17	0.12				-0.33
	(5.02)				(-1.16)	(1.23)				(-1.31)
log$bm^{Eisfeldt}$		0.33			-0.30		0.18			-0.15
		(4.57)			(-1.96)		(2.21)			(-0.83)
logbm^{Peters}			0.47		-0.11			0.29		0.02
			(5.85)		(-0.35)			(3.09)		(0.06)
logbm^{Ewens}				0.49	1.07				0.29	0.75
				(6.12)	(3.25)				(3.04)	(1.56)
β	0.01	-0.05	-0.01	-0.01	-0.01	-0.03	-0.01	0.01	0.01	-0.04
	(0.07)	(-0.27)	(-0.06)	(-0.07)	(-0.07)	(-0.09)	(-0.02)	(0.02)	(0.02)	(-0.14)
Size	-0.06	-0.03	-0.03	-0.03	-0.04	-0.10	-0.07	-0.06	-0.06	-0.05
	(-0.83)	(-0.4)	(-0.45)	(-0.39)	(-0.58)	(-1.07)	(-0.75)	(-0.63)	(-0.63)	(-0.58)
Intercept	1.14	0.71	0.84	0.86	1.01	1.50	1.18	1.16	1.18	1.06
	(2.31)	(1.35)	(1.66)	(1.68)	(2.11)	(2.21)	(1.72)	(1.74)	(1.76)	(1.79)
Adj. R^2	0.02	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.03	0.04
n	4088	4176	4161	4156	4034	3590	3669	3688	3683	3511

Table 2.19: Summary Statistics within the High Tech Sector and the Traditional Sector
This table presents summary statistics for four versions of book-to-market ratios (bm^{FF}, $bm^{Eisfeldt}$, bm^{Peters} and bm^{Ewens} and $logbm^{FF}$ (the natural log of bm^{FF}) and market value of equity (ME). Panel A covers the sample of the high-tech sector stocks and panel B covers the sample of the traditional sector. ME is adjusted to reflect the 2021 dollar using the consumer price index and recorded in millions of dollars.

Panel A: High tech sector												
	mean	std	skew	kurtosis	min	5%	25%	median	75%	95%	max	count
logbm^{FF}	-0.91	0.94	-0.80	2.54	-6.37	-2.51	-1.45	-0.83	-0.27	0.43	1.96	1113
bm^{FF}	0.62	0.63	4.72	60.05	0.00	0.09	0.25	0.46	0.80	1.60	8.73	1113
$bm^{Eisfeldt}$	2.09	3.16	6.18	74.08	0.02	0.21	0.60	1.20	2.40	6.67	48.70	1149
bm^{Peters}	1.28	1.54	5.07	55.72	0.02	0.19	0.47	0.85	1.54	3.69	21.83	1154
bm^{Ewens}	1.30	1.55	4.84	48.56	0.02	0.19	0.47	0.86	1.58	3.77	20.47	1152
ME	2399	12456	12	203	1	8	42	164	663	7529	233763	1205
Panel B: Traditional sector												
	mean	std	skew	kurtosis	min	5%	25%	median	75%	95%	max	count
logbm^{FF}	-0.41	0.85	-0.93	3.58	-6.38	-1.89	-0.86	-0.32	0.12	0.78	2.78	2736
bm^{FF}	0.95	0.94	6.52	111.37	0.00	0.16	0.44	0.76	1.17	2.30	19.21	2736
$bm^{Eisfeldt}$	2.87	5.60	11.11	251.27	0.01	0.31	0.85	1.53	2.99	9.08	143.54	2783
bm^{Peters}	1.54	2.07	9.09	185.15	0.01	0.25	0.64	1.07	1.75	4.21	47.15	2780
bm^{Ewens}	1.39	1.66	7.52	127.66	0.01	0.23	0.60	1.00	1.61	3.67	34.77	2777
ME	2145	9249	14	279	1	11	61	260	1082	8689	226719	3191

Table 2.20: Correlation between Book-to-Market Ratios within the High Tech Sector and the Traditional Sector
This table presents the time-series averages of cross-sectional Pearson product-moment (below-diagonal entries) and Spearman rank (above-diagonal entries) correlations between the four versions of the book-to-market ratio for sample period from June 1976 to December 2019 within the high-tech sector (Panel A) and the traditional sector (Panel B).

Panel A: High Tech Sector				
	bm^{FF}	$bm^{Eisfeldt}$	bm^{Peters}	bm^{Ewens}
bm^{FF}	1.00	0.77	0.84	0.86
$bm^{Eisfeldt}$	0.70	1.00	0.87	0.92
bm^{Peters}	0.77	0.88	1.00	0.98
bm^{Ewens}	0.80	0.95	0.96	1.00
Panel B: Traditional Sector				
	bm^{FF}	$bm^{Eisfeldt}$	bm^{Peters}	bm^{Ewens}
bm^{FF}	1.00	0.71	0.88	0.91
$bm^{Eisfeldt}$	0.65	1.00	0.92	0.89
bm^{Peters}	0.83	0.94	1.00	0.99
bm^{Ewens}	0.87	0.90	0.98	1.00

Table 2.21: Persistence of Book-to-Market Ratio within the High Tech Sector and the Traditional Sector
This table presents the results of persistence analyses of four versions of the book-to-market ratio for the sample period from June 1976 to December 2019 within the high-tech sector (Panel A) and the traditional sector (Panel B). Each month t, the cross-sectional Pearson product–moment correlation between the month t and month $t+\tau$ values of the given variable is calculated. The table presents the time-series averages of the monthly cross-sectional correlations. The column labeled τ indicates the lag at which the persistence is measured.

Panel A: High tech sector				
	bm^{FF}	$bm^{Eisfeldt}$	bm^{Peters}	bm^{Ewens}
12	0.73	0.77	0.72	0.73
24	0.59	0.64	0.56	0.58
36	0.49	0.55	0.46	0.49
48	0.43	0.48	0.39	0.42
60	0.37	0.42	0.32	0.36
120	0.24	0.27	0.18	0.22
Panel B: Traditional sector				
	bm^{FF}	$bm^{Eisfeldt}$	bm^{Peters}	bm^{Ewens}
12	0.76	0.83	0.79	0.78
24	0.61	0.71	0.65	0.64
36	0.50	0.63	0.56	0.54
48	0.43	0.56	0.49	0.47
60	0.38	0.51	0.44	0.41
120	0.30	0.39	0.33	0.31

Table 2.22: Univariate Portfolio Analysis within the High Tech Sector and the Traditional Sector
This table presents the averages of value-weighted one-month-ahead excess returns and corresponding Newey and West (1987) t-statistics using 6 lags within portfolios formed by sorting different book-to-market ratios. The first row indicates the sorting ratio used, and the first column indicates the portfolio with ratios from low to high, and also for the long-short portfolio (10-1). The sample period is from June 1976 to December 2019. Panel A covers the sample of the high-tech sector stocks and panel B covers the sample of the traditional sector.

Panel A: High tech sector								
	bm^{FF}		$bm^{Eisfeldt}$		bm^{Peters}		bm^{Ewens}	
	VW excess return	t VW	VW excess return	t VW	VW excess return	t VW	VW excess return	t VW
1	0.60	(1.85)	0.71	(2.1)	0.68	(2.04)	0.68	(2.04)
2	0.78	(2.95)	0.67	(2.5)	0.61	(2.19)	0.64	(2.3)
3	0.74	(2.94)	0.68	(2.79)	0.75	(3.1)	0.75	(3.1)
4	0.68	(2.67)	0.81	(3.37)	0.72	(2.93)	0.76	(3.16)
5	0.82	(3.44)	0.79	(3.49)	0.77	(3.16)	0.79	(3.33)
6	0.89	(3.55)	0.94	(3.8)	0.93	(3.8)	0.91	(3.6)
7	0.86	(3.38)	0.81	(3.18)	0.95	(3.74)	0.92	(3.59)
8	0.87	(3.17)	0.97	(3.6)	0.85	(3.12)	0.85	(3.1)
9	0.95	(3.55)	1.08	(3.81)	1.02	(3.74)	1.08	(3.96)
10	1.02	(3.77)	1.13	(4.01)	1.12	(4.01)	1.09	(3.89)
10-1	0.42	(1.9)	0.42	(2.44)	0.44	(2.33)	0.41	(2.31)

Panel B: Traditional sector								
	bm^{FF}		$bm^{Eisfeldt}$		bm^{Peters}		bm^{Ewens}	
	VW excess return	t VW	VW excess return	t VW	VW excess return	t VW	VW excess return	t VW
1	0.59	(2.77)	0.54	(2.4)	0.57	(2.67)	0.58	(2.74)
2	0.57	(2.8)	0.61	(2.95)	0.59	(2.78)	0.60	(2.86)
3	0.65	(3.12)	0.58	(2.97)	0.61	(3.0)	0.59	(2.95)
4	0.68	(3.35)	0.64	(3.34)	0.65	(3.36)	0.67	(3.37)
5	0.65	(3.11)	0.67	(3.34)	0.63	(3.18)	0.62	(3.14)
6	0.67	(3.27)	0.73	(3.69)	0.65	(3.27)	0.64	(3.15)
7	0.64	(3.08)	0.70	(3.32)	0.73	(3.53)	0.71	(3.44)
8	0.67	(3.27)	0.74	(3.38)	0.68	(3.17)	0.70	(3.26)
9	0.67	(3.23)	0.71	(3.15)	0.71	(3.13)	0.69	(3.03)
10	0.69	(3.08)	0.68	(2.99)	0.66	(2.89)	0.66	(2.89)
10-1	0.09	(1.18)	0.14	(1.82)	0.09	(1.14)	0.08	(1.03)

Table 2.23: Fama-MacBeth Regression within the High Tech Sector and the Traditional Sector

This table presents the results of Fama and MacBeth (1973) regression analyses of the relation between expected stock returns and book-to-market ratio using the sample period from June 1976 to December 2019 within the high-tech sector and the traditional sector. Each column in the table presents results for a different cross-sectional regression specification. The dependent variable is the one-month-ahead excess stock return. The independent variables are indicated in the first column and are winsorized at the 0.5% level on a monthly basis. The table presents average slope and intercept coefficients along with t-statistics (in parentheses), adjusted following Newey and West (1987) using six lags, testing the null hypothesis that the average coefficient is equal to zero. The rows labeled Adj. Adj. R^2 and n present the average adjusted R-squared and the number of data points, respectively, for the cross-sectional regressions.

	High tech sector					Traditional Sector				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
$\log bm^{FF}$	0.36				0.09	0.33				0.02
	(5.57)				(0.86)	(5.65)				(0.16)
$\log bm^{Eisfeldt}$		0.29			0.02		0.29			-0.02
		(4.81)			(0.16)		(5.41)			(-0.19)
$\log bm^{Peters}$			0.49		1.34			0.37		0.15
			(7.32)		(2.78)			(5.9)		(0.41)
$\log bm^{Ewens}$				0.43	-0.98				0.39	0.23
				(6.58)	(-2.08)				(6.04)	(0.59)
β	-0.05	-0.07	-0.03	-0.03	-0.04	-0.07	-0.11	-0.10	-0.09	-0.07
	(-0.34)	(-0.44)	(-0.17)	(-0.21)	(-0.25)	(-0.42)	(-0.63)	(-0.55)	(-0.52)	(-0.42)
Size	-0.16	-0.13	-0.11	-0.12	-0.13	-0.03	-0.01	-0.01	-0.01	-0.02
	(-2.36)	(-1.79)	(-1.63)	(-1.68)	(-2.0)	(-0.63)	(-0.14)	(-0.21)	(-0.19)	(-0.33)
Intercept	2.16	1.63	1.66	1.68	1.84	1.01	0.65	0.77	0.80	0.83
	(3.94)	(2.83)	(3.01)	(3.0)	(3.6)	(2.82)	(1.8)	(2.16)	(2.22)	(2.43)
Adj. R^2	0.03	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.03	0.03
n	1113	1149	1154	1153	1094	2738	2785	2782	2778	2691

Table 2.24: Variable Definitions
This table displays the acronym of variables and their long description.

Variables	Description
ChInvIA	Industry adjusted change in capital investment
Herf	Industry concentration (Herfindahl) sales
covana	Number of eps estimates in IBES for one-quarter ahead earnings for the firm
fgr5yrLag	Long-term EPS forecast
UpForecast	Binary variable equal to 1 if the mean estimation for eps increases from lagged mean estimates
sfe	The ratio of mean estimated earnings to price
sueana	Standardized unexpected earnings calculated as the difference between the mean estimation and actual eps, divided by fiscal annual closing price
FDLT	Long-term EPS forecast dispersion, standardized by mean estimation
FD	EPS forecast dispersion, standardized by mean estimation
CashProd	Cash Productivity, calculated as the difference between market value of equity and total assets, divided by cash and short-term investments
AdExp	Advertising expense adjusted by market value
SP	Ratio of annual sales to the market value of equity
EP	Ratio of earnings to the market value of last December
RMom6	Residual momentum of past 11 months
RMom11	Residual momentum of past 6 months
UpsideBeta	Beta on adjusted market returns when the excess market return is above its mean
DownsideBeta	Beta on adjusted market returns when the excess market return is below its mean
PriceDelayRsq	Price delay measured by the fraction of variation of contemporaneous individual stock returns explained by lagged market returns
PriceDelayAdj	Price delay adjusted by standard error
PriceDelay	Price delay measured using betas on past four periods' market returns
betaVIX	Systematic volatility measured as beta on changes in the VIX index
skew3F	Skewness of daily idiosyncratic returns measured using residuals from FF three-factor model
skewCAPM	Skewness of daily idiosyncratic returns measured using residuals from CAPM
rmse3F	The mean of root-mean-square error from FF three-factor model
rmsecapm	The mean of root-mean-square error from CAPM
ill	Illiquidity measured as the mean of absolute return over market value
ShareIss5Y	The growth in the number of shares during the past five years
ShareIss1Y	The growth in the number of shares in the past year
roa	Return on assets defined as the net income over total assets
roe	Return on equity defined as the net income over the market value of equity
roi	Return on investments defined as the net income over invested capital
SurpriseRD	Binary variable equal to 1 if there is an unexpected R&D increase
ChNNCOA	Change in net noncurrent operating assets
ChNWC	Change in net working capital
NOA	Net working capital adjusted by previous total assets
cashdebt	Cash flow to debt

Table 2.24 continued from previous page	
Cash	Cash to assets
AssetTurnover	Asset Turnover calculated as sales over average assets
DelNetFin	Change in net financial assets
AssetGrowth	Asset growth
DelSTI	Change in short-term investment
Accruals	Annual change in current total assets, with the annual change in cash and short-term investments, in current liabilities, in income taxes (txp) excluded, and then divided by average total assets
Investment	Investment to revenue
GrSaleToGrInv	Sales growth minus inventory growth
MRankRevgrw	Weighted mean revenue growth rank
payout	Payout ratio calculated as the total payout over earnings
ShareRepurchase	Binary variable equal to 1 if firm repurchases stocks in cash
op	Operating profitability over book value of equity
prof	Profitability over total assets
gp	Gross profitability over total assets
pm	Profit margin
sueac	Standardized unexpected earnings calculated as the difference between current and previous eps, divided by fiscal annual closing price
ErnSm	Earnings smoothness
ErnPred	Earnings predictability
ErnPers	Earnings persistence
eps	Earnings per share
ZScore	Altman Z score which measures the financial strength
OScore	Ohlson O score which measures the financial distress
EquityDuration	Measure of a share's cash-flow maturity
XFIN	Net external financing, scaled by total assets
OPLeverage	Operating leverage, the sum of administrative expenses and cost of goods sold, scaled by total assets
BookLeverage	Total assets divided by book value of equity plus deferred taxes and preferred stock
BrandCapital	Brand capital to assets
mispscore_avg	The average of the monthly mispricing score which is constructed based on averaging anomaly rankings

Table 2.25: Firms' Characteristics
All stocks are grouped into ten portfolios based on their bm^{FF} (suffix v_1) or bm^{Ewens} (suffix v_4). We report the average of characteristics shown in the first column for stocks in the long leg, i.e. the group of stocks with the highest book-to-market ratio in columns long_v1 and long_v4 separately for using bm^{FF}, and bm^{Ewens} as the sorting variable, in the short leg, i.e. the group of stocks with the lowest book-to-market ratio in columns short_v1 and short_v4 separately for using bm^{FF}, and bm^{Ewens} as the sorting variable. Columns delta_v1, and delta_v4 display the difference between the average characteristics of long and short legs. Columns p_delta_v1, and p_delta_v4 display the p-value of testing if the delta is significantly different from 0. Columns ratio_v1, ratio_v4 display the ratio of delta_v1 over short_v1, delta_v4 over short_v4 separately.

	long_v1	short_v1	ratio_v1	delta_v1	p_delta_v1	long_v4	short_v4	ratio_v4	delta_v4	p_delta_v4
ChInvIA	-0.8179	-0.6278	0.3029	-0.1902	0.4810	-0.9927	-0.5409	0.8354	-0.4518	0.0064
Herf	0.3652	0.3114	0.1728	0.0538	0.0000	0.3581	0.3298	0.0856	0.0282	0.0000
covana	4.9439	8.4439	-0.4145	-3.5000	0.0000	3.6760	9.0694	-0.5947	-5.3933	0.0000
fgr5yr	11.8069	24.4295	-0.5167	-12.6226	0.0000	13.1314	24.1981	-0.4573	-11.0667	0.0000
UpForecast	0.4026	0.7747	-0.4803	-0.3721	0.0000	0.4409	0.8012	-0.4497	-0.3603	0.0000
sfe	-0.2017	-0.1727	0.1680	-0.0290	0.5231	-0.3922	-0.0485	7.0774	-0.3436	0.0004
sueana	0.1808	0.0475	2.8079	0.1333	0.0002	0.1895	0.0447	3.2427	0.1448	0.0001
FDLT	0.3473	0.2289	0.5175	0.1184	0.0000	0.3091	0.2250	0.3738	0.0841	0.0187
FD	0.4804	0.1870	1.5687	0.2934	0.0000	0.5330	0.1411	2.7771	0.3919	0.0000
CashProd	-47.4783	28.0521	-2.6925	-75.5305	0.0000	-40.1379	30.9008	-2.2989	-71.0387	0.0000
AdExp	0.1805	0.0237	6.6152	0.1568	0.0000	0.2044	0.0130	14.6947	0.1914	0.0000
SP	6.5244	0.6319	9.3245	5.8925	0.0000	7.5407	0.4065	17.5494	7.1342	0.0000
EP	-0.2793	-0.0501	4.5788	-0.2292	0.0001	-0.4023	-0.0023	170.3313	-0.4000	0.0000
RMom6	0.0988	-0.1001	-1.9875	0.1989	0.0000	0.1115	-0.1353	-1.8236	0.2468	0.0000
RMom11	0.0632	0.0289	1.1837	0.0342	0.2845	0.0723	0.0173	3.1810	0.0550	0.1612
UpsideBeta	0.0044	0.0083	-0.4761	-0.0040	0.0000	0.0042	0.0088	-0.5248	-0.0046	0.0000
DownsideBeta	0.0070	0.0119	-0.4096	-0.0049	0.0000	0.0072	0.0120	-0.3965	-0.0047	0.0000
PriceDelayRsq	0.7121	0.6180	0.1522	0.0941	0.0000	0.7354	0.5937	0.2388	0.1417	0.0000
PriceDelayAdj	2.5150	2.5602	-0.0176	-0.0452	0.5470	2.5211	2.5161	0.0020	0.0050	0.9376
PriceDelay	1.5380	1.2554	0.2251	0.2826	0.0022	1.6030	1.1635	0.3777	0.4395	0.0000
betaVIX	0.0004	0.0004	-0.0435	0.0000	0.9182	0.0004	0.0004	-0.1576	-0.0001	0.7147
skew3F	0.8017	0.7089	0.1309	0.0928	0.0197	0.9404	0.5984	0.5716	0.3420	0.0000
skewCAPM	0.8003	0.6866	0.1656	0.1137	0.0044	0.9407	0.5734	0.6405	0.3673	0.0000
rmse3F	0.0438	0.0382	0.1468	0.0056	0.0023	0.0497	0.0339	0.4670	0.0158	0.0000
rmsecapm	0.0442	0.0388	0.1390	0.0054	0.0036	0.0500	0.0345	0.4517	0.0156	0.0000
ill	0.0000	0.0000	5.6428	0.0000	0.0000	0.0000	0.0000	12.8380	0.0000	0.0000
roa	-0.0417	-0.1344	-0.6897	0.0927	0.0000	-0.0984	-0.0426	1.3089	-0.0558	0.0005
roe	-0.0965	-0.4634	-0.7918	0.3669	0.0000	-0.2890	-0.1346	1.1471	-0.1544	0.0003
roi	-0.0684	-0.2647	-0.7416	0.1963	0.0000	-0.1878	-0.0771	1.4346	-0.1107	0.0001
SurpriseRD	0.2382	0.3357	-0.2904	-0.0975	0.0000	0.2459	0.3183	-0.2273	-0.0724	0.0000
ChNNCOA	0.0025	-0.0006	-5.0185	0.0032	0.0987	-0.0043	0.0099	-1.4341	-0.0142	0.0000
ChNWC	-0.0023	-0.0036	-0.3588	0.0013	0.3603	-0.0068	0.0015	-5.5969	-0.0082	0.0000
NOA	0.5595	0.4932	0.1346	0.0664	0.0014	0.4938	0.6063	-0.1856	-0.1125	0.0000
cashdebt	-0.0579	-0.3173	-0.8177	0.2595	0.0000	-0.1822	-0.0706	1.5793	-0.1116	0.0141
Cash	0.1057	0.2939	-0.6403	-0.1882	0.0000	0.1352	0.2785	-0.5146	-0.1433	0.0000
AssetTurnover	2.2003	4.1358	-0.4680	-1.9355	0.0000	3.2318	3.5518	-0.0901	-0.3200	0.0298
DelNetFin	-0.0088	-0.0301	-0.7080	0.0213	0.0000	-0.0081	-0.0241	-0.6623	0.0159	0.0065
AssetGrowth	0.0287	0.2801	-0.8975	-0.2513	0.0000	-0.0207	0.3743	-1.0554	-0.3950	0.0000
DelSTI	-0.0028	0.0067	-1.4235	-0.0095	0.0004	-0.0071	0.0159	-1.4448	-0.0229	0.0000
Accruals	-0.0039	0.0140	-1.2786	-0.0179	0.0000	-0.0124	0.0212	-1.5841	-0.0336	0.0000
Investment	0.9158	0.9582	-0.0443	-0.0424	0.0002	0.9086	0.9941	-0.0860	-0.0855	0.0000
GrSaleToGrInv	-0.0724	0.0423	-2.7113	-0.1148	0.0003	-0.0460	0.0148	-4.1067	-0.0608	0.0193

Table 2.25 continued from previous page

MRankRevgrw	3412.7610	2770.1946	0.2320	642.5663	0.0000	3549.7422	2494.2944	0.4231	1055.4479	0.0000
payout	0.2128	0.2612	-0.1854	-0.0484	0.1323	0.1495	0.2974	-0.4975	-0.1480	0.0002
ShareRepurchase	0.3241	0.2782	0.1648	0.0459	0.0019	0.2815	0.3105	-0.0934	-0.0290	0.1314
op	0.0379	-0.0537	-1.7065	0.0916	0.0276	-0.0626	0.1276	-1.4903	-0.1902	0.0000
prof	-0.0430	-0.1371	-0.6865	0.0942	0.0000	-0.0999	-0.0450	1.2213	-0.0549	0.0007
gp	0.2401	0.3591	-0.3315	-0.1190	0.0000	0.3802	0.3130	0.2145	0.0671	0.0000
pm	-0.3252	-1.6388	-0.8016	1.3136	0.0000	-0.6387	-1.1177	-0.4285	0.4790	0.0013
sueac	-0.0446	0.2608	-1.1709	-0.3054	0.0235	0.3728	0.0235	14.8530	0.3493	0.2707
ErnSm	0.3321	0.5986	-0.4452	-0.2665	0.0000	0.3248	0.7218	-0.5501	-0.3971	0.0000
ErnPred	0.6469	0.7933	-0.1845	-0.1463	0.0000	0.7141	0.7215	-0.0102	-0.0073	0.5826
ErnPers	1863.3591	2110.3577	-0.1170	-246.9986	0.4691	2551.1611	909.7246	1.8043	1641.4365	0.0000
eps	-3.2537	-3.9969	-0.1859	0.7431	0.1610	-5.2805	-1.7981	1.9367	-3.4824	0.0001
ZScore	2.4464	9.9147	-0.7533	-7.4684	0.0000	1.7051	13.5605	-0.8743	-11.8554	0.0000
OScore	-0.7833	-0.9922	-0.2105	0.2089	0.2075	-0.1185	-2.3603	-0.9498	2.2418	0.0000
EquityDuration	-21.0499	15.1144	-2.3927	-36.1642	0.0000	-19.0594	14.6181	-2.3038	-33.6775	0.0000
XFIN	0.0093	0.1754	-0.9469	-0.1661	0.0000	0.0156	0.1635	-0.9045	-0.1479	0.0000
OPLeverage	0.9674	1.0645	-0.0912	-0.0971	0.0430	1.3454	0.8729	0.5413	0.4725	0.0000
BookLeverage	3.7554	4.7794	-0.2142	-1.0240	0.0000	3.9128	3.5240	0.1103	0.3888	0.0565
BrandCapital	0.0586	0.0796	-0.2633	-0.0209	0.0000	0.0866	0.0626	0.3842	0.0240	0.0000
mispscore_avg	50.7295	54.4542	-0.0684	-3.7247	0.0001	48.4602	54.7235	-0.1145	-6.2633	0.0000

Table 2.26: Univariate Portfolio Analysis for Comparison Between the BM^{FF} and IM Ratios This table presents the averages of value-weighted one-month-ahead excess returns and corresponding Newey and West (1987) t-statistics using 6 lags within portfolios formed by sorting on BM^{FF} or IM ratios. The first row indicates the sorting ratio used, and the first column indicates the portfolio with ratios from low to high, and also for the long-short portfolio (10-1). The sample period is from June 1976 to December 2019.

	bm^{FF}		$im^{Eisfeldt}$		im^{Peters}		im^{Ewens}	
	VW excess return	t VW	VW excess return	t VW	VW excess return	t VW	VW excess return	t VW
1	0.60	(2.6)	0.53	(2.46)	0.50	(2.36)	0.49	(2.41)
2	0.62	(2.96)	0.59	(3.14)	0.62	(3.24)	0.66	(3.37)
3	0.68	(3.4)	0.57	(3.19)	0.58	(3.09)	0.57	(3.19)
4	0.69	(3.36)	0.64	(3.57)	0.60	(3.31)	0.62	(3.34)
5	0.70	(3.45)	0.68	(3.66)	0.66	(3.67)	0.67	(3.58)
6	0.67	(3.38)	0.68	(3.64)	0.67	(3.6)	0.67	(3.69)
7	0.70	(3.41)	0.67	(3.56)	0.74	(4.07)	0.73	(3.85)
8	0.67	(3.26)	0.68	(3.64)	0.70	(3.68)	0.70	(3.67)
9	0.74	(3.67)	0.66	(3.42)	0.69	(3.47)	0.68	(3.39)
10	0.72	(3.27)	0.64	(3.35)	0.65	(3.23)	0.65	(3.22)
10-1	0.13	(1.4)	0.12	(1.45)	0.15	(1.64)	0.16	(1.69)

Table 2.27: Fama-MacBeth Regression Using Int-to-Market Ratios
This table presents the results of Fama and MacBeth (1973) regression analyses of the relation between expected stock returns and BM^{FF}, intangible capital over market value (IM) ratios using the sample period from June 1976 to December 2019. Each column in the table presents results for a different cross-sectional regression specification. The dependent variable in all specifications is the one-month-ahead excess stock return. The independent variables are indicated in the first column. Independent variables are winsorized at the 0.5% level on a monthly basis. The table presents average slope and intercept coefficients along with t-statistics (in parentheses), adjusted following Newey and West (1987) using six lags, testing the null hypothesis that the average coefficient is equal to zero. The rows labeled Adj. R^2 and n present the average adjusted R-squared and the number of data points, respectively, for the cross-sectional regressions.

	(1)	(2)	(3)	(4)	(5)
$logbm^{FF}$	0.27				0.22
	(3.98)				(3.48)
$logim^{Eisfeldt}$		0.19			-0.22
		(4.78)			(-2.49)
$logim^{Peters}$			0.21		-0.04
			(5.27)		(-0.33)
$logim^{Ewens}$				0.22	0.43
				(5.52)	(2.83)
β	0.00	-0.03	-0.04	-0.05	-0.04
	(0.02)	(-0.14)	(-0.23)	(-0.25)	(-0.24)
Size	-0.09	-0.05	-0.05	-0.04	-0.05
	(-1.49)	(-0.92)	(-0.87)	(-0.75)	(-0.93)
Intercept	1.35	1.08	1.30	1.32	1.69
	(3.27)	(2.72)	(3.19)	(3.22)	(3.72)
$Adj.R^2$	0.03	0.03	0.03	0.03	0.03
n	3851	3774	3968	3967	3213

Table 2.28: Spanning Tests for the Four Versions of HML Factor
This table presents the spanning tests for the four versions of the HML factor constructed using bm^{FF}, $bm^{Eisfeldt}$, bm^{Peters} and bm^{Ewens} respectively. The dependent variable is indicated by the column and the non-constant dependent variables are the other four factors from Fama and French (2015). The sample period is from June 1976 to December 2019. p-values are shown in parentheses.

