THREE ESSAYS ON CONDITIONAL FACTOR PREMIUMS AND ASSET PRICING

THREE ESSAYS ON CONDITIONAL FACTOR PREMIUMS AND ASSET PRICING

By LAITE GUO,

A Thesis Submitted to the School of Graduate Studies in the Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

McMaster University

Doctor of Philosophy (2022)

Hamilton, Ontario (School of Business)

TITLE: THREE ESSAYS ON CONDITIONAL FACTOR PREMIUMS AND ASSET PRICING

AUTHOR: Laite Guo (McMaster University)

SUPERVISOR: Dr. Ron Balvers

NUMBER OF PAGES: 156, 224

Abstract

While existing studies have provided rich insights for understanding asset pricing factors or anomalies, this thesis refreshes the understanding of several major asset pricing phenomena from novel perspectives. Specifically, the refreshed understanding is achieved by conditioning on time-series or cross-sectional information.

The first chapter resurrects the size effect, the first challenge to market efficiency under the CAPM. Motivated by the extensive evidence of information barriers facing small stocks and the consequential underreaction of small-stock prices to information shocks, I conjecture that a big-minus-small effect comparable in strength to the small-minus-big effect exists. Conditional on ex-ante signals capturing information shocks, I uncover the two faces of the size effect. The finding not only translates into a size investment strategy with a remarkable improvement in risk-return trade-off but also sheds new light on the long-standing argument about the size effect's validity.

The second chapter refreshes the understanding of the low-beta anomaly. While recent studies have resolved the previously known low-beta anomaly from different perspectives, this chapter discovers a new low-beta anomaly not explained by these studies. I show that, theoretically, the new and known low-beta anomalies differ in their underlying factors. The known low-beta anomaly is driven by factors directly correlated with the market risk, while the new low-beta anomaly is driven by factors only partially correlated with the market risk. Besides showing that the low-beta anomaly is still unexplained by existing studies, the significance of the new low-beta anomaly lies in that it identifies partial-correlated factors, which are

unnoticeable but are important for driving asset returns.

The last chapter extends the analysis of the low market-beta anomaly to all factors or anomalies. Unlike the negative market beta-alpha relationship, which directly violates equilibrium theories such as the Consumption-CAPM, the implication of a factor-beta anomaly is unclear. After all, there is no consensus on the origins of most factors or anomalies, making it more attractive to clarify this class of phenomena as a whole to pave the way for economic interpretations. I explicitly model the mechanism for a factor-beta anomaly to emerge under a factor-model framework, which rationalizes the pervasiveness of low factor-beta anomalies and provides a paradigm for inferring information from an observed factor-beta anomaly.

Through the stochastic discount factor, asset pricing factors are at the core of the interaction between the financial market and the real economy. This thesis takes a crucial step toward understanding this interaction by enhancing the understanding of these asset pricing factors and anomalies.

Acknowledgements

I would like to express my sincere appreciation to my supervisor, Dr. Ronald Balvers. His experience, knowledge and integrity have been of great help. He was there to listen, discuss, and check my papers patiently, and provided valuable comments on my research. Moreover, I am grateful that he spent valuable time arranging my dissertation defense to make it proceed quickly. I thank my committee members, Dr. Guoying Luo and Dr. Peter Miu, for their valuable feedback on my research and help with my job applications. I also thank Dr. Clarence Kwan for his constant guidance and encouragement throughout my Ph.D. stage. Additionally, I would like to acknowledge Dr. Paulo Maio for agreeing to be the external reviewer and providing insightful comments and suggestions on my papers.

I thank my parents for their unconditional love and support during every stage of my life. I thank my girlfriend for her accompany and belief in me. Lastly, I also thank all my friends for their help. The Ph.D. journey, especially those shocks during the process, has ultimately changed my life trajectory. Most of the time, the expected payoff is not positively correlated with ex-ante efforts or endowments. However, I am still willing to endeavor to promote the progress of financial economics.

Contents

List of Figures

List of Tables

Declaration of Authorship

I, Laite Guo, declare that this thesis, entitled, "THREE ESSAYS ON CONDITIONAL FACTOR PREMIUMS AND ASSET PRICING", and the work presented in it, are my own. I confirm that the thesis comprises the following chapters:

- Two Faces of the Size Effect
- Uncovering Omitted Factors through a New Low-Beta Anomaly
- Demystifying Factor-Beta Anomalies

This thesis is entirely my own original work unless otherwise indicated. Any use of the work of other authors is acknowledged at their point of use.

Introduction

Before the formal analyses, I briefly introduce the major findings of the three chapters.

The first chapter finds that the size effect has two faces predictable from forward-looking signals: conditional on the up signal, small stocks outperform big stocks; conditional on the down signal, big stocks outperform small stocks. Size strategies based on the signals significantly outperform the original size strategy, and the enhanced performance is unexplained by common risk factors. Moreover, the two-faced size effect remains robust to stylized facts that invalidate the unconditional size effect. The conditional performance is consistent with the evidence that prices of small stocks respond slowly to information shocks and is mainly driven by the conditional difference in cash-flow shocks between small and big stocks. Overall, my findings not only show that both small-minus-big and big-minus-small premiums exist but also foreground the necessity of accounting for these two faces when assessing the validity of the size effect.

The second chapter uncovers a new low-beta anomaly (LBA) driven by factors different from those underlying the previously known LBA. I first show that estimated betas, besides capturing the market risk, also capture exposures to omitted factors; these omitted factors are the source of LBA. Then, I show that the known LBA reflects *factors directly correlated with the market risk* as they are the only omitted factors captured by estimated betas; however, *factors not directly but partially correlated with the market risk* (partial-correlation factors) can also induce LBA, but their impact was uncaptured. Based on recent studies resolving the known LBA, I allow estimated betas to capture partial-correlation factors and obtain a new LBA as a manifestation of these factors. The new LBA exhibits remarkable performance unexplained by extant factors or anomalies. This performance indicates that partial-correlation factors, previously not identified, are important in driving asset returns.

The third chapter clarifies the mechanism for factor-beta anomalies (i.e., nonzero cross-sectional factor beta-alpha relationships) to emerge under a standard factor-model framework. The explicit modeling generates the following findings. First, factor-beta anomalies do not challenge the covariance-based asset pricing framework but are induced by omitted systematic factors. Second, negative rather than positive factor beta-alpha relationships should be pervasive when betting against betas of priced factors with positive premiums. Third, omitted priced (unpriced) factors tend to (can only) induce a negative factor beta-alpha relationship. Last, omitted-priced factors should usually dominate unpriced factors in contributing to a factor-beta anomaly, and there is an upper bound for a factor-beta anomaly that unpriced factors can reconcile. Empirically, the betting-against-beta performance on a large set of factor and anomaly portfolios is consistent with these findings.

Chapter 1

Two Faces of the Size Effect

1.1 Introduction

The size effect, first documented in Banz (1981), refers to the tendency for small stocks to outperform big stocks. Although the phenomenon is widely regarded as an important challenge to market efficiency under the CAPM and has received considerable attention from academia and industry, its validity has been hotly debated. Some studies argue that the size effect is weak or its initial discovery is problematic (e.g., Brown et al. 1983; Shumway 1997). Others argue that the size effect has disappeared since its discovery (e.g., Dimson and Marsh 1999; Horowitz et al. 2000a,b; Hirshleifer 2001; Amihud 2002; Alquist et al. 2018). Indeed, as I confirm, the size premium is weak and has disappeared since the 1980s. However, this conclusion only holds for the unconditional size effect. In this study, I find that the size effect has two faces: a big-minus-small (BMS) effect and a small-minus-big (SMB) effect. Accounting for the two faces and assessing the validity of the size effect conditionally, I show that the size effect is strong and significant for both the full-sample and post-discovery periods.

My investigation is motivated by the extensive evidence of information costs or frictions facing small stocks, such as high information processing costs, low analyst coverage, institutional holding constraints, high information uncertainty, or limited investor attention (e.g., Merton 1987; Brennan et al. 1993; Badrinath et al. 1995; Cohen et al. 2002; Zhang 2006; Bali et al. 2013). A common impact of these barriers is that they prevent the prices of small stocks from incorporating new information timely; hence, the prices of small stocks exhibit an "underreaction" pattern. Consistent with this mechanism, previous studies document empirical evidence such as the lead-lag pattern between the returns of big and small stocks (e.g., Lo and MacKinlay 1990; Hou 2007) and the underreaction of small-stock prices to information shocks (e.g., Hong et al. 2000; Jiang and Zhu 2017).

In light of the evidence that the prices of small stocks incorporate new information slowly, I conjecture that big stocks outperform small stocks routinely. When there is bad news (information shocks negatively affecting prices) for the overall market or small stocks, the prices of small stocks may not adjust sufficiently to incorporate the information. In this case, small stocks tend to perform poorly in subsequent periods until the information is fully incorporated into their prices, underperforming big stocks and hence inducing a BMS effect. As news is largely reflected in realized returns, I use realized big- and small-stock returns to construct forward-looking signals to uncover the effect. A "down" signal, followed typically by a positive BMS return, is defined when both returns are negative, as this scenario usually reflects bad market-wide and small-stock-specific news. In contrast, the signal is “up” when both returns are positive, as this scenario usually corresponds to periods either with good news or less affected by bad news, which tends to be followed by a positive SMB return. The signal is “uncertain" when the two returns have opposite signs, conditional on which I do not expect small stocks to underperform or outperform big stocks notably. My empirical focus is to uncover the two-faced size effect and its economic benefits using the up and down signals.

I first provide portfolio- and firm-level evidence to show that the size effect has two faces and is valid if evaluated conditionally. The average unconditional SMB return is positive (0.32% per month, although not significant at the 1% level) for the full sample period (August 1926 to December 2019) but becomes negligible (0.01% per month, insignificant) for the post-1982 period. The performance is consistent with the critiques on the weakness and post-discovery disappearance of the size effect in previous studies. Once its two faces are considered, however, the size effect revives. Conditional on the up signal, the average SMB return is 1.20% per month, and conditional on the down signal, the average BMS return is 1.02% per month, both significant at the 1% level. The average monthly SMB and BMS returns remain considerable (0.74% and 1.18%) and significant (at the 1% level) for the post-discovery period. The firm-level evidence leads to the same conclusion, which remains intact when controlling for other characteristics capturing cross-sectional differences in returns.

Next, I investigate the economic benefits of the two-faced effect. Managing the SMB portfolio according to the two-faced size effect achieves an average monthly return of 0.96% and an annualized Sharpe ratio of 0.74 over the full sample period, significantly outperforming the original SMB portfolio (with an annualized Sharpe ratio of 0.22). The outperformance survives transaction costs and persists for the post-1980s period. When regressed against Fama and French factor models and the augmented q-factor model (Fama and French 1993; Carhart 1997; Fama and French 2015; Fama and French 2018; Hou et al. 2020a), the managed size strategy generates significantly (at the 1% level) positive monthly alphas ranging from 0.84% to 0.99%. The managed portfolio is weakly and insignificantly exposed to the size factor, which is reasonable as it captures both the SMB and BMS effects. The managed portfolio does not have a significantly positive momentum factor exposure, indicating that the two-faced size effect identified from past return-based signals does not profit from the momentum effect. The managed portfolio is not notably exposed to other factors, either. The result that common risk factors do not explain the two-faced size strategy is consistent with the slow information incorporation mechanism.

After confirming the major implications, I further investigate the economic mechanism of the two-faced size effect. If small stocks' prices respond to information shocks slowly, on average, the SMB premium after a down (up) signal should be negative (higher than usual) for a while before returning to a positive (normal) level, which I confirm by observing the performance pattern of an SMB portfolio. If the two-faced size effect stems from the mechanism that information barriers prevent small stocks' prices from incorporating new information timely, it should be stronger among stocks with a higher level of information barriers, which I confirm through three different proxies for information barriers. Additionally, I find that the two-faced size effect is well identified using intra-industry information, which is also consistent with the information-incorporation mechanism, as the industry channel is a major channel for information diffusion. These pieces of evidence support my conjecture that small stocks' slow response to information shocks induces the two-faced size effect. Then, I investigate which component of information shocks plays a major role. According to Campbell and Shiller (1988) and Campbell (1991), information shocks are composed of cash-flow and discount-rate shocks. By examining the two shocks extracted through a VAR-based return decomposition, I find that the SMB spread in cash-flow shocks notably changes with the up and down signals, while discount-rate shocks change trivially. The result is consistent with previous findings about the dominant impact of cash-flow shocks on the time variation of the size effect (e.g., Hou and van Dijk 2019; Lochstoer and Tetlock 2020). Overall, the results suggest that the two-faced size effect originates from smaller stocks' slower response to cash-flow shocks.

Finally, I confirm the robustness of the two-faced size effect. Many studies document stylized facts that invalidate the unconditional size effect, contending that it disappears if controlling for the January effect (e.g., Keim 1983; Lamoureux and Sanger 1989; Van Dijk 2011), is limited to the smallest stocks (e.g., Horowitz et al. 2000b; Crain 2011; Alquist et al. 2018), does not exist among junk stocks (e.g., Asness et al. 2018), is not observable in other equity markets (e.g., Alquist et al. 2018; Asness et al. 2018), and is indistinguishable from zero if excluding the trough stage of the business cycle (e.g., Ahn et al. 2019). While the unconditional size effect barely survives these critiques, I show that the two-faced size effect remains strong and significant when controlling for these effects.

Besides studies arguing against the size effect's validity or documenting barriers facing small stocks, my findings directly contribute to the literature about the size effect's conditional validity. Asness et al. (2018) find that the size effect becomes significant once firms' quality/junk characteristics are controlled. Esakia et al. (2019) find that the size premium is significant after adjusting for the implicit exposures to other factors. While these two studies exploit the interaction between size and other factors/characteristics, my study revives the size effect by recognizing its two faces. Hou and van Dijk (2019) uncover a significant SMB effect by removing the cash-flow shocks extracted using contemporaneous information, while I identify a two-faced size effect using ex-ante information. Studies about factor momentum (e.g., Ehsani and Linnainmaa 2019; Gupta and Kelly 2019), the interaction of information uncertainty and factor returns (e.g., Zhang 2006), or the interaction of the business cycle and the size effect (e.g., Ahn et al. 2019) also document the conditional realization of the size effect. My study contributes to the literature by uncovering a BMS effect comparable in strength to the SMB effect and showing that recognizing the two faces not only resurrects the size effect conditionally but also generates significant economic benefits.

The remainder of this study is structured as follows. Section 1.2 describes the motivation, signal construction, and data source. Section 1.3 presents the portfolio and firm-level evidence of the two-faced size effect. Section 1.4 examines the economic benefits, followed by the economic mechanism investigation in Section 1.5. Section 1.6 conducts a series of additional tests, and Section 1.7 provides the conclusion.

1.2 Motivation and Construction of the Forward-looking Signals

This section introduces the motivation of this study, the construction of the forward-looking signals, and the data.

1.2.1 Motivation

The validity of the size effect has been hotly debated since its first documentation in Banz (1981) due to its poor track record and disappearance since the 1980s (e.g., Brown et al. 1983; Horowitz et al. 2000b; Hirshleifer 2001; Amihud 2002; Alquist et al. 2018). Meanwhile, previous studies document various information costs or barriers facing small stocks, such as low analyst coverage, limited investor attention, high information uncertainty, institutional constraints on holding small stocks, and high information acquisition and dissemination costs (e.g., Merton 1987; Brennan et al. 1993; Badrinath et al. 1995; Hong et al. 2000; Cohen et al. 2002; Zhang 2006; Bali et al. 2013; Jiang and Zhu 2017). In this study, I connect the two strands of literature to analyze the validity of the size effect.

A prominent impact of these barriers is that they prevent the prices of small stocks from incorporating new information timely. For example, due to these barriers, investors may pay less attention to or need more time to process information relevant to small stocks, and hence do not sufficiently adjust their supply and demand for small stocks immediately upon the arrival of new information. Consequently, the prices of small stocks continue to respond to ex-ante information that they have not fully incorporated; that is, they exhibit an “underreaction" pattern. There is ample empirical evidence consistent with this slow information incorporation mechanism. For example, Hong et al. (2000) find that small stocks react sluggishly to bad news. Zhang (2006) finds that the high information uncertainty of small stocks results in their underreaction to new information. Jiang and Zhu (2017) find that small stocks exhibit a stronger underreaction to information shocks. Another piece of empirical support is the pattern that the returns of big stocks lead the returns of small stocks (i.e., the lead-lag effect), first documented in Lo and MacKinlay (1990). These authors suggest that the cause is that information shocks are incorporated into big stocks first and then transmitted to small stocks with a lag, which is supported empirically by later studies (e.g., Mech 1993; McQueen et al. 1996; Chordia and Swaminathan 2000; Hou 2007).

The extensive evidence of small stocks' slow incorporation of information implies that a big-minus-small (BMS) effect should also exist. The rationale is as follows. When there is bad news for the market or specifically for small stocks, small stocks perform poorly in response to the news. However, due to their slow information incorporation, the prices of small stocks do not adjust sufficiently to incorporate the news and thus will continue to perform poorly in subsequent periods until the information is fully incorporated. Therefore, conditional on bad news, there should be a BMS premium instead of an SMB premium. In contrast, if good news relevant to small stocks arrives in a period, small stocks will continue to perform well in subsequent periods until the news is fully incorporated into their prices, contributing to the SMB effect. The unconditional size effect performs poorly or even disappears if the BMS effect is as strong as the SMB effect; if this is the case, the size effect‘s validity should be assessed conditionally.

As the conditional performance stems from small stocks' insufficient incorporation of concurrent news, which can be easily identified, my major hypothesis is that the size effect has two faces that can be identified from ex-ante information. Moreover, given this slow information incorporation mechanism, managing a simple SMB portfolio according to the two-faced size effect should generate abnormal performance unexplained by common risk factors.

1.2.2 Forward-looking signals of the two-faced size effect

To uncover the two-faced size effect, I use monthly big- and small-stock returns as input to construct a forward-looking indicator (IG). The big(small)-stock return refers to the return of the portfolio composed of big (small) stocks in excess of the risk-free rate (one-month T-bill rate). Small- and big-stock portfolios are constructed by sorting stocks into size quantiles according to the NYSE 30th and 70th percentile breakpoints at the end of each June, and portfolio returns are recorded at the end of each month on a value-weighted basis. By construction, the IG indicator gives a down signal when both returns in the prior month are negative, an up signal when both are positive, and an uncertain signal when the two returns have opposite signs. The down signal predicts a positive BMS effect, and the up signal predicts a positive SMB effect. Constructing the signals simply using the signs of the two returns minimizes data-mining concerns and makes my analysis trackable.

More importantly, using big- and small-stock returns to construct the signals is consistent with my motivation. Stock prices move in response to information shocks; hence, information shocks are largely reflected in returns. If the prices of small stocks do not adjust sufficiently, their subsequent performance will be consistent with the signals conveyed by the big- and small-stock returns of the preceding period. For periods in which both big- and small-stock returns are negative (i.e., a down signal), the news relevant to small stocks is usually bad, and hence small stocks tend to perform poorly in subsequent periods, underperforming big stocks. In contrast, the news relevant to small stocks in periods with positive big- and small-stock returns (i.e., an up signal) is more likely to be good, and hence small stocks tend to outperform big stocks in subsequent periods. Following a similar rationale, we can understand the implication of the uncertain signal by considering a period with a negative big-stock return, which usually reflects bad market-wide news, and a positive small-stock return, which may reflect good small-stock-specific news. After such a period, small stocks tend to perform poorly if the bad market-wide news has a dominant impact but tend to perform well if the impact of the good small-stock-specific news is stronger. Therefore, conditional on the uncertain signal, I do not expect small stocks to underperform or outperform big stocks significantly.

1.2.3 Data

I obtain individual stock return and market capitalization data from CRSP. Only common stocks listed on NYSE, AMEX, and NASDAQ with share codes 10 or 11 are included (ADRs, ETFs, and REITs are excluded), and the original returns are adjusted for delistings following Shumway (1997). Firm accounting data are obtained from COMPUSTAT. Factor return data are from Kenneth French's and Lu Zhang's websites. For a firm, size from July of year t to June of year $t + 1$ is measured as its market equity (using the CRSP data) at the end of June of year t.

The full sample period is July 1926 to December 2019, with the data of July 1926 used for the initial signal calculation.

My main analysis uses the signals extracted from past one-month returns to utilize information effectively at the monthly frequency. For the full sample period, the percentages of time that up, down, and uncertain signals occur are 51.74%, 32.92%, and 15.34%, respectively; the percentages are respectively 52.63%, 30.48%, and 16.89% for the post-1982 period. Figure 1.1 provides a visual depiction of the realization of the forward-looking signals (the dotted vertical lines) and the SMB portfolio return (the solid vertical lines). As a robustness check, I also uncover the two-faced size effect based on different variations of the indicators, including using the past big-stock return or small-stock return alone, using different levels of past returns, using the returns of different past horizons, or using the aggregate market and SMB portfolio returns. These additional results are reported in Appendix A1.

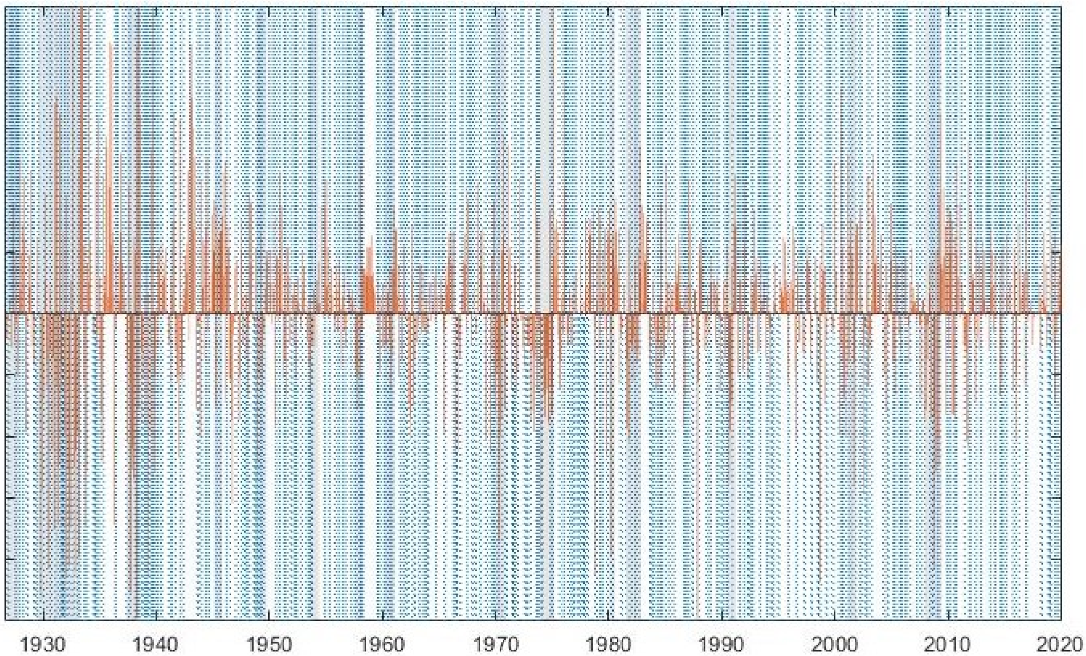
1.3 Empirical Evidence of the Two-Faced Size Effect

This section conducts portfolio- and firm-level examinations to confirm that the size effect has two faces.

1.3.1 Portfolio-level evidence

I start by analyzing the returns of ten size-decile portfolios and a small-minus-big (SMB) portfolio. To construct size-sorted portfolios, I sort stocks into size deciles according to the NYSE breakpoints at the end of each June and record the

Figure 1.1: A Visual Depiction of the Two-Faced Size Effect
The figure depicts the realized returns of the small-minus-big portfolio (the solid vertical line) and the forward-looking signals (the dotted vertical line) from August 1926 to December 2019. The signals are provided by the IG indicator at the beginning of each month, and the realized returns correspond to the end of each month. Dotted vertical lines above (below) the zero horizontal line refer to up (down) signals. Solid lines above (below) the zero horizontal line correspond to positive (negative) realized returns. Returns of the periods when the signal is uncertain are not plotted.



value-weighted returns of each portfolio every month until the next June. SMB refers to the zero-investment portfolio that buys small stocks and sells big stocks. The long and short sides of the SMB portfolio are the same as the two small- and big-stock portfolios used for constructing the forward-looking signals (see Section 1.2.2). Table 1.1 reports the unconditional and conditional returns of these decile portfolios (in excess of the risk-free rate) and the SMB portfolio (the last column).[1]

As reported in the first row of Panel A, for the full sample period from August

[1]For brevity, I only report the performance of value-weighted portfolios as that of equal-weighted portfolios leads to the same conclusions.

1926 to December 2019, unconditional average returns exhibit a descending pattern from the smallest (Q1) to largest (Q10) size portfolios. The average SMB return is non-negligible (0.32%, monthly), although its t-statistic is not high. While the full-sample return pattern indicates the existence of the size effect, the return pattern for the post-1982 period makes it difficult to justify the size effect's existence (the first row of Panel B). For this post-discovery period, the average excess returns of size-sorted portfolios no longer exhibit a decreasing pattern, and the average monthly SMB return is near zero (0.01%, insignificant). These results are consistent with the well-documented argument in existing studies that the unconditional size effect is weak and has disappeared since its discovery.

The conclusions are sharply different when the validity of the size effect is assessed conditionally. The remaining rows in Panel A report the average returns conditional on different signals (up, down, and uncertain) provided by the IG indicator. Consistent with my hypothesis, the size effect resurrects conditional on these signals. Over the full period, from Q1 to Q10, the average monthly excess returns of size-sorted portfolios conditional on the up signal decrease from 2.54% to 0.81%. The monthly SMB return conditional on the up signal (1.20%) is significant at the 1% level. A reverse size effect conditional on the down signal also emerges. The average monthly excess returns increase from -1.16% (Q1) to 0.26% (Q10), and the conditional SMB return (-1.02%, monthly) is significant at the 1% level. The cross-sectional difference in returns conditional on the uncertain signal is insignificant, consistent with how the uncertain signal is defined.

Panel B reports the conditional size effects for the post-Banz (1981) period, from which we can draw the same conclusion about the validity of the two-faced size effect. The monthly average excess returns of size-sorted portfolios decrease from 1.77% (Q1) to 0.77% (Q10) conditional on the up signal and increase from -1.34% (Q1) to 0.51% (Q10) conditional on the down signal. The average monthly SMB and BMS returns (conditional on the up and down signals) are respectively 0.74% and 1.18%, both significant at the 1% level. The cross-sectional difference in returns is insignificant when past big- and small-stock returns have opposite signs (i.e., when the signal is uncertain). These return patterns indicate that although the unconditional size effect is negligible over this period, the SMB and BMS effects remain considerable and significant.

In Appendix A1.1, I also compare the relative importance of the big- and small-stock returns in predicting the realization of the size effect. Both returns provide independent information for predicting the performance of the SMB portfolio, confirming that it is reasonable to combine information from both returns to identify the two-faced size effect. Overall, through the forward-looking signals constructed from past big- and small-stock returns, the portfolio-level evidence confirms that the size effect has two faces and revives once it is evaluated conditionally. Over the full-sample period, the realized SMB effect conditional on the up signal is stronger than the realized BMS effect conditional on the down signal, and hence an unconditional size effect is observed. For the post-discovery period, the realized SMB effect conditional on the up signal is not stronger than the realized BMS effect conditional on the down signal, which is why the unconditional size effect disappears.

TABLE 1.1: Portfolio-Level Evidence for the Two-Faced Size Effect

This table reports the unconditional and conditional excess returns (monthly, %) of size-sorted portfolios and the small-minus-big (SMB) portfolio. Panel A is for the full sample performance, and Panel B is for the post-1980s performance. Q1 to Q10 refer to the ten market-capitalization-decile portfolios. SMB refers to the portfolio that buys/sells stocks with a market capitalization below/above the NYSE 30/70th percentile breakpoint. Stocks are sorted into quantiles at the end of each June based on NYSE breakpoints to construct these portfolios, and portfolio returns are calculated on a value-weighted basis at the end of each month. The t-statistics are based on Newey and West (1987) adjusted standard errors. To get robust standard errors when evaluating conditional returns, I create a dummy variable following Cooper et al. (2004) to preserve the full sample time series. For example, when evaluating the mean return conditional on the down signal, I regress the time series of returns on a dummy variable (with no intercept) that takes the value of one when the signal is down and zero otherwise. The conditional averages in the remaining tables are evaluated using a similar approach. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The full sample period is from August 1926 to December 2019.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	SMB
Panel A: August 1926 - December 2019											
Unconditional	1.07***	0.94***	0.98***	0.91***	0.87***	0.88***	0.82***	0.79***	0.73***	0.62***	0.32*
	(3.23)	(3.24)	(3.66)	(3.72)	(3.78)	(4.04)	(3.98)	(4.10)	(4.04)	(3.99)	(1.93)
Up	2.54***	2.10***	1.88***	1.71***	1.57***	1.52***	1.38***	1.19***	1.05***	0.81***	1.20***
	(5.55)	(5.40)	(5.44)	(5.47)	(5.29)	(5.58)	(5.46)	(5.12)	(4.94)	(4.46)	(5.16)
Down	-1.16**	-0.77*	-0.47	-0.33	-0.14	-0.05	-0.04	0.17	0.23	0.26	-1.02***
	(-2.33)	(-1.67)	(-1.07)	(-0.79)	(-0.35)	(-0.13)	(-0.11)	(0.48)	(0.67)	(0.83)	(-4.48)
Uncertain	0.87	0.69	1.03	0.84	0.73	0.76	0.77	0.80	0.67	0.71*	0.21
	(1.31)	(1.07)	(1.59)	(1.42)	(1.31)	(1.37)	(1.52)	(1.58)	(1.44)	(1.74)	(0.62)
Panel B: January 1982 - December 2019											
Unconditional	0.62**	0.72**	0.82***	0.73***	0.81***	0.76***	0.82***	0.80***	0.79***	0.69***	0.01
	(1.97)	(2.38)	(2.95)	(2.69)	(3.05)	(3.08)	(3.34)	(3.40)	(3.54)	(3.50)	(0.06)
Up	1.77***	1.53***	1.43***	1.30***	1.22***	1.12***	1.17***	1.06***	1.01***	0.77***	0.74***
	(5.10)	(4.52)	(4.68)	(4.32)	(4.09)	(4.18)	(4.30)	(4.05)	(4.13)	(3.27)	(3.42)
Down	-1.34**	-0.67	-0.20	-0.16	0.15	0.22	0.20	0.35	0.46	0.51	-1.18***
	(-2.53)	(-1.18)	(-0.36)	(-0.30)	(0.28)	(0.42)	(0.39)	(0.68)	(0.94)	(1.13)	(-4.49)
Uncertain	0.57	0.68	0.75	0.54	0.71	0.60	0.83	0.82	0.69	0.79*	-0.11
	(0.83)	(0.98)	(1.13)	(0.86)	(1.14)	(1.12)	(1.51)	(1.63)	(1.52)	(1.95)	(-0.25)

1.3.2 Firm-level evidence

This subsection provides firm-level evidence to confirm the conditional relationship between return and size through the Fama and MacBeth (1973) regression. Besides complementing the portfolio-level evidence, the firm-level cross-sectional regression examination can also show whether the two-faced size effect remains robust after controlling for other characteristics. The firm-level test of the two-faced size effect is based on the cross-sectional regression:

$$r_{i,t} - r_f = \gamma_{0t} + \gamma_t \times Size_{i,t-1} + c_t \times Z_{i,t-1} + e_{i,t} \tag{1.1}$$

where $r_{i,t} - r_f$ is the realized return of stock i at time t in excess of the risk-free rate (r_f), γ_{0t} is the cross-sectional regression intercept, reflecting the return common to left-hand-side assets but uncaptured by the size effect or other control variables, $Size_{i,t-1}$ is the natural log of firm i's market capitalization, $Z_{i,t-1}$ is the vector of other characteristics informative about the cross-sectional differences in equity returns, and γ_t and c_t are the cross-sectional regression slopes at time t.

I estimate regression coefficients for each month and evaluate the size effects through two-tailed tests on the average unconditional and conditional values of γ_t. A positive γ_t means that smaller stocks have lower excess returns than bigger stocks in month t, i.e., a negative realized SMB premium. When evaluating the unconditional size effect, I test the unconditional value of γ_t, denoted by $\gamma = E(\gamma_t)$:

$$H_0 : \gamma = 0; \ \ H_1 : \gamma \neq 0$$

where H_0 and H_1 denote the null and alternative hypotheses, respectively. When evaluating the two-faced size effect, I test the conditional values of γ_t, denoted by $\gamma_s = E(\gamma_t|\text{Signal=s})$:

$$H_0 : \gamma_s = 0;\ \ H_1 : \gamma_s \neq 0$$

where s is up, down, or uncertain. Essentially, γ_{up}, γ_{down}, and $\gamma_{\text{uncertain}}$ are estimated using the γ_ts of periods with an ex-ante signal that is up, down, or uncertain, respectively. The BMS effect exists if γ_{down} is positive and significantly different from zero, and the SMB effect exists if γ_{up} is negative and significantly different from zero.

Table 1.2 presents the time-series averages of the cross-sectional regression slopes. Different columns correspond to regression specifications that differ in control variables or sample periods. The unconditional coefficients of size (γ) reported in the first row have low absolute values and are insignificant if observations from the first several decades are excluded. These values are consistent with the portfolio-level evidence that the unconditional size effect is weak over the full sample period and has disappeared since its discovery. The next several rows report the coefficients (γ_s) conditional on the signals from the IG indicator. Strong size premiums emerge once the size effect is evaluated conditionally. For example, when using the data of the full sample period and not including control variables, γ_{up} is -0.38 and γ_{down} is 0.33, both significant at the 1% level. These values show that the cross-sectional size-return relationship is significantly negative conditional on the up signal and significantly positive conditional on the down signal, indicating the existence of the SMB and BMS effects. The slopes conditional on the uncertain signal are mostly indistinguishable from zero.

TABLE 1.2: Firm-Level Evidence for the Two-Faced Size Effect

This table reports the firm-level cross-sectional regression slopes of return on size and other characteristics. Regressions for different columns differ in control variables and sample periods. The up, down, and uncertain signals are provided by the IG indicator. The control variables include the book-to-market ratio (BM), prior two-to-twelve-month cumulative return ($\text{Ret}_{2,12}$), operational profitability (OP), and investment (INV). The t-statistics are based on Newey and West (1987) adjusted standard errors. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The sample ends in December 2019.

	192608	198201	192608	198201	192701	198201	196307	198201
Unconditional	-0.11***	-0.02	-0.06*	0.04	-0.09***	-0.01	-0.09**	-0.04
	(-2.82)	(-0.55)	(-1.80)	(0.89)	(-2.89)	(-0.31)	(-2.49)	(-1.13)
Up	-0.38***	-0.27***	-0.31***	-0.19***	-0.34***	-0.23***	-0.32***	-0.23***
	(-7.24)	(-4.72)	(-7.45)	(-3.79)	(-8.31)	(-4.59)	(-7.06)	(-5.14)
Down	0.33***	0.39***	0.32***	0.42***	0.28***	0.35***	0.30***	0.28***
	(6.70)	(6.88)	(6.75)	(6.66)	(6.32)	(5.95)	(6.04)	(5.05)
Uncertain	-0.12	0.00	-0.03	0.06	-0.06	0.01	-0.15**	-0.03
	(-1.58)	(0.01)	(-0.36)	(0.75)	(-0.86)	(0.17)	(-2.16)	(-0.40)
BM			0.71***	1.12***	0.66***	0.98***	0.65***	0.64***
			(5.90)	(5.53)	(6.05)	(5.41)	(4.59)	(3.93)
$\text{Ret}_{2,12}$					1.13***	0.72***	0.83***	0.57***
					(9.12)	(4.40)	(6.31)	(3.79)
OP							0.16***	0.15**
							(2.78)	(2.39)
INV							-0.41***	-0.39***
							(-4.75)	(-4.68)

The bottom rows report the unconditional coefficients of other characteristics, including the book-to-market ratio (BM), prior two-to-twelve-month cumulative return ($\text{Ret}_{2,12}$), operational profitability (OP), and investment (INV). These characteristics are measured following Davis et al. (2000) and Fama and French (2015). The signs of these coefficients are consistent with previous studies on how these characteristics predict returns. The two-faced size effect remains significant when controlling for these characteristics. Overall, the firm-level evidence also confirms the existence of the two-faced size effect.

1.4 Economic Benefits of the Two-Faced Size Effect

A direct implication of the two-faced size effect is that managing a small-minus-big (SMB) portfolio according to the forward-looking signals can achieve a better risk-return tradeoff. This section investigates how a two-faced size strategy outperforms a simple unconditional size strategy and whether prominent asset pricing models subsume the enhanced performance.

1.4.1 Enhancing the size strategy performance according to the two-faced size effect

I compare the performance of the original small-minus-big (SMB) portfolio (OR, the same as the SMB portfolio in Section 1.3.1) and the managed SMB portfolio (LSI). At the beginning of each month, LSI buys OR when the signal is up, sells OR when the signal is down, and quits the risky investment when the signal is uncertain (i.e., the return of LSI is zero during periods with an uncertain signal). Panel A of Table 1.3 reports the performance for the full sample period. The first two columns present the performance of OR and LSI without adjusting for transaction costs. The average monthly return of LSI is 0.96%, higher than that of OR (0.32%). LSI achieves a much higher annualized Sharpe ratio than OR (0.74 versus 0.22). Through the Sharpe ratio improvement test of DeMiguel et al. (2009), I confirm that the improvement in the risk-return trade-off is significant. The significant spanning regression alpha (0.88%, monthly) also indicates that the enhanced strategy achieves returns unexplained by the original strategy (the t-statistic is 8.14).

Compared with the original size strategy (OR), the position of the managed size strategy (LSI) changes more frequently, depending on the time variation of the signals. For example, after portfolio formation at the end of each June, while the position of OR remains unchanged, the position of LSI changes from one to zero if the signal changes from up to uncertain. Such position changes imply that LSI is associated with higher transaction costs, which makes it necessary to investigate whether LSI outperforms OR after adjusting for transaction costs. Following Frazzini et al. (2015) and Moreira and Muir (2017), I set the transaction cost as ten basis points (10 bps). The remaining columns in Panel A report the transaction cost-adjusted performance of LSI. The Sharpe ratio of LSI decreases to 0.67 but is still significantly higher than that of OR. The outperformance remains significant when the transaction cost is doubled to 20 bps. The improvement is also significant for the post-1982 sample period when the unconditional size effect disappears. As shown in Panel B, while the OR portfolio has a near-zero annualized Sharpe ratio (0.01), the LSI portfolio still achieves positive Sharpe ratios comparable in magnitude to that of the full sample period. When transaction costs are not considered, the LSI strategy has a Sharpe ratio of 0.87 (annualized) and a spanning regression alpha (against OR) of 0.75% (monthly), significantly outperforming the OR portfolio. The outperformance remains when considering transaction costs, as shown in the last two columns. In Appendix A2, I confirm that the enhanced performance extends to other SMB portfolios.

Figure 1.2 depicts the cumulative returns (in logarithm) of the original SMB portfolio (OR) and the managed portfolio (LSI). To better understand how well

TABLE 1.3: Performance of the Original and Managed Size Strategies
This table reports the performance of the original (OR) and managed (LSI) size strategies. Average monthly returns (μ), annualized Sharpe ratios (SR), and spanning regression alphas of LSI on OR (α) are reported. The OR portfolio is the same as the SMB portfolio in Table 1.1. The position of LSI is adjusted based on the signals from the IG indicator: when the signal is up (down), the managed portfolio buys (shorts) OR; when the signal is uncertain, LSI quits the risky investment (and hence earns zero returns during such periods). The LSI performance adjusting for the transaction cost (TC) of ten or twenty basis points is also reported. The t-statistic to evaluate whether the improvement relative to OR is significant is reported under each alpha and Sharpe ratio. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The full sample period is from August 1926 to December 2019.

	Panel A: Full Sample				Panel B: January 1982 - December 2019			
	OR	LSI			OR	LSI		
		0 bps	10 bps	20 bps		0 bps	10 bps	20 bps
μ(%)	0.32*	0.96***	0.87***	0.79***	0.01	0.75***	0.66***	0.58***
	(1.93)	(6.68)	(6.05)	(5.43)	(0.06)	(5.37)	(4.47)	(4.11)
SR	0.22	0.74***	0.67***	0.60***	0.01	0.87***	0.77***	0.67***
		(4.09)	(3.57)	(3.04)		(4.01)	(3.56)	(3.11)
α(%)		0.88***	0.79***	0.71***		0.75***	0.66***	0.58***
		(8.14)	(7.29)	(6.46)		(5.57)	(4.91)	(4.26)

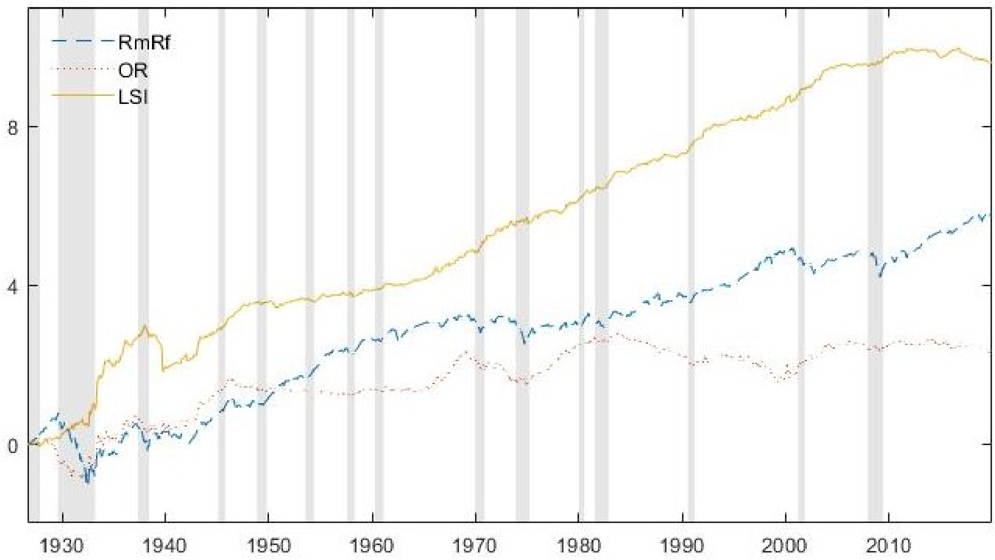
the enhanced strategy performs, I also plot the cumulative excess market return (RmRf). The comparison indicates that timing according to the two-faced size effect generates a considerable wealth increment relative to the original SMB portfolio. Moreover, while the original size strategy underperforms the market, the managed strategy performs better than the market.

1.4.2 Do common risk factors explain the enhanced performance?

This subsection investigates whether prominent asset pricing models explain the enhanced performance of LSI. The benefits of the two-faced size strategy are

FIGURE 1.2: Cumulative Returns of the Original and Managed Size Strategies

The figure depicts the logarithm of the cumulative returns of the original size portfolio (OR) and managed size portfolio (LSI). The OR portfolio is the same as the SMB portfolio in Table 1.1. The position of LSI is adjusted based on the signals given by the IG indicator: when the signal is up (down), the managed portfolio buys (shorts) OR; when the signal is uncertain, LSI quits the investment (and hence has zero returns). The cumulative excess market return (RmRf) is also plotted. The shaded bars correspond to the NBER recession periods. The sample period is from August 1926 to December 2019.



less desirable if they are the compensation for risks than when they are abnormal. Moreover, if the two-faced size effect stems from the slow information incorporation of small stocks' prices, common risk factors should not explain the enhanced performance. I use the three-factor model (Fama and French 1993), four-factor model (Carhart 1997), five-factor model (Fama and French 2015), five-factor plus momentum model (Fama and French 2018), and augmented q-factor model (Hou et al. 2020a) to examine whether the two-faced size strategy generates abnormal returns. Table 1.4 reports the factor loadings and alphas of the original and managed size portfolios (OR and LSI, respectively). MKT, SIZE, HML, RMW, CMA, and MOM are the market, size, value, profitability, investment, and momentum factors from Kenneth French's website. IA, ROE, and EG are the investment, return on equity, and expected growth factors from Lu Zhang's website.

The OR portfolio exhibits a significant loading on the size factor, which is expected given that they capture the same underlying effect. When controlling for these factor models, OR has insignificant or negative alphas, indicating that its performance is subsumed. After enhancing OR using the two-faced effect, however, the exposure to the size factor decreases remarkably. The exposures of LSI to the size factor range from -0.02 to 0.24 (insignificant), depending on which factor model is controlled. Additionally, LSI does not exhibit a significantly positive momentum factor exposure, indicating that the two-faced size effect identified from the past big- and small-stock performance differs from the momentum effect. LSI does not exhibit predominant exposures to any other factors, either. When regressing LSI on asset pricing models, the adjusted R-squares are much smaller (1.20% to 5.45%) than when regressing OR on asset pricing models (all above 90%). The monthly alphas of LSI are significantly positive (at the 1% level) and of considerable magnitude (ranging from 0.84% to 0.99%). In summary, the results in Table 1.4 suggest that the two-faced size strategy generates abnormal performance unexplained by common risk factors. The same conclusion holds for the post-discovery period, which I do not report for brevity.

Table 1.4: Time-Series Asset Pricing Tests of the Original and Managed Size Strategies

This table presents the spanning regression results of the original and managed size portfolios (OR and LSI) on asset pricing models. The OR portfolio is the same as the SMB portfolio in Table 1.1. The position of LSI is adjusted according to the signals from the IG indicator: when the signal is up (down), LSI buys (shorts) OR; when the signal is uncertain, LSI quits the risky investment (and hence earns zero returns during such periods). Monthly alphas, factor loadings (under each factor notation), and adjusted R-squares are reported. The factor models include the Fama-French three-factor (from August 1926), Carhart four-factor (from January 1927), Fama-French five-factor (from July 1963), Fama-French five-factor plus momentum (from July 1963), and augmented q-factor (from January 1967) models. MKT, SIZE, HML, RMW, CMA, and MOM are the market, size, value, profitability, investment, and momentum factors. IA, ROE, and EG are the investment, return on equity, and expected growth factors. The t-statistics are based on Newey and West (1987) adjusted standard errors. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The sample period depends on the data availability of these factors. The full sample period is August 1926 to December 2019.

	α(%)	MKT	SIZE	HML	MOM	CMA	RMW	IA	ROE	EG	$\bar{R}^2$(%)
OR	-0.15***	0.06***	1.32***	0.47***							94.22
	(-4.33)	(4.18)	(39.99)	(12.03)							
LSI	0.84***	0.01	0.24	0.18							5.45
	(7.85)	(0.43)	(1.19)	(1.43)							
OR	-0.11***	0.06***	1.31***	0.45***	-0.04						94.37
	(-2.95)	(4.60)	(38.29)	(15.25)	(-1.55)						
LSI	0.84***	0.01	0.24	0.18	-0.01						5.36
	(6.26)	(0.43)	(1.18)	(1.58)	(-0.01)						
OR	-0.07**	-0.01	1.23***	0.24***		-0.02	-0.07**				95.02
	(-2.26)	(-0.85)	(68.93)	(11.31)		(-0.62)	(-2.43)				
LSI	0.92***	-0.08*	0.02	-0.08		0.14	-0.07				1.20
	(7.12)	(-1.87)	(0.21)	(-0.86)		(1.21)	(-1.08)				
OR	-0.08**	-0.01	1.23***	0.24***	0.00	-0.02	-0.07**				95.01
	(-2.16)	(-0.80)	(68.60)	(10.69)	(0.13)	(-0.63)	(-2.50)				
LSI	0.99***	-0.09**	0.02	-0.13	-0.09*	0.17	-0.05				2.39
	(7.36)	(-2.28)	(0.29)	(-1.40)	(-1.85)	(1.61)	(-0.73)				
OR	0.07	-0.04**	1.15***					0.00	-0.20***	-0.09***	92.72
	(1.55)	(-2.44)	(53.11)					(0.13)	(-5.43)	(-2.63)	
LSI	0.95***	-0.08*	-0.02					0.01	-0.16*	0.07	1.67
	(5.85)	(-1.93)	(-0.22)					(0.07)	(-1.85)	(0.72)	

1.5 Economic Source of the Two-Faced Size Effect

This section provides additional evidence to show that the performance of the two-faced size effect is consistent with the mechanism that the prices of small stocks respond slowly to information shocks.

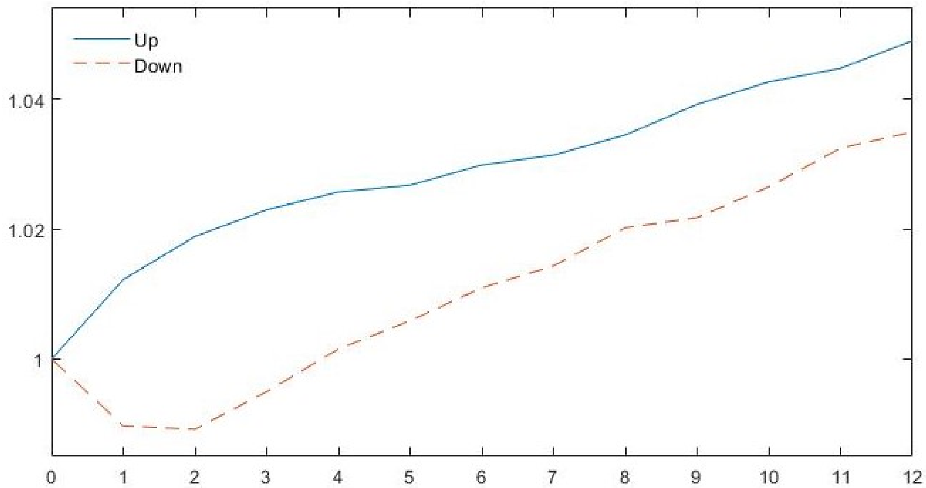
1.5.1 Cumulative performance after up and down signals

If the two-faced size effect stems from small stocks' slow response to new information, the performance of an SMB portfolio should exhibit an "underreaction" pattern, regardless of whether the slow response is caused by information costs, behavioral biases, or other barriers. In other words, the returns of an SMB portfolio should be abnormally low (high) over the first several periods after a down (up) signal and then go back to normal levels when the news is fully incorporated into stock prices. This subsection examines the cumulative performance of an SMB portfolio after up and down signals to confirm this mechanism.

Figure 1.3 depicts the cumulative return of an SMB portfolio (same as the SMB portfolio in Section 1.3.1) after the up or down signal. To get cumulative returns over time, at each time point, I record the signal (up or down) and the cumulative returns of the SMB portfolio over the next twelve months. I then calculate the n-th month cumulative return after the up (or down) signal as the average of all n-th month cumulative returns after up signals (or down signals) in the sample period. As the dashed line suggests, the cumulative return decreases over the first several months after the down signal and then increases steadily. This pattern is consistent with the mechanism that the prices of small stocks respond to information shocks slowly: after a period of bad news (down signal), the SMB portfolio continues to perform poorly until the bad news is fully incorporated into prices. In contrast, the SMB portfolio performs better than usual after the up signal, as shown by the solid line, and then exhibits normal cumulative performance after the good news is fully incorporated into prices. This pattern is also consistent with the slow information incorporation mechanism. Except for the first several months, the growth rates of the cumulative returns after up and down signals are roughly similar, reflecting the size effect less affected by information shocks.

FIGURE 1.3: Cumulative Returns after Up and Down Signals

This figure depicts the cumulative returns after different signals. At each time point, I record the signal (up or down) and the cumulative returns of an SMB portfolio (same as the SMB portfolio in Section 1.3.1) over the next twelve months. I then calculate the average n-th month cumulative return after the up (or down) signal as the average of all n-th month cumulative returns after up signals (or down signals) in the sample period, where $n = 1, 2, \cdots, 12$. The sample period is from August 1926 to December 2019.



1.5.2 Two-faced size effect for stocks with different information barriers

If the two-faced size effect stems from the fact that many barriers prevent small stocks' prices from incorporating new information timely, it should be stronger for stocks with a higher level of information barriers than for those with a lower level of information barriers. To confirm this implication, I divide the sample into a higher information-barrier group and a lower information-barrier group, and compare the two-faced size effect between the two groups. In each period, stocks are sorted into two groups according to the NYSE median breakpoint based on a specific measure that captures different levels of information barriers. I use three firm-level proxies to capture information barriers: analyst coverage, firm age, and volatility.[2] For example, when using analyst coverage as the proxy, the higher/lower information-barrier group contains stocks with a lower/higher level of analyst coverage.

Table 1.5 reports the average conditional cross-sectional regression slopes of return on size for the two groups (following Section 1.3.2). Owing to the availability of analyst coverage data, I first report the results based on the post-1980s sample in Panel A. When analyst coverage is used to group stocks, the average monthly slope for the higher information-barrier group is -0.40 conditional on the up signal and 0.38 conditional on the down signal. The average monthly slopes for the lower information-barrier group are -0.23 and 0.19, which have smaller absolute values. Similar results hold when volatility or firm age is used to capture different levels of

[2]Analyst coverage is measured as the number of analysts following a firm in the previous year (the data are available from the IBES database); volatility is measured using daily returns of the prior year; firm age is measured as the number of months since the CRSP database first covered a firm. These proxies are related to different levels of information barriers, as shown in previous studies (e.g., Barry and Brown 1985; Hong et al. 2000; Zhang 2006).

TABLE 1.5: Two-Faced Size Effect by Information-Barrier Groups
This table reports the firm-level cross-sectional regression slopes of return on size for the group of stocks with a higher or lower level of information barrier. In each period, stocks are sorted into higher and lower information-barrier groups according to the NYSE median breakpoint based on a measure that captures different levels of barriers. Three proxies are used: analyst coverage (AC), volatility (Vol), and firm age (Age). Average slopes conditional on the up and down signals (from the IG indicator) are reported. The last two columns report the slope differences between the two groups conditional on the up or down signal. The t-statistics are based on Newey and West (1987) adjusted standard errors. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The full sample period is from August 1926 to December 2019, and data from the first year are used for calculating the initial volatility and age of sample stocks.

Grouped	Higher-Barrier Group		Lower-Barrier Group		Diff	
By	Up	Down	Up	Down	Up	Down
			Panel A: The Post-1980s Sample			
AC	-0.40***	0.38***	-0.23***	0.19***	-0.18**	0.19**
	(-4.74)	(5.44)	(-2.73)	(2.68)	(-2.27)	(2.06)
Vol	-0.36***	0.38***	-0.11***	0.19***	-0.25***	0.19***
	(-5.24)	(4.39)	(-3.77)	(3.60)	(-4.10)	(3.13)
Age	-0.31***	0.40***	-0.22***	0.29***	-0.10**	0.11**
	(-3.92)	(6.05)	(-4.19)	(4.59)	(-2.07)	(1.97)
			Panel B: The Full Sample			
Vol	-0.58***	0.23***	-0.22***	0.16***	-0.36***	0.07***
	(-7.93)	(4.76)	(-9.66)	(4.74)	(-5.65)	(2.72)
Age	-0.48***	0.29***	-0.40***	0.23***	-0.08***	0.06**
	(-4.77)	(4.73)	(-4.22)	(3.86)	(-3.01)	(2.14)

information barriers. The last two columns report the difference in the conditional size effects between the two groups. All these differences are statistically significant, showing that the two-faced size effect is significantly stronger among stocks with higher information barriers. Panel B reports the results using the full sample data (volatility and firm age data are available; data from the first year are used for calculating the initial volatility and age of sample stocks), from which we can draw similar conclusions.

1.5.3 Two-faced size effect based on intra-industry information

Under the slow information incorporation mechanism, using industry-based rather than aggregate information should identify the two-faced size effect well, as the industry channel is a major channel for information diffusion for various reasons. For example, some information shocks may be relevant to certain industries, or some aggregate information shocks may affect different industries differently. To examine whether intra-industry information identifies the two-faced size effect, I construct an industry-based indicator (IG_{Ind}) using the past period small- and big-stock returns of each industry. To get intra-industry small- and big-stock returns, I sort the stocks of each industry into portfolios at the end of each June according to the NYSE 30th and 70th percentile size breakpoints at that time point; then, I record portfolio returns on a value-weighted basis at the end of each month over the next twelve months and subtract the risk-free rate from the original portfolio returns. Similar to the definition of the IG indicator, IG_{Ind} provides an up signal for a stock if both the past small- and big-stock returns of the industry to which that stock belongs are positive, a down signal when both are negative, and an uncertain signal when the two returns have opposite signs.

Panels A and B of Table 1.6 report the results based on the 49 and 12 industry classification systems (available from Kenneth French's website), respectively. For a cross-sectional regression coefficient to be valid, I require at least one-third of the sample stocks at that time point to be included in the regression. The average cross-sectional regression slopes are significantly negative conditional on the up signal and significantly positive conditional on the down signal. The results

TABLE 1.6: Two-Faced Size Effect Based on Industry Signals
This table reports the firm-level cross-sectional regression slopes of return on size conditional on signals provided by an industry-based indicator, IG_{Ind}. The 49 and 12 industry classifications from Kenneth French's website are used. IG_{Ind} provides an up signal for a stock if both the past small- and big-stock returns of the industry that stock belongs to are positive, a down signal when both are negative, and an uncertain signal when the two returns have opposite signs. Controls-Yes refers to that the book-to-market ratio (BM), prior two-to-twelve-month cumulative return ($\text{Ret}_{2,12}$), operational profitability (OP), and investment (INV) are controlled in the regression. The t-statistics are based on Newey and West (1987) adjusted standard errors. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The sample period is from August 1926 to December 2019.

	Panel A: Results Based on 49 Industries				Panel B: Results Based on 12 Industries			
	192608	198201	196307	198201	192608	198201	196307	198201
Up	-0.42***	-0.31***	-0.34***	-0.25***	-0.39***	-0.28***	-0.33***	-0.25***
	(-7.28)	(-5.03)	(-6.98)	(-5.05)	(-6.85)	(-4.69)	(-7.02)	(-5.26)
Down	0.33***	0.37***	0.27***	0.26***	0.31***	0.35***	0.26***	0.26***
	(6.86)	(6.36)	(5.76)	(4.65)	(7.08)	(6.13)	(5.56)	(4.57)
Uncertain	-0.04	-0.03	-0.09**	-0.04	-0.09**	-0.01	-0.07	0.01
	(-0.99)	(-0.72)	(-2.46)	(-1.07)	(-1.97)	(-0.01)	(-1.44)	(0.14)
Controls	No	No	Yes	Yes	No	No	Yes	Yes

indicate that using intra-industry information directly can identify the two-faced size effect well, supporting my slow information incorporation-based motivation. Given the importance of the industry channel for information diffusion, I regard a comprehensive industry-based analysis as a promising future extension of this study.

1.5.4 Cash-flow or discount-rate shocks?

The empirical evidence in previous sections is consistent with my motivation that the two-faced size effect stems from small stocks' slow response to information shocks. According to Campbell (1991), the information shock component of a realized return consists of cash-flow and discount-rate shocks. In this subsection, I further investigate which component of information shocks accounts for the two-faced size effect.

Extracting cash-flow and discount-rate shocks from realized returns

The Campbell (1991) return decomposition is expressed as

$$r_{i,t+1} - E_t(r_{i,t+1}) \approx CF_{i,t+1} - DR_{i,t+1} \tag{1.2}$$

where $r_{i,t+1} - E_t(r_{i,t+1})$ is the information-shock component of the log realized return ($r_{i,t+1}$), $CF_{i,t+1}$ is the cash-flow shock, and $DR_{i,t+1}$ is the discount-rate shock. A VAR model is widely used to extract the two shock components of returns (e.g., Vuolteenaho 2002):

$$Z_{t+1} = A + BZ_t + U_{t+1} \tag{1.3}$$

where Z_{t+1} is a K-by-1 vector of VAR variables, A is a K-by-1 vector of constants, B is a K-by-K matrix of VAR slopes, and U_{t+1} is a K-by-1 vector of zero-mean innovations. Cash-flow and discount-rate shocks are extracted using coefficients and return innovations estimated from the VAR model:

$$DR_{t+1} = E_{t+1}\sum_{j=2}^{\infty}\rho^{j-1}r_{t+j} - E_t\sum_{j=2}^{\infty}\rho^{j-1}r_{t+j} = l'\rho B(I-\rho B)^{-1}U_{t+1} \tag{1.4}$$

$$CF_{t+1} = l'[I + \rho B(I-\rho B)^{-1}]U_{t+1} \tag{1.5}$$

where l is a K-by-1 vector with one as its first element and zeros as its remaining elements, ρ is a log-linearization constant set as 0.95 (e.g., Cohen et al. 2003;

Lochstoer and Tetlock 2020), and I is a K-by-K identity matrix.

I follow the approach in Lochstoer and Tetlock (2020), which is modified for analyzing characteristics-sorted portfolios, to decompose returns. The VAR variables (measured in logarithm) are the realized return, book-to-market ratio, market capitalization (measured as the five-year change in log market equity to ensure stationarity), cumulative return of the past two-to-twelve months, operational profitability, and investment. This approach assumes that the expected return is linear in observable variables:

$$E_t(r_{i,t+1}) = c_0 + c_1 X_t + c_2 X^e_{it} \tag{1.6}$$

where X^e_{it} is the vector of firm-level characteristics above the aggregate-level characteristics, X_t (measured as the value-weighted average of firm-level characteristics). Accordingly, the aggregate-level and firm-specific return components are extracted through a standard VAR model and a panel VAR model separately, allowing the VAR coefficients to vary over time and across firms to match the data.[3]

A return component of firm i at time $t+1$, for example, the discount-rate stock ($DR_{i,t+1}$), is the summation of the aggregate-level and firm-specific discount-rate shocks (DR^m_{t+1} and DR^i_{t+1}, respectively). The former (DR^m_{t+1}) is extracted using the standard VAR model, and the latter (DR^i_{t+1}) is extracted using the panel

[3]For the panel VAR model, the coefficients are estimated through weighted least squares following Cohen et al. (2003) and Lochstoer and Tetlock (2020). The weight for each observation equals the inverse of the number of firms at that time point to weight each time point equally.

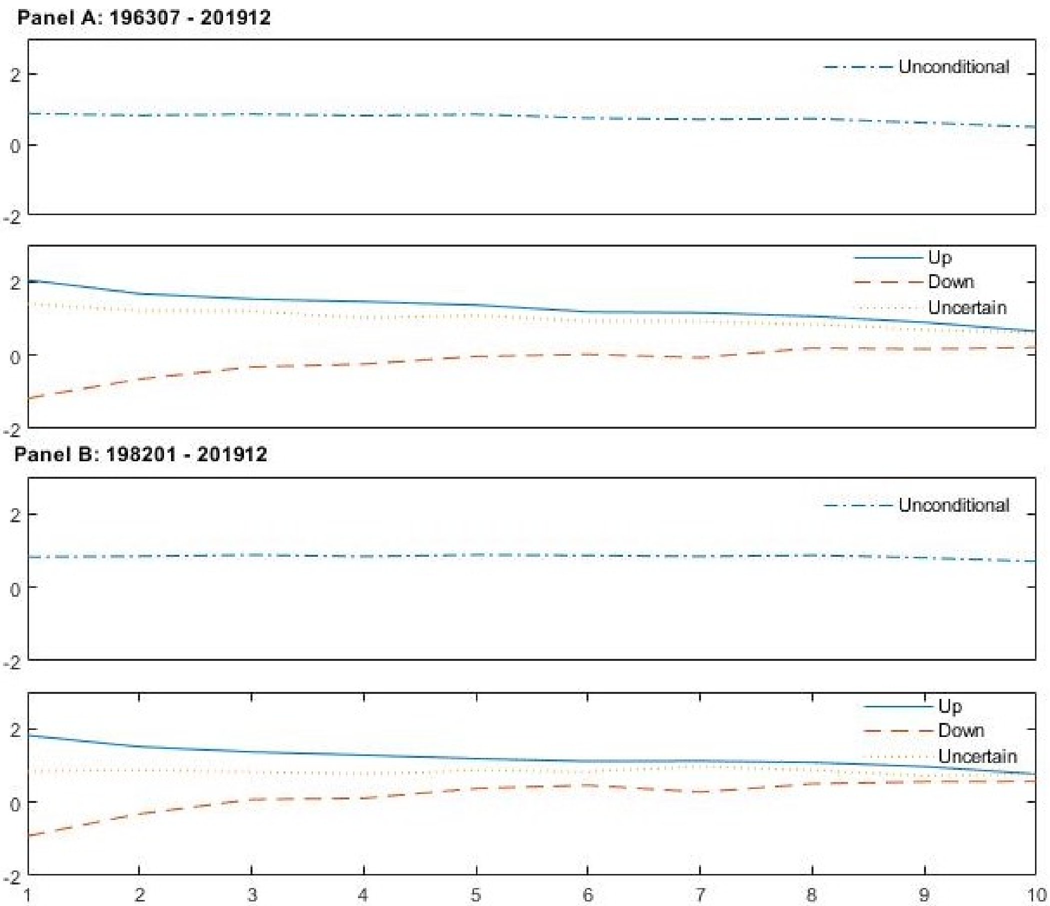
VAR model. VAR variables in the standard/panel VAR model are the aggregate-level/firm-specific returns and characteristics. Finally, each return component of a portfolio is calculated on a value-weighted basis using all stocks in that portfolio.

The difference in cash-flow and discount-rate shocks between small and big stocks after different signals

The sample period for this subsection is from July 1963 to December 2019 due to the data availability of VAR variables. To be included in the VAR estimation, a stock should have return and characteristic data with no missing values for at least two adjacent time points. The data requirements make the number of sample firms (and observations) different from that in previous sections. Therefore, I first check whether the two-faced size effect exists for this sample. Figure 1.4 depicts the unconditional and conditional returns of size-decile portfolios. Similar to those reported in Table 1.1, the returns of size-sorted portfolios decrease from the smallest to the largest deciles conditional on the up signal and increase conditional on the down signal, indicating that the conclusion on the existence of the two-faced size effect holds for this sample. The conditional SMB returns and cross-sectional return-size relationships are also similar to those reported in Section 1.3, which I do not report for brevity.

Table 1.7 reports the difference in cash-flow (CF) and discount-rate (DR) shocks between small and big stocks conditional on the forward-looking signals. I focus on examining whether CF or DR exhibits two faces. Panel A presents the CF and DR shocks of the SMB portfolio (constructed by buying/shorting stocks with a market capitalization below/above the NYSE 30th/70th percentile breakpoint

Figure 1.4: Unconditional and Conditional Returns of Size-Decile Portfolios
This figure depicts the average excess returns (monthly, %) of size-sorted portfolios using the sample of Section 1.5.4. At the end of each June, stocks are sorted into size deciles according to NYSE breakpoints; value-weighted returns are recorded at the end of each month. Panel A reports the returns for July 1963 to December 2019, and Panel B reports the returns for January 1982 to December 2019. The first subfigure of each panel depicts unconditional average returns. The second subfigure of each panel depicts average returns conditional on different signals (up/down/uncertain) provided by the IG indicator.



using the sample stocks). For the sample period of July 1963 to December 2019, conditional on the down signal, the average monthly cash-flow shock is -1.41%, and conditional on the up signal, the average monthly cash-flow shock is 0.66%, showing that the cash-flow shock has two faces. In contrast, the average discount-rate shock is of small magnitude and does not exhibit two faces notably (0.01% and 0.05% per month conditional on the up and down signals, respectively), indicating that the discount-rate channel has a negligible impact on the two-faced size effect. For the post-1980s period, we can also observe that the DR shock of the SMB portfolio trivially changes with the up and down signals (-0.03% and -0.04%, monthly), while a considerable two-faced size effect in CF shocks exists (0.40% and -1.41%, monthly).

Panel B presents the firm-level evidence to confirm the cash-flow channel. Following Section 1.3.2, I run the cross-sectional regressions of CF and DR shocks on firm size and other characteristics each month and evaluate the average slopes. Note that for the firm-level evidence, a positive slope means that small stocks are associated with a lower level of the left-hand-side variable than are big stocks. For the full sample period, the average monthly cross-sectional regression slope of CF on size is -0.15 conditional on the up signal and 0.46 conditional on the down signal. The average conditional slopes when the left-hand-side variable is DR have small absolute values. Controlling for other characteristics does not alter the conclusion. The results show again that the two-faced size effect is driven by the two-faced cash-flow shock.

The results in Table 1.7 also shed further light on the disappearance of the unconditional size effect since the 1980s. While the BMS spread in cash-flow shocks

TABLE 1.7: Source of the Two-Faced Size Effect

This table reports the cash-flow (CF) and discount-rate (DR) shocks conditional on different signals. The unconditional averages and averages conditional on the signals from the IG indicator are reported. CF and DR are extracted following Section 1.5.4. Panel A is for the portfolio-level evidence, that is, the CF and DR shocks of the SMB portfolio. The SMB portfolio is constructed by buying/shorting stocks with a market capitalization below/above the NYSE 30th/70th percentile breakpoint; the CF and DR shocks of the SMB portfolio are calculated on a value-weighted basis each month. Panel B is for the firm-level evidence, that is, the average cross-sectional regression slopes of CF and DR shocks on size and other characteristics. The regressions of different columns in Panel B differ in control variables or sample periods. Controls-Yes refers to that the book-to-market ratio (BM), prior two-to-twelve-month cumulative return ($\text{Ret}_{2,12}$), operational profitability (OP), and investment (INV) are controlled in the regression. Note that for the portfolio-level evidence, a positive value denotes that small stocks have higher CF (or DR) shocks than big stocks, while for the firm-level evidence, a negative value denotes that small stocks have higher CF (or DR) shocks than big stocks. The t-statistics are based on Newey and West (1987) adjusted standard errors. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The full sample period is July 1963 to December 2019.

	Panel A: Portfolio-Level Evidence				Panel B: Firm-Level Evidence							
	CF		DR		CF				DR			
	196307	198201	196307	198201	196307	198201	196307	198201	196307	198201	196307	198201
Unconditional	-0.12	-0.29*	0.03	-0.02	0.08**	0.13***	0.10***	0.15***	0.01	0.02***	0.03***	0.05***
	(-0.77)	(-1.66)	(1.15)	(-0.71)	(2.24)	(3.56)	(2.95)	(4.19)	(1.56)	(4.21)	(6.25)	(7.90)
Up	0.66***	0.40**	0.01	-0.03	-0.15***	-0.09*	-0.12***	-0.03	0.01**	0.02***	0.03***	0.04***
	(3.76)	(2.02)	(0.20)	(-0.90)	(-3.62)	(-1.81)	(-3.13)	(-0.65)	(2.36)	(2.68)	(4.65)	(5.28)
Down	-1.41***	-1.41***	0.05	-0.04	0.46***	0.48***	0.46***	0.45***	-0.01	0.03***	0.03***	0.05***
	(-5.62)	(-5.15)	(1.13)	(-1.00)	(8.59)	(7.70)	(8.93)	(6.75)	(-0.57)	(3.18)	(4.11)	(5.01)
Uncertain	0.08	-0.39	0.05	0.04	0.03	0.16**	0.06	0.16**	0.02	0.02**	0.04***	0.05***
	(0.19)	(-0.78)	(0.87)	(0.73)	(0.42)	(2.20)	(1.09)	(2.14)	(1.53)	(2.05)	(2.66)	(3.27)
Controls					No	No	Yes	Yes	No	No	Yes	Yes

conditional on the down signal for the post-1980s sample is comparable in magnitude to that of the full sample period (1.41% versus 1.41%), the magnitude of the SMB spread in cash-flow shocks conditional on the up signal for the post-1980s sample becomes smaller than that of the full sample period (0.40% versus 0.66%). A similar pattern can be observed from the firm-level evidence. Consequently, the average BMS spread in cash-flow shocks conditional on the down signal is offset less by the average SMB spread in cash-flow shocks conditional on the up signal over the sample period, which accounts for the disappearance of the unconditional size effect. Horowitz et al. (2000b) suggest that extensive financial products such as small-cap mutual funds emerging during this period cause the disappearance of the unconditional size effect. Following this rationale, it is probable that investors' pursuit of SMB profits drives down the difference in CF shocks between small and big stocks conditional on the up signal, resulting in the disappearance of the unconditional size effect.

Overall, the evidence in this section suggests that the two-faced size effect originates from smaller stocks' slower response to cash-flow shocks, consistent with the impact of information barriers on small stocks documented in previous studies (see Section 1.2.1). The dominant role of cash-flow shocks is consistent with the finding of Hou and van Dijk (2019) that the negative average cash-flow shock causes the disappearance of the size premium since the 1980s and the finding of Lochstoer and Tetlock (2020) that the time-variation of anomaly returns is mainly driven by cash-flow shocks.

1.6 Additional Tests

This section examines whether the two-faced size effect remains robust when controlling for stylized facts that invalidate the unconditional size effect.

1.6.1 Interaction with the January effect

It is argued that the size effect is extraordinarily strong in January and nearly disappears during other months (e.g., Keim 1983; Brown et al. 1983; Lamoureux and Sanger 1989). Due to the consequential impact of the January effect on the size effect, it is important to investigate whether the two-faced size effect remains robust when considering it. The first two columns of Table 1.8 present the unconditional size effect for different months. The average January SMB return is 3.79% per month and is significant at the 1% level, confirming the significance of the January size effect. In contrast, the average February-to-December SMB return is negligible (0.01%), consistent with the January effect's dominance on the unconditional size effect documented in previous studies. The size effect over February to December revives once I consider the two-faced effect. Conditional on the up (down) signal, for the months from February to December, I uncover a considerable SMB (BMS) spread in monthly returns of 0.92% (1.30%). The results indicate that although the unconditional size effect does not survive after removing the January effect, the two-faced size effect remains robust. The same conclusion holds for the post-1982 period.

TABLE 1.8: Two-Faced Size Effect after Controlling for the January Size Effect
This table reports the average January and February-to-December returns of the SMB portfolio. The SMB portfolio is the same as that in Table 1.1. Panel A is for the result of the full sample, and Panel B is for the result of the post-1982 sample. The first two columns report unconditional returns, and the remaining columns report average February-to-December returns conditional on different signals. The t-statistics are based on Newey and West (1987) adjusted standard errors. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The sample period is from August 1926 to December 2019.

Unconditional		Up	Down	Uncertain
Jan	Feb-Dec		Feb-Dec	
Panel A: The Full Sample				
3.79***	0.01	0.92***	-1.30***	-0.14
(5.44)	(0.03)	(3.89)	(-5.72)	(-0.39)
Panel B: The Post-1982 Sample				
1.12	-0.09	0.60***	-1.20***	-0.12
(1.49)	(-0.52)	(2.76)	(-4.34)	(-0.26)

1.6.2 Interaction with the business cycle

Ahn et al. (2019) find that the size effect is indistinguishable from zero during the all-but-trough stage of the business cycle (i.e., explosion, peak, and recession stages). This subsection shows that the two-faced size effect remains strong and significant for the all-but-trough state of the business cycle. Different stages of the business cycle are defined according to the business cycle turning points identified by the National Bureau of Economic Research (NBER). The NBER Business Cycle Dating Committee determines turning points, each denoted by one particular month, based on its ex-post judgment on absolute declines in a wide spectrum of economic measures. Following Ahn et al. (2019), I define the trough stage by including three months before and after each trough month identified by the NBER, which ensures that a sufficient number of months are assigned to this stage. For August 1926 to December 2019, 105 of the 1121 months are in the trough stage,

TABLE 1.9: Two-Faced Size Effect after Controlling for the Business Cycle Effect
This table reports the average returns of the SMB portfolio during the trough and all-but-trough (i.e., explosion, peak, and recession stages) stages of the business cycle. The SMB portfolio is the same as that in Table 1.1. Different stages of the business cycle are defined according to NBER business cycle turning points following Ahn et al. (2019). Panel A is for the result of the full sample, and Panel B is for the result of the post-1982 sample. The first two columns report unconditional returns, and the remaining columns report the returns of the all-but-trough period (Others) conditional on different signals. The t-statistics are based on Newey and West (1987) adjusted standard errors. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The sample period is from August 1926 to December 2019.

Unconditional		Up	Down	Uncertain
Trough	Others		Others	
Panel A: The Full Sample				
2.33***	0.11	0.93***	-1.13***	0.24
(2.69)	(0.73)	(4.64)	(-5.03)	(0.66)
Panel B: The Post-1982 Sample				
1.76***	-0.10	0.55**	-1.19***	-0.08
(3.84)	(-0.58)	(2.45)	(-4.44)	(-0.18)

and for the post-1982 period, 28 of the 428 months are in the trough stage. Table 1.9 reports the unconditional size effect and the size effects conditional on the business cycle. As shown in the first two columns, the unconditional size effect is considerable and significant for the trough stage but is indistinguishable from zero for the all-but-trough stage, confirming the finding of Ahn et al. (2019). The size effect revives once I consider the two-faced effect, as suggested by the remaining columns. Conditional on the up (down) signal, for the all-but-trough stage, I uncover an SMB (BMS) spread in monthly returns of 0.93% (1.13%) for the full sample period and 0.55% (1.19%) for the post-1982 period. The results confirm the robustness of the two-faced size effect.

1.6.3 Excluding microcap stocks

Another common critique of the size effect is that it is limited to the smallest stocks (e.g., Horowitz et al. 2000b; Crain 2011; Alquist et al. 2018). In response to this critique, I remove the whole universe of microcap stocks (stocks with a market capitalization below the NYSE 20th percentile breakpoint) and investigate whether the two-faced size effect can still be observed. As the overall descending and ascending patterns of size-sorted portfolio returns conditional on the up and down signals shown in Table 1.1 have already shed light on this issue, this subsection provides firm-level evidence for further confirmation. Table 1.10 reports the unconditional and conditional cross-sectional regression slopes of return on size and other characteristics using all-but-microcap stocks. The significantly negative slope conditional on the up signal (-0.21) and the positive slope conditional on the down signal (0.21) indicate that the two-faced size effect exists among these stocks. This conclusion does not change when I control for other characteristics. Even for the post-1980s period, the two-faced size effect still exists and remains significant. Therefore, the two-faced size effect is not limited to the smallest stocks. In an unreported test, I find that compared with the result of Table 1.2, the absolute values of the slopes here are significantly smaller, which raises the question of whether the two-faced size effect among all-but-microcap stocks is economically significant. To address this issue, I show in Appendix A3 that the two-faced size strategy on the SMB portfolio constructed from all-but-microcap stocks generates abnormal returns unexplained by prominent asset pricing models.

Table 1.10: Two-Faced Size Effect on All-But-Microcap Stocks

This table reports the firm-level cross-sectional regression slopes of return on size using all-but-microcap stocks. Microcap stocks refer to stocks with a market capitalization below the NYSE 20th percentile breakpoint. Regressions for different columns differ in control variables and sample periods. The control variables include the book-to-market ratio (BM), prior two-to-twelve-month cumulative return ($Ret_{2,12}$), operational profitability (OP), and investment (INV). The signals are from the IG indicator. The t-statistics are based on Newey and West (1987) adjusted standard errors. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The full sample period is from August 1926 to December 2019.

	192608	198201	192608	198201	192701	198201	196307	198201
Unconditional	-0.06**	0.00	-0.04	0.01	-0.05*	0.00	-0.05	-0.02
	(-1.96)	(0.12)	(-1.51)	(0.38)	(-1.79)	(0.08)	(-1.43)	(-0.46)
Up	-0.21***	-0.10**	-0.22***	-0.13***	-0.23***	-0.14***	-0.20***	-0.13***
	(-5.36)	(-1.98)	(-6.40)	(-2.59)	(-7.06)	(-2.90)	(-5.00)	(-2.77)
Down	0.21***	0.20***	0.22***	0.23***	0.22***	0.22***	0.21***	0.16***
	(4.08)	(3.04)	(4.57)	(3.55)	(4.85)	(3.41)	(4.06)	(2.75)
Uncertain	-0.11	-0.02	-0.02	0.06	-0.01	0.06	-0.09	0.02
	(-1.56)	(-0.27)	(-0.26)	(0.62)	(-0.15)	(0.63)	(-1.19)	(0.18)
BM			0.35**	0.41	0.44***	0.52*	0.54***	0.47*
			(2.32)	(1.38)	(3.20)	(1.93)	(2.74)	(1.90)
$Ret_{2,12}$					1.34***	0.89***	1.07***	0.77***
					(9.47)	(4.08)	(6.04)	(3.70)
OP							0.28***	0.22**
							(4.03)	(3.17)
INV							-0.38***	-0.28**
							(-3.23)	(-2.39)

1.6.4 Interaction with the quality/junk effect

Asness et al. (2018) explore the interaction between the size effect and the quality/junk effect and have two findings: 1) the connection between return and size breaks down among junk stocks; 2) small-junk stocks tend to underperform big-quality stocks. This subsection examines whether the two-faced size effect exists among junk stocks and whether the two-faced size effect can recover a size premium between small-junk and big-quality portfolios. I use the three-by-two quality-size-sorted portfolios available from the AQR website. The full sample period is from July 1957 to December 2019. Table 1.11 presents the unconditional and conditional monthly returns of the small-junk-minus-big-junk (SL-BL) and small-junk-minus-big-quality (SL-BH) portfolios. As shown in the first column of Panel A, the unconditional monthly returns of SL-BL and SL-BH are 0.08% and -0.18%, consistent with the two implications from Asness et al. (2018). However, as shown by the second and third columns, the two-faced size effect exists not only among junk stocks but also between small-junk and big-quality portfolios. For SL-BL, the return is 0.79% per month conditional on the up signal and -1.13% conditional on the down signal, both significant at the 1% level. For SL-BH, the return is 0.72% conditional on the up signal and -1.67% conditional on the down signal, also significant at the 1% level. The results for the post-1980s period reported in Panel B lead to the same conclusion. As the two-faced size effect remains robust when controlling for the interaction between size and quality/junk, the size effect is not limited to quality stocks if evaluated conditionally. Instead of attributing the unconditional size effect's disappearance to a subgroup of stocks, my findings suggest a conditional manifestation mechanism.

1.6.5 International evidence

Some studies, such as Alquist et al. (2018) and Asness et al. (2018), find that the unconditional size effect does not hold in other equity markets. This subsection evaluates the unconditional and conditional returns of the size factors of different regions, including the size factor of the twenty-two developed markets without the U.S. (DpExUS), the size factor of the twenty-five emerging markets (Emerging), the size factor of the sixteen European markets (Europe), the size factor of the

Table 1.11: Two-Faced Size Effect and Quality/Junk Effect
This table reports the unconditional and conditional average returns of the small-junk-minus-big-junk (SL-BL) and small-junk-minus-big-quality (SL-BH) portfolios. Panel A is for the results of the full sample, and Panel B is for the result of the post-1982 sample. SL-BL and SL-BH are constructed from the three-by-two quality-size-sorted portfolios available from the AQR website. The signals are from the IG indicator. The t-statistics are based on Newey and West (1987) adjusted standard errors. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The full sample period is from July 1957 to December 2019.

	Unconditional	Up	Down	Uncertain
Panel A: The Full Sample				
SL-BL	0.08	0.79***	-1.13***	0.28
	(0.63)	(5.13)	(-5.54)	(0.85)
SL-BH	-0.18	0.72***	-1.67***	-0.03
	(-1.11)	(3.61)	(-5.86)	(-0.06)
Panel B: The Post-1982 Sample				
SL-BL	-0.16	0.57***	-1.30***	-0.38
	(-1.03)	(2.75)	(-5.09)	(-0.95)
SL-BH	-0.48**	0.56**	-2.11***	-0.76
	(-2.23)	(2.06)	(-5.59)	(-1.53)

Japanese market (Japan), and the size factor of the four Asia-Pacific markets excluding Japan (APExJapan). The data are available from Kenneth French's website, and the sample period is from July 1990 to December 2019. The IG indicator of each region is constructed using its corresponding big- and small-stock returns.

As suggested in Table 1.12, the unconditional size premiums are slightly positive or even negative (all of them are less than 0.10% per month), indicating that the unconditional size effect is weak or does not exist in these markets during this period. While the empirical evidence does not support the existence of the unconditional size effect, there is a strong and significant two-faced size effect. Most average returns are significantly positive conditional on the up signal

(0.48%, 0.48%, 0.61%, 0.54%, and 0.35%, respectively) and significantly negative conditional on the down signal (-0.77%, -0.45%, -0.98%, -0.60%, and -1.24%, respectively). The international evidence also shows the robustness of the two-faced size effect.

TABLE 1.12: International Evidence for the Two-Faced Size Effect

The table reports the unconditional and conditional average returns of the size factors of different international markets. DpExUS refers to the size factor of developed countries except for the U.S. APExJapan refers to Asia-Pacific countries except for Japan. The data are available from Kenneth French's website. The signals are given by the IG indicator constructed using the excess small- and big-stock returns of each region. The original data are the value-weighted returns of two-by-three size-BM-sorted portfolios. I calculate the small(big)-stock return as the average of the three BM portfolios of stocks with a small (big) market capitalization in excess of the risk-free rate. The t-statistics are based on Newey and West (1987) adjusted standard errors. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The sample period is July 1990 to December 2019.

	DpExUS	Emerging	Europe	Japan	APExJapan
Unconditional	0.01	0.09	-0.01	0.03	-0.28*
	(0.07)	(0.71)	(-0.09)	(0.20)	(-1.80)
Up	0.48***	0.48***	0.61***	0.54**	0.35
	(3.60)	(2.80)	(4.10)	(2.38)	(1.61)
Down	-0.77***	-0.45**	-0.98***	-0.60**	-1.24***
	(-4.33)	(-2.05)	(-5.47)	(-2.07)	(-4.77)
Uncertain	0.44	0.07	0.23	0.17	-0.15
	(1.35)	(0.18)	(0.72)	(0.41)	(-0.43)

1.7 Conclusion

This study uncovers a positive big-minus-small premium conditional on a down signal (the BMS effect) and a positive small-minus-big premium conditional on an up signal (the SMB effect). Managing a simple SMB portfolio according to the two-faced size effect achieves considerable and significant performance improvements over the unmanaged SMB portfolio, and the enhanced performance is unexplained by common risk factors. The BMS effect is comparable in strength to the SMB effect, especially for the post-1980s period, accounting for the unconditional size effect's poor track record and post-discovery disappearance.

In additional tests, I show that the two-faced effect remains robust when controlling for stylized facts that usually refute the validity of the unconditional size effect. The two-faced size effect remains significant after controlling for the January effect or the business cycle, is statistically and economically significant among all-but-microcap stocks, exists in other financial markets outside the U.S., and exists among junk stocks or between small-junk and big-quality stocks. The performance of the two-faced size effect over time and across different groups of stocks is consistent with the mechanism that the prices of small stocks slowly respond to information shocks. By extracting cash-flow and discount-rate shocks from realized returns employing a VAR-based return decomposition, I further find that the two-faced size effect is mainly driven by cash-flow shocks.

My findings show that the slow information incorporation of the prices of small stocks translates into predictable SMB and BMS effects. The BMS and SMB effects are much stronger than the unconditional size effect, and hence their relative realized strength over a sample period directly determines whether the unconditional size effect is observed. Therefore, future studies about the size effect should consider its two faces when assessing its validity or economic benefits.

Chapter 2

Uncovering Omitted Factors through a New Low-Beta Anomaly

2.1 Introduction

Contrary to the positive risk-return trade-off prediction of the CAPM (Sharpe 1964; Lintner 1965; Mossin 1966), empirical studies find that low market-beta stocks outperform high market-beta stocks on a risk-adjusted basis (e.g., Friend and Blume 1970; Brennan 1971; Black et al. 1972; Fama and MacBeth 1973; Fama and French 2004). The negative beta-alpha relationship spurs the low-beta anomaly (LBA) literature and various refinements of the CAPM. In the past decade, various explanations have been proposed to fully subsume the anomaly, including factor models with investment and profitability factors (e.g., Fama and French 2015; Hou et al. 2015; Novy-Marx and Velikov 2022) or specific economic sources such as leverage constraints, lottery demand, arbitrage asymmetry or coskewness risk (e.g., Frazzini and Pedersen 2014; Bali et al. 2017; Liu et al. 2018; Schneider et al. 2020).

While recent studies resolve the LBA known in the literature (the known LBA), conditional on these studies, I find a new LBA unexplained by these studies. For example, while Novy-Marx and Velikov (2022) show that LBA derives its abnormal performance purely from the positive exposures to investment and profitability factors, the new LBA I obtain is not exposed to the two factors but still achieves a high Sharpe ratio.[1]

I argue that the new LBA is driven by return sources different from those underlying the known LBA. To demonstrate the rationale, I first show that LBA is a manifestation of factors unknown to econometricians in a multifactor world. Then I show that, theoretically, the known and new LBAs are manifestations of different types of omitted factors that differ in their correlation with market risk. The known LBA is driven by *factors directly correlated with market risk*. In contrast, the new LBA is driven by *factors not directly but only partially correlate with market risk* (partial-correlation factors). The new LBA's significance is that it uncovers the existence of partial-correlation factors, which are also important for driving asset returns but were unnoticed before.

The formal analysis starts by clarifying the general mechanism for LBA to emerge under a factor-model framework. Specifically, I adopt the view that econometricians with only public information are less informed than investors with both

[1] Over the sample period of this study, the annualized Sharpe ratio is 0.48 for the known LBA and 0.56 for the new LBA.

public and private information (e.g., Hansen and Richard 1987; Andrei et al. 2021). Accordingly, I set the econometricians' model used for beta estimation as a factor model with fewer factors than the ideal model that captures all information relevant to asset returns. As a result, estimated betas also capture exposures to factors unknown to econometricians. These extra factor exposures are the origin of LBA.

Specifically, when econometricians estimate betas through the CAPM, an estimated market beta is a weighted average of the true market beta (that purely reflects market risk) and exposures to omitted factors; the weights depend on the correlations between omitted factors and the market risk. Suppose omitted factors negatively correlate with market risk.[2] In this case, sorting stocks into estimated-beta quantiles tends to pick up stocks with high/low omitted factors exposures in the low/high-beta side. As a result, a portfolio buying/selling low/high-beta stocks will be positively exposed to omitted factors. This positive exposure is compensated by an extra premium unexplained by market risk, accounting for the LBA known in the literature.

The explicit factor-based explanation of LBA predicts that LBA is not limited to what has been documented in the literature. As a manifestation of omitted factors, the economic content of LBA depends on the omitted factors it reflects, which in turn depends on the information captured by estimated betas. I show that estimated betas only capture factors directly correlated with market risk. The rationale is that how an estimated beta captures the exposure to an omitted

[2]In this study, "a factor negatively correlates with market risk" refers to both a negative time-series correlation between factor returns and a negative cross-sectional correlation between factor exposures. See Section 2.3.1 for the rationality of this setup.

factor purely depends on how the omitted factor correlates with market risk. Consequently, the LBA known in the literature is a manifestation of factors directly correlated with market risk.

In contrast, factors not directly correlated with market risk are not captured by estimated betas. Accordingly, these factors are not reflected in the known LBA, regardless of how important they are for driving asset returns. However, I find that these factors can also induce an LBA if they are partially correlated with market risk. The rationale is that partial-correlation factors start to correlate with market risk once the impact of factors directly correlated with market risk is removed. In this case, estimated betas, besides capturing true market risk, no longer capture factors directly correlated with market risk but only capture exposures to partial-correlation factors. These extra factor exposures still tend to induce an LBA (see Corollary 2.4.2). If such a new LBA indeed exists, it is a pure manifestation of partial-correlation factors.

Recent studies fully resolve the known LBA from different perspectives, which essentially provide good proxies for factors directly correlated with market risk. This advance in the literature enables direct examinations of the major theoretical findings of this study. My empirical analysis starts by interpreting recent explanations of the known LBA under the framework of this study. For example, consider the investment and profitability factor-based explanation (e.g., Hou et al. 2015; Fama and French 2016; Barroso et al. 2020; Novy-Marx and Velikov 2022). Consistent with my model, sorting stocks according to CAPM-betas inversely picks up exposures to investment and profitability factors as the two factors negatively correlate with the market factor. These extra factor exposures are associated with a positive betting-against-beta alpha. The positive alpha is fully subsumed by the investment and profitability factors, indicating that betas estimated from the CAPM, besides capturing the true market risk, mainly capture exposures to the two factors. Therefore, the LBA known in the literature is a pure manifestation of investment and profitability factors.

As predicted by my model, a new LBA can be uncovered with the impact of factors underlying the known LBA removed. I confirm this prediction by constructing a betting-against-beta portfolio based on market betas estimated from the six-factor model (Fama and French 2018). The new betting-against-beta portfolio generates abnormal performance again, which, however, is orthogonal to investment and profitability factors. As the impact of factors underlying the known LBA is removed, according to my model, this abnormal performance must be driven by partial correlation factors. Examining other sufficient explanations of LBA (e.g., leverage constraints, lottery demands, coskewness risk and arbitrage asymmetry) leads to the same conclusion. The results indicate that existing explanations of LBA are all proxies for factors directly correlated with market risk but do not capture information related to partial-correlation factors. This finding is not surprising. Before recognizing that LBA can be manifestations of different types of factors, the focus of existing studies is to resolve the known LBA. As my model suggests, only conditional on resolving the known LBA can we uncover the new LBA as a manifestation of partial-correlation factors.

Essentially, factors directly correlated with market risk (captured by the known LBA) correspond to direct violations of the cross-sectional beta-return relationship regulated by the CAPM. In contrast, partial-correlation factors (captured by the new LBA) do not directly violate the CAPM. To the extent that asset pricing factors or anomalies are discovered through violations of the CAPM, the existence of partial-correlation factors should hardly be noticed. Empirically, I find that the new LBA's abnormal performance is not explained by time-varying risk exposures, nor is it notably related to any of the factors/anomalies from Hou et al. (2020b). The results confirm that the partial-correlation factors' impact was not noticed in previous studies.

The finding signifies a unique advantage of documenting factors or anomalies using betas: betas enable learning about the unknowns. Since the documentation of the size effect (Banz 1981), asset pricing studies have been capturing factors or anomalies predominantly through characteristics (e.g., Fama and French 1992; Jegadeesh and Titman 1993; Cooper et al. 2008); Novy-Marx 2013). Characteristics are reliably available over time for individual stocks and are usually not subject to estimation errors, hence allowing updating information efficiently. However, the characteristic approach only works after certain characteristics relevant to the cross-section of returns are discovered by coincidence or advancing theories. Instead, the finding that LBA is a manifestation of omitted factors suggests that, through betas, we can learn about unknown factors before recognizing the corresponding characteristics. The new LBA suggests that a certain type of omitted factor has been overlooked, and we can allow estimated betas to capture it. To the extent that the new LBA captures the mean-variance efficiency and time-series performance of partial-correlation factors, we can infer the properties of these overlooked factors.

Regarding learning about the unknowns, the finding is also related to the factor model literature. The existence of LBA indicates that the standard CAPM omits pricing information. Accordingly, previous studies discover additional factors to update the model. From a pure factor-model perspective, for example, although the investment and profitability factors were formally introduced into asset-pricing models in the 2010s, their existence was already implied by the discovery of LBA in the 1970s. Given this history, it is reasonable to conjecture that the new LBA, which has a higher Sharpe ratio than the known LBA, indicates the existence of factors not covered by prevailing asset pricing models (e.g., Fama and French 2015, 2018; Hou et al. 2015; Barillas and Shanken 2018; Hou et al. 2020a). This implication is consistent with studies suggesting that more factors are needed in factor models (e.g., Lewellen 2022). Moreover, the finding that the new LBA identifies overlooked factors also implies that it is necessary to include portfolios sorted by multifactor betas as standard test assets when evaluating asset-pricing models.

While theoretical studies of LBA usually start with a general equilibrium economy to generate LBA endogenously, I start with a reduced-form multifactor model. As such, my analysis is not about modeling an equilibrium economy but is about inferring an equilibrium economy's properties using observed LBAs. The general practice in the LBA literature of building an equilibrium economy is to release one assumption of the standard CAPM. For example, Jagannathan and Wang (1996) allows the market premium and betas to be time-varying and update the CAPM with a factor that captures beta instability. Frazzini and Pedersen (2014) add leverage constraints to investors' mean-variance optimization and generate an expected return linear in both the market premium and tightness of funding constraints.

Hong and Sraer (2016) allow investors to disagree on the stock market's prospects, while Andrei et al. (2021) set expected returns to vary across investors. This general practice essentially introduces a direct violation of the standard CAPM, which thus can only generate additional factors that directly correlate with market risk. The fact that no previous studies uncovered the new LBA supports this implication. The remarkable empirical performance of the new LBA indicates that factors only violating the CAPM in a partial sense (i.e., violating the CAPM when direct violations are controlled) are also important for asset returns. Therefore, a more realistic equilibrium economy should also reconcile partial-correlation factors.

This study attributes LBA to the information gap between investors and econometricians, which is consistent with the insight of Andrei et al. (2021). However, instead of setting the investors' model as a dynamic CAPM, I set it as a general multifactor model with one market factor and at least two additional types of factors. Another difference is that there are two types of econometricians in my model. The first type, pre-resolution econometricians, observes a violation of the CAPM, the same as that in Andrei et al. (2021). The second type, post-resolution econometricians, analyzes the economy conditional on the resolution of the known LBA and hence observes a new LBA. These setups are crucial for unveiling the novel finding that LBA is not limited to what has been documented in the literature.

Uncovering a new LBA also refreshes the understanding of LBA in many other aspects. For example, previous studies regard low-beta and low-volatility anomalies as the same. In this regard, Bali et al. (2011) and Bali et al. (2017) respectively find that the lottery-demand effect explains the low-volatility and low-beta anomalies. Stambaugh et al. (2015) and Liu et al. (2018) respectively show that the arbitrage-asymmetry effect explains the low-volatility and low-beta anomalies. Schneider et al. (2020) find that a single principal component drives returns of beta and volatility-sorted portfolios. My finding indicates that while the known LBA shares a similar driver with the low-volatility effect, the new LBA is orthogonal to the low-volatility effect. The remarkable Sharpe ratio of the new LBA also makes it practically important as it still generates alphas, while the known low-risk strategies no longer generate abnormal performance if adjusting for recent asset-pricing models. To the least, the finding translates into a higher bar for future studies to resolve LBA: a sufficient explanation should also resolve the new LBA.

The remainder of this paper is structured as follows. Section 2.2 previews the major propositions through empirical examples. Section 2.3 clarifies the mechanism for LBA to emerge. Section 2.4 shows that, theoretically, a new LBA different from the LBA known in the literature can exist. Section 2.5 conducts the empirical analysis, followed by additional discussions in Section 2.6. Section 2.7 provides concluding remarks.

2.2 Preview of Major Propositions

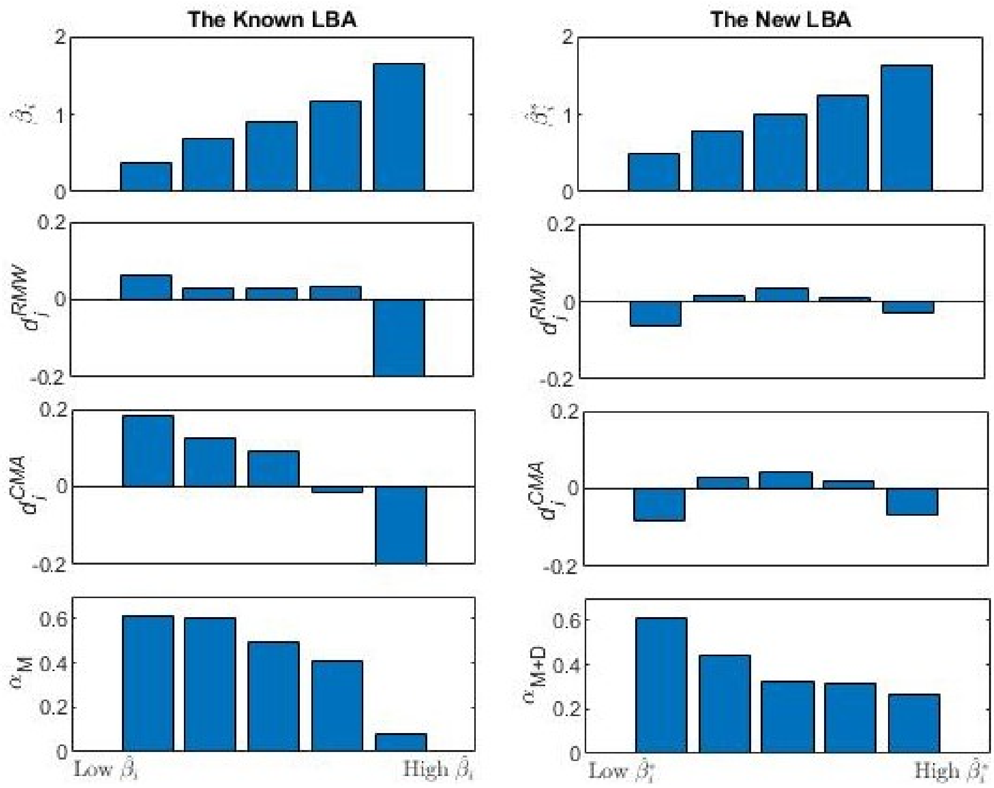
Before the formal analysis, this section presents the known and new LBAs empirically, setting the stage for understanding this study's major propositions. The left panel of Figure 2.1 presents the LBA known in the literature (see Appendix B2.1 for details). Sorting stocks into quantiles according to betas estimated from the CAPM (the top-left subfigure) generates a negative beta-alpha relationship (the bottom-left subfigure). Consistent with recent studies (e.g., Novy-Marx and Velikov 2022), this LBA is explained by investment and profitability factors, as shown by its positive exposures to the two factors (higher/lower d_i^{RMW} and d_i^{CMA} in the low/high-$\hat{\beta}_i$ side as shown in the middle two subfigures of the left panel).

I show that the pattern in the left panel of Figure 2.1 emerges because LBA is a manifestation of factors unknown to econometricians (see Proposition 1 in Section 2.3). In the context of the left panel of Figure 2.1, LBA is a manifestation of investment and profitability factors; this proposition also reconciles other explanations of LBA. Based on the analytical framework of this proposition, I further unveil a novel feature of LBA: it is not limited to what is known previously, but a new LBA can exist.

The right panel of Figure 2.1 presents the new LBA uncovered in this study. I re-estimate market betas with factors underlying the known LBA controlled to remove their impact. Using these new-estimated betas, sorting stocks into beta quantiles (the top-right subfigure) still generates a negative beta-alpha relationship (the bottom-right subfigure). However, this new LBA is no longer exposed to investment and profitability factors (similar d_i^{RMW} and d_i^{CMA} in the low/high-$\hat{\beta}_i$ side as shown by the middle two subfigures of the right panel). The result suggests that the LBA in the right panel should be driven by other economic sources.

FIGURE 2.1: The Known and New Low-Beta Anomalies (LBAs)

The left panel reports the LBA known in the literature, and the right panel reports the new LBA uncovered in this study. $\hat{\beta}_i$ is the market beta estimated from the CAPM. $\hat{\beta}_i^*$, d_i^{RMW} and d_i^{CMA} are market betas and exposures to investment and profitability factors estimated from the six-factor model (Fama and French 2018). α_M is the spanning regression alpha (monthly, %) against the CAPM (α_M). α_{M+D} is the spanning regression alpha (monthly, %) against the six-factor model. The sample period is from July 1968 to December 2019. Empirical details are in Appendix B2.1. Section 2.4 provides the theoretical foundation for the new LBA.

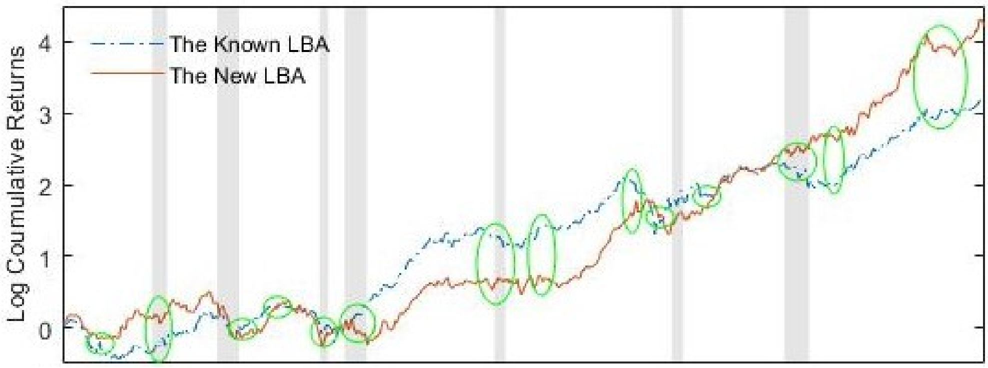


The two LBAs are driven by different types of factors (see Proposition 2 in Section 2.4). While the known LBA (in the left panel) is a manifestation of *factors directly correlated with market risk*, the new LBA (in the right panel) is a manifestation of *factors only partially correlated with market risk* (referred to as partial-correlation factors).[3] A major significance of the new LBA is that it uncovers partial-correlation factors, which are important for driving asset returns but were not noticed before.