	HML^{FF}	$HML^{Eisfeldt}$	HML^{Peters}	HML^{Ewens}
SMB	-0.05	0.19	0.12	0.11
	(0.406)	(0.0)	(0.004)	(0.009)
RMW	0.23	0.36	0.15	0.11
	(0.063)	(0.0)	(0.08)	(0.155)
CMA	0.99	0.86	0.92	0.91
	(0.0)	(0.0)	(0.0)	(0.0)
mkt	-0.02	0.07	0.04	0.03
	(0.659)	(0.054)	(0.228)	(0.384)
Intercept	-0.10	0.04	0.04	0.06
	(0.431)	(0.705)	(0.67)	(0.574)
Adj. R^2	0.53	0.50	0.50	0.48

Table 2.29: Spanning Tests for Pre-1999 and Post-1999 Periods
This table presents the spanning tests for the four versions of the HML factor constructed using bm^{FF}, $bm^{Eisfeldt}$, bm^{Peters} and bm^{Ewens} respectively. The sample periods are from June 1976 to December 1999 (Pre 1999) and from January 1999 to December 2019 (Post 1999). The dependent variables are indicated by the column and the non-constant dependent variables are the other four factors from Fama and French (2015). p-values are shown in parentheses.

	Pre 1999				Post 1999			
	HML^{FF}	$HML^{Eisfeldt}$	HML^{Peters}	HML^{Ewens}	HML^{FF}	$HML^{Eisfeldt}$	HML^{Peters}	HML^{Ewens}
SMB	-0.10	0.19	0.10	0.08	0.03	0.23	0.18	0.17
	(0.021)	(0.0)	(0.014)	(0.043)	(0.659)	(0.001)	(0.001)	(0.003)
RMW	-0.32	0.12	-0.15	-0.15	0.55	0.55	0.36	0.32
	(0.0)	(0.219)	(0.085)	(0.083)	(0.0)	(0.0)	(0.0)	(0.0)
CMA	0.89	0.90	0.89	0.90	0.83	0.73	0.79	0.78
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
mkt	-0.13	0.01	-0.03	-0.04	0.15	0.18	0.17	0.16
	(0.0)	(0.791)	(0.221)	(0.11)	(0.023)	(0.0)	(0.0)	(0.003)
Intercept	0.27	0.17	0.26	0.27	-0.34	-0.07	-0.12	-0.12
	(0.032)	(0.205)	(0.034)	(0.024)	(0.042)	(0.586)	(0.345)	(0.372)
Adj. R^2	0.68	0.44	0.57	0.57	0.57	0.59	0.54	0.49

Table 2.30: Spanning Tests within the High Tech Sector and the Traditional Sector
This table presents the spanning tests for the four versions of the HML factor constructed using bm^{FF}, $bm^{Eisfeldt}$, bm^{Peters} and bm^{Ewens} respectively within the high-tech sector and the traditional sector. The sample period is from June 1976 to December 2019. The dependent variables are indicated by the column and the non-constant dependent variables are the other four factors from Fama and French (2015). p-values are shown in parentheses.

	High tech sector				Traditioanl sector			
	HML^{FF}	$HML^{Eisfeldt}$	HML^{Peters}	HML^{Ewens}	HML^{FF}	$HML^{Eisfeldt}$	HML^{Peters}	HML^{Ewens}
SMB	0.02	0.15	0.20	0.22	-0.05	0.19	0.12	0.09
	(0.753)	(0.037)	(0.007)	(0.005)	(0.337)	(0.0)	(0.015)	(0.051)
RMW	-0.03	0.13	-0.04	0.03	-0.09	0.08	-0.04	-0.07
	(0.744)	(0.057)	(0.629)	(0.737)	(0.325)	(0.368)	(0.72)	(0.437)
CMA	0.76	0.78	0.81	0.79	0.73	0.60	0.71	0.74
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
mkt	-0.03	0.06	0.03	0.06	-0.05	0.04	0.04	0.03
	(0.506)	(0.113)	(0.421)	(0.14)	(0.211)	(0.225)	(0.284)	(0.498)
Intercept	0.18	0.23	0.29	0.17	0.09	0.14	0.10	0.10
	(0.261)	(0.138)	(0.058)	(0.291)	(0.411)	(0.173)	(0.377)	(0.417)
Adj. R^2	0.21	0.19	0.23	0.20	0.39	0.28	0.31	0.32

Table 2.31: Spanning Tests for IHML Factors
This table presents the spanning tests for the HML factors constructed using bm^{FF}, $im^{Eisfeldt}$, im^{Peters}, and im^{Ewens} respectively. The dependent variable is indicated by the column and the non-constant dependent variables are the other four factors from Fama and French (2015). The sample period is from June 1976 to December 2019. p-values are shown in parentheses.

	$IHML^{FF}$	$IHML^{Eisfeldt}$	$IHML^{Peters}$	$IHML^{Ewens}$
SMB	-0.05	0.20	0.24	0.22
	(0.406)	(0.0)	(0.0)	(0.0)
RMW	0.23	0.31	-0.02	-0.11
	(0.063)	(0.0)	(0.765)	(0.138)
CMA	0.99	0.76	0.65	0.53
	(0.0)	(0.0)	(0.0)	(0.0)
mkt	-0.02	0.05	0.03	0.03
	(0.659)	(0.082)	(0.174)	(0.292)
Intercept	-0.10	0.15	0.31	0.40
	(0.431)	(0.144)	(0.001)	(0.0)
$Adj.R^2$	0.53	0.42	0.37	0.29

Chapter 3

What Drives Short Duration Premium

3.1 Introduction

Equity duration, defined as the present value weighted average maturity of the equity cash flow, measures the discount rate risk. In the fixed-income market, securities with a longer maturity usually exhibit a larger yield to maturity, exhibiting an increasing yield curve. However, in the stock market, a firm's life is assumed to be infinite, and we do not have a well-defined concept of maturity. Dechow, Sloan, and Soliman (2004), borrowing the idea of Macaulay duration from the fixed-income market, propose a new measure for stock risk, the equity duration. Unlike the increasing term structure in the fixed-income market, many papers document a decreasing term structure for equity (see Van Binsbergen and Koijen (2017) for a summary) and suggest a negative relation between equity duration and future stock returns–the so-called short duration premium (see, e.g., Lettau and Wachter (2007), Weber (2018)).

This paper explores the short duration premium, trying to answer the following questions: Are stocks with high duration empirically more sensitive to discount rate (DR) news? Is the duration premium just a reflection of the value premium, as the duration and the book-to-market ratio are negatively related? Is this anomaly risk-based or behavior-based? We contribute to the literature by applying the Campbell-

Shiller present value decomposition to look at the influence of DR news and cash flow (CF) news on the returns directly, showing that all portfolios with different durations are mainly driven by CF news. In most cases, the variance of CF news over the variance of unexpected return news is larger than that of DR news. We also provide new evidence that the stable duration premium subsumes the predictive power of the traditional book-to-market (BM) ratio but not the intangible-adjusted BM. The coefficient associated with the traditional BM in the Fama-MacBeth regression is no longer significant when equity duration is included as an independent variable. In contrast, the coefficient of intangible adjusted BM still positively predicts future stock returns. Using the short interest rate over institutional ownership (SIIO) as a proxy for short constraints, we find a result consistent with Weber (2018) that the duration premium is larger and more significant for stocks more likely to be constrained, and this premium is concentrated on small stocks.

Intuitively, a shorter duration implies that the stock's most cash flows come early; thus, the stock price is not sensitive to discount rate news. We first examine this intuition using the Campbell-Shiller decomposition. This approach enables us to study the contribution of cash flow news and discount rate news directly to the unexpected return news of stocks with different durations. Theoretically, as duration is the measure of price sensitivity to the interest rate change, we expect that the discount rate news becomes a more critical driver of unexpected returns as stocks' duration increases. Besides, we also assume that for portfolios composed of stocks with high (low) duration, the contribution of DR news should be more (less) significant than that of CF news. To test these arguments, we first estimate the equity duration following Santa-Clara (2004). Unlike fixed-income securities, the stock is (theoretically) long-lived with infinite life, and we do not have fixed cash flows in the future as the dividends that firms pay are highly uncertain. Santa-Clara (2004) tackle these two problems by assuming a level perpetuity for the terminal cash flow stream after T periods and forecasting future cash flows based on financial performance measures. Then we rank stocks in ascending order of duration and group them into ten portfolios composed of the same number of stocks. Since Campbell-Shiller decomposition requires dividend information, we aggregate firm-level information to portfolio-level and analyze each portfolio. The logic of aggregation is identical to that of forming market-level information. The results indicate that cash flow news is the primary driver of returns across different portfolios, as its contribution is larger than that of discount rate news. This is quite surprising. But, as argued in Dechow, Erhard, Sloan, and Soliman (2021), the duration can also measure the sensitivity to unexpected macroeconomic events that mainly impact short-term cash flows. One possible explanation for the significant contribution of cash flow news is that investors are short-sighted and care more about near-term cash flows. Another potential interpretation is that there are more events affecting short-run cash flows than events influencing long-run discount rates. Notice that we follow the vast literature by measuring the contribution of CF news and DR news by their variance over the variance of unexpected returns, as it is difficult to disentangle the covariance of CF news and DR news.

Next, we focus on the short duration premium and its relationship with the value premium. As shown theoretically and empirically in Dechow, Sloan, and Soliman (2004), the book-to-market ratio can work as a proxy for equity duration and they are negatively related. Some papers (see Schröder, David and Esterer, Florian (2016), Weber (2018), Gormsen and Lazarus (2023)) attribute cash-flow relevant anomalies like the value factor and profitability factor to duration. However, the value strategy's performance has deteriorated in recent decades. Does it imply that the duration premium is also decreasing or not significant in the later sample period? We explore the duration premium in further detail using sub-sample analysis and compare its power to predict future stock returns with different book-to-market ratios. The results demonstrate that the short-long value-weighted duration premium is relatively stable across time, and the premium is larger within stocks with a higher possibility of being short-selling constrained. We run the Fama-MacBeth regression with equity duration and each of the traditional or intangible-adjusted BM ratios. The results show that the equity duration subsumes the predictive power of the conventional BM. However, the intangible adjusted book-to-market ratio still strongly predicts future stock returns even when the equity duration is included, as the coefficient of the intangible adjusted book-to-market ratio is significant. Chen (2017) also suggests that duration alone is unlikely to explain the value premium by looking at the future cash flow growth rates. Thus, we challenge the duration-based explanation for value strategy as Chen (2017) does, but from a different perspective.

Finally, we would like to explore the explanation for the short duration premium. Specifically, we investigate if one of the market frictions, the short-sale constraint which prevents investors from freely selling stocks, leads to abnormal returns. The price of any security is fairly valued in a complete and perfect market, and no investors can earn excess returns. But in reality, many situations like expensive lending fees, excess demand for borrowing stocks, and short selling among others hinder investors from freely holding a short position. If this constraint causes the short duration premium, we expect to see a significant and large premium for stocks more likely to be short-sale constrained. Weber (2018) uses the residual institutional ownership (RIOR) as a proxy for short-sale constraint and corroborates this source of mispricing. However, the RIOR only represents the constraint from the supply side. Instead, we also use the short interest rate over institutional ownership, which proxies for the relative demand over supply, to stand for the possibility of being short-selling constrained. Consistent with Weber (2018), we find the duration premium is larger and more significant for stocks more likely to be constrained. Stock size also plays a role. Therefore, we tend to reject the explanation proposed by Gonçalves, Andrei (2021). Gonçalves, Andrei (2021) leverages the intertemporal capital asset pricing model (ICAPM) to argue that the short duration premium is compensation for exposure to reinvestment risk (the wealth change caused by the discount rate). If this rationale is correct, we should anticipate the existence of the short duration premium across stocks with different possibilities of being short-sale constrained because the constraint does not affect the reinvestment risk.

The rest of this chapter is organized as follows: Section 2 provides a literature review on duration premium, present value decomposition, and short-sale constraints. Section 3 records the data sources, the variable definitions, and the procedure of Campbell-Shiller decomposition adapted to our paper. Section 4 displays the results of variance decomposition. Section 5 studies the duration premium in sub-samples and compares its predictive power with book-to-market ratios. Section 6 provides double-sorting analyses exploring the reasons for the duration premium. Section 7 concludes.

3.2 Literature review

Stock return predictability has long been a focal point of finance. Many characteristics, including size, book-to-market ratio, liquidity, momentum, etc., have been proposed to be predictors of future stock returns in the cross-section, leading to the so-called 'factor zoo' (see De Nard and Zhao (2022), Chen and Zimmermann (2022) for example). In the time series, it has been demonstrated that the term spread, dividend yield, consumption growth, and aggregate short interest rate, among others, have a strong predictive ability for future market returns (see Koijen and Nieuwerburgh

(2011), Cooper and Gulen (2006)). This chapter focuses on the role of equity duration in cross-sectional stock market predictability but also relates to the methodology and variables used in the time-series analysis. To be more specific, we build our work on the following three branches: duration premium, present value decomposition, and short-sale constraints.

3.2.1 Duration premium

Duration, a concept from fixed-income markets, is the weighted average of the time until those fixed cash flows are received. It measures the price sensitivity to a change in interest rates. Dechow, Sloan, and Soliman (2004) introduce and adapt this definition to equity markets. They describe the procedures to estimate the duration of a single stock using financial information. Perceived as a measure of risk, the stock duration is shown to be positively associated with price volatility and beta but negatively related to future stock returns. Besides, the duration measure subsumes the book-to-market factor in stock returns and exhibits a downward-sloping equity yield. van Binsbergen, Brandt, and Koijen (2012), on the other hand, recover the prices of dividend strips associated with the index and argue that the term structure of equity is indeed downward sloping.

Some papers propose new estimation methods for the duration. Da (2009), borrowing the idea of the consumption-based asset pricing model, measures the stocks' exposure to risk using fundamental cash flow characteristics instead of returns and prices. The duration the author proposes is entirely cash-flow based, which differs from the price-based Macaulay duration. He also points out that duration affects risk premia via its interaction with cash flow covariance between long-run accounting returns and long-run consumption innovations. Chen (2011) considers the possibility of bankruptcy when calculating the duration and attributes the stronger book-to-market effect among small stocks to their shorter equity duration. Schröder, David and Esterer, Florian (2016) use analysts' forecasts and the implied cost of capital to estimate the duration and provide a risk-based explanation for the value premium as compensation for the value firms' higher exposure to cash-flow risk. Mullins (2021) derives and summarizes three models from different theoretical underpinnings to estimate the duration and find strong co-movements among them. Gonçalves, Andrei (2021) leverages vector autoregression (VAR) to estimate the firm-level duration and shows that the short duration premium (8.6% per year in value-weighted decile portfolios) subsumes the value and profitability premia. Instead of relying on fundamental cash flow information to estimate the duration and assuming irrelevance between discount rates and expected cash flow growth as most papers do, Chen (2022) leverages the Federal Open Market Committee (FOMC) announcements as events to measure the sensitivity of stock prices to changes in discount rates, which also captures the effects of expected cash flow growth changes associated with changes in discount rates. Using this novel duration measure, the author finds a hump-shaped equity yield curve and argues that this duration captures information other than monetary policy risk.

Is the duration premium compensation for risk or due to behavioral mispricing? Weber (2018) focuses on the downward-sloping term structure. He finds that the short duration premium only exists for short-sale constrained stocks proxied by lower institutional ownership and that the premium is larger after periods of high sentiment. These results suggest both market friction and sentiment-based mispricing are the reasons for this short duration premium. In contrast, Dechow, Erhard, Sloan, and Soliman (2021) suggest a novel role for equity duration in measuring the sensitivity of prices to short-term cash flow changes rather than only measuring the sensitivity to discount rate changes. They show that short-duration stocks are more heavily influenced by the pandemic, as the pandemic mainly impacts short-term cash flow. For this reason, the underperformance is rational, as investors expect the shock to have a larger negative impact on short-duration firms than on long-duration firms whose main value comes from cash flow in the long run. The discussion above is empirically based. Gonçalves, Andrei (2021), however, manages to apply the intertemporal capital asset pricing model (ICAPM) to argue that the premium is compensation for exposure to reinvestment risk which is undesirable from the perspective of long-term investors. Therefore, stocks with a long duration, as they are more sensitive to return changes, provide a hedge when future expected returns decrease because their present value will increase, leading to large wealth effects. By contrast, short-duration stocks are exposed to this reinvestment risk and thus require a higher return to be held.

As shown in Dechow, Sloan, and Soliman (2004), the book-to-market ratio is a special case of duration by imposing assumptions on future cash flows, and they are negatively related. Consistent with that, Santa-Clara (2004) and Schröder, David and Esterer, Florian (2016), among others, demonstrate that the value factor is subsumed by duration. Gormsen and Lazarus (2023) explain several cash-flow relevant anomalies, including value, profitability, investment, low-risk, and payout factors, using duration. Their paper is promising as the data of single-stock dividend futures, which are claims on dividends of individual firms, ensures the same characteristics but changing duration of underlying assets; thus, the return difference is only caused by the duration difference while separating the effects of other characteristics. It is also encouraging as it provides an economic intuition to reduce the number of factors rather than rely on statistical analysis. Mullins (2021) also shows that the duration is closely related to the value factor, but the former outperformed a value strategy in the period following the Great Financial Crisis. However, Chen (2017) argues that the duration-based explanation alone is unlikely to resolve the value premium by directly comparing the future cash flow growth of value stocks and growth stocks. It shows that the growth rates of dividends paid by growth stocks are not substantially larger than those of value stocks. Thus, growth stocks do not have a substantially longer duration than value stocks and the average long-run growth rate difference between value and growth stocks is much smaller than assumed in duration estimation.

The only paper that links duration and present value decomposition is written by Golez and Koudijs (2020). Instead of exploring the role played by duration in equity term structure or cross-sectional differences in returns, it suggests that equity duration explains the relative predictability of returns and growth rates. They argue that the relative contribution of expected returns to stock price variation is large after 1945, as the duration of the equity market as a whole has increased substantially over time.

3.2.2 Present value decomposition

The present value identity literature establishes theoretical guidance for the predictability of returns rather than those obtained based on empirical findings which are possibly subject to the data mining problem.

Starting with the definition of return, Campbell and Shiller (1988) propose a log-linear return approximation between the relations of returns, dividend yield, and dividend growth. Based on this approximation, Campbell (1991) further shows that the unexpected return news can be decomposed into cash flow news (N_{CF}) and discount rate news (N_{DR}):

$$r_{t+1} - \mathrm{E}_t r_{t+1} = (\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=0}^{\infty}\rho^j \Delta d_{t+1+j} - (\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=1}^{\infty}\rho^j r_{t+1+j} = N_{CF,t+1} - N_{DR,t+1} \quad (3.1)$$

where r_t is the stock return, ρ is the discount rate, set to be 0.96 for annual data, Δd is the change in dividends, $(\mathrm{E}_{t+1} - \mathrm{E}_t)(X)$ means the difference between expected X at time $t+1$ and t.

Ever since then, academics have seen a growing number of papers leveraging the decomposition through VAR (see Campbell (2008), Campbell, Polk, and Vuolteenaho (2010) among others), examining the accuracy of the approximation (see Chen and Zhao (2009), Engsted, Pedersen, and Tanggaard (2012a), Engsted, Pedersen, and Tanggaard (2012b), Gao and Martin (2021)), extending the present value identity (Callen and Segal (2004), Cho, Kremens, Lee, and Polk (2022)), measuring the news component directly through data instead of computing from the VAR system (see Knox and Vissing-Jorgensen (2022), De La O and Myers (2021)).

The decomposition is important as it helps us understand what drives the price fluctuation and its underlying economic implications. For example, the CF news is regarded as a permanent shock and is a fundamental component of firm returns. In contrast, the discount rate news is viewed as a temporary shock related to the investor's risk aversion or sentiment (see Campbell, Polk, and Vuolteenaho (2010) for example).

There are also some variants of decomposition. Larrain and Yogo (2008) modify the measure of cash flow to include interest and net repurchases of equity and debt, besides the dividend, and study the present value identity between net payout and asset value. Gonçalves, Andrei (2019) relates the stock return with equity strips (i.e., dividends with different maturities) and develops a term structure return decomposition. He finds roughly 60% of equity volatility comes from the present value of dividends with maturities beyond 20 years; cash flow shocks drive volatility in short-term present values, whereas discount rate news is responsible for volatility in long-term present value. Knox and Vissing-Jorgensen (2022), on the other hand, propose a new decomposition approach for stock returns that does not rely on log-linearization or VAR estimation and can be implemented at a daily frequency using observable data. Antolin-Diaz, Petrella, and Rubio Ramírez (2021) highlight the necessity of including dividend growth as a state variable and develop methods to estimate the system from a Bayesian perspective.

The problem with Campbell (1991) decomposition is that it requires dividend data, which is missing in most periods for individual firms as many publicly listed firms do not pay dividends. Therefore, this decomposition is feasible only at the aggregate level (either portfolio level or market level). To solve this problem, Vuolteenaho (2002) starts from the definition of return and clean surplus identity for book value and enables the decomposition at the firm level by replacing dividend growth with return on equity. The findings show that cash-flow news primarily drives firm-level stock returns and that cash-flow news can be largely diversified away in aggregate portfolios. Many papers apply this approach to explore firm-level information (see Cohen, Polk, and Vuolteenaho (2003), Chaves (2009), Lochstoer and Tetlock (2020), and Cho, Kremens, Lee, and Polk (2022) for example).

3.2.3 Short-sale constraints

The short-sale constraints prevent investors from freely selling stocks and thus lead to overpricing and excess returns. Most studies use the short interest rate, institutional holdings, or loan fees on the equity lending market to proxy or estimate the short-sale constraints.

Stock exchanges measure and report the short interest rate, which is the proportion of shares sold short but has not yet covered or closed out over the total number of shares outstanding. It can be viewed as a rough proxy for short-selling demand. Asquith and Meulbroek (1995) is the first to empirically propose the relation between short interest and future stock returns at the individual stock level. Before that, as a large percentage of firms have little or no short interest in any given month, the academic could not lead to a consistent and strong connection between them. Ever since Asquith and Meulbroek (1995), many papers (see Arnold, Butler, Crack, and Zhang (2005), Boehmer, Huszar, and Jordan (2010), for example) have recorded the same pattern. Chung, Liu, and Wang (2021) confirm the predictability of the short interest rate at the industry level. At the aggregate market level, Rapach, Ringgenberg, and Zhou (2016) use the equal-weighted short interest to construct market-level short interest and show that it is arguably the strongest known predictor of aggregate stock returns both in and out of the sample. Priestley (2019), however, refutes Rapach, Ringgenberg, and Zhou (2016) by claiming that the predictability of short interest disappears once the financial crisis of 2008 is excluded from the sample. Akbas, Boehmer, Erturk, and Sorescu (2017) justify the predictability of short interest by illustrating the information content of short interest regarding future fundamental events as it is associated with negative earnings surprises, bad public news, and downgrades in analyst earnings forecasts several months ahead.

Nagel (2005) and Chen, Hong, and Stein (2002), instead use institutional ownership as a proxy for the short-selling supply and relevant estimation to proxy for the short-sale constraints. They point out that the majority of stocks have little to no outstanding short interest and that a low short interest may not necessarily indicate a less-constrained condition but instead be the result of the high transaction costs associated with short selling. Asquith, Pathak, and Ritter (2005) define short-sale constrained stocks as those in the 99th percentile of short interest ratios and the lowest third of institutional ownership. The constrained stocks underperform significantly on an equally-weighted basis but insignificantly on a value-weighted basis. Engelberg, Reed, and Ringgenberg (2018) explain the short interest premium from the point of view of short-selling risk. Specifically, they estimate the short-selling risk by regressing the variance of the daily loan fees on the equity lending market and firm characteristics. The predictability of short interest is stronger among stocks with a larger risk of short-selling. Beneish, Lee, and Nichols (2015) compute a measure of "specialness" (hard-to-borrow) that captures the extent to which short-sale constraints are binding for each firm-month observation using loan supply and demand conditions in the lending market. Moreover, they show that the abnormal returns to the short side of nine well-known market anomalies, including profits to assets, payout ratio, O-score, and so on, are attributable solely to "special" stocks.

3.3 Data and methodology

We use data from the Center for Research in Security Prices (CRSP) for stocks' prices, returns, delisting returns, shares outstanding, and so on, from Compustat for firms' financial statistics like book value of equity, goodwill, short interest, etc. We include all stocks with available data from the NYSE, Amex, and NASDAQ. Institutional ownership is constructed using Refinitiv Eikon which provides information about IBES and 13F-fillings. We obtain the aggregate predictors on Amit Goyal's website, the Fama-French factors, and the small stock value spread from French Kenneth's library.

3.3.1 Variables calculation

For fixed-income securities like a 15-year company bond or a treasury bill, the price of the contract, the payment in each period, and the maturity are all decided when designing or signing the trade contract. Therefore, the Macaulay duration for a security that pays annually is easily calculated as follows:

$$MacD = \sum_{f=1}^{n} \frac{f \times (CF_f/(1+r)^f)}{P} \tag{3.2}$$

where CF_f is the cash flow paid in period f, r is the yield to maturity, n is the number of years to maturity, P is the par value of the contract.

However, for the stock, we only know the price (P) and, under the best case, dividends (CF_f) for up to future several years if the firm initiates a dividend payment and strictly executes it. More generally, the future cash flows of stock are uncertain as the dividend payment is subject to change, including initiation, increase, decrease, or even omission. The maturity is, therefore, also affected, in addition to the fact that the stock might be delisted. But usually, we assume that the stock has an infinite life, though Chen (2011) also considers the possibility of bankruptcy when estimating the stock duration.

The estimation of stock duration in this chapter follows Dechow, Sloan, and Soliman (2004) closely and relies only on Compustat items. For each firm-year observation, we estimate future cash flows as follows:

$$CF_t = Earnings_t - \Delta BE_t = BE_{t-1}(ROE_t - g_t) \tag{3.3}$$

for $t = 1, 2, ..., 10$, where $Earnings_t$ is the earnings, ΔBE_t is the change in the book value of equity, the ROE_t is the predicted value from running AR(1) process with a long-run mean of 0.12 and persistence coefficient of 0.57, and g_t is the predicted value from running AR(1) with a long run mean of 0.06 and persistence coefficient of 0.24. ROE_0 is the current year's income before extraordinary items (IB) divided by last year's book value of equity (CEQ). g_0 is the current year's sales (SALE) divided by last year's sales.

Then the duration is calculated as

$$Dur = \frac{\sum_{t=1}^{T} t \times CF_t/(1+r)^t}{ME} + (T + \frac{1+r}{r}) \times (1 - \frac{\sum_{t=1}^{T} CF_t/(1+r)^t}{ME}) \tag{3.4}$$

where $T = 10$, $r = 0.12$, ME is the market capitalization calculated as the product of common shares outstanding (CSHO) and price close (PRCC).

The short interest rate (SIR) is the short interest divided by the common shares outstanding. The calculation of institutional ownership (IOR) follows Chen, Hong, and Stein (2002), and is the total shares owned by institutions divided by common shares outstanding. The book-to-market ratios, with or without intangible capital adjustment, are summarized in Chapter 2. Instead of directly using SIR or IOR, we also calculate the ratio of SIR over IOR (SIIO) and the residual institutional ownership in Nagel (2005) to proxy for short-selling constraints. When combining Compustat and CRSP, we ensure that the financial information or sorting variables are available when forming the portfolios.

Our final sample consists of 204,285 annual firm-year observations from 1965 to 2020, and 2,346,298 firm-month observations from July 1965 to June 2020. When the short interest rate and the institutional ownership are used, the sample usually starts from 1980 to ensure data availability. Table 3.1 displays the mean, standard deviation, minimum, 25, 50, 75 percentiles, and max for equity duration, book-to-market ratios, annualized returns, and S&P index value-weighted returns. We winsorize equity duration at 1% in both tails because it contains extreme values that affect the standard deviation heavily. Before doing that, we carefully check that these extreme values are not caused by miscalculations but attribute to a significant change in financial information. For example, the company Mr. Cooper, with gvkey 13888, had a common equity of 93,592 thousand dollars in 2017, and it jumped to 1,945 million dollars in 2018 due to a merger. Its stock prices soar from less than 1 dollar to more than 100 dollars, leading to an equity duration of -12458.8 that year. Regardless, the winsorization does not affect the portfolio sorting as we use deciles to form portfolios. The table shows that the average equity duration is 16.09 years, and the minimum is -17.22 years. Duration should be positive theoretically. The negative value comes from the case that the present value of future cash flows exceeds the current market value so that the last term in parentheses of equation 3.4 is negative:

$$(1 - \frac{\sum_{t=1}^{T} CF_t/(1+r)^t}{ME}) < 0$$

Dechow, Sloan, and Soliman (2004) suggest that it might be caused by underpricing.

3.3.2 Present value decomposition

The stock return decomposition starts from the standard definition of returns,

$$R_t = \frac{P_t + D_t}{P_{t-1}} \tag{3.5}$$

where R_t is the stock return at period t, P_t is the stock price at t, D_t is the dividend paid by firm at time t.

Campbell (1991) shows that, after a series of transformations (taking logs, iterating, first-order Taylor approximation, taking expectations...), we can obtain equation 3.1. It implies the following variance decomposition:

$$\mathrm{Var}(r_{t+1} - \mathrm{E}_t r_{t+1}) = \mathrm{Var}(CF_{t+1}) + \mathrm{Var}(DR_{t+1}) - 2\mathrm{Cov}(CF_{t+1}, DR_{t+1}) \tag{3.6}$$

Therefore the portions of the variance of unexpected returns that are attributed to the variance of discount rate news (contr (DR)), the variance of cash flows news (contr (CF)), and the covariance of cash flow news and discount rate news are defined respectively as:

$$\text{contr (DR)} = \frac{\text{Var}(DR)}{\text{Var}(\Delta\text{E}(r))} \tag{3.7}$$

$$\text{contr (CF)} = \frac{\text{Var}(CF)}{\text{Var}(\Delta\text{E}(r))} \tag{3.8}$$

$$\text{cocontr (DR,CF)} = \frac{\text{Cov}(DR, CF)}{\text{Var}(\Delta\text{E}(r))} \tag{3.9}$$

where $\Delta\text{E}(r)$) represents the unexpected return news ($r_{t+1} - \text{E}_t r_{t+1}$).