Figure 2.2 presents the cumulative returns of the known and new LBAs to show their difference intuitively (see Appendix B2.2 for details). The two LBAs have different performance for many periods, especially during recession periods (the shaded bars). For example, they behave inversely during the two worst economic downturns since the Great Depression.[4]

FIGURE 2.2: Performance of the known and new Low-Beta Anomalies (LBAs)

This figure reports the log of the cumulative returns of the known LBA (the dotted-dashed line) and new LBA (the solid line). The shaded bars correspond to NBER recession periods. The sample period is from July 1968 to December 2019. Empirical details are in Section B2.2.



Next, in the theoretical analysis of this study, I first use a succinct factor-model framework to rationalize LBA, showing that it is a manifestation of factors

[3]In the context of this study, partial-correlation factors become correlated with the market risk once the impact of factors directly correlated with the market risk is removed. See Definition 1 in Section 2.4 for details.

[4]See https://www.federalreservehistory.org/essays/recession-of-1981-82. The first is the recession of 1981-82, triggered by tight monetary policy to fight inflation, and the second is the Great Recession of 2007-09, triggered by mortgage-related financial assets.

omitted by market risk. Then, under this framework, I demonstrate why a new LBA driven by partial correlation factors should exist and how to uncover the new LBA empirically.

2.3 Rationalizing the Low-Beta Anomaly (LBA)

This section rationalizes LBA in a succinct factor-model framework. Adopting the view that empiricists have less information than investors, the model econometricians use for beta estimation should have fewer factors than the ideal asset-pricing model that fully captures asset returns. Accordingly, estimated betas capture factors omitted by the econometricians' model, which is the origin of LBA. This analytical framework paves the way for uncovering a new LBA in Section 2.4.

2.3.1 The latent ideal asset-pricing model

To ensure that this study unveils an intrinsic feature of LBA, the ideal asset-pricing model capturing all pricing information should be sufficiently general. Based on the discussion in Appendix B1.1, I set the ideal model as a multifactor model, Equation (2.1). Then, I show that Equation (2.1) can be transformed into a succinct version, Equation (2.5), without changing the information it captures. Under Equation (2.5), different return sources only interact through the time-series correlations between factor returns. Equation (2.5) is used as the benchmark model for the theoretical analysis.

Preparatory step 1: starting with a general multifactor model

The information set of investors should contain all information relevant to pricing if we adopt the view that they together (retail and institutional investors, etc.) are both publicly and privately informed (e.g., Hansen and Richard 1987; Andrei et al. 2021). I use a multifactor model with a strict factor structure to summarize this complete information set:

$$r_{it} - r_f = \alpha_i + \beta_i^o M_t + \boldsymbol{d}_i^{oT} \boldsymbol{D}_t^o + \boldsymbol{c}_i^{oT} \boldsymbol{C}_t^o + \epsilon_{it} \tag{2.1}$$

where $r_{it} - r_f$ is the excess return of security i; M_t is the market factor; $\boldsymbol{D}_t^o = [D_{1t}^o \dots D_{kt}^o \dots D_{Kt}^o]^T$ and $\boldsymbol{C}_t^o = [C_{1t}^o \dots C_{st}^o \dots C_{St}^o]^T$ are the K-by-1 and S-by-1 vectors of factors capturing return sources beyond market risk. Without loss of generality, assume all priced factors have positive expected returns. $\boldsymbol{D}_t^o$ and $\boldsymbol{C}_t^o$ differ in their connection with the market, which will be defined in Section 2.4. β_i^o, $\boldsymbol{d}_i^o = [d_{1i}^o \dots d_{ki}^o \dots d_{Ki}^o]^T$ and $\boldsymbol{c}_i^o = [c_{1i}^o \dots c_{si}^o \dots c_{Si}^o]^T$ are the true factor exposures.

α_i and ϵ_{it} are the pure idiosyncratic components of security i's return. ϵ_{it} does not correlate with factor returns in the time series. α_i and ϵ_{it} do not correlate with factors exposures in the cross-section, that is,

$$Cov([M_t \ \boldsymbol{D}_t^{oT} \ \boldsymbol{C}_t^{oT}],\ \epsilon_{it}) = \boldsymbol{0};\ Cov_{CS}([\beta_i^o \ \boldsymbol{d}_i^{oT} \ \boldsymbol{c}_i^{oT}],\ \alpha_i + \epsilon_{it}) = \boldsymbol{0} \tag{2.2}$$

where Cov_{CS} refers to the cross-sectional covariance. $[M_t \ \boldsymbol{D}_t^{oT} \ \boldsymbol{C}_t^{oT}]$ and $[\beta_i^o \ \boldsymbol{d}_i^{oT} \ \boldsymbol{c}_i^{oT}]$ are $(1+K+S)\times 1$ vectors of factors and factor exposures. $\boldsymbol{0}$ is the $(1+K+S)\times 1$ vector of zeros. Otherwise, the components of α_i and ϵ_{it} related to factors can be captured by adding additional factors to Equation (2.1). Note that $\alpha_i = 0$

should hold if Equation (2.1) is derived from a general equilibrium framework or an arbitrage model (e.g., both Merton 1973 and Ross 1976 imply that the alpha is zero).

Preparatory step 2: simplifying Equation (2.1)

A problem with a multifactor model is that it reflects time-series correlations between factor returns but does not reflect cross-sectional correlations between factor exposures. Modeling the cross-sectional correlation between market betas and other factor exposures (or characteristics/behavioral biases, depending on the interpretation) is more popular in the literature. For example, under the leverage constraints-based explanation (e.g., Black 1972; Frazzini and Pedersen 2014; Jylhä 2018), stocks with higher betas are associated with higher demand from leverage-constrained investors. Simultaneously analyzing the cross-sectional and time-series correlations will complicate the analysis. This subsection shows that we can simplify the analysis by only focusing on time-series correlations between factor returns through the following transformation:

Suppose that market betas and other factor exposures are correlated in the cross-section:

$$\beta_i^o = \beta_i + \boldsymbol{\xi}_d^T \boldsymbol{d}_i^o + \boldsymbol{\xi}_c^T \boldsymbol{c}_i^o \tag{2.3}$$

where $\boldsymbol{\xi}_d$ and $\boldsymbol{\xi}_c$ capture the cross-sectional correlation between factor exposures. β_i is the true market beta in the sense that it only reflects market risk. The cross-sectional correlations between β_i and other factor exposures are zero, that is, $Cov_{CS}(\beta_i, \boldsymbol{d}_i^o) = \boldsymbol{0}$ and $Cov_{CS}(\beta_i, \boldsymbol{c}_i^o) = \boldsymbol{0}$. For example, investors may have higher demands for high β_i^o stocks due to lottery demands; in this case, β_i is the market-risk component with this extra demand spun off.

By combining Equations (2.1) and (2.3), we have

$$r_{it} - r_f = \alpha_i + \beta_i M_t + \boldsymbol{d}_i^{oT} \boldsymbol{D}_t + \boldsymbol{c}_i^{oT} \boldsymbol{C}_t + \epsilon_{it} \tag{2.4}$$

where $\boldsymbol{D}_t = \boldsymbol{D}_t^o + \boldsymbol{\xi}_d M_t$ and $\boldsymbol{C}_t = \boldsymbol{C}_t^o + \boldsymbol{\xi}_c M_t$ hold. After the transformation, we have a model (Equation 2.4) reflecting the same information as Equation 2.1, but the time-series correlation between a factor's return and the market return subsumes the cross-sectional correlation between the exposure to this factor and the market beta. In other words, the time-series correlation between $\boldsymbol{D}_t$ (or $\boldsymbol{C}_t$) and M_t fully reflects the interaction between their underlying economic sources. Appendix B1.2 illustrates the benefit of this transformation using an example.

The benchmark ideal asset-pricing model for analysis

For ease of exposition, I further simplify Equation 2.4 by using a single factor D_t to capture $\boldsymbol{D}_t$ and C_t to capture $\boldsymbol{C}_t$:

$$r_{it} - r_f = \alpha_i + \beta_i M_t + d_i D_t + c_i C_t + \epsilon_{it} \tag{2.5}$$

As shown in Appendix B1.3, this simplification is achieved without affecting any analytical conclusions. Following Equation (2.4), Equation (2.5) satisfies the condition $Cov_{CS}(\beta_i, [d_i\ c_i]) = [0\ 0]$, which is not an assumption but is a general setup guided by Section 2.3.1.

$$\begin{aligned} &Cov([M_t\ D_t\ C_t],\ \epsilon_{it}) = [0\ 0\ 0];\ Cov_{CS}([\beta_i\ d_i\ c_i],\ \alpha_i\ + \epsilon_{it}) = [0\ 0\ 0] \\ &Cov_{CS}(\beta_i, [d_i\ c_i]) = [0\ 0] \end{aligned} \tag{2.6}$$

where the first line follows Equation (2.2). The second line follows Equation (2.4), which is not an assumption but is a general setup guided by Section 2.3.1.[5]

The connection between factors is denoted by

$$\begin{aligned} D_t &= \alpha_D + \rho_D M_t + u_{D,t} \\ C_t &= \alpha_C + \rho_C M_t + u_{C,t} \\ D_t &= \alpha_{DC} + \rho_{DC} C_t + u_{DC,t} \end{aligned} \tag{2.7}$$

For brevity, all factors are scaled to have the same variance, $Var(M_t) = Var(D_t) = Var(C_t)$, so that ρ_D, ρ_C and ρ_{DC}, which are exposures of one factor to another factor, are also correlation coefficients. This linear relationship is a general setup. For example, ρ_C is zero if factor C_t is uncorrelated with the market.

Given the generality of Equation (2.5), I use it as the latent ideal asset-pricing model that captures all pricing information. Using this model greatly simplifies the remaining analysis. Hereafter, under Equation (2.5), the correlation between D_t (or C_t) and M_t really reflects both the time-series correlation between factor returns and the cross-sectional correlation between factor exposures. The difference

[5] I do not further combine D_t and C_t because, as discussed in Section 2.4, they will not be captured by econometricians simultaneously under certain scenarios, which is the precondition for analyzing a new LBA. There may also exist other types of additional factors beyond D_t and C_t. However, as the finding of this study will suggest, econometricians will not perceive their existence before uncovering C_t; thus, they are not included in the benchmark model for brevity.

between D_t and C_t will be defined by Definition 1 in Section 2.4. In this section, I focus on analyzing the impact of a single omitted factor, D_t.

2.3.2 The emergence of LBA

Given the ideal asset-pricing model denoted by Equation (2.5), I clarify how pioneer econometricians examining the CAPM discover LBA. The econometricians' model is:

$$r_{it} - r_f = \hat{\alpha}_i + \hat{\beta}_i M_t + \hat{\epsilon}_{it} \tag{2.8}$$

that is, the model econometricians have is a degenerated version of the model that fully captures asset returns. As this model differs from Equation (2.5), the estimated market beta also captures exposures to factors omitted by market risk. In this section, I only consider the impact of D_t for ease of exposition. For example, we can assume $\rho_C \approx 0$ to remove C_t; in this case, the estimated beta is

$$\hat{\beta}_i = \beta_i + \rho_D \times d_i + \rho_C \times c_i \approx \beta_i + \rho_D \times d_i \tag{2.9}$$

The deviation of the estimated beta ($\hat{\beta}_i$) from the true beta (β_i) is systematically affected by omitted factors as econometricians' betas also capture exposures to factors omitted by market risk.[6] The extra factor exposure captured by estimated betas is the origin of a non-zero market beta-alpha relationship. Next, I explain how the extra information ($\rho_D \times d_i$) induces LBA.

[6]Correspondingly, $\hat{\alpha}_i + \hat{\epsilon}_{it} = (\beta_i - \hat{\beta}_i)M_t + (d_i D_t + c_i C_t) + (\alpha_i + \epsilon_{it})$ holds; that is, $\hat{\alpha}_i + \hat{\epsilon}_{it}$ is uncorrelated with the market return in the time series as forced by OLS but is correlated with the market beta in the cross section.

The betting-against-beta (BAM) performance

LBA refers to the phenomenon that low-beta stocks outperform high-beta stocks on a risk-adjusted basis. Statistically, it is identified by a positive alpha of a betting-against-beta (BAM) portfolio's return against the market return. To understand the economic content of LBA, I construct a BAM portfolio by buying low-$\hat{\beta}_i$ stocks and selling high-$\hat{\beta}_i$ stocks at the beginning of each time t,

$$BAM_t = \sum_{\{i:\ \hat{\beta}_i < \beta_L\}} w_i r_{it} - \sum_{\{i:\ \hat{\beta}_i > \beta_H\}} w_i r_{it} \tag{2.10}$$

where β_L and β_H are the thresholds for stocks to be categorized as low- or high-beta stocks; w_i is the weight of security i in the long (low-beta) or short (high-beta) side, satisfying $\sum_{\{i:\ \hat{\beta}_i<\beta_L\}} w_i = \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} w_i = 1$. The time subscripts of betas and weights are dropped for brevity.[7]

Next, I remove the BAM portfolio's market exposure:

$$BAM_t - \hat{\beta}_M M_t = \hat{\alpha}_M + \hat{\epsilon}_{M,t} \tag{2.11}$$

where $\hat{\alpha}_M + \hat{\epsilon}_{M,t}$ is the component of the BAM return orthogonal to the market factor; $\hat{\beta}_M = \sum_{\{i:\ \hat{\beta}_i<\beta_L\}} w_i\hat{\beta}_i - \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} w_i\hat{\beta}_i$ is BAM_t's exposure to M_t (persistent over time; allowing it to be time-varying does not affect the analysis). As

[7]Note that at the beginning of time t, β_L and β_H are determined by the latest estimated betas ($\hat{\beta}_i$s), and w_is are determined by the relative market equity of individual stocks; that is, these variables are time-varying empirically. The time subscripts of these variables are dropped for brevity as betas are persistent over time, especially for a portfolio containing a large number of stocks. Moreover, given the finding of Lewellen and Nagel (2006), it is reasonable not to consider the time variation of betas as they have a trivial impact on asset returns empirically. Explicitly modeling the time variation of these variables does not affect analytical conclusions, and in the empirical analysis, I also report results based on time-varying factor exposures.

proved in Appendix B1.4, the BAM alpha is

$$\hat{\alpha}_M = \Big(\sum_{\{i:\ \hat{\beta}_i < \beta_L\}} w_i d_i - \sum_{\{i:\ \hat{\beta}_i > \beta_H\}} w_i d_i \Big) \alpha_D \tag{2.12}$$

where $(\sum_{\{i:\ \hat{\beta}_i < \beta_L\}} w_i d_i - \sum_{\{i:\ \hat{\beta}_i > \beta_H\}} w_i d_i)$ is the BAM portfolio's true latent exposure to factor D_t. Equation (2.12) shows that the BAM alpha ($\hat{\alpha}_M$) can only be driven by D_t.

The condition for LBA to emerge

Equations (2.12) and (2.9) clarify the condition for LBA to be observed, which can be summarized as

Proposition 1 *LBA is a manifestation of factors omitted by market risk in a multifactor world. Specifically, LBA is observed in the literature because omitted factors overall negatively correlate with market risk.*[8]

As discussed in Section B1.1, when the standard CAPM does not capture all the information investors have, the extra information can be captured by adding more factors to the CAPM. As such, the ideal latent asset-pricing model containing all information can be summarized by a multifactor model (Equation 2.5). As shown in Equations (2.8) to (2.9), factors beyond market risk will be captured by market betas estimated from the CAPM, which is the origin of LBA. Moreover, Equations (2.10) and (2.12) suggest that returns unrelated to return comovement

[8]For brevity, I assume all omitted factors have positive premiums. If all omitted factors have negative risk premiums, a low-beta anomaly (LBA) emerges when the impacts of omitted factors positively correlated with the market factor dominate. As this scenario does not affect any conclusions, this study only analyzes the assumption that all priced factors have positive premiums.

do not enter the betting-against-beta alpha. Therefore, LBA only reflects returns related to omitted factors.

Specifically, LBA is observed because factors omitted by market risk overall negatively correlate with market risk.[9] We can understand why low-beta stocks outperform high-beta stocks from Equation (2.9). Suppose the market factor is negatively correlated with D_t and uncorrelated with C_t, that is, $\rho_D < 0$ and $\rho_C = 0$. In this case, according to $\hat{\beta}_i = \beta_i + \rho_D \times d_i$, $\hat{\beta}_i$ negatively captures d_i as ρ_D is negative. Consequently, sorting stocks into $\hat{\beta}_i$ quantiles tends to pick up stocks with high exposures to D_t in the low-$\hat{\beta}_i$ side and low exposures to D_t in the high-$\hat{\beta}_i$ side; on the other hand, as C_t is uncorrelated with the market, stocks in the high and low-$\hat{\beta}_i$ sides tend to have similar exposures to C_t. Therefore, a BAM portfolio tends to be positively exposed to D_t and not exposed to C_t (that is, $(\sum_{\{i:\ \hat{\beta}_i<\beta_L\}} w_i d_i - \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} w_i d_i) > 0$ and $(\sum_{\{i:\ \hat{\beta}_i<\beta_L\}} w_i c_i - \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} w_i c_i) = 0$). In this case, LBA emerges because a BAM portfolio tends to be positively exposed to D_t

When a large number of stocks are included in the analysis, this tendency becomes a definitive relationship. Figure 2.3 demonstrates how this tendency becomes definitive (simulation details are in Appendix B3.2). The horizontal axis is the "number of stocks" used for constructing BAM portfolios. Corresponding to each horizontal value, the results of 1000 simulations are reported. Consistent with Equation (2.9), a BAM portfolio negatively exposed to the market ($\hat{\beta}_M < 0$) tends to have a negative true market exposure ($\beta_M < 0$) and positive true exposure to

[9]Note that the statement that "an omitted factor negatively correlates with the market factor" refers to the scenario that 1) factor returns are negatively correlated in the time series and 2) factor exposures are negatively correlated in the time series (see Section 2.3.1).

D_t ($\beta_D > 0$). When the BAM portfolio is constructed using a large number of stocks, it is always positively exposed to D_t. This positive exposure translates into a positive alpha against the market return. In the literature, when a study finds the economic source underlying D_t and develops a corresponding empirical proxy that subsumes $\hat{\alpha}_M$, it concludes that LBA is resolved.[10]

Proposition 1 nests how previous studies explain LBA. For example, the lottery-demand-based explanation (Bali et al. 2017) can be interpreted under this framework. The factor-mimicking portfolio (with a positive premium) that fully captures lottery demands has low lottery-demand stocks in the long side and high lottery-demand stocks in the short side. Since high (low) lottery-demand stocks coincide with high (low) market-beta stocks, the lottery-demand factor that accounts for the low-beta anomaly is negatively correlated with the market factor. In other words, LBA exists since there exists a lottery-demand factor unknown to econometricians and negatively correlated with the market factor. Besides nesting existing explanations of LBA, more importantly, the framework regulated by Proposition 1 provides the foundation for uncovering the new LBA in the next section.

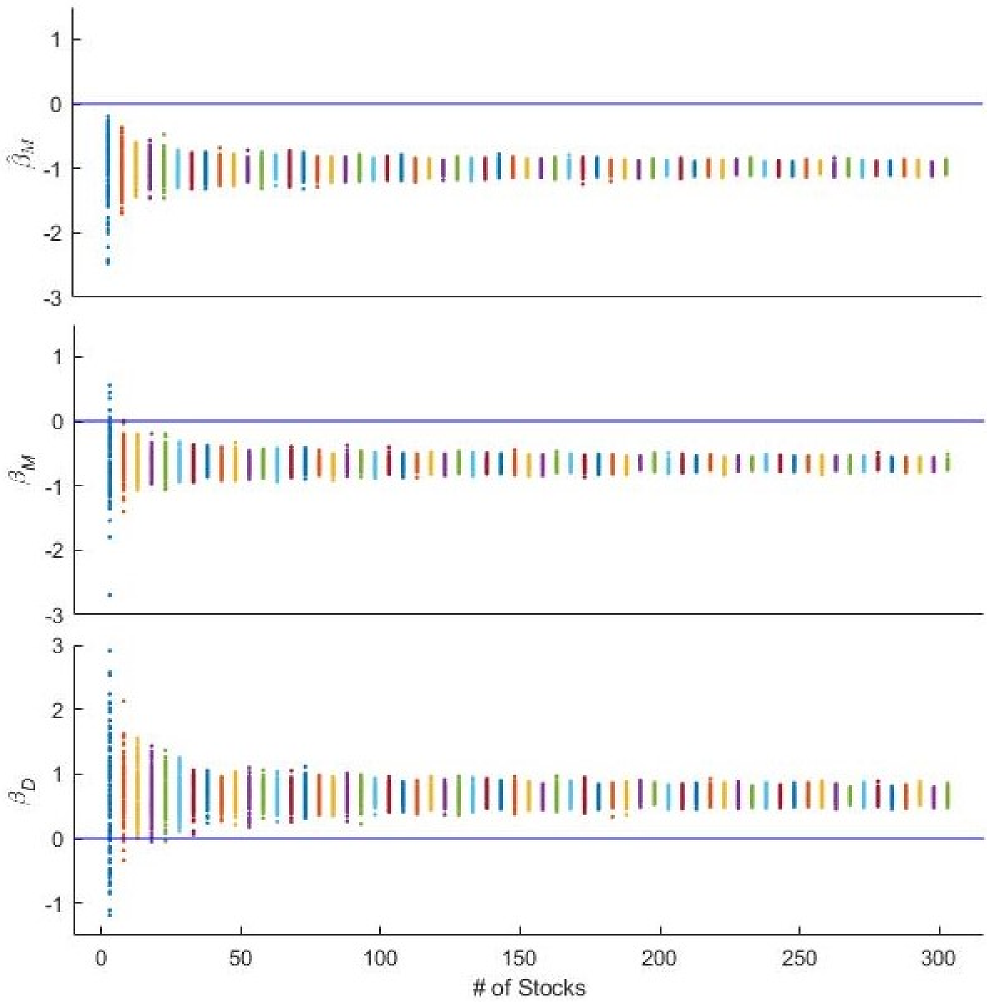
[10]As the BAM portfolio only captures M_t and D_t when C_t is uncorrelated with market risk (see Appendix B1.4), that is,

$$BAM_t = \left(\sum_{\{i:\ \hat{\beta}_i < \beta_L\}} w_i \beta_i - \sum_{\{i:\ \hat{\beta}_i > \beta_H\}} w_i \beta_i \right) M_t + \left(\sum_{\{i:\ \hat{\beta}_i < \beta_L\}} w_i d_i - \sum_{\{i:\ \hat{\beta}_i > \beta_H\}} w_i d_i \right) D_t$$

adjusting for M_t and D_t should fully subsume the average BAM return.

FIGURE 2.3: Number of Stocks and a Betting-Against-Beta (BAM) Portfolio's Factor Exposures

The figure depicts the factor exposures of a BAM portfolio that buy/sell low/high beta stocks through simulation evidence. The true data-generating process is $r_{it} - r_f = \alpha_i + \beta_i M_t + d_i D_t + \epsilon_{it}$. A BAM portfolio is constructed using market betas estimated from the CAPM (i.e., $r_{it} - r_f = \hat{\alpha}_i + \hat{\beta}_i M_t + \hat{\epsilon}_{it}$). $\hat{\beta}_M$ is the BAM portfolio's exposure to the market factor (M_t) in a single-factor spanning regression, which is negative by construction. β_M and β_D are the BAM portfolio's true latent exposures to factors M_t and D_t. The horizontal axis corresponds to the number of stocks used for constructing a BAM portfolio. The vertical line corresponding to each horizontal-axis value is composed of 1000 dots, with each dot denoting the result of one simulation. Appendix B3.2 provides the simulation details.



2.4 A New Low-Beta Anomaly (LBA)

Based on the rationalization of LBA in Section 2.3, this section shows that, theoretically, the new LBA presented in Section 2.2 must differ from the LBA known in the literature. To demonstrate the difference, I classify the factors omitted by market risk (see Equation 2.5) into two types:

Definition 1 *The two factors omitted by market risk (D_t and C_t; see Equation 2.5) differ in their connections with market risk (M_t):*
1) D_t summarizes factors directly correlated with market risk, $|Corr(M_t, D_t)| > 0$;
2) C_t summarizes factors only partially correlated with market risk, $|Corr(M_t, C_t)| \approx 0$ and $|Corr(M_t, C_t|D_t)| \neq 0$.

C_t does not (or weakly) correlate with market risk (M_t); however, once the impact of D_t is removed, C_t becomes correlated with market risk. Given Definition 1, the difference between the known and new LBAs lies in that they are manifestations of different factors, as summarized by the following proposition:

Proposition 2 *The low-beta anomaly (LBA) known in the literature is a manifestation of factors directly correlated with market risk (D_t), while a new LBA, if it exists, must be a manifestation of partial-correlation factors (C_t).*

Hereafter, this section demonstrates in turn 1) why previous studies do not observe the LBA as a manifestation of C_t, 2) how to uncover the new LBA as a manifestation of C_t, and 3) why C_t tends to induce a new LBA.

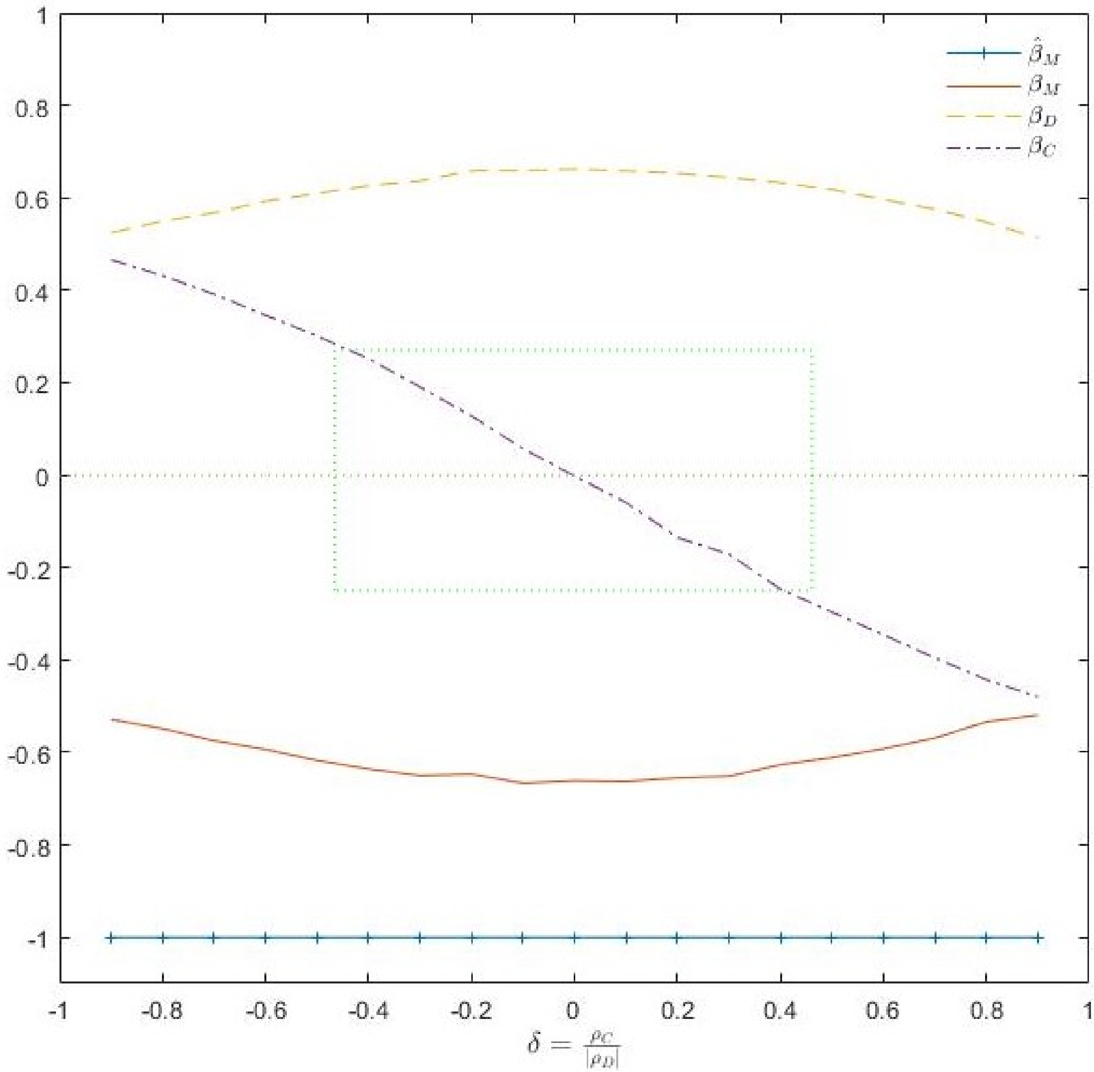
2.4.1 The known LBA does not reflect partial-correlation factors

First, I demonstrate the necessity of analyzing the new LBA. As a manifestation of factors omitted by market risk (Proposition 1), LBA provides an effective approach to identifying important omitted factors. However, the LBA known in the literature cannot reflect partial-correlation factors: Before the known low-beta anomaly (LBA) is resolved, LBA does not reflect partial-correlation factors (C_t), regardless of how important they are for asset pricing. Before the resolution of the known LBA, estimated betas only capture factors directly correlated with market risk (D_t) besides capturing market risk; hence, partial-correlation factors (C_t) cannot be noticed. The reason is that to what extent exposure to an omitted factor is captured by the estimated beta entirely depends on the correlation between that omitted factor and the market factor (see Equation 2.9).

Figure 2.4 demonstrates the limitation of estimated betas through the true factor exposures of a BAM portfolio (constructed following Equation 2.10; betas are estimated from the CAPM). β_M, β_D and β_C are the BAM portfolio's true latent exposures to M_t, D_t and C_t, respectively. The correlation between D_t and M_t is set as $\rho_D = -0.5$ to reflect the known LBA. The correlation between C_t and M_t is set as $\rho_C = \delta|\rho_D|$. As shown by the dash-dotted line (β_C), only when C_t's connection with M_t is as strong as D_t's, that is, $|\delta|$ approaching one, will the BAM portfolio notably capture C_t. In contrast, the BAM portfolio barely contains information about C_t when $|\delta|$ is small (the rectangular area).

Therefore, regardless of how important partial-correlation factors (C_t) are,

Figure 2.4: Limitation of Pre-Resolution Econometricians' Betas
The figure depicts the limitation of pre-resolution empiricists' betas (β_i) through the factors exposures of the betting-against-beta (BAM) portfolio. $\hat{\beta}_M$ is the BAM portfolio's exposure to market risk in a single-factor spanning regression. β_M, β_D and β_C are the BAM portfolio's true latent factor exposures. $\delta = \frac{\rho_C}{|\rho_D|}$ is the strength of C_t's connection with the market relative to that of D_t's. ρ_D is set to be -0.5 to reflect the LBA known in the literature. Appendix B3.1 provides the details.



econometricians can only perceive the existence of factors directly correlated with market risk (D_t) before the resolution of the known LBA, that is,

$$\hat{\beta}_i \approx \beta_i + \rho_D \times d_i \;\Rightarrow\; BAM_t - \hat{\beta}_M M_t \approx \Big(\sum_{\{i:\ \hat{\beta}_i < \beta_L\}} w_i d_i - \sum_{\{i:\ \hat{\beta}_i > \beta_H\}} w_i d_i \Big)(\alpha_D + u_{D,t}) \tag{2.13}$$

where $\alpha_D + u_{D,t}$ is the component of D_t orthogonal to the market factor (see Equation 2.7).

2.4.2 Uncovering the new LBA driven by partial-correlation factors

Next, I demonstrate how to uncover the new LBA as a manifestation of partial-correlation factors (C_t). Should a new LBA exist, it can only be uncovered conditional on a sufficient explanation of the known LBA. Moreover, the new LBA must be a manifestation of partial-correlation factors (C_t). To uncover the impact of C_t, we need to allow estimated betas, besides capturing market risk, to capture exposures to C_t. A resolution of the known LBA (as a manifestation of D_t) enables estimated betas to capture new information. To demonstrate the rationale, I consider the information perceived by econometricians if they investigate the market beta-return relationship conditional on the resolution of known LBA. Their model is,

$$r_{it} - r_f = \hat{\alpha}_i^* + \hat{\beta}_i^* M_t + \hat{d}_i^* D_t + \hat{\epsilon}_{it}^* \tag{2.14}$$

where D_t is included as several recent studies have fully resolved the LBA known in the literature; $\hat{\beta}_i^*$ and $\hat{d}_i^*$ are the market beta and factor exposure obtained by post-resolution econometricians. Conditional on a resolution of the known LBA, betas obtained by econometricians become

$$\hat{\beta}_i^* = \beta_i + \frac{Var(M_t)}{Var(M_t^\epsilon)}(\rho_C - \rho_D\rho_{DC})c_i \tag{2.15}$$

where M_t^ϵ is the component of M_t orthogonal to D_t. β_i and c_i are a stock's true exposure to M_t and C_t (Equation 2.5). ρ_C, ρ_D and ρ_{DC} are the correlations between factors (Equation 2.7). This equation indicates that how $\hat{\beta}_i^*$ captures the exposure to C_t is determined by the connections among M_t, D_t and C_t (see Equations B.16 and B.17 in Section B1.4 for the derivation).

Equations (2.14) and (2.15) show that when the impact of D_t is removed, estimated betas only capture information about C_t besides capturing market risk. Under the same rationale for $\hat{\beta}_i$ (Equation 2.9) to translate into the LBA driven by D_t, $\hat{\beta}_i^*$ can translate into an LBA driven by C_t; that is, as long as $\rho_C - \rho_D\rho_{DC}$ is negative, low (high) estimated market betas tend to be associated with high (low) exposures to C_t, translating into a negative beta-alpha relationship. In other words, a factor need not relate to market risk to induce LBA. For example, if C_t is uncorrelated with M_t but is negatively correlated with D_t ($\rho_C = 0$ and $\rho_{DC} < 0$), C_t can still induce an LBA when D_t is negatively correlated with the market factor ($\rho_D < 0$).[11]

[11]The essence for a new LBA to exist is clarified by rewriting the post-resolution econometricians' model (Equation 2.14) in following format:

$$r_{it} - r_f = \hat{\alpha}_i + \hat{\beta}_i^* M_t^\epsilon + \hat{\epsilon}_{it}$$

Following the same derivation as Section 2.3.1, it can be shown that the estimated market beta from this model is the same as that of Equation (2.15) without making any additional assumptions. $\frac{Var(M_t)}{Var(M_t^\epsilon)}(\rho_C - \rho_D\rho_{DC})$ in Equation (2.15) is the loading of C_t on M_t^ϵ. A new LBA exists as long as $Cov(M_t^\epsilon, C_t)$ is negative.

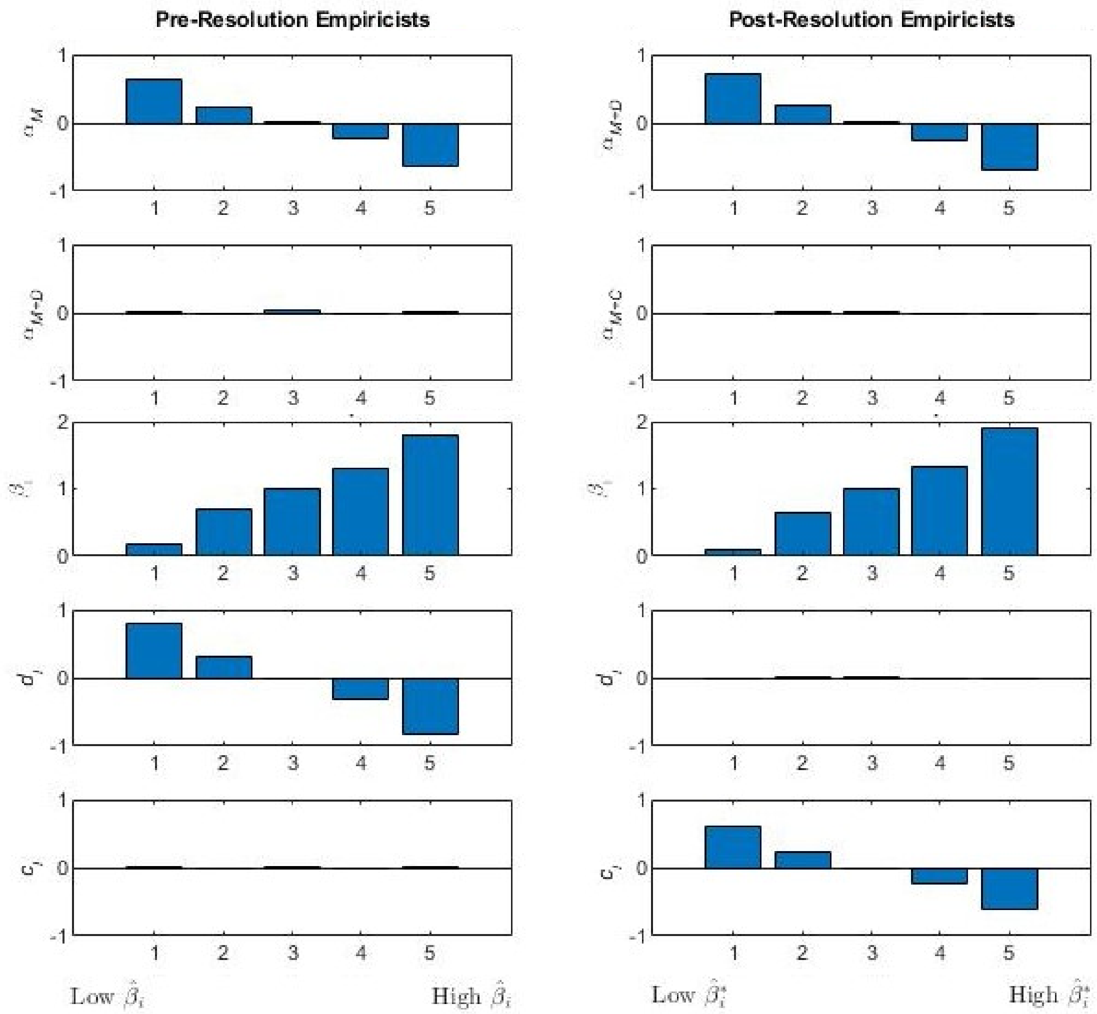
To clarify the economic content of the new LBA, we can construct a BAM portfolio (following Equation 2.10) using betas obtained by post-resolution econometricians and remove the BAM portfolio's factor exposures (see Appendix B1.4 for detailed derivations):

$$BAM_t^* - \hat{\beta}_M^* M_t - \hat{\beta}_D^* D_t = \Big(\sum_{\{i:\ \hat{\beta}_i^* < \beta_L\}} w_i c_i - \sum_{\{i:\ \hat{\beta}_i^* > \beta_H\}} w_i c_i \Big)(\alpha_C^* + u_{C,t}^*) \quad (2.16)$$

where $\alpha_C^* + u_{C,t}^*$ is the component of C_t orthogonal to M_t and D_t. If the expectation of $\alpha_C^* + u_{C,t}^*$ is significantly positive, a new LBA is identified. Therefore, once the impact of factors directly correlated with market risk is removed, we can uncover a new LBA. The new LBA is driven by factors different from the drivers of the known LBA.

Figure 2.5 provides a numerical example of an economy satisfying Proposition 2. Unlike Figure 2.1, we know the factors underlying the new LBA and the true latent factor exposures in this environment. When sorting stocks into quantiles using their respective estimated betas, both pre- and post-resolution econometricians observe an LBA, as shown by the negative beta-alpha relationship in the top subfigures. The second subfigure in the left panel shows that D_t fully subsumes the LBA observed by pre-resolution econometricians. Accordingly, this LBA is a manifestation of D_t. On the other hand, as shown by the top two subfigures of the right panel, the LBA observed by post-resolution econometricians is not explained by D_t but is fully subsumed by C_t. Therefore, the two LBAs have different economic content as D_t and C_t capture different return sources.

FIGURE 2.5: Economic Content of the Known and New Low-Beta Anomalies (LBAs)
The figure depicts the changeable economic content of LBA in an economy with factors *directly correlated with market risk* (D_t) and factors *not directly but partially correlated with market risk* (C_t). In the left and right panels, stocks are sorted into quintiles according to betas from Equations (2.8) and (2.14), respectively. α_M, α_{M+D} and α_{M+C} are the alphas against M_t, $[M_t,\ D_t]$ and $[M_t,\ C_t]$, respectively. β_i, c_i and d_i are each quintile portfolio's true latent factor exposures. Appendix B3.1 provides the details.



2.4.3 Partial-correlation factors tend to induce a new LBA

A remaining question is why a new non-zero beta-alpha relationship should be negative (i.e., LBA) rather than positive. The following corollary sheds light on this question. A new non-zero beta-alpha relationship, should it exist, tends to be a negative relationship and hence a low-beta anomaly (LBA).

To demonstrate Corollary 2.4.3, I decompose the new BAM alpha of Equation (2.16) by combining it with Equations (2.14) and (2.15) to get:

$$\begin{aligned}\hat{\alpha}^*_M = &\Big(\sum_{\{i:\ \hat{\beta}^*_i<\beta_L\}} w_i c_i - \sum_{\{i:\ \hat{\beta}^*_i>\beta_H\}} w_i c_i\Big) E(C_t)\\ &+\Big[\Big(\sum_{\{i:\ \hat{\beta}^*_i<\beta_L\}} w_i \beta_i - \sum_{\{i:\ \hat{\beta}^*_i>\beta_H\}} w_i \beta_i\Big) - \Big(\sum_{\{i:\ \hat{\beta}^*_i<\beta_L\}} w_i \hat{\beta}^*_i - \sum_{\{i:\ \hat{\beta}^*_i>\beta_H\}} w_i \hat{\beta}^*_i\Big)\Big] E(M_t)\\ &+\Big[\Big(\sum_{\{i:\ \hat{\beta}^*_i<\beta_L\}} w_i d_i - \sum_{\{i:\ \hat{\beta}^*_i>\beta_H\}} w_i d_i\Big) - \Big(\sum_{\{i:\ \hat{\beta}^*_i<\beta_L\}} w_i \hat{d}^*_i - \sum_{\{i:\ \hat{\beta}^*_i>\beta_H\}} w_i \hat{d}^*_i\Big)\Big] E(D_t)\end{aligned} \tag{2.17}$$

where $E(M_t)$, $E(D_t)$ and $E(C_t)$ are the expected returns of the three factors, which are all positive; $\beta_M = \sum_{\{i:\ \beta^*_i<\beta_L\}} w_i\beta_i - \sum_{\{i:\ \beta^*_i>\beta_H\}} w_i\beta_i$, $\beta_D = \sum_{\{i:\ \beta^*_i<\beta_L\}} w_i d_i - \sum_{\{i:\ \beta^*_i>\beta_H\}} w_i d_i$ and $\beta_C = \sum_{\{i:\ \beta^*_i<\beta_L\}} w_i c_i - \sum_{\{i:\ \beta^*_i>\beta_H\}} w_i c_i$ are the BAM portfolio's true latent exposures to M_t, D_t and C_t; $\hat{d_i}^*$ is asset i's exposure to factor D_t estimated by econometricians using Equation (2.14); $\hat{\beta}^*_M = \sum_{\{i:\ \beta^*_i<\beta_L\}} w_i\hat{\beta}^*_i - \sum_{\{i:\ \beta^*_i>\beta_H\}} w_i\hat{\beta}^*_i$ and $\hat{\beta}^*_D = \sum_{\{i:\ \beta^*_i<\beta_L\}} w_i\hat{d}^*_i - \sum_{\{i:\ \beta^*_i>\beta_H\}} w_i\hat{d}^*_i$ are the BAM portfolio's estimated factor exposures when regressing it on M_t and D_t. Note that the fact that $\hat{\alpha}^*_M$ can be decomposed into three components as in Equation (2.17) does not mean that the new LBA captures returns related to M_t and D_t. As shown in Equation (2.16), the combination of the three components is the component of C_t

orthogonal to M_t and D_t.

We can understand Corollary 2.4.3 by analyzing the signs of the three elements on the right-hand side of Equation (2.17). The sign of the first element depends on how C_t partially correlates with M_t, that is, $Cov(C_t, M_t|D_t)$. The coefficient before the third element is zero as, according to Equation (2.15), betas estimated by post-resolution econometricians do not capture exposures to D_t. The second element is always positive as a BAM portfolio's estimated absolute market exposure ($|\hat{\beta}^*_M|$) is larger than its true absolute market exposure ($|\beta_M|$) for two reasons. First, $\hat{\beta}^*_M$ and β_M are negative since a BAM portfolio has low/high-beta stocks in the long/short side. Second, $h(\rho_C - \rho_D\rho_{DC})$ and β_C always have opposite signs, as discussed following Equation (2.15). Therefore, $\beta_M - \hat{\beta}^*_M = -\frac{Var(M_t)}{Var(M_t^\epsilon)}(\rho_C - \rho_D\rho_{DC})\beta_C$ is always positive.

When C_t negatively contributes to $\hat{\alpha}^*_M$, its impact is weakened by the second component. On the other hand, the impact when C_t positively contributes to $\hat{\alpha}^*_M$ is enhanced by the second component, which in turn translates into a strong LBA. Appendix B3.2 provides simulation examples to demonstrate Corollary 2.4.3. Note that this discussion only suggests that the probability of observing a new negative rather than positive beta-alpha relationship is higher. Essentially, whether a new LBA exists still depends on the partial correlation between C_t and M_t. Once a strong new LBA is observed, we can learn about partial-correlation factors before future studies uncover their economic fundamentals.

2.5 Empirical Analysis

This section confirms that the low-beta anomaly (LBA) is a manifestation of factors directly correlated with market risk and, more importantly, uncovers a new LBA as a manifestation of factors partially correlated with market risk (partial-correlation factors). I also provide evidence to show that the impact of partial-correlation factors was not noticed before.

2.5.1 Data and the betting-against-beta (BAM) portfolio

I collect individual stock return and market capitalization data from CRSP. Only common stocks listed on NYSE, AMEX and NASDAQ with share codes 10 or 11 are included (ADRs, ETFs, and REITs are excluded). Raw returns are adjusted for delisting following Shumway (1997). Factor return data are obtained from Kenneth French's website. The sample period is July 1968 to December 2019. Data before July 1968 are used for the initial beta estimation.

The empirical examinations are based on both the simple and market-neutral betting-against-beta portfolios (BAM and BAB, respectively). To construct a BAM portfolio, I double-sort stocks into two-by-three size-beta quantiles independently according to NYSE breakpoints (50th percentile for size, 30th and 70th percentiles for beta) at the end of each month. The BAM portfolio return is calculated as:

$$BAM_t = r_t^L - r_t^H \tag{2.18}$$

where $r_t^L = \frac{1}{2}(r_t^{Small,\ Low\ \beta} + r_t^{Big,\ Low\ \beta})$ holds and $r_t^{Small,\ Low\ \beta}$ $(r_t^{Big,\ Low\ \beta})$ is the value-weighted return of stocks in the small(big)-size and low-beta portfolio;

$r_t^H = \frac{1}{2}(r_t^{Small,\ High\ \beta} + r_t^{Big,\ High\ \beta})$ holds and $r_t^{Small,\ High\ \beta}$ $(r_t^{Big,\ High\ \beta})$ is the value-weighted return of stocks in the small(big)-size and high-beta portfolio. The portfolio is size-balanced to avoid small/microcap biases. Size is the market capitalization at the end of the previous June.[12]

A BAB portfolio is constructed by leveraging up/down the low/high-beta side of the simple BAM portfolio so that its ex-ante beta is zero (Frazzini and Pedersen 2014). Specifically, I use a value-weighting scheme following Novy-Marx and Velikov (2022) to avoid small/microcap bias:

$$BAB_t = \frac{1}{\beta_{t-1}^L}(r_t^L - r_f) - \frac{1}{\beta_{t-1}^H}(r_t^H - r_f) \tag{2.19}$$

where β_{t-1}^L and β_{t-1}^H are the ex-ante betas (value-weighted) of the low and high-beta sides, respectively. Essentially, a BAB portfolio explicitly allows betas to be time-varying.[13]

The BAM (or BAB) portfolio is constructed using $\hat{\beta}_i$ (Equation 2.8) if it aims to capture the known LBA and is constructed using $\hat{\beta}_i^*$ (Equation 2.14) if it aims to capture the new LBA. For the main result, I estimate betas of individual stocks using daily returns of the prior year at each time t, with a minimum of 100 observations required for an estimated beta to be valid. Appendix B2.3 also reports the results based on different beta estimates.

[12]It is well known that empirical asset pricing models do not perform well in explaining small stocks (e.g., Fama and French 1996, 2016). Constructing the BAM portfolio in a size-balanced way enables a fair comparison between the empirical proxies for LBA and common risk factors.

[13]Details of the BAB portfolio are provided in Appendix B2.4. I expect that results based on BAM and BAB lead to the same conclusion. As shown in Novy-Marx and Velikov (2022), a BAB portfolio constructed under the value-weighting scheme is essentially a BAM portfolio hedged with a market index. As such, they capture the same information.

2.5.2 The known and new low-beta anomalies (LBAs)

Recent studies find that multi-factor models with profitability and investment factors subsume the abnormal performance of LBA (e.g., Fama and French 2015; Hou et al. 2015; Barroso et al. 2020; Novy-Marx and Velikov 2022). On the other hand, proxies for the existing explanations of LBA are found to contain overlapping information with investment and profitability factors (e.g., Fama and French 2016; Hou et al. 2020b). Therefore, the six-factor model (Fama and French 2018) can be regarded as a comprehensive measure of existing explanations of LBA. This subsection interprets the way the six-factor model resolves the known LBA following Section 2.3 and then uncovers a new LBA based on this model following Section 2.4.

Settlement of the known low-beta anomaly (LBA)

Table 2.1 depicts the known LBA (BAM_t or BAB_t) and its settlement. BAM_t is constructed following Equation (3.1) using $\hat{\beta}_i$s estimated from the CAPM (Sharpe 1964). When reporting the spanning regression results, I scale each BAM portfolio by a constant so that their exposures to the market are minus one. In this case, a BAM alpha reflects the abnormal performance corresponding to negative one unit exposure to the market. Measuring the BAM alpha in a Treynor (1966)-ratio manner enables the comparison of the alphas from different spanning regressions. This scaling operation does not affect the statistical significance of a BAM alpha or the Sharpe ratio of an LBA.

As reported by the first row, the BAM portfolio generates a significant alpha (0.65% per month, at the 1% level) against the CAPM, referred to as LBA in previous studies. This BAM alpha attenuates when the CAPM model is updated with the value and momentum factors but remains significant (0.57% and 0.44% per month), suggesting that these two factors help explain the known LBA but do not account for its major economic content. Consistent with the left panel of Figure 2.1, the alpha of the known LBA is subsumed once the investment (CMA) and profitability (RMW) factors are controlled. The alpha declines to 0.09% per month and becomes insignificant, as reported in the last row of Panel A. Panel B presents the time-series asset-pricing tests of the BAB portfolio. For ease of comparison, I scale the BAB portfolio so that it has the same variance as the market portfolio. The BAB portfolio performance leads to the same conclusion as BAM.

The result confirms the finding in recent studies that investment and profitability factors subsume LBA. Consistent with Proposition 1, Panel C confirms that the two factors negatively correlate with the market factor. From a pure factor-model perspective, the result indicates that the LBA known in the literature is mainly a manifestation of investment and profitability factors (consistent with the finding of previous studies such as Fama and French 2015, Hou et al. 2015, Novy-Marx and Velikov 2022).

Uncovering a new low-beta anomaly (LBA)

As the six-factor model fully explains the known LBA, it is a good proxy for factors directly correlated with market risk. According to Corollary 2.4.2, controlling for the six-factor model should remove the impact of factors directly correlated with market risk. As a result, betas estimated from the six-factor model should capture piratical-correlation factors besides capturing market risk. Table 2.2 reports the

TABLE 2.1: The Low-Beta Anomaly (LBA) Known in the Literature

This table presents the spanning regression alphas (α), factor exposures and adjusted R-squares ($\bar{R}^2$) of the LBA known in the literature (captured by BAM_t or BAB_t, constructed using market betas estimated from the CAPM) against the CAPM, three, four, five and six-factor models (Sharpe 1964 and Lintner 1965; Carhart 1997; Fama and French 1993, 2015, 2018). BAM_t refers to the betting-against-beta portfolio constructed following Equation (3.1). BAB_t refers to the market-neutral betting-against-beta portfolios constructed following Equation (2.19) (based on Frazzini and Pedersen 2014 and Novy-Marx and Velikov 2022). MKT, SMB, HML, RMW, CMA and MOM are the market, size, value, profitability, investment and momentum factors from Kenneth French's website. The t-statistics are based on Newey and West (1987) adjusted standard errors. Panel C reports the correlation coefficients between the market factor and other factors. The sample period is from July 1968 to December 2019.

	α(%)	MKT	SMB	HML	RMW	CMA	MOM	$\bar{R}^2$(%)
Panel A: BAM_t on Factor Models								
	0.65	-1.00						55.17
	(3.81)	(-16.35)						
BAM_t	0.57	-1.00	-0.52	0.60				64.79
	(2.95)	(-12.72)	(-5.66)	(4.51)				
	0.44	-1.00	-0.54	0.69			0.18	65.73
	(2.00)	(-13.10)	(-4.98)	(5.29)			(1.71)	
	0.18	-1.00	-0.41	0.30	0.85	0.82		69.92
	(0.87)	(-12.50)	(-4.64)	(1.64)	(6.91)	(3.73)		
	0.09	-1.00	-0.42	0.39	0.84	0.77	0.14	70.32
	(0.39)	(-12.58)	(-4.55)	(2.25)	(6.76)	(3.63)	(1.25)	
Panel B: BAB_t on Factor Models								
	0.73	0.00						-0.16
	(3.64)	(0.03)						
BAB_t	0.45	0.13	-0.06	0.71				19.34
	(2.45)	(1.82)	(-0.54)	(5.24)				
	0.35	0.16	-0.05	0.75			0.12	20.37
	(1.77)	(2.29)	(-0.46)	(5.80)			(1.19)	
	0.05	0.25	0.11	0.41	0.76	0.63		32.51
	(0.29)	(3.98)	(1.54)	(2.78)	(6.92)	(3.60)		
	0.01	0.26	0.11	0.45	0.74	0.60	0.06	32.72
	(0.04)	(4.20)	(1.50)	(3.20)	(6.56)	(3.51)	(0.72)	
Panel C: Correlation Coefficients Between the Market Factor and Other Factors								
Daily returns			-0.08	-0.15	-0.19	-0.36	-0.11	
Monthly returns			0.28	-0.25	-0.23	-0.38	-0.14	

performance of the BAM (or BAB) portfolio based on market betas estimated from the six-factor model ($\hat{\beta}_i^*$). As shown by the first row of Panel A, the BAM_t^* portfolio has a considerable and significant alpha against the CAPM (0.75%, monthly). The alphas remain significant without attenuation (0.77%, monthly) when investment and profitability factors are introduced, as shown in the last row of Panel A. Moreover, the new LBA achieves a stronger Sharpe ratio (as indicated by the t-statistic in the parenthesis under BAM_t^*'s alpha), making its abnormal performance unlikely a technical improvement relative to the known LBA in Table 2.1. The BAB_t^* performance reported in Panel B is consistent with the BAM_t^* performance in Panel A.

The result that BAM_t^* (or BAB_t^*) generates a considerable and significant alpha against the factors that fully subsume BAM_t (or BAB_t) suggests that the new LBA derives its abnormal performance from return sources beyond the investment and profitability effects. Specifically, as proved in Section 2.4, the new LBA captures partial-correlation factors. Therefore, Proposition 2 that LBA's is not limited to what has been documented in the literature is confirmed. In an unreported test, I confirm that other prevailing asset-pricing models do not explain the new LBA, either (e.g., Hou et al. 2015; Stambaugh and Yuan 2017; Barillas and Shanken 2018; Hou et al. 2020a).

The new LBA cannot be uncovered before resolving the known LBA

This subsection reports the performance of betting-against-beta portfolios using market betas estimated from the three- or four-factor model (Fama and French

Table 2.2: The New Low-Beta Anomaly (LBA)

This table presents the spanning regression alphas (α), factor exposures and adjusted R-squares ($\bar{R}^2$) of the new LBA (captured by BAM_t^* or BAB_t^*, constructed using market betas estimated from the six-factor model) against the CAPM, three, four, five and six-factor models (Sharpe 1964 and Lintner 1965; Carhart 1997; Fama and French 1993, 2015, 2018). BAM_t^* refers to the betting-against-beta portfolio constructed following Equation (3.1). BAB_t^* refers to the market-neutral betting-against-beta portfolios constructed following Equation (2.19) (based on Frazzini and Pedersen 2014 and Novy-Marx and Velikov 2022). MKT, SMB, HML, RMW, CMA and MOM are the market, size, value, profitability, investment and momentum factors from Kenneth French's website. The *t*-statistics are based on Newey and West (1987) adjusted standard errors. The sample period is from July 1968 to December 2019.

	α(%)	MKT	SMB	HML	RMW	CMA	MOM	$\bar{R}^2$(%)
Panel A: BAM_t^* on Factor Models								
	0.75	-1.00						51.12
	(3.89)	(-18.61)						
	0.84	-1.00	-0.62	0.07				57.62
	(4.21)	(-16.85)	(-6.55)	(0.74)				
BAM_t^*	0.80	-1.00	-0.63	0.09			0.06	57.67
	(4.11)	(-16.48)	(-6.30)	(0.97)			(0.85)	
	0.80	-1.00	-0.55	0.19	0.27	-0.28		58.52
	(4.11)	(-18.06)	(-6.19)	(1.56)	(2.28)	(-1.47)		
	0.77	-1.00	-0.55	0.23	0.26	-0.31	0.06	58.59
	(3.94)	(-17.41)	(-6.07)	(1.85)	(2.16)	(-1.65)	(0.98)	
Panel B: BAB_t^* on Factor Models								
	0.83	0.07						0.30
	(4.24)	(1.18)						
	0.81	0.10	-0.14	0.07				0.94
	(4.22)	(1.78)	(-1.56)	(0.72)				
BAB_t^*	0.76	0.11	-0.13	0.09			0.06	1.09
	(4.10)	(1.99)	(-1.48)	(0.96)			(0.96)	
	0.82	0.08	-0.09	0.20	0.18	-0.31		2.39
	(4.32)	(1.56)	(-1.01)	(1.79)	(1.44)	(-1.72)		
	0.77	0.09	-0.09	0.24	0.16	-0.35	0.07	2.66
	(4.15)	(1.72)	(0.99)	(2.10)	(1.33)	(-1.92)	(1.16)	

1993; Carhart 1997). As shown in Panel A of Table 2.3, the monthly spanning regression alphas of BAM_t^{FF3} and BAM_t^{FF4} against the CAPM are respectively 0.57% and 0.71%, indicating the existence of LBA. When controlling for the six-factor model, the alphas attenuate greatly (0.18% and 0.42%, not significant at the 1% level), although they are stronger than those reported in Table 2.1. This pattern arises because the value (HML) and momentum (MOM) factors do not fully capture the drivers of the known LBA; hence, controlling for them does not fully drive out the impacts of investment and profitability factors. As a result, BAM_t^{FF3} and BAM_t^{FF4} still largely capture the known LBA, which is why the two BAM portfolios are still explained by the six-factor model. The results in Panel B lead to the same conclusion.

To sum up, the betting-against-beta performance based on betas estimated from the CAPM, three- or four-factor model confirms that 1) before the resolution of the known LBA, estimated betas only capture factors directly correlated with market risk (Corollary 2.4.1). The results based on betas estimated from the six-factor model (Table 2.2) confirm that conditional on a full resolution of the known LBA, we can uncover a new LBA if it exists (Corollary 2.4.2). Together, this section confirms Proposition 2 that the LBA known in the literature is a manifestation of factors directly correlated with market risk, while a new LBA as a manifestation of partial-correlation factors exists.

TABLE 2.3: The New Low-Beta Anomaly (LBA) Cannot Be Identified before Resolving the Known LBA

This table presents the spanning regression alphas (α), factor exposures and adjusted R-squares ($\bar{R}^2$) of betting-against-beta portfolios against the CAPM and the six-factor model (Fama and French 2018). BAM refers to the betting-against-beta portfolio constructed following Equation (3.1). BAB refers to the market-neutral betting-against-beta portfolios constructed following Equation (2.19) (based on Frazzini and Pedersen 2014 and Novy-Marx and Velikov 2022). BAM_t^{FF3} and BAB_t^{FF3} are constructed using market betas estimated from the three-factor model (Fama and French 1993). BAM_t^{FF4} and BAB_t^{FF4} are constructed using market betas estimated from the four-factor model (Carhart 1997). MKT, SMB, HML, RMW, CMA and MOM are the market, size, value, profitability, investment and momentum factors from Kenneth French's website. The t-statistics are based on Newey and West (1987) adjusted standard errors. The sample period is from July 1968 to December 2019.

	α(%)	MKT	SMB	HML	RMW	CMA	MOM	$\bar{R}^2$(%)
Panel A: BAM_t^{FF3} and BAM_t^{FF4} on Factor Models								
BAM_t^{FF3}	0.57	-1.00						54.36
	(3.22)	(-16.31)						
	0.18	-1.00	-0.56	0.18	0.93	0.41	0.07	69.20
	(0.87)	(-13.76)	(-7.36)	(1.39)	(7.30)	(2.06)	(0.63)	
BAM_t^{FF4}	0.71	-1.00						51.69
	(3.75)	(-17.17)						
	0.42	-1.00	-0.56	0.26	0.89	0.27	0.03	65.65
	(1.87)	(-15.77)	(-5.90)	(2.11)	(6.31)	(1.49)	(0.42)	
Panel B: BAB_t^{FF3} and BAB_t^{FF4} on Factor Models								
BAB_t^{FF3}	0.63	-0.26						6.80
	(3.77)	(-7.04)						
	0.17	-0.04	-0.17	0.15	0.72	0.49	0.02	27.30
	(1.10)	(-1.01)	(-3.19)	(1.97)	(9.66)	(4.43)	(0.46)	
BAB_t^{FF4}	0.75	-0.11						0.99
	(4.35)	(-2.77)						
	0.33	0.08	-0.10	0.21	0.67	0.33	0.02	17.30
	(1.99)	(2.02)	(-1.81)	(2.60)	(8.45)	(2.80)	(0.52)	

2.5.3 The impact of partial-correlation factors was overlooked

While the CAPM is developed to summarize asset returns, previous studies document ample evidence of violations of this benchmark model, which is largely how factors or anomalies beyond market risk were documented in the literature. As partial-correlation factors do not violate the positive cross-sectional beta-return relationship prescribed by the CAPM, their impact should be hardly noticed. This subsection examines if the impact of partial-correlation factors was noticed in previous studies.

Was the new LBA included in the factor zoo?

It is possible that the driving force of the new LBA (i.e., partial-correlation factors) has already been documented in previous studies but has not been linked to LBA. To investigate this possibility, I update the six-factor model (Fama and French 2018) with each of the factors/anomalies available from Hou et al. (2020b) and examine if the abnormal performance of the new LBA is subsumed. The original anomaly data from Lu Zhang's website are three-by-five double-sorted portfolios (size and the corresponding characteristic), and the return of each portfolio is measured on a value-weighted basis. I calculate each anomaly return as the return spread between the high- and low-characteristic sides, where the return of each side is calculated as the average of the three size-sorted portfolios. There are 187 anomalies in total, which are categorized into six groups: momentum (MOM), value-versus-growth (VG), investment (INV), profitability (PROF), intangibles (ITAN) and trading frictions (FRIC).

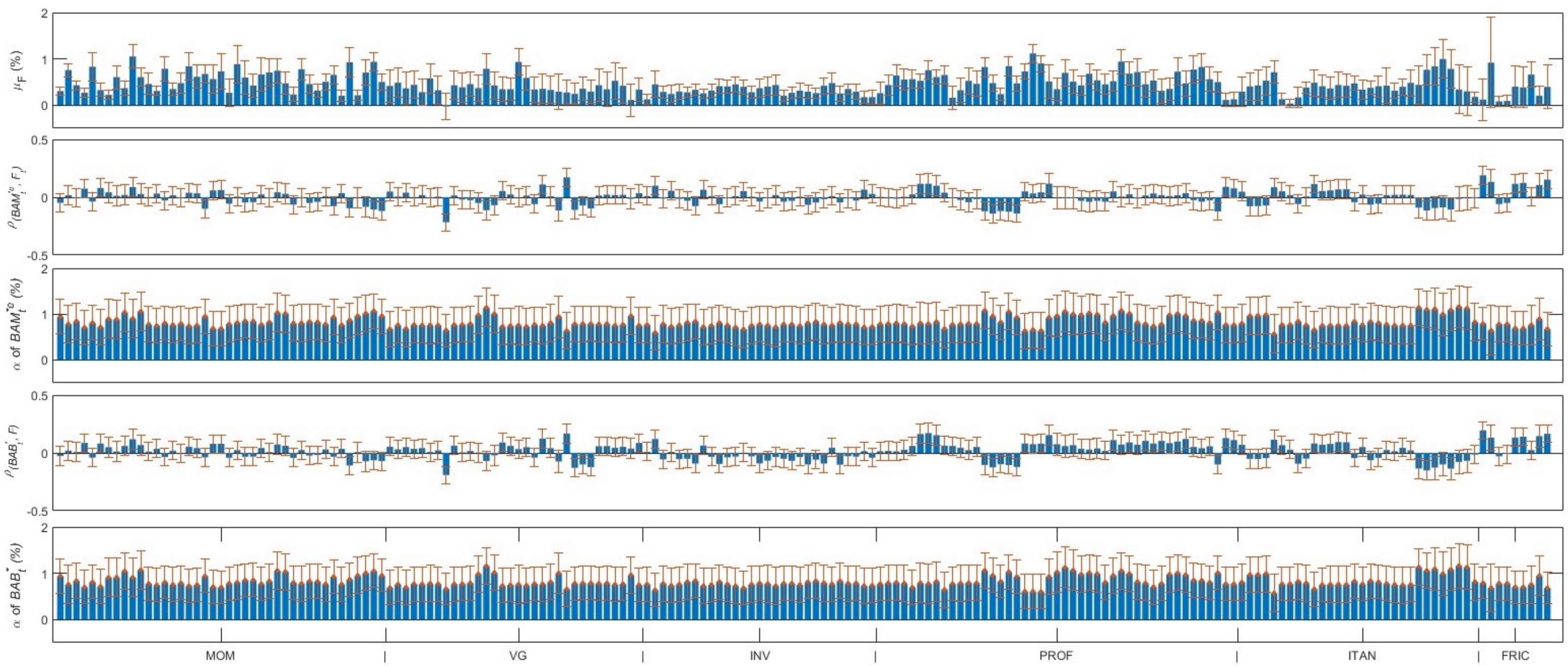
The first subfigure of Figure 2.6 reports the average monthly returns of these factor and anomaly portfolios. Most average returns are significantly positive, as the 95% confidence intervals suggest, indicating that they capture certain dimensions of return sources. The second subfigure presents the correlation coefficient between the residual BAM return (i.e., the return of BAM_t^* orthogonal to the six-factor model) and each factor/anomaly. These correlation coefficients are of small magnitude, and most are insignificant according to the 95% confidence intervals, indicating that the new LBA is not closely related to these factors/anomalies. Updating the six-factor model with any of these factors/anomalies cannot subsume the abnormal BAM_t^* performance, as shown by the considerable and significant alphas in the third subfigure. The remaining two subfigures examine the BAB version of the new LBA, which exhibit similar performance.

Given that extant factors or anomalies do not explain the new LBA (as a manifestation of partial-correlation factors), we can reliably conclude that the impact of partial-correlation factors probably was not documented in previous studies.

Time-varying factor exposures

This subsection examines if the abnormal performance of the new LBA is related to time-varying factor exposures.

FIGURE 2.6: Examining if Other Factors/Anomalies Explain the New Low-Beta Anomaly (LBA)
This figure reports the connection between the new LBA and each of the 187 factors/anomalies of Hou et al. (2020b). Two portfolios capturing the new LBA are used: BAM_t^{*e} and BAB_t^*. BAM_t^{*e} is the residual return from the spanning regression of a simple BAM portfolio (constructed following Equation 3.1) on the six-factor model of Fama and French (2018). BAB_t^* is the ex-ante market-neutral version of the new LBA (constructed following Equation 2.19). These portfolios are constructed using betas estimated from the six-factor model. The first subfigure reports the average monthly returns (μ_F) of these factors and anomalies. The second and fourth subfigures report the correlation coefficients. The third and fifth subfigures report the spanning regression alphas against the six-factor model updated with each factor/anomaly portfolio. The 95% confidence intervals based on Newey and West (1987) adjusted standard errors are also reported. The full sample period is from July 1968 to December 2019.



A conditional CAPM test I first conduct a conditional CAPM test following Petkova and Zhang (2005):

$$BAM_t = \alpha + (b_0 + b_1 DY_{t-1} + b_2 DEF_{t-1} + b_3 TED_{t-1} + b_4 TB_{t-1}) MKT_t + b_{SMB} SMB_t + b_{HML} HML_t + b_{RMW} RMW_t + b_{CMA} CMA_t + b_{MOM} MOM_t + \epsilon_t \tag{2.20}$$

where DY, DEF, TED and TB are respectively the dividend yield of the S&P 500 index, the yield spread between Baa and Aaa-rated corporate bonds, the yield spread between ten-year T-bonds and three-month T-bills and the three-month T-bill yield. The data are available from Amit Goyal's website. Following Wang (2019), I also control for other factors from the six-factor model. BAM_t is used when examining the known LBA and BAM_t^* is used when examining the new LBA. The estimated alpha should be indistinguishable from zero if the conditional CAPM can explain the new LBA. As shown in Panel A of Table 2.4, while the known LBA does not generate a significant alpha, the alpha of the new LBA remains considerable and significant. The results indicate that this conditional CAPM cannot explain the new LBA. Moreover, the different coefficients before DY and TED provide evidence that the known and new LBAs are oppositely related to macro-state variables, implying again that the two anomalies have different drivers.

Adjusting for ex-ante factor exposures over time Next, I remove the factor exposures of the new LBA over time using ex-ante data and examine whether the risk-adjusted portfolio still generates positive returns. At the beginning of each month t, I estimate a BAM portfolio's factor exposures using monthly returns of the previous T years up to the end of month $t-1$ and then hedge the BAM portfolio according to the estimated factor exposures. I record the return of the hedged portfolio at the end of month t. Rolling windows of three, five and ten years are considered. The first six columns of Panel B report the performance of the known and new LBAs after adjusting for ex-ante factor exposures over time. The two anomalies generate positive returns after adjusting for the market exposures using the CAPM. In contrast, when using the six-factor model, while the known LBA no longer generates significantly positive returns after risk adjustments, the new LBA's abnormal performance remains considerable and significant.

Time-varying alphas I also adjust for factor exposures using concurrent data; that is, I estimate the alpha of a BAM portfolio against the CAPM or the six-factor model at the end of each month t using returns up to that time point to get the alpha for that period. Rolling windows of three, five and ten years are considered. I obtain alphas over time and evaluate if the average alphas are significantly positive. The last six columns of Panel B report the average time-varying alphas. Similar to the results above, the average alphas of both anomalies are significant when adjusting for the CAPM. When adjusting for the six-factor model, while the known LBA does not generate significant alphas, the average alphas of the new LBA are considerable and significant.

The results indicate that the abnormal performance of the new LBA is not subsumed under the three different approaches to adjusting for time-varying risk exposures. The bottom two rows of each panel present the same tests for the market-neutral BAB portfolios, which leads to the same conclusion.

Table 2.4: Adjusting for Time-Varying Factor Exposures

This table presents the performance of the known LBA (captured by BAM_t or BAB_t) and the new LBA (captured by BAM_t^* or BAB_t^*) after controlling for time-varying factor exposures. Panel A reports the alphas (α), factor exposures and adjusted R-squares ($\bar{R}^2$) against a conditional CAPM (see Equation 2.20). DY, DEF, TED and TB are the dividend yield, default yield spread, term spread and three-month T-bill yield. SMB, HML, RMW, CMA and MOM are the size, value, profitability, investment and momentum factors. Panel B reports the average returns after adjusting for ex-ante factor exposures (the first six columns) or concurrent factor exposures (the last six columns) to the CAPM or six-factor model. Rolling windows of 3, 5 or 10 years are considered. CAPM and FF6 under each rolling window notation (e.g., 3Y) denote the model used for removing time-varying factor exposures. BAM refers to the betting-against-beta portfolio constructed following Equation (3.1). BAB refers to the market-neutral betting-against-beta portfolio constructed following Equation (2.19). BAM_t and BAB_t are constructed using market betas estimated from the CAPM. BAM_t^* or BAB_t^* are constructed using market betas estimated from the six-factor model. The t-statistics are based on Newey and West (1987) adjusted standard errors. The sample period is from July 1968 to December 2019.

Panel A: Conditional CAPM

BAB	α(%)	b_0	DY	DEF	TED	TB	SMB	HML	RMW	CMA	MOM	$\bar{R}^2$(%)
BAM_t	0.17	-1.00	0.02	0.06	-0.42	0.06	-0.50	0.46	0.74	0.93	0.03	71.10
	(0.77)	(-5.93)	(0.17)	(0.86)	(-2.23)	(1.43)	(-5.60)	(2.94)	(5.58)	(4.34)	(0.29)	
BAM_t^*	0.80	-1.00	-0.26	0.00	0.31	0.09	-0.52	0.18	0.35	-0.04	0.02	59.59
	(4.32)	(-5.71)	(-3.63)	(-0.05)	(2.66)	(3.33)	(-6.40)	(1.60)	(3.06)	(-0.23)	(0.36)	
BAB_t	0.10	0.13	0.08	0.07	-0.43	0.05	0.07	0.55	0.60	0.61	-0.03	37.62
	(0.63)	(1.04)	(1.18)	(1.51)	(-3.42)	(1.74)	(1.02)	(4.75)	(5.50)	(3.94)	(-0.39)	
BAB_t^*	0.84	-0.05	-0.23	0.04	0.16	0.12	-0.05	0.26	0.19	-0.12	0.01	6.30
	(4.66)	(-0.28)	(-3.41)	(0.86)	(1.33)	(4.79)	(-0.69)	(2.38)	(1.70)	(-0.69)	(0.12)	

Panel B: Adjusting for factor exposures over time

	Ex-Ante						Concurrent					
	3Y		5Y		10Y		3Y		5Y		10Y	
	CAPM	FF6	CAPM	FF6	CAPM	FF6	CAPM	FF6	CAPM	FF6	CAPM	FF6
BAM_t	0.66	0.15	0.57	0.12	0.60	0.08	0.61	0.12	0.64	0.12	0.61	0.04
	(3.91)	(0.57)	(3.25)	(0.47)	(3.21)	(0.33)	(4.12)	(0.85)	(5.00)	(0.74)	(8.93)	(0.20)
BAM_t^*	0.82	0.88	0.75	0.89	0.79	0.94	0.73	0.78	0.74	0.76	0.66	0.74
	(4.00)	(3.71)	(3.62)	(4.36)	(3.83)	(4.54)	(3.70)	(4.47)	(3.94)	(5.79)	(3.35)	(5.19)
BAB_t	0.74	0.07	0.74	0.09	0.70	0.07	0.71	0.13	0.74	0.12	0.76	0.06
	(3.98)	(0.46)	(3.79)	(0.60)	(3.25)	(0.41)	(5.13)	(1.43)	(6.58)	(1.17)	(10.31)	(0.48)
BAB_t^*	0.95	0.82	0.91	0.81	0.95	0.88	0.88	0.72	0.88	0.73	0.82	0.74
	(4.91)	(4.54)	(4.60)	(4.31)	(4.66)	(4.52)	(4.94)	(4.49)	(4.71)	(5.43)	(3.96)	(5.18)

2.5.4 Uncovering the new LBA through LBA's specific explanations

Besides using empirical factor models to explain LBA, recent studies also propose return sources that generate the phenomenon endogenously, such as lottery demand (Bali et al. 2017), arbitrage asymmetry (Liu et al. 2018), coskewness risk (Schneider et al. 2020) and leverage constraints (Black 1972). The basic logic is that low(high)-beta stocks are associated with characteristics rewarded with higher (lower) returns, translating into a positive BAM alpha. As discussed in Section 2.3.2, these explanations are consistent with the multifactor-model framework.

Given that these studies fully explain the known LBA, controlling for any of them should uncover a new LBA (Corollary 2.4.2). To further confirm the new LBA's existence, I first report ex-ante factor exposures and ex-post alphas of portfolios sorted on $\hat{\beta}_i$s and $\hat{\beta}_i^*$s. $\hat{\beta}_i$s are estimated from the CAPM (Equation 2.8). $\hat{\beta}_i^*$s and exposures to each specific theory's proxy are estimated from Equation 2.14.

Ex-ante factor loadings and ex-post alphas

Figure 2.7 reports ex-ante factor exposures and ex-post alphas when examining each of the following specific explanations.

Lottery Demand Bali et al. (2017) propose that the lottery-demand effect accounts for LBA. According to this explanation, high-beta stocks coincide with lottery-like stocks; hence, their expected returns are driven down due to investors' lottery demand. This extra demand generates a return spread between low- and high-$\hat{\beta}_i$ stocks unexplained by the market factor. Following Bali et al. (2011) and

Bali et al. (2017), I measure a stock's lottery demand as the average of the five highest daily returns during the prior month, with a minimum of fifteen daily return observations required, and then measure the return of the lottery-demand factor (MAX) as the return spread between the value-weighted returns of low and high lottery-demand quintiles.

As suggested in the middle-left subfigure of Panel A, sorting stocks into $\hat{\beta}_i$ quintiles picks up stocks with higher MAX-factor exposures in the low-$\hat{\beta}_i$ side and lower MAX-factor exposures in the high-$\hat{\beta}_i$ side. This extra factor exposure explains the LBA depicted in the bottom-left subfigure. The right panel depicts the performance of $\hat{\beta}_i^*$-sorted portfolios. As suggested by the bottom-right subfigure, the low(high)-$\hat{\beta}_i^*$ side generates a higher (lower) alpha against the market-plus-MAX-factor model, indicating that sorting stocks according to $\hat{\beta}_i^*$ still produces an LBA. However, the low(high)-$\hat{\beta}_i^*$ side no longer contains stocks with high(low)-exposures to the lottery-demand factor, as shown in the middle-right subfigure, which indicates that the new LBA profits from other unidentified return sources.

Arbitrage Asymmetry Stambaugh et al. (2015) propose an arbitrage asymmetry-based explanation of the low-volatility anomaly. According to this explanation, idiosyncratic volatility (IVol) is a proxy for arbitrage risk. For mispriced stocks, high-IVol stocks' prices are less corrected than low-IVol stocks' due to limits to arbitrage. As a result, among underpriced stocks, high-IVol stocks have higher expected returns than low-IVol stocks, while among overpriced stocks, high-IVol stocks have lower expected returns than low-IVol stocks. Since it is easier to buy than short, the overall pattern is a low-volatility anomaly. Liu et al. (2018) further propose that arbitrage asymmetry also explains LBA. The logic is that low-beta stocks have higher risk-adjusted returns since they coincide with low-IVol stocks.

Following Ang et al. (2006) and Liu et al. (2018), I measure idiosyncratic volatility as the standard deviation of the prior month's daily residual returns relative to the three-factor model (Fama and French 1993), and then construct the $IVol$ factor as the difference between the lowest and highest IVol-sorted quintile portfolios. Panel B reports the $\hat{\beta}_i$-sorted portfolio performance, which leads to the same conclusion as that in the case of the lottery demand-based explanation. As suggested in the middle-left subfigure, the low(high)-$\hat{\beta}_i$ stocks' high (low) exposures to the $IVol$ factor account for the known LBA. However, once the $IVol$ factor is controlled in beta estimation, LBA still exists, but the low(high)-$\hat{\beta}_i^*$ side is no longer associated with a high (low) exposure to the $IVol$ factor, as suggested by the right panel. Therefore, the new LBA is orthogonal to the volatility effect and thus is unexplained by the arbitrage asymmetry theory.

Coskewness Risk Schneider et al. (2020) attribute LBA to the coskewness risk. Besides demanding compensation for covariance risk, investors accept lower expected returns on assets with positive skewness. High-beta stocks are associated with higher coskewness and thus lower CAPM-adjusted returns. They construct alternative factors as proxies of the coskewness risk and find they subsume LBA. Since results based on these alternative factors lead to the same conclusion, I only report the result based on the SKEW factor obtained from Schneider et al. (2020) for brevity. The data is available from February 1996 to August 2008. Given that returns of the SKEW factor are available at the monthly frequency, I estimate exposures to the two factors in the market-plus-SKEW model using monthly returns of the prior two years, requiring at least 12 observations for a beta estimate to be valid.

Panel C examines if a new LBA orthogonal to the coskewness effect exists. The conclusion is the same as that under the previous two explanations: while the SKEW factor explains the known LBA, it does not explain the new LBA uncovered following Proposition 2. Low(high)-$\hat{\beta}_i$ stocks are associated with high (low) loadings on the SKEW factor, which accounts for the known LBA. In contrast, sorting stocks into $\hat{\beta}_i^*$ quantiles no longer picks up high (low) SKEW-factor exposures in the low(high)-$\hat{\beta}_i^*$ side, but the LBA still exists, indicating that the new LBA is unrelated to the coskewness risk.

Leverage Constraints and BAB^{FP} Frazzini and Pedersen (2014) attribute LBA to leverage constrained investors' higher demand for high-beta stocks. Instead of constructing a factor using characteristics capturing leverage constraints, they construct a market-neutral betting-against-beta (BAB^{FP}) portfolio using a rank-weighting scheme to capture the effect.[14] The left of Panel D shows that sorting stocks into $\hat{\beta}_i$ quintiles picks up stocks with high (low) BAB^{FP}-factor exposures in the low(high)-$\hat{\beta}_i$ side, indicating that the BAB^{FP} factor captures the known LBA. On the other hand, sorting stocks into $\hat{\beta}_i^*$ (estimated from the market-plus-BAB^{FP} model) quintiles generates a trivial cross-sectional difference in alphas against the market-plus-BAB^{FP} model, as shown in the bottom-right subfigure.

However, the result here that no new LBA emerges after controlling for the

[14]The factor is denoted as BAB in Frazzini and Pedersen (2014). Here I use BAB^{FP} to distinguish from the market-neutral betting-against-beta portfolio used in this study. The data for BAB^{FP} is available from the AQR website.

BAB^{FP} factor does not mean that the BAB^{FP} factor also captures the new LBA. As shown in the middle-right subfigure, sorting stocks according to the newly estimated market betas ($\hat{\beta}_i^*$s) still largely picks up exposures to the BAB^{FP} factor, indicating that the impact of the BAB^{FP} factor is not removed in beta estimation. Consequently, besides capturing market risk, the newly estimated betas are still mainly driven by the driver of the known LBA.

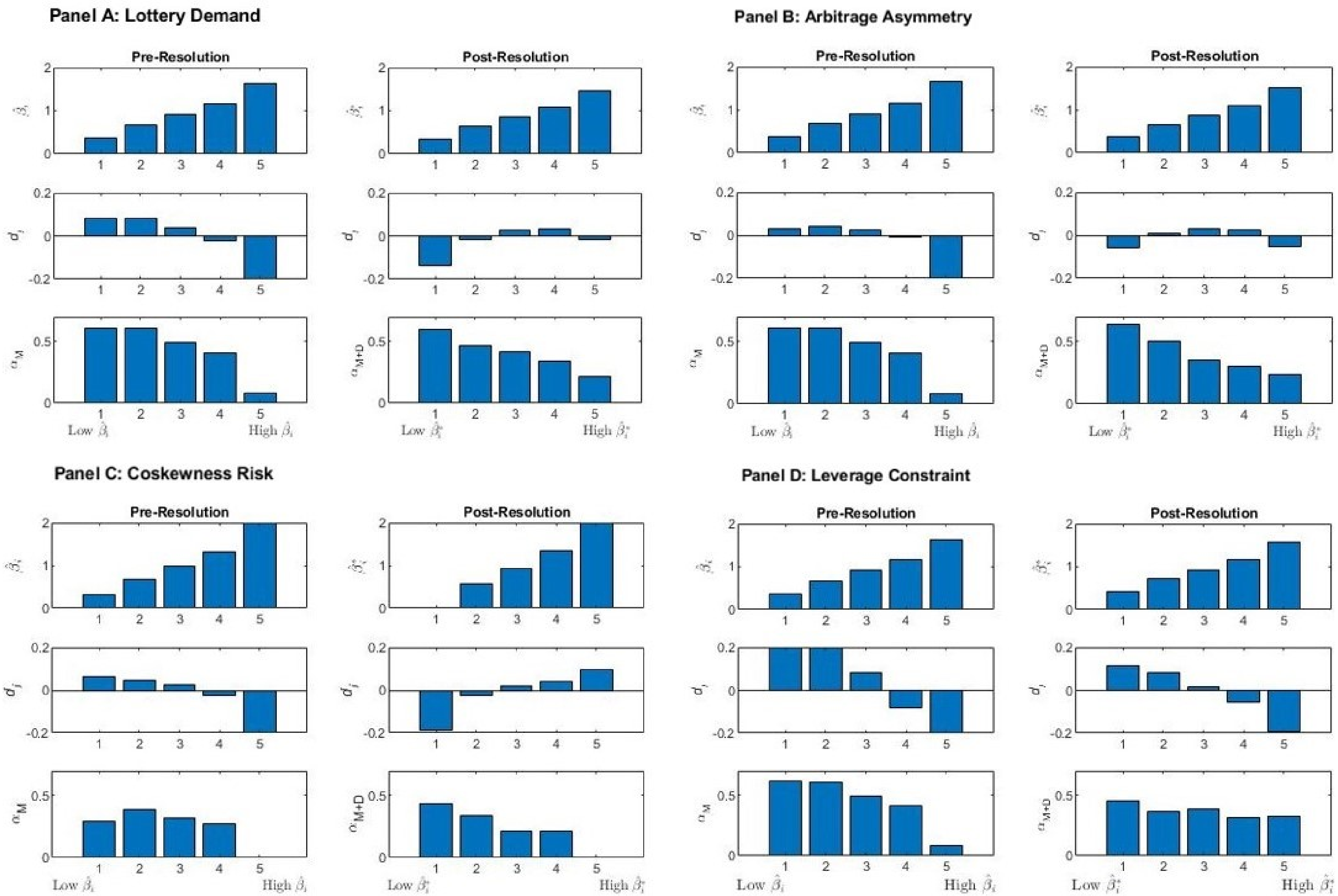
Time-Series Regressions

This section further confirms that the theories in Section 2.5.4 all reflect factors directly correlated with market risk, while a new LBA driven by partial-correlation factors can be identified by conditioning on them. I test the BAM portfolio constructed from market betas estimated from the six-factor model (BAM_t^*). As explained in Section 2.5.2, the six-factor model can be regarded as a comprehensive measure of existing explanations of LBA; accordingly, the BAM_t^* portfolio should capture information different from these return sources. I update the six-factor model with each of the four factors examined in Section 2.5.4, and examine if BAM_t^* generates abnormal returns against them. I also report the performance of BAM_t (based on $\hat{\beta}_i$s from the CAPM) for comparison.

Panels A, B and C of Table 2.5 report the spanning regression results when updating the six-factor model with the lottery-demand factor (MAX), idiosyncratic-volatility factor ($IVol$) or coskewness factor ($SKEW$). Consistent with the patterns of ex-ante factor exposures reported in Panels A to C of Figure 2.7, BAM_t has considerable and significant exposures to these factors and does not generate significantly positive alphas against these factor models. In contrast, BAM_t^* has

Figure 2.7: Testing Against Specific Explanations of the Low-Beta Anomaly (LBA)

Each panel of this figure reports the ex-ante factor exposures and ex-post alphas of portfolios sorted on $\hat{\beta}_i$s (Equation 2.8) on the left and those of portfolios sorted on $\hat{\beta}_i^*$s (Equation 2.14) on the right. $\hat{\beta}_i$s are market betas estimated from the CAPM and $\hat{\beta}_i^*$s are market betas estimated from the market-plus-D model. D factors of Panels A to D are the lottery-demand factor proposed in Bali et al. (2017), idiosyncratic-volatility factor constructed following Ang et al. (2006) and Liu et al. (2018), coskewness factor of Schneider et al. (2020) and BAB^{FP} factor of Frazzini and Pedersen (2014), respectively. In each panel, the top two subfigures report ex-ante betas, the middle two report ex-ante exposures to D and the bottom two report the ex-post spanning regression alphas (monthly, %). Factor exposures and returns of each portfolio are measured on a value-weighted basis. The sample period is from February 1996 to August 2008 for Panel C and from July 1968 to December 2019 for other panels.



much smaller exposures to these factors than BAM_t does and generates considerable and significant alphas in these spanning regressions, ranging from 0.74% to 1.16% (monthly).

Panel D reports the performance when updating the six-factor model with the BAB^{FP} factor. Although BAB^{FP} is not a direct leverage-constraint factor, Frazzini and Pedersen (2014) show that the BAB^{FP} factor's time-series performance is consistent with the leverage-constraint theory, indicating that the BAB^{FP} factor can be regarded as an indirect proxy of the leverage-constraint effect.[15] While the known LBA (BAM_t) is fully explained by the six-factor model updated with the BAB^{FP} factor, the new LBA (BAM_t^*) still generates a considerable and significant alpha (0.59%, monthly). Compared with BAM_t's exposure to the BAB^{FP} factor (0.96, significant at the 1% level), BAM_t^*'s exposure to the BAB^{FP} factor reduces drastically. As the six-factor model largely captures the return source underlying BAB^{FP} (see Novy-Marx and Velikov 2022), information related to BAB^{FP} is largely removed when estimating market betas from the six-factor model. Consequently, the BAM_t^* portfolio is able to capture return sources beyond what underlies BAB^{FP}, which is why the performance is in sharp contrast to that in Panel D of Figure 2.7. BAM_t^* is still significantly exposed to the BAB^{FP} factor since we do not control for a leverage-constraint factor when estimating market betas from the six-factor model. However, the results overall indicate that the BAB^{FP} factor, which aims to capture the economic source underlying the known LBA, cannot explain the new LBA. The conclusion remains the same when using BAB_t^* as the proxy for the new LBA.

[15] For a similar point, see Boguth and Simutin (2018), Chen and Lu (2019), Akbari et al. (2021), etc.

TABLE 2.5: The Known and New Low-Beta Anomalies (LBAs)

This table presents the spanning regression alphas (α), factor exposures and adjusted R-squares ($\bar{R}^2$) of the known and new LBAs against the six-factor model (Fama and French 2018) updated with one additional factor, D. D is chosen from the lottery-demand factor (MAX), idiosyncratic-volatility factor ($IVol$), coskewness factor ($SKEW$) or BAB^{FP} factor, all introduced in Section 2.5.4. BAM refers to the betting-against-beta portfolio constructed following Equation (3.1). BAB refers to the market-neutral betting-against-beta portfolios constructed following Equation (2.19) (based on Frazzini and Pedersen 2014 and Novy-Marx and Velikov 2022). BAM_t and BAB_t are constructed using market betas estimated from the CAPM. BAM_t^* and BAB_t^* are constructed using market betas estimated from the six-factor model. MKT, SMB, HML, RMW, CMA and MOM are the market, size, value, profitability, investment and momentum factors from Kenneth French's website. The t-statistics are based on Newey and West (1987) adjusted standard errors. The sample period is from July 1968 to December 2019.

	α(%)	MKT	SMB	HML	RMW	CMA	MOM	D	$\bar{R}^2$(%)
Panel A: $D = MAX$									
BAM_t	-0.15	-1.00	-0.10	0.31	0.32	0.70	0.06	0.57	75.41
	(-0.59)	(-11.90)	(-0.77)	(1.63)	(1.75)	(3.21)	(0.62)	(7.11)	
BAM_t^*	0.75	-1.00	-0.46	0.19	0.06	-0.42	0.03	0.18	59.49
	(3.49)	(-15.53)	(-4.40)	(1.34)	(0.41)	(-2.05)	(0.45)	(3.10)	
BAB_t	-0.16	0.42	0.39	0.33	0.28	0.43	-0.01	0.36	44.40
	(-0.97)	(8.08)	(4.65)	(2.77)	(2.24)	(3.03)	(-0.20)	(7.78)	
BAB_t^*	0.70	0.17	0.04	0.19	-0.05	-0.42	0.04	0.16	4.96
	(3.68)	(2.98)	(0.37)	(1.57)	(-0.40)	(-2.34)	(0.58)	(3.18)	
Panel B: $D = IVol$									
BAM_t	-0.11	-1.00	0.01	0.29	0.34	0.64	0.02	0.49	73.49
	(-0.48)	(-13.91)	(0.07)	(1.45)	(1.58)	(2.86)	(0.17)	(4.43)	
BAM_t^*	0.74	-1.00	-0.42	0.19	0.08	-0.39	0.02	0.16	59.07
	(3.59)	(-16.47)	(-3.72)	(1.41)	(0.53)	(-2.00)	(0.31)	(2.40)	
BAB_t	-0.13	0.35	0.43	0.35	0.34	0.45	-0.03	0.32	38.97
	(-0.79)	(6.84)	(3.95)	(2.46)	(2.21)	(2.80)	(-0.45)	(4.36)	
BAB_t^*	0.72	0.13	0.03	0.20	0.01	-0.40	0.04	0.12	3.42
	(3.83)	(2.29)	(0.30)	(1.72)	(0.07)	(-2.21)	(0.56)	(2.08)	
Panel C: $D = SKEW$									
BAM_t	-0.09	-1.00	-0.28	0.00	1.15	0.76	-0.07	0.26	82.62
	(-0.21)	(-6.75)	(-1.59)	(-0.01)	(5.51)	(2.88)	(-0.53)	(4.15)	
BAM_t^*	1.16	-1.00	-0.57	0.34	0.40	-0.50	0.26	-0.04	62.40
	(3.54)	(-8.80)	(-4.67)	(2.14)	(2.15)	(-1.88)	(2.72)	(-1.04)	
BAB_t	-0.10	0.30	0.21	0.21	0.84	0.53	-0.11	0.14	53.62
	(-0.35)	(3.12)	(1.62)	(1.36)	(5.42)	(2.68)	(-1.04)	(3.04)	
BAB_t^*	1.18	0.06	-0.09	0.35	0.27	-0.61	0.25	-0.05	8.44
	(3.78)	(0.53)	(-0.74)	(2.37)	(1.55)	(-2.47)	(2.87)	(-1.34)	
Panel D: $D =$ BAB^{FP}									
BAM_t	-0.24	-1.00	-0.52	0.06	0.23	0.40	-0.06	0.96	82.66
	(-1.58)	(-23.41)	(-7.20)	(0.62)	(2.70)	(3.17)	(-0.76)	(15.58)	
BAM_t^*	0.59	-1.00	-0.59	0.08	0.01	-0.43	-0.02	0.44	61.90
	(3.35)	(-18.87)	(-6.39)	(0.57)	(0.07)	(-2.42)	(-0.37)	(4.98)	
BAB_t	-0.28	0.18	-0.02	0.18	0.26	0.33	-0.11	0.87	62.72
	(-2.18)	(4.92)	(-0.23)	(2.18)	(3.32)	(2.89)	(-1.74)	(17.29)	
BAB_t^*	0.62	0.05	-0.15	0.10	-0.09	-0.49	-0.02	0.46	10.80
	(3.46)	(0.94)	(-1.65)	(0.76)	(-0.76)	(-2.72)	(-0.26)	(4.99)	

In an unreported test, I confirm that BAM_t^* (or BAB_t^*) still generates a considerable and significant alpha when updating the six-factor model with these four factors simultaneously. The time-series regression results again confirm that while previous explanations subsume the known LBA, they do not explain the new LBA. On the one hand, the way these theories resolve the known LBA is consistent with Proposition 1. On the other hand, the result indicates that all these existing theories capture factors directly correlated with market risk, while there exist partial-correlation factors that can also induce an LBA, supporting Proposition 2.

2.6 Additional Discussion

This section discusses some additional implications of the new low-beta anomaly (LBA).

2.6.1 The new LBA reflects new return sources

According to the mimicking-portfolio theorem (Cochrane 2005), risk factors can be well captured by factor-mimicking portfolios, which suggests that factor-mimicking portfolios are proxies for different economic sources. Although a betting-against-beta portfolio is not constructed using true betas or underlying characteristics, it can only originate from factor-related returns, as shown in Section 2.3.2. In this sense, an LBA can be regarded as an indirect proxy for an underlying return source; hence, uncovering a new LBA essentially identifies the existence of new economic sources.

A problem for concluding that the new LBA is a manifestation of factors different from what underlies the known LBA is the empirical flexibility of constructing factor-mimicking portfolios. Due to this empirical flexibility, different proxies for the same underlying return source may perform differently. For example, a momentum portfolio can be constructed using returns of different past horizons, and their performance difference may be significant (e.g., Jegadeesh and Titman 1993; George and Hwang 2004; Jegadeesh and Titman 2011; Novy-Marx 2012). This scenario may discredit the claim of identifying a new economic source based on a better-performing factor-mimicking portfolio.

However, two features of the new LBA can largely alleviate this concern. First, the new LBA, theoretically, is orthogonal to factors underlying the known LBA (see Equation 2.16), and it is indeed so empirically. Second, the new LBA achieves a Sharpe ratio higher than that of the known LBA. The first feature suggests that the known and new LBAs are likely to capture different return sources, as different proxies for the same return source, such as momentum, are highly correlated. The second feature also indicates that the two LBAs are driven by different underlying factors given the prior that there should be an upper bound for factor Sharpe ratios (e.g., MacKinlay 1995; Pukthuanthong et al. 2019); otherwise, the factor that explains both LBAs will have a very high Sharpe ratio. Moreover, the empirical finding in Section 2.5.3 that the large set of factors/anomalies, which largely spans existing return sources, cannot subsume the new LBA also indicates the new LBA captures return sources unnoticed by previous studies.

2.6.2 A higher bar for future studies to resolve LBA

Section 2.3 clarifies the mechanism for LBA to emerge under the multifactor-model framework. Under this framework, the economic content of LBA is not fixed, as shown in Section 2.4. The empirical results confirm that the LBA known in the literature is a manifestation of factors directly correlated with market risk, while there exists a new LBA as a manifestation of partial-correlation factors. Given the consistency of the empirical results and model implications, I propose to raise the bar for future studies to resolve LBA.

The Higher Bar: A sufficient explanation of the low-beta anomaly (LBA) should explain its different versions. No new LBAs should emerge once a sufficient explanation is controlled.

Accordingly, one additional test containing the following two steps should be conducted when an explanation is proposed:

Step 1: re-estimate market betas with the proposed explanation controlled and construct a new betting-against-beta (BAM) portfolio using the newly estimated market betas.

Step 2: test whether the new BAM portfolio generates abnormal performance when controlling for the market factor and the proposed explanation.

Step 1 enables estimated market betas to capture exposures to return sources beyond what underlies the proposed explanation, which is the precondition for a new LBA, if existing, to be identified. As discussed in Corollary 2.4.2, only when the impact from the existing explanation is removed will estimated market betas capture other return sources. **Step 2** not only reveals whether an explanation is sufficient but also identifies the existence of new LBAs. If the new BAM portfolio corresponding to a specific explanation generates abnormal returns, the explanation does not pass this test and hence is insufficient. If all previous explanations do not pass this test, then there is a new LBA whose economic content differs from the previous understanding of LBA.

2.6.3 To uncover the new LBA, removing the impact of the known LBA is crucial

According to Corollary 2.4.2, removing the impact of factors directly correlated with market risk is the precondition for uncovering the impact of partial-correlation factors. When re-estimating market betas, controlling for a factor that subsumes the known LBA can usually achieve this purpose. However, the empirical flexibility of constructing factor-mimicking portfolios may generate misleading results. A factor-mimicking portfolio may be constructed in a way that its overlapping information with the market factor is removed. In this case, the factor still captures factors directly correlated with market risk, but these factors' impact on estimated market betas is not removed when re-estimating market betas.

The performance after controlling for the BAB^{FP} factor of Frazzini and Pedersen (2014) reported in Figure 2.7 depicts this scenario. Since the BAB^{FP} factor is constructed to be market-neutral, it cannot remove the overlapping information between the market portfolio and factors directly correlated with market risk. As a result, the newly estimated betas still capture factors directly correlated with market risk rather than partial-correlation factors. To avoid misleading empirical results, we need to conduct an additional check immediately after **Step 1** of Section 2.6.2; that is, confirm if the new BAM portfolio is no longer predominantly exposed to the proposed explanation. For example, if we check Panel D of Figure 2.7, we can find that sorting stocks into quintiles according to new estimated betas still picks up cross-sectional difference in the exposure to the BAB^{FP} factor (the middle-right subfigure). This pattern indicates that the impact of BAB^{FP} on estimated betas is not removed when re-estimating betas. Further, as shown in Panel D of Table 2.5, BAB^{FP} does not explain the new LBA if we remove the impact of factors underlying the known LBA, which confirms that BAB^{FP} only captures the known LBA. Therefore, to uncover the impact of partial-correlation factors, it is crucial to **remove the impact of factors underlying the known LBA**.

2.7 Conclusion

While recent studies propose different theories that resolve the low-beta anomaly (LBA), this study shows that LBA is not limited to what has been documented in the literature. Accordingly, I uncover a new LBA unexplained by existing studies. I show that, theoretically, the known and new LBAs differ in their underlying economic sources.

To demonstrate the rationale, I first show that LBA is a manifestation of factors omitted by market risk as estimated betas, besides capturing the true market risk, also capture exposures to omitted factors. Then, I show that two types of factors can induce LBA: 1) *factors directly correlated with market risk* and 2) *factors not directly but partially correlated with market risk* (referred to as partial-correlation factors). Before the known LBA is resolved, only factors directly correlated with market risk are captured by estimated betas; hence, the known LBA only reflects these factors. Based on recent studies resolving the known LBA, I allow estimated betas to capture partial-correlation factors and hence achieve a new LBA as a manifestation of these factors.

As a manifestation of omitted factors, LBA provides an effective approach to identifying important asset-pricing factors before their underlying characteristics are recognized. For example, although the investment and profitability factors were formally introduced into benchmark factor models in the 2010s, the initial discovery of LBA in the 1970s already indicated their existence. In this sense, the new LBA suggests the omission of important asset-pricing factors. The significance of the new LBA lies in that the factors it identifies are hardly noticed otherwise. Partial-correlation factors do not directly bias the cross-sectional beta-return relationship regulated by the CAPM. To the extent that asset pricing factors or anomalies are discovered through violations of a benchmark model, the impact of partial-correlation factors was likely not noticed. I confirm this conjecture by showing that extant factors or anomalies do not explain the new LBA.

Essentially, factors directly correlated with market risk correspond to direct violations of the CAPM (i.e., violating the cross-sectional beta-return relationship prescribed by the CAPM). Instead, partial-correlation factors correspond to partial violations of the CAPM (i.e., not violating the CAPM unconditionally but violating the CAPM conditionally). Given this correspondence, the discovery of the new LBA shed light on theoretical studies that refine the CAPM. These studies usually release assumptions of the standard CAPM (e.g., Jagannathan and Wang 1996; Hong and Sraer 2016; Andrei et al. 2021), which essentially reconcile a direct violation of the CAPM. However, the discovery of the new LBA suggests that reconciling a direct violation of the CAPM cannot unveil all the pricing information it omits, but future studies also need to introduce partial violations of the CAPM.