The variance decomposition literature mainly concentrates on the role played by the variance of DR news (contr (DR)) and CF news (contr (CF)). When contr (DR) is the largest, we usually say the (variance of) the DR news contributes more to the unexpected return news, or the DR news predominates over returns.

To recover the CF news and DR news, we use the following vector autoregression (VAR) to obtain the infinite sum terms:

$$Z_t = AZ_{t-1} + \varepsilon_t \tag{3.10}$$

with $Z_t = [r_t, \Delta d_t, pd_t, ty_t, dfy_t, valuespread_t]$, where r_t is return, Δd_t is the dividends, pd_t is market capitalization over dividends. The other variables, *ty*, *dfy*, are term yield and default yield constructed using predictors from Amit Goyal's website, and *valuespread* is the small stocks' value spread constructed using factors information from French's library. They are most often controlled state variables in previous papers studying variance decomposition.

Define $e1' = [1, 0, ..., 0]$, then the unexpected return news is $e1'\varepsilon$, and the DR news is $e1'\rho A(I - \rho A)^{-1}\varepsilon$. The CF news can be directly calculated or backed out as $e1'\varepsilon$+ DR news. We use the latter approach to ensure the total contribution sums up to one.

Since firms may or may not pay dividends, we solve this problem by focusing on portfolio-level decomposition. Firms are grouped into ten portfolios each year based on their equity duration. Within each portfolio, we get the return as the value-weighted return, the dividends as the sum of all dividends paid, and the market capitalization as the sum of all firms' market capitalization. Then the Campbell-Shiller decomposition is applied to each portfolio. We run the time-series VAR, calculate the news, and compute the news contribution for each portfolio.

3.4 Variance decomposition

In this section, we use Campbell-Shiller variance decomposition to check if the contribution of discount rate news to unexpected returns is higher within stocks with larger duration, and if the contribution of cash flow news is higher within stocks with shorter duration. As duration increases, stocks are more sensitive to yield change, so we expect that the contribution of DR news increases as stocks' duration increases while the contribution of CF news decreases. More than that, we should see the contribution of DR news is larger than that of CF news among stocks with longer duration.

Table 3.2 presents the contribution of discount rate news (varDR) and cash flow news (varCF) using returns and other variables contemporary with duration. At the end of June in year y, we have equity duration available; the returns of each portfolio are the value-weighted cumulative returns of each stock in that portfolio from July in year $y-1$ to June in year y. It is not the usual way of constructing portfolios where we use information in June to form portfolios and hold them for the next 12 months. The justification for doing so is that the duration may change significantly in the next year.

From Portfolio 1 to Portfolio 10, the duration increases. Column 'Full sample' contains all the firms in our sample. Column '*sir* >= *med*(*sir*)' limits the sample to firms with sir larger than or equal to the contemporary median of sir and firms with missing sir. We include firms with missing sir in this subsample because high sir implies high short-selling demand and missing sir means that even if there is a demand, it is impossible to short-sell the stocks for various reasons. Nevertheless, excluding firms with a missing short interest rate does not affect the conclusion significantly. The sample is confined to firms with fewer short-sale constraints in column '*sir* < *med*(*sir*)'. The same logic applies to column '*ior* >= *med*(*ior*)' where we only include firms with ior larger than or equal to contemporary median ior, and column '*ior* < *med*(*ior*)' where firms with smaller ior or missing ior are contained. We see that CF news is the main driver for the unexpected return news as the contribution of CF news is larger than that of DR news in most cases. Take portfolio 1, full sample, for example, the contribution of discount rate news is 56%, and the contribution of cash flow news is 106%. Notice that the contribution of each news can be larger than 1, and it simply suggests that the covariance of CF news and DR news is positive. There are cases where the contribution of discount rate news is larger than cash flow news, but they are rare and concentrate on portfolios 5, 6, or 7. We do not observe the pattern we expect at the beginning of this part.

Table 3.3 displays the results using the next period's returns and other variables. It accords with the common way of forming a portfolio to be held during the next 12 months. We see that the comments on Table 3.2 still apply.

The observation that CF news is the main driver at the portfolio level contradicts the fact that DR news is the main driver at the aggregate level, as documented by many papers. Table 3.4 reports the Campbell-Shiller decomposition at the market level across different periods. The contribution of discount rate news is larger than that of cash flow news. And the variances of computed DR news and CF news are much larger than the variance of unexpected return news during the sample period from 1965 to 1993.

We can make two conclusions. First, there is no trend of increasing (decreasing) contribution of discount rate (cash flow) news to unexpected returns, which is quite surprising, as we expect stocks with long duration to be sensitive to discount rate changes. Second, in most cases, the cash flow news is the main driver of returns across different portfolios, though discount rate news remains most important at the aggregate market level.

3.5 Superiority over (traditional) value strategy

Dechow, Sloan, and Soliman (2004), Schröder, David and Esterer, Florian (2016), Gormsen and Lazarus (2023), among others, support the view that the duration premium subsumes the value premium. We also know that the value strategy's performance has deteriorated in recent decades. Since the book-to-market ratio is a crude measure for the duration, does it indicate a smaller duration premium as well in recent decades? However, except Dechow, Erhard, Sloan, and Soliman (2021) and Weber (2018) which check the duration premium during the 2008-2009 crisis and the Covid pandemic, no other papers check the duration premium in the later sample period. It might be caused by the fact that the duration is less influential than the book-to-market ratio. Therefore, in this part, we explore the premium in further detail using sub-sample analysis and compare its power to predict future stock returns with different book-to-market ratios.

The returns of the portfolios sorted by duration during different time periods are shown in Table 3.5. We use 1999 as the division year because the number of listed firms peaked in 1997 and decreased rapidly afterward, achieving a steady number of around 3,800. We see that the long-short value-weighted return is -0.1050 for the full sample period, -0.1050, and -0.1049, respectively, for periods from 1965 to 1998 and from 1999 to 2020. Though the premium is only marginally significant in the later sample period, the magnitude of the value-weighted returns remains similar. The premium is much smaller in the later sample period for the equal-weighted portfolios. At least for long-short value-weighted returns, the short duration premium is more stable and robust than the value premium (Recall that the value premium is around 0.2 for the early sample period and around 0.04 for the later sample period). The difference between value-weighted and equal-weighted premia suggests that a size effect plays a role in the later period.

Table 3.6 displays the univariate portfolio analyses for stocks with high SIR and low SIR. If short-sale constraints cause the duration premium, we would expect a larger premium (in the absolute magnitude) for stocks with a higher possibility of being short-sale constrained (highSIR). It is so though 0.0636 is not much smaller than 0.0943. The small difference might be attributed to the poor proxy for constraints using the short interest rate. Table 3.7 confirms this argument. When we combine short interest and institutional ownership, a higher SIIO implies relatively high demand over supply, and the stock is thus more likely to be short-sale constrained. The long-short duration premium for short-sale constrained stocks is around 0.23, while the premium for not constrained stocks is only around 0.03.

We regress the excess return on duration and different book-to-market ratios separately while controlling for size and short-sale proxy. Table 3.8 exhibits the Fama-MacBeth regression results. Only the traditional book-to-market is subsumed by duration as its associated coefficient is no longer significant (t value is only -0.35). For other intangible adjusted book-to-market ratios, the coefficient associated with the book-to-market ratios is larger in magnitude and more significant than the coefficient associated with duration. Column (5) exhibits confounding figures. However, as noted by Cochrane (2011), running multiple panel-data forecasting regressions is full of pitfalls, of course. Besides, the different book-to-market ratios are highly related to each other, raising more statistical problems. There are two possible explanations for why duration cannot succeed in replacing intangible adjusted book-to-market ratio. One is that the calculation of duration involves the book value of equity. As we follow Dechow, Sloan, and Soliman (2004), the intangible capital is missing from the book value of equity, leading to a mismeasurement in the equity duration. Another explanation is that the intangible adjusted book-to-market ratio contains extra information to predict future returns.

To conclude, the magnitude of the short-long value-weighted duration premium is relatively stable even in the later sample period when the value premium is shown to be not significantly different from zero. The premium is larger within stocks with a higher possibility of short-selling constraints. However, the difference between value-weighted and equal-weighted in the later sample period suggests that the size may play a role. This phenomenon diverges from the recent underperformance of value stocks recorded in Chapter 2, suggesting that the duration premium is more robust than the value premium. If we emphasize the power to predict future returns, then the duration can only subsume the traditional book-to-market ratio (bm_{FF}), which is consistent with previous studies. But it fails to make the intangible adjusted book-to-market ratios redundant.

3.6 Duration and short-sale constraints

What explains the duration premium? Is it compensation for exposure to risk or behavioral mispricing? The cross-sectional sub-sample analyses in the section above suggest that the short duration premium is larger for stocks with high SIIO, and the market capitalization may affect it. We explore these two factors, the short-sale constraint and size, in more detail in this part.

The short interest rate is a proxy for short-sale demand. But a higher short interest rate does not necessarily implies a more constrained situation. The short interest rate can be high because there is a large and active lending market for it, and thus it may instead imply that this stock is not constrained at all. Table 3.6 and Table 3.7 confirm this argument and suggest that the relative demand and supply be a better proxy for the possibility of short-sale constraints. When SIIO is larger, the excess demand is more likely, and so is the constraint. Nagel (2005) and Weber (2018) also employ residual institutional ownership (RIOR) as the proxy for the constraints, so we present double-sorting results using these two variables.

Table 3.9 replicates Table 10 in Weber (2018), displaying the independent sorting based on RIOR and duration. Recall that from RIOR1 to RIOR5, the RIOR increases as the stocks are less likely to be short-sale constrained. We would at least expect that the duration premium is large and significant within constrained stocks (in portfolio RIOR1). Or more than that, the magnitude of the short duration premium may decrease in this direction.

In panel A of Table 3.9, the value-weighted short duration premium is 0.0197 for the most constrained group (RIOR1), but it is not significant. The duration premium is 0.0197, 0.0502, 0.0099, and 0.0552, respectively, for portfolios with higher residual institutional ownership. There is no decreasing trend in the duration premium. What surprises us is the non-significance of 0.0197 in the first row. RIOR only reflects the possibility of being constrained, so stocks in RIOR2 to RIOR5 might not be constrained at all. Therefore, no apparent trend is not disappointing. However, for stocks most prone to be mispriced due to the incomplete market, the t value is only -0.6359, far from the critical value to be significantly different from zero. The only significant duration premium appears in the portfolio with the highest RIOR. Panel B of Table 3.9 exhibits the equal-weighted returns. Here, we see a decreasing trend in the magnitude of duration premium. For the unconstrained portfolio, we even see a long duration premium, though it is not significant. The duration premium is large and significant within stocks divided into the lowest RIOR group, 0.1262, with the t statistics of 4.2952. The long-short duration return is also significant for RIOR2 and RIOR3 but not for stocks with higher RIOR. Compared with results from Weber (2018) in which only the equal-weighted method is used, we have a consistent conclusion.

Table 3.10 reports the returns of double sorting based on SIIO and duration. Since from SIIO1 to SIIO5, the relative demand over supply increases, stocks are more likely to be short-sale constrained. We would at least expect that the duration premium is large and significant within stocks classified into portfolio SIIO5. Or more than that, the magnitude of the short duration premium may increase in this direction. Panel A says that no long-short duration return is significantly different from 0. However, in panel B, we do observe that the duration premium is large and only significant for stocks with large SIIO (0.0691 with a t-statistics of 2.8576).

Table 3.5, Table 3.9, and Table 3.10 all present the divergence between value-weighted and equal-weighted duration premia. Hence it is natural to do independent sorts conditional on stock size. Table 3.11 reports 3×3 independent sorting using duration and SIIO, conditional on whether the market capitalization of the stocks is not less than its cross-sectional median (Large stocks) or less than the median (Small stocks).

All the duration premia are insignificant if the analysis is based on the sub-sample of large stocks. For the sub-sample composed of small stocks, we see that the absolute magnitude of duration premium increases as stocks are more short-sale constrained and that the premium becomes significant only under the highest SIIO portfolio when using the value-weighted method. Therefore, we suggest that the duration premium is mainly concentrated in small and short-sale-constrained stocks.

3.7 Conclusion

In this chapter, we apply the Campbell-Shiller decomposition to study the sources of unexpected returns for portfolios with different durations. The present value decomposition demonstrates that the cash flow news is the main driver of unexpected returns across portfolios with different durations, and there is no monotonic increase or decrease in the contribution cash flow news or discount rate news as duration increases. This pattern is contrary to our intuition that the variation of unexpected returns of a portfolio with a long duration should be attributed more to discount rate news, as a long duration implies a higher sensitivity to discount rate news. The result is brand new in the literature as it directly examines the role of both cash flow and discount rate news in the context of equity duration. Possible explanations for the significant contribution of cash flow news can be due to either the shortsightedness of investors who care more about near-term cash flows or the more frequent events that impact short-run cash flows than events that influence the long-run discount rate.

The downward-sloping equity term structure is puzzling. Stocks with a long duration generate smaller returns than stocks with a short duration. Is this negative relationship between duration and future stock returns caused by the value premium? We show that unlike the value strategy, which has suffered from significant and persistent drawdowns in recent decades, the short duration premium is relatively persistent and robust. Previous studies claim that duration explains many cash flow-based anomalies. We challenge this argument by showing that the intangible adjusted book-to-market ratios still have strong predictive power for future stock returns even when duration is included in the Fama-MacBeth regression. However, the duration premium mainly comes from stocks prone to being short-sale constrained. The reinvestment risk proposed by Gonçalves, Andrei (2021) can hardly explain why the duration premium disappears among stocks that are not likely to be short-sale constrained.

3.8 Tables

Table 3.1: Summary Statistics
This table presents the summary statistics for the estimated stock equity (EquityDuration), the traditional book-to-market ratio defined by Fama-French (bm_{FF}), the three intangible adjusted book-to-market ratios ($bm_{Eisfeldt}$, bm_{Peters}, bm_{Ewens}) respectively, the annual returns of stocks from July to next June (ret_y), the S&P 500 index value-weighted return in June timing twelve ($CRSP_{SPvw}$). The sample period is from 1965 to 2019. The mean (mean), standard deviation (sd), minimum(min), 25th percentile (p25), median (p50), 75th percentile (p75), and maximum (max) of each variable are calculated. The accounting variables used when calculating equity duration and book-to-market ratios for year y are of the fiscal year ending in calendar year y-1 to ensure the availability of information.

	mean	sd	min	p25	p50	p75	max
EquityDuration	16.09	5.70	-17.22	14.03	16.14	17.81	137.14
bm_{FF}	0.93	2.43	0.00	0.37	0.68	1.13	900.75
$bm_{Eisfeldt}$	2.74	5.89	0.00	0.76	1.41	2.85	977.53
bm_{Peters}	1.57	3.18	0.00	0.59	1.01	1.75	923.79
bm_{Ewens}	1.46	2.99	0.00	0.57	0.97	1.66	916.27
ret_y	0.15	0.69	-1.00	-0.21	0.06	0.34	42.52
$CRSP_{SPvw}$	0.07	0.40	-0.99	-0.19	0.05	0.36	0.83

Table 3.2: Variance Decomposition of Contemporary Returns
This table presents the results of variance decomposition for portfolios formed using equity duration. The returns and other variables in the VAR system we use are contemporary with the duration. The numbers denote the contribution of the variance of discount rate news (varDR) and cash flow news (varCF) to the total variance of unexpected returns. The columns in the first row indicate the sample we use: the full sample (Full sample), the subsample composed of stocks whose short interest rate is no smaller than its cross-sectional median or whose short interest rate is missing ($sir >= med(sir)$, the subsample composed of stocks whose short interest rate is smaller than the cross-sectional mean ($sir < med(sir)$, the subsample composed of stocks whose institutional ownership ratio is no smaller than the cross-sectional median ($ior >= med(ior)$, the subsample composed of stocks whose institutional ownership ratio is smaller than its cross-sectional median or whose institutional ownership ratio is missing. The first column shows the portfolio groups sorted by equity duration with portfolio 10 consisting of stocks whose equity duration is larger than the top decile.

	Full sample		$sir >= med(sir)$		$sir < med(sir)$		$ior >= med(ior)$		$ior < med(ior)$	
port	varDR	varCF	varDR	varCF	varDR	varCF	varDR	varCF	varDR	varCF
1	0.56	1.06	0.84	1.00	0.45	0.54	0.50	0.75	0.30	0.53
2	0.37	1.13	0.97	1.00	0.34	0.48	0.74	0.66	0.39	0.68
3	0.83	0.92	0.58	0.43	0.41	0.64	0.66	0.46	0.42	0.60
4	0.32	0.61	0.48	0.74	0.42	0.59	0.37	0.65	0.47	0.57
5	0.42	0.42	0.48	0.36	0.48	0.48	0.38	0.42	0.52	0.42
6	0.53	0.37	0.38	0.35	0.59	0.40	0.49	0.38	0.33	0.48
7	0.42	0.37	0.22	0.45	0.59	0.39	0.38	0.51	0.30	0.51
8	0.12	0.57	0.15	0.55	0.28	0.40	0.21	0.53	0.18	0.63
9	0.16	0.63	0.16	0.55	0.60	0.54	0.17	0.62	0.38	0.50
10	0.77	0.87	0.37	0.46	0.34	0.74	0.11	0.55	0.20	0.70

Table 3.3: Variance Decomposition of Future Returns
This table presents the results of variance decomposition for portfolios formed using the past year's equity duration. Every year in June, we sort stocks based on available information on equity duration into 10 groups and hold each portfolio for the next 12 months. The returns and other variables in the VAR system we use are therefore subsequent to the equity duration. The numbers denote the contribution of the variance of discount rate news (varDR) and cash flow news (varCF) to the total variance of unexpected returns. The columns in the first row indicate the sample we use: the full sample (Full sample), the subsample composed of stocks whose short interest rate is no smaller than its cross-sectional median or whose short interest rate is missing ($sir >= med(sir)$, the subsample composed of stocks whose short interest rate is smaller than the cross-sectional mean ($sir < med(sir)$, the subsample composed of stocks whose institutional ownership ratio is no smaller than the cross-sectional median ($ior >= med(ior)$, the subsample composed of stocks whose institutional ownership ratio is smaller than its cross-sectional median or whose institutional ownership ratio is missing. The first column shows the portfolio groups sorted by equity duration with portfolio 10 consisting of stocks whose equity duration is larger than the top decile.

	Full sample		$sir >= med(sir)$		$sir < med(sir)$		$ior >= med(ior)$		$ior < med(ior)$	
port	varDR	varCF	varDR	varCF	varDR	varCF	varDR	varCF	varDR	varCF
1	0.23	1.13	0.62	0.94	0.28	0.65	0.38	1.09	0.33	1.10
2	0.21	0.96	0.18	0.73	0.25	1.35	0.69	0.62	0.07	0.99
3	0.36	0.81	0.79	1.27	0.48	1.16	0.43	1.07	0.39	1.42
4	0.42	0.86	0.46	0.86	0.23	0.95	0.24	0.79	0.28	0.65
5	0.35	0.64	0.35	0.90	0.32	0.99	0.35	0.85	0.18	0.57
6	0.12	0.65	0.09	1.07	0.33	0.95	0.31	0.82	0.16	0.51
7	0.12	1.00	0.46	0.74	0.19	0.92	0.64	0.98	0.11	1.06
8	0.17	0.69	0.57	1.01	0.59	1.03	0.34	0.75	0.51	1.03
9	0.27	0.84	0.66	1.05	0.17	1.21	0.99	0.81	0.38	0.56
10	0.57	1.01	0.46	0.69	0.24	0.64	0.52	1.50	0.31	0.66

Table 3.4: Variance Decomposition of the Market Portfolio
This table presents the results of variance decomposition for the market portfolio which is the value-weighted returns. To be consistent with previous results, we define year y as from July in year y to June in year $y+1$. The numbers denote the contribution of the variance of discount rate news (varDR) and cash flow news (varCF) to the total variance of unexpected returns. The columns in the first row indicate the sample we use: the full sample from July 1965 to June 2020 (1965-2020), the first half of the sample from July 1965 to June 1993 (1965-1993), the second half of the full sample from July of 1993 to June of 2020 (1993-2020).

	1965-2020		1965-1993		1993-2020	
	varDR	varCF	varDR	varCF	varDR	varCF
contr	0.83	0.32	4.03	2.72	0.78	0.43

Table 3.5: Univariate Portfolio Analysis
This table displays the univariate portfolio analyses for the full sample (1965-2020) and for the two subsamples (1965-1998, 1999-2020) using monthly data. At the end of June each year, we sort stocks based on their equity duration, and then divide them into ten portfolios using deciles. Each portfolio is held for the next twelve months. We calculate the value-weighted (ret_{vw}) and equal-weighted (ret_{ew}) stock returns for each portfolio from 1 to 10, and also for the long-short portfolio ($10-1$). The numbers in this table are the time-series averages of these returns, except the last row in which the t statistics for the time-series of long-short portfolio returns are shown.

	1965-2020		1965-1998		1999-2020	
port	ret_{vw}	ret_{ew}	ret_{vw}	ret_{ew}	ret_{vw}	ret_{ew}
1	0.1373	0.1921	0.1715	0.2249	0.0818	0.1390
2	0.1317	0.1687	0.1616	0.1968	0.0834	0.1234
3	0.1324	0.1596	0.1606	0.1818	0.0868	0.1237
4	0.1206	0.1534	0.1431	0.1711	0.0842	0.1248
5	0.1204	0.1386	0.1409	0.1524	0.0872	0.1162
6	0.1027	0.1333	0.1213	0.1400	0.0725	0.1225
7	0.1144	0.1341	0.1302	0.1391	0.0888	0.1261
8	0.1045	0.1145	0.1176	0.1104	0.0832	0.1211
9	0.0993	0.1153	0.1098	0.0968	0.0824	0.1452
10	0.0323	0.1159	0.0665	0.1199	-0.0231	0.1095
10-1	-0.1050	-0.0762	-0.1050	-0.1049	-0.1049	-0.0296
$10-1$ (t)	5.0007	6.8415	5.2113	6.6338	1.6995	2.8460

Table 3.6:
Univariate Portfolio Analysis for Subsamples Based on Short Interest Rate
This table displays the univariate portfolio analyses for the subsamples based on the magnitude of the short interest rate using monthly data. The column HighSIR (LowSIR) indicates that the subsample is composed of stocks with short interest rate larger (smaller) than its cross-sectional mean. At the end of June each year, we rank stocks' equity duration in the subsample in ascending order, and then divide them into ten portfolios using deciles. Each portfolio is held for the next twelve months. We calculate the value-weighted (ret_{vw}) and equal-weighted (ret_{ew}) stock returns for each portfolio from 1 to 10, and also for the long-short portfolio (10 – 1). The numbers in this table are the time-series averages of these returns, except the last row in which the t statistics for the time-series of long-short portfolio returns are shown.

	HighSIR		LowSIR	
port	ret_{vw}	ret_{ew}	ret_{vw}	ret_{ew}
1	0.1156	0.0965	0.1587	0.1864
2	0.1156	0.1173	0.1431	0.1643
3	0.1245	0.1136	0.1272	0.1444
4	0.1270	0.1141	0.1199	0.1443
5	0.1370	0.1324	0.1236	0.1480
6	0.1189	0.1090	0.1214	0.1377
7	0.1159	0.1015	0.1154	0.1442
8	0.1350	0.1111	0.1381	0.1400
9	0.0946	0.0610	0.1246	0.1339
10	0.0212	-0.0032	0.0952	0.1312
10 – 1	-0.0943	-0.0997	-0.0636	-0.0552
10 – 1 (t)	3.1572	2.3521	5.1710	6.2766

Table 3.7: Univariate Portfolio Analysis for Subsamples Based on SIIO Ratio
This table displays the univariate portfolio analyses for the subsamples based on the magnitude of SIIO (the short interest rate over the ownership ratio) using monthly data. The column HighSIIO (LowSIIO) indicates that the subsample is composed of stocks with SIIO larger (smaller) than its cross-sectional mean. At the end of June each year, we rank stocks' equity duration in the subsample in ascending order, and then divide them into ten portfolios using deciles. Each portfolio is held for the next twelve months. We calculate the value-weighted (ret_{vw}) and equal-weighted (ret_{ew}) stock returns for each portfolio from 1 to 10, and also for the long-short portfolio (10 – 1). The numbers in this table are the time-series averages of these returns, except the last row in which the t statistics for the time-series of long-short portfolio returns are shown.

	HighSIIO		LowSIIO	
port	ret_{vw}	ret_{ew}	ret_{vw}	ret_{ew}
1	0.1289	0.1324	0.1371	0.1773
2	0.0472	0.0721	0.1328	0.1579
3	0.0833	0.1007	0.1389	0.1519
4	0.1536	0.1233	0.1217	0.1400
5	0.0635	0.0390	0.1272	0.1438
6	0.0224	0.0254	0.1208	0.1386
7	0.0529	-0.0098	0.1293	0.1393
8	-0.0207	0.0210	0.1323	0.1341
9	-0.0918	0.0130	0.1305	0.1326
10	-0.1118	-0.0865	0.1109	0.1380
10 – 1	-0.2472	-0.2254	-0.0262	-0.0394
10 – 1 (t)	2.0170	2.3933	4.4284	5.8757

Table 3.8: Fama-MacBeth Regression
This table presents the results of Fama and MacBeth (1973) regression analyses of the relation between expected stock returns and equity duration, book-to-market ratios using the sample period from 1975 to 2020 using monthly data. Each column in the table presents results for a different cross-sectional regression specification. The dependent variable in all specifications is the one-month-ahead excess stock return. The independent variables are indicated in the first column. Independent variables are winsorized at the 1% level on a monthly basis. The table presents average slope and intercept coefficients along with t-statistics (in parentheses), adjusted following Newey and West (1987) using six lags, testing the null hypothesis that the average coefficient is equal to zero. The row labeled R^2 shows the average R-squared of the cross-sectional regressions.

	(1)	(2)	(3)	(4)	(5)
	ret	ret	ret	ret	ret
EquityDuration	-0.0334**	-0.0275**	-0.0249*	-0.0249*	-0.0446***
	(-2.34)	(-2.02)	(-1.85)	(-1.82)	(-3.38)
bm_FF	-0.0340				-0.718***
	(-0.35)				(-3.92)
bm_Eisfeldt		0.0467**			-0.117**
		(2.09)			(-2.13)
bm_Peters			0.106**		0.0715
			(2.12)		(0.46)
bm_Ewens				0.124**	0.620***
				(2.21)	(3.38)
loglmktcap	-0.0666*	-0.0413	-0.0399	-0.0384	-0.0412
	(-1.83)	(-1.17)	(-1.16)	(-1.12)	(-1.17)
siior	-0.00179***	-0.00178***	-0.00176***	-0.00176***	-0.00185***
	(-5.54)	(-5.48)	(-5.45)	(-5.46)	(-5.78)
_cons	1.819***	1.468***	1.375***	1.352***	1.810***
	(5.45)	(4.68)	(4.35)	(4.30)	(5.38)
R^2	0.03	0.03	0.03	0.03	0.03

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.9: Bivariate Independent-Sort Portfolio Analysis—Control for RIOR
This table presents the results of bivariate independent-sort portfolio analyses of the relation between Equity duration (Dur) and future stock returns after controlling for the effect of residual institutional ownership ratio (RIOR) from 1980 to 2020. Each month, all stocks in the CRSP sample are sorted into five groups based on an ascending sort of RIOR. All stocks are independently sorted into five groups based on an ascending sort of equity duration. The quintile breakpoints used to create the groups are calculated using all stocks in the CRSP sample. The intersections of the RIOR and duration groups are used to form twenty-five portfolios. The table presents the average one-month-ahead return (in percent per month) and the associated t statistics (below) for each of the twenty-five portfolios. Also shown are the average return and t statistics of a long-short zero-cost portfolio that is long in the fifth duration (RIOR) quintile portfolio and short in the first duration (RIOR) quintile portfolio in each RIOR (duration) quintile. The t statistics are Newey and West adjusted using six lags, testing the null hypothesis that the average return equal to zero. Panel A presents results for equal-weighted portfolios. Panel B presents results for value-weighted portfolios.