This finding also refreshes the understanding of LBA in many other dimensions. For example, while previous studies regard the low-beta and low-volatility anomalies as the same (e.g., Bali et al. 2011; Bali et al. 2017; Stambaugh et al. 2015; Liu et al. 2018; Schneider et al. 2020), the new LBA indicates the existence of low-beta effect orthogonal to the low-volatility effect. In this sense, the new LBA revives the low-risk strategy. The finding also translates into a higher bar for future studies to resolve LBA: a sufficient explanation should also resolve the new LBA.

Chapter 3

Demystifying Factor-Beta Anomalies

3.1 Introduction

Most asset pricing theories come down to a factor with positive premiums summarizing a positive expected return-beta relationship (e.g., Sharpe 1964; Lintner 1965; Merton 1973; Ross 1976; Chamberlain and Rothschild 1983). However, empirical studies find that buying/selling low/high-beta stocks (BAF, betting-against-factor-beta) generates a positive alpha against the factor-mimicking portfolio (FMP) from which betas are estimated.[1] While earlier evidence in this strand concentrates on the low market-beta anomaly (e.g., Black et al. 1972; Frazzini and Pedersen 2014), recent studies find that the phenomenon also exists for other

[1]According to the mimicking-portfolio theorem (Cochrane 2005), risk factors can be well captured by factor-mimicking portfolios. Hereafter, I use FMP (F_t) to denote a factor-mimicking portfolio from which betas are estimated.

factors (e.g., Herskovic et al. 2019). Some studies regard the phenomenon as a challenge to the covariance-based asset pricing framework (e.g., Murray 2020). Other studies attribute the phenomenon to the unpriced-risk component embedded in an FMP (e.g., Pukthuanthong et al. 2019; Daniel et al. 2020b).

Different BAF portfolios usually correlate oppositely, making it hard to reconcile different factor-beta anomalies through a specific risk or behavioral bias. Instead, the pervasiveness of factor-beta anomalies indicates that a common mechanism drives the phenomenon. The challenge facing the clarification of the mechanism is that the empirical proxies for the factors that fully capture return comovement are unclear, making it unpersuasive to draw rigorous conclusions from empirical analyses. To circumvent this empirical limitation, this study adopts an analytical approach, explicitly modeling the mechanism under a standard factor-model framework to demystify factor-beta anomalies.

For two reasons, employing an explicit modeling approach to address factor-beta anomalies is reasonable. First, by construction, a BAF portfolio captures information contained in estimated factor betas; hence, the pervasiveness of factor-beta anomalies indicates that betas indeed capture information relevant to pricing, making it natural to study factor-beta anomalies in the framework under which betas are defined. Second, unlike a violation of the positive market beta-return relationship that directly suggests a violation of equilibrium asset pricing theories such as the CAPM, the implication of a non-zero factor beta-alpha relationship is unclear as there is no consensus on how other factors or anomalies emerge, which makes it more attractive to clarify the general mechanism for factor-beta anomalies to emerge so as to pave the way for further economic interpretations.

I start by modeling the mechanism for a factor-beta anomaly to emerge under a standard factor-model framework. When an FMP is proposed to capture a beta-return relationship, factors beyond the FMP usually exist, which, if not controlled, are omitted factors when estimating individual assets' FMP betas. If some omitted factors correlate with the FMP, estimated FMP betas will also capture individual assets' loadings on these factors.[2] In this case, sorting stocks into quantiles according to estimated FMP betas will pick up the cross-sectional difference in loadings on omitted factors, and hence a BAF portfolio based on estimated FMP betas is exposed to omitted factors. The BAF portfolio's additional factor exposures can translate into a factor-beta anomaly. Explicitly modeling the content of estimated FMP betas makes it immediately clear that a factor-beta anomaly should be an outcome of a multi-factor model but not a challenge, as estimated FMP betas only capture return sources related to systematic factors. Accordingly, the cross-sectional difference in alphas unrelated to factors will not be captured when sorting stocks into estimated FMP-beta quantiles. Therefore, a factor-beta anomaly can only be caused by omitted factors correlated with the FMP from which betas are estimated.

Next, I clarify the tendency for low factor-beta anomalies to emerge. The connection between an FMP and omitted factors has two impacts on the estimated beta: first, the estimated FMP beta deviates from the true FMP beta, that is, beta deviation; second, the estimated FMP beta also captures loadings on omitted

[2]An FMP correlates with an omitted unpriced factor may be because the FMP contains an unpriced component. For example, characteristics used for constructing factors also capture the unpriced comovement of returns, and hence a characteristics-sorted factor portfolio also contains an unpriced component (e.g, Roll and Ross 1984; Asness et al. 2000). An FMP correlates with an omitted priced factor may be because there exists an intrinsic connection between two factors or a characteristic captures multiple dimensions of return sources.

factors, that is, extra premium. Beta deviation always induces a negative beta-alpha relationship. We can understand this through the BAF portfolio. As the estimated FMP beta also captures loadings on omitted factors, a BAF portfolio's observed exposure to its corresponding FMP has a larger magnitude than its true exposure to the FMP if all omitted factors were known. In a spanning regression, the BAF's exposure to the FMP is identified as the explained component of the BAF return, and the BAF's true exposure to the FMP is associated with a negative premium; their difference is identified as a positive BAF alpha. On the other hand, whether extra premium makes the BAF alpha positive or negative depends on the correlations between the FMP and omitted factors. Given that a negative beta-alpha relationship is contributed by both beta deviation and extra premium while a positive beta-alpha relationship is contributed by extra premium but mitigated by beta deviation, however, low but not high factor-beta anomalies should be pervasive when betting against betas of priced factors with positive premiums.

Third, I address how omitted unpriced and priced factors contribute to a low factor-beta anomaly. Unpriced factors can only induce low factor-beta anomalies as they only come into effect through beta deviation. Priced factors induce a factor-beta anomaly through both beta deviation and extra premium, whose impact on average thus also tends to be a low factor-beta anomaly. As exposures to omitted priced factors carry premiums, the strength of a low factor-beta anomaly from the priced-factor channel is flexible. In contrast, The strength of a low factor-beta anomaly from the unpriced-factor channel should be weak under the typical definition of unpriced factors. An unpriced factor is unrewarded since it is not pervasive, and investors can diversify, for example, across industries to avoid such risks (e.g., Roll and Ross 1984; Daniel and Titman 1997; Pukthuanthong et al. 2019; Daniel et al. 2020b; Clarke 2022). Under this definition, most stocks should not be notably exposed to an unpriced factor, and hence an unpriced factor should only cause a weak level of beta deviation, which in turn can only induce a weak low factor-beta anomaly. Moreover, there is an upper bound for the strength of a low factor-beta anomaly from the unpriced-factor channel, and a BAF alpha exceeding this bound indicates the existence of omitted priced factors.

Although the proxies of the full set of priced and unpriced factors that fully capture return comovement are unclear, the existence of a large set of factor and anomaly portfolios can still shed light on the rationality of the aforementioned analytical conclusions. In empirical analyses, I first examine the BAF portfolios constructed from 116 anomaly portfolios that previous studies find to capture priced return sources (e.g., Hou et al. 2020b). 106 of the 116 BAF alphas are positive, and many are significant, confirming the analytical conclusion that low factor-beta anomalies should be pervasive when betting against betas of priced factors with positive premiums. In contrast, when betting against betas of industry portfolios, which are widely referenced as proxies for unpriced factors, the BAF alphas are no longer pervasively positive, and most are insignificant. The result is consistent with the analytical conclusion that systematic factors cause beta anomalies. To shed light on to what extent unpriced factors contribute to BAF alphas in general, I remove the beta deviation caused by unpriced factors by controlling for industry returns in the spanning regressions of BAFs on their corresponding FMPs. Most BAF alphas attenuate but remain considerable and significant. This result shows that unpriced factors positively contribute to a BAF alpha, but their impact is usually not strong. These BAF alphas change remarkably when the six-factor model of Fama and French (2018) is controlled, supporting the analytical conclusion that priced factors should dominate unpriced factors in driving factor-beta anomalies.

This study is directly related to recent studies on the pervasiveness of beta anomalies among risk factors (e.g., Herskovic et al. 2019; Murray 2020). While these studies show that low-beta anomalies exist for factor portfolios, I also confirm that low-beta anomalies exist for a large set of anomaly portfolios. More importantly, I clarify the tendency for low factor-beta anomalies to emerge when betting against betas of priced factors with positive premiums. This study is also directly related to studies that use BAF portfolios for no-cost hedges (e.g., Herskovic et al. 2019; Murray 2020; Daniel et al. 2020b). My findings indicate that beta anomalies are mostly compensations for priced-factor exposures; hence, a BAF hedge is unlikely to be cost-free. In terms of utilizing residual information, this study is related to MacKinlay and Pástor (2000) and Giglio and Xiu (2021). The former study investigates the implications of mispricing embedded within the residual covariance matrix on portfolio selection, and the latter study utilizes omitted factors to improve risk premium estimation of observable factors, while we investigate the mechanism for omitted factors to bias beta-return relationships and its implications for factor-beta anomalies.

The remainder of this paper is structured as follows. Section 3.2 describes the background and motivation. Section 3.3 demystifies factor-beta anomalies under a standard factor-model framework. Section 3.4 examines the performance of a large set of betting-against-beta portfolios. Section 3.5 provides concluding remarks.

3.2 Background and Motivation

A basic implication of a covariance-based asset-pricing model is that assets with a higher risk, measured as their comovement with a factor (i.e., factor beta), tend to have higher expected returns. This study refers to a violation of this implication as a factor-beta anomaly. This section introduces the definition of the factor-beta anomaly from the perspective of the betting-against-factor-beta (BAF) portfolio and provides empirical examples.[3]

3.2.1 Defining the factor-beta anomaly - A non-zero BAF alpha

In empirical studies, a factor-mimicking portfolio (FMP) is constructed as the proxy of an underlying return source. Consistent with the definition of systematic risk, factor betas are estimated as the comovement between individual assets and the corresponding FMP:

$$r_{it}^{e} = \alpha_i + \beta_i F_t + \epsilon_{it} \tag{3.1}$$

where r_{it}^{e} is the excess return of security i and F_t is the FMP from which betas are estimated.

After obtaining estimated betas on an FMP, a betting-against-factor-beta (BAF) portfolio is constructed by buying/selling low/high-β_i stocks

$$BAF_t = \sum_{\{i:\ \beta_i<\beta_L\}} w_i r_{it}^{e} - \sum_{\{i:\ \beta_i>\beta_H\}} w_i r_{it}^{e} \tag{3.2}$$

[3]Strictly speaking, a beta anomaly usually refers to a significant violation of a positive expected return-beta relationship. I use a loose definition as this study primarily focuses on analyzing the sign of a factor beta-alpha relationship.

where β_L and β_H are the thresholds for assets to be categorized as low or high beta assets; w_i is the weight of security i in the long (low-beta) or short (high-beta) side, satisfying $\sum_{\{i:\ \beta_i<\beta_L\}} w_i = \sum_{\{i:\ \beta_i>\beta_H\}} w_i = 1$.[4]

Whether a factor-beta anomaly (i.e., a non-zero cross-sectional beta-alpha correlation) exists depends on if a non-zero BAF alpha is identified in the following spanning regression:

$$BAF_t = \alpha_{BAF} + \Delta_\beta F_t + \epsilon_{BAF,t} \tag{3.3}$$

where Δ_β is the loading of the BAF portfolio on the FMP, $\Delta_\beta E(F_t)$ is the explained component of the BAF return, α_{BAF} is the component of the BAF return unexplained by the FMP, that is, the BAF alpha. If the BAF portfolio does not capture returns unrelated to F_t, α_{BAF} should be negligible; otherwise, a non-zero α_{BAF} will be observed.

3.2.2 Examples of (low) factor-beta anomalies

I use the five factors of Fama and French (2015) and the momentum factor as examples to introduce factor-beta anomalies. Table 3.1 reports the performance of

[4]that at the beginning of each time t, β_is are estimated using returns up to that time point, and the market capitalization of stocks determines w_is; that is, the BAF return is

$$BAF_t = \sum_{\{i:\ \beta_{i\bar{t}}<\beta_{L\bar{t}}\}} w_{i\bar{t}} r^e_{it} - \sum_{\{i:\ \beta_{i\bar{t}}>\beta_{H\bar{t}}\}} w_{i\bar{t}} r^e_{it}$$

where $\bar{t}$ refers to the beginning of t; $\beta_{L\bar{t}}$ and $\beta_{H\bar{t}}$ are the thresholds; e.g., NYSE 30/70th percentile breakpoints at the beginning of t. As betas are persistent, I drop the time subscripts of w_i and β_i for brevity. Allow betas to be time-varying does not affect the analytical conclusions as, in this case, we can adjust for factor exposures over time.

these factors and their corresponding BAF portfolios. Details of the data used are introduced in Section 3.4.1.

The pervasiveness of (low) factor-beta anomalies The first row of Panel A reports BAF returns (μ_{BAF}). Consistent with previous studies (e.g., Herskovic et al. 2019; Murray 2020; Daniel et al. 2020b), the return-beta relationship is flat for all these factors. The BAF returns of size, value and investment factors are slightly negative but indistinguishable from zero. The remaining rows of Panel A report the spanning regression results (Equation 3.3). The slopes (b_{BAF}) are significantly negative, and the adjusted R-squares ($\bar{R}^2$) are considerable. As BAF portfolios are constructed using betas estimated from pre-formation period returns and spanning regressions are based on post-formation returns, the results show that pre-formation betas well capture post-formation betas. The significantly positive alphas (α_{BAF}) indicate the existence of beta anomalies, indicating that BAF portfolios capture returns unexplained by the original factor-mimicking portfolios. Moreover, as suggested by these positive BAF alphas, it is notable that all these factor-beta anomalies are low factor-beta anomalies.

The strength of a typical low factor-beta anomaly An overlooked fact is the strength of a typical low factor-beta anomaly. For ease of comparison, this study refers to the strength of a factor-beta anomaly as a BAF alpha (α_{BAF}) divided by the magnitude of the explained component of the BAF return ($|b_{BAF}E(F_t)|$). To better perceive this, I report scaled BAF alphas, $\frac{\alpha_{BAF}}{|b_{BAF}|}$, in Panel D of Table 3.1. By comparing with average factor returns (μ_F), we can observe that for BAFs that exhibit a low factor-beta anomaly (RMW, CMA and MOM), the scaled BAF alpha is as large as or even larger than the corresponding factor return. For example, for the investment factor (CMA), the scaled BAF alpha is 0.29%, similar to the average CMA factor return (0.31%). As the large magnitude of BAF alphas is pervasive, a mechanism to reconcile factor-beta anomalies should account for the strength of a typical low factor-beta anomaly.

Analysis Studies on the CAPM-beta anomaly usually attribute it to market frictions, behavioral biases or other priced risks (e.g., Black 1972; Bali et al. 2017; Liu et al. 2018; Schneider et al. 2020). However, the fact reported in Panel B of Table 3.1 that different BAF portfolios are negatively correlated makes an explanation in this strand hard to explain different factor-beta anomalies jointly. For example, suppose that a portfolio capturing a certain behavioral bias fully explains the CAPM-beta anomaly. In that case, it cannot explain the investment-beta anomaly since the two BAF portfolios are strongly negatively correlated. Moreover, unlike a negative market beta-alpha relationship that directly suggests a violation of an equilibrium model, the economic implication of a violation of a positive factor beta-alpha relationship is unclear. As such, it is necessary to clarify the general mechanism for factor-beta anomalies to emerge. This study focuses on analyzing whether a standard covariance-based asset-pricing framework reconciles the observed factor-beta anomalies and establishes the paradigm for inferring information from an observed factor-beta anomaly.

TABLE 3.1: Factor and Betting-Against-Factor-Beta Portfolios

This table presents the summary statistics of factor portfolios (F_t) and their corresponding betting-against-factor-beta portfolios (BAF_t). Panel A reports monthly average BAF returns, spanning regression alphas (α_{BAF}), factor loadings (Δ_β), and adjusted R-squares ($\bar{R}^2$) of BAF_t on F_t. Panel B reports the correlation coefficients among BAF portfolios. Panel C reports the correlation coefficients among original factor portfolios. Panel D reports average monthly factor returns and scaled BAF alphas. SMB, HML, RMW, CMA and MOM are the size, value, profitability, investment and momentum factors from Kenneth French's website. The t-statistics are based on Newey and West (1987) adjusted standard errors. The full sample period is from July 1963 to December 2019, with data from the first five years used for the initial beta estimation.

	SMB	HML	RMW	CMA	MOM
Panel A: *BAF* Returns and Spanning Regressions					
$\mu_{BAF}(\%)$	-0.09	-0.23	0.05	-0.02	0.20
	(-0.67)	(-1.47)	(0.31)	(-0.14)	(1.38)
α_{BAF}	0.02	0.08	0.37	0.33	0.40
	(0.23)	(0.74)	(3.64)	(2.74)	(2.89)
Δ_β	-0.77	-1.01	-1.19	-1.13	-0.33
	(-23.27)	(-18.40)	(-17.72)	(-13.57)	(-5.68)
$\bar{R}^2(\%)$	53.44	57.63	53.03	37.28	15.98
Panel B: Correlation Coefficients among *BAF*s					
SMB		-0.40	-0.54	-0.65	-0.18
HML			0.18	0.82	-0.47
RMW				0.37	0.21
CMA					-0.20
Panel C: Correlation Coefficients among FMPs					
SMB		-0.07	-0.37	-0.08	-0.07
HML			0.09	0.70	-0.19
RMW				0.01	0.10
CMA					0.01
Panel D: Factor Returns and Scaled *BAF* Alphas					
$\mu_F(\%)$	0.14	0.31	0.28	0.31	0.62
	(1.15)	(2.43)	(2.87)	(3.67)	(3.49)
$\frac{\alpha_{BAF}}{\lvert\Delta_\beta}\rvert(\%)$	0.03	0.08	0.32	0.29	1.23

3.3 Demystifying Factor-Beta Anomalies Under a Standard Factor-Model Framework

This section clarifies the mechanisms for factor-beta anomalies to emerge, explains the pervasiveness of low factor-beta anomalies, and analyzes the role of unpriced and priced omitted factors in contributing to a factor-beta anomaly.

3.3.1 The emergence of factor-beta anomalies

Whenever a factor-mimicking portfolio (FMP) is proposed, theoretically, we can update it with true latent factors to get a multi-factor model that fully captures return comovement, that is, updating the empirical model of Equation (3.1) as:

$$r_{it}^e = \alpha_i^* + \beta_i^* F_t + \boldsymbol{\gamma}_i^T \boldsymbol{\Gamma_t} + \epsilon_{it}^* \tag{3.4}$$

where r_{it}^e is the excess return of security i, F_t is the FMP from which betas are estimated, $\boldsymbol{\Gamma_t} = [\Gamma_{1t} \dots \Gamma_{kt} \dots \Gamma_{Kt}]^T$ is the K-by-1 vector of omitted factors, β_i^* is the true FMP beta if all factors correlated with F_t were known, $\boldsymbol{\gamma}_i = [\gamma_{1i} \dots \gamma_{ki} \dots \gamma_{Ki}]^T$ is the vector of the true loadings on omitted factors, and $\alpha_i^* + \epsilon_{it}^*$ is the idiosyncratic component of return. Note that α_i^* and ϵ_{it}^* are unrelated to factors or factor loadings, that is $Cov(\alpha_i^* + \epsilon_{it}^*, [F_t \; \boldsymbol{\Gamma_t}^T]) = \mathbf{0}$ and $Cov_{CS}(\alpha_i^* + \epsilon_{it}^*, [\beta_i^* \; \boldsymbol{\gamma}_i^T]) = \mathbf{0}$; otherwise, we can add additional factors to the factor model to capture the correlated part.

For ease of exposition, I assume without loss of generality that all factors have the same variance and set the omitted factors' connections with the FMP to be linear:

$$\boldsymbol{\Gamma_t} = \boldsymbol{\alpha_\Gamma} + \boldsymbol{\rho} F_t + \boldsymbol{e_t} \tag{3.5}$$

where $\boldsymbol{\rho} = [\rho_1 \ldots \ \rho_k \ldots \ \rho_K]^T$ is the K-by-1 vector of the loadings of omitted factors on F_t; $\boldsymbol{\rho}$ is also the vector of correlation coefficients under the assumption that all factors have the same variance.

Essentially, an asset pricing anomaly is caused by the deviation of an empirical model (Equation 3.1) used for beta estimation from the model that fully captures return comovement (Equation 3.4). To understand how this affects beta estimation, we can write Equation (2.9) as

$$r^e_{it} = (\alpha^*_i + \boldsymbol{\gamma}^T_i \boldsymbol{\alpha_\Gamma}) + (\beta^*_i + \boldsymbol{\gamma}^T_i \boldsymbol{\rho}) F_t + (\boldsymbol{\gamma}^T_i \boldsymbol{e_t} + \epsilon^*_{it}) \tag{3.6}$$

By comparing with Equation (3.1), we have

$$\beta_i = \beta^*_i + \boldsymbol{\rho}^T \boldsymbol{\gamma}_i \tag{3.7}$$

When constructing a BAF portfolio following Equation (3.2) by buying/shorting low/high-β_i assets, the BAF portfolio's exposure to F_t satisfies

$$\Delta_\beta = \Delta_{\beta^*} + \boldsymbol{\rho}^T \boldsymbol{\Delta_\gamma} \tag{3.8}$$

where $\Delta_\beta = \sum_{\{i:\ \beta_i<\beta_L\}} w_i\beta_i - \sum_{\{i:\ \beta_i>\beta_H\}} w_i\beta_i$ is the spanning regression slope of BAF_t on F_t under the assumption that betas are persistent, and $\boldsymbol{\Delta_\gamma} = [\Delta_{\gamma_1} \ldots \ \Delta_{\gamma_k} \ldots \ \Delta_{\gamma_K}]^T$ is the K-by-1 vector of the BAF portfolio's exposures to omitted factors. Δ_β is observable empirically, while Δ_{β^*} and $\boldsymbol{\Delta_\gamma}$ are latent true factor exposures.[5]

For a BAF portfolio, the unsystematic component of returns $(\alpha_i^* + \epsilon_{it}^*)$ of its long and short sides cancel out, as this component is unrelated to factor loadings and thus will not be picked up by estimated betas (β_i) in the cross-section. On the other hand, as β_i picks up loadings on omitted factors, a BAF return, beyond negatively capturing the FMP return, will only reflect returns related to omitted factors, that is,

$$BAF_t - \Delta_\beta F_t = \underbrace{\boldsymbol{\Delta_\gamma}^T \boldsymbol{\alpha_\Gamma}}_{\alpha_{BAF}} + \underbrace{\boldsymbol{\Delta_\gamma}^T \boldsymbol{e_t}}_{\epsilon_{BAF,t}} \tag{3.9}$$

A factor-beta anomaly emerges if α_{BAF} is not zero.[6]

Analytical Conclusion 1: A factor-beta anomaly is not a challenge to the

[5] Given Equation (3.3), we have

$$\begin{aligned}\Delta_\beta =& \frac{Cov(BAF_t, F_t)}{Var(F_t)} = \frac{Cov(\sum_{\{i:\ \beta_i<\beta_L\}} w_i r_{it}^e - \sum_{\{i:\ \beta_i>\beta_H\}} w_i r_{it}^e, F_t)}{Var(F_t)} \\ =& \sum_{\{i:\ \beta_i<\beta_L\}} w_i\beta_i - \sum_{\{i:\ \beta_i>\beta_H\}} w_i\beta_i\end{aligned}$$

Given Equation (3.7), we have

$$\begin{aligned}\Delta_\beta =& \sum_{\{i:\ \beta_i<\beta_L\}} w_i(\beta_i^* + \boldsymbol{\rho}^T\boldsymbol{\gamma}_i) - \sum_{\{i:\ \beta_i>\beta_H\}} w_i(\beta_i^* + \boldsymbol{\rho}^T\boldsymbol{\gamma}_i) \\ =& \Big(\sum_{\{i:\ \beta_i<\beta_L\}} w_i\beta_i^* - \sum_{\{i:\ \beta_i>\beta_H\}} w_i\beta_i^*\Big) + \boldsymbol{\rho}^T\Big(\sum_{\{i:\ \beta_i<\beta_L\}} w_i\gamma_i - \sum_{\{i:\ \beta_i>\beta_H\}} w_i\gamma_i\Big)\end{aligned}$$

Defining $\Delta_{\beta^*} = \sum_{\{i:\ \beta_i<\beta_L\}} w_i\beta_i^* - \sum_{\{i:\ \beta_i>\beta_H\}} w_i\beta_i^*$ and $\boldsymbol{\Delta_\gamma} = \sum_{\{i:\ \beta_i<\beta_L\}} w_i\gamma_i - \sum_{\{i:\ \beta_i>\beta_H\}} w_i\gamma_i$, we have Equation (3.8).

[6] Given Equations (3.2) and (3.4), we have

$$\begin{aligned}BAF_t =& \sum_{\{i:\ \beta_i<\beta_L\}} w_i(\alpha_i^* + \beta_i^* F_t + \boldsymbol{\gamma}_i^T\boldsymbol{\Gamma_t} + \epsilon_{it}^*) - \sum_{\{i:\ \beta_i>\beta_H\}} w_i(\alpha_i^* + \beta_i^* F_t + \boldsymbol{\gamma}_i^T\boldsymbol{\Gamma_t} + \epsilon_{it}^*) \\ =& \sum_{\{i:\ \beta_i<\beta_L\}} w_i(\beta_i^* F_t + \boldsymbol{\gamma}_i^T\boldsymbol{\Gamma_t}) - \sum_{\{i:\ \beta_i>\beta_H\}} w_i(\beta_i^* F_t + \boldsymbol{\gamma}_i^T\boldsymbol{\Gamma_t}) \\ &+ \sum_{\{i:\ \beta_i<\beta_L\}} w_i(\alpha_i^* + \epsilon_{it}^*) - \sum_{\{i:\ \beta_i>\beta_H\}} w_i(\alpha_i^* + \epsilon_{it}^*)\end{aligned}$$

covariance-based asset-pricing framework but only emerges when estimated betas capture exposures to omitted factors. In other words, a betting-against-factor-beta (BAF) alpha can only be induced by factors that capture return comovement.

3.3.2 The pervasiveness of low factor-beta anomalies

Next, I decompose the BAF alpha of Equation (3.9) further to analyze its sign:

$$\alpha_{BAF} = (\Delta_{\beta^*} - \Delta_{\beta})E(F_t) + \boldsymbol{\Delta_{\gamma}}^T \boldsymbol{E(\Gamma_t)} \tag{3.10}$$

There are two components of the BAF alpha, $(\Delta_{\beta^*} - \Delta_{\beta})E(F_t)$ and $\boldsymbol{\Delta_{\gamma}}^T \boldsymbol{E(\Gamma_t)}$. The former originates from the deviation of the estimated FMP beta from the true FMP beta, and the latter is the compensation for the BAF portfolio's extra exposures to omitted factors. The following analyzes how these two components contribute to a BAF alpha.[7]

As $\alpha_i^* + \epsilon_{it}^*$ is not correlated with factor exposures, sorting stocks into β_i quantiles does not pick not cross-sectional difference in $\alpha_i^*+\epsilon_{it}^*$; hence, $\sum_{\{i:\ \beta_i<\beta_L\}} w_i(\alpha_i^*+\epsilon_{it}^*) \approx \sum_{\{i:\ \beta_i>\beta_H\}} w_i(\alpha_i^*+\epsilon_{it}^*)$ holds. Given Equation (3.7), we have

$$\begin{aligned}
BAF_t = &\sum_{\{i:\ \beta_i<\beta_L\}} w_i[\beta_i^* F_t + \boldsymbol{\gamma}_i^T(\boldsymbol{\alpha_\Gamma} + \boldsymbol{\rho} F_t + \boldsymbol{e_t})] - \sum_{\{i:\ \beta_i>\beta_H\}} w_i[\beta_i^* F_t + \boldsymbol{\gamma}_i^T(\boldsymbol{\alpha_\Gamma} + \boldsymbol{\rho} F_t + \boldsymbol{e_t})] \\
=&\Big(\sum_{\{i:\ \beta_i<\beta_L\}} w_i \boldsymbol{\gamma}_i^T - \sum_{\{i:\ \beta_i>\beta_H\}} w_i \boldsymbol{\gamma}_i^T\Big)\boldsymbol{\alpha_\Gamma} \\
&+ \Big[\sum_{\{i:\ \beta_i<\beta_L\}} w_i(\beta_i^* + \boldsymbol{\rho}^T \boldsymbol{\gamma}_i) - \sum_{\{i:\ \beta_i>\beta_H\}} w_i(\beta_i^* + \boldsymbol{\rho}^T \boldsymbol{\gamma}_i)\Big] F_t \\
&+ \Big(\sum_{\{i:\ \beta_i<\beta_L\}} w_i \boldsymbol{\gamma}_i^T - \sum_{\{i:\ \beta_i>\beta_H\}} w_i \boldsymbol{\gamma}_i^T\Big)\boldsymbol{e}_t
\end{aligned}$$

Therefore, we have Equation (3.9).

[7]To get Equation (3.10), I start with $\alpha_{BAF} = \boldsymbol{\Delta_\gamma}^T \boldsymbol{\alpha_\Gamma}$. Replacing $\boldsymbol{\alpha_\Gamma}$ with Equation (3.5) leads to $\alpha_{BAF} = \boldsymbol{\Delta_\gamma}^T[E(\boldsymbol{\Gamma_t}) - \boldsymbol{\rho} E(F_t)] = \boldsymbol{\Delta_\gamma}^T E(\boldsymbol{\Gamma_t}) - \boldsymbol{\Delta_\gamma}^T \boldsymbol{\rho} E(F_t)$. Finally, replacing $\boldsymbol{\Delta_\gamma}^T \boldsymbol{\rho}$ with Equation (3.8) leads to Equation (3.10).

Beta deviation: $(\Delta_{\beta^*} - \Delta_\beta)E(F_t)$

Due to the existence of omitted factors correlated with the FMP, the estimated FMP betas deviate from the true FMP betas, which induces a negative factor beta-alpha relationship. We can analyze the signs and magnitudes of Δ_β and Δ_{β^*} to understand the beta-deviation component:

Δ_β and Δ_{β^*} are negative: Sorting stocks into quantiles according to estimated FMP betas (β_i) tends to pick up stocks with high/low true FMP betas (i.e., high/low β^*) in the high/low-β side. As a result, the BAF portfolio's true exposure to the FMP is negative, that is, $\Delta_{\beta^*} < 0$. Note that there may be a decoupling between the sign of Δ_β and Δ_{β^*} when the number of stocks used for portfolio construction is small. However, the decoupling will not happen if a large number of stocks are included in the BAF portfolio, as discussed in Chapter 2. The true exposure to the FMP induces a negative return, $\Delta_{\beta^*}E(F_t) < 0$. The negative exposure to the FMP, $\Delta_\beta E(F_t)$, is identified as the explained component of the BAF return in the spanning regression of BAF_t on F_t. Whether a BAF alpha is positive depends on the relative magnitude of Δ_β and Δ_{β^*}.

$|\Delta_\beta|$ is larger than $|\Delta_{\beta^*}|$: According to the connection among factor loadings (Equation 3.7), when ρ is positive, sorting stocks into quantiles according to estimated betas (β_i) tends to pick up stocks with high/low omitted-factor loadings (high/low γ_i) in the high/low-β side, resulting in a negative Δ_γ. Similarly, Δ_γ is positive when ρ is negative. This opposite relationship in signs holds as long as the BAF portfolio is constructed using a large number of. The opposite signs of ρ and Δ_γ indicate that the magnitude of Δ_{β^*} is always smaller than that of Δ_β (see

Equation 3.8). In other words, the magnitude of the explained component of the BAF return ($|\Delta_\beta E(F_t)|$) is larger than the magnitude of the return picked up by the BAF portfolio ($|\Delta_{\beta^*} E(F_t)|$), whose difference is identified as a positive BAF alpha in the spanning regression of BAF_t on F_t.

Extra premium: $\boldsymbol{\Delta}_\gamma{}^T E(\boldsymbol{\Gamma}_t)$

Extra exposures to factors omitted by the FMP may carry extra premiums. To understand how omitted priced factors contribute to a BAF alpha, we can consider the single omitted-factor case of Equation (3.7)

$$\beta_i = \beta_i^* + \rho\gamma_i \tag{3.11}$$

When there exists an omitted factor negatively correlated with F_t ($\rho < 0$), sorting stocks into β quaniles tends to pick up stocks with high omitted-factor loadings in the low-β side and low omitted-factor loadings in the high-β side, that is, $\sum_{\{i:\ \beta_i<\beta_L\}} w_i\gamma_i > \sum_{\{i:\ \beta_i>\beta_H\}} w_i\gamma_i$. Consequently, the BAF portfolio is positively exposed to the omitted factor (Γ_t). This extra exposure is compensated if the omitted factor has a positive premium, which is unexplained by F_t as the omitted factor captures return sources beyond F_t. The extra-premium component of the BAF alpha is thus positive, that is, $\Delta_\gamma^T E(\Gamma_t) > 0$.

The pervasiveness of low factor-beta anomalies

Having clarified how beta deviation and extra premium contribute to a BAF alpha, it is clear why low factor-beta anomalies (i.e., negative cross-sectional factor beta-alpha relationships) should be pervasive. The rationale is as follows:

When an omitted factor correlated with the FMP is priced (say, $E(\Gamma_t) > 0$), the omitted factor loading picked up by the BAF portfolio, Δ_γ, will be associated with an extra premium. The BAF portfolio is negatively exposed to the omitted factor if the omitted factor is positively correlated with F_t, which drives down the BAF alpha since $\Delta_\gamma E(\Gamma)$ is negative in this case. If the impact exceeds the impact from beta deviation, the BAF alpha may be negative. However, given that a positive BAF alpha is contributed by both beta deviation and extra premium ($(\Delta_{\beta^*} - \Delta_\beta)E(F_t) > 0$ and $\boldsymbol{\Delta_\gamma}^T E(\boldsymbol{\Gamma_t}) > 0$), while a negative BAF alpha is contributed by extra premium but mitigated by beta deviation ($(\Delta_{\beta^*} - \Delta_\beta)E(F_t) > 0$ and $\boldsymbol{\Delta_\gamma}^T E(\boldsymbol{\Gamma_t}) < 0$), the collective impact tends to be a positive BAF alpha.

The same conclusion applies when omitted priced factors have negative risk premiums. Therefore, positive BAF alphas should be widely observed when betting against betas of priced factors with positive premiums. Note that for this conclusion to hold, a key is that a large number of stocks should be included in the BAF portfolio construction, which is the precondition for the tendency regulated by Equation (3.7) to become a definitive relationship.

Analytical Conclusion 2: When betting against betas of priced factors with positive premiums, low factor-beta anomalies (i.e., negative cross-sectional factor beta-alpha correlations) should be pervasive, although high factor-beta anomalies may emerge occasionally.

3.3.3 The unpriced- and priced-factor channels of the low factor-beta anomaly

I further analyze how unpriced and priced factors induce positive BAF alphas through beta deviation, $(\Delta_{\beta^*} - \Delta_{\beta})E(F_t)$, and extra premium, $\boldsymbol{\Delta_{\gamma}}^T \boldsymbol{E(\Gamma_t)}$.[8] As shown by Equations (3.7) and (3.8), any omitted factors correlated with the FMP from which betas are estimated will be captured by estimated FMP betas and hence cause an estimated FMP beta to deviate from the true FMP beta it is supposed to capture. As the beta-deviation component always tilts the factor beta-alpha relationship toward positive and any factors correlated with the FMP will cause beta deviation, both omitted unpriced and priced factors drive up a BAF alpha through beta deviation. Besides beta deviation, omitted priced factors also contribute to a BAF alpha through extra factor exposures, which positively (negatively) contribute to a BAF alpha if omitted priced factors with positive premiums are negatively (positively) correlated with the FMP. Following the same rationale for low factor-beta anomalies to be pervasive, as discussed in the previous subsection, omitted priced factors tend to induce a low factor-beta anomaly, although they can also induce a negative BAF alpha.

Analytical Conclusion 3: Omitted unpriced factors, which contribute to a BAF alpha through beta deviation, always drive up a BAF alpha, while omitted priced factors, which contribute to a BAF alpha through both beta deviation and extra premium, also tend to drive up a BAF alpha.

[8] A factor-model framework can also reconcile previous studies relating beta anomalies to unpriced risks. The rationale invoked in previous studies that use betting-against-factor-beta (BAF) portfolios for cost-free hedges is that a BAF portfolio captures the unpriced component of an FMP (e.g., Pukthuanthong et al. 2019; Daniel et al. 2020b). The unpriced component will induce a non-zero correlation between an FMP and omitted unpriced factors.

3.3.4 Strength of a factor-beta anomaly

This subsection analyzes the strength of a BAF alpha reconcilable by the unpriced or priced factor channel. As the weighting scheme of portfolio construction affects the magnitude of a BAF alpha, I define the strength of a factor-beta anomaly as the BAF alpha divided by the magnitude of the explained component of the BAF return:

$$m = \frac{\alpha_{BAF}}{|b_{BAF}E(F_t)|} = \frac{\Delta_{\beta^*} - \Delta_\beta}{-\Delta_\beta} + \frac{\boldsymbol{\Delta_\gamma}^T \boldsymbol{E(\Gamma_t)}}{-\Delta_\beta E(F_t)} \tag{3.12}$$

m reflects the strength of a factor-beta anomaly. For example, $m = 50\%$ refers to the BAF alpha's magnitude being half that of the explained component of the BAF return.

As shown in Equation (3.8), conditional on the properties of the FMP and relevant omitted factors, we can infer Δ_{β^*} and $\boldsymbol{\Delta_\gamma}$ from Δ_β. Specifically, $\frac{\Delta_{\beta^*}-\Delta_\beta}{-\Delta_\beta}$ and $\frac{\boldsymbol{\Delta_\gamma}^T}{-\Delta_\beta}$, which reflect beta deviation, is entirely determined by the connections between the FMP and omitted factors and their ability to capture the time variation of asset returns (see Appendix C1.1). Hereafter, I replace these factor loadings with the properties of relevant factors to analyze the strength of a factor-beta anomaly.

The unpriced-factor channel

Unpriced factors have zero premiums, that is, $\boldsymbol{E(\Gamma_t)} = \mathbf{0}$. In other words, unpriced factors induce low factor-beta anomalies through beta deviation; hence, the BAF alpha from this channel is $\alpha_{BAF} = (\Delta_{\beta^*} - \Delta_\beta)E(F_t)$. By replacing the BAF portfolio's factors loadings with the properties of relevant factors (see Appendix

C1.2), Equation (3.12) becomes

$$m = \frac{\Delta_{\beta^*} - \Delta_\beta}{-\Delta_\beta} = \frac{1 - \frac{Var(F_t^*)}{Var(F_t)}}{\left[1 - \frac{Var(F_t^*)}{Var(F_t)}\right] + \frac{Var_{CS}(\beta_i^*)}{Var_{CS}(\gamma_i)}} \tag{3.13}$$

where F_t^* refers to the priced component of F_t (i.e., $F_t = F_t^* + \Theta_t$; Θ_t is the unpriced component). $1 - \frac{Var(F_t^*)}{Var(F_t)}$ reflects the inefficiency of the FMP, that is, to what extent the unpriced component drives the time variation of F_t. For example, $1 - \frac{Var(F_t^*)}{Var(F_t)} = 40\%$ refers to that the embedded unpriced component accounts for 40% of the FMP's time variation. Var_{CS} is the cross-sectional variance of individual assets' loadings on a factor; $\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)}$ captures the relative importance of the FMP and an average unpriced factor in driving the time-variation of asset returns. The correlation coefficients between the FMP and omitted unpriced factors and the number of relevant omitted unpriced factors do not appear in Equation (3.13) as they are all captured by $1 - \frac{Var(F_t^*)}{Var(F_t)}$.

Based on Equation (3.13), I analyze why a low factor-beta anomaly induced by unpriced factors should be weak.

The magnitude of $1 - \frac{Var(F_t^*)}{Var(F_t)}$ should not be high. Indeed, an FMP is an imperfect proxy of the true underlying factor (F_t^*) and may embed an unpriced component. However, as shown in previous studies such as Fama and French (2015) and Hou et al. (2015), empirical factor models indeed explain a large portion of cross-sectional differences in returns and capture underlying return sources well. This line of empirical evidence implies that, although imperfect, these proxies are not inefficient in capturing the underlying true priced factors. Therefore, we regard the condition that an FMP constructed under reasonable procedures is overall informative as a reasonable condition.

$\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)}$ **should have a small magnitude.** As discussed in Appendix C1.3, $\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)}$ captures the relative importance of two factors in driving the time variation of asset returns. The value of $\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)}$ should be small conditional on the definition of unpriced factors. According to the literature, the unpriced factor captures return comovement but has a zero risk premium. This conception traces back to the imaginary cosmetics factor in Roll and Ross (1984). Roll and Ross argue that although capturing return comovement, the factor is unpriced since it is not pervasive and investors can diversify across industries. Later studies related to unpriced factors follow this conception (e.g., Daniel and Titman 1997; Asness et al. 2000; Pukthuanthong et al. 2019; Daniel et al. 2020b; Clarke 2022). These studies suggest that firm characteristics also capture comovement related to unpriced risks; hence, characteristic-sorted FMPs contain an unpriced-risk component. According to the conception, the rationale for the unpriced-risk factor, which captures return comovement, to have a zero premium is that it is avoidable. The fact that an unpriced factor is avoidable indicates that only a small set of assets are exposed to it while the remaining are not. By contrast, a priced factor should drive the time variation of most assets' returns. As such, the magnitude of $Var_{CS}(\gamma_i)$ should be smaller than that of $Var_{CS}(\beta_i^*)$.

An upper bound for a BAF alpha under the unpriced-factor channel While a low factor-beta anomaly induced purely by unpriced factors should be weak, it is empirically unclear what the value of m should be. However, an upper bound exists for a BAF alpha that the unpriced-factor channel can reconcile. To clarify this upper bound, I force the BAF alpha to be no less than the magnitude of the explained component of the BAF return, that is, forcing m in Equation (3.13) to be less than one:

$$\Delta_{\beta^*} - \Delta_{\beta} \geq -\Delta_{\beta} \tag{3.14}$$

this inequality leads to an unreasonable condition immediately, that is,

$$|\Delta_{\beta^*}| \leq 0 \tag{3.15}$$

For the unpriced-factor channel to reconcile the case of $m = 1$, estimated FMP betas should entirely fail in capturing what they are supposed to capture so that $|\Delta_{\beta^*}| = 0$ holds. However, according to Equation (3.7), estimated FMP betas (β) capture true FMP betas (β^*) as long as the FMP is not pure noise. Therefore, the condition $|\Delta_{\beta^*}| = 0$ is unreasonable. Moreover, in the case that a BAF alpha is larger than the explained component of the BAF return ($\alpha_{BAF} > |\Delta_{\beta} E(F_t)|$), $|\Delta_{\beta^*}|$ needs to be negative. This scenario is impossible since estimated FMP betas should contain the information of true FMP betas as long as the FMP is not pure noise. A direct message is that a BAF alpha exceeding this upper bound indicates the existence of omitted priced factors.

The Priced-Factor Channel

Besides driving up a BAF alpha through beta deviation, the priced-factor channel also affects a BAF alpha through extra premiums. The strength of a low factor-beta anomaly from this channel depends not only on the extent to which omitted priced factors can cause the estimated FMP beta to deviate from the true FMP beta but also on the risk-return profiles of omitted priced factors. A low factor-beta anomaly depends on how omitted factors affect the FMP, and hence the interaction of omitted factors does not matter. Given this precondition, I made several assumptions that simplify the analysis without essentially affecting the analytical framework: all omitted factors are uncorrelated with each other, have the same correlation with the FMP, have the same variance and expected return, and are equally important in driving the time variation of asset returns. These simplified assumptions enable a further simplification of Equation (3.12):

$$m = \frac{\Delta_{\beta^*} - \Delta_\beta}{-\Delta_\beta} + K\frac{\Delta_\gamma E(\Gamma_t)}{-\Delta_\beta E(F_t)} \tag{3.16}$$

where K is the number of omitted factors, Δ_γ is the BAF portfolio's exposure to an average omitted factor, and $E(\Gamma_t)$ is the expected return of an average omitted factor. By replacing the BAF portfolio's factors loadings with the properties of relevant factors (see Appendix C1.4), m is expressed as:

$$m = \frac{K\rho^2}{K\rho^2 + \frac{Var_{CS}(\beta_i^*)}{Var_{CS}(\gamma_i)}}(1 - \frac{1}{\rho}\frac{SR_\Gamma}{SR_F}) \tag{3.17}$$

where $\frac{SR_\Gamma}{SR_F}$ is the relative Sharpe ratio of factors F_t and Γ_t, which is the same as $\frac{E(\Gamma)}{E(F_t)}$) under the assumption that factor variances are the same. As discussed in the previous subsection, $\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)}$ reflects the relative ability of the two factors to capture the time variation of asset returns. In the context of priced factors, $\frac{SR_\Gamma}{SR_F}$ and $\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)}$ together also reflects the ability of the two factors to capture cross-sectional differences in returns.

No bound for a BAF alpha under the priced-factor channel Intuitively, depending on the properties of the relevant omitted priced factors, the priced-factor channel can generate a strong BAF alpha with no upper bound. I demonstrate the strength of a BAF alpha from this channel formally by forcing Equation (3.17) to be no less than one, $m \geq 1$, which leads to

$$-K\rho\frac{Var_{CS}(\gamma_i)E(\Gamma_t)}{Var_{CS}(\beta_i^*)E(F_t)} \geq 1 \tag{3.18}$$

The strength of the BAF alpha induced by an omitted priced factor depends on the relative importance of the omitted factor and FMP in capturing the cross-sectional differences in returns ($\frac{Var_{CS}(\gamma_i)E(\Gamma_t)}{Var_{CS}(\beta_i^*)E(F_t)}$) and the connection between the two factors (ρ). When an omitted factor strongly correlates with the FMP, the estimated FMP beta capture more information about the loading on the omitted factor. When the omitted factor explains more cross-sectional differences in returns, sorting on estimated betas tends to pick up a higher spread in loadings on the omitted factor, which is associated with a higher premium. Moreover, in sharp contrast to the unpriced-factor channel under which the number of omitted factors related to the FMP does not matter, the number of omitted factors matters under the priced-factor channel. When there are more omitted factors (K is larger), each omitted factor does not need to strongly correlate with the FMP to induce a strong BAF alpha. The rationale is that omitted priced factors correlated with an FMP increase the inefficiency of the FMP through ρ^2 but induce extra premiums through ρ. Given that the magnitude of ρ is always lower than one, more omitted factors make it easier for the priced-factor channel to come into effect.

Analytical Conclusion 4: Omitted priced factors should dominate unpriced factors in driving factor-beta anomalies, and there is an upper bound for a factor-beta anomaly that unpriced factors can reconcile. When a low factor-beta anomaly with a strength of $m \geq 1$ emerges, there must exist omitted priced factors contributing to it.

3.3.5 Numerical Examples

In the last of this section, I provide numerical examples to explain the aforementioned analytical conclusions further. These numerical examples also illustrate how to infer information from an observed factor-beta anomaly based on our ex-ante understanding of the properties of unpriced and priced factors. I also provide simulation examples in Appendix C2 to explain intuitively how the analytical framework works.

The low factor-beta anomaly from the unpriced-factor channel

Figure 3.1 plots Equation (3.13) , that is, the values of $\frac{Var(F_t^*)}{Var(F_t)}$ and $\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)}$ when m taking different values. The dashed horizontal line corresponds to $\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)} = 1$. The upper-left subfigure depicts the scenario close to a typical low factor-beta anomaly; that is, a BAF alpha's magnitude is as large as the explained component of the BAF return ($\alpha_{BAF} = -0.99\Delta_\beta E(F_t)$). According to this subfigure, for example, to make $m = 1$ hold, $\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)}$ needs to be very high when $\frac{Var(F_t^*)}{Var(F_t)}$ equals 90%. This condition means that when the unpriced component accounts for 10% of the FMP's time variation, the time variation of asset returns explained by the unpriced factor should be nearly 1000 times that explained by the true priced factor underlying the FMP to make this channel come into effect. While $\frac{Var(F_t^*)}{Var(F_t)} = 90\%$ is reasonable for an FMP, $\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)} > 1000$ is not expected for an unpriced factor. On the other hand, when $\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)}$ takes a small value, $\frac{Var(F_t^*)}{Var(F_t)}$ has to be close to zero, indicating a completely uninformative FMP. Similar observations hold for weaker low factor-beta anomalies. Even when the BAF alpha is only 30% of the explained component of the BAF return ($\alpha_{BAF} = -0.3\Delta_\beta E(F_t)$), unpriced factors need to overwhelm the true priced factor underlying the FMP in driving the time variation of asset returns (large $\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)}$) if the FMP is informative (large $\frac{Var(F_t^*)}{Var(F_t)}$).

Next, I plot Equation (3.13) in another format to analyze the level of the BAF alpha this channel can reconcile. The middle of Figure 3.2 depicts the scenario when the omitted unpriced factor is equally important as the true priced factor underlying the FMP in driving the time variation of asset returns. To generate a BAF alpha with $m = 30\%$, 40% of the variance of the FMP should be driven by its embedded unpriced component. To generate a BAF alpha with $m = 50\%$, The variance of the FMP should be entirely driven by its embedded unpriced component. The pattern indicates again that to generate a strong BAF alpha, the FMP from which betas are estimated should be inefficient. Therefore, although the unpriced-factor channel can induce a negative beta-alpha relationship, its contribution to a BAF alpha should be weak.

The priced-factor channel

Figure 3.3 plots Equation (3.17) to analyze 1) how the strength of the BAF alpha (capture by m) changes with $\frac{SR_\Gamma}{SR_F}$ when the average omitted priced factor and FMP (F_t) are equally important in driving the time variation of asset returns ($\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)} = 1$), and 2) how m changes with $\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)}$ when the average omitted priced factor have the same Sharpe ratio as the FMP. I also consider the scenario

Figure 3.1: Low Factor-Beta Anomalies Reconcilable by Unpriced Factors

The figure depicts the values of two parameters ($\frac{Var_{CS}(\gamma_i)}{Var_{VS}(\beta_i^*)}$ and $\frac{Var(F_t^*)}{Var(F_t)}$) for the unpriced-factor channel to generate a low factor-beta anomaly with the strength of $m = \frac{\alpha_{BAF}}{|\Delta_\beta E(F_t)|}$. For example, $m = 90\%$ refers to that in the spanning regression of BAF_t on F_t, the BAF alpha is 90% of the magnitude of the explained component of the BAF return. As discussed in Section 3.3.4, $\frac{Var(F_t^*)}{Var(F_t)}$ reflects the percent of a factor-mimicking portfolio's variance driven by its priced component. $\frac{Var_{CS}(\gamma_i)}{Var_{VS}(\beta_i^*)}$ reflects the relative importance of an average omitted unpriced factor and F_t in driving the time variation of asset returns. The dashed line in each subfigure corresponds to $\frac{Var_{CS}(\gamma_i)}{Var_{VS}(\beta_i^*)} = 1$.