Panel A: value-weighted return						
	Dur1	Dur2	Dur3	Dur4	Dur5	Dur5 – 1
RIOR1	0.1055	0.1230	0.1152	0.1349	0.0858	-0.0197
	3.3830	5.3144	4.9024	5.4422	2.4567	-0.6359
RIOR2	0.1325	0.1389	0.1329	0.1368	0.1128	-0.0197
	4.5191	5.4916	4.9667	4.9694	2.8504	-0.7127
RIOR3	0.1616	0.1507	0.1468	0.1374	0.1114	-0.0502
	5.0691	5.2840	5.1206	4.4499	2.7205	-1.9391
RIOR4	0.1493	0.1482	0.1415	0.1504	0.1394	-0.0099
	4.5253	4.9786	4.4405	4.3147	3.1790	-0.3728
RIOR5	0.1708	0.1483	0.1551	0.1508	0.1156	-0.0552
	4.7263	4.6133	4.4320	3.8647	2.4697	-2.1291
RIOR5 – 1	0.0653	0.0253	0.0398	0.0160	0.0298	
	2.4928	1.1115	1.5780	0.5700	0.9216	

Panel B: equal-weighted return						
	Dur1	Dur2	Dur3	Dur4	Dur5	Dur5 – 1
RIOR1	0.1807	0.1470	0.1303	0.1243	0.0546	-0.1262
	6.6229	7.0737	5.5751	4.6188	1.1474	-4.2952
RIOR2	0.1648	0.1472	0.1408	0.1356	0.1056	-0.0592
	6.4685	6.4987	5.5717	4.7655	2.3111	-1.9151
RIOR3	0.1740	0.1552	0.1501	0.1348	0.1304	-0.0436
	6.1669	5.8944	5.3697	4.2443	2.9729	-1.6515
RIOR4	0.1625	0.1597	0.1451	0.1492	0.1528	-0.0096
	5.3700	5.6053	4.6409	4.3245	3.3591	-0.3519
RIOR5	0.1926	0.1633	0.1652	0.1624	0.2108	0.0182
	5.5000	5.1888	4.6634	4.1607	4.1811	0.6256
RIOR5 – 1	0.0119	0.0163	0.0349	0.0381	0.1563	
	0.5628	0.9446	1.7441	1.8599	6.5871	

Table 3.10: Bivariate Independent-Sort Portfolio Analysis–Control for SIIO
This table presents the results of bivariate independent-sort portfolio analyses of the relation between Equity duration (Dur) and future stock returns after controlling for the effect of the proxy for relative short-sale demand over supply (SIIO) from 1980 to 2020. Each month, all stocks in the CRSP sample are sorted into five groups based on an ascending sort of SIIO. All stocks are independently sorted into five groups based on an ascending sort of equity duration. The quintile breakpoints used to create the groups are calculated using all stocks in the CRSP sample. The intersections of the SIIO and duration groups are used to form twenty-five portfolios. The table presents the average one-month-ahead return (in percent per month) and the associated t statistics (below) for each of the twenty-five portfolios. Also shown are the average return and t statistics of a long-short zero-cost portfolio that is long in the fifth duration (SIIO) quintile portfolio and short in the first duration (SIIO) quintile portfolio in each SIIO (duration) quintile. The t statistics are Newey and West adjusted using six lags, testing the null hypothesis that the average return is equal to zero. Panel A presents results for equal-weighted portfolios. Panel B presents results for value-weighted portfolios.

Panel A: value-weighted return						
	Dur1	Dur2	Dur3	Dur4	Dur5	Dur5 – 1
SIIO1	0.1387	0.1469	0.1249	0.1496	0.1567	0.0180
	4.9239	6.0859	4.9513	5.9250	4.8258	0.6308
SIIO2	0.1333	0.1172	0.1349	0.1266	0.1237	-0.0096
	4.4951	4.7232	5.2656	5.0631	4.3404	-0.3950
SIIO3	0.1487	0.1337	0.1083	0.1299	0.1142	-0.0345
	4.5955	4.9632	4.0963	4.8677	3.5817	-1.3673
SIIO4	0.1191	0.1346	0.1328	0.1217	0.1168	-0.0023
	3.7359	5.0072	4.6279	4.3586	3.4189	-0.0897
SIIO5	0.0819	0.1314	0.1240	0.1263	0.0892	0.0074
	2.2183	4.3959	3.6905	3.8780	2.2076	0.2484
SIIO5 – 1	-0.0568	-0.0155	-0.0009	-0.0233	-0.0675	
	-2.3806	-0.7430	-0.0377	-0.9235	-2.2354	

Panel B: equal-weighted return						
	Dur1	Dur2	Dur3	Dur4	Dur5	Dur5 – 1
SIIO1	0.1854	0.1643	0.1639	0.1658	0.1724	-0.0130
	7.2105	7.0182	6.3211	6.0874	4.6989	-0.5441
SIIO2	0.1707	0.1376	0.1539	0.1370	0.1509	-0.0198
	5.9465	5.2455	5.6237	5.1611	4.2894	-0.9658
SIIO3	0.1710	0.1396	0.1437	0.1373	0.1429	-0.0281
	5.3365	5.0520	5.0914	4.9039	3.9492	-1.4139
SIIO4	0.1617	0.1484	0.1250	0.1287	0.1360	-0.0257
	4.9968	5.3001	4.1508	4.2001	3.6158	-1.2702
SIIO5	0.1002	0.1111	0.0958	0.0934	0.0310	-0.0691
	2.5320	3.4899	2.9237	2.7271	0.7290	-2.8576
SIIO5 – 1	-0.0852	-0.0532	-0.0681	-0.0725	-0.1414	
	-3.5188	-2.9833	-3.7468	-3.7886	-6.0446	

Table 3.11: Bivariate Independent-Sort Portfolio Analysis in Subsamples
This table presents the results of subsample bivariate independent-sort portfolio analyses of the relation between Equity duration (Dur) and future stock returns after controlling for the effect of the proxy for relative short-sale demand over supply (SIIO) from 1980 to 2020. We divide stocks into two subsamples: Small stocks, and large stocks. The former contains stocks whose market capitalization is smaller than or equal to its cross-sectional median, while the latter contains stocks whose market capitalization is larger than its cross-sectional median. Each month, all stocks in the subsample are sorted into five groups based on an ascending sort of SIIO. All stocks are independently sorted into five groups based on an ascending sort of equity duration. The tertile breakpoints used to create the groups are calculated using all stocks in the subsample. The intersections of the SIIO and duration groups are used to form nine portfolios. The table presents the average one-month-ahead return (in percent per month) and the associated t statistics (below) for each of the nine portfolios. Also shown are the average return and t statistics of a long-short zero-cost portfolio that is long the fifth duration (SIIO) tertile portfolio and short the first duration (SIIO) tertile portfolio in each SIIO (duration) tertile. The t statistics are Newey and West adjusted using six lags, testing the null hypothesis that the average return is equal to zero. Panel A presents results for equal-weighted portfolios. Panel B presents results for value-weighted portfolios.

Panel A: value-weighted return								
	Small stocks				Large stocks			
	Dur1	Dur2	Dur3	Dur3 – 1	Dur1	Dur2	Dur3	Dur3 – 1
SIIO1	0.1604	0.1615	0.1303	-0.0301	0.1249	0.1306	0.1359	0.0110
	6.0324	5.5976	3.5386	-1.2465	4.9394	5.4650	5.6655	0.6811
SIIO2	0.1767	0.1238	0.1384	-0.0383	0.1354	0.1170	0.1146	-0.0207
	5.4455	3.8036	3.4804	-1.6568	4.9633	4.5575	4.2344	-1.2307
SIIO3	0.0878	0.0971	0.0281	-0.0597	0.1166	0.1270	0.1158	-0.0008
	2.0723	2.5856	0.6528	-2.2300	4.0769	4.4978	3.5894	-0.0420
SIIO3 – 1	-0.0726	-0.0645	-0.1022		-0.0083	-0.0036	-0.0201	
	-2.6260	-2.8174	-3.6872		-0.6026	-0.2347	-1.1046	
Panel B: equal-weighted return								
	Small stocks				Large stocks			
	Dur1	Dur2	Dur3	Dur3 – 1	Dur1	Dur2	Dur3	Dur3 – 1
SIIO1	0.1842	0.1612	0.1700	-0.0142	0.1567	0.1512	0.1414	-0.0153
	6.2090	5.4633	4.1711	-0.4874	6.1010	5.9621	5.4053	-1.2669
SIIO2	0.1757	0.1365	0.1607	-0.0149	0.1491	0.1376	0.1357	-0.0134
	5.4273	4.119	3.7166	-0.5957	5.2155	4.9958	4.6272	-1.0169
SIIO3	0.1040	0.1094	0.0582	-0.0459	0.1237	0.1218	0.0938	-0.0299
	2.4230	2.9489	1.2011	-1.5536	4.0249	4.0205	2.7012	-1.8198
SIIO3 – 1	-0.0802	-0.0518	-0.1118		-0.0330	-0.0295	-0.0476	
	-2.7749	-2.4194	-3.7371		-2.6862	-2.5435	-3.3124	

Chapter 4

What Drives Behavioral Factors

4.1 Introduction

We employ the present value decomposition to analyze the cash flow and discount rate components of the newly proposed behavioral factors. These factors are motivated by behavioral theory, and are claimed to outperform alternative factor models such as the Fama-French three-factor model and the q-factor model. This chapter contributes to the understanding of the economic rationale behind these factors and adds to the existing literature on the 'factor war' by adopting a fresh perspective that explores the underlying cash flow (CF) news and discount rate (DR) news of factors, instead of relying solely on GRS tests, spinning tests, or how many anomalies the factor model is able to explain. Our findings reveal that the primary driver of factors' returns and their mean-variance efficient frontier is the cash flow news, as the variance of the CF news takes up a large portion of the total return news. Moreover, when regressing anomalies' return news on factors' CF news or DR news, we observe a significantly higher adjusted R-squared value when CF news is considered as an independent variable. Additionally, the cross-sectional variation of market betas closely corresponds to the variation of beta components associated with CF news. Therefore, we highlight the importance of the fundamental CF of these factors. These results challenge the foundational behavioral models as we consistently demonstrate significant effects of cash flow news, which embodies fundamental risk rather than sentiment risk or misvaluation represented by discount rate news.

The last two decades have witnessed an explosion of anomalies. As documented by Harvey and Liu (2019), the production of factors is out of control, leading to nearly 400 factors published in top journals by January 2019. Despite the possibility of data mining, these anomalies present a challenge for the asset pricing models, as these models fail to explain the anomalies' returns. In the meantime, some new factor models have emerged, claiming superior performance due to their ability to expand the mean-variance frontier and explain a greater number of anomalies (see Chen and Zhang (2010), Fama and French (2015), Hou, Xue, and Zhang (2015), Stambaugh and Yuan (2017), Barillas and Shanken (2018)). The past dominant factor models are mainly characteristics-based, using firms' accounting information like BM, size, and others, or past return and volatility information. However, in recent years, some authors have introduced factor models based on behavioral aspects or mispricing, which have attracted much attention due to their better performance.

Given the power of these factors, it is essential to know what drives them, as this understanding not only contributes to theoretical models but also has implications for practical investment. However, as noted by Scientific Background on the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2013, "A weakness of the (Fama-French) three-factor model is that it is primarily an empirical model that describes stock returns, but it is silent on the underlying economic reasons for why these risk factors have nonzero prices." Similarly, we know little about these new factors. Do they genuinely stand for a kind of risk that requires a higher premium, or do they stand for the animal spirits driven by sentiment, over/under-reaction?

Various studies have attempted to explore these reasons, with some employing short-sale constraints as proxies.

In this chapter, we explore the CF and discount rate (DR) news components of these behavioral factors by taking the present value decomposition approach. Lochstoer and Tetlock (2020) have done similar work analyzing the five well-known anomalies—value, size, profitability, investment, and momentum—which are most frequently included in factor models. However, there are some minor differences in construction in Lochstoer and Tetlock (2020) from the original version, like using a lower updating frequency. They find that the fundamental cash flow news drives the returns of anomaly portfolios as well as their mean-variance-efficient (MVE) portfolios. Instead, we focus on the behavioral factors from the predominant behavioral factor models proposed by Stambaugh and Yuan (2017) and Daniel, Hirshleifer, and Sun (2020).

First, we run panel vector autoregression (VAR) to extract each firm's return news, CF news, and DR news, and run time-series VAR to obtain news at the aggregate market level. Despite the existence of other approaches for estimating the news (see the discussion in Khimich (2017)), we select the VAR approach to enlarge the sample size (as the requirement of analysts' coverage to proxy for market expectations when estimating news components shrinks the sample size dramatically) and maintain alignment with Lochstoer and Tetlock (2020). Next, we construct the anomalies and behavioral factors to get anomaly-level, factor-level, and MVE frontier news. Though the original papers provide factor returns, we need to reconstruct these factors in our study for two reasons: First, to aggregate firm-level news to other levels, we need to know their characteristics to group stocks into portfolios. Second, some factors, like the PEAD factor from Daniel, Hirshleifer, and Sun (2020), are rebalanced monthly, but due to the availability of accounting data, the CF news, DR news, and unexpected return news components are at an annual frequency. Therefore, the best we can do is mimic the logic of these monthly updated factors and reconstruct them at an annual frequency. This might harm the ability of these factors to explain other anomalies to some extent, but our empirical results suggest that the impact is acceptable. Then we aggregate the firm-level news to the desired level, like portfolio level or anomaly level, using the same procedure as constructing the factor returns, except that we use stock news instead of stock returns.

With news components ready, we can analyze the main driver of these returns by comparing the variance of CF and DR news over that of unexpected return news. We find that the most variance in return news falls within the variation of CF news. To explore factors' explanatory ability, we run anomalies return news on the news component of these factors to identify which news yields the largest R squared. Again, the CF news contributes the most. Finally, we decompose the market beta using the news components and find that the variation of market beta across portfolios is significantly affected by the CF news.

Our findings question the behavioral rationale under these models, as our results all broadcast the significant effects of cash flow news that represents fundamental risk instead of sentiment or misvaluation proxied by discount rate news.

Firstly, variance decomposition indicates that the variation in cash flow news accounts for the majority of the unexpected return news variation across all behavioral factors. This holds true for the firm-level news decomposition and the anomalies considered in this chapter. Even though behavioral explanations such as delayed agent reactions or managerial market timing may help interpret the abnormal returns of these factors or anomalies, it's the cash flow news that captures real fundamental risk and drives most of the return news variation.

Consider, for example, the FIN factor from Daniel, Hirshleifer, and Sun (2020) and the momentum factor from Jegadeesh and Titman (1993). Stock issuance is empirically shown to be negatively correlated with future stock returns. The FIN factor is constructed using information about stock issuance. The momentum factor reflects the fact that past medium-term (six to twelve months) winners continue to outperform in the future. Two strands of reasoning are plausible: risk-based or behavioral-based. The risk-based perspective posits that high returns compensate for the inherent high risk of the underlying stocks. Conversely, the behavioral explanation suggests potential profits due to misvaluation caused by cognitive, and psychological aspects, or limits to arbitrage due to market frictions like asymmetric information, transaction costs, etc. A high net issuance could result from a promising investment opportunity or from managers' intent to sell the overpriced stock, leading to lower future returns. However, while the former reflects fair pricing, the latter indicates an inefficient market. From the risk perspective, high momentum stocks earn high future returns because these stocks are more likely to have their surging prices plunge back to earth, and therefore investors require a risk premium—higher returns for bearing this additional risk. However, it can also be the case that investors are slow to respond to the news, and therefore it takes much longer for the price to increase to the actual intrinsic value, leading the winners to win and the losers to lose.

Secondly, we regress anomalies' unexpected return news on factors' cash flow news, discount rate news, or unexpected return news, respectively. The adjusted R squared is considerably larger when we include the cash flow news of these factors as an independent variable, compared to when we include only the discount rate news. Thus, the predictability of these factors predominantly stems from the cash flow news component. It differs from our first result. Here we use factor news to explain other anomaly news; instead of concentrating on the portion of factor CF and DR news variance over its unexpected return news variance.

Lastly, we regress the news components of the portfolios sorted and grouped by the mispricing or behavioral characteristics on the news components of the market to analyze the composition of the market beta of these groups. The market beta reveals the systematic risk that cannot be diversified away. We are interested in the variation of the beta across portfolios. Our results suggest that, while the covariance between discount rate news of the portfolio and the market constitutes a significant portion of the market beta, it's the covariance related to cash flow news that leads to market beta variation across portfolios.

In conclusion, our findings underscore the fundamental role of CF news. We contribute to the literature by applying the present value decomposition approach to the newly proposed behavioral factors from Stambaugh and Yuan (2017) and Daniel, Hirshleifer, and Sun (2020). We identify CF news as the main driver for these factors and explore the predictability source from the news perspective, contrasting previous studies that solely relied on return regression-based results examining coefficients or intercepts. We offer evidence supporting the fundamental and risk-based explanations for these factors.

The rest of this chapter is organized as follows: Section 2 provides a literature review on the factor model and return variance decomposition. Section 3 discusses the underlying theoretical framework that inspires the empirical design. Section 4 records the data sources, the variable construction, and the procedure we use to obtain news. We provide summary statistics and regression results testing the validity of these reconstructed factors as well in this part. Section 5 presents all the results from return decomposition, news regression, and beta decomposition. Section 6 concludes.

4.2 Literature review

My paper builds on two branches of literature: factor models and present value decomposition. The former motivates our study of the newly proposed behavioral anomalies, while the latter provides us with the tool to explore them from the perspective of cash flow and discount rate news.

4.2.1 Factor models

In asset pricing, the central question is what determines the asset price. The most intuitive way to look at the price is that it should be an expected discount value of tomorrow's total payoff. So determining the price falls under resolving the stochastic discount factor (SDF). In the most famous consumption-based Capital Asset Pricing Model (CCAPM), the SDF is just the discounted marginal rate of substitute on consumption.

However, since the end of the last century, both the academic and empirical worlds have recorded many anomalies and claimed the failure of some classic asset pricing models like CCAPM. Up to today, there is even an anomaly zoo in which over 400 anomalies are found. After the notion of "factor zoo" from Cochrane (2011), Harvey, Liu, and Zhu (2016), Harvey (2017), Harvey and Liu (2019) also call our caution to the finding of new anomalies. Are they true anomalies, merely the outcome of data mining, wrong statistics, or caused by other reasons? How could we manage to explain them? We have some factor models developed to rescue this situation. For example, the Fama-French three-factor model (Fama and French (1992)) uses market excess returns, HML, and size factors in their model to explain the cross-section of stock returns. The Fama-French five-factor model (Fama and French (2015)) adds profitability and investment factors to their previous three-factor model. Hou, Xue, and Zhang (2015) propose the q-factor model inspired by the investment-based asset pricing, containing market excess returns, investment, and roe factors. These factors are mainly based on the firms' accounting characteristics.

To reduce the space of anomalies to parsimonious factors or to compare among factors, there are also papers applying statistical or machine learning models like LASSO, Principle Component Analysis, similarity analysis, and proposing new Bayesian models; see Feng, Giglio, and Xiu (2020), Bryzgalova, Huang, and Julliard (2023), Kozak, Nagel, and Santosh (2020), Lettau and Pelger (2020), and Giglio, Kelly, and Xiu (2022) among others.

Diverging from rational asset pricing, some authors propose new models with factors motivated by behavior or mispricing. They claim it works better in spinning the mean-variance frontier and explaining more anomalies. Daniel, Hirshleifer, and Sun (2020) propose a financing factor (FIN) to capture the long-run mispricing based on the managers' decisions to issue or repurchase equity, and post-earning announcement drift factor (PEAD) to capture the short-run mispricing based on investors' inattention and slow response to the news, along with a market factor in their model. The FIN factor is constructed using 1-year net-share-issuance (NSI) and 5-year composite-share-issuance (CSI) to account for the managers' timing to exploit the stock's mispricing in the interest of other investors. The PEAD factor uses the earnings surprise after firms' announcements to address investors' limited attention. Stambaugh and Yuan (2017), on the other hand, include market factor, size factor, and a MISPRICING (UMO, the underpriced minus overpriced in their paper) factor which aggregates information across eleven prominent anomalies by averaging their rankings[1]. By doing so, they aim to achieve a less noisy measure of stock's mispricing.

[1]They also propose a four-factor model with two "mispricing" factors. But to save space, we only use their three-factor model.

In this chapter, instead of proposing new factors, we take a different point of view to look at these factors. When explaining the cause of these factors, papers rely on the relation of these factors' characteristics with other variables like Gormsen and Lazarus (2023), if the sentiment index predicts the factors like Stambaugh and Yuan (2017), or if the factor or underlying characteristics predict future cash flows like Chen (2017) among others. The motivation for using the present value decomposition is that, in the end, the returns are driven either by cash flow news or discount rate news. Many characteristics like ROE, investment rate, O-score, and so on contain information about either cash flow news, discount rate news, or both. We would like to see if the variance decomposition could add more value to the shrinkage of anomalies or the comparison between factor models.

4.2.2 Present value decomposition

The present-value decomposition is based on the work of Campbell and Shiller (1988). They provide a log-linear approximation between the returns, dividend yield, and dividend growth starting from the definition of returns. In order to study the sources of variation in the returns, Campbell (1991) further shows that the unexpected return news can be decomposed into cash flow news which contains information on unexpected future dividend growth, and discount rate news which contains information on unexpected future returns. He finds that the future excess returns' volatility takes up to around 70% of the total variation in the unexpected return news for the aggregate market from 1952 to 1988. However, since this decomposition requires information on dividend growth, while most firms do not pay dividends, it prevents us from applying this approach to the firm level. To solve this problem, Vuolteenaho (2002) uses the accounting identity to reach a similar return decomposition where he uses return on equity (roe) instead of dividend growth. It, therefore, enables decomposition at the firm level. Vuolteenaho (2002) finds that at the firm level, the CF news is the main driver, but when aggregating to the market level, the cash-flow news gets diversified, leaving the discount rate news as the main driver, which is consistent with the conclusion from Campbell (1991).

The present value decomposition was previously applied when studying the predictability of returns, systematic risk, and the driving elements of returns' variance. It has also seen new developments in recent years. For example, Cochrane (2008) discusses the predictability of returns and dividend growth. Maio and Xu (2020) generalize the Campbell-Shiller decomposition and study the prediction power of aggregate earnings yield. Campbell and Mei (1993), Campbell and Vuolteenaho (2004), Campbell, Polk, and Vuolteenaho (2010) use the news components to analyze the sources of market betas. Mao and Wei (2014) explain price and earnings momentum by investigating the dynamics of cash flow (CF) news and discount rate (DR) news. Chen and Zhao (2009), and Engsted, Pedersen, and Tanggaard (2012b) discuss the VAR-based decomposition. Callen and Segal (2004) extends the decomposition by adding an accruals news, while Cho, Kremens, Lee, and Polk (2022) take into account the investment in stock issuance and then present a new CF news brought by that. De La O and Myers (2021), instead of using the VAR approach to estimate the news components, employ the subjective cash flow and discount rate expectations from survey forecasts and find that cash flow growth expectations explain 93% and 63% of the variation in the S&P 500 price-dividend and price-earnings ratios. Gao and Martin (2021) exploit a measure of dividend yield to derive a new decomposition that resembles the Gordon growth model more closely and has certain other advantages. Gonçalves, Andrei (2019) relates the stock return with equity strips (i.e., dividends with different maturities) and develops a term structure return decomposition.

The decomposition is important as it helps us understand what drives the price fluctuation or even payout policy and its underlying economic implications. The cash-flow news is regarded as a permanent shock and is the fundamental component of firm returns, while the discount rate news is viewed as a temporary shock, related to the investor's risk aversion or sentiment. Michaely, Rossi, and Weber (2021) argue that the CF news drives payout policy, and payout policy conveys information about future cash-flow volatility.

While previous research mainly applies this decomposition to firm-level and aggregate market-level, Lochstoer and Tetlock (2020) study the five well-known anomalies (which form the basis factors of the traditional factor models) under this approach, and they find that CF news contributes more to the unexpected returns of all these five anomalies. We extend this literature by studying the newly proposed behavioral factors.

4.3 Theoretical settings

This part briefly exhibits the theoretical backgrounds to support our empirical approaches. The present value decomposition explains why the return news equals CF news minus DR news and thus enables the variance decomposition. The beta decomposition allows us to explore the systematic beta by looking separately at the risks brought by CF or DR news.

4.3.1 Present value decomposition

Starting from the definition of returns, Campbell and Shiller (1988) reaches a dividend-ratio model, i.e., dynamic Gordon model where they show mathematically that log dividend price ratio equals to a constant plus the sum of expected discounted value of all future one-period "growth-adjusted discount rates" ($e_{t+j} - \Delta d_{t+j}$). Build on this work, Campbell (1991) shows that unexpected return news equals CF news minus DR news as follows:

$$\begin{aligned} r_{t+1} - E_t r_{t+1} &= (E_{t+1} - E_t) \sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} + (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j r_{t+1+j} \\ &= N_{CF,t+1} - N_{DR,t+1} \end{aligned} \tag{4.1}$$

where Δd_{t+1+j} is the dividend growth in period $t+1+j$, $(E_{t+1} - E_t)x_j = E_{t+1}[x_j] - E_t[x_j]$, and r_t is the log return at period t.

The most inspiring step is the Taylor first-order expansion which enables one to write $log(A+B)$ as a linear combination of $log(A)$ and $log(B)$.

However, we now document the deduction from Vuolteenaho (2002) because this approach makes it possible to decompose returns at the firm level instead of the aggregate level by using return on equity (ROE) instead of dividend growth as the basic cash-flow fundamental (as many firms do not pay dividends and this prevents us from constructing dividend growth needed in equation (4.1) and it also sheds light on the variables included in the vector autoregression system.

By assuming zero equity issuance[2], we have the clean-surplus identity:

$$B_{t+1} = B_t + Y_{t+1} - D_{t+1} N_t \tag{4.2}$$

where B_t is the book value of equity at time t, Y_t is the total earnings at t, D_t is the dividend per share, and N_t is the number of shares.

[2]That is, if N_t is the number of shares, then we have $N_t = N_{t-1} = N_{t+1}$. This assumption ignores the future book equity investment through share issuance or repurchase. Cho, Kremens, Lee, and Polk (2022) relax this assumption by allowing the change in the issuance. They then introduce a new variable as a source of cash flow and divide the cash flow news into two components, one brought by investment and another by profitability. We do not follow this approach because it makes CF news more complex, and we consider other variables in the VAR to account for the investment.

Multiplying $\frac{N_t}{B_t}$ on both sides of the definition equation on return $P_t = \frac{1}{1+R_{t+1}}(D_{t+1} + P_{t+1})$ gives us:

$$\begin{aligned} \frac{P_t N_t}{B_t} &= \frac{1}{1+R_{t+1}}(D_{t+1}+P_{t+1})\frac{N_t}{B_t} \\ \frac{M_t}{B_t} &= \frac{1}{1+R_{t+1}}(D_{t+1}+P_{t+1})\frac{N_t}{B_t} \end{aligned} \tag{4.3}$$

as $M_t = P_t N_t$ and it is the market value of equity.

Multiplying $\frac{N_t}{N_{t+1}}$ which equals 1 to the left-hand side of equation (4.2) and solve for N_t we can obtain N_t as $N_t = \frac{B_t+Y_{t+1}}{D_{t+1}+B_{t+1}/N_{t+1}}$. Plugging it into (4.3) leads to:

$$\begin{aligned} \frac{M_t}{B_t} &= \frac{1}{1+R_{t+1}}(D_{t+1}+P_{t+1})\frac{1}{B_t}\frac{B_t+Y_{t+1}}{D_{t+1}+B_{t+1}/N_{t+1}} \\ &= \frac{1}{1+R_{t+1}}\frac{B_t+Y_{t+1}}{B_t}\frac{D_{t+1}+P_{t+1}}{D_{t+1}+B_{t+1}/N_{t+1}} \\ &= \frac{1}{1+R_{t+1}}\frac{B_t+Y_{t+1}}{B_t}\frac{1}{D_{t+1}+B_{t+1}/N_{t+1}}\left[D_{t+1}+\frac{P_{t+1}N_{t+1}B_{t+1}/N_{t+1}}{B_{t+1}}\right] \\ &= \frac{1}{1+R_{t+1}}\left(\frac{B_t+Y_{t+1}}{B_t}\right)\left(\frac{D_{t+1}}{D_{t+1}+B_{t+1}/N_{t+1}}+\frac{P_{t+1}N_{t+1}}{B_{t+1}}\times\frac{B_{t+1}/N_{t+1}}{D_{t+1}+B_{t+1}/N_{t+1}}\right) \\ &= \frac{1}{1+R_{t+1}}\left(\frac{B_t+Y_{t+1}}{B_t}\right)\left(1-\frac{B_{t+1}/N_{t+1}}{D_{t+1}+B_{t+1}/N_{t+1}}+\frac{P_{t+1}N_{t+1}}{B_{t+1}}\times\frac{B_{t+1}/N_{t+1}}{D_{t+1}+B_{t+1}/N_{t+1}}\right) \end{aligned} \tag{4.4}$$

Define $\Lambda_{t+1} = \frac{B_{t+1}/N_{t+1}}{D_{t+1}+B_{t+1}/N_{t+1}}$ which is the plowback ratio, then (4.4) becomes:

$$\begin{aligned} \frac{M_t}{B_t} &= \frac{1}{1+R_{t+1}}\times(1+ROE_{t+1})\times\left(1-\Lambda_{t+1}+\frac{M_{t+1}}{B_{t+1}}\Lambda_{t+1}\right) \\ &= \frac{1}{1+R_{t+1}}\times(1+ROE_{t+1})\times\left(1+\left(\frac{M_{t+1}}{B_{t+1}}-1\right)\Lambda_{t+1}\right) \end{aligned} \tag{4.5}$$

where $ROE_{t+1} = Y_{t+1}/B_t$ is the return on equity.

Taking log on both sides of (4.5):

$$mb_t = -r_{t+1} + roe_{t+1} + log[1+(exp(mb_{t+1})-1)exp(\lambda_{t+1})] \tag{4.6}$$

where $mb_t = log(M_t/B_t)$, $r_{t+1} = log(1+R_{t+1})$, $roe_{t+1} = log(a+ROE_{t+1})$, $\lambda_{t+1} = log(\Lambda_{t+1})$.

Apply the first-order Taylor approximation[3], we can get the log-linear present-value identity in Vuolteenaho (2002):

$$mb_t \approx -r_{t+1} + roe_{t+1} + \rho mb_{t+1} \tag{4.7}$$

where ρ is set to 0.967.