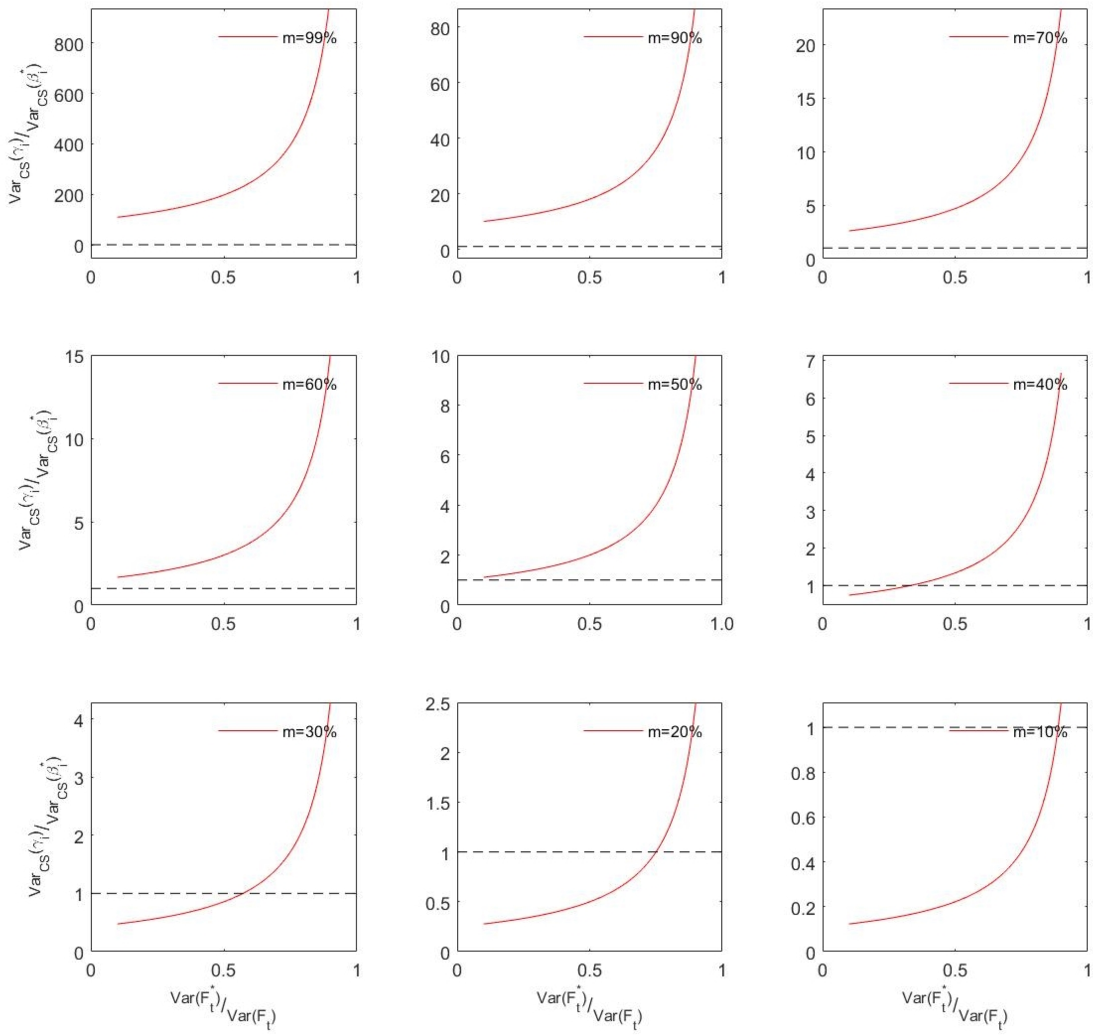
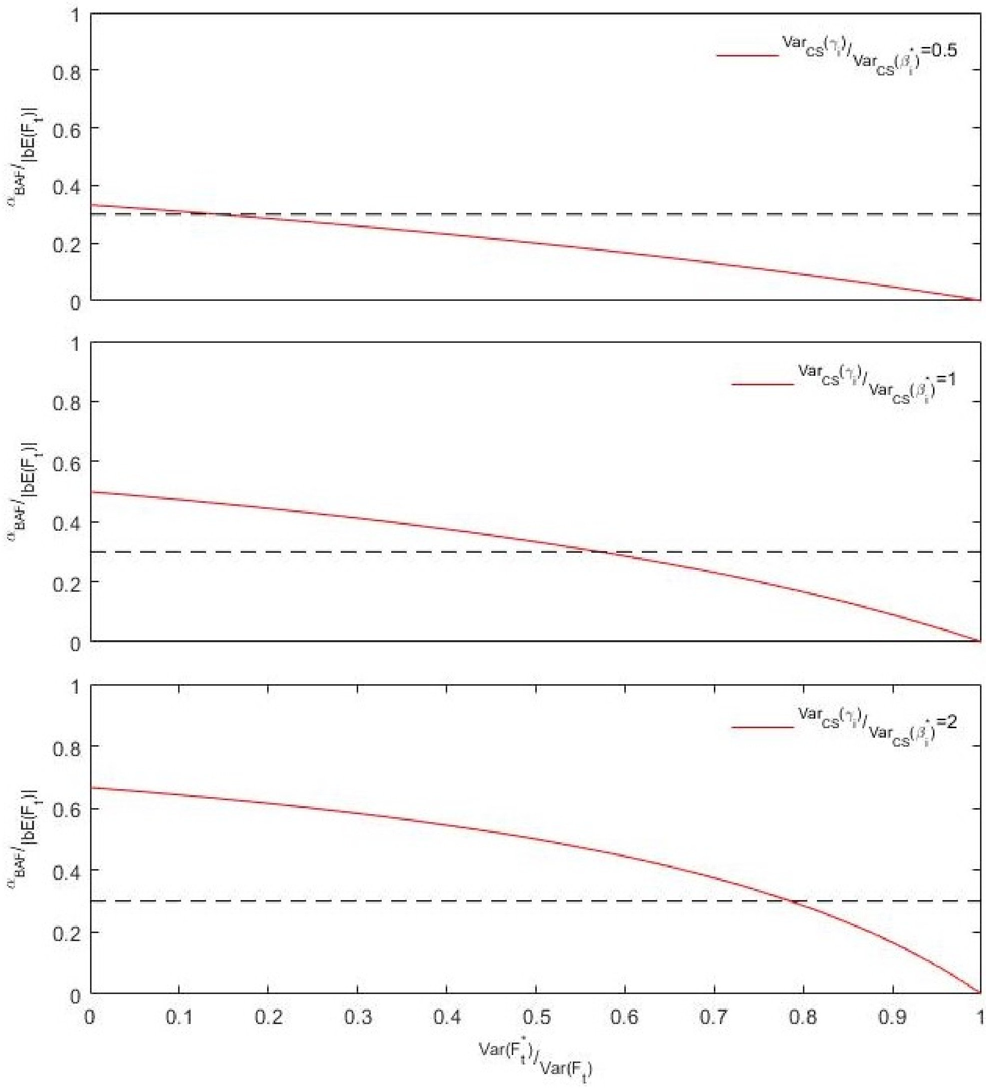


FIGURE 3.2: Low Factor-Beta Anomalies Reconcilable by Unpriced Factors
The figure depicts the strength of a BAF alpha, denoted by $m = \frac{\alpha_{BAF}}{|\Delta_\beta E(F_t)|}$, under different values of two parameters ($\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)}$ and $\frac{Var(F_t^*)}{Var(F_t)}$). As discussed in Section 3.3.4, $\frac{Var(F_t^*)}{Var(F_t)}$ reflects the percent of a factor-mimicking portfolio's variance driven by its priced component. $\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)}$ reflects the relative importance of the omitted factor and F_t in driving the time variation of asset returns. The dashed line in each subfigure corresponds to $m = 30\%$.



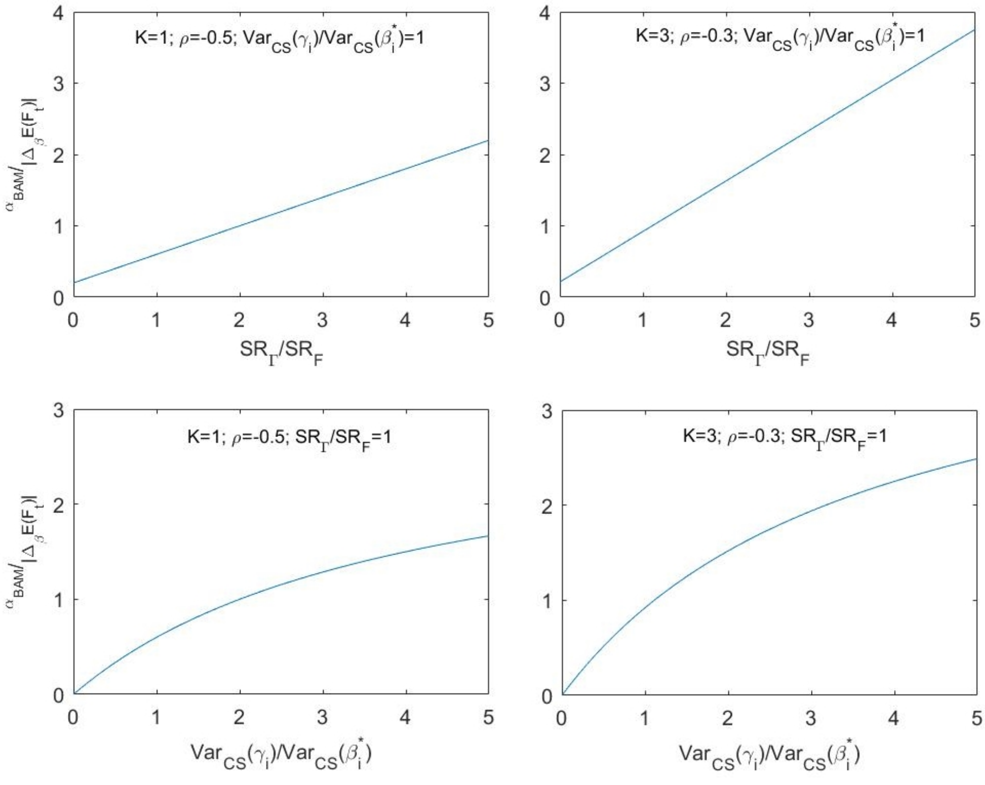
of several omitted factors.

The upper two subfigures depict the impact of the relative Sharpe ratio when factors are equally important in driving the time variation of asset returns. The upper-left subfigure depicts the scenario of one omitted factor negatively correlated with the FMP ($K = 1$, $\rho = -0.5$). In this scenario, the strength of the BAF alpha increases with the Sharpe ratio of the omitted priced factor. To generate a low factor-beta anomaly whose strength is $m = 1$, the omitted priced factor should have a larger Sharpe ratio than F_t. The upper-right subfigure depicts the scenario of three omitted factors, each with a weaker correlation with F_t ($\rho = -0.3$). When more omitted factors are negatively correlated with F_t, it is easier for the priced-factor channel to come into effect. A low factor-beta anomaly whose strength is $m = 1$ emerges when each omitted has a similar Sharpe ratio to F_t.

The bottom two subfigures depict the impact of $\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)}$ when factor Sharpe ratios are the same. The BAF alpha is larger when the cross-sectional variance of loadings on the omitted factor, $Var_{CS}(\gamma_i)$, is larger. When there is only one omitted factor negatively correlated with F_t, the omitted factor should be more important in driving the time variation of asset returns to induce a typical low factor-beta anomaly. Similarly, omitted factors do not need to be more important than F_t in driving the time variation of asset returns when there are several omitted factors negatively correlated with F_t, as indicated by the bottom-right subfigure. The result that we can easily find values of the four parameters (K, $\frac{SR_\Gamma}{SR_F}$, $\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)}$ and ρ) to make $m = 1$ hold indicates that the priced-factor channel can easily generate strong low factor-beta anomalies under realistic conditions.

FIGURE 3.3: Low Factor-Beta Anomalies Reconcilable by Priced Factors

The figure depicts the strength of the low factor-beta anomaly (captured by $m = \frac{\alpha_{BAF}}{|\Delta_\beta E(F_t)|}$) induced by priced factors under different values of K, ρ, $\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)}$ and $\frac{SR_\Gamma}{SR_F}$. K is the number of omitted factors (Γ_t) with a negative correlation ($\rho < 0$) with F_t. $\frac{SR_\Gamma}{SR_F}$ reflects the relative Sharpe ratio of the omitted factor and F_t, and $\frac{Var_{CS}(\gamma_i)}{Var_{VS}(\beta_i^*)}$ reflects their relative importance in driving the time variation of asset returns. The correlation between each omitted factor and F_t is assumed to be the same.



3.4 Empirical Evidence

This section provides empirical support for the aforementioned analytical conclusions. As the proxies for unpriced and priced factors that fully explain asset returns are unclear, I avoid drawing conclusions from any specific BAF performance but analyze whether the performance of a large set of betting-against-factor-beta (BAF) portfolios is consistent with the multifactor-model framework. Unless mentioned, this study uses a single-factor model to estimate factor betas empirically. The benefit is that conditional on the fact that the six-factor model of Fama and French (2018) well captures asset returns, the omitted factors relative to any single factor are well known, which thus enables a better assessment of the rationality of the analytical framework is reasonable.

3.4.1 Data

I collect individual stock return and market capitalization data from CRSP. Only common stocks listed on NYSE, AMEX and NASDAQ with share codes 10 or 11 are included (ADRs, ETFs, and REITs are excluded). I adjust raw returns for delisting following Shumway (1997). Factor, anomaly and industry return data are obtained from Kenneth French's and Lu Zhang's websites. The sample period differs due to data availability or initial samples used for beta estimation. The full sample period spans from July 1963 to December 2019, and data from the first five years are used for the initial beta estimation.

I estimate betas of individual stocks using monthly returns of the previous sixty months at each time t, with a minimum of twenty-four observations required for a calculated beta to be valid. For brevity, I only report results based on betas estimated from monthly returns of the previous sixty months. To avoid small/microcap biases, I construct size-balanced BAF portfolios. At the end of each month, I double-sort stocks into two-by-three size-beta quantiles independently according to NYSE breakpoints (50th percentile for size, 30th and 70th percentiles for beta). Size is the market capitalization at the end of the previous June. The long side return in Equation (3.2), for example, is calculated as $\sum_{\{i:\ \beta_i<\beta_L\}} w_i r_{it} = \frac{1}{2}(r_t^{Small,\ Low\ \beta} + r_t^{Big,\ Low\ \beta})$, where $r_t^{Small,\ Low\ \beta}$ is the value-weighted return of the stocks in the small size and low factor-beta portfolio.[9]

3.4.2 The mechanism for low factor-beta anomalies to emerge

The first piece of empirical evidence is that positive betting-against-beta alphas is widely observed when betting against priced factors with positive premiums but is not widely observed when betting against betas of unpriced factors.

Betting against betas of priced factors (anomalies)

This subsection examines the BAF performance of a large set of anomaly portfolios obtained from Lu Zhang's website. The original anomaly data are three-by-five double-sorted portfolios (size and the corresponding characteristic), and the return of each portfolio is measured on a value-weighted basis. I calculate each anomaly return as the return spread between the high and low characteristic sides, where the return of each side is calculated as the average of micro, small and big-size portfolios. To include as many anomalies as possible while maintaining a long

[9] In the case of betting against betas of the size factor, I sort stocks into three-by-two BM-beta quantiles independently. The BAF return is calculated as the spread between the average of the three low-beta portfolios and the average of the three high-beta portfolios. The book-to-market (BM) ratio is measured following Fama and French (2015).

enough sample period, I keep those anomalies with data available from February 1967 to December 2019, which leads to 116 anomalies remaining in the sample.

I estimate loadings of individual stocks on an anomaly following Equation (3.1) and then construct size-balanced BAF portfolios under Equation (3.2). Figure 3.4 reports the performance along with the 95% confidence intervals of these anomaly portfolios and their BAF portfolios. The first subfigure reports the average returns (μ_F) of these anomalies, most of which are positive, consistent with previous studies. The second subfigure reports the average returns of BAF portfolios, most suggesting a flat return-beta relationship. The third subfigure reports the spanning regression loading of each BAF portfolio on its corresponding factor/anomaly portfolio. Most loadings (b_{BAF}) are significantly negative with a considerable magnitude, indicating that estimated FMP betas well capture return comovement. Many of these portfolios are interpreted as anomalies in previous studies. The fact that most b_{BAF} are significantly negative indicates that these anomaly portfolios also capture return comovement, which is consistent with the finding of Kozak et al. (2018) that investor sentiments must be aligned with factor betas to be relevant for pricing. In other words, anomaly portfolios should also capture return comovement if they have positive expected returns. Therefore, positive BAF alphas should also widely emerge when betting against betas of anomaly portfolios.

The bottom subfigure reports the BAF alphas. 106 of the 116 BAF alphas are positive. The pervasiveness of positive BAF alphas is consistent with the analytical conclusion that low factor-beta anomalies should be pervasive when betting against betas of priced factors with positive premiums. By observing BAF alphas (α_{BAF}) and BAF portfolios' exposures to FMP ($|b_{BAF}|$), we can find that if scaling α_{BAF}

by $|b_{BAF}|$, most $\alpha_{BAF}/|b_{BAF}|$s are no less than original anomaly returns (μ_F). According to the finding on the upper bound of a factor-beta anomaly from the unpriced-factor channel, this large magnitude of BAF alphas should be largely driven by priced factors omitted by the corresponding FMP.

Betting against betas of industry portfolios

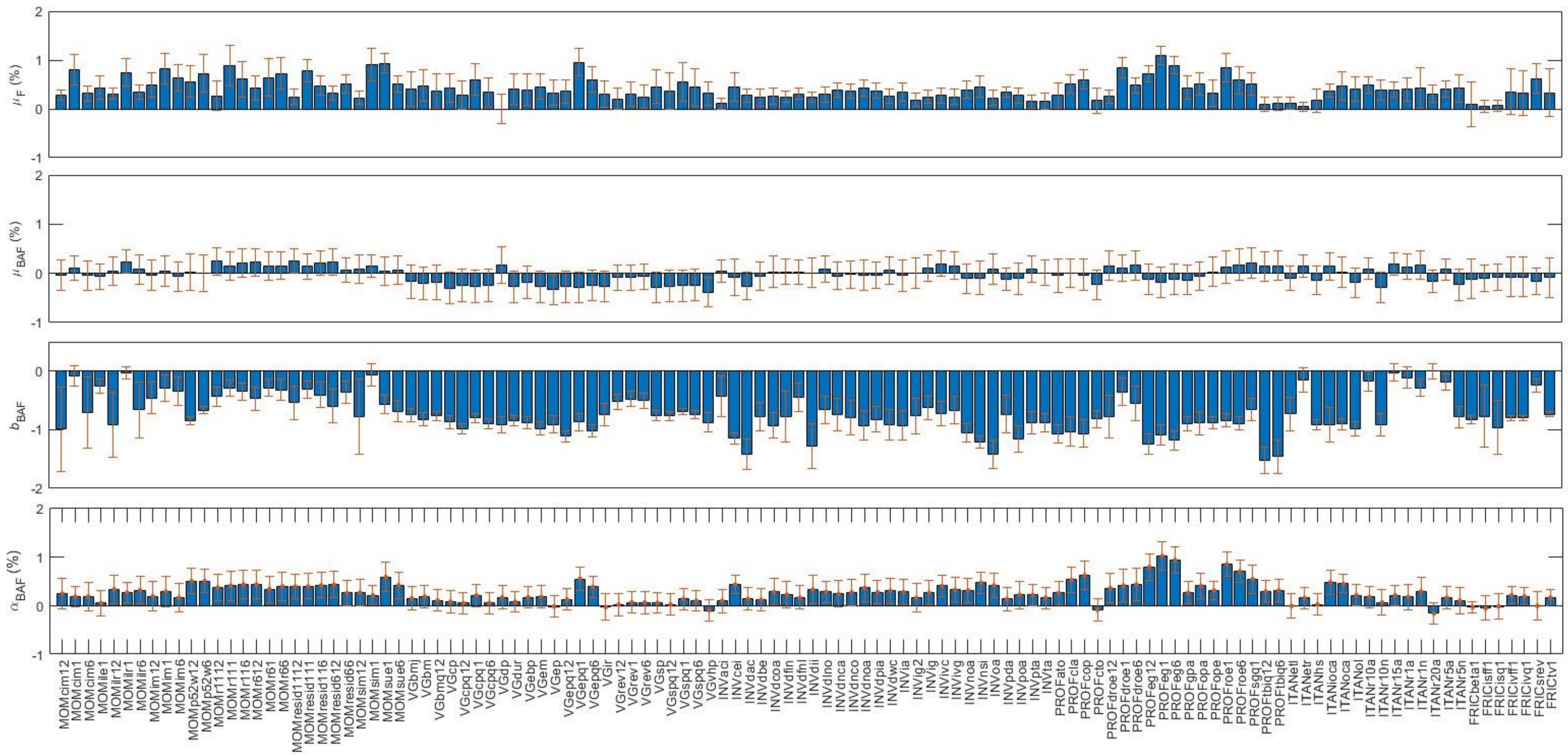
The model in Section 3.3 indicates that for low factor-beta anomalies to be pervasive, an FMP should be a priced factor that correlates with omitted factors. If betting against betas of unpriced factors, strong low-beta anomalies should not emerge, and the sign of BAF alphas should be random. An unpriced factor has zero expected returns, meaning that there is no beta-deviation component to contribute to the BAF alpha when betting against betas of unpriced factors. As a result, the precondition for the positive BAF alpha to be pervasive under the model of this study does not hold. Additionally, an unpriced factor should not relate widely to omitted priced factors; otherwise, the unpriced factor will be priced. Given these priors, I examine the BAF performance on unpriced factors. I use the twelve industry portfolios as proxies for unpriced factors, as the time variation of industry returns is widely regarded as unpriced risks (e.g., Roll and Ross (1984); Daniel and Titman 1997; Asness et al. 2000; Daniel et al. 2020b).[10]

As industry returns are highly correlated with the market return, directly estimating industry betas following Equation (3.1) produces betas close to CAPM betas, and the corresponding BAF portfolios essentially capture the low-CAPM beta anomaly. To better capture exposures to unpriced factors, I estimate industry

[10]The industry classification system is available from Kenneth French's website.

FIGURE 3.4: The Pervasiveness of Low Factor-Beta Anomalies

This figure reports the performance of 116 factors/anomalies (from Hou et al. 2020b) and their corresponding BAF portfolios. I estimate loadings of individual stocks on a factor/anomaly following Equation (3.1) and then construct size-balanced BAF portfolios under Equation (3.2). The first and second subfigures report the average returns. The third subfigure reports the loading of each BAF portfolio on its corresponding factor/anomaly portfolio. The fourth subfigure reports the BAF alphas. The 95% confidence intervals based on Newey and West (1987) adjusted standard errors are also reported. The full sample period is from February 1967 to December 2019, with data from the first five years used for the initial beta estimation.



loadings with the market return controlled:

$$r_{it}^e = \alpha_i + \beta_{MKT,i} MKT_t + \beta_{Ind,i} Ind_{I,t} + \epsilon_{it} \tag{3.19}$$

where r_{it}^e is the excess return of security i; MKT_t is the excess market return; $Ind_{I,t}$ is the excess return of the Ith industry. I then use the estimated industry betas ($\beta_{Ind,i}$) to construct the BAF portfolio for each industry following Equation (3.2).

Figure 3.5 reports the summary statistics along with the 95% confidence intervals of these industry portfolios and their corresponding BAF portfolios. The first subfigure reports these industry portfolios' average excess returns after adjusting for market exposures (μ_F^{eMKT}). After removing their market exposures, most industry returns are indistinguishable from zero, indicating that they mainly capture unpriced risks. Three of the twelve industry returns are significantly (at the 5% level) different from zero, indicating that they are still exposed to some priced risks. As these three industry returns have a small magnitude, their time variation is still largely driven by unpriced risks, and hence the time variations of these industry portfolios mainly reflect unpriced risks. The second subfigure reports the average returns of BAF portfolios, almost indistinguishable from zero. The third subfigure reports the spanning regression loading of each BAF portfolio on its corresponding market-adjusted industry portfolio (Ind_{It}^e), that is,

$$BAF_{It} = \alpha_{BAF,I} + b_{BAF,I} Ind_{It}^e + \epsilon_{It} \tag{3.20}$$

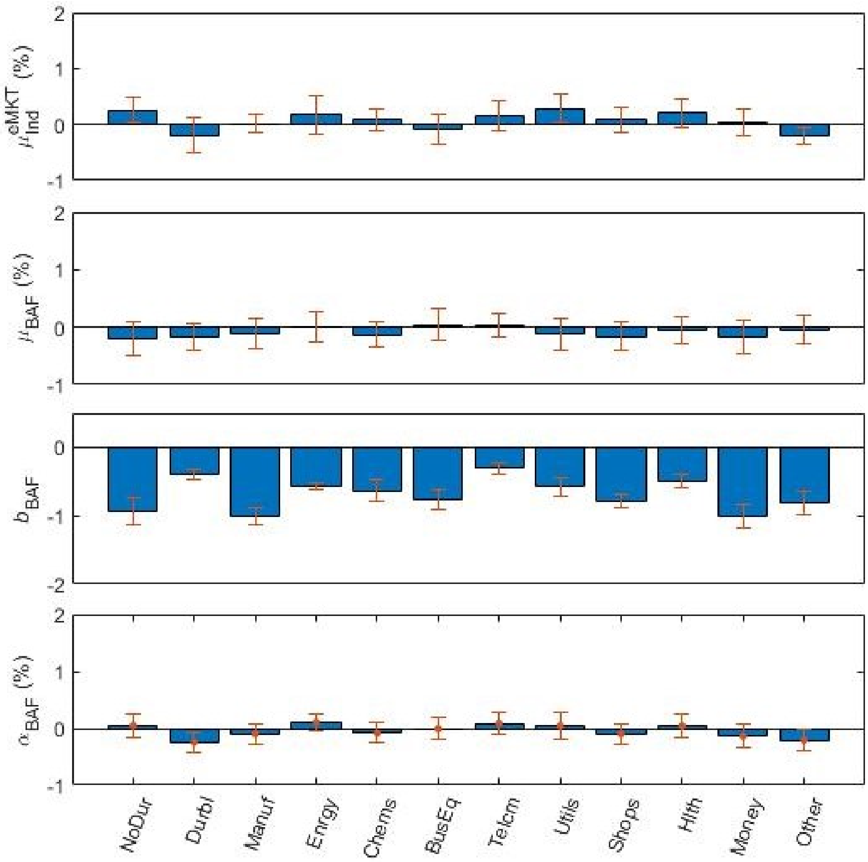
These loadings (b_{BAF}) are significantly negative, indicating that these industry portfolios also capture return comovement. Compared with those of priced factors, the magnitudes of these spanning regression slopes are not higher, indicating that, in general, unpriced factors do not drive more return comovement than priced factors. The bottom subfigure reports the BAF alphas (α_{BAF}). Unlike the pervasiveness of positive BAF alphas when betting against factor/anomaly betas, whether positive or negative BAF alphas emerge is random in the case of betting against industry betas. Moreover, ten of the twelve BAF alphas are indistinguishable from zero. The two significant BAF alphas (durable and other industries) are negative and have a small magnitude. There are two reasons for these BAF alphas to be weak. First, there is no beta-deviation component to contribute to a BAF alpha when betting against betas of unpriced factors. Second, the market-adjusted industry portfolios are not strongly related to omitted priced factors. Therefore, to the extent that industry returns are good proxies of unpriced factors, the result supports the rationality of this study's multi-factor-based explanation of low factor-beta anomalies.

3.4.3 Contribution of priced and unpriced factors to a BAF alpha

This subsection shows that the BAF performance of a large set of factor portfolios when controlling for priced and unpriced factors is consistent with the analytical conclusion that priced factors dominate unpriced factors in contributing to a BAF alpha.

FIGURE 3.5: Betting against Betas of Unpriced Factors

This figure reports the performance of the twelve industry portfolios and their corresponding BAF portfolios. I estimate loadings of individual stocks on each industry portfolio with the market return controlled (using the monthly returns of the prior five years) and then construct size-balanced BAF portfolios following Equation (3.2) using the estimated industry portfolio loadings. The first subfigure reports the market-adjusted industry portfolio returns. The second subfigure reports the average returns of the BAF portfolios. The third subfigure reports the loading of each BAF portfolio on its corresponding industry portfolio in the spanning regression. The fourth subfigure reports the BAF alphas. The 95% confidence intervals based on Newey and West (1987) adjusted standard errors are also reported. The full sample period is from July 1963 to December 2019, with data from the first five years used for the initial beta estimation.



Controlling for unpriced factors

Under the unpriced-factor channel (e.g., Pukthuanthong et al. 2019; Daniel et al. 2020b), a factor-beta anomaly is purely caused by the deviation of the estimated FMP beta from the true FMP beta and hence correcting for beta deviation can resolve the anomaly. We can understand this through a simplified version of the model in Section 3.3. Suppose a priced factor and an unpriced factor (F_t^* and Γ_t) generate asset returns; the empirical proxy (F_t) of the priced factor contains information about the unpriced factor so that they are correlated, $|Cov(F_t, \Gamma_t)| > 0$. In this case, the BAF portfolio constructed from estimated FMP betas generates alphas against F_t, and the alpha is entirely contributed by beta deviation, $\alpha_{BAF} = (\Delta_{\beta^*} - \Delta_{\beta})E(F_t)$. The alpha exists since the estimated exposure (Δ_β) of BAF_t on F_t in a spanning regression is different from the BAF portfolio's true exposure (Δ_{β^*}) on the underlying priced factor (F_t^*). This deviation is corrected when controlling for the corresponding unpriced factor in the spanning regression:

$$BAF_t = \alpha_{BAF}^{Corrected} + bF_t + c\Gamma_t + e_t \tag{3.21}$$

Once the unpriced factor with zero risk premium ($E(\Gamma_t) = 0$) is controlled, the BAF portfolio's exposure to the underlying priced factor is corrected, $b = \frac{Cov(BAF_t, F_t - \rho\Gamma_t)}{Var(F_t - \rho\Gamma_t)} = \Delta_{\beta^*}$, and the BAF alpha becomes

$$\alpha_{BAF}^{Corrected} = E(BAF_t) - bE(F_t) - cE(\Gamma_t) = E(BAF_t) - bE(F_t) = 0 \tag{3.22}$$

that is, the BAF alpha is subsumed. Therefore, under the unpriced-factor channel, a Γ_t with a zero risk premium should subsume the factor-beta anomaly. Equation

(3.22) only holds when a BAF portfolio mainly captures unpriced risks beyond capturing F_t. If omitted priced factors impact the BAF portfolio more than unpriced factors, only eliminating beta deviation cannot subsume the factor-beta anomaly.[11]

To ensure that only the beta-deviation component is corrected, I control for the mean-removed market-adjusted industry returns: I first remove the market exposure of each industry portfolio following the previous subsection and then demean the residual return of each portfolio. If a factor-beta anomaly is mainly caused by unpriced factors, to the extent that industry portfolios capture unpriced risks, such correction should subsume the anomaly. I use the following spanning regression to correct for unpriced factor exposures

$$BAF_t = \alpha_{BAF}^{CrIND} + bF_t + \boldsymbol{c}^T \tilde{\boldsymbol{Ind}}_t^e + e_t \tag{3.23}$$

where $\tilde{\boldsymbol{Ind}}^e$ is the vector of mean-removed market-adjusted industry returns.

The top of Figure 3.6 reports the BAF alphas from Equation (3.23). Compared with the bottom of Figure 3.5, most BAF alphas decrease after controlling for industry exposures, indicating that these FMPs indeed contain unpriced components. The beta deviation caused by unpriced factor exposures is largely remedied when industry returns are controlled, which is why BAF alphas decrease. Some BAF alphas are positive in Figure 3.5 but become negative in the top subfigure of Figure 3.6. This change is consistent with the analytical conclusion that unpriced factors can only positively contribute to a BAF alpha. Unpriced factors' impact is

[11]This approach tends to overstate the importance of unpriced factors as some industry portfolios also capture priced return sources.

removed when they are controlled, and hence the remaining component of a BAF alpha is negative if the impact from omitted priced factors with positive premiums but negatively correlated with an FMP dominates. The result that no BAF alphas become larger in magnitude is also consistent with the analytical framework, as controlling for unpriced factors only remedies beta deviation and hence can only drive down a BAF alpha. Overall, the result is consistent with the finding in the previous section that although omitted unpriced factors can drive up the BAF alpha, their impact should be weak under the general definition of unpriced factors.

Controlling for priced factors

Next, I examine how controlling for omitted priced factors affects a BAF alpha. To the extent that the six factors of Fama and French (2018) explain asset returns, these six factors should largely capture priced factors omitted by each of the factor/anomaly portfolios examined in Figure 3.5. Therefore, BAF alphas should change notably after controlling for the six-factor model:

$$BAF_t = \alpha_{BAF}^{CrFF6} + bF_t + \boldsymbol{\beta}^T \boldsymbol{\Gamma}_t^{FF6} + v_t \tag{3.24}$$

where $\boldsymbol{\Gamma}_t^{FF6} = [MKT_t,\ SMB_t,\ HML_t,\ RMW_t,\ CMA_t,\ MOM_t]^T$ is the vector of the six factors of Fama and French (2018). The middle of Figure 3.6 reports the BAF alphas from Equation (3.24). The change of these BAF alphas is substantial compared with the top subfigure. In sharp contrast to the case of controlling for unpriced factors, many positive BAF alphas become larger in magnitude than those reported in Figure 3.5. According to the analytical framework, such a change occurs when the impact of omitted factors (with positive premiums) positively correlated with an FMP is removed when the six-factor model is controlled, and hence the corresponding BAF alpha becomes larger. On the contrary, when the six factors, on average, are negatively correlated with an FMP, controlling for them will make the impact of omitted factors positively correlated with the FMP dominate, and hence the corresponding BAF alpha becomes negative. Overall, the result indicates that omitted priced factors strongly impact factor-beta anomalies.

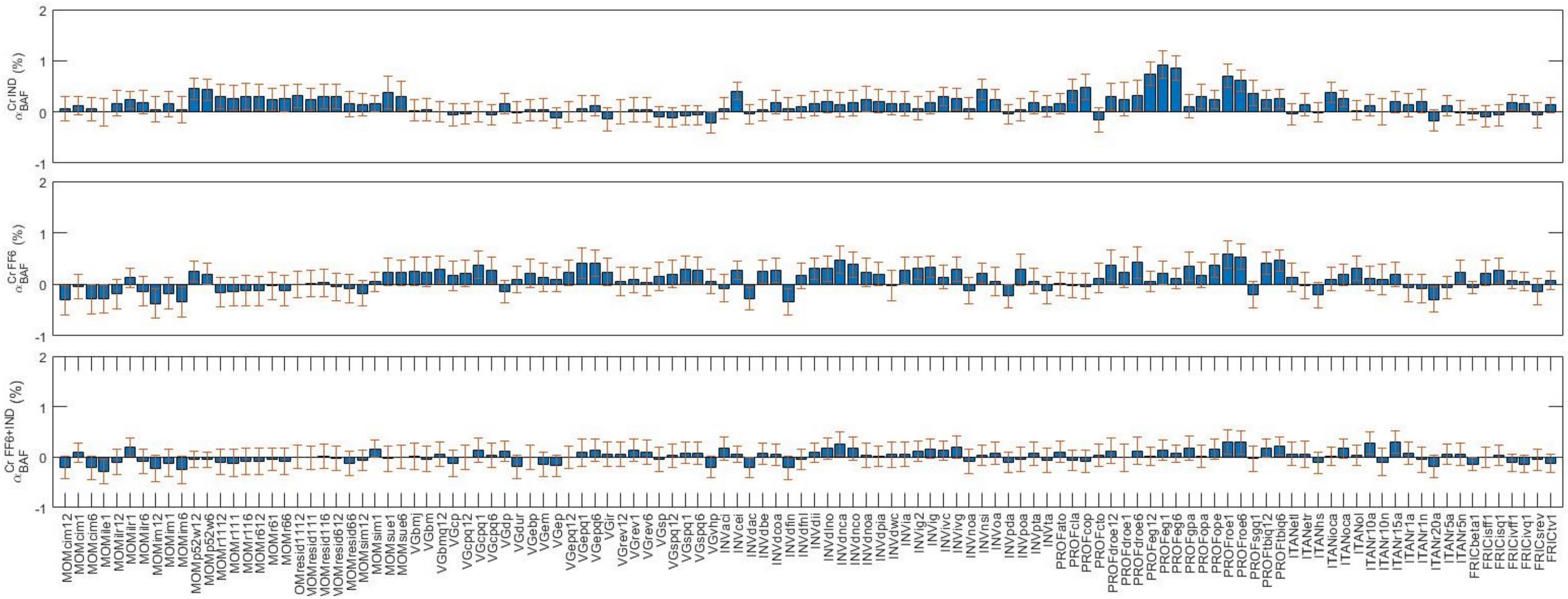
Last, I examine the BAF alphas after controlling for unpriced and priced factors simultaneously:

$$BAF_t = \alpha_{BAF}^{CrFF6+IND} + bF_t + \boldsymbol{\beta}^T \boldsymbol{\Gamma}_t^{FF6} + \boldsymbol{c}^T \tilde{\boldsymbol{Ind}}_t^e + v_t \tag{3.25}$$

As reported by the bottom of Figure 3.6, most BAF alphas become negligible. Two forces drive this result: first, by controlling for priced and unpriced factors at the same time, both the beta-deviation and extra-premium components of a BAF alpha are controlled; second, controlling for the two sets of factors at the same time makes the efficient version of the six factors come into effect, as controlling for variations related to industry returns improves their mean-variance efficiency. It is reasonable to find that these priced factors subsume most BAF alphas, as the six-factor model, which captures asset returns well, should largely capture priced factors omitted by these anomaly portfolios. Some BAF alphas are still significant after the adjustment, indicating there are priced factors omitted by the six-factor model. The empirical results are highly consistent with the analytical conclusions, supporting the rationality of analyzing factor-beta anomalies under the standard covariance-based asset-pricing framework.

FIGURE 3.6: Controlling for Unpriced and Priced Factors

This figure reports the performance of the anomalies from Figure 3.4 with unpriced or priced factors controlled. The first subfigure reports the BAF alpha from the spanning regressions of each BAF on its corresponding FMP and the mean-removed market-adjusted industry returns (the twelve industries). The second subfigure reports the BAF alpha from the spanning regressions of each BAF on its corresponding FMP and the six factors of Fama and French (2018). The third subfigure reports the BAF alphas when the FMP, mean-removed market-adjusted industry returns and six factors are controlled simultaneously. The 95% confidence intervals based on Newey and West (1987) adjusted standard errors are also reported. The full sample period is from February 1967 to December 2019, with data from the first five years used for the initial beta estimation.



3.5 Conclusion

This study shows that factor-beta anomalies do not challenge the covariance-based asset-pricing framework but are induced by factors beyond the factor-mimicking portfolio (FMP) from which betas are estimated. Both omitted unpriced and priced factors correlated with the FMP cause estimated FMP betas to deviate from true FMP betas, as estimated FMP betas also capture loadings on omitted factors. This beta deviation positively contributes to the alpha from the spanning regression of the betting-against-factor-beta portfolio (BAF) on the FMP. Exposures to omitted priced factors captured by estimated betas can also contribute to a BAF alpha, as these exposures are compensated by extra premiums. The impact from extra premiums is attenuated by the beta-deviation component when they negatively contribute to a BAF alpha but enhanced when they positively contribute to a BAF alpha. Therefore, low factor-beta anomalies should be pervasive when betting against betas of priced factors with positive premiums.

Although both unpriced and priced-factor can generate a positive BAF alpha, the strength of a BAF alpha unpriced factors can reconcile is different from that priced factors can reconcile. As the unpriced-factor channel induces a negative beta-alpha relationship only through beta deviation, the strength of a BAF alpha from this channel depends on the deviation of estimated FMP betas from true FMP betas. As long as an FMP is not entirely inefficient, estimated betas should capture information about true betas and hence the BAF alpha from this channel is bounded. Even to generate a beta anomaly weaker than those observed empirically, this channel requires either an FMP to be extremely inefficient or unpriced factors to overwhelm priced factors in driving the time variation of asset returns.

The two requirements are inconsistent with previous empirical findings that FMPs well capture the underlying latent return sources and the conception of the unpriced factor under which it is avoidable (e.g., Roll and Ross 1984; Daniel and Titman 1997; Clarke 2022). Instead, the priced-factor channel can generate strong BAF alphas as, besides inducing beta deviation, this channel also carries extra premiums. Since the impacts from many omitted priced factors can stack up, a set of omitted priced factors weakly capturing asset returns or weakly correlated with an FMP can also generate a strong low factor-beta anomaly. Therefore, omitted priced factors dominate unpriced factors in driving factor-beta anomalies.

Factor-beta anomalies contain information about latent factors. The practical difficulty of extracting the information is that many arbitrary choices largely impact BAF performance. For example, how an FMP is constructed or how betas are estimated (the data frequency for beta estimation). These choices affect the information captured by estimated betas, which, in turn, affects the information captured by a BAF portfolio and its performance. Alleviating these concerns is of first-order importance to systematically extract the relevant information. Even with these caveats, we can still infer information from an observed factor-beta anomaly under the analytical framework of this study. A prominent example is that a factor-beta anomaly whose strength exceeds a certain bound must indicate the existence of omitted priced factors.

Conclusion

This thesis refreshes the understanding of well-documented factors and anomalies from novel conditional perspectives. The first chapter is motivated by the long-standing argument about the validity of the size effect and the extensive evidence on the slow information incorporation of small-stock prices. The finding shows that the size effect comprises small-minus-big and big-minus-small effects, which resurrects the size investment strategy and raises the necessity of assessing the size effect conditionally. The second chapter is motivated by the commonality of existing explanations for the low-beta anomaly. While many theoretical or empirical explanations appear to resolve the anomaly, this chapter shows that the anomaly is more abnormal than documented in the literature. Specifically, the discovery of the new low-beta anomaly makes three contributions. First, the new low-beta anomaly identifies partial-correlated factors, which are unnoticeable but are important for driving asset returns. Second, methodologically, the finding translates into a stricter standard for future studies to resolve the low-beta anomaly. Third, empirically, the finding translates into a new low-beta investment strategy orthogonal to existing low-risk strategies. The third chapter extends the analysis of beta anomalies to all factors and anomalies, showing that a negative beta-alpha relationship is not a challenge to the covariance-based asset-pricing framework but is a result of it. The finding predicts that low factor-beta anomalies must be pervasive, for which this chapter documents solid empirical support by examining a large set of factor and anomaly portfolios.

Besides contributing to the literature, these chapters also pave the way for further extensions. The finding in the first chapter shows that the conditional small-minus-big and big-minus-small effects are much stronger than the unconditional size effect, and hence the observed size effect is determined by the relative realized strength of the two conditional size effects. However, it is still interesting to investigate if an unconditional size risk premium exists. Claiming the existence of a risk premium relies on well-designed econometric techniques, robust empirical evidence and solid economic reasoning, which is one important extension of this chapter in the next stage. A question directly following the second chapter is: what is the new low-beta anomaly's underlying economic source? It is promising to develop a model to generate the new low-beta anomaly endogenously, for which the key is that the new version should be orthogonal to the version known in the literature. The third chapter, which documents the pervasiveness of factor-beta anomalies, also has potential extensions. For example, the finding relates to recent studies that advocate using individual assets to evaluate factor models. While these studies develop techniques to correct the error-in-variable bias, the omitted factor problem, as implied by the pervasive low beta anomalies, is severe and makes tests using individual assets problematic. It is necessary to systematically evaluate the trade-off between the low test power and omitted variable bias when using individual stocks or portfolios as test assets.

This thesis aims to better understand the interaction between the financial market and the real economy. As asset pricing factors are at the core of the interaction (through the stochastic discount factor), the current findings, which enhance the understanding of several asset pricing factors and anomalies in the time series and cross section, take a crucial step towards this goal.

Bibliography

Ahn, D.-H., Min, B.-K., and Yoon, B. (2019). Why has the size effect disappeared? *Journal of Banking and Finance* 102(C), 256–276.

Akbari, A., Carrieri, F., and Malkhozov, A. (Dec. 20, 2021). Can Cross-Border Funding Frictions Explain Financial Integration Reversals? *The Review of Financial Studies* 35(1). Publisher: Oxford Academic, 394–437. ISSN: 0893-9454.

Alquist, R., Israel, R., and Moskowitz, T. (2018). Fact, Fiction, and the Size Effect. *The Journal of Portfolio Management* 45(1), 34–61.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets* 5(1), 31–56.

Andrei, D., Cujean, J., and Wilson, M. I. (Dec. 2021). *The Lost Capital Asset Pricing Model*. en. SSRN Scholarly Paper. Rochester, NY.

Ang, A., Hodrick, R., Xing, Y., and Zhang, X. (2006). The Cross-Section of Volatility and Expected Returns. *Journal of Finance* 61(1), 259–299.

Asness, C., Frazzini, A., Israel, R., Moskowitz, T. J., and Pedersen, L. H. (2018). Size matters, if you control your junk. *Journal of Financial Economics* 129(3), 479–509.

Asness, C. S., Porter, R. B., and Stevens, R. L. (Feb. 2000). *Predicting Stock Returns Using Industry-Relative Firm Characteristics*. SSRN Scholarly Paper ID 213872. Rochester, NY: Social Science Research Network.

Badrinath, S., Kale, J., and Thomas, N. (Feb. 1995). Of Shepherds, Sheep, and the Cross-Autocorrelations in Equity Returns. *Review of Financial Studies* 8, 401–30.

Bali, T. G., Brown, S., Murray, S., and Tang, Y. (2017). A Lottery-Demand-Based Explanation of the Beta Anomaly. *Journal of Financial and Quantitative Analysis* 52(6), 2369–2397.

Bali, T. G., Cakici, N., and Whitelaw, R. F. (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics* 99(2), 427–446.

Bali, T. G., Peng, L., Shen, Y., and Tang, Y. (Nov. 2013). Liquidity Shocks and Stock Market Reactions. *The Review of Financial Studies* 27(5), 1434–1485. ISSN: 0893-9454.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics* 9(1), 3–18.

Barillas, F. and Shanken, J. (2018). Comparing Asset Pricing Models. *The Journal of Finance* 73(2), 715–754.

Barroso, P., Detzel, A. L., and Maio, P. F. (June 2020). Managing the risk of the low risk anomaly. *SSRN Scholarly Paper* (ID 2876450).

Barry, C. B. and Brown, S. (1985). Differential Information and Security Market Equilibrium. *Journal of Financial and Quantitative Analysis* 20(4), 407–422.

Black, F. (1972). Capital Market Equilibrium with Restricted Borrowing. *The Journal of Business* 45(3), 444–55.

Black, F., Jensen, M. C., and Scholes, M. (1972). The Capital Asset Pricing Model: Some Empirical Tests. *In Studies in the Theory of Capital Markets, edited by M. C. Jensen. New York: Praeger.*

Boguth, O. and Simutin, M. (2018). Leverage constraints and asset prices: Insights from mutual fund risk taking. *Journal of Financial Economics* 127(2), 325–341.

Brennan, M. J. (1971). Capital Market Equilibrium with Divergent Borrowing and Lending Rates. *The Journal of Financial and Quantitative Analysis* 6(5), 1197–1205. ISSN: 00221090, 17566916.

Brennan, M., Jegadeesh, N., and Swaminathan, B. (1993). Investment Analysis and the Adjustment of Stock Prices to Common Information. *Review of Financial Studies* 6(4), 799–824.

Brown, P., Kleidon, A. W., and Marsh, T. A. (1983). New evidence on the nature of size-related anomalies in stock prices. *Journal of Financial Economics* 12(1), 33–56.

Campbell, J. (1991). A Variance Decomposition for Stock Returns. *Economic Journal* 101(405), 157–79.

Campbell, J. Y. and Shiller, R. (1988). The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors. *Review of Financial Studies* 1(3), 195–228.

Carhart, M. (1997). On Persistence in Mutual Fund Performance. *Journal of Finance* 52(1), 57–82.

Chamberlain, G. and Rothschild, M. (1983). Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets. *Econometrica* 51(5), 1281–1304.

Chen, Z. and Lu, A. (Dec. 1, 2019). A Market-Based Funding Liquidity Measure. *The Review of Asset Pricing Studies* 9(2), 356–393. ISSN: 2045-9920.

Chordia, T. and Swaminathan, B. (2000). Trading Volume and Cross-Autocorrelations in Stock Returns. *Journal of Finance* 55(2), 913–935.

Clarke, C. (2022). The level, slope, and curve factor model for stocks. *Journal of Financial Economics* 143(1), 159–187.

Cochrane, J. H. (Jan. 23, 2005). *Asset Pricing: Revised Edition*. Revised ed. edition. Princeton, N.J: Princeton University Press. 560 pp. ISBN: 978-0-691-12137-6.

Cohen, R. B., Gompers, P., and Vuolteenaho, T. (2002). Who underreacts to cash-flow news? evidence from trading between individuals and institutions. *Journal of Financial Economics* 66(2-3), 409–462.

Cohen, R. B., Polk, C., and Vuolteenaho, T. (2003). The Value Spread. *The Journal of Finance* 58(2), 609–641.

Cooper, I., Ma, L., Maio, P., and Philip, D. (June 2021). Multifactor Models and Their Consistency with the APT. *The Review of Asset Pricing Studies* 11(2), 402–444. ISSN: 2045-9920.

Cooper, M., Gulen, H., and Schill, M. (Feb. 2008). Asset Growth and the Cross-Section of Stock Returns. *Journal of Finance* 63, 1609–1651.

Cooper, M. J., Gutierrez, R. C., and Hameed, A. (2004). Market States and Momentum. *The Journal of Finance* 59(3), 1345–1365.

Crain, M. A. (2011). *A Literature Review of the Size Effect*. SSRN Scholarly Paper ID 1710076. Rochester, NY: Social Science Research Network.

Daniel, K., Hirshleifer, D., and Sun, L. (Apr. 2020a). Short- and Long-Horizon Behavioral Factors. en. *Rev Financ Stud* 33(4). Publisher: Oxford Academic, 1673–1736. ISSN: 0893-9454.

Daniel, K., Mota, L., Rottke, S., and Santos, T. (2020b). The Cross-Section of Risk and Returns. *Review of Financial Studies* 33(5), 1927–1979.

Daniel, K. and Titman, S. (1997). Evidence on the Characteristics of Cross Sectional Variation in Stock Returns. *Journal of Finance* 52(1), 1–33.

Davis, J. L., Fama, E. F., and French, K. R. (2000). Characteristics, Covariances, and Average Returns: 1929 to 1997. *Journal of Finance* 55(1), 389–406.

DeMiguel, V., Garlappi, L., and Uppal, R. (2009). Optimal Versus Naive Diversification: How Inefficient is the 1-N Portfolio Strategy? *Review of Financial Studies* 22(5), 1915–1953.

Dimson, E. and Marsh, P. (1999). *Murphy's Law and Market Anomalies*. SSRN Scholarly Paper ID 137337. Social Science Research Network.

Ehsani, S. and Linnainmaa, J. T. (2019). *Factor Momentum and the Momentum Factor*. SSRN Scholarly Paper ID 3014521. Rochester, NY: Social Science Research Network.

Esakia, M., Goltz, F., Luyten, B., and Sibbe, M. (2019). Size Factor in Multifactor Portfolios: Does the Size Factor Still Have Its Place in Multifactor Portfolios? *The Journal of Index Investing*. Publisher: Institutional Investor Journals Umbrella.

Fama, E. and French, K. (1992). The Cross-Section of Expected Stock Returns. *Journal of Finance* 47(2), 427–65.

Fama, E. and French, K. (2004). The Capital Asset Pricing Model: Theory and Evidence. *Journal of Economic Perspectives* 18(3), 25–46.

Fama, E. and MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy* 81(3), 607–36.

Fama, E. F. (1996). Multifactor Portfolio Efficiency and Multifactor Asset Pricing. *The Journal of Financial and Quantitative Analysis* 31(4). Publisher: [Cambridge University Press, University of Washington School of Business Administration], 441–465. ISSN: 0022-1090.

Fama, E. F. and French, K. R. (Feb. 1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33(1), 3–56.

Fama, E. F. and French, K. R. (1996). Multifactor Explanations of Asset Pricing Anomalies. *Journal of Finance* 51(1), 55–84.

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

Fama, E. F. and French, K. R. (2016). Dissecting Anomalies with a Five-Factor Model. *Review of Financial Studies* 29(1), 69–103.

Fama, E. F. and French, K. R. (2018). Choosing factors. *Journal of Financial Economics* 128(2), 234–252.

Fama, E. F. and French, K. R. (2020). Comparing Cross-Section and Time-Series Factor Models. *Review of Financial Studies* 33(5), 1891–1926.

Frazzini, A., Israel, R., and Moskowitz, T. (2015). Trading Costs of Asset Pricing Anomalies. *Working Paper*.

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

Friend, I. and Blume, M. E. (1970). Measurement of Portfolio Performance Under Uncertainty. *American Economic Review* 60(4), 561–75.