Iterating mb forward and assuming $lim_{j\to\infty}\rho^j mb_{t+j} = 0$, we can obtain:

$$\begin{aligned}
mb_t &\approx -r_{t+1} + roe_{t+1} + \rho mb_{t+1}\\
&= -r_{t+1} + roe_{t+1} + \rho(-r_{t+2} + roe_{t+2} + \rho mb_{t+2})\\
&= (-r_{t+1} - \rho r_{t+2}) + (roe_{t+1} + \rho roe_{t+1}) + \rho^2 mb_{t+2}\\
&= (-r_{t+1} - \rho r_{t+2}) + (roe_{t+1} + \rho roe_{t+2}) + \rho^2(-r_{t+3} + roe_{t+3} + \rho mb_{t+3})\\
&= (-r_{t+1} - \rho r_{t+2} + \rho^2 r_{t+3}) + (roe_{t+1} + \rho roe_{t+2} + \rho^2 roe_{t+3}) + \rho^3 mb_{t+3}\\
&= (-r_{t+1} - \rho r_{t+2} + \rho^2 r_{t+3} + \ldots) + (roe_{t+1} + \rho roe_{t+2} + \rho^2 roe_{t+3} + \ldots) + \rho^\infty mb_{t+\infty}\\
&= -\sum_{j=1}^{\infty}\rho^{j-1} r_{t+j} + \sum_{j=1}^{\infty}\rho^{j-1} roe_{t+j}
\end{aligned} \tag{4.8}$$

[3]Recall the first order Taylor approximation of $f(x,y)$ around (x_k, y_k):

$$f(x,y) \approx f(x_k,y_k) + (x-x_k)f'_x(x_k,y_k) + (y-y_k)f'_y(x_k,y_k)$$

Taking the first order Taylor approximation of $log[1+(exp(x)-1)exp(y)]$ around $x=0$ and $y = log(\rho)$ gives us $log[1+(exp(x)-1)exp(y)] \approx log[1+(exp(0)-1)exp(log(\rho))] + (x-0)\left[\frac{exp(x)exp(y)}{1+(exp(x)-1)exp(y)}\right]_{x=0,y=log(\rho)} + (y-log(\rho))\left[\frac{(exp(x)-1)exp(y)}{1+(exp(x)-1)exp(y)}\right]_{x=0,y=log(\rho)} = log(1+0) + x\left[\frac{1\times\rho}{1+0}\right] + (y-log(\rho))\times 0 = x\rho$. Use mb_{t+1} to substitute x, and we end at $log[1+(exp(mb_{t+1})-1)exp(\lambda_{t+1})] \approx mb_{t+1}\rho$.

Using (4.8), we have the following difference between the conditional expectation of bm_{t-1} at time t and t-1:

$$
\begin{aligned}
&bm_{t-1} = -mb_{t-1} \approx \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j-1} - \sum_{j=1}^{\infty} \rho^{j-1} roe_{t+j-1} = \sum_{j=0}^{\infty} \rho^{j} r_{t+j} - \sum_{j=0}^{\infty} \rho^{j} roe_{t+j} \\
&E_t\left[bm_{t-1}\right] - E_{t-1}[bm_{t-1}] = 0 = E_t\left[\sum_{j=0}^{\infty} \rho^{j} r_{t+j} - \sum_{j=0}^{\infty} \rho^{j} roe_{t+j}\right] - E_{t-1}\left[\sum_{j=0}^{\infty} \rho^{j} r_{t+j} - \sum_{j=0}^{\infty} \rho^{j} roe_{t+j}\right] \\
&0 = E_t\left[r_t + \sum_{j=1}^{\infty} \rho^{j} r_{t+j} - \sum_{j=0}^{\infty} \rho^{j} roe_{t+j}\right] - E_{t-1}\left[r_t + \sum_{j=1}^{\infty} \rho^{j} r_{t+j} - \sum_{j=0}^{\infty} \rho^{j} roe_{t+j}\right] \\
&E_t(r_t) - E_{t-1}(r_t) = (E_t - E_{t-1}) \sum_{j=0}^{\infty} \rho^{j} roe_{t+j} - (E_t - E_{t-1}) \sum_{j=1}^{\infty} \rho^{j} r_{t+j}
\end{aligned} \tag{4.9}
$$

It gives us a similar decomposition as in equation (4.1):

$$
\begin{aligned}
r_{t+1} - E_t r_{t+1} &= (E_{t+1} - E_t) \sum_{j=0}^{\infty} \rho^{j} \Delta d_{t+1+j} + (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^{j} r_{t+1+j} \\
&= N_{CF,t+1} - N_{DR,t+1}
\end{aligned} \tag{4.10}
$$

which implies that the unexpected stock return is due to two components: news about future cash flow (roe) and news about the future discount rate.

4.3.2 Beta decomposition

Though the CAPM does not price many assets well, its idea of describing stock returns as a risk measure times risk premium is influential, and we still care about the market beta as it measures the asset's risk relative to the aggregate market. Campbell and Vuolteenaho (2004) decompose the market beta into good beta and bad beta to explain the size and value anomalies. Campbell, Polk, and Vuolteenaho (2010) further decompose the bad beta and good beta into four beta components, which breaks the returns of asset and market into CF and DR news components, and then analyzes their covariance.

The market beta of asset i is:

$$\beta_{i,M} \equiv \frac{\mathrm{Cov}_t\left(r_{i,t+1}, r_{M,t+1}\right)}{\mathrm{Var}_t\left(r_{M,t+1}\right)} = \frac{\mathrm{Cov}_t\left(N_{i,t+1}, N_{M,t+1}\right)}{\mathrm{Var}_t\left(r_{M,t+1}\right)} \tag{4.11}$$

where $r_{i,t+1}$, $r_{M,t+1}$ are asset excess return and market factor return at time $t+1$ respectively, $N_{i,t+1}$ is the unexpected return news of asset i at time $t+1$ and $N_{M,t+1}$ is the unexpected return news of market factor at time $t+1$. We use the return news in place of returns because the expectation is just a constant.

As we can decompose return news as CF news minus DR news shown in the last section, breaking market return news gives us bad cash flow beta ($\beta_{i,CFM}$) and good discount rate beta ($\beta_{i,DRM}$):

$$\beta_{i,CFM} \equiv \frac{\mathrm{Cov}_t\left(N_{i,t+1}, N_{M,CF,t+1}\right)}{\mathrm{Var}_t\left(r_{M,t+1}\right)} \tag{4.12}$$

$$\beta_{i,DRM} \equiv \frac{\mathrm{Cov}_t\left(N_{i,t+1}, -N_{M,DR,t+1}\right)}{\mathrm{Var}_t\left(r_{M,t+1}\right)} \tag{4.13}$$

where $N_{M,CF,t+1}$ and $N_{M,DR,t+1}$ are the market CF news and DR news at time $t+1$

We can further decompose asset return news similarly and get the four beta components:

$$\beta_{CFi,CFM} = \frac{\mathrm{Cov}_t\left(N_{i,CF,t+1}, N_{M,CF,t+1}\right)}{\mathrm{Var}_t\left(r_{M,t+1}\right)} \tag{4.14}$$

$$\beta_{DRi,CFM} = \frac{\mathrm{Cov}_t\left(-N_{i,DR,t+1}, N_{M,CF,t+1}\right)}{\mathrm{Var}_t\left(r_{M,t+1}\right)} \tag{4.15}$$

$$\beta_{CFi,DRM} = \frac{\text{Cov}_t\left(N_{i,CF,t+1}, -N_{M,DR,t+1}\right)}{\text{Var}_t\left(r_{M,t+1}\right)} \tag{4.16}$$

$$\beta_{DRi,DRM} = \frac{\text{Cov}_t\left(-N_{i,DR,t+1}, -N_{M,DR,t+1}\right)}{\text{Var}_t\left(r_{M,t+1}\right)} \tag{4.17}$$

Through this process, we decompose the market beta into four components:

$$\beta_{i,M} = \beta_{i,CFM} + \beta_{i,DRM} = \beta_{CFi,CFM} + \beta_{DRi,CFM} + \beta_{CFi,DRM} + \beta_{DRi,DRM} \tag{4.18}$$

4.4 Empirical design

In this section, we document the data source, the construction of variables, sample filtration, and the procedure to extract news.

4.4.1 Data and factor construction

Various sources of data are employed. In summary, we get common stocks information in NYSE, Nasdaq, and AMEX, and inflation to calculate returns from CRSP; firms' accounting information to get variables used in the VAR system and some anomaly characteristics from Compustat; Fama-French three-factors, and value spread from French's data library; aggregate predictors provided by Welch and Goyal (2008); historical book equity data from Davis, Fama, and French (2000); some anomaly characteristics provided by Chen and Zimmermann (2022); and treasury yield fromLochstoer and Tetlock (2020).

For variables to be contained in the VAR, i.e., returns, return on equity (ROE), book-to-market ratio (BM), Profitability (Prof), Investment (Inv_M5), five-year change in log market equity ME_D5), and six-month momentum (Mom6), we construct them exactly the same as in Lochstoer and Tetlock (2020) by modifying their codes to establish the final sample. Stock returns are annualized from July to June next year, adjusted by deducting the inflation rate. ROE is the return on equity, calculated as earnings available for common over last year's book equity. BM is the book equity defined by Fama and French (1992) in December of year t[4] divided by the market capitalization in June of year $t+1$. When the book value of equity is missing, the historical book equity is used to supplement. Prof is annual revenues minus costs of goods sold, interest expense, and selling, general, and administrative expenses, divided by book equity. Inv_M5 is the five-year growth of the total asset. ME_D5 is the five-year change in market equity of June. Mom6 is the six-month cumulative return from January to June. We take the log to returns and Mom6, and transform the other variables mentioned above by adding one and taking the log. To avoid problems of extreme values when taking log (say, when ROE is near -1), we define pseudo-firms as a portfolio of ninety percent common stock and ten percent Treasury bills and adjust the variables accordingly.

For variables used to construct the anomalies, we obtain the announcement return (announcementreturn), five-year composite issuance (compequiss), accruals (accruals), net operating asset (noa), one-year asset growth (assetgrowth), distress probability (failureprobability), O-score (OScore), momentum in last 12 months (mom12), gross profitability (gp), return on asset (roaq) from the data provided by Chen (2022). We construct the net stock issues (lnNS) as annual log change in split-adjusted outstanding shares following Fama and French (2008), and investment-to-asset (ioa) as changes in gross property, plant, and equipment plus changes in inventory, divided by lagged total assets following Stambaugh and Yuan (2017). We carefully align these variables with annual returns to ensure data availability and timeliness when constructing anomalies.

[4]We restrict our sample to firms with a fiscal-year end in December to make the variables in the VAR logical in timing, and it simplifies the time alignment of book value as a bonus.

Several restrictions are applied to the sample. Firms do not get delisted, have trading information, have market capitalization larger than 10 million, and have BM between 0.01 and 100 in June. We also only keep firms with the fiscal-year end in December, with prices larger than or equal to five, and exclude tiny firms in the bottom NYSE quintile. These filtrations are common, especially when forming anomalies to avoid microstructure effects. Our final sample covers the period from 1964 to 2019.

All the twelve anomalies (named the same as the characteristic variables used to construct them) are value-weighted long-short portfolio returns. We group all stocks into ten groups but only use stocks listed on NYSE to get the breakpoints. For factors FIN, PEAD from Daniel, Hirshleifer, and Sun (2020), and MISPRICING, SIZE from Stambaugh and Yuan (2017), we follow the same procedure of sorting stock into 2×3 portfolios (except SIZE factor) and calculate the value-weighted returns. The most distinctive difference is that instead of a monthly updated PEAD, MISPRICING, and SIZE, we rebalance the factors yearly at the end of June. The reason for doing so is that the news components are extracted by regressing VAR at the yearly frequency, as the accounting variables in the VAR are at the annual frequency. If we have monthly rebalanced factors, what is the corresponding monthly news? It is not available. To investigate the potential adverse consequences of this practice, we provide the summary statistics for the factors in table 4.1 and the regression intercepts of using these models to explain anomaly returns in table 4.2.

From table 4.1, the MISPRCING factor (misp) delivers the highest average return of 0.0704, and the SIZE factor (size) has the lowest return of 0.0323. All returns are statistically different from zero as their t values are larger than 2.38. We check the explanatory power of the mispricing model of Stambaugh and Yuan (2017) in panel A of table 4.2 and of the behavioral model of Daniel, Hirshleifer, and Sun

(2020) in panel B. In each column, we regress the anomaly returns on the factors included in each model. If the model explains the anomaly, the intercept should be insignificant from zero. The mispricing factors we construct cannot explain announcement return and accruals, and the behavioral factors we construct cannot explain net issuance, accruals, and asset growth. Many papers document that using more timely information benefits the factors (see Asness and Frazzini (2013), Barillas and Shanken (2018)); given that we only update our factors at a yearly frequency, it is reasonable and acceptable to see certain anomalies not explained. So we regard the construction of these factors as satisfactory and continue explorations on them.

4.4.2 Recover news at different levels

There are two possible approaches to extracting the news components at the portfolio, anomaly or mean-variance efficient portfolio levels. One is to find the corresponding portfolio-level variables to be included in the VAR system and then get the news components directly at the portfolio level directly. Another is to aggregate firm-level news components to portfolio-level. The potential problem with the first approach, as explained in Lochstoer and Tetlock (2020), is that the cash flows and discount rates of rebalanced portfolios can differ substantially from those of the underlying firms in the portfolios. If we are only interested in the properties of the portfolio level instead of its underlying firms' properties, the concern is minor. But the main issue that hinders us is that we do not know how to construct the corresponding variables in the long-short portfolio. What is the book-to-market ratio or the dividend yield for a portfolio which is long in firms with high book-to-market ratios and short in firms with low book-to-market ratios? If it is reasonable to calculate the book-to-market ratio for a portfolio that takes a long position in its constituting stocks as the sum of the book value of all these stocks over the sum of the market capitalization or as the simple-weighted or value-weighted averages of the stocks constituting the portfolio, it by no means is sensible to do so for a long-short portfolio. This leaves us nothing else to do but to apply the second approach—aggregating firms' news into the portfolios' news following Lochstoer and Tetlock (2020).

To this end, we first need to extract the firm-level news in the following three steps.

First, estimate a time-series Vector Autoregression (VAR) for the aggregate market-level return decomposition. We can do so because the composition of the market portfolio is not rebalanced unless a firm quits the stock market.

$$Z_{t+1} = \mu^{agg} + A^{agg} Z_t + \varepsilon_{t+1}^{agg}$$

where $Z_t = [r_t^{agg}, roe_t^{agg}, bm_t^{agg}, prof_t^{agg}, inv_t^{agg}, me_t^{agg}, mom6_t^{agg}]'$ with each variable as the cross-sectional value-weighted average of the firm-level variables (all in log form). We do not restrict the roe_t^{agg} only as a dependent variable as Lochstoer and Tetlock (2020) do because our sample period is after 1964, and we can distinguish between ROE and profitability.

Then the unexpected return news, discount rate news, and cash flow news at the aggregate level are calculated as[5]:

$$r_{t+1}^{agg} - E_t r_{t+1}^{agg} = e_1' \varepsilon_{t+1}^{agg}$$

$$DR_{t+1}^{agg} = E_{t+1} \sum_{j=2}^{\infty} \kappa^{j-1} r_{t+j}^{agg} - E_t \sum_{j=2}^{\infty} \kappa^{j-1} r_{t+j}^{agg} = e_1' \kappa A^{agg} \left(I_7 - \kappa A^{agg}\right)^{-1} \varepsilon_{t+1}^{agg}$$

$$CF_{t+1}^{agg} = r_{t+1}^{agg} - E_t \left[r_{t+1}^{agg}\right] + DR_{t+1}^{agg} = e_1' \left(I_7 + e_1' \kappa A^{agg} \left(I_7 - \kappa A^{agg}\right)^{-1}\right) \varepsilon_{t+1}^{agg}$$

[5]To see how we obtain the formula for calculating DR news, see Callen and Segal (2010). But shortly speaking, the components (difference in expectations for each period) form a geometric series with a common ratio of $\kappa \times A^{agg}$.

where e_1 is a 7×1 column vector with one as its first element and zeros elsewhere, I_7 is a 7×7 identity matrix, $\kappa = 0.967$ as in Vuolteenaho (2002). As implied in the equations above, the CF news is calculated residually from the present-value identity, which states that the unexpected return news equals the CF news minus DR news. There exists the possibility that the unexpected return news contains shocks other than CF and DR news. Therefore, the CF news we get may be larger than actual CF news, considering that other shocks are included in the CF news. Another approach that may solve this problem is to calculate CF news directly, as we do with DR news. But as explained in Lochstoer and Tetlock (2020), it does not capture the CF for stockholders correctly.

Second, we estimate a panel VAR for the demeaned firm-level return decomposition.

$$Z_{i,t+1} = \mu^{ma} + A^{ma} Z_{i,t} + \varepsilon^{ma}_{i,t+1}$$

where $Z_t = [r^{ma}_{i,t}, roe^{ma}_{it}, bm^{ma}_{it}, prof^{ma}_{it}, inv^{ma}_{it}, me^{ma}_{it}, mom6^{ma}_{it}]'$ with each variable as the firm-level variable demeaned by the cross-sectional value-weighted average of that variable, say $r^{ma}_{it} = r_{it} - r^{agg}_{t}$.

Then the demeaned firm-level news components are calculated similarly as at the aggregate level:

$$r^{ma}_{i,t+1} - E_t r^{ma}_{i,t+1} = e_1' \varepsilon^{ma}_{i,t+1}$$

$$DR^{ma}_{i,t+1} = E_{t+1} \sum_{j=2}^{\infty} \kappa^{j-1} r^{ma}_{i,t+j} - E_t \sum_{j=2}^{\infty} \kappa^{j-1} r^{ma}_{i,t+j} = e_1' \kappa A^{ma} \left(I_7 - \kappa A^{ma}\right)^{-1} \varepsilon^{ma}_{i,t+1}$$

$$CF^{ma}_{i,t+1} = r^{ma}_{i,t+1} - E_t \left[r^{ma}_{i,t+1}\right] + DR^{ma}_{i,t+1} = e_1' \left(I_7 + e_1' \kappa A^{ma} \left(I_7 - \kappa A^{ma}\right)^{-1}\right) \varepsilon^{ma}_{i,t+1}$$

where e_1, I_7, κ are the same defined vectors or scalar. We use the inverse of the number of firms in each year as a weight to each firm in that year, so the weighted least square regression weights each year equally following Vuolteenaho (2002).

Third, the firms' total news components are defined as the sum of the corresponding aggregate-level news and demeaned firms' news:

$$\begin{aligned} r_{i,t+1} - E_t r_{i,t+1} &= (r^{agg}_{t+1} - E_t r^{agg}_{t+1}) + (r^{ma}_{i,t+1} - E_t r^{ma}_{i,t+1}) \\ &= e_1' \varepsilon^{agg}_{t+1} + e_1' \varepsilon^{ma}_{i,t+1} \\ DR_{i,t+1} &= DR^{agg}_{t+1} + DR^{ma}_{i,t+1} \\ CF_{i,t+1} &= CF^{agg}_{t+1} + CF^{ma}_{i,t+1} \end{aligned}$$

As Lochstoer and Tetlock (2020) explains, this procedure allows the VAR coefficients to differ for the common movement (A^{agg}) and firms' idiosyncratic movement (A^{ma}) to match the data.

Next, we need to aggregate the firm-level news components into the different portfolio levels.

To form portfolio/anomaly/factor level news, we follow the same procedure of calculating returns except that we use firms' news to replace firms' returns. At the end of June of year t, all information needed to form portfolios is known. We must assume that all the (long-short) portfolios are held for one year. The assumption is necessary because the availability of firms' CF or DR news is subject to financial information in the VAR system, and they only reflect yearly news instead of monthly or quarterly news. Therefore, for factors that are rebalanced every month or every quarter, like those in Hou, Xue, and Zhang (2015), there is no way to organize the analysis in this way unless we only keep the construction spirit but lower the updating frequency to yearly.

For example, we sort firms on their book-to-market ratios and group them evenly into ten portfolios in June of year t. Then the ten portfolios' returns are the simple-weighted or value-weighted returns using next year's returns. The calculation for the news is similar. The CF news of each portfolio is the simple-weighted or value-weighted CF news of the firms in each portfolio in the next holding year. The unexpected return news and DR news are calculated in the same way.

For the long-short portfolio, we use the news of the portfolio in the long position minus the news of the portfolio in the short position. For the factor, it follows the same logic. Throughout this chapter, we use NYSE breakpoints to divide all stocks in the sample into ten groups and use market capitalization as the weights. For factors, we follow the exact procedure in the original papers except that we only update them once a year, instead of a monthly updated version of PEAD factor from Daniel, Hirshleifer, and Sun (2020), for example.

For the news components at the MVE level, we need to calculate the weights for each factor included by maximizing the full-sample Sharpe ratio. For example, for the MVE portfolio composed of FIN, PEAD, and Market factors, the corresponding weight is calculated as the inverse of their covariance matrix times the mean of these factors. The MVE news components are then the value-weighted sum of the factor news components, say the cash flow news of the MVE portfolio is $CF_t^{MVE} = \sum_{i \in \{FIN, PEAD, Market\}} w_i \times CF_t^i$, where w_i is the weight associated with factor i to achieve the maximal Sharpe ratio, and CF_t^i is the CF news for factor i.

The output from all the aggregation is a time-series news data for each portfolio.

4.5 Empirical results

Now we present the estimation results for the VAR specified above, decompose the returns to CF news and DR news to study the main driver for the variation of returns at different levels, explore the source of the behavioral factors' explanatory power, and analyze the systematic risk associated with these factors.

4.5.1 VAR estimation

As Chen and Zhao (2009) point out, the specification of VAR affects the relevant importance of CF news and DR news. The conclusion may contradict each other under different model settings. Our specification of the variables included in the vector follows the main specification in Lochstoer and Tetlock (2020) except that we do not restrict the coefficients on the lagged ROE to zero, as explained in the empirical design part. Returns, book-to-market ratio, and ROE are necessary to be contained in the system, especially for the firm-level decomposition, because they are specified in the deduction of return decomposition in Vuolteenaho (2002). The other variables add value to the prediction of returns. Lochstoer and Tetlock (2020) shows that this specification provides a reasonable approximation o the long-run dynamics of returns and earnings. At the aggregate level, the more frequently seen variables are dividend growth, dividend yield, and other variables like value spread, eqis (the ratio of equity issuing activity as a fraction of total issuing activity), etc. However, they mainly apply the decomposition to the stock index like the value-weighted NYSE stock index, in which case we have well-defined relevant variables available. In our paper, it is necessary to study the firm-level returns, and as most firms do not pay dividends, we can only apply the firm-level decomposition from Vuolteenaho (2002). Besides, we are interested in the common movement of firms in our sample, so we use the value-weighted variables to learn firms' return decomposition. But we also include other variables: term yield spread, the default yield spread, and the small stock defined in the data part as a robustness check, and it turns out they do not affect the results that much.

Table 4.3 shows coefficients and t statistics (in parenthesis) for the time-series aggregate VAR, and table 4.4 shows those for the panel VAR. The sample period for

VAR estimation is from 1964 to 2019 to enlarge the data used in the regression instead of from 1974 when some anomalies like PEAD are available due to data issues.

From table 4.3, we can see that the accounting variables—ROE, BM, profitability, investment, and market capitalization, as usually more persistent, are better predicted using past information. The adjusted R^2s for these variables are moderate or large. For return variables, i.e., the value-weighted real returns, and the value-weighted momentum returns, the adjusted R^2s are much smaller, 6% and 5%, respectively. It reveals that the returns are less predictable, at least for the short term. Last year's return does not convey any information for this year's return. Among all the other dependent variables, only investment (asset growth in the past five-year) exhibits a significant coefficient, manifesting the long-run predictability of market returns.

At the firm-level panel regression, we adopt the common practice of weighting each year equally by using weighted least squares. As firm-level variables contain more idiosyncratic information, the adjusted R^2s are smaller than the aggregate VAR R^2. Only 2% variation in returns is explained by the other characteristics, which is routine when predicting firm returns. For the value-weighted market returns, the signs of the coefficients on dependent variables are consistent with previous studies, though some are insignificant.

4.5.2 Decomposing returns

Once we obtain the coefficients and errors from the VARs, we can compute the news at our desired level and analyze the relevant importance of the news components for the variation of unexpected return news and explanatory power.

Table 4.5 shows the variance decomposition for firm-level and market-level returns. We see from the first row that, consistent with previous studies like Campbell

(1991) and Vuolteenaho (2002) among others, the discount rate news is the main driver of the market. The variance of DR news over the variance of unexpected return news is 85.49%, quite a large proportion. The variance of CF news or the covariance components only takes up to 17.88% and -3.37%, respectively. In total, the sum of them equals one, as we use the present value identity to extract the CF news. The correlation between CF news and DR news is only 0.0432. It is quite reasonable as the CF news, which is usually idiosyncratic, gets diversified away at the aggregate market level. What is left is the variation in the DR news which generally stands for the market sentiment. The second row shows that for the demeaned firm-level returns, CF news accounts for 95.74% of the total variation in the unexpected returns, leaving DR news and their covariance accounting for less than 10% of the change. The correlation between CF news and DR news is again very small, less than 0.1. The decomposition for total firm news is shown in the third row, where we observe that though the contribution of DR news increases as we add back the common movement components among firms, it is still the CF news that is the main driver of returns.

We also look at the anomalies' variance decomposition. For the underlying eleven anomalies used to construct the behavioral factors, we report the results in table 4.6. All anomalies are rebalanced yearly, using NYSE breakpoints to be divided into ten groups, and formed using the long-short value-weighted method. Consistent with Lochstoer and Tetlock (2020) in which they study five well-known traditional anomalies—value, size, profitability, investment, and momentum, and find that the systematic CF news drives the returns of anomaly portfolios, we find similar results using a different set of anomalies. The ratio of the variance of CF news over the variance of unexpected news ranges from 84.51% to 118.34%. The variation of DR news ranges from 3.34% for the composite equity issuance anomaly to 12.74% for the momentum anomaly. The correlation between the CF news and DR news ranges from 0.0049 to 0.40 in absolute values.

When it comes to our most concerning factors and their MVE portfolios in this chapter, table 4.7 presents the results for the decomposition.

Panel A shows the decomposition for every single factor. For FIN which stands for the long-run managers' equity issuance, CF news takes up nearly 90% of the variation in the unexpected return news. Many possible explanations exist for the negative relationship between stock issuance of future returns. From the behavioral-based view, one possible interpretation is that firms' managers, as insiders who know the intrinsic value of their stocks better, are willing to issue stocks when they think their stock is currently overpriced and can earn money at the expense of outside investors, and therefore future stock return decreases. A possible rational-based explanation is that when a firm has steady cash flow revenues, managers have spare money to pay back to the equity holders by repurchasing stocks (a form of payout policy that is more often applied nowadays to pay dividends, see Farre-Mensa, Michaely, and Schmalz (2014)), thus the issuance decreases but the future return increases. Our result of CF news driving the most variation in returns provides evidence favoring the latter view, though it does not reject the existence of behavioral motivation. The PEAD factor is usually regarded as an outcome of short-run investors' underreaction to market information. However, there are also rational explanations for this; see Fink (2021) for a review of all these explanations. In our table, the CF news takes up 77.01% of the variation in the return news. As CF news epitomizes the fundamental risk, it indicates that the risk-based explanation plays a large role in the presence of this anomaly.

The MISP factor synthesizes the underlying eleven anomalies and aggregates the information in management and firms' performance. Stambaugh and Yuan (2017) argues that this factor is consistent with a mispricing interpretation as it is predicted by sentiment. However, the return decomposition says that the CF news is the main driver of the return news with a portion of 92.71% in the variation of return news. For the SIZE factor from Stambaugh and Yuan (2017), they construct it using a subset of stocks that are most likely not subject to mispricing. We see again that the CF news is the main driver. The correlation between CF news and DR news is negative for individual anomalies, and the magnitude is larger compared with firm-level and market-level decomposition.

Panel B exhibits the decomposition for MVE portfolios in which the market factor is always included. The row "All Factors" means that we incorporate FIN, PEAD, MISP, SIZE factors, and market factors in the composition of the MVE portfolio. The row "Behavioral Factors" implies that we have FIN, PEAD, and market factors included, while the row "Mispricing Factors" means the MVE portfolio only includes MISP, SIZE, and market factors. Across all the specifications, the variation of CF news takes up 68.25% to 94.95% in the variance of return news. The DR news plays a more important role in the portfolio consisting of behavioral factors, as they are more prone to be affected by the market timing of managers or the underreaction of investors compared with mispricing factors. Panel C presents the results for the MVE portfolio without the market factor. The pattern is similar to that in panel B. When the market factor whose DR news contributes most of the return's variation is excluded, we see that the var(DR) decreases in general.

Along with the discussion of the results, we see that the CF news is always the significant driver for returns at every factor level and MVE portfolio level except the market factor. This trend highlights the common fundamental risk contained in these anomalies/factors and thus supports the risk-based explanation.

Next, we turn to the analysis of the explanatory power of these factors. Previous studies usually argue that their factor model is better because it expands the efficient frontier spanned by other factor models and because they explain many anomalies' returns as delivering insignificant intercepts in the factor model regression. Yet nobody knows where their explanatory power comes from. In this chapter, since the unexpected returns are decomposed into CF news and DR news components, it enables us to investigate from which news components these factors explain the anomalies.

Table 4.8 demonstrates the adjusted R^2 when regressing anomalies' return news on the CF news, DR news, and return news of the factors in each factor model. The CAPM in which only the market factor is included does not explain these anomalies. As a result, the news components of the market factor have little explanatory power, as shown in panel A. We take the noa anomaly as an illustration. When we regress the unexpected return news of noa on the CF news of the market factor, the adjusted R^2 is even negative, with a statistic of –1.88%. When we regress the same return news on the DR news of the market factor, 0.02% of the variation is explained. If the independent variable is the return news of the market, then the adjusted R^2 is 0.39%. The explanatory power of the CF news of the market factor ranges from -2.16% to 8.29%. The power of the DR news ranges from -2.17% to 7.87%, and the power of the total return news ranges from -2.13% to 9.02%. It implies that consistent with the failure of CAPM to explain anomalies, the news components of the market factor generally do not explain much of the return news of these anomalies.

Panel B shows the adjusted R^2 for the behavioral model. We regress anomaly news on the news components of FIN, PEAD, and market factors. This model explains many anomalies much better than the CAPM does. Taking stock's net issuance (lnNS), for example, the CF (DR) news of these factors explains 62.05% (14.6%) of the return news of this anomaly, and the total return news of them explains 64.94% of the anomaly. This large explanatory power is not only attributable to the close relation between this anomaly and the FIN factor which uses the information of issuance anomaly. The assetgrowth anomaly is not involved in the construction of the behavioral factors, yet the CF news of these behavioral factors still explains it well. The portion of return news of this anomaly explained by the CF news, DR news, and total return news of the factors is 40.3%, 13.4%, and 43.11%, respectively. The observation that the CF news of factors contributes most of the explanatory power of their return news to explain the anomaly's return news is widespread. All the patterns shown in panel B also hold in panel C where we use news components from the mispricing factors, and in panel D, in which we use news components of all the factors (FIN, PEAD, MISPRICING, SIZE, Market). The adjusted R^2s are the highest when we combine all factors except anomaly announcementreturn, and roaq.

We conclude from this section that the CF news is the main driver of all these factors we study, though they are considered more subject to behavioral phenomena. Their explanatory power mainly comes from the CF news components as well. Notice that it is not caused by the large portion of CF news variation over return news variation, as in table 4.8, the adjusted R^2 captures the covariance between CF news and return news of each anomaly.

4.5.3 Systematic risks

We now turn to study the systematic risks of these factors with respect to the market return as Campbell, Polk, and Vuolteenaho (2010) who explore the value anomaly. The factors are also anomalies. The beta decomposition allows us to investigate why the portfolios sorted on these factor characteristics have different systematic risks. Table 4.9, 4.10, 4.11 reports for PEAD, FIN, and MISPRICING factors, respectively.

Portfolios are indicated in the first row of each table. We do not sort on factor characteristics and then group stocks into five portfolios based on the quintiles by convention. Instead, we follow the construction of these factors by sorting them into three groups—L, M, H. One reason is that the interaction of two variables determines the grouping of the FIN factor, and it is hard to sort FIN into five groups. Portfolio 1 refers to the portfolio of stocks with the lowest factor characteristic; Portfolio 3 refers to the one with the highest factor characteristic; Portfolio 2 is the group of stocks not used when forming the factor returns; and Portfolio 3-1 refers to the long-short portfolio. We would like to note here that for FIN and PEAD, the factor characteristics are positively related to future stock returns, so the long-short portfolios are also the factor portfolios. For the MISPRICING, however, the mispricing measure is negatively related to future stock returns. Therefore the long-short portfolio 3-1 shown in the last column of table 4.11 is the reverse of forming the factor returns.

i, CFm represents the bad cash-flow beta, which reflects the covariance of the portfolio's returns news and the CF news of the market factor over the variance of market returns. We obtain the estimated bad beta by regressing each portfolio's return news on the scaled CF news of the market by timing $\frac{var(r_m^e)}{var(CF_m)}$ where $var(r_m^e)$ is the variance of the market factor we construct in this chapter and $var(CF_m)$ is the variance of CF news of the market factor we extract. The coefficient associated with market CF news is the beta desired. We further decompose the bad beta into two components driven either by the portfolio's DR news (DRi, CFm) or CF news (CFi, DRm) by regressing the DR news and the CF news of the portfolio on the scaled market CF news respectively.

i, DRm represents the good discount-rate beta, which reflects the covariance of the portfolio's returns news and the DR news of the market factor over the variance of market returns. We obtain the estimated bad beta by regressing each portfolio's return news on the scaled DR news of the market by timing $\frac{var(r_m^e)}{var(DR_m)}$ where $var(DR_m)$ is the variance of DR news of the market factor we extract. Similarly, it is further decomposed into DRi, DRm and CFi, DRm.

Table 4.9 displays all the beta components associated with factor PEAD. We observe the following patterns:

First, comparing bad beta and good beta in portfolios except for the long-short portfolio, we see that consistent with Campbell, Polk, and Vuolteenaho (2010), the good discount-rate beta takes up a larger portion of the total beta. For portfolio 1, the bad beta is 0.1124, while the good beta is 0.6480, which gives a total CAPM beta of 0.7604. The magnitude of good beta takes up around 85% of the total beta. All these betas are significant in each portfolio. Yet the good beta does not exhibit much variation across portfolios. The long-short factor's good beta, 0.0372, is not significantly different from 0 as its standard error is 0.0429. On the other hand, the bad beta for the long-short factor is 0.0335 with a standard error of 0.0191.

Second, when we look at the elements (betas of DRi, CFm, and of CFi, CFm) of the bad beta (betas of i, CFm), we see that the magnitude of the portfolio's CF-driven beta is larger than that of DR-driven beta across portfolio 1, 2, and 3. The betas of DRi_CFm across portfolios 1, 2, and 3 are small and insignificant. Besides, they do not vary much across portfolios and lead to an insignificant number of 0.0008 with a standard error of 0.0054 in the long-short portfolio. On the contrary, the betas of CFi, CFm are large and significant in all portfolios, the long-short one included. It means that the cash flows of stocks with a large characteristic of PEAD are particularly sensitive to permanent movements in aggregate stock prices proxied by the CF news of the market factor.

Third, if we turn to the elements (betas of DRi, DRm, and of CFi, DRm) of the good beta (betas of i, DRm), we see that the magnitude of the portfolio's DR-driven beta is larger than that of CF-driven beta across portfolio 1, 2, and 3. The betas of DRi_DRm across portfolios 1, 2, and 3 are large and significant, but there is little variation across portfolios. The betas of CFi, DRm, though smaller and insignificant across portfolios 1, 2, and 3, take up most variation (0.0391/0.0372 v.s. -0.0019/0.0372) in the good beta in the long-short portfolio.

As shown in table 4.10, these patterns also hold for the FIN factor. We have a changing pattern for the MISPRICING factor, and one additional phenomenon calls our attention in table 4.11. First, the good beta (betas of i, DRm) and the beta components of good beta—betas of DRi, DRm also show variation across portfolios now, leading to a significant number of 0.1276, 0.0369 in the long-short portfolio. Second, unlike PEAD and FIN whose good betas and bad betas contribute to the total beta in the same direction, now we see that the bad beta for the MISPRICING factor is 0.0495 (negative of returns in 3-1 portfolio, as noted earlier in this section), while the good beta for the MISPRICING factor is -0.1276, leading to a negative total beta. The reason why the MISPRICING factor earns a positive return though its total beta is negative corresponds exactly to the explanations in Campbell and Vuolteenaho (2004) where they argue that the price for bad beta is higher than for good beta. The good beta does not benefit the returns of the factor.

Figure 4.1 draws the total beta and its four subcomponents. The variation across portfolios is now more apparent. The trend of the total beta is mainly driven by the betas of CFi, CFm or of CFi, DRm, which suggests that the variation of beta stems from the fundamental part of the stocks.

In sum, we have three comments. First, the bad beta always exhibits variation across portfolios. In contrast, the good beta, though constituting a large portion of the total beta, either is stable or does not contribute to the positive returns of the factor. Second, when analyzing the components of bad beta or good beta, the magnitude of betas associated with portfolios' CF news is substantial. Third, the variation of total beta is driven by the beta components relevant to the firms' fundamental CF news.

4.6 Conclusion

We explore what drives the newly proposed behavioral factors. These factors, which are portfolio returns, have been demonstrated to better describe the return variation across different stocks or portfolios. However, it is controversial about the underlying economic rationale behind these factors. Are they proxies for fundamental risk or misvaluation, where does their explanatory power come from, and what determines their systematic exposure to market returns? This chapter attempts to answer these questions by decomposing the factor returns into cash flow news and discount rate news to identify which component primarily influences these factor returns, and conducting regression analyses utilizing news components.

Our findings lead to three key conclusions. Firstly, CF news plays a pivotal role in the variance of returns for these factors and their MVE portfolio. Secondly, their explanatory power largely originates from their CF news explaining the return news of other anomalies. Thirdly, though the good beta accounts for a significant part of the market beta, it is consistently the beta components associated with CF that contribute to the variation of the market beta across portfolios.

The implications of these results suggest that even though behavioral factors are often considered to be driven by behavioral considerations, indicating an inefficient market or irrational investor, our study presents evidence that they are principally driven by CF fundamentals. A promising future direction for this research might be to investigate whether this approach could further illuminate the anomaly shrinkage by examining the news components across a broader range of anomalies.

4.7 Figures and tables

4.7.1 Figures

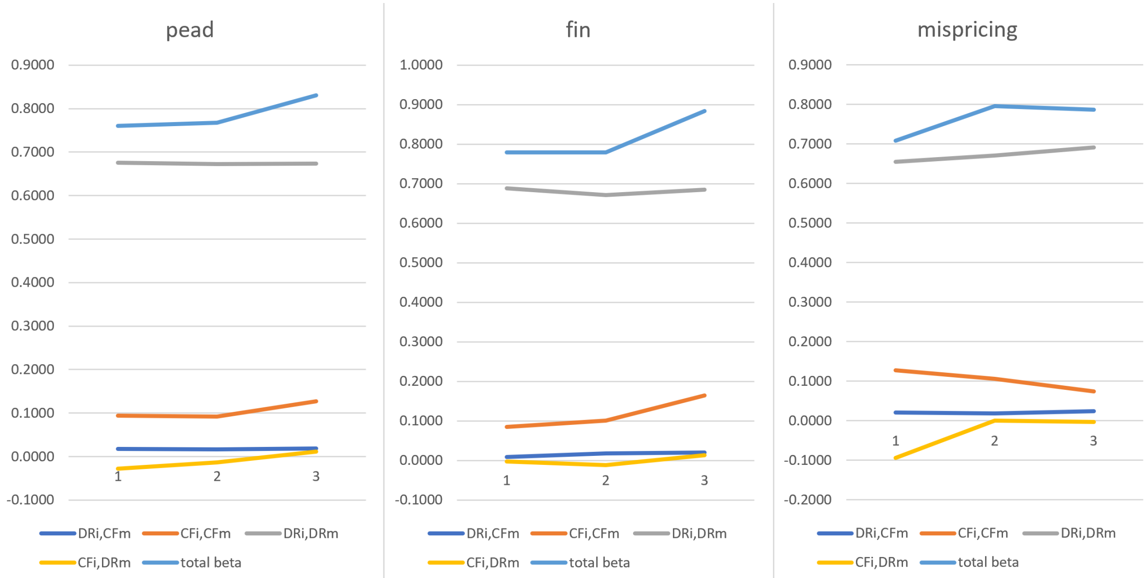

Fig. 4.1: Cross-sectional variation in the components of beta

This figure plots the total beta and its four components for pead factor (the left one), fin factor (the middle one), and mispricing factor (the right one).

4.7.2 Tables

Table 4.1: Summary Statistics for Factor Returns
This table presents the mean (Mean), standard deviation (SD), minimum (Min), maximum (Max), and Newey-West corrected t-statistics with six lags (t) for annual value-weighted market excess returns (mkt), annual size factor returns constructed in the spirit of Stambaugh and Yuan (2017) (size), annual mispricing factor returns constructed in the spirit of Stambaugh and Yuan (2017) (misp), annual financing factor constructed in the spirit of Daniel, Hirshleifer, and Sun (2020) (fin), and annual post-earnings announcement drift factor constructed in the spirit of Daniel, Hirshleifer, and Sun (2020) (pead). The sample period is from 1972 July to 2020 June mainly due to the availability of accounting information needed to construct factors.

	mkt	size	misp	fin	pead
Mean	0.0609	0.0323	0.0704	0.0645	0.0361
SD	0.1399	0.0942	0.0929	0.1503	0.0582
Min	-0.2687	-0.2302	-0.0736	-0.2215	-0.0782
Max	0.4601	0.2481	0.3963	0.5627	0.2139
t	3.0157	2.3773	5.2512	2.9733	4.3036

Table 4.2: Explanatory Power of Factor Model
This table shows the coefficients of regressing twelve anomalies separately on the mispricing factor model (Panel A) and on the behavioral model (Panel B). The numbers in parenthesis are Newey-West corrected t-statistics with six lags.

Panel A: Mispricing Factor Model												
	lnNS	announcementreturn	compequiss	accruals	noa	assetgrowth	ioa	failureprobability	OScore	mom12m	gp	roaq
mkt	-0.1604	0.0463	0.4820***	0.0709	0.1703	-0.0655	0.0271	-0.1543	-0.0986	0.0898	0.0861	-0.2584
	(-1.5350)	(0.5505)	(4.3404)	(1.0604)	(1.5958)	(-0.5326)	(0.3833)	(-1.0173)	(-1.2621)	(0.7095)	(0.6319)	(-1.9538)
misp	0.6355***	-0.0804	0.1132	-0.2802**	0.4906**	0.2858**	0.3400	0.5152**	0.7393***	0.5122**	0.7390***	0.8105***
	(5.5149)	(-0.8109)	(0.8198)	(-2.2912)	(2.4757)	(2.2729)	(1.8948)	(2.1424)	(5.4018)	(2.0348)	(3.1763)	(3.1587)
size	-0.0472	-0.3719	-0.6582***	-0.3439**	0.2967	0.5788***	0.6539***	-0.3829	-0.4804***	-0.2835	-0.0215	-0.3795
	(-0.3129)	(-1.9265)	(-3.7198)	(-2.1018)	(1.7634)	(4.5087)	(8.5419)	(-1.6477)	(-2.8616)	(-1.2614)	(-0.0720)	(-1.6574)
_cons	0.0310	0.0499**	0.0240	0.0552***	-0.0094	0.0128	-0.0397	0.0179	-0.0300	0.0294	-0.0310	-0.0255
	(1.6161)	(2.5893)	(0.8699)	(3.6385)	(-0.3478)	(0.6144)	(-1.5082)	(0.5160)	(-1.7548)	(0.8329)	(-0.9621)	(-0.9300)
Panel B: Behavioral Factor Model												
	lnNS	announcementreturn	compequiss	accruals	noa	assetgrowth	ioa	failureprobability	OScore	mom12m	gp	roaq
mkt	-0.0833	-0.1016	0.3027***	-0.0046	0.2542**	0.0638	0.1075	-0.1247	-0.1192	-0.1098	0.1013	-0.2892
	(-1.0038)	(-1.5693)	(3.0637)	(-0.0692)	(2.5634)	(0.6367)	(1.1683)	(-0.6661)	(-0.9797)	(-0.8766)	(0.5932)	(-1.6929)
pead	-0.2514	1.2961***	1.4280***	0.2775	-0.0130	-0.6628***	-0.0495	-0.7394	-0.0333	1.4876***	-0.5429**	0.2931
	(-1.7149)	(8.2312)	(4.6118)	(0.9063)	(-0.0482)	(-3.5977)	(-0.1897)	(-1.7106)	(-0.1472)	(3.5993)	(-2.3421)	(0.4673)
fin	0.4987***	-0.1704***	-0.1222	-0.2494***	0.4589***	0.3029**	0.2777**	0.1794	0.2861**	-0.1653	0.1509	0.3412**
	(8.5186)	(-4.2873)	(-0.7071)	(-3.1846)	(3.2066)	(2.1454)	(2.6019)	(1.0609)	(2.4335)	(-0.7265)	(1.1479)	(2.3953)
_cons	0.0465***	0.0054	-0.0221	0.0350**	0.0005	0.0482***	-0.0157	0.0551	-0.0095	0.0253	0.0293	-0.0114
	(4.4573)	(0.4174)	(-0.9144)	(2.1576)	(0.0324)	(2.8858)	(-0.7940)	(1.5524)	(-0.4038)	(0.8893)	(1.0143)	(-0.3852)

t statistics in parentheses

** p<0.05 *** p<0.01"

Table 4.3: Aggregate VAR
This table presents the time-series aggregate VAR results. The variables in the first row are dependent variables while the variables in the first column are independent variables. Each column represents a regression. The variables are all value-weighted averages of the corresponding firm-level variables. The sample is from 1964 to 2019. Heteroskedasticity-adjusted (White) standard errors appear in parentheses.

	lnRealRet_Jun	lnROE_V02	lnBM	lnProf	lnInv_M5	lnME_D5	lnMom6
lag1_lnRealRet_Jun	-0.0137	-0.00586	0.0998	0.00988	-0.00882	0.0510	-0.0723
	(-0.09)	(-0.24)	(0.77)	(0.57)	(-0.81)	(0.21)	(-0.52)
lag1_lnROE_V02	-1.575	0.242	1.226	-0.285***	0.220***	0.136	-0.923
	(-1.39)	(1.34)	(1.34)	(-3.44)	(3.29)	(0.09)	(-0.99)
lag1_lnBM	0.147	-0.00175	0.854***	0.00415	-0.0105	0.0407	0.142**
	(1.24)	(-0.17)	(8.72)	(0.45)	(-1.92)	(0.25)	(2.29)
lag1_lnProf	1.018	0.417***	-0.582	1.076***	0.151***	1.887	0.934
	(0.87)	(2.92)	(-0.67)	(13.65)	(2.93)	(1.35)	(1.35)
lag1_lnInv_M5	-1.625**	-0.212	1.567**	-0.120	0.739***	-2.731**	-1.484**
	(-2.36)	(-1.56)	(2.50)	(-1.37)	(9.73)	(-2.17)	(-2.20)
lag1_lnME_D5	-0.0156	-0.00874	-0.0597	-0.00501	0.0102**	0.753***	0.0694
	(-0.18)	(-0.84)	(-0.81)	(-0.54)	(2.18)	(4.97)	(1.30)
lag1_lnMom6	0.133	0.0192	0.0259	-0.0376	-0.000965	-0.109	-0.0629
	(0.56)	(0.52)	(0.13)	(-1.55)	(-0.06)	(-0.29)	(-0.31)
_cons	0.220	0.0145	-0.221	0.0323	-0.0455***	-0.0399	0.122
	(1.03)	(0.57)	(-1.29)	(1.69)	(-4.05)	(-0.12)	(0.86)
N	56	56	56	56	56	56	56
adj. R^2	0.060	0.396	0.797	0.778	0.903	0.444	0.054

t statistics in parentheses
** $p < 0.05$, *** $p < 0.01$

Table 4.4: Panel VAR
This table presents the panel VAR results. The variables in the first row are dependent variables while the variables in the first column are independent variables. Each column represents a regression. The variables are all market adjusted by deducting the cross-sectional value-weighted averages. The sample is from 1964 to 2019. Heteroskedasticity-adjusted (White) standard errors appear in parentheses.

	lnRealRet_Jun	lnROE_V02	lnBM	lnProf	lnInv_M5	lnME_D5	lnMom6
lag1_lnRealRet_Jun	0.0563	0.107***	0.102***	0.0449***	0.00982***	0.194***	0.0190
	(1.47)	(7.33)	(2.78)	(6.27)	(4.34)	(4.11)	(0.71)
lag1_lnROE_V02	0.0638***	0.256***	0.109***	-0.0262	0.0333***	-0.148**	0.0198
	(2.79)	(10.41)	(4.09)	(-1.48)	(7.05)	(-2.59)	(1.87)
lag1_lnBM	0.0187**	-0.0218***	0.932***	-0.00754***	-0.00319***	-0.00848	0.00930
	(2.32)	(-9.64)	(93.04)	(-3.98)	(-6.49)	(-0.75)	(1.59)
lag1_lnProf	0.0910***	0.219***	0.0228	0.617***	-0.00271	0.147***	0.0559**
	(2.95)	(12.71)	(0.93)	(23.76)	(-0.85)	(3.25)	(2.62)
lag1_lnInv_M5	-0.160***	-0.117***	0.129***	-0.0803***	0.712***	0.00663	-0.0497**
	(-6.18)	(-9.81)	(4.17)	(-9.32)	(77.98)	(0.14)	(-2.26)
lag1_lnME_D5	-0.0201**	0.00528***	0.0445***	-0.000281	0.0203***	0.719***	-0.0157**
	(-2.35)	(3.97)	(5.60)	(-0.20)	(20.02)	(51.81)	(-2.26)
lag1_lnMom6	0.101**	0.0248	-0.0318	0.0283***	-0.0000739	0.0437	0.0288
	(2.64)	(1.67)	(-0.91)	(3.36)	(-0.03)	(0.92)	(1.28)
_cons	0.000896	-0.0126***	0.0350***	-0.0143***	-0.00291***	-0.0222	0.00732
	(0.12)	(-6.59)	(4.03)	(-10.14)	(-4.14)	(-1.83)	(1.37)
N	80233	80570	80571	80571	80571	80571	80571
adj. R^2	0.026	0.211	0.779	0.396	0.774	0.561	0.009

t statistics in parentheses
** $p < 0.05$, *** $p < 0.01$

Table 4.5: Variance Decomposition for Firm-Level and Market Return
The table displays the variance decomposition of the firm- and market-level real returns. var(CF) stands for the portion of the variance of cash flow news on the variance of unexpected return news, var(DR) stands for the portion of the variance of discount rate news on the variance of unexpected return news, -2cov(CF,DR) stands for the portion of covariance between cash flow news and discount rate news on the variance of unexpected return news, corr(CF,DR) stands for the correlation between cash flow news and discount rate news. mkt return refers to the decomposition from aggregate VAR. firm mkt-adj return refers to the decomposition from panel VAR. firm return refers to the decomposition by combining components of market returns and firm market-adjusted returns. The sample is from 1972 to 2019.

	var(CF)	var(DR)	-2cov(CF,DR)	corr(CF,DR)
mkt return	17.88%	85.49%	-3.37%	4.32%
firm mkt-adj return	95.74%	8.48%	-4.21%	7.40%
firm return	85.74%	16.93%	-2.67%	7.40%

Table 4.6: Variance Decomposition for Anomaly-Level Return
The table displays the variance decomposition of anomaly-level (anomalies indicated in the first column) returns. var(CF) stands for the portion of the variance of cash flow news on the variance of unexpected return news, var(DR) stands for the portion of the variance of discount rate news on the variance of unexpected return news, -2cov(CF,DR) stands for the portion of covariance between cash flow news and discount rate news on the variance of unexpected return news, corr(CF,DR) stands for the correlation between cash flow news and discount rate news. The sample is from 1972 to 2019.

	var(CF)	var(DR)	-2cov(CF,DR)	corr(CF,DR)
lnNS	89.86%	3.49%	6.65%	-18.76%
announcementreturn	92.87%	9.92%	-2.79%	4.60%
compequiss	84.51%	3.34%	12.15%	-36.17%
accruals	90.34%	3.62%	6.04%	-16.71%
noa	87.26%	6.32%	6.42%	-13.66%
assetgrowth	91.60%	3.30%	5.10%	-14.68%
ioa	92.53%	4.77%	2.70%	-6.42%
failureprobability	93.65%	6.11%	0.23%	-0.49%
OScore	100.49%	3.50%	-3.99%	10.64%
mom12m	118.34%	12.74%	-31.08%	40.02%
gp	85.06%	3.78%	11.17%	-31.15%
roaq	104.59%	6.23%	-10.82%	21.20%

Table 4.7: Variance Decomposition for Factor-Level and Mean-Variance Efficient Portfolio Return
The table displays the variance decomposition of individual factor-level returns (Panel A), MVE portfolios with market factor (Panel B), and MVE portfolios without market factor (Panel C). var(CF) stands for the portion of the variance of cash flow news on the variance of unexpected return news, var(DR) stands for the portion of the variance of discount rate news on the variance of unexpected return news, -2cov(CF,DR) stands for the portion of covariance between cash flow news and discount rate news on the variance of unexpected return news, corr(CF,DR) stands for the correlation between cash flow news and discount rate news. "All Factors" means that we include all individual factors in Panel A with or without market factor. "Behavioral Factors" include "FIN" and "PEAD". "Mispricing Factors" include "MISP" and "SIZE". The sample is from 1972 to 2019.

	var(CF)	var(DR)	-2cov(CF,DR)	corr(CF,DR)
Panel A: Individual Factors				
FIN	89.63%	2.52%	7.85%	-26.14%
PEAD	77.01%	7.35%	15.64%	-32.87%
MISP	92.71%	2.92%	4.37%	-13.27%
SIZE	82.97%	5.60%	11.43%	-26.53%
Panel B: MVE portfolios with market factor				
All Factors	93.91%	10.99%	-4.90%	7.63%
Behavioral Factors	68.25%	22.97%	8.79%	-11.10%
Mispricing Factors	94.95%	15.01%	-9.96%	13.19%
Panel C: MVE portfolios without market factor				
All Factors	98.48%	4.71%	-3.19%	7.40%
Behavioral Factors	85.67%	8.06%	6.27%	-11.93%
Mispricing Factors	94.20%	2.70%	3.10%	-9.74%

Table 4.8: Explanatory Power of Factor News
This table shows the adjusted squares of regressing twelve anomalies unexpected return news separately on different components of market news (Panel A), on different components of behavioral factor model news (Panel B), on different components of mispricing factor model news (Panel C), on different components of all factors' news (Panel D). Row name "CFnews" means that we regress total news of anomaly on the cash flow news components of factors indicated by panels, say in panel B, it implies that we regress unexpected anomaly news on cash flow news of market factor cash flow news, fin factor cash flow news, and pead factor cash flow news.

	lnNS	announcementreturn	compequiss	accruals	noa	assetgrowth	ioa	failureprobability	OScore	mom12m	gp	roaq
Panel A: Anomaly News Explained by Components of Market News												
CFnews	2.88%	-0.79%	4.44%	-1.66%	-1.88%	0.29%	-1.84%	-2.09%	-2.16%	8.29%	-0.15%	-1.21%
DRnews	0.85%	-1.57%	2.13%	-1.86%	0.02%	-2.17%	-2.03%	3.95%	2.61%	2.21%	-0.80%	7.87%
Totalnews	4.37%	-2.13%	6.86%	-2.13%	0.39%	-1.68%	-1.82%	3.66%	1.75%	-1.85%	-1.94%	9.02%
Panel B: Anomaly News Explained by Components of Behavioral Factor News												
CFnews	62.05%	50.55%	31.23%	13.91%	25.53%	40.30%	14.98%	3.02%	19.33%	26.47%	8.20%	13.48%
DRnews	14.66%	11.54%	22.60%	1.53%	13.93%	13.40%	5.00%	2.86%	0.52%	3.88%	-0.14%	5.97%
Totalnews	64.94%	54.34%	38.09%	12.01%	29.63%	43.11%	14.15%	6.57%	19.04%	24.00%	2.69%	20.14%
Panel C: Anomaly News Explained by Components of Mispricing Factor News												
CFnews	37.41%	-1.24%	26.02%	17.09%	9.52%	28.42%	35.16%	31.11%	60.70%	19.29%	33.32%	37.62%
DRnews	3.64%	12.23%	5.40%	-2.10%	1.55%	-4.72%	-4.73%	3.39%	4.67%	1.29%	2.04%	6.87%
Totalnews	32.62%	0.83%	26.32%	16.44%	15.44%	22.52%	32.36%	32.72%	61.69%	11.70%	35.11%	38.63%
Panel D: Anomaly News Explained by Components of All Factor News												
CFnews	66.53%	50.90%	44.05%	19.27%	25.49%	48.38%	37.74%	36.29%	64.90%	53.92%	38.30%	34.88%
DRnews	15.87%	19.04%	24.36%	-0.88%	11.39%	9.34%	0.86%	4.59%	4.30%	0.85%	4.93%	5.47%
Totalnews	69.50%	54.26%	48.96%	19.45%	30.68%	48.37%	37.11%	37.42%	66.94%	54.81%	42.85%	36.37%

Table 4.9: Beta Analysis for PEAD Factor
This table reports firm-level news components of bad beta (i,CFm) and good beta (i,DRm) measured for pead-sorted portfolios. From portfolio 1 to portfolio 3, the pead measure increases. These components are $\beta_{DRi,CFm}$, $\beta_{CFi,CFm}$, $\beta_{DRi,DRm}$, $\beta_{CFi,DRm}$ obtained by regressing corresponding news component of PEAD factor on corresponding news component of the market return.

β	1	2	3	3-1
i,CFm	0.1124	0.1091	0.1458	0.0335
se	0.0483	0.0477	0.0512	0.0191
DRi,CFm	0.0181	0.0170	0.0189	0.0008
se	0.0462	0.0457	0.0460	0.0054
CFi,CFm	0.0943	0.0921	0.1269	0.0326
se	0.0154	0.0156	0.0189	0.0167
i,DRm	0.6480	0.6587	0.6852	0.0372
se	0.0577	0.0518	0.0675	0.0429
DRi,DRm	0.6759	0.6720	0.6740	-0.0019
se	0.0174	0.0140	0.0162	0.0117
CFi,DRm	-0.0278	-0.0134	0.0112	0.0391
se	0.0452	0.0452	0.0582	0.0375

Table 4.10: Beta Analysis for FIN Factor
This table reports firm-level news components of bad beta (i,CFm) and good beta (i,DRm) measured for fin-sorted portfolios. From portfolio 1 to portfolio 3, the fin measure increases. These components are $\beta_{DRi,CFm}$, $\beta_{CFi,CFm}$, $\beta_{DRi,DRm}$, $\beta_{CFi,DRm}$ obtained by regressing corresponding news component of FIN factor on corresponding news component of the market return.