George, T. J. and Hwang, C.-Y. (2004). The 52-Week High and Momentum Investing. en. *The Journal of Finance* 59(5). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.11
6261.2004.00695.x, 2145–2176. ISSN: 1540-6261.

Giglio, S. and Xiu, D. (2021). Asset Pricing with Omitted Factors. *Journal of Political Economy* 129(7), 1947–1990.

Gospodinov, N. and Robotti, C. (2021). Common pricing across asset classes: Empirical evidence revisited. *Journal of Financial Economics* 140(1), 292–324.

Gupta, T. and Kelly, B. (2019). Factor Momentum Everywhere. *The Journal of Portfolio Management* 45(3), 13–36.

Haddad, V., Kozak, S., and Santosh, S. (May 2020). Factor Timing. en. *Rev Financ Stud* 33(5). Publisher: Oxford Academic, 1980–2018. ISSN: 0893-9454.

Hansen, L. and Jagannathan, R. (1991). Implications of Security Market Data for Models of Dynamic Economies. *Journal of Political Economy* 99(2), 225–62.

Hansen, L. P. and Richard, S. F. (1987). The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models. *Econometrica* 55(3). Publisher: [Wiley, Econometric Society], 587–613. ISSN: 0012-9682.

Herskovic, B., Moreira, A., and Muir, T. (Jan. 2019). *Hedging Risk Factors.* SSRN Scholarly Paper. Rochester, NY: Social Science Research Network.

Hirshleifer, D. (2001). Investor Psychology and Asset Pricing. *Journal of Finance* 56(4), 1533–1597.

Hong, H., Lim, T., and Stein, J. (2000). Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies. *Journal of Finance* 55(1), 265–295.

Hong, H. and Sraer, D. A. (2016). Speculative Betas. *The Journal of Finance* 71(5), 2095–2144.

Horowitz, J. L., Loughran, T., and Savin, N. E. (2000a). Three analyses of the firm size premium. *Journal of Empirical Finance* 7(2), 143–153.

Horowitz, J. L., Loughran, T., and Savin, N. E. (2000b). The disappearing size effect. *Research in Economics* 54(1), 83–100.

Hou, K. (2007). Industry Information Diffusion and the Lead-lag Effect in Stock Returns. *The Review of Financial Studies* 20(4). Publisher: Oxford Academic, 1113–1138.

Hou, K., Mo, H., Xue, C., and Zhang, L. (Feb. 2020a). An Augmented q-Factor Model with Expected Growth. *Review of Finance* 25(1), 1–41.

Hou, K. and van Dijk, M. A. (2019). Resurrecting the Size Effect: Firm Size, Profitability Shocks, and Expected Stock Returns. *The Review of Financial Studies* 32(7). Publisher: Oxford Academic, 2850–2889.

Hou, K., Xue, C., and Zhang, L. (2015). Digesting Anomalies: An Investment Approach. *Review of Financial Studies* 28(3), 650–705.

Hou, K., Xue, C., and Zhang, L. (2020b). Replicating Anomalies. *Review of Financial Studies* 33(5), 2019–2133.

Jagannathan, R. and Wang, Z. (1996). The Conditional CAPM and the Cross-Section of Expected Returns. *The Journal of Finance* 51(1), 3–53.

Jegadeesh, N. and Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance* 48(1), 65–91.

Jegadeesh, N. and Titman, S. (Nov. 2011). Momentum. *Annu. Rev. Financ. Econ.* 3(1). Publisher: Annual Reviews, 493–509. ISSN: 1941-1367.

Jiang, G. J. and Zhu, K. X. (2017). Information Shocks and Short-Term Market Underreaction. *Journal of Financial Economics* 124(1), 43–64.

Jylhä, P. (2018). Margin Requirements and the Security Market Line. en. *The Journal of Finance* 73(3). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.12616, 12
1321. ISSN: 1540-6261.

Keim, D. B. (1983). Size-related anomalies and stock return seasonality: Further empirical evidence. *Journal of Financial Economics* 12(1), 13–32.

Kozak, S., Nagel, S., and Santosh, S. (2018). Interpreting Factor Models. *Journal of Finance* 73(3), 1183–1223.

Lamoureux, C. G. and Sanger, G. C. (1989). Firm size and turn-of-the-year effects in the OTC/NASDAQ market. *The Journal of Finance* 44(5), 1219–1245.

Lettau, M. and Pelger, M. (Sept. 2020a). Estimating latent asset-pricing factors. en. *Journal of Econometrics* 218(1), 1–31. ISSN: 0304-4076.

Lettau, M. and Pelger, M. (May 2020b). Factors That Fit the Time Series and Cross-Section of Stock Returns. *The Review of Financial Studies* 33(5), 2274–2325. ISSN: 0893-9454.

Lewellen, J. (2022). How many factors? *Working Paper.*

Lewellen, J. and Nagel, S. (2006). The conditional CAPM does not explain asset-pricing anomalies$. en. *Journal of Financial Economics*, 26.

Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics* 47(1), 13–37.

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

Lo, A. W. and MacKinlay, A. C. (1990). When are Contrarian Profits Due to Stock Market Overreaction? *The Review of Financial Studies* 3(2), 175–205.

Lochstoer, L. A. and Tetlock, P. C. (2020). What Drives Anomaly Returns? *The Journal of Finance* 75(3), 1417–1455.

MacKinlay, A. C. (1995). Multifactor models do not explain deviations from the CAPM. *Journal of Financial Economics* 38(1), 3–28.

MacKinlay, A. C. and Pástor, Ľ. (2000). Asset Pricing Models: Implications for Expected Returns and Portfolio Selection. *The Review of Financial Studies* 13(4), 883–916.

McQueen, G., Pinegar, M., and Thorley, S. (1996). Delayed Reaction to Good News and the Cross-Autocorrelation of Portfolio Returns. *The Journal of Finance* 51(3), 889–919. ISSN: 00221082, 15406261.

Mech, T. S. (1993). Portfolio return autocorrelation. *Journal of Financial Economics* 34(3), 307–344.

Merton, R. (1973). An Intertemporal Capital Asset Pricing Model. *Econometrica* 41(5), 867–87.

Merton, R. (1987). A Simple Model of Capital Market Equilibrium with Incomplete Information. *Journal of Finance* 42(3), 483–510.

Moreira, A. and Muir, T. (2017). Volatility-Managed Portfolios. *Journal of Finance* 72(4), 1611–1644.

Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica* 34(4), 768–783. ISSN: 00129682, 14680262.

Murray, S. (Feb. 2020). *Betting Against Other Betas*. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network.

Newey, W. and West, K. (1987). A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55(3), 703–08.

Novy-Marx, R. (Mar. 2012). Is momentum really momentum? en. *Journal of Financial Economics* 103(3), 429–453. ISSN: 0304-405X.

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

Novy-Marx, R. and Velikov, M. (2022). Betting against betting against beta. *Journal of Financial Economics* 143(1), 80–106.

Petkova, R. and Zhang, L. (2005). Is value riskier than growth? *Journal of Financial Economics* 78(1), 187–202.

Pukthuanthong, K., Roll, R., and Subrahmanyam, A. (2019). A Protocol for Factor Identification. *Review of Financial Studies* 32(4), 1573–1607.

Roll, R. (1977). A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory. *Journal of Financial Economics* 4(2), 129–176.

Roll, R. and Ross, S. (1984). A Critical Reexamination of the Empirical Evidence on the Arbitrage Pricing Theory: A Reply. *Journal of Finance* 39(2), 347–50.

Ross, S. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory* 13(3), 341–360.

Schneider, P., Wagner, C., and Zechner, J. (2020). Low-Risk Anomalies? *Journal of Finance* 75(5), 2673–2718.

Sharpe, W. (1964). Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk. *Journal of Finance* 19(3), 425–442.

Shumway, T. (1997). The Delisting Bias in CRSP Data. *Journal of Finance* 52(1), 327–40.

Stambaugh, R. F., Yu, J., and Yuan, Y. (2015). Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle. *The Journal of Finance* 70(5), 1903–1948.

Stambaugh, R. F. and Yuan, Y. (2017). Mispricing Factors. *The Review of Financial Studies* 30(4), 1270–1315.

Treynor, J. L. (1966). Can Mutual Funds Outguess the Market? *Harvard business review.* 44(4). ISSN: 0017-8012.

Van Dijk, M. A. (2011). Is size dead? A review of the size effect in equity returns. *Journal of Banking and Finance* 35(12), 3263–3274.

Vuolteenaho, T. (2002). What Drives Firm-Level Stock Returns? *The Journal of Finance* 57(1), 233–264.

Wang, B. (2019). The cash conversion cycle spread. *Journal of Financial Economics* 133(2), 472–497.

Zhang, X. F. (2006). Information Uncertainty and Stock Returns. *Journal of Finance* 61(1), 105–137.

Appendix A

Chapter 1 Supplement

A1 Additional Tests of the Signals

This appendix section examines the robustness of the two-faced size effect using different variations of the forward-looking signals.

A1.1 The relative importance of past big- and small-stock returns in predicting the SMB portfolio performance

To complement the portfolio-level analysis of Table 1.1, we exercise the predictive regressions of the SMB return on the past big-stock return (M), small-stock return (F), or both. Big- and small-stock returns are the same as those used to construct the forward-looking signals. In these predictive regressions, a predictor takes the value of one if it is positive and zero otherwise. The first two rows of Table A1.1 report the single-predictor regression results. The slopes in both cases are significant, and the adjusted R-squares are around 3%, implying that either the big-stock return or the small-stock return can predict the subsequent realization of the size effect. The bottom row reports the predictive regression result when M and F are included as predictors simultaneously. Both slopes (b_M and b_F) are significant, indicating that the two returns provide independent information for predicting the realization of the size effect. Therefore, using information from both returns to identify the two-faced size effect is reasonable.

TABLE A1.1: Predicting SMB Performance Using Past Big- and Small-Stock Performance

This table reports the predictive regression intercepts (a), slopes (b_M or b_F), and adjusted R-squares of monthly SMB returns on past big-stock (M) or small-stock (F) returns. Big- and small-stock returns are the same as those used to construct indicators. The SMB portfolio is the same as that in Table 1.1. A predictor takes the value of one if it is positive and zero otherwise. The t-statistics are based on Newey and West (1987) adjusted standard errors. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The sample period is August 1926 to December 2019.

	a	b_M	b_F	$\bar{R}^2(\%)$
$SMB_t = a + b_M M_{t-1} + e_t$	-0.75***	1.76***		2.93
	(-3.47)	(5.96)		
$SMB_t = a + b_F F_{t-1} + e_t$	-0.82***		1.96***	3.73
	(-4.17)		(6.66)	
$SMB_t = a + b_M M_{t-1} + b_F F_{t-1} + e_t$	-0.99***	0.76**	1.45***	3.95
	(-4.56)	(2.07)	(3.91)	

A1.2 The levels of past big- and small-stock returns

In the main analysis, we construct the IG indicator based on the signs of past big- and small-stock returns. Following the slow information incorporation mechanism that motivates this study, the levels of past big- and small-stock returns should also be informative about the performance of the size effect. Extremely high or low past returns are more likely to be associated with information shocks since the non-shock component of realized returns, the expected return, is usually bounded. This appendix evaluates the impact of the magnitude of past big- and small-stock returns to confirm this point. We sort the negative past big-stock returns into two levels (M1 and M2, below and above the median negative big-stock return) and positive past big-stock returns into two levels (M3 and M4, below and above the median positive big-stock return). Similarly, we sort negative and positive past small-stock returns into two levels separately, denoted by F1, F2, F3, and F4.

Table A1.2 reports the SMB portfolio returns conditional on different combinations of past small- and big-stock returns. No values are available for the combinations M1-F4 and M4-F1, as observations corresponding to these two cases are too scarce (fewer than ten); hence, we cannot get reliable statistics. The overall pattern is that returns increase from the upper left to the lower right: higher/lower past big- and small-stock returns tend to be followed by better/worse realization of the unconditional size effect, indicating that the levels of past returns are also informative. Moreover, the result again shows that combining both returns to identify the two-faced size effect is reasonable, given that the performance is more remarkable when the two returns provide consistent signals.

A1.3 Returns of different past horizons

Our slow information incorporation-based motivation indicates that concurrent information relevant to the size effect should be informative about its subsequent performance. In the main analysis, the signals are constructed using past one-month returns to utilize the timeliest information at the monthly frequency. However, neither the literature nor the current study specifies the exact past horizon to use. We expect that the returns of other short-term horizons can also identify

TABLE A1.2: SMB Performance Based on the Levels of Past Big- and Small-Stock Returns

This table reports the average SMB returns conditional on different levels of past big- and small-stock returns. Big- and small-stock returns are the same as those used to construct indicators. The SMB portfolio is the same as that in Table 1.1. M1/M2 refers to past big-stock returns below/above the median negative big-stock return. M3/M4 refers to past big-stock returns below/above the median positive big-stock return. Similarly, F1/F2 refers to past small-stock returns below/above the median negative small-stock return, and F3/F4 refers to past small-stock returns below/above the median positive small-stock return. The t-statistics are in parentheses; *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The sample period is from August 1926 to December 2019.

	M1	M2	M3	M4
F1	-0.69	-2.25***	1.97	
	(-1.58)	(-4.51)	(1.23)	
F2	-0.91	-0.95***	-0.32	0.34
	(-1.41)	(-3.16)	(-0.70)	(0.24)
F3	1.57	1.05	0.60**	0.34
	(0.96)	(1.53)	(2.34)	(1.18)
F4		-1.54	1.00***	2.12***
		(-1.04)	(2.79)	(4.63)

the two-faced size effect. Panel A of Table A1.3 confirms this conjecture by investigating the performance of the IG indicators constructed from big-stock (M) and small-stock (F) returns of different past horizons. For example, the case of M=3 and F=3 corresponds to the IG indicator constructed using the previous three-month cumulative big- and small-stock returns. We run cross-sectional regressions of return on size following Section 1.3.2 and evaluate the slopes conditional on up and down signals. The results show a significant SMB effect conditional on the up signal and a significant BMS effect conditional on the down signal for all these combinations, leading to the same conclusion as Table 1.2.

Panel B reports how the past big-stock (the M indicator) or small-stock (the F indicator) return alone identifies the two-faced size effect. The M (or F) indicator is defined according to the sign of the past period return. For example, in the case of M=1, M gives an up/down signal when the prior one-month big-stock return is positive/negative; in the case of F=1, F gives an up/down signal when the prior one-month small-stock return is positive/negative. For most cases, the average slopes are significantly negative when the M (F) indicator gives an up signal and significantly positive when the M (F) indicator gives a down signal. The results indicate that both M and F provide information for identifying the two-faced size effect, consistent with the portfolio-level evidence.

Panel C examines the two-faced size effect based on signals constructed from the past market return (in excess of the risk-free rate) and SMB return. Using aggregate market and SMB returns is also consistent with our motivation, as they reflect market-wide information and information relevant to small stocks. For example, a negative market return indicates the existence of bad news for the market; meanwhile, the news specifically for small stocks is also bad if the SMB return is negative. Accordingly, the size effect should have bad performance subsequently when continuing to incorporate the two sets of information. As shown in Panel B, the two returns also identify the two-faced size effect well. The key is that any measures capturing information shocks relevant to small stocks are informative about the size effect's subsequent performance.

Table A1.3: Signals Based on the Returns of Different Past Horizons

This table reports the firm-level cross-sectional regression slopes of return on size. The average cross-sectional regression slopes conditional on the up, down, and uncertain (Uncer) signals are reported. In Panel A, the signals are given by the IG indicator constructed using big-stock (M) and small-stock (F) returns. In Panel B, the signals are given by the M or F indicator constructed using the big-stock or small-stock return alone. In Panel C, the signals are given by the IG indicator constructed using the excess market return (RmRf) and the SMB portfolio return. Returns of different past horizons are used when constructing indicators. For example, the case of M=3 and F=3 corresponds to the IG indicator constructed using the previous three-month cumulative big- and small-stock returns. The t-statistics are based on Newey and West (1987) adjusted standard errors. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The full sample period is from August 1926 to December 2019.

	Up	Down	Uncer		Up	Down	Uncer
Panel A: IG based on Small-Stock (F) and Big-Stock (M) returns							
M=1, F=3	-0.47***	0.36***	0.04	M=3, F=1	-0.43***	0.30***	0.07
	(-8.41)	(5.68)	(0.77)		(-7.93)	(5.30)	(1.18)
M=1, F=6	-0.40***	0.28***	0.01	M=3, F=3	-0.35***	0.23***	0.07
	(-7.00)	(3.75)	(0.20)		(-6.95)	(3.98)	(0.91)
M=1, F=12	-0.36***	0.26***	0.00	M=3, F=6	-0.33***	0.18**	0.03
	(-5.99)	(3.06)	(0.01)		(-5.98)	(2.46)	(0.53)
Panel B: Signals based on Small-Stock (F) or Big-Stock (M) return alone							
M=1	-0.33***	0.23***		F=1	-0.37***	0.26***	
	(-6.72)	(4.83)			(-7.42)	(5.89)	
M=3	-0.27***	0.18***		F=3	-0.31***	0.21***	
	(-5.72)	(3.28)			(-6.78)	(4.27)	
M=6	-0.16***	0.00		F=6	-0.22***	0.08	
	(-3.42)	(0.01)			(-4.63)	(1.35)	
Panel C: IG based on Market (RmRf) and SMB returns							
RmRf=1, SMB=1	-0.51***	0.37***	-0.07	RmRf=3, SMB=1	-0.47***	0.27***	-0.01
	(-7.97)	(6.10)	(-1.58)		(-8.14)	(4.49)	(-0.33)
RmRf=1, SMB=3	-0.52***	0.30***	-0.04	RmRf=3, SMB=3	-0.39***	0.24***	-0.07
	(-7.82)	(4.63)	(-0.81)		(-6.17)	(3.75)	(-1.54)
RmRf=1, SMB=6	-0.47***	0.25***	-0.04	RmRf=3, SMB=6	-0.37***	0.19***	-0.06
	(-7.06)	(3.52)	(-0.87)		(-5.68)	(2.64)	(-1.23)

A2 Two-Faced Size Strategy Using Different SMB Portfolios

As a robustness check, this subsection examines whether the two-faced effect enhances the performance of other SMB portfolios. We consider the BM-balanced size portfolio (SMB^{FF3}, Fama and French 1993), the BM, operational profitability, and investment-balanced size portfolio (SMB^{FF5}, Fama and French 2015), the investment and profitability-balanced size portfolio from the q-factor model (ME^{q}, Hou et al. 2015), the size portfolio constructed using the smallest and largest deciles or quintiles (SMB^{1Q} and SMB^{QI}, respectively), and the size portfolio constructed using the smallest and largest five deciles (SMB^{5Q}). We compare the performance of the original (OR) and managed (LSI, following Section 1.4.1) portfolios. The full sample period is July 1963 to December 2019 for SMB^{FF5}, January 1967 to December 2019 for ME^{q}, and August 1926 to December 2019 for the remaining size portfolios. Table A2 reports the performance of these size portfolios and their managed performance. Compared with the original size portfolios (OR), LSI portfolios achieve better performance that survives transaction costs, as suggested by their higher Sharpe ratios (SR) and significant spanning regression alphas. The results confirm that managing size strategies according to the two-faced size effect improves their mean-variance efficiency.

TABLE A2: Performance of Original and Managed Size Strategies

This table reports the performance of different SMB portfolios and their corresponding managed portfolios (LSI). The position of each LSI portfolio is adjusted based on the signals given by the IG indicator: when the signal is up (down), an LSI buys (shorts) OR; when the signal is uncertain, an LSI quits the risky investment (and hence earns zero returns during such periods). Average monthly returns (μ), annualized Sharpe ratios (SR), and spanning regression alphas of LSI on OR (α) are reported. The LSI performance adjusting for the transaction cost (TC) of ten or twenty basis points is also reported. Details of these SMB portfolios and their sample periods are provided in Section A2. The t-statistic to evaluate whether the improvement relative to OR is significant is reported under each alpha and Sharpe ratio. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	OR		LSI, TC=0			LSI, TC=10			LSI, TC=20		
	μ(%)	SR	μ(%)	SR	α(%)	μ(%)	SR	α(%)	μ(%)	SR	α(%)
SMB^{1Q}	0.45*	0.21	1.36***	0.69***	1.24***	1.28***	0.65***	1.15***	1.19***	0.60***	1.07***
	(1.93)		(6.36)	(3.85)	(8.81)	(5.94)	(3.50)	(8.17)	(5.52)	(3.15)	(7.53)
SMB^{QI}	0.37*	0.22	1.15***	0.73***	1.05***	1.06***	0.68***	0.96***	0.98***	0.62***	0.88***
	(1.93)		(6.63)	(4.12)	(8.55)	(6.11)	(3.68)	(7.81)	(5.59)	(3.24)	(7.07)
SMB^{5Q}	0.19*	0.20	0.63***	0.69***	0.58***	0.54***	0.60***	0.49***	0.46***	0.50**	0.40***
	(1.80)		(6.33)	(3.96)	(8.22)	(5.43)	(3.20)	(6.93)	(4.54)	(2.44)	(5.66)
SMB^{FF3}	0.20**	0.22	0.59***	0.73***	0.57***	0.51***	0.63***	0.48***	0.42***	0.52**	0.40***
	(1.98)		(6.79)	(3.76)	(7.50)	(5.76)	(2.99)	(6.30)	(4.74)	(2.21)	(5.12)
SMB^{FF5}	0.23*	0.26	0.61***	0.83***	0.61***	0.53***	0.72***	0.53***	0.44***	0.60**	0.44***
	(1.87)		(6.48)	(3.88)	(6.54)	(5.54)	(3.11)	(5.60)	(4.61)	(2.34)	(4.66)
ME^{q}	0.27**	0.31	0.54***	0.73***	0.55***	0.46***	0.62**	0.47***	0.37***	0.50	0.39***
	(2.22)		(5.63)	(2.76)	(5.86)	(4.72)	(2.02)	(4.93)	(3.82)	(1.28)	(4.02)

A3 Two-Faced Size Strategy Using All-But-Microcap Stocks

This appendix section examines the performance of the two-faced size strategy using all-but-microcap stocks. The small-minus-big portfolio ($SMB_{ExMicro}$) is constructed by buying the third decile and selling the tenth decile of size-sorted portfolios (based on NYSE breakpoints). The return of each side is calculated on a value-weighted basis. We manage the $SMB_{ExMicro}$ portfolio according to the two-faced size effect following Section 1.4.1, holding/selling $SMB_{ExMicro}$ when the signal is up/down and quitting the risky investment when the signal is uncertain. Table A3 reports the performance of the original $SMB_{ExMicro}$ portfolio, denoted by OR, and the managed $SMB_{ExMicro}$ portfolio, denoted by LSI. Similar to when all stocks are included, while the original $SMB_{ExMicro}$ portfolio does not generate significantly positive alphas, the managed $SMB_{ExMicro}$ portfolio generates considerable and significant abnormal performance unexplained by common risk factors. The result indicates that the two-faced size effect among all-but-microcap stocks is economically meaningful.

Table A3: Two-Faced Size Strategy on All-But-Microcap Stocks

This table presents the spanning regression results of the original and managed size portfolios (OR and LSI) on asset pricing models. The OR portfolio is the small-minus-big portfolio constructed without including microcap stocks ($SMB_{ExMicro}$). The position of LSI is adjusted according to the signals given by the IG indicator: when the signal is up (down), LSI buys (shorts) OR; when the signal is uncertain, LSI quits the risky investment (and hence earns zero returns during such periods). Monthly alphas, factor loadings (under each factor notation), and R-squares are reported. Details of these factor models and their data availability are introduced in Table 1.4. The t-statistics are based on Newey and West (1987) adjusted standard errors. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The sample period depends on the data availability of these factors. The full sample period is from August 1926 to December 2019.

	α(%)	MKT	SIZE	HML	MOM	CMA	RMW	IA	ROE	EG	$\bar{R}^2$(%)
OR	-0.10***	0.10***	1.24***	0.42***							92.02
	(-2.63)	(5.18)	(36.28)	(8.79)							
LSI	0.69***	-0.01	0.24	0.17							4.89
	(6.47)	(-0.33)	(1.26)	(1.45)							
OR	-0.06	0.09***	1.24***	0.40***	-0.05*						92.23
	(-1.39)	(6.16)	(34.88)	(10.52)	(-1.88)						
LSI	0.70***	-0.01	0.24	0.16	-0.01						4.80
	(5.36)	(-0.38)	(1.26)	(1.53)	(-0.09)						
OR	-0.04	0.06***	1.20***	0.25***		-0.01	0.00				92.43
	(-0.98)	(4.54)	(54.56)	(9.31)		(-0.21)	(-0.02)				
LSI	0.76***	-0.12***	0.03	-0.09		0.13	-0.03				2.37
	(5.98)	(-2.77)	(0.36)	(-0.99)		(1.12)	(-0.47)				
OR	-0.03	0.06***	1.20***	0.24***	-0.02	0.00	0.00				92.46
	(-0.66)	(4.35)	(55.58)	(8.32)	(-1.20)	(-0.05)	(0.09)				
LSI	0.82***	-0.13***	0.03	-0.14	-0.09*	0.16	-0.01				3.43
	(6.18)	(-3.22)	(0.43)	(-1.52)	(-1.82)	(1.49)	(-0.15)				
OR	0.12**	0.02	1.13***					0.02	-0.11***	-0.14***	91.55
	(2.41)	(1.64)	(38.59)					(0.62)	(-3.50)	(-3.83)	
LSI	0.80***	-0.13***	-0.02					-0.03	-0.16*	0.07	2.99
	(5.01)	(-2.88)	(0.32)					(-0.36)	(-1.82)	(0.69)	

Appendix B

Chapter 2 Supplement

B1 Additional Explanations of the Theoretical Analysis

B1.1 Rationality of the analytical framework

The major finding should be generated from a sufficiently general framework to ensure that it is an intrinsic feature of LBA but not specific to certain economies. The analytical framework of this study promotes this generality for three reasons. First, a factor-model framework naturally unveils the information content of LBA as betas, defined under this framework, determine what LBA captures. Second, distinguishing investors' and econometricians' information mimics the natural process of how the literature discovers LBA. Third, setting the ideal model with all pricing information as a multifactor model (i.e., Equation 2.1) can nest different formats of information relevant to pricing.

Setting Equation 2.1 as a model with a market factor and additional unknown factors is reasonable for the following reasons. A factor model captures the structure of asset returns without relying on specific constraints, as it is not deduced from an equilibrium economy. Although a model like an APT-type model does not specify what factors to include (e.g., Ross 1976; Chamberlain and Rothschild 1983), it is without loss of generality to set the first factor as the market factor, given the consensus that the first principal component extracted from return comovements is an aggregate risk factor (e.g., Lettau and Pelger 2020a,b; Haddad et al. 2020; Cooper et al. 2021). The remaining factors in Equation (2.1), $\boldsymbol{D}_t^o$ and $\boldsymbol{C}_t^o$, are factors whose economic origins are not specifically specified, which enables them to capture different economic sources in a general way.

A multifactor model can nest different formats of pricing information in a general way without involving specific economic details. Dynamic models can be summarized by adding additional factors to the market factor (e.g., Jagannathan and Wang 1996). Under the same insight, other violations of the standard CAPM assumptions can be captured by additional factors. Characteristic-related pricing can be summarized by including cross-sectional factors in a factor model (e.g., Fama and French 2020). A multifactor model can also capture the scenario of Roll (1977)'s critique, as a factor can be added to a factor model to capture the inefficiency component of a factor-mimicking portfolio (e.g., Daniel et al. 2020b). It is also a general practice to summarize behavioral-based return sources by factors (e.g., Daniel et al. 2020a). The rationale is provided by the finding in Stambaugh and Yuan (2017) that there is a common component of mispricing and the finding in Kozak et al. (2018) that investors' sentiments must align with factor exposures to be relevant for pricing. To the least, we can add infinitely many factors to a multifactor model if one does not agree that asset returns follow a succinct factor structure.

B1.2 Benefit of the Transformation in Section 2.3.1

Observation 1: *A multifactor model can be transformed so that the time-series correlations between factor returns subsume the cross-sectional correlations between factor exposures without changing the information the model captures.*

The transformation in Section 2.3.1 implies that, without changing the pricing information, we can adjust a factor model so that the times-series correlations between factor returns sufficiently capture all interactions between different return sources. We can understand this benefit by using the theoretical prediction of Frazzini and Pedersen (2014) as an example. The basic prediction of their model (their Equation 8) is that a two-factor model summarizes asset returns, that is,

$$r_{it} - r_f = \beta_i^o M_t + (1 - \beta_i^o)\psi_t + \epsilon_{it} \tag{B.1}$$

where ψ_t captures the leverage-constraint effect.

Obviously, the exposures to ψ_t and M_t are correlated in the cross-section. Following *Observation 1*, we can transform Equation (B.1) to

$$r_{it} - r_f = M_t + (1 - \beta_i^o)(\psi_t - M_t) + \epsilon_{it} \tag{B.2}$$

The case of Frazzini and Pedersen (2014) is special as assets' exposures to the market factor all equal one in the transformed model (while in general, assets' market exposures under the transformed model should still be different). However, the rationale of *Observation 1* still applies. In this case, the time-series correlation between M_t and $\psi_t - M_t$ fully captures the interaction between market risk and leverage-constraint effect; hence, there is no cross-sectional correlation between exposures to the two transformed factors.

Under this transformed model (Equation B.2), we can regard the market factor (M_t) as the driver of the level of returns (see Kozak et al. 2018 and Haddad et al. 2020 for the same view) and the leverage-constraint factor ($\psi_t - M_t$) as the driver of the cross-sectional difference in returns. To analyze how the leverage-constraint effect impacts the CAPM alpha, analyzing $Cov(M_t, \psi_t - M_t)$ under the transformed model leads to the same conclusion as analyzing $Cov_{CS}(\beta_i^o, 1 - \beta_i^o)$ under the original model (suppose $Cov(M_t, \psi_t) = 0$ holds). If $Cov(M_t, \psi_t) \neq 0$ holds, analyzing the transformed model will greatly simplify the analysis.

B1.3 Simplifying Equation (2.4)

This appendix section demonstrates how to simplify Equation (2.4) without affecting the information it captures. I start by rewriting Equation (2.4) as:

$$r_{it} - r_f = \alpha_i + \beta_i M_t + \frac{\boldsymbol{d}_i^{oT} \boldsymbol{l}_K}{K} \times \frac{K}{\boldsymbol{d}_i^{oT} \boldsymbol{l}_K} \boldsymbol{d}_i^{oT} \boldsymbol{D}_t + \frac{\boldsymbol{c}_i^{oT} \boldsymbol{l}_S}{S} \times \frac{S}{\boldsymbol{c}_i^{oT} \boldsymbol{l}_S} \boldsymbol{c}_i^{oT} \boldsymbol{C}_t + \epsilon_{it} \quad \text{(B.3)}$$

where $\boldsymbol{l}_K$ is a K-by-1 vector of ones and $\boldsymbol{l}_S$ is an S-by-1 vector of ones. Using the following notations

$$\begin{aligned} \frac{\boldsymbol{d}_i^{oT}\boldsymbol{l}_K}{K} = d_i;\ \frac{K}{\boldsymbol{d}_i^{oT}\boldsymbol{l}_K}\boldsymbol{d}_i^{oT}\boldsymbol{D}_t = D_t; \\ \frac{\boldsymbol{c}_i^{oT}\boldsymbol{l}_S}{S} = c_i;\ \frac{S}{\boldsymbol{c}_i^{oT}\boldsymbol{l}_S}\boldsymbol{c}_i^{oT}\boldsymbol{C}_t = C_t \end{aligned} \tag{B.4}$$

we can get Equation (2.5); that is, we can use a linear combination of factors in $\boldsymbol{D}_t$ (or $\boldsymbol{C}_t$) to replace these factors without affecting the information captured by the model. Equation (2.5) and (2.4) span the same SDF as they lead to the same maximum Sharpe ratio combination or multifactor-efficient portfolio (e.g., Hansen and Jagannathan 1991; Fama 1996).

Moreover, transforming Equation (2.4) to Equation (2.5) does not affect any analytical conclusions. As shown by Equation (2.9), LBA depends on the information captured by estimated betas.

$$\hat{\beta}_i = \beta_i + \frac{Cov(D_t, M_t)}{Var(M_t)} \times d_i \tag{B.5}$$

Replacing D_t and d_i by $\frac{K}{\boldsymbol{d}_i^{oT}\boldsymbol{l}_K}\boldsymbol{d}_i^{oT}\boldsymbol{D}_t$ and $\frac{\boldsymbol{d}_i^{oT}\boldsymbol{l}_K}{K}$, we have

$$\hat{\beta}_i = \beta_i + \frac{Cov(\frac{K}{\boldsymbol{d}_i^{oT}\boldsymbol{l}_K}\boldsymbol{d}_i^{oT}\boldsymbol{D}_t, M_t)}{Var(M_t)} \times \frac{\boldsymbol{d}_i^{oT}\boldsymbol{l}_K}{K} = \beta_i + \frac{Cov(\boldsymbol{D}_t^T, M_t)}{Var(M_t)} \times \boldsymbol{d}_i^o \tag{B.6}$$

where $\beta_i + \frac{Cov(\boldsymbol{D}_t^T, M_t)}{Var(M_t)} \times \boldsymbol{d}_i^o$ is the same as the information captured by $\hat{\beta}_i$ if analyzing Equation (2.4) directly.

B1.4 Decomposing the betting-against-beta (BAM) return

This appendix section decomposes the returns of two BAM portfolios, BAM_t and BAM_t^*. BAM_t is constructed by econometricians before the resolution of LBA, and BAM_t^* is constructed by econometricians conditioning on the resolution of the known LBA.

Decomposing the pre-resolution econometricians' betting-against-beta (BAM) return

Econometricians, before the resolution of the LBA known in the literature, construct a BAM portfolio using the market betas estimated from the CAPM following Equation (2.10). I start by replacing individual stock returns in Equation (2.10) with Equation (2.5):

$$BAM_t = \sum_{\{i:\ \hat{\beta}_i<\beta_L\}} w_i(\alpha_i+\beta_i M_t+d_i D_t+c_i C_t+\epsilon_{it}) - \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} w_i(\alpha_i+\beta_i M_t+d_i D_t+c_i C_t+\epsilon_{it}) \tag{B.7}$$

The idiosyncratic terms of the long and short sides cancel out as they are uncorrelated with factor exposures in the cross-section (see Equation 2.2). As a result, their values should be similar for different quantiles of beta-sorted portfolios, that is,

$$\sum_{\{i:\ \hat{\beta}_i<\beta_L\}} w_i(\alpha_i+\epsilon_{it}) - \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} w_i(\alpha_i+\epsilon_{it}) \approx 0 \tag{B.8}$$

The observation is that even if there exist returns unrelated to factors, they will not be picked up by estimated betas in the cross-section.

Next, I replace D_t and C_t by their connections with the market factor (Equation 2.7):

$$\begin{aligned} BAM_t =&\Big(\sum_{\{i:\ \hat{\beta}_i<\beta_L\}} w_i\beta_i - \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} w_i\beta_i\Big)M_t + \Big(\sum_{\{i:\ \hat{\beta}_i<\beta_L\}} w_i d_i - \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} w_i d_i\Big)\alpha_D \\ &+ \Big(\sum_{\{i:\ \hat{\beta}_i<\beta_L\}} w_i d_i - \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} w_i d_i\Big)\rho_D M_t + \Big(\sum_{\{i:\ \hat{\beta}_i<\beta_L\}} w_i d_i - \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} w_i d_i\Big)u_{D,t} \\ &\Big(\sum_{\{i:\ \hat{\beta}_i<\beta_L\}} w_i c_i - \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} w_i c_i\Big)\alpha_C + \Big(\sum_{\{i:\ \hat{\beta}_i<\beta_L\}} w_i c_i - \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} w_i c_i\Big)\rho_C M_t \\ &+ \Big(\sum_{\{i:\ \hat{\beta}_i<\beta_L\}} w_i c_i - \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} w_i c_i\Big)u_{C,t} \end{aligned} \tag{B.9}$$

Given how pre-resolution econometricians' betas capture true factor exposures (Equation 2.9), the above equation is simplified as

$$\begin{aligned} BAM_t =&\Big(\sum_{\{i:\ \hat{\beta}_i<\beta_L\}} w_i d_i - \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} w_i d_i\Big)\alpha_D + \Big(\sum_{\{i:\ \hat{\beta}_i<\beta_L\}} w_i c_i - \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} w_i c_i\Big)\alpha_C \\ &+ \Big(\sum_{\{i:\ \hat{\beta}<\beta_L\}} w_i\hat{\beta}_i - \sum_{\{i:\ \hat{\beta}>\beta_H\}} w_i\hat{\beta}_i\Big)M_t \\ &+ \Big(\sum_{\{i:\ \hat{\beta}_i<\beta_L\}} w_i d_i - \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} w_i d_i\Big)u_{D,t} + \Big(\sum_{\{i:\ \hat{\beta}_i<\beta_L\}} w_i c_i - \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} w_i c_i\Big)u_{C,t} \end{aligned} \tag{B.10}$$

When adjusting the BAM portfolio's market exposure, the first two elements on the right-hand side are identified as the BAM alpha, the third element is the explained component of the BAM return, and the last two elements are identified as the residual BAM return.

As shown by Equation Equation 2.9, how the exposure to a factor is captured by an estimated beta depends on how that factor relates to the market factor. Suppose $\rho_C \approx 0$ holds. The market-adjusted BAM return becomes

$$BAM_t - \left(\sum_{\{i:\ \hat{\beta}<\beta_L\}} w_i \hat{\beta}_i - \sum_{\{i:\ \hat{\beta}>\beta_H\}} w_i \hat{\beta}_i \right) M_t = \left(\sum_{\{i:\ \hat{\beta}_i<\beta_L\}} w_i d_i - \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} w_i d_i \right) (\alpha_D + u_{D,t}) \tag{B.11}$$

This equation shows that a BAM portfolio, besides negatively capturing the market return, also captures the return of the omitted factors directly related to market risk.

Impact of the weighting scheme

As long as it does not capture effects such as the microcap effect, the weighting scheme used for constructing a BAM portfolio should not consequentially affect the risk-return trade-off of an observed LBA. In other words, a BAM alpha's sign and information ratio are not notably affected by the weighting scheme as long as the BAM portfolio has low/high-$\hat{\beta}_i$ stocks in its long/short side.

To illustrate the impact of weighting scheme, I assume that low estimated market betas ($\hat{\beta}_i$) are associated with higher d_is (that is, $Cov(D_t, M_t) < 0$ and $Cov(C_t, M_t) \approx 0$) and rewrite the omitted factor exposure of the BAM portfolio as

$$\beta_D = \sum_{\{i:\ \hat{\beta}_i<\beta_L\}} w_i d_i - \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} w_i d_i = \sum_{\{i:\ \hat{\beta}_i<\beta_L\}} \tilde{w}_i d_i + \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} \tilde{w}_i d_i = \tilde{\boldsymbol{w}}^T \boldsymbol{d} = N \times E_{CS}(\tilde{w}_i d_i) \tag{B.12}$$

where $\tilde{\boldsymbol{w}}$ is the vector of weights and $\boldsymbol{d}$ is the vector of individual stocks' exposures to D_t; $\tilde{w}_i$ is positive (negative) for stocks with $\hat{\beta}_i < \beta_L$ ($\hat{\beta}_i > \beta_H$); N is the number of stocks in the BAM portfolio; E_{CS} denotes the cross-sectional expectation. Decompose the BAM portfolio's exposure to D_t into two components:

$$\beta_D = N \times [E_{CS}(\tilde{w}_i)E_{CS}(d_i) + Cov_{CS}(\tilde{w}_i, d_i)] \tag{B.13}$$

where Cov_{CS} denotes the cross-sectional covariance. Since the weights of stocks in the zero-investment BAM portfolio add up to zero, $E_{CS}(\tilde{w}_i) = 0$, the BAM portfolio's exposure to D_t is entirely determined by the second component.

A BAM portfolio by construction has high-d_i stocks in the long side and low-d_i stocks in the short side; in other words, high-d_i tends to receive higher weights than low-d_i stocks, $Cov_{CS}(\tilde{w}_i, d_i) > 0$; this relationship becomes definitive when a large number of stocks are involved (see Figure 2.3). Therefore, the sign of the BAM alpha is determined by the relationship between D_t and M_t. The information ratio of the BAM alpha is not affected by the weighting scheme since it is determined by α_D and $u_{D,t}$, while the weighting scheme only affects the magnitude of $(\sum_{\{i:\ \hat{\beta}_i<\beta_L\}} w_i d_i - \sum_{\{i:\ \hat{\beta}_i>\beta_H\}} w_i d_i)$. Therefore, if not picking up microcap effects, the weighting scheme should not consequentially impact the betting-against-beta performance.

Decomposing the post-resolution econometricians' betting-against-beta (BAM) return

The appendix section decomposes the BAM portfolio constructed by post-resolution econometricians. Econometricians estimate betas using Equation (2.14) and construct a BAM portfolio as

$$BAM_t^* = \sum_{\{i:\ \hat{\beta_i}^* < \beta_L\}} w_i r_{it} - \sum_{\{i:\ \hat{\beta_i}^* > \beta_H\}} w_i r_{it} \tag{B.14}$$

Replace r_{it} by Equation (2.5) to get

$$\begin{aligned} BAM_t^* = &\Big(\sum_{\{i:\ \hat{\beta}_i^* < \beta_L\}} w_i \beta_i - \sum_{\{i:\ \hat{\beta}_i^* > \beta_H\}} w_i \beta_i \Big) M_t + \Big(\sum_{\{i:\ \hat{\beta}_i^* < \beta_L\}} w_i d_i - \sum_{\{i:\ \hat{\beta}_i^* > \beta_H\}} w_i d_i \Big) D_t \\ &+ \Big(\sum_{\{i:\ \hat{\beta}_i^* < \beta_L\}} w_i c_i - \sum_{\{i:\ \hat{\beta}_i^* > \beta_H\}} w_i c_i \Big) C_t + \Big(\sum_{\{i:\ \hat{\beta}_i^* < \beta_L\}} w_i \epsilon_{it} - \sum_{\{i:\ \hat{\beta}_i^* > \beta_H\}} w_i \epsilon_{it} \Big) \end{aligned} \tag{B.15}$$

To decompose BAM_t^*, we need to know the composition of estimated factors exposures. Rewrite Equation (2.14) as

$$\begin{aligned} r_{it} - r_f &= \hat{\alpha}_i^* + \hat{\beta}_i^* (M_t - \rho_D D_t) + (\hat{\beta}_i^* \rho_D + \hat{d}_i^*) D_t + \hat{\epsilon}_{it}^* \\ &= \hat{\alpha}_i^* + \hat{\beta}_i^* M_t^\epsilon + (\hat{\beta}_i^* \rho_D + \hat{d}_i^*) D_t + \hat{\epsilon}_{it}^* \end{aligned} \tag{B.16}$$

where M_t^ϵ is the component of M_t orthogonal to D_t. Given that $Cov(M_t^\epsilon, D_t) = 0$ and replacing $r_{it} - r_f$ by Equation (2.5), we have

$$\begin{aligned} \hat{\beta}_i^* = &\frac{Cov(r_{it} - r_f, M_t^\epsilon)}{Var(M_t^\epsilon)} = \frac{Cov(\alpha_i + \beta_i M_t + d_i D_t + c_i C_t + \epsilon_{it}, M_t^\epsilon)}{Var(M_t^\epsilon)} = \beta_i + c_i \frac{Cov(C_t, M_t^\epsilon)}{Var(M_t^\epsilon)} \\ &= \beta_i + c_i \frac{Cov(C_t, M_t - \rho_D D_t)}{Var(M_t^\epsilon)} = \beta_i + c_i \frac{Var(M_t)}{Var(M_t^\epsilon)} \frac{Cov(C_t, M_t - \rho_D D_t)}{Var(M_t)} \end{aligned} \tag{B.17}$$

This equation leads to Equation (2.15). Then, we have

$$\beta_i = \hat{\beta}_i^* - \frac{Var(M_t)}{Var(M_t^\epsilon)}(\rho_C - \rho_D\rho_{DC})c_i \tag{B.18}$$

where M_t^ϵ is the component of M_t orthogonal to D_t.

Similarly, we can estimate $\hat{d}_i^*$ from Equation (2.14) and get d_i

$$d_i = \hat{d}_i^* - \frac{Var(D_t)}{Var(D_t^\epsilon)}(\rho_{DC} - \rho_D\rho_C)c_i \tag{B.19}$$

where ρ_D, ρ_C and ρ_{DC} are the correlation coefficients between factors defined in Equation (2.7); D_t^ϵ is the component of D_t orthogonal to M_t. Based on Equations (2.7), (B.18) and (B.19), C_t can be decomposed as

$$C_t = \alpha_C^* + \frac{Var(M_t)}{Var(M_t^\epsilon)}(\rho_C - \rho_D\rho_{DC})M_t + \frac{Var(D_t)}{Var(D_t^\epsilon)}(\rho_{DC} - \rho_D\rho_C)D_t + u_{C,t}^* \tag{B.20}$$

where $E(u_{C,t}^*,\ [M_t\ D_t]) = [0\ 0]$ holds.

Combining Equations (B.15), (B.18), (B.19) and (B.20) leads to the new BAM portfolio's return

$$\begin{aligned} BAM_t^* = &\Big(\sum_{\{i:\ \hat{\beta}_i^*<\beta_L\}} w_i c_i - \sum_{\{i:\ \hat{\beta}_i^*>\beta_H\}} w_i c_i\Big)\alpha_C^* + \Big(\sum_{\{i:\ \hat{\beta}_i^*<\beta_L\}} w_i \hat{\beta}_i^* - \sum_{\{i:\ \hat{\beta}_i^*>\beta_H\}} w_i \hat{\beta}_i^*\Big)M_t \\ &+ \Big(\sum_{\{i:\ \hat{\beta}_i^*<\beta_L\}} w_i \hat{d}_i^* - \sum_{\{i:\ \hat{\beta}_i^*>\beta_H\}} w_i \hat{d}_i^*\Big)D_t + \Big(\sum_{\{i:\ \hat{\beta}_i^*<\beta_L\}} w_i c_i - \sum_{\{i:\ \hat{\beta}_i^*>\beta_H\}} w_i c_i\Big)u_{C,t}^* \end{aligned} \tag{B.21}$$

The second and third elements on the right-hand side will be subsumed in a spanning regression of BAM_t^* on M_t and D_t under the assumption that factor exposures are persistent. The BAM portfolio constructed by post-resolution econometricians leads to a new LBA that only captures the component of C_t orthogonal to M_t and D_t, that is, $\alpha_C^* + u_{C,t}^*$.

B2 Additional Explanations of the Empirical Analysis

B2.1 Comparing the known and new low-beta anomalies (LBAs): ex-ante factor exposures and ex-post alphas

This appendix subsection presents the empirical procedures for generating Figure 2.1, which examines the ex-ante factor exposures and ex-post alphas of $\hat{\beta}_i$- and $\hat{\beta}_i^*$-sorted portfolios. $\hat{\beta}_i$s are estimated from the CAPM following Equation(2.8). $\hat{\beta}_i^*$s are estimated from the six-factor model (Fama and French 2018) following Equation(2.14). At the beginning of each time t, I sort stocks into $\hat{\beta}_i$ or $\hat{\beta}_i^*$ quintiles according to NYSE breakpoints and record ex-ante factor exposures of each quintile on a value-weighted basis. Value-weighted returns of each quintile are recorded at the end of time t. Profitability and investment factor exposures (b_i^{RMW} and b_i^{CMA}) are estimated from the six-factor model.

The left panel of Figure 2.1 reports the results of $\hat{\beta}_i$-sorted portfolios. Ex-ante $\hat{\beta}_i$s reported in the top-left subfigure exhibit an ascending pattern from Q1 to Q5 since these portfolios are constructed by sorting on $\hat{\beta}_i$s. The bottom-left subfigure reports the ex-post alphas (monthly, %) against the CAPM model (α_M). The negative beta-alpha relationship is referred to as LBA in the literature. The middle two subfigures in the left panel report the ex-ante exposures to investment and profitability factors (d_i^{RMW} and d_i^{CMA}). The way that $\hat{\beta}_i$s capture d_i^{RMW}s and d_i^{CMA}s is consistent with Proposition 1; that is, low (high) estimated betas pick up high (low) exposures to other factors when those factors negatively correlate with the market factor. Panel C of Table 2.1 confirms that investment and profitability factors negatively correlate with the market factor.

Corollary 2.4.2 suggests that a new LBA should be uncovered by conditioning on factors that fully explain the known LBA. Accordingly, I estimate $\hat{\beta}_i^*$s following Equation (2.14) to remove the impact of these two factors, and then use $\hat{\beta}_i^*$s to analyze the beta-alpha relationship by repeating the above procedures. The top-right subfigure of Figure 2.1 reports the value-weighted $\hat{\beta}_i^*$s of each quintile, which exhibits an ascending pattern by construction. The bottom-right subfigure reports each quintile portfolio's ex-post spanning regression alphas (monthly, %) against the six-factor model (α_{M+D}). The descending pattern from Q1 to Q5 shows that LBA still exists. However, as shown by the middle two subfigures in the right panel, different $\hat{\beta}_i^*$ stocks have similar exposures to the investment and profitability factors. The difference between $\hat{\beta}_i$- and $\hat{\beta}_i^*$-sorted portfolios indicates that the six-factor model no longer explains the LBA identified in the right panel. Therefore, a new LBA with different economic content from the known LBA exists.

B2.2 Comparing the known and new low-beta anomalies (LBAs): cumulative performance and realized returns

This appendix subsection presents the empirical procedures for generating Figure 2.2. The LBA known in the literature is captured by the residual return of regressing BAM_t against the CAPM. The new LBA is captured by the residual return of regressing BAM_t^* against the six-factor model of Fama and French (2018). BAM_t and BAM_t^* are the same as those in Table 2.5. The top of Figure 2.2 plots the cumulative performance of the two anomalies. The shaded bars correspond to NBER recession periods.

One may argue that the performance is not practically meaningful since, for the two portfolios in Figure 2.2, the full-sample information is used when removing their factor exposures (i.e., look-ahead bias). To address this concern, I also report the cumulative returns of BAB-type portfolios; that is, the BAM portfolios' market exposures are removed using ex-ante information (see Appendix B2.4 for details). BAB_t captures the known LBA and BAB_t^* captures the new LBA. The BAB-type portfolios can be regarded as real-time investment strategies. As shown by Figure B2.1, the performance of the BAB portfolios is similar to that of the BAM portfolios reported in Figure 2.2. The opposite performance indicates that the new LBA differs from the LBA known in the literature.

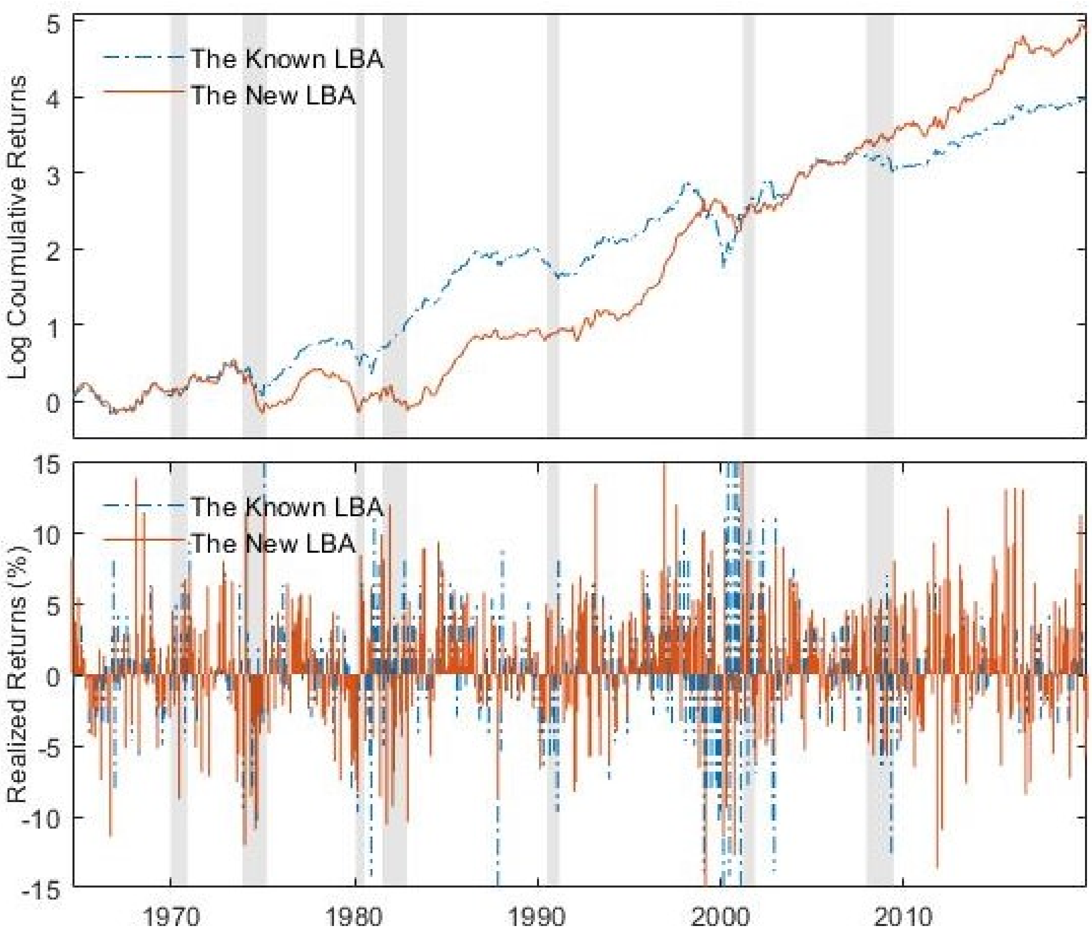
B2.3 Results based on different beta-estimation methods

This appendix section examines whether different beta-estimation methods affect the discovery of the new LBA. I estimate betas using daily and monthly returns of different horizons, including daily returns of the prior three months, daily returns of the prior six months, daily returns of the prior two years and monthly returns of the prior five years.

I also estimate betas following Frazzini and Pedersen (2014), under which correlations and volatilities are estimated separately. Under this method, the first step

FIGURE B2.1: Performance of the known and new Low-Beta Anomalies (LBAs)

The figure reports the log of the cumulative returns of the known LBA (the dotted lines) and new LBA (the solid lines) based on the BAB-type portfolios. BAB_t captures the known LBA and BAB_t^* captures the new LBA. BAB_t and BAB_t^* are the market-neutral portfolios constructed by leveraging up/down the low/high-beta side so that the ex-ante CAPM-beta or market-beta (estimated from the six-factor model) is zero (see Appendix B2.4). The shaded bars correspond to NBER recession periods. The sample period is from July 1968 to December 2019.



is estimating correlations between individual stocks and factors using overlapping three-trading-day log returns over a five-year window. The second step is estimating the volatilities of individual stocks and factors using daily log returns over a one-year window. The third step is combining the estimated correlations and volatilities to calculate betas. Using overlapping returns in the first step considers the non-synchronicity issue of trading. Separately estimating correlations and volatilities considers the empirical evidence that correlations move more slowly than volatilities. When estimating market betas from a multi-factor model, I control for other factors in correlation estimation following Daniel et al. (2020b).

Under each of the above beta-estimation methods, I estimate market betas from the CAPM and the six-factor model (Fama and French 2018) and construct corresponding BAM portfolios following Equation (3.1). Table B2.1 reports the performance of these BAM portfolios against the six-factor model. The first column presents the data frequency and horizon used for beta estimation. The second column reports the BAM portfolio examined. The remaining columns report time-series regression alphas, factor exposures and adjusted R-squares. The results lead to the same conclusion as that in Section 2.5.2: while the six-factor model explains the low-beta anomaly known in the literature, there is a new LBA different from the known LBA and unexplained by the six-factor model.

B2.4 Construction of the BAB-type portfolios

This appendix section illustrates the construction of a BAB-type portfolio following Frazzini and Pedersen (2014). Specifically, returns of a BAB portfolio's long and short sides are calculated on a value-weighted basis following Novy-Marx and

Table B2.1: The Known and New Low-Beta Anomalies (LBAs) Based on Different Betas

This table presents the spanning regression alphas (α), factor exposures and adjusted R-squares ($\bar{R}^2$) of BAM_t and BAM_t^* against the six-factor model (Fama and French 2018). BAM refers to the betting-against-beta portfolio constructed following Equation (3.1). BAM_t and BAM_t^* are respectively constructed using market betas estimated from the CAPM and six-factor model. The first column presents the data frequency and horizon for estimating betas. Betas for the last two rows are estimated following Frazzini and Pedersen (2014), which estimate correlation coefficients and variances separately. MKT, SMB, HML, RMW, CMA and MOM are the market, size, value, profitability, investment and momentum factors from Kenneth French's website. The t-statistics are based on Newey and West (1987) adjusted standard errors. The sample period is from July 1968 to December 2019.

Beta Estimation	BAM	α(%)	MKT	SMB	HML	RMW	CMA	MOM	$\bar{R}^2$(%)
Daily, 3-Month	BAM_t	-0.10	-1.00	-0.35	0.41	0.95	0.80	0.14	64.66
		(-0.39)	(-10.89)	(-3.34)	(2.25)	(6.58)	(3.63)	(0.96)	
	BAM_t^*	0.77	-1.00	-0.36	0.14	0.18	-0.17	0.02	45.40
		(2.97)	(-13.37)	(-3.89)	(1.00)	(1.57)	(-0.75)	(0.25)	
Daily, 6-Month	BAM_t	-0.05	-1.00	-0.37	0.40	0.87	0.77	0.13	68.08
		(-0.21)	(-11.67)	(-3.99)	(2.20)	(7.42)	(3.66)	(1.14)	
	BAM_t^*	0.84	-1.00	-0.46	0.20	0.16	-0.28	-0.01	50.50
		(3.56)	(-13.93)	(-4.95)	(1.57)	(1.27)	(-1.33)	(-0.20)	
Daily, 24-Month	BAM_t	0.16	-1.00	-0.48	0.43	0.77	0.71	0.16	70.76
		(0.71)	(-13.27)	(-5.77)	(2.46)	(5.64)	(3.61)	(1.48)	
	BAM_t^*	0.76	-1.00	-0.57	0.18	0.19	-0.15	0.13	60.26
		(3.84)	(-16.68)	(-7.23)	(1.53)	(1.86)	(-0.80)	(2.09)	
Monthly, 5-Year	BAM_t	-0.01	-1.00	-0.88	0.49	0.85	0.59	0.37	74.60
		(-0.04)	(-13.56)	(-10.28)	(3.45)	(6.34)	(3.83)	(3.09)	
	BAM_t^*	0.55	-1.00	-0.60	0.23	-0.06	-0.68	0.32	51.01
		(2.05)	(-13.11)	(-6.15)	(1.40)	(-0.50)	(-3.06)	(2.90)	
Frazzini and	BAM_t	0.39	-1.00	-0.64	0.39	0.75	0.37	0.26	70.18
Pedersen (2014)		(1.72)	(-13.94)	(-7.29)	(2.33)	(4.91)	(2.10)	(2.35)	
	BAM_t^*	0.91	-1.00	-0.73	-0.05	0.04	-0.45	0.27	59.46
		(4.40)	(-15.55)	(-8.30)	(-0.41)	(0.38)	(-2.39)	(3.06)	

Velikov (2022) to ensure that extremely small stocks do not drive the portfolio performance. A BAB portfolio is constructed by leveraging up the low-beta side and leveraging down the high-beta side so that the ex-ante beta of the betting-against-beta portfolio is zero, as shown by Equation (2.19). Due to the persistence of betas, the ex-post market exposure of such a portfolio is also near zero.

I construct a BAB_t portfolio using market betas estimated from the CAPM and a BAB_t^* portfolio using market betas estimated from the six-factor model. When constructing the BAB portfolio, at the beginning of each time t, I shrink betas (β_i^{TS}) towards their cross-sectional mean (β^{CS}) using the latest estimated betas following Frazzini and Pedersen (2014), $\beta_i = w_i\beta_i^{TS} + (1 - w_i)\beta^{CS}$. This beta-shrinkage procedure ensures that the ex-post market exposure of a BAB portfolio is near zero. I set w_i as 0.6 (same as Frazzini and Pedersen 2014) when constructing the BAB_t portfolio and 0.4 when constructing the BAB_t^* portfolio (or BAB portfolios based on betas from three- or four-factor model). Within a reasonable range, the value of w_i does not have a consequential impact on the BAB portfolio performance. For example, if setting w_i as 0.6 when constructing BAB_t^*, BAB_t^* will be positively exposed to the market factor. However, this BAB portfolio still generates abnormal returns after adjusting for market risk. For ease of comparison, I scale the BAB portfolios so that they have the same variance as the market portfolio.

B3 Simulation Examples

This section provides simulation evidence as illustrative examples to support the propositions in Sections 2.3 and 2.4.

B3.1 Simulation examples of the changeable nature of the low-beta anomaly (LBA)

This subsection provides simulation examples to intuitively demonstrate how the economic content of LBA changes.

Simulation setup

Following Equation (2.5), I set the data-generating process as a three-factor model with the market, D and C factors:

$$DGP: \begin{cases} r_{it} - r_f = \beta_i M_t + d_i D_t + c_i C_t + \epsilon_{it} \\ D_t = \alpha_D + \rho_D M_t + u_{D,t} \\ C_t = \alpha_C + \rho_C M_t + u_{C,t} \\ D_t = \alpha_{DC} + \rho_{DC} C_t + u_{DC,t} \end{cases} \qquad i = 1, \cdots, N \text{ and } t = 1, \cdots, T \tag{B.22}$$

where $r_{it} - r_f$ is the excess return of security i at time t; β_i, d_i and c_i are the true exposures to these factors; $\epsilon_{it} \sim N(\alpha_i, \sigma_\epsilon^2)$ is the residual return. The market factor (M_t) and the omitted factors (D_t and C_t) have the same expected return and variance.[1] Following MacKinlay and Pástor (2000), I set the average returns and standard deviations of these factors to be the same as those of the market factor (monthly) over the sample period ($\mu_M = \mu_D = \mu_C = 0.54\%$, $\sigma_M = \sigma_D = \sigma_C = 4.39\%$). Factor returns are set to follow normal distributions. Factor exposures follow normal distributions in the cross-sectional, $\beta_i \sim N(1, 0.5)$, $d_i \sim N(0,1)$ and $c_i \sim N(0,1)$. The cross-sectional variance of d_i and c_i is two times that of

[1] In an unreported test, I confirm that the conclusion does not change if we set returns to follow t distributions (e.g., Gospodinov and Robotti 2021)

β_i to reflect the empirical evidence that other factors usually explain more cross-sectional return differences than the market factor. The residual variance (σ_ϵ^2) is set as two times the factor variance to reflect that the idiosyncratic part of the individual stock variance is usually larger than the systematic part. The values of T and N affect the likelihood of observing a specific simulation result. For example, if T is large, the results are less affected by the random realization of returns. I set T as 618 and N as 3000. The former is to mimic that there are 618 months over the sample period, and the latter is to mimic that there are 3000 stocks on average at each time point. These parameter values do not materially impact the conclusions within a wide but reasonable range.

I explicitly consider the correlation between the two omitted factors (ρ_{DC}) here since, after updating the CAPM model with D_t, to what extent estimated market betas capture c_i is affected by ρ_{DC}. After setting the values of ρ_D, ρ_C and ρ_{DC}, I generate factors returns and individual stock returns under Equation (C.20) and then conduct the simulation analysis mimicking empiricists before and after the resolution of the known LBA. First, I estimate market betas using Equation (2.8) and analyze the $\hat{\beta}_i$-$\hat{\alpha}_i$ relationship. Second, I estimate market betas using Equation (2.14) and analyze the $\hat{\beta}_i^*$-$\hat{\alpha}_i^*$ relationship. Appendix B3.3 describes the simulation procedures in detail. I consider two scenarios: first, C_t does not contribute to the known LBA but induces a new LBA; second, C_t also contributes to the known LBA, but its impact is overwhelmed by D_t's. Under both scenarios, D_t is the return source underlying the known LBA, while C_t is the return source underlying the new LBA. The new LBA is weaker than the known LBA in the second scenario but can be stronger in the first scenario.

The limitation of estimated betas before the resolution of the known low-beta anomaly (LBA)

This subsection provides the details for generating Figure 2.4. Under the data-generating process of Equation (C.20), the correlations are set as

$$\rho_D = -0.5;\ \rho_C = \delta|\rho_D|;$$

where ρ_D is set to be negative to mimic the known LBA. $\delta \sim [-0.9, 0.9]$ captures the the strength of C_t's connection with the market relative to D_t's. For brevity, ρ_{DC} is not considered in this subsection.

Under the above setup, I choose a value of δ and then generate true factor returns and individual stock returns following Appendix B3.3. With the generated data, I estimate market betas using Equation (2.8) (the CAPM), sort stocks into three quantiles according to estimated market betas ($\hat{\beta}_i$), and construct a BAM portfolio by buying/selling the lowest/highest-$\hat{\beta}_i$ tercile. Next, I record the BAM portfolio's true latent factor exposures (β_M, β_D and β_C) scaled by its estimated market exposure ($\hat{\beta}_M$). Finally, I repeat the procedures for different values of δ, and plot the BAM portfolio's factor exposures against δ.

C_t uncorrelated with the market factor

First, I consider the scenario that C_t is uncorrelated with the market factor by setting the correlation coefficients as

$$\rho_D = -0.5;\ \rho_C = 0;\ \rho_{DC} = -0.5$$

This scenario depicts the economy underlying Figure 2.5. In this subsection, I interpret the results in detail to illustrate how the economic content of LBA changes from D_t to C_t.

The left panel reports the ex-ante factor loadings and ex-post alphas of $\hat{\beta}_i$-sorted portfolios. The third subfigure reports the true market betas (β_i), that is, the exposures of stocks to the market factor if all the factors were known. They exhibit an increasing pattern from Q1 to Q5 since $\hat{\beta}_i$s still capture true market betas well, although they also reflect omitted factor exposures. The bottom two subfigures report exposures to the two omitted factors. Consistent with Proposition 1, sorting stocks into $\hat{\beta}_i$ quantiles picks out stocks with high (low) exposures to D_t in the low(high)-$\hat{\beta}_i$ side when D_t negatively correlates with the market factor. The extra exposures to D_t induce a positive alpha against the CAPM model, as shown in the top subfigure in the left panel. Since C_t is uncorrelated with the market factor, $\hat{\beta}_i$-sorted portfolios do not exhibit a cross-sectional difference in exposures to C_t, which is why the known LBA only reflects the economic content of D_t. Alphas entirely disappear if D_t is controlled in the spanning regression, as shown by the second subfigure in the left panel.

The right panel reports the performance of $\hat{\beta}_i^*$-sorted portfolios. In this case, sorting stocks into $\hat{\beta}_i^*$ quantiles no longer picks up exposures to D_t but instead picks up exposures to C_t, as described by the bottom two subfigures. $\hat{\beta}_i^*$s are able to capture information about C_t because its impact is not squeezed out by D_t once D_t is controlled in beta estimation. The top subfigure reports the alphas of each quintile portfolio against the market-plus-D_t model. The exposures to C_t generate a strong LBA unexplained by D_t, indicating that the new LBA does not reflect the economic content of D_t but is driven by C_t. The new LBA is explained if C_t were known, as shown by the second subfigure. Therefore, D_t is not the full picture of LBA, as C_t can also induce an LBA.

C_t weakly correlated with the market factor

Next, I consider a scenario that C_t are weakly unconditionally correlated with the market, setting the correlation coefficients as

$$\rho_D = -0.6;\ \rho_C = -0.3;\ \rho_{DC} = 0$$

In this case, D_t and C_t both positively contribute to the known LBA as ρ_D and ρ_C are negative. ρ_{DC} is set as zero for brevity.

I follow the same approach as Section B3.1 to interpret the results of Figure B2.2. The top three lines report true latent factor exposures. As shown in the left panel, sorting stocks into $\hat{\beta}_i$ quintiles picks up the cross-sectional difference in exposures to both omitted factors (the second and third subfigures). This pattern arises because both D_t and C_t negatively relate to the market. Consequently, the known LBA is contributed by both (the fourth subfigure). However, as D_t's connection with the market is stronger than C_t's, the known LBA largely disappears when D_t is controlled (the bottom subfigure). The right panel reports the results for portfolios sorted on $\hat{\beta}_i^*$s. From Q1 to Q5, there is no cross-sectional difference in exposures to D_t since it is controlled in beta estimation. Instead, unlike sorting according to $\hat{\beta}_i$s, the cross-sectional difference in exposures to C_t becomes more remarkable when sorting stocks into $\hat{\beta}_i^*$ quintiles. This pattern arises because empiricists' betas can capture information about C_t when the impact of

D_t is squeezed out. Consequently, C_t induces a new LBA orthogonal to D_t (the fourth subfigure). The alphas disappear only if C_t were known, as shown by the bottom-right subfigure.

The two illustrative examples depicted in Figures B2.2 and 2.5 also confirm the necessity to conduct one additional test following the higher bar of Section 2.6.2 to verify whether an explanation for LBA is sufficient.

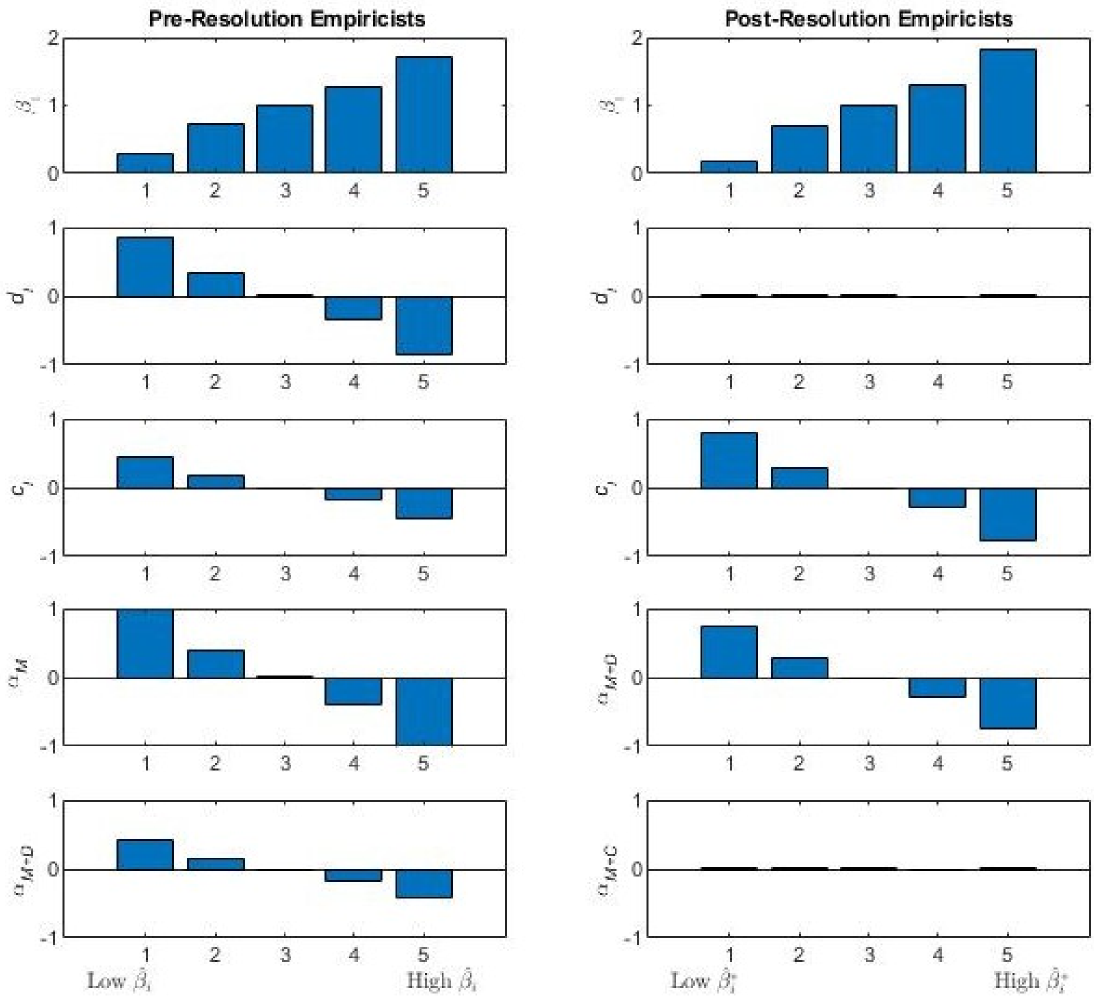
B3.2 Simulation examples of the single omitted factors cases

This subsection investigates 1) how the betting-against-beta (BAM) portfolio captures factor exposures and 2) how the BAM performance changes with the correlation between the market factor and the omitted factor. The model fully summarizing asset returns is set as the market factor plus an omitted factor for ease of exposition:

$$DGP: \begin{cases} r_{it} - r_f = \beta_i M_t + d_i D_t + \epsilon_{it} \\ D_t = \alpha_D + \rho_D M_t + u_{D,t} \end{cases} \quad i = 1, \cdots, N \text{ and } t = 1, \cdots, T \quad \text{(B.23)}$$

Unless mentioned, the parameters as the same as those in Section B3.1. Appendix B3.3 provides the detailed simulation procedures.

Figure B2.2: The Known and New Low-Beta Anomalies (LBAs): $Cov(C_t, M_t) < 0$
The figure depicts how econometricians before and after the resolution of the known LBA observe different LBAs in an economy with both D_t and C_t negatively correlated with M_t, but C_t's connection with the market is weaker than D_t's. In the left and right panels, stocks are sorted into quintiles according to betas from Equations (2.8) and (2.14), respectively. α_M, α_{M+D} and α_{M+C} are the alphas against M_t, $[M_t,\ D_t]$ and $[M_t,\ C_t]$, respectively. β_i, c_i and d_i are each quintile portfolio's true latent factor exposures. Appendix B3.1 provides the details.



How a BAM portfolio's factor exposures change with the number of stocks

This subsection provides the details for generating Figure 2.3. Under the data-generating process of Equation (C.19), the correlation is set as $\rho_D = -0.5$ to mimic the known LBA. Under this setup, for each simulation, I choose the number of stocks (N) and then generate true factor returns and individual stock returns following Equation (C.19). With the generated data, I estimate market betas following Equation (2.8), sort stocks into three quantiles according to estimated market betas ($\hat{\beta}_i$), and construct a BAM portfolio by buying/selling lowest/highest-$\hat{\beta}_i$ terciles. Finally, I observe the BAM portfolio's estimated exposure to M_t and true latent exposures to M_t and D_t. For each number of stocks, I repeat the simulation 1000 times.

In Figure 2.3, the horizontal axis is for "different number of stocks". The vertical line corresponding to each "# of stocks" is composed of 1000 dots, and each dot depicts the result of one simulation. The top subfigure reports the BAM portfolio's market exposure in a spanning regression ($\hat{\beta}_M$), which is negative by construction since the BAM portfolio is constructed by buying/shorting low/high-$\hat{\beta}_i$ stocks. For ease of exposition, I scale the value of each dot by the average $|\hat{\beta}_M|$ of the 1000 simulations corresponding to each horizontal value.

The bottom two subfigures report the BAM portfolio's true factor exposures. The true market exposure (β_M) is what estimated betas ($\hat{\beta}_M$) are supposed to capture. β_D is the BAM portfolio's true latent exposure to the omitted factor. β_D tends to be positive as an estimated beta negatively captures the exposure to D_t. When the number of stocks is small, it happens that a BAM portfolio is negatively exposed to D_t. However, this tendency becomes a definitive relationship when the number of stocks increases, as shown by the positive β_Ds corresponding to larger horizontal-axis values.

The tendency for the low-beta anomaly (LBA) to emerge

To demonstrate the tendency for LBA to emerge (Corollary 2.4.3), I choose a value of ρ_D and then generate true factor returns and individual stock returns following the data-generating process of Equation (C.19). With the generated data, I estimate market betas following Equation (2.8), sort stocks into three quantiles according to estimated market betas ($\hat{\beta}_i$), and construct a BAM portfolio by buying/selling lowest/highest-β terciles. Finally, I observe the BAM portfolio's factor exposures and alphas against the CAPM model.

Figure B2.3 reports the simulation results. The horizontal axis is for different values of ρ_D. The vertical line corresponding to each ρ_D is composed of 1000 dots, and each dot depicts the result of one simulation. The left panel depicts the mechanism for LBA to emerge when the omitted factor negatively correlates with the market factor. The top subfigure reports the BAM portfolio's market exposure in a spanning regression ($\hat{\beta}_M$), which is negative by construction since the BAM portfolio is constructed by buying/shorting low/high-$\hat{\beta}_i$ stocks. I scale the value of each dot by the average $|\hat{\beta}_M|$ of the 1000 simulations corresponding to each ρ so that the results reflect the BAM performance per one-unit negative market exposure. Owing to this scaling, the average $\hat{\beta}_M$ is always minus one.

The middle two subfigures report the BAM portfolio's true factor exposures. The true market exposure (β_M) is what estimated betas ($\hat{\beta}_M$) are supposed to capture. When the correlation between the omitted factor and the market factor is weak (e.g., $\rho_D = -0.1$), the magnitude of β_M is large since $\hat{\beta}_M$ mainly reflects β_M. $|\beta_M|$ decreases with $|\rho_D|$ since a larger $|\rho_D|$ is associated with a stronger impact of the omitted factor on the estimated market beta. β_D is the exposure of the BAM portfolio to the omitted factor. β_D is positive when ρ_D is negative, as when a factor is negatively correlated with the market, sorting stocks into $\hat{\beta}_i$ quantiles picks up stocks with higher (lower) loadings on that factor in the low(high)-$\hat{\beta}_i$ side. $|\beta_D|$ increases with $|\rho_D|$ since $\hat{\beta}_i$ captures more information about the omitted factor when the omitted factor's connection with the market is stronger. The positive exposure to the omitted factor translates into a positive BAM alpha, as shown in the bottom-left subfigure.

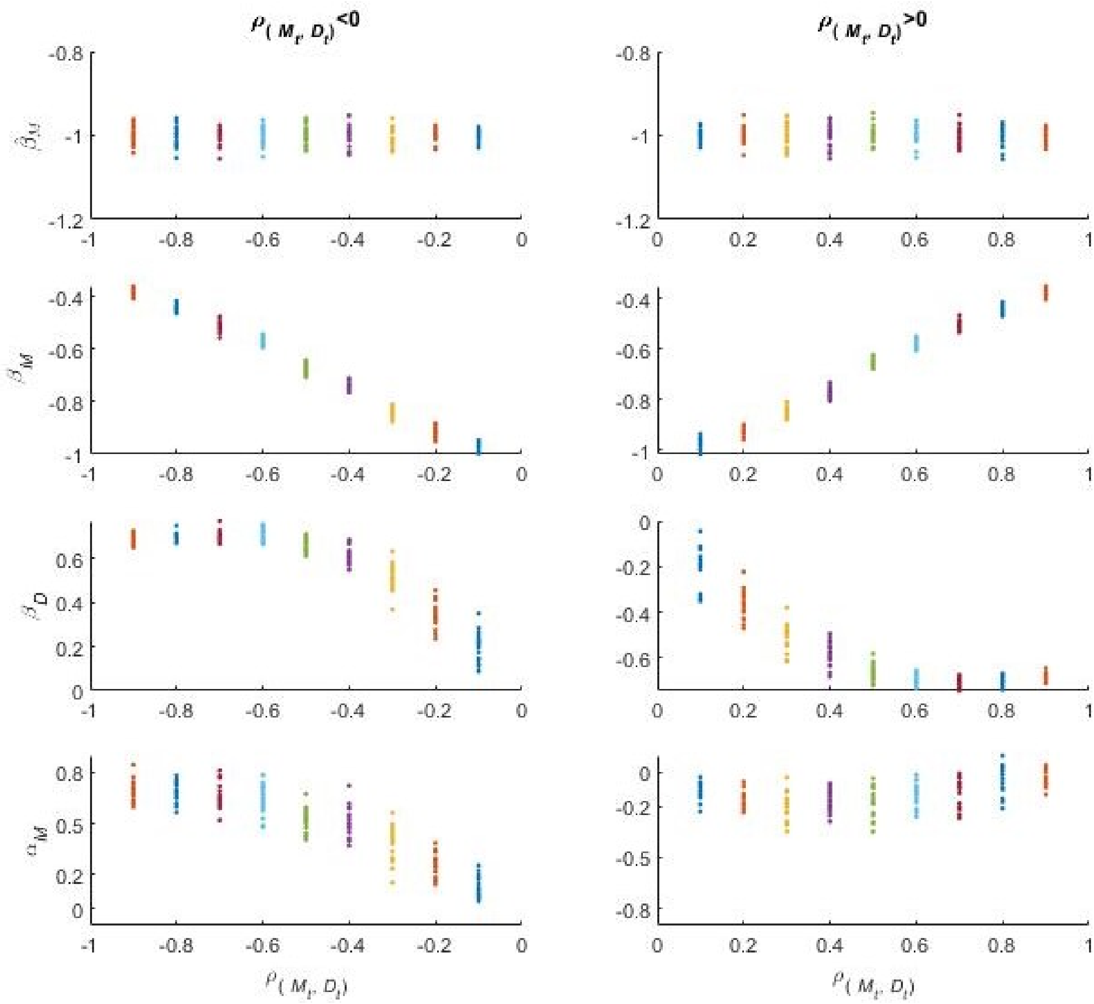
The case of positive-ρ_Ds in the right panel sheds light on why an LBA tends to be observed. As shown in the bottom-right subfigure, the magnitude of the BAM alpha increases trivially with ρ_D and maintains at a low level when the omitted factor is positively correlated with the market factor. The same conclusion applies when analyzing C_t. Given that the impact from an omitted factor negatively correlated with the market tends to be stronger than that from an omitted factor positively correlated with the market (see Equation 2.17), it is reasonable to expect that the new market-beta anomaly, if existing, is more likely an LBA.

B3.3 Simulation Procedures

This appendix section introduces the detailed simulation procedures. The setup of the parameters for the data-generating process is introduced in Section B3.1.

Figure B2.3: The Tendency for the Low-Beta Anomaly (LBA) to Emerge

This figure depicts how the betting-against-beta (BAM) portfolio performance changes with the correlation coefficient between the omitted and market factors. The data-generating process follows Equation(C.19), and the detailed simulation procedures are introduced in Appendix B3.3. The horizontal axis corresponds to the correlation coefficient between the market and omitted factors. The vertical line corresponding to each horizontal-axis value is composed of 1000 dots, with each dot denoting the result of one simulation. $\hat{\beta}_M$ reflects the estimated exposure of the BAM portfolio to the market factor. β_M and β_D reflect the BAM portfolio's true market and omitted-factor exposures. α_M is the monthly alpha (%) of the BAM portfolio against the CAPM model.



Simulation procedures to generate the results of Section B3.1

This appendix section introduces the detailed simulation procedures for Section B3.1:

1) set the values of ρ_D, ρ_C and ρ_{DC}, which describe different scenarios for a new LBA to emerge;

2) generate the time series of M_t, $u_{D,t}$ and $u_{C,t}$ with a length of $T = 618$, and then generate D_t and C_t accordingly;

3) generate β_i^*, d_i and c_i for individual stocks; generate N-by-T ϵ_{it}s as the residual returns of individual stocks; calculate individual stock returns accordingly;

4) estimate CAPM betas ($\hat{\beta}_i$) of individual stocks, sort stocks into $\hat{\beta}_i$ quantiles, and calculate value-weighted ex-ante factor exposures of each portfolio; estimate alphas against the CAPM and $[M_t, D_t]$ models;

5) estimate market betas ($\hat{\beta}_i^*$) of individual stocks using the $[M_t, D_t]$ model, sorting stocks into $\hat{\beta}_i^*$ quantiles and calculate value-weighted ex-ante factor exposures of each portfolio; estimate alphas against the $[M_t, D_t]$ and $[M_t, C_t]$ models;

6) repeat 1) to 4) 1000 times and report the average simulation results.

7) change the values of ρs and repeat 1) to 5).

Simulation procedures to generate the results of Section B3.2

This appendix section introduces the detailed simulation procedures for Section B3.2:

1) set the value of ρ_D and N; generate the time series of M_t and e_t with a length of $T = 618$, and then generate D_t accordingly;

2) generate factor returns and individual stock returns following steps 2) and 3) in Appendix B3.3;

3) estimate market betas through the CAPM to get $\hat{\beta}_i$;

4) sort stocks into quantiles according to $\hat{\beta}_i$s, and aggregate β_is and d_is of each tercile on a value-weighted basis; calculate spreads in these variables between low and high-$\hat{\beta}_i$ terciles, which are the true factor exposures of the betting-against-beta (BAM) portfolio; calculate the spanning regression alpha of the BAM portfolio on the simulated market factor;

5) repeat 1) to 4) 1000 times to get 1000 sets of simulation results;

6) change the value of N (if for Appendix B3.2) or ρ (if for Appendix B3.2) and repeat 1) to 5).

Appendix C

Chapter 3 Supplement

C1 Derivations

C1.1 Estimated beta, true beta and exposures to omitted factors

This appendix section illustrates how Δ_{β^*} and Δ_γ are inferred from Δ_β. Δ_β is determined by the thresholds used for sorting stocks into high and low-β quantiles. As Equation (3.8) suggests, once Δ_β is picked up, Δ_{β^*} is fixed. To quantify the relationship between Δ_β and Δ_{β^*}, I assume that β and β^* follow a bivariate normal distribution. Given Equation (3.7), the true beta can be inferred as

$$E[\beta_i^*|\beta_i] = x_1\beta_i + y_1 \tag{C.1}$$

where x_1 and y_1 are coefficients, satisfying

$$x_1 = \frac{Cov_{CS}\left(\beta_i^*, \beta_i\right)}{Var_{CS}(\beta_i)} = \frac{Var_{CS}(\beta_i^*)}{Var_{CS}(\beta_i)} \tag{C.2}$$

where Cov_{CS} and Var_{CS} refer to cross-sectional covariance and variance, respectively. $Cov_{CS}\Big(\beta_i^*, \beta_i\Big) = Cov_{CS}\Big(\beta_i^*, \beta_i^* + \rho\gamma_i\Big) = Var_{CS}(\beta_i^*)$ holds as different factor loadings are set to be cross-sectionally uncorrelated. Equation (C.1) can be rewritten as

$$E[\beta_i^*|\beta_i] = \frac{Var_{CS}(\beta_i^*)}{Var_{CS}(\beta_i)}\beta_i + y_1 \tag{C.3}$$

Then,

$$\Delta_{\beta^*} = \frac{Var_{CS}(\beta_i^*)}{Var_{CS}(\beta_i)}\Delta_\beta \tag{C.4}$$

Similarly, the loadings on omitted factors can be inferred as

$$E[\boldsymbol{\rho}^T\boldsymbol{\gamma}|\beta_i] = x_2\beta_i + y_2 \tag{C.5}$$

where x_2 and y_2 are coefficients, satisfying

$$x_2 = \frac{Cov_{CS}\Big(\boldsymbol{\rho}^T\boldsymbol{\gamma}, \beta_i\Big)}{Var_{CS}(\beta_i)} \tag{C.6}$$

where Cov_{CS} and Var_{CS} refer to cross-sectional covariance and variance, respectively.

$$E[\boldsymbol{\rho}^T\boldsymbol{\gamma}|\beta_i] = \frac{Var_{CS}(\boldsymbol{\rho}^T\boldsymbol{\gamma})}{Var_{CS}(\beta_i)}\beta_i + y_2 \tag{C.7}$$

Then,

$$\boldsymbol{\rho}^T\boldsymbol{\Delta\gamma} = \frac{Var_{CS}(\boldsymbol{\rho}^T\boldsymbol{\gamma})}{Var_{CS}(\beta_i)}\Delta_\beta \tag{C.8}$$

Suppose that all omitted factors, which are uncorrelated with each other, have the same correlation with F_t; then, the BAF portfolio's factor loadings are

$$\begin{aligned}\Delta_{\beta^*} &= \frac{Var_{CS}(\beta_i^*)}{Var_{CS}(\beta_i^*) + K\rho^2 Var_{CS}(\gamma_i)}\Delta_\beta \\ \boldsymbol{\rho}^T\boldsymbol{\Delta}_\gamma &= \frac{K\rho^2 Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*) + K\rho^2 Var_{CS}(\gamma_i)}\Delta_\beta\end{aligned} \tag{C.9}$$

where K is the number of omitted factors. To under this equation, we can consider a single omitted factor case:

$$\Delta_\gamma = \rho\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^* + \boldsymbol{\rho}^T\boldsymbol{\gamma}_i)}\Delta_\beta \tag{C.10}$$

where ρ is the correlation between an average omitted factor and the FMP, and Δ_γ is the BAF portfolio's exposure to the average omitted factor. Equation (C.10) indicates that to what extent estimated beta captures loadings on an omitted factor depends on the connection between that omitted factor and the FMP, and the cross-sectional variances (Var_{CS}) of factor loadings.

C1.2 The determinants of the strength of a BAF alpha from the unpriced-factor channel

This appendix section re-expresses Equation (3.12) in the format that reflects factor efficiency and the relative importance of factors in driving the time variation of asset returns. The second item on the right-hand side of Equation (3.12) is removed as unpriced factors have zero premiums. Given Equation (C.9), Equation

(3.12) can be rewritten as

$$m = K\rho^2 \times \frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i)} = \frac{K\rho^2}{K\rho^2 + \frac{Var_{CS}(\beta_i^*)}{Var_{CS}(\gamma_i)}} \tag{C.11}$$

Next, I show that $K\rho^2$ reflects factor inefficiency by decomposing F_t along omitted factors

$$F_t = \alpha_F + \rho_1\Gamma_{1t} + \rho_2\Gamma_{2t} + \cdots + \rho_K\Gamma_{Kt} + \epsilon_{F,t} \tag{C.12}$$

where $\alpha_F + \epsilon_{F,t}$ is the component of F_t orthogonal to all omitted factors.

Under the assumption that these omitted unpriced factors are uncorrelated and have the same correlation with F_t, $K\rho^2$ can be expressed as

$$K\rho^2 = 1 - \frac{Var(\alpha_F + \epsilon_{F,t})}{Var(F_t)} = 1 - \frac{Var(F_t^*)}{Var(F_t)} \tag{C.13}$$

where $F_t^* = \alpha_F + \epsilon_{F,t}$ denotes the priced component of F_t. Therefore, $K\rho^2$ evaluates the inefficiency of F_t, which is intuitive as F_t relates to unpriced factors through its unpriced component. The correlations between F_t and omitted unpriced factors reflect the extent to which the time variation of F_t is driven by its unpriced component. Combining Equations (C.11) and (C.13) leads to Equation (3.13).

C1.3 Information captured by the cross-sectional variance of factor loadings

This appendix section illustrates the information reflected by the cross-sectional variance of loadings on a factor. For brevity, I consider a single-factor model:

$$r^e_{it} = \alpha_i + \beta_i F_t + \epsilon_{it} \tag{C.14}$$

where r^e_{it} is the excess return of security i, β_i is the loading on factor F_t. ϵ_{it} is the residual return.

Take the variance of both sides

$$Var(r_{it}) = \beta_i^2 Var(F_t) + Var(\epsilon_{it}) \tag{C.15}$$

$\beta_i^2 Var(F_t)$ captures the systematic component of asset i's return variation.

Add up the variance of all stocks and take the average

$$\frac{\sum_i^N Var(r_{it})}{N} = \frac{\sum_i^N \beta_i^2}{N} Var(F_t) + \frac{\sum_i^N Var(\epsilon_{it})}{N} \tag{C.16}$$

Under the assumption that all factors have the same variance, $\frac{\sum_i^N \beta_i^2}{N}$ reflects the extent to which factor F_t captures the time variation of asset returns on average. I further decompose $\frac{\sum_i^N \beta_i^2}{N}$ as

$$\frac{\sum_i^N \beta_i^2}{N} = E_{CS}(\beta_i^2) = Var_{CS}(\beta_i) + E^2_{CS}(\beta_i) \tag{C.17}$$

where E_{CS} denotes the cross-sectional average.

Except for the market and size factors, $E_{CS}(\beta_i)$ is much smaller than $\sqrt{Var_{CS}(\beta_i)}$ for most factors, which indicates that $E_{CS}(\beta_i^2) \approx Var_{CS}(\beta_i^2)$ holds for most zero-investment factor-mimicking portfolios. Therefore, in general, $Var_{CS}(\beta_i^2)$ reflects the ability of a factor to capture the time variation of asset returns on average. For example, we can interpret $\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)} = 2$ as that the return variation captured by factor Γ_t is twice that captured by factor F_t.

C1.4 The determinants of the strength of a BAF alpha from the priced-factor channel

This appendix section re-express Equation (3.12) in the format that reflects the characteristics of the relevant omitted priced factors. Rewrite Equation (3.12) as

$$m = \frac{\boldsymbol{\Delta_\gamma}^T \boldsymbol{\alpha_\Gamma}}{-\Delta_\beta E(F_t)} = \frac{\boldsymbol{\Delta_\gamma}^T [\boldsymbol{E(\Gamma_t)} - \boldsymbol{\rho} E(F_t)]}{-\Delta_\beta E(F_t)} \tag{C.18}$$

For ease of exposition, assume that all omitted priced factors, which are uncorrelated, have the same expected return, variance, and correlation with the FMP. These assumptions do not affect the major analytical conclusions as the strength of a factor-beta anomaly depends on how omitted factors impact an FMP but not on the interaction of omitted factors. Given these assumptions, I simplify $\boldsymbol{\Delta_\gamma}^T \boldsymbol{E(\Gamma_t)}$; also, given Equation (C.9), I simplify $\boldsymbol{\Delta_\gamma}^T \boldsymbol{\rho}/\Delta_\beta$. Finally, I reach Equation (3.17).

C2 Simulation Examples

This appendix section provides simulation examples to intuitively demonstrate the emergence and strength of a low factor-beta anomaly. I generate factor and individual-stock returns following the prespecified data-generating process, estimate the betas of individual stocks on an FMP and construct a corresponding BAF portfolio. With these generated data, I analyze the BAF portfolio's factor exposures, spanning regression alpha and the composition of the BAF alpha. Details of the simulation procedures are provided in Section C2.3.

C2.1 The single omitted-factor case

I first consider the two-factor case to demonstrate how an omitted factor negatively correlated with F_t induces a positive BAF alpha. The data-generating process is as follows:

$$DGP: \begin{cases} r^e_{it} = \beta^*_i F_t + \gamma_i \Gamma_t + \epsilon^*_{it} \\ \Gamma_t = \alpha_\Gamma + \rho F_t + e_t \end{cases} \quad i = 1, \cdots, N \text{ and } t = 1, \cdots, T \tag{C.19}$$

Figure C3.1 reports the simulation result. The horizontal axis is for different $|\rho|$ values. The vertical line corresponding to each $|\rho|$ is composed of 1000 dots, with each dot denoting the result of one simulation. To generate the left panel, I set the two factors to be equally important in driving the time-variation of asset returns and have the same Sharpe ratio, $Var_{CS}(\beta^*_i) = Var_{CS}(\gamma_i)$ and $SR_\Gamma = SR_F$. Under this setup, the two factors are equally important in driving the cross-sectional return differences. The first subfigure in the left panel depicts the spread in estimated FMP betas, which is negative by construction since the BAF portfolio is constructed by buying/shorting low/high-β stocks. The second subfigure in the left panel depicts the BAF portfolio's true exposure (Δ_{β^*}) to F_t, which is what Δ_β is supposed to capture. When the correlation between the omitted factor and

F_t is weak (e.g., $\rho = -0.1$), The magnitude of Δ_{β^*} is large since β mainly reflects β^*. $|\Delta_{\beta^*}|$ decreases with $|\rho|$. However, even when ρ is minus 0.8, $|\Delta_{\beta^*}|$ is still not small, indicating that β still well captures β^*. The third subfigure in the left panel depicts the BAF portfolio's exposure to the omitted factor. Δ_γ is negative since the omitted factor negatively correlates with F_t. The fourth subfigure in the left panel depicts the component of the BAF alpha induced by beta deviation. As discussed, the existence of omitted factors correlated with F_t causes the estimated FMP beta (β) to deviate from the true FMP beta (β^*), and this beta deviation will generate a negative beta-alpha relationship. The fifth subfigure in the left panel depicts the component of the BAF alpha induced by the extra factor exposure, which is also positive due to the negative correlation. The negative beta-alpha relationships induced by beta deviation and extra-factor exposure constitute the BAF alpha reported in the bottom-left subfigure. The solid horizontal line above the zero horizontal line reflects the expected return of F_t. Under the setup that the omitted factor is equally important in driving the cross-sectional differences in returns as F_t, only when the connection between the two factors is strong enough ($\rho > 0.8$) will a low factor-beta anomaly with the strength of $m = 1$ emerge.

As the alpha depicted in the fourth subfigure in the left panel is induced by beta deviation, it also reflects the BAF alpha when the omitted factor has a zero premium: the BAF alpha from the unpriced-factor channel. This is a favorable setup for the unpriced-factor channel since the omitted factor is set as equally important as F_t in driving the time variation of asset returns ($Var_{CS}(\beta_i^*) = Var_{CS}(\gamma_i)$). The result indicates that the unpriced-factor channel can only induce a weak BAF alpha.

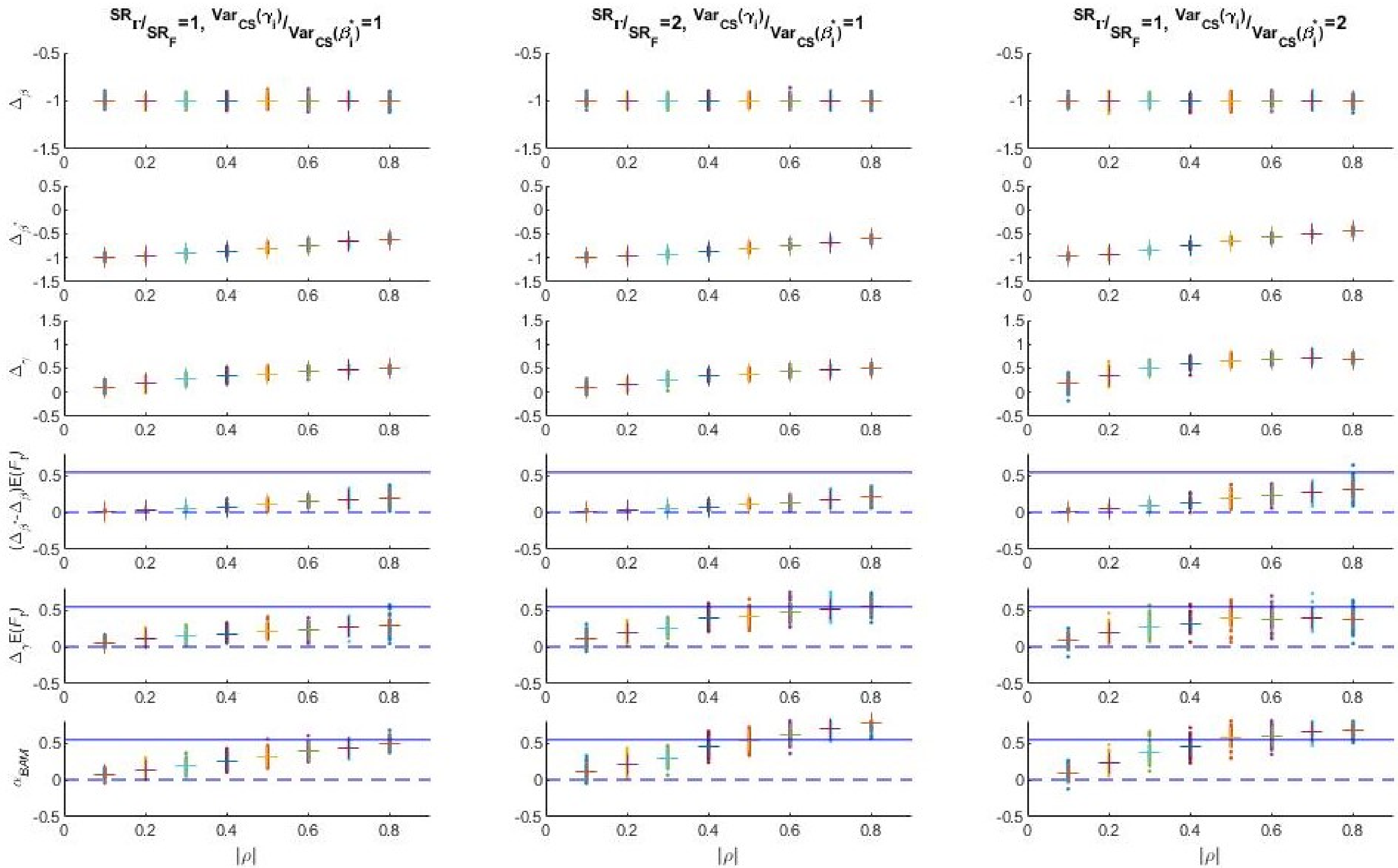
The middle panel depicts the scenario where the omitted factor has a higher Sharpe ratio than F_t ($\frac{SR_\Gamma}{SR_F} = 2$ and $Var_{CS}(\beta_i^*) = Var_{CS}(\gamma_i)$). In this scenario, the composition of estimated betas does not change (as indicated by the first three subfigures in the middle panel), but the compensation for the same level of omitted-factor exposure increases due to the higher expected return of the omitted factor. The right panel depicts the scenario that the omitted factor drives more time variation of asset returns than F_t ($\frac{SR_\Gamma}{SR_F} = 1$ and $\frac{Var_{CS}(\gamma_i)}{Var_{CS}(\beta_i^*)} = 2$). In this scenario, the omitted factor exposure accounts for a larger proportion of the estimated FMP beta (as indicated by the third subfigure in the right panel), which is also associated with a higher compensation. The BAF alphas in the middle and right panels are larger than in the left panel, and a low factor-beta anomaly with the strength of $m = 1$ emerges when the correlation is minus 0.5. The result is consistent with Equation (3.18): the strength of the low factor-beta anomaly increases when the omitted priced factor captures more cross-sectional differences in returns.

C2.2 The multiple omitted-factor case

I further analyze the emergence of low factor-beta anomalies when multiple omitted factors are correlated with F_t. This analysis also illustrates why a positive

FIGURE C3.1: Simulation Results on the Emergence of the Low Factor-Beta Anomaly - One Omitted Factor

The figure depicts the simulation results of the emergence of the low factor-beta anomaly when there exists one omitted factor negatively correlated with the factor-mimicking portfolio (F_t) from which betas are estimated. The detailed simulation procedures are introduced in Appendix C2.3. The three panels differ in the two factors' relative Sharpe ratio and importance in driving the time variation of asset returns ($\frac{SR_\Gamma}{SR_F}$ and $\frac{Var_{CS}(\gamma_i)}{Var_{CS}(^*_i)}$). The horizontal axis corresponds to the absolute correlation coefficient ($|\rho|$) between F_t and the omitted factor. The vertical line corresponding to each horizontal-axis value is composed of 1000 dots, with each dot denoting the result of one simulation. The top three subfigures report the spread in estimated betas (Δ_β) and true spreads in F_t and omitted-factor loadings (Δ_{β^*}, Δ_γ). The fourth and fifth subfigures report the composition of the BAF alpha. The bottom subfigure reports the alphas of the BAF portfolio against F_t



BAF alpha is more likely to be observed than a negative BAF alpha. The data-generating process is a multi-factor model with four factors:

$$DGP: \begin{cases} r_{it} = \beta_i^* F_t + \gamma_{1i}\Gamma_{1t} + \gamma_{2i}\Gamma_{2t} + \gamma_{3i}\Gamma_{3t} + \epsilon_{it}^* \\ \Gamma_{1t} = \alpha_{\Gamma_1} + \rho_1 F_t + e_{1t} \\ \Gamma_{2t} = \alpha_{\Gamma_2} + \rho_2 F_t + e_{2t} \\ \Gamma_{2t} = \alpha_{\Gamma_3} + \rho_3 F_t + e_{3t} \end{cases} \qquad i = 1, \cdots, N \text{ and } t = 1, \cdots, T \tag{C.20}$$

Figure C3.2 reports the simulation result. The