β	1	2	3	3-1
i,CFm	0.0938	0.1186	0.1851	0.0913
se	0.0593	0.0473	0.0611	0.0620
DRi,CFm	0.0085	0.0177	0.0204	0.0119
se	0.0475	0.0457	0.0469	0.0099
CFi,CFm	0.0853	0.1009	0.1647	0.0794
se	0.0362	0.0132	0.0345	0.0589
i,DRm	0.6858	0.6604	0.6990	0.0131
se	0.0865	0.0515	0.1038	0.1386
DRi,DRm	0.6880	0.6714	0.6851	-0.0029
se	0.0222	0.0144	0.0184	0.0220
CFi,DRm	-0.0022	-0.0109	0.0139	0.0161
se	0.0837	0.0434	0.0922	0.1312

Table 4.11: Beta Analysis for MISPRICING Factor
This table reports firm-level news components of bad beta (i,CFm) and good beta (i,DRm) measured for mispricing-sorted portfolios. From portfolio 1 to portfolio 3, the mispricing measure increases. These components are $\beta_{DRi,CFm}$, $\beta_{CFi,CFm}$, $\beta_{DRi,DRm}$, $\beta_{CFi,DRm}$ obtained by regressing corresponding news component of MISPRICING factor on corresponding news component of the market return.

β	1	2	3	3-1
i,CFm	0.1479	0.1235	0.0983	-0.0495
se	0.0423	0.0488	0.0569	0.0399
DRi,CFm	0.0204	0.0180	0.0238	0.0033
se	0.0445	0.0457	0.0474	0.0069
CFi,CFm	0.1274	0.1055	0.0746	-0.0528
se	0.0192	0.0148	0.0298	0.0382
i,DRm	0.5608	0.6718	0.6884	0.1276
se	0.0633	0.0563	0.0788	0.0866
DRi,DRm	0.6544	0.6712	0.6913	0.0369
se	0.0149	0.0150	0.0203	0.0141
CFi,DRm	-0.0936	0.0006	-0.0030	0.0907
se	0.0571	0.0469	0.0695	0.0843

Chapter 5

Appendix

5.1 Replication tables

The tables are my replication of all the tables of Chapter 10 in Bali, Engle, and Murray (2016) with the titles indicating the corresponding tables.

Table A.1: Replication of Table 10.1 in Bali, Engle, and Murray (2016)

	mean	std	skew	kurt	min	5%	25%	median	75%	95%	max	count
BM	0.90	0.86	6.00	109.38	0.01	0.15	0.41	0.72	1.14	2.22	18.64	3357
lnBM	-0.49	0.86	-0.75	2.43	-6.09	-1.98	-0.96	-0.40	0.07	0.73	2.68	3357
BE	1028	4702	17	453	1	11	50	158	539	3909	129749	3357
ME	2060	9752	15	342	2	17	77	272	1032	7265	246203	3357
Mktcap	1249	6111	15	332	1	7	34	129	535	4598	159004	3356

Table A.2: Replication of Table 10.1 in Bali, Engle, and Murray (2016) Using Exactly the Same Procedure Described in the Book

	mean	std	skew	kurtosis	min	5%	25%	median	75%	95%	max	count
BM	0.94	1.13	9.27	239.53	0.01	0.15	0.42	0.72	1.15	2.31	31.18	3393
lnBM	-0.47	0.87	-0.66	2.42	-6.09	-1.97	-0.95	-0.39	0.08	0.77	3.02	3393
BE	1043.78	4796.39	16.94	448.32	0.82	11.20	50.22	159.17	544.44	3950.07	129749.12	3393
ME	2039.75	9666.34	15.48	346.93	2.31	16.98	76.79	270.21	1024.87	7225.90	246203.29	3393
Mktcap	1232.37	6035.45	15.27	337.19	0.67	6.62	33.58	127.87	529.61	4567.99	159003.98	3392

Table A.3: Replication of Table 10.2 in Bali, Engle, and Murray (2016)

	BM	lnBM		size
BM		1.00	-0.23	-0.26
lnBM	0.85		-0.23	-0.26
β	-0.20	-0.22		0.32
Size	-0.28	-0.24	0.30	

Table A.4: Replication of Table 10.3 in Bali, Engle, and Murray (2016)

τ	BM	lnBM
12	0.778	0.812
24	0.633	0.686
36	0.537	0.599
48	0.471	0.540
60	0.423	0.491
120	0.324	0.382

Table A.5: Replication of Table 10.4 in Bali, Engle, and Murray (2016)

Panel A: NYSE/AMEX/NASDAQ Breakpoints												
	Value	1	2	3	4	5	6	7	8	9	10	10-1
Characteristics	BM	0.15	0.29	0.41	0.53	0.65	0.79	0.94	1.14	1.46	2.65	
	lnBM	-2.17	-1.32	-0.96	-0.71	-0.49	-0.31	-0.12	0.07	0.31	0.83	
	Mktcap	2195	2223	1698	1437	1295	1038	937	816	619	283	
	β	1.06	0.98	0.91	0.85	0.81	0.78	0.74	0.70	0.66	0.60	
	%NYSE	29.05%	35.87%	40.78%	43.60%	45.93%	47.45%	45.68%	44.43%	40.51%	32.53%	
	n	334	334	333	334	334	333	333	334	333	334	
EW portfolios	Excess return	0.01	0.33	0.45	0.57	0.70	0.78	0.89	0.93	1.07	1.35	1.35
		(0.02)	(1.05)	(1.59)	(2.12)	(2.59)	(3.04)	(3.47)	(3.51)	(3.98)	(4.03)	(6.24)
	CAPM α	-0.64	-0.28	-0.12	0.04	0.18	0.29	0.42	0.46	0.60	0.88	1.51
		(-3.25)	(-1.79)	(-0.84)	(0.29)	(1.33)	(2.17)	(3.06)	(3.14)	(3.82)	(4.04)	(6.77)
VW portfolios	Excess return	0.31	0.32	0.44	0.44	0.43	0.50	0.62	0.58	0.74	0.84	0.53
		(1.2)	(1.51)	(2.2)	(2.11)	(2.14)	(2.64)	(3.15)	(2.73)	(3.77)	(3.36)	(2.25)
	CAPM α	-0.19	-0.16	-0.01	-0.02	0.00	0.09	0.22	0.17	0.31	0.36	0.55
		(-1.54)	(-2.1)	(-0.19)	(-0.34)	(-0.02)	(1.04)	(2.09)	(1.38)	(2.47)	(2.39)	(2.27)
Panel B: NYSE Breakpoints												
	Value	1	2	3	4	5	6	7	8	9	10	10-1
Characteristics	BM	0.18	0.35	0.47	0.58	0.69	0.80	0.93	1.09	1.35	2.37	
	lnBM	-1.88	-1.12	-0.82	-0.60	-0.43	-0.28	-0.12	0.04	0.24	0.73	
	Mktcap	2383	1941	1607	1341	1223	1085	945	868	673	315	
	/beta	1.05	0.94	0.88	0.83	0.80	0.76	0.73	0.70	0.67	0.61	
	%NYSE	32.23%	39.27%	43.33%	45.05%	47.16%	48.02%	47.93%	45.53%	42.23%	34.05%	
	n	514	353	313	292	280	273	273	284	315	425	
EW portfolios	Excess return	0.17	0.49	0.60	0.59	0.78	0.79	0.83	0.90	1.04	1.27	1.27
		(0.52)	(1.69)	(2.17)	(2.15)	(2.92)	(3.05)	(3.35)	(3.54)	(3.91)	(3.95)	(5.83)
	CAPM α	-0.46	-0.09	0.05	0.06	0.27	0.30	0.37	0.44	0.57	0.80	1.25
		(-2.65)	(-0.62)	(0.35)	(0.44)	(1.97)	(2.25)	(2.85)	(3.14)	(3.78)	(3.94)	(6.4)
VW portfolios	Excess return	0.36	0.47	0.46	0.40	0.46	0.49	0.62	0.62	0.63	0.83	0.47
		(1.53)	(2.32)	(2.3)	(1.89)	(2.41)	(2.58)	(3.09)	(3.14)	(3.13)	(3.49)	(2.3)
	CAPM α	-0.12	0.00	0.00	-0.06	0.05	0.09	0.22	0.21	0.21	0.35	0.48
		(-1.31)	(0.04)	(0.04)	(-0.72)	(0.52)	(0.97)	(2.02)	(1.83)	(1.84)	(2.5)	(2.25)

Table A.6: Replication of Table 10.5 in Bali, Engle, and Murray (2016)

Panel A: BM Difference Portfolios								
Control	Weights	Value	Control1	Control2	Control3	Control4	Control5	ControlAvg
β	EW	Return	0.90	0.79	0.75	0.87	1.28	0.92
			(5.06)	(4.74)	(5.02)	(5.26)	(6.85)	(6.56)
		CAPM α	0.95	0.84	0.78	0.91	1.31	0.96
			(5.28)	(5.06)	(5.17)	(5.37)	(7.03)	(6.76)
	VW	Return	0.52	0.47	0.36	0.43	0.39	0.43
			(2.51)	(2.76)	(2.22)	(2.23)	(1.69)	(3.14)
		CAPM α	0.47	0.51	0.34	0.43	0.41	0.43
			(2.28)	(3.02)	(2.09)	(2.18)	(1.72)	(3.07)
Mktcap	EW	Return	0.80	1.19	0.94	0.69	0.28	0.78
			(3.41)	(5.17)	(4.07)	(3.09)	(1.41)	(4.11)
		CAPM α	0.95	1.33	1.11	0.87	0.43	0.94
			(4.14)	(5.76)	(4.97)	(3.94)	(2.17)	(5.02)
	VW	Return	0.90	1.15	0.89	0.68	0.17	0.76
			(3.77)	(4.87)	(3.84)	(3.09)	(0.91)	(4.03)
		CAPM α	1.04	1.29	1.07	0.86	0.24	0.90
			(4.38)	(5.41)	(4.74)	(3.93)	(1.24)	(4.78)
Panel B: Avarage Control Variable Portfolios								
Control	Weights	Value	BM1	BM2	BM3	BM4	BM5	BM5-1
β	EW	Return	0.26	0.57	0.70	0.86	1.18	0.92
			(0.84)	(2.09)	(2.68)	(3.26)	(3.74)	(6.56)
		CAPM α	-0.31	0.04	0.18	0.36	0.65	0.96
			(-1.87)	(0.3)	(1.51)	(2.7)	(3.61)	(6.76)
	VW	Return	0.32	0.39	0.47	0.57	0.75	0.43
			(1.54)	(2.0)	(2.47)	(3.02)	(3.45)	(3.14)
		CAPM α	-0.12	-0.04	0.05	0.17	0.31	0.43
			(-1.84)	(-0.74)	(0.87)	(2.14)	(2.78)	(3.07)
Mktcap	EW	Return	0.27	0.60	0.80	0.86	1.05	0.78
			(0.79)	(2.07)	(2.99)	(3.35)	(3.85)	(4.11)
		CAPM α	-0.36	0.05	0.29	0.39	0.58	0.94
			(-1.87)	(0.33)	(2.17)	(2.86)	(3.74)	(5.02)
	VW	Return	0.17	0.48	0.68	0.75	0.92	0.76
			(0.5)	(1.76)	(2.64)	(3.09)	(3.49)	(4.03)
		CAPM α	-0.44	-0.06	0.18	0.30	0.46	0.90
			(-2.48)	(-0.43)	(1.47)	(2.39)	(3.08)	(4.78)

Table A.7: Replication of Table 10.6 in Bali, Engle, and Murray (2016)

Panel A: Equal-Weighted Portfolios								
	β1	β2	β3	β4	β5	βAvg	β5-1	β5-1CAPMα
BM1	0.36	0.35	0.41	0.19	0.01	0.26	-0.35	-0.72
							(-1.59)	(-3.26)
BM2	0.51	0.57	0.60	0.59	0.48	0.55	-0.03	-0.43
							(-0.16)	(-2.2)
BM3	0.84	0.65	0.75	0.81	0.74	0.76	-0.10	-0.52
							(-0.46)	(-2.47)
BM4	0.83	0.89	0.89	0.90	0.92	0.89	0.09	-0.35
							(0.38)	(-1.67)
BM5	1.19	1.19	1.26	1.14	1.18	1.19	-0.01	-0.41
							(-0.06)	(-1.84)
BMAvg	0.75	0.73	0.78	0.73	0.66		-0.08	-0.49
							(-0.41)	(-2.68)
BM5-1	0.83	0.84	0.85	0.95	1.17	0.93		
	(3.99)	(4.48)	(5.06)	(5.53)	(5.5)	(6.15)		
BM 5-1	0.91	0.89	0.88	0.99	1.22	0.98		
CAPMα	(4.52)	(4.79)	(5.2)	(5.59)	(5.74)	(6.44)		

Panel B: Value-Weighted Portfolios								
	β1	β2	β3	β4	β5	βAvg	β5-1	β5-1CAPMα
BM1	0.34	0.27	0.41	0.25	0.21	0.29	-0.13	-0.52
							(-0.43)	(-1.94)
BM2	0.22	0.46	0.38	0.60	0.32	0.40	0.09	-0.26
							(0.39)	(-1.17)
BM3	0.45	0.41	0.50	0.45	0.58	0.48	0.13	-0.25
							(0.53)	(-1.04)
BM4	0.59	0.56	0.61	0.62	0.62	0.60	0.03	-0.40
							(0.12)	(-1.98)
BM5	0.75	0.72	0.88	0.83	0.60	0.76	-0.16	-0.58
							(-0.59)	(-2.2)
BMAvg	0.47	0.48	0.55	0.55	0.47		-0.01	-0.40
							(-0.02)	(-2.11)
BM5-1	0.42	0.45	0.47	0.58	0.39	0.46		
	(1.81)	(2.66)	(2.85)	(2.99)	(1.62)	(3.34)		
BM 5-1	0.42	0.49	0.46	0.57	0.35	0.46		
CAPMα	(1.82)	(2.97)	(2.76)	(2.83)	(1.46)	(3.3)		

Table A.8: Replication of Table 10.7 in Bali, Engle, and Murray (2016)

Panel A: Equal-Weighted Portfolios

	Mktcap 1	Mktcap 2	Mktcap 3	Mktcap 4	Mktcap 5	Mktcap Avg	Mktcap 5-1	Mktcap 5-1 CAPM α
BM1	1.02	-0.18	-0.01	0.25	0.37	0.29	-0.65	-0.57
							(-1.68)	(-1.55)
BM2	1.02	0.30	0.51	0.56	0.49	0.58	-0.52	-0.49
							(-1.69)	(-1.61)
BM3	1.31	0.73	0.69	0.66	0.58	0.79	-0.73	-0.67
							(-2.59)	(-2.49)
BM4	1.36	0.71	0.86	0.89	0.62	0.89	-0.74	-0.69
							(-2.96)	(-2.84)
BM5	1.62	0.97	0.88	0.91	0.71	1.02	-0.90	-0.90
							(-3.36)	(-3.34)
BMAvg	1.27	0.51	0.59	0.65	0.56		-0.71	(-0.66)
							(-2.58)	(-2.51)
BM5-1	0.60	1.15	0.89	0.67	0.34	0.73		
	(2.2)	(4.9)	(3.71)	(2.87)	(1.76)	(3.68)		
BM 5-1	0.79	1.30	1.06	0.82	0.46	0.88		
CAPM α	(3.03)	(5.57)	(4.5)	(3.44)	(2.37)	(4.55)		

Panel B: Valuel-Weighted Portfolios

	Mktcap 1	Mktcap 2	Mktcap 3	Mktcap 4	Mktcap 5	Mktcap Avg	Mktcap 5-1	Mktcap 5-1 CAPM α
BM1	0.36	-0.14	0.04	0.26	0.35	0.17	-0.01	0.13
							(-0.02)	(0.36)
BM2	0.50	0.31	0.50	0.55	0.43	0.46	-0.07	0.02
							(-0.23)	(0.06)
BM3	0.81	0.72	0.71	0.64	0.43	0.66	-0.38	-0.27

Table A.8 continued from previous page

							(-1.33)	(-1.03)
BM4	0.92	0.72	0.86	0.88	0.51	0.78	-0.41	-0.34
							(-1.56)	(-1.33)
BM5	1.16	0.96	0.89	0.90	0.61	0.90	-0.55	-0.53
							(-2.12)	(-2.08)
BMAvg	0.75	0.51	0.60	0.65	0.46		-0.28	-0.20
							(-1.02)	(-0.76)
BM5-1	0.80	1.10	0.85	0.64	0.25	0.73		
	(2.98)	(4.63)	(3.48)	(2.75)	(1.32)	(3.67)		
BM 5-1	0.97	1.25	1.02	0.78	0.31	0.86		
CAPM α	(3.74)	(5.23)	(4.28)	(3.33)	(1.55)	(4.41)		

Table A.9: Replication of Table 10.8 in Bali, Engle, and Murray (2016)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BM	0.48	0.41	0.34	0.27				
	(5.63)	(5.46)	(3.75)	(3.67)				
lnBM					0.44	0.40	0.33	0.29
					(6.3)	(6.58)	(4.07)	(4.51)
β		-0.11		0.03		-0.07		0.08
		(-0.85)		(0.17)		(-0.53)		(0.49)
Size			-0.12	-0.13			-0.12	-0.13
			(-2.37)	(-2.2)			(-2.29)	(-2.25)
Intercept	0.31	0.47	0.99	1.07	0.93	0.96	1.42	1.42
	(1.1)	(2.09)	(1.98)	(2.48)	(3.4)	(4.08)	(3.14)	(3.46)
Adj. R^2	0.01	0.02	0.02	0.04	0.01	0.02	0.02	0.04
n	3356	3337	3356	3337	3356	3337	3356	3337

5.2 Note on intangible capital

5.2.1 Why do we need to estimate intangible assets

Accounting principle

In current accounting standards, the conservatism principle prevents internally generated intangible capital from being recorded in the balance sheets. Conservatism requires company accounts to be prepared with caution and high degrees of verification. Intangible assets, as the name implies, are not physical in nature and are not financial instruments, making it hard to measure their value. Though current accounting items do have intangible items like patent, brand, trademark, copyright listed under long-term assets on the balance sheet, they are actually mostly acquired from firms outside, and the amount is the cost to buy those externally generated items. Only neglectable direct costs incurred in developing the intangible, such as legal costs, are capitalized, while the rest are all expensed. From now on, we regard this item (INTAN) as the value of externally obtained intangible assets. Goodwill, which is the excess price higher than the sum of the fair value of another firm being bought, is also recorded as an intangible asset. On the other hand, if firms plan to generate intangible assets inside, like developing a new system, machine, medicine, applying for a new patent, and so on, the expenses from this process will be expensed, which decreases the net book value of the firm. But the intangible assets shall not be recorded in the accounting.

While we cannot deny the revenue and importance of intangible assets, the process of creating them faces too much uncertainty or risk, and they do not usually have a fair value. Allowing the inclusion of internally generated intangible assets tends to cause an overestimation of their value since the recording of these intangible assets is susceptible to the manipulation of management, as the managers have the incentive to present perfect, though inappropriate, accounting reports. This approach would, as a result, seriously impair the quality of reports and market confidence in making investment choices based on these reports. It may justify the reasons why current accounting principles still impede the listing of internally generated intangible assets.

The need for an estimation

In spite of their absence in financial reports, it is an indisputable fact that intangible assets have become increasingly important nowadays. In practice, while traditional firms mainly rely on plant, property and equipment as their main assets, more and more firms like Amazon, Alibaba, Facebook, Apple, Dell, and other technology or e-commerce firms nowadays take intangible assets as their crucial assets. Human resources, technology updates, and new platform-based transactions, like web-based transactions, are the essential competitive power among firms. In theory, many models endeavor to integrate intangible capital in addition to traditional physical capital. Besides, what comes to the topic of this note is the measure of book-to-market ratio and its affiliated value strategy.

The value strategy, which can be simply understood as being long in stocks with a high book-to-market ratio and short in stocks with a low book-to-market ratio, has experienced persistent and large drawdowns in recent decades. One possible explanation for this phenomenon is that the current accounting principle prevents us from recording and calculating book value precisely as the costs used to develop internally generated intangible assets are expensed, reducing firms' assets and hence the book value of equity. But we know that employee training, research and development costs, among others, usually bring potential future earnings and shall also be considered to be part of firms' assets. Therefore, some researchers propose to estimate internally generated intangible assets and adjust the book value accordingly.

5.2.2 Literature review on intangible assets

Intangible capital is estimated to take up to about 50% of firms' capital stock (see Falato, Kadyrzhanova, and Sim (2013)). Though it is now raising more and more people's attention, there is no strict and explicit definition for it. The initial awareness of intangible capital is mainly associated with the emergence of firms' acquisitions of intangibles and firms' activities of research and development (see Griliches (1979)). Some researchers understand it mainly from the perspective of human resources like high-skilled staff (key talent, organization capital) and inputs like equity-based compensation to maintain this labor (see Corrado, Hulten, and Sichel (2009), Eisfeldt and Papanikolaou (2013), Eisfeldt, Falato, and Xiaolan (2021)). Some focus on information technology (digital capital) (see Tambe, Hitt, Rock, and Brynjolfsson (2020)). Others include both human capital and technology development.

Below I offer a literature review classified into three aspects: intangible assets' effects on the macroeconomy, their application in corporate finance, and asset pricing.

Macroeconomy

This strand of papers mainly focuses on the inclusion of intangible capital into classical models and explores its importance and influences on the macroeconomy.

Corrado, Hulten, and Sichel (2009) add the estimated aggregate intangible capital stocks to the standard sources-of-growth framework used by the BLS and find that this inclusion makes a significant difference in the observed patterns of US economic growth. Lall and Zeng (2020) incorporate the intangible capital into the standard AS-AD framework to explain the causes underlying low inflation after the global financial crisis and highlight the possibility that technological change and a large portion of intangible investment lead to wage stagnation and greater market concentration. Gareis and Mayer (2022) develop an extended real business cycle model with intangible capital and study the relative dynamics of tangible and intangible investment in response to financial shocks. Döttling and Ratnovski (2023) document that the stock prices of firms with more intangible assets react less to monetary policy shocks show that the total investment in firms with more intangible assets responds less to monetary policy shocks than tangible assets and investments. Chiavari and Goraya

(2020) augment a standard production function with intangible capital and find that its input share has increased at the expense of labor in production. Intangible capital entails higher investment adjustment costs than traditional capital and tends to be misallocated. They further show that the shift of input in production can explain some of the major trends for the US economy, such as the rising average firm size.

Corporate finance

When it comes to the capital structure of firms, intangible capital is examined for its adjustment costs and its resistance to financial constraints, most often from the wage, equity compensation, and human resources aspects.

Sun and Xiaolan (2019) provide a theoretical and empirical analysis of firms' dynamic capital structure decisions in the presence of intangible capital accumulation and find that the intangible capital overhang effect dominates the precautionary effect. Eisfeldt and Papanikolaou (2013) argue that organization capital is a production factor embodied in the firm's key talent and that both shareholders and key talent have a claim to its cash flows. They develop a model and document that firms with more intangible capital have higher average returns as the outside option of key talent determines the share of firm cash flow that accrues to shareholders and thus represents a higher risk from a shareholders' perspective.

For the neoclassical theory of investment, Peters and Taylor (2017) show that Tobin's Q also explains intangible investment in a better way than it explains tangible investment. They suggest a simple, new Tobin's Q proxy that accounts for intangible capital, and they show that it is a superior proxy for both physical and intangible investment opportunities. The new proxy is just to adjust the denominator, the book value of equity, by adding intangible assets.

Asset pricing

Research relevant to asset pricing either modifies and develops new asset pricing models by combining intangible capital, or revises the book-to-market ratio to remedy the value strategy.

Ahn (2016) proposes an investment-based asset pricing model augmented with intangible capital and a transient volatility shock. The author argues that the intangible capital mitigates the negative impact of temporary volatility shock on output and thus, the physical-capital-intensive value firms are more exposed to volatility risk and require more premium (which is contrary to the conclusion of Eisfeldt and Papanikolaou (2013)).

Gulen, Li, Peters, and Zekhnini (2020) include intangible capital in the factor model. Several recent papers apply intangible capital to adjust the book-to-market ratio and show that it improves the performance of the value strategy significantly (see Eisfeldt, Kim, and Papanikolaou (2022), Arnott, Harvey, Kalesnik, and Linnainmaa (2021), Amenc, Goltz, and Luyten (2020), Park (2022)). Dugar and Pozharny (2021) show that the relationship between financial variables and contemporaneous stock prices has weakened for high intangible intensity companies.

5.2.3 An overview of estimation methods

Corrado, Hulten, and Sichel (2009) use the input-based approach to estimate the investment in certain broad groups of business intangibles and then estimate the intangible capital at the aggregate level. Squicciarini and Le Mouel (2012) use the microdata to develop a task-based approach to quantify investment in organizational capital and find that previous measures seemingly underestimate investment in organizational capital at the macro level.

At the firm level, researchers usually capitalize relevant expenditures to estimate the intangible capital, which will be discussed in detail in this note (see Lev and Sougiannis (1996), Eisfeldt and Papanikolaou (2013), Peters and Taylor (2017), Park (2022)). Lev and Radhakrishnan (2009) simultaneously estimate the production function and the selling, general administrative expenses to get the annual extra revenues, the difference between the predicted revenue with and without organization capital. Then they capitalize on the annual extra revenue to get an estimate of organizational capital. When estimating the firms' missing internally generated intangible assets, researchers can impute intangible capital as a latent variable based on other moments or use inventory or exploit asset prices to compute intangible capital directly. The most popular one applied in the field of value strategy is the perpetual inventory method described as follows:

$$X_{t+1} = (1-\delta)X_t + \gamma \cdot G_t$$

where X_t is the intangible capital stock, G_t is the relevant expenditures, δ is the depreciation rate and γ is the transformation rate. The diversification among the estimates employed by different authors mainly lies in the selection of initial stock X_0, further classification of X_{it} into knowledge K_{it} and organization capital O_{it}, and the parameters δ, γ.

5.2.4 Preliminary on Compustat items

When calculating the intangible adjusted book value of equity, we use several items from Compustat:

- Those related to the book value of equity defined by Fama and French are parent stockholders' equity (SEQ), deferred taxes (TXDB), investment tax credit (ITCB), and the book value of preferred stocks (PRTKRV, PSTKL, PSTK).
- Those related to intangible assets are total intangible assets (INTAN), goodwill (GDWL) which is part of INTAN, selling, general and administrative expense (XSGA), and research and development expense (XRD)

INTAN item contains mostly the value of externally obtained intangible capital but a negligible portion of internally generated one. So we usually regard it as the value of externally acquired intangible capital. Besides, as it is under long-term assets which is part of the book value of equity, it is not separately listed or cited to calculate the intangible adjusted book-to-market ratio.

XSGA item displays the costs not directly related to making a product or performing a service, like indirect selling expenses such as advertising, marketing, telephone bills, travel costs, salaries of staff, commissions, utilities, and rent, which are not part of manufacturing. Its portion of revenue is usually highest among the health care and financial industries. This item is essential for the estimation of organization capital as it can be viewed as a proxy for investment in human capital, brand, customer relationships, etc.

XRD item displays the expenses relevant to the process of exploring and creating new products, services, and technologies. Tech firms incur large expenses on this item as they develop facial recognition and AI techniques.

One thing calls for our attention: while firms typically report SG&A and R&D expenses separately, Compustat, however, almost always adds them together under item XSGA. Therefore, when dealing with knowledge capital and organization capital separately, we need to subtract XRD from XSGA.

Another issue with Compustat records is that it incorporates firms not as soon as firms are founded but may contain a several-year gap between a firm's foundation and the same firm's appearance in Compustat.

5.2.5 Three measures of internally generated intangible assets

Here I first offer a sketch of the differences among the three measures and then detail their processes to estimate intangible assets one by one.

Peters and Taylor (2017) and Ewens, Peters, and Wang (2020) use the sampling method to account for the initial value of intangible assets and accumulate organization (O_{it}) and knowledge capital (K_{it}) separately but use a different set of parameters. Eisfeldt, Kim, and Papanikolaou (2022) on the other hand, initialize internally generated intangible assets as $\frac{SG\&A}{0.3}$ where SG&A is the first observation for selling and general administrative expenses when the firm appears in Compustat for the first time, and they do not consider research and development expenses separately in their main measure. The authors in their paper sometimes consider alternative approaches, like if goodwill should be included, if expenses on research and development should be separately adjusted, or if other often applied parameters should be used, and they argue that their main results are robust to these variations. Given the length of this note, only the ways to construct their main measures of internally generated intangible assets are used here.

Eisfeldt, Kim, and Papanikolaou (2020)'s estimation of intangible assets:

$$Int_{it}^{Eisfeldt} = 0.8 \cdot Int_{i,t-1}^{Eisfeldt} + SG\&A$$

where SG&A is Compustat item XSGA, replaced by 0 if missing. The initial stock is assumed to be $\frac{SG\&A_1}{0.3}$ with $SG\&A_1$ is the first record of this firm in Compustat under item XSGA. Therefore, $Int_{i1}^{Eisfeldt}$, the first estimation in the first record of this firm would be $0.8 \cdot \frac{SG\&A_1}{0.3} + SG\&A_1$.

As the authors note in their paper, goodwill is subtracted. Then the overall intangible capital used to adjust the book equity of Fama and French in the same fiscal year is:

$$Intit = Int_{it}^{Eisfeldt} - GDWL$$

Peters and Taylor (2017)'s estimation of intangible assets:

As pointed out above, Peter and Taylor (2017) and Ewens, Peters, and Wang (2020) further divide internally generated intangible capital into knowledge capital and organization capital.

$$K_{it} = (1 - \delta_{Li,2012} K_{i,t-1} + R\&D)$$

$$O_{it} = 0.8 \cdot O_{i,t-1} + 0.3 \cdot SG\&A$$

$$Int_{it}^{Peters} = K_{it} + O_{it}$$

where R&D is Compustat item XRD. SG&A is $XSGA - XRD - RDIP$ (XRD, RDIP should be replaced by 0 if missing), or XSGA itself replaced by 0 if missing when $COGS > XRD > XSGA$. The detailed argument for the implementation can be found in the original paper. The table for $\delta_{Li,2012}$ is attached at the end.

The adjustment for missing XRD and XSGA for existing records in Compustat is subtle: Peters and Taylor set them to 0 when missing except for the years when the firm's assets are also missing. For these years, they interpolate these two variables using their nearest non-missing values.

The estimation of firms' initial capital stock is a little bit tricky. Basically speaking, it includes the following steps:
1, Estimate AgeSinceIPO-specific (age) growth rates and PreIPO growth rates for R&D and SG&A;
2, Use the estimated growth rates to estimate R&D and SG&A for years when firms are founded yet not listed in Compustat;
3, Assume the firm is founded with no intangible capital and apply the perpetual inventory method to the estimated statistics to calculate the stock of intangible capital at the first Compustat record.

The detailed process for estimating the initial knowledge-capital stock below is copied from Peters and Taylor (2017). The method for organization capital is similar:
1, Define age since IPO as the number of years that have elapsed since a firm's IPO. Using the full Compustat database, compute the average log change in R&D in each yearly category of age since IPO. Apply these age-specific growth rates to fill in missing R&D observations before 1977.
2, Using the full Compustat database, isolate records for firms' IPO years and the previous two years. (Not all firms have pre-IPO data in Compustat.) Compute the average log change in R&D within this pre-IPO subsample, which equals 0.348. (The corresponding pre-IPO average log change in SG&A equals 0.333.)
3. If firm i's IPO year is in Compustat, go to Step 5. Otherwise, go to the next step.
4. This step applies almost exclusively to firms with IPOs before 1950. Estimate firm i's R&D spending in each year between the firm's IPO year and first Compustat year, assuming the firm's R&D grows at the average age-specific rates estimated in Step 1.
5. Obtain data on the firm i's founding year from Jay Ritter's website. For firms with missing founding year, estimate the founding year as the minimum of (a) the year of the firm's first Compustat record and (b) the firm's IPO year minus eight, which is the median age between founding and IPO for IPOs from 1980 to 2012 (from Jay Ritter's website).
6. Estimate the firm i's R&D spending in each year between the firm's founding year and IPO year assuming the firm's R&D grows at the estimated pre-IPO average rate from Step 2.
7. Assume the firm was founded with no capital. Apply the perpetual inventory method to the estimated R&D spending from the previous steps to obtain K_{i0}, the stock of knowledge capital at the beginning of the firm's first Compustat record.

As explained above, INTAN is a long-term asset and already included in the book equity of Fama and French, then we simply add Int_{it}^{Peters} to the book equity of the firm i in the same fiscal year t to get the intangible adjusted book value of equity.

Ewens, Peters, and Wang (2020)'s estimation of intangible assets

The only difference between Ewens, Peters, and Wang (2020) and Peters and Taylor (2017) is the parameter selection. The computation of initial stock capital and the measure for the current period's inputs are exactly the same.

$$K_{it} = (1 - \delta_{Ewens})K_{i,t-1} + R\&D$$

$$O_{it} = 0.8 \cdot O_{i,t-1} + \gamma_{Ewens} \cdot SG\&A$$

$$Int_{it}^{Ewens} = K_{it} + O_{it}$$

The table for δ_{Ewens} and γ_{Ewens} is attached at the end. We only need to add Int_{it}^{Ewens} to obtain the intangible adjusted book value of equity.

Table A.10: The Parameters Used by Peters and Taylor (2017)

$\delta_{Li,2012}$, R&D depreciate rate for estimating knowledge capital		
Industry	SIC codes	$\delta_{Li,2012}$
Computers and peripheral equipment	3570-3579, 3680-3689 and 3695	40%
Software	7372	22%
Pharmaceuticals	2830, 2831 and 2833 - 2836	10%
Semiconductor	3661-3666 and 3669-3679	25%
Aerospace product and parts	3720, 3721, 3724, 3728 and 3760	22%
Communication equipment	3576, 3661, 3663, 3669 and 3679	27%
Computer system design	7370, 7371 and 7373	36%
Motor vehicles, bodies, trailers, and parts	3585, 3711, 3713 and 3716	31%
Navigational, measuring, electromedical, and control instruments	3812, 3822, 3823, 3825, 3826, 3829, 3842, 3844 and 3845	29%
Scientific research and development	8731	16%
Others	Others	15%

Table A.11: The Parameters Used by Ewens, Peters, and Wang (2020)

$\delta_{Ewens}, \gamma_{Ewens}$			
Industry	SIC codes	δ_{Ewens}	γ_{Ewens}
Consumer	0100-0999, 2000-2399, 2700-2749, 3100-3199, 3940-3989, 2500-2519, 2590-2599, 3630-3659, 3710-3711, 3714, 3716, 3750, 3751, 3792, 3900-3999, 5000-5999, 7200-7299, 7600-7699, 8000-8099, 4813, 4812, 4841, 4833, 4832	0.33	0.19
Manufacturing	2520-2589, 2600-2699, 2750-2769, 2800-2829, 3000-3099, 3200-3569, 3580-3621, 3623-3629, 3700-3709, 3712-3713, 3715, 3717-3749, 3752-3791, 3793-3799, 3860-3899, 1200-1399, 2900-2999, 4900-4949	0.42	0.22
High Tech	3570-3579, 3622, 3660-3692, 3694-3699, 3810-3839, 7370-7379, 7391, 8730-8734, 4800-4899	0.46	0.44
Health	2830-2839, 3693, 3840-3859	0.34	0.49
Other	Other	0.3	0.34

References

Ahn, Yongkil. 2016. “Capital heterogeneity, volatility shock, and the value premium.” In *Paris December 2016 Finance Meeting EUROFIDAI-AFFI*.

Akbas, Ferhat, Ekkehart Boehmer, Bilal Erturk, and Sorin Sorescu. 2017. “Short interest, returns, and unfavorable fundamental information.” *Financial Management* 46 (2): 455–486.

Ali, Ashiq, Lee-Seok Hwang, and Mark A Trombley. 2003. “Arbitrage risk and the book-to-market anomaly.” *Journal of Financial Economics* 69 (2): 355–373.

Amenc, Noël, Felix Goltz, and Ben Luyten. 2020. “Intangible capital and the value factor: Has your value definition just expired?” *The Journal of Portfolio Management* 46 (7): 83–99.

Ang, Andrew. 2022. “Trends and cycles of style factors in the 20th and 21st centuries.” *Available at SSRN*.

Antolin-Diaz, Juan, Ivan Petrella, and Juan Rubio Ramírez. 2021. “Dividend momentum and stock return predictability: A bayesian approach.” *WBS Finance Group Research Paper*.

Arnold, Tom, Alexander W Butler, Timothy Falcon Crack, and Yan Zhang. 2005. “The information content of short interest: A natural experiment.” *The Journal of Business* 78 (4): 1307–1336.

Arnott, Robert D, Campbell R Harvey, Vitali Kalesnik, and Juhani T Linnainmaa. 2021. “Reports of value’s death may be greatly exaggerated.” *Financial Analysts Journal* 77 (1): 44–67.

Asness, Clifford, and Andrea Frazzini. 2013. “The devil in hml’s details.” *The Journal of Portfolio Management* 39 (4): 49–68.

Asquith, Paul, and Lisa K Meulbroek. 1995. *An empirical investigation of short interest*. Division of Research, Harvard Business School.

Asquith, Paul, Parag A Pathak, and Jay R Ritter. 2005. “Short interest, institutional ownership, and stock returns.” *Journal of Financial Economics* 78 (2): 243–276.

Bali, T.G., R.F. Engle, and S. Murray. 2016. *Empirical Asset Pricing: The Cross Section of Stock Returns*. Online access with DDA: Askews. Wiley.

Barillas, Francisco, and Jay Shanken. 2018. “Comparing asset pricing models.” *The Journal of Finance* 73 (2): 715–754.

Barr Rosenberg, Kenneth Reid, and Ronald Lanstein. 1984. “Persuasive evidence of market inefficiency.” *Journal of Portfolio Management* 11: 9–17.

Basu, Sanjoy. 1977. "Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis." *The Journal of Finance* 32 (3): 663–682.

Beneish, Messod Daniel, Charles MC Lee, and D Craig Nichols. 2015. "In short supply: Short-sellers and stock returns." *Journal of Accounting and Economics* 60 (2-3): 33–57.

Bhandari, Laxmi Chand. 1988. "Debt/equity ratio and expected common stock returns: Empirical evidence." *The Journal of Finance* 43 (2): 507–528.

van Binsbergen, Jules, Michael Brandt, and Ralph Koijen. 2012. "On the timing and pricing of dividends." *American Economic Review* 102 (4): 1596–1618.

Boehmer, Ekkehart, Zsuzsa R Huszar, and Bradford D Jordan. 2010. "The good news in short interest." *Journal of Financial Economics* 96 (1): 80–97.

Böll, Julian, Julian Thimme, and Marliese Uhrig-Homburg. 2022. "Anomalies and optionability." *Available at SSRN 4300137*.

Bryzgalova, Svetlana, Jiantao Huang, and Christian Julliard. 2023. "Bayesian solutions for the factor zoo: We just ran two quadrillion models." *The Journal of Finance* 78 (1): 487–557.

Callen, Jeffrey L, and Dan Segal. 2004. "Do accruals drive firm-level stock returns? a variance decomposition analysis." *Journal of Accounting Research* 42 (3): 527–560.

Callen, Jeffrey L, and Dan Segal. 2010. "A variance decomposition primer for accounting research." *Journal of Accounting, Auditing & Finance* 25 (1): 121–142.

Campbell, John Y. 1991. "A variance decomposition for stock returns." *The Economic Journal* 101 (405): 157–179.

Campbell, John Y. 2008. "Estimating the equity premium." *Canadian Journal of Economics/Revue canadienne d'économique* 41 (1): 1–21.

Campbell, John Y, Stefano Giglio, and Christopher Polk. 2023. "What drives booms and busts in value?" *Available at SSRN 4391054*.

Campbell, John Y, and Jianping Mei. 1993. "Where do betas come from? asset price dynamics and the sources of systematic risk." *The Review of Financial Studies* 6 (3): 567–592.

Campbell, John Y, Christopher Polk, and Tuomo Vuolteenaho. 2010. "Growth or glamour? fundamentals and systematic risk in stock returns." *The Review of Financial Studies* 23 (1): 305–344.

Campbell, John Y, and Robert J Shiller. 1988. "The dividend-price ratio and expectations of future dividends and discount factors." *The Review of Financial Studies* 1 (3): 195–228.

Campbell, John Y, and Tuomo Vuolteenaho. 2004. "Bad beta, good beta." *American Economic Review* 94 (5): 1249–1275.

Chaves, DB. 2009. "What explains the variance of prices and returns." *Time-series Vs. Crosssection. SSRN eLibrary*.

Chen, Andrew Y., and Tom Zimmermann. 2022. "Open source cross-sectional asset pricing." *Critical Finance Review* 27 (2): 207–264.

Chen, Huafeng Jason. 2011. "Firm life expectancy and the heterogeneity of the book-to-market effect." *Journal of Financial Economics* 100 (2): 402–423.

Chen, Huafeng Jason. 2017. "Do cash flows of growth stocks really grow faster?" *The Journal of Finance* 72 (5): 2279–2330.

Chen, Joseph, Harrison Hong, and Jeremy C Stein. 2002. "Breadth of ownership and stock returns." *Journal of Financial Economics* 66 (2-3): 171–205.

Chen, Long, and Lu Zhang. 2010. "A better three-factor model that explains more anomalies." *The Journal of Finance* 65 (2): 563–595.

Chen, Long, and Xinlei Zhao. 2009. "Return decomposition." *The Review of Financial Studies* 22 (12): 5213–5249.

Chen, Nai-fu, and Feng Zhang. 1998. "Risk and return of value stocks." *The Journal of Business* 71 (4): 501–535.

Chen, Zhanhui. 2022. "Inferring stock duration around fomc surprises: Estimates and implications." *Journal of Financial and Quantitative Analysis* 57 (2): 669–703.

Chiavari, Andrea, and Sampreet Goraya. 2020. "The rise of intangible capital and the macroeconomic implications."Technical report, Working Paper.

Cho, Thummim, Lukas Kremens, Dongryeol Lee, and Christopher Polk. 2022. "Scale or yield? a present-value identity." *A Present-Value Identity (June 1, 2021)*.

Chung, Chune Young, Chang Liu, and Kainan Wang. 2021. "The big picture: The industry effect of short interest." *International Review of Financial Analysis* 76: 101760.

Cochrane, John H. 2008. "The dog that did not bark: A defense of return predictability." *The Review of Financial Studies* 21 (4): 1533–1575.

Cochrane, John H. 2011. "Presidential address: Discount rates." *The Journal of Finance* 66 (4): 1047–1108.

Cohen, Randolph B, Christopher Polk, and Tuomo Vuolteenaho. 2003. "The value spread." *The Journal of Finance* 58 (2): 609–641.

Cooper, Michael, and Huseyin Gulen. 2006. "Is time-series-based predictability evident in real time?" *The Journal of Business* 79 (3): 1263–1292.

Corrado, Carol, Charles Hulten, and Daniel Sichel. 2009. "Intangible capital and us economic growth." *Review of income and wealth* 55 (3): 661–685.

Da, Zhi. 2009. "Cash flow, consumption risk, and the cross-section of stock returns." *The Journal of Finance* 64 (2): 923–956.

Daniel, Kent, David Hirshleifer, and Lin Sun. 2020. "Short-and long-horizon behavioral factors." *The Review of Financial Studies* 33 (4): 1673–1736.

Davis, James L, Eugene F Fama, and Kenneth R French. 2000. "Characteristics, covariances, and average returns: 1929 to 1997." *The Journal of Finance* 55 (1): 389–406.

Davydiuk, Tetiana, Brent Glover, and Rachel Szymanski. 2020. "The decline in public firms."Technical report, working paper.

De La O, Ricardo, and Sean Myers. 2021. "Subjective cash flow and discount rate expectations." *The Journal of Finance* 76 (3): 1339–1387.

De Nard, Gianluca, and Zhao Zhao. 2022. "A large-dimensional test for cross-sectional anomalies: Efficient sorting revisited." *International Review of Economics & Finance* 80: 654–676.

Dechow, Patricia M., Ryan D. Erhard, Richard G. Sloan, and Mark T. Soliman. 2021. "Implied equity duration: A measure of pandemic shutdown risk." *Journal of Accounting Research* 59 (1): 243–281.

Dechow, Patricia M, Richard G Sloan, and Mark T Soliman. 2004. "Implied equity duration: A new measure of equity risk." *Review of Accounting Studies* 9 (2): 197–228.

Dong, Xi, Qi Liu, Lei Lu, Bo Sun, and Hongjun Yan. 2022. "Anomaly discovery and arbitrage trading." *Available at SSRN*.

Döttling, Robin, and Lev Ratnovski. 2023. "Monetary policy and intangible investment." *Journal of Monetary Economics* 134: 53–72.

Dugar, Amitabh, and Jacob Pozharny. 2021. "Equity investing in the age of intangibles." *Financial Analysts Journal* 77 (2): 21–42.

Eisfeldt, Andrea L, Antonio Falato, and Mindy Z Xiaolan. 2021. "Human capitalists."Technical report, National Bureau of Economic Research.

Eisfeldt, Andrea L, Edward T Kim, and Dimitris Papanikolaou. 2022. "Intangible value." *Critical Finance Review* 11 (2): 299–332.

Eisfeldt, Andrea L, and Dimitris Papanikolaou. 2013. "Organization capital and the cross-section of expected returns." *The Journal of Finance* 68 (4): 1365–1406.

Engelberg, Joseph E, Adam V Reed, and Matthew C Ringgenberg. 2018. "Short-selling risk." *The Journal of Finance* 73 (2): 755–786.

Engsted, Tom, Thomas Q Pedersen, and Carsten Tanggaard. 2012a. "The log-linear return approximation, bubbles, and predictability." *Journal of Financial and Quantitative Analysis* 47 (3): 643–665.

Engsted, Tom, Thomas Q Pedersen, and Carsten Tanggaard. 2012b. "Pitfalls in var based return decompositions: A clarification." *Journal of Banking & Finance* 36 (5): 1255–1265.

Ewens, Michael, Ryan H Peters, and Sean Wang. 2020. "Measuring intangible capital with market prices." *Working Paper*.

Falato, Antonio, Dalida Kadyrzhanova, and Jae W Sim. 2013. "Rising intangible capital, shrinking debt capacity, and the us corporate savings glut."Technical report, Board of Governors of the Federal Reserve System (US).

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

Fama, Eugene F, and Kenneth R French. 1993. "Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics* 33: 3–56.

Fama, Eugene F, and Kenneth R French. 2008. "Dissecting anomalies." *The Journal of Finance* 63 (4): 1653–1678.

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

Fama, Eugene F, and Kenneth R French. 2021. "The value premium." *The Review of Asset Pricing Studies* 11 (1): 105–121.

Fama, Eugene F, and James D MacBeth. 1973. "Risk, return, and equilibrium: Empirical tests." *Journal of Political Economy* 81 (3): 607–636.

Farre-Mensa, Joan, Roni Michaely, and Martin Schmalz. 2014. "Payout policy." *Annual Review of Financial Economics* 6 (1): 75–134.

Feng, Guanhao, Stefano Giglio, and Dacheng Xiu. 2020. "Taming the factor zoo: A test of new factors." *The Journal of Finance* 75 (3): 1327–1370.

Fink, Josef. 2021. "A review of the post-earnings-announcement drift." *Journal of Behavioral and Experimental Finance* 29: 100446.

Gao, Can, and Ian WR Martin. 2021. "Volatility, valuation ratios, and bubbles: An empirical measure of market sentiment." *The Journal of Finance* 76 (6): 3211–3254.

Gareis, Johannes, and Eric Mayer. 2022. "Financial shocks and the relative dynamics of tangible and intangible investment: Evidence from the euro area." *Macroeconomic Dynamics*: 1–26.

Gibbons, Michael R, Stephen A Ross, and Jay Shanken. 1989. "A test of the efficiency of a given portfolio." *Econometrica: Journal of the Econometric Society*: 1121–1152.

Giglio, Stefano, Bryan Kelly, and Dacheng Xiu. 2022. "Factor models, machine learning, and asset pricing." *Annual Review of Financial Economics* 14: 337–368.

Golez, Benjamin, and Peter Koudijs. 2020. "Equity duration and predictability." *SSRN Electronic Journal*: 1–51.

Golubov, Andrey, and Theodosia Konstantinidi. 2019. "Where is the risk in value? evidence from a market-to-book decomposition." *The Journal of Finance* 74 (6): 3135–3186.

Gonçalves, Andrei. 2019. "What moves equity markets? a term structure decomposition for stock returns." *SSRN Electronic Journal* (May 2019).

Gonçalves, Andrei. 2021. "The short duration premium." *Journal of Financial Economics* 141 (3): 919–945.

Gormsen, Niels Joachim, and Eben Lazarus. 2023. "Duration-driven returns." *The Journal of Finance* 78 (3): 1393–1447.

Griliches, Zvi. 1979. "Issues in assessing the contribution of research and development to productivity growth." *The Bell Journal of Economics*: 92–116.

Gulen, Huseyin, Dongmei Li, Ryan H Peters, and Morad Zekhnini. 2020. "Intangible capital in factor models." *Available at SSRN*.

Gungor, Sermin, and Richard Luger. 2015. "Bootstrap tests of mean-variance efficiency with multiple portfolio groupings." *L'Actualité économique* 91 (1-2): 35–65.

Gungor, Sermin, and Richard Luger. 2016. "Multivariate tests of mean-variance efficiency and spanning with a large number of assets and time-varying covariances." *Journal of Business & Economic Statistics* 34 (2): 161–175.

Harvey, Campbell R. 2017. "Presidential address: The scientific outlook in financial economics." *The Journal of Finance* 72 (4): 1399–1440.

Harvey, Campbell R, and Yan Liu. 2019. "A census of the factor zoo." *Available at SSRN 3341728*.

Harvey, Campbell R, Yan Liu, and Heqing Zhu. 2016. "… and the cross-section of expected returns." *The Review of Financial Studies* 29 (1): 5–68.

Hasler, Mathias. 2021. "Is the value premium smaller than we thought?" *Available at SSRN 3886984*.

Hou, Kewei, Chen Xue, and Lu Zhang. 2015. "Digesting anomalies: An investment approach." *The Review of Financial Studies* 28 (3): 650–705.

Huberman, Gur, and Shmuel Kandel. 1987. "Mean-variance spanning." *The Journal of Finance* 42 (4): 873–888.

Jaffe, Jeffrey F, Jan Jindra, David J Pedersen, and Torben Voetmann. 2020. "Can mispricing explain the value premium?" *Financial Management* 49 (3): 615–633.

Jaffe, Jeffrey, Donald B Keim, and Randolph Westerfield. 1989. "Earnings yields, market values, and stock returns." *The Journal of Finance* 44 (1): 135–148.

Jagannathan, Ravi, Robert Korajczyk, and Kai Wang. 2023. "An intangibles-adjusted profitability factor."Technical report, National Bureau of Economic Research.

Jegadeesh, Narasimhan, and Sheridan Titman. 1993. "Returns to buying winners and selling losers: Implications for stock market efficiency." *The Journal of Finance* 48 (1): 65–91.

Kahle, Kathleen M, and René M Stulz. 2017. "Is the us public corporation in trouble?" *Journal of Economic Perspectives* 31 (3): 67–88.

Kan, Raymond, and GuoFu Zhou. 2012. "Tests of mean-variance spanning." *Annals of Economics and Finance* 13 (1): 139–187.

Kazemi, Maziar. 2022. "Intangible investment, displacement risk, and the value discount." *Available at SSRN*.

Khimich, Natalya. 2017. "A comparison of alternative cash flow and discount rate news proxies." *Journal of Empirical Finance* 41: 31–52.

Knox, Benjamin, and Annette Vissing-Jorgensen. 2022. "A stock return decomposition using observables."

Kogan, Leonid, and Dimitris Papanikolaou. 2014. "Growth opportunities, technology shocks, and asset prices." *The Journal of Finance* 69 (2): 675–718.

Koh, Dongya, Raül Santaeulàlia-Llopis, and Yu Zheng. 2020. "Labor share decline and intellectual property products capital." *Econometrica* 88 (6): 2609–2628.

Koijen, Ralph S.J., and Stijn Van Nieuwerburgh. 2011. "Predictability of returns and cash flows." *Annual Review of Financial Economics* 3: 467–491.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh. 2018. "Interpreting factor models." *The Journal of Finance* 73 (3): 1183–1223.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh. 2020. "Shrinking the cross-section." *Journal of Financial Economics* 135 (2): 271–292.

Lakonishok, Josef, Andrei Shleifer, and Robert W Vishny. 1994. "Contrarian investment, extrapolation, and risk." *The Journal of Finance* 49 (5): 1541–1578.

Lall, Subir, and Zeng. 2020. "Intangible investment and low inflation: A framework and some evidence."Technical report, International Monetary Fund.

Larrain, Borja, and Motohiro Yogo. 2008. "Does firm value move too much to be justified by subsequent changes in cash flow?" *Journal of Financial Economics* 87 (1): 200–226.

Lettau, Martin, and Sydney Ludvigson. 2001. "Resurrecting the (c) capm: A cross-sectional test when risk premia are time-varying." *Journal of Political Economy* 109 (6): 1238–1287.

Lettau, Martin, and Markus Pelger. 2020. "Estimating latent asset-pricing factors." *Journal of Econometrics* 218 (1): 1–31.

Lettau, Martin, and Jessica A. Wachter. 2007. "Why is long-horizon equity less risky? a duration-based explanation of the value premium." *The Journal of Finance* 62 (1): 55–92.

Lev, Baruch, and Suresh Radhakrishnan. 2009. "3. the valuation of organization capital." In *Measuring Capital in the New Economy,* 73–110. University of Chicago Press.

Lev, Baruch, Suresh Radhakrishnan, and Weining Zhang. 2009. "Organization capital." *Abacus* 45 (3): 275–298.

Lev, Baruch, and Theodore Sougiannis. 1996. "The capitalization, amortization, and value-relevance of r&d." *Journal of Accounting and Economics* 21 (1): 107–138.

Li, CY. 2012. "Depreciation of business R&D capital. Bureau of Economic Analysis/National Science Foundation R&D Satellite Account Paper."

Li, WC, and BH Hall. 2016. "Depreciation of business R&D capital (No. w22473)."

Lochstoer, Lars A, and Paul C Tetlock. 2020. "What drives anomaly returns?" *The Journal of Finance* 75 (3): 1417–1455.

Maio, Paulo, and Danielle Xu. 2020. "Cash-flow or return predictability at long horizons? the case of earnings yield." *Journal of Empirical Finance* 59: 172–192.

Mao, Mike Qinghao, and KC John Wei. 2014. "Price and earnings momentum: An explanation using return decomposition." *Journal of Empirical Finance* 28: 332–351.

McLean, R David, and Jeffrey Pontiff. 2016. “Does academic research destroy stock return predictability?” *The Journal of Finance* 71 (1): 5–32.

Michaely, Roni, Stefano Rossi, and Michael Weber. 2021. “Signaling safety.” *Journal of Financial Economics* 139 (2): 405–427.

Mullins, Gary. 2021. “Equity duration.” *SSRN Electronic Journal*.

Nagel, Stefan. 2005. “Short sales, institutional investors and the cross-section of stock returns.” *Journal of Financial Economics* 78 (2): 277–309.

Newey, Whitney K, and Kenneth D West. 1987. “Hypothesis testing with efficient method of moments estimation.” *International Economic Review*: 777–787.

Park, Hyuna. 2022. “An intangible-adjusted book-to-market ratio still predicts stock returns.” *Critical Finance Review* 11 (2): 265–297.

Peters, Ryan H, and Lucian A Taylor. 2017. “Intangible capital and the investment-q relation.” *Journal of Financial Economics* 123 (2): 251–272.

Porta, Rafael La, Josef Lakonishok, Andrei Shleifer, and Robert Vishny. 1997. “Good news for value stocks: Further evidence on market efficiency.” *the Journal of Finance* 52 (2): 859–874.

Priestley, Richard. 2019. “Short interest, macroeconomic variables and aggregate stock returns.” *Macroeconomic Variables and Aggregate Stock Returns (May 8, 2019)*.

Rapach, David E, Matthew C Ringgenberg, and Guofu Zhou. 2016. “Short interest and aggregate stock returns.” *Journal of Financial Economics* 121 (1): 46–65.

Rhodes-Kropf, Matthew, David T Robinson, and Sean Viswanathan. 2005. “Valuation waves and merger activity: The empirical evidence.” *Journal of Financial Economics* 77 (3): 561–603.

Rizova, Savina, and Namiko Saito. 2021. “Internally developed intangibles and expected stock returns.” *Available at SSRN 3697452*.

Santa-Clara, Pedro. 2004. “Discussion of "implied equity duration: A new measure of equity risk".” *Review of Accounting Studies* 9 (2-3): 229–231.

Schröder, David and Esterer, Florian. 2016. “A new measure of equity and cash flow duration: The duration-based explanation of the value premium revisited.” *Journal of Money, Credit and Banking* 48 (5): 857–900.

Soebhag, A, Bart Van Vliet, and Patrick Verwijmeren. 2022. “Non-standard errors in asset pricing: Mind your sorts.”Technical report, SSRN Working Paper.

Squicciarini, Mariagrazia, and Marie Le Mouel. 2012. “Defining and measuring investment in organisational capital: using us microdata to develop a task-based approach.”

Stambaugh, Robert F, and Yu Yuan. 2017. “Mispricing factors.” *The Review of Financial Studies* 30 (4): 1270–1315.

Sun, Qi, and Mindy Z Xiaolan. 2019. “Financing intangible capital.” *Journal of Financial Economics* 133 (3): 564–588.

Tambe, Prasanna, Lorin Hitt, Daniel Rock, and Erik Brynjolfsson. 2020. “Digital capital and superstar firms.”Technical report, National Bureau of Economic Research.

Van Binsbergen, Jules H., and Ralph S.J. Koijen. 2017. “The term structure of returns: Facts and theory.” *Journal of Financial Economics* 124 (1): 1–21.

Vincenz, Stefan. 2023. “Intangible value: An international perspective.” *Available at SSRN 4344729*.

Vuolteenaho, Tuomo. 2002. “What drives firm-level stock returns?” *The Journal of Finance* 57 (1): 233–264.

Weber, Michael. 2018. “Cash flow duration and the term structure of equity returns.” *Journal of Financial Economics* 128 (3): 486–503.

Welch, Ivo, and Amit Goyal. 2008. “A comprehensive look at the empirical performance of equity premium prediction.” *The Review of Financial Studies* 21 (4): 1455–1508.

Zhang, Lu. 2005. “The value premium.” *The Journal of Finance* 60 (1): 67–103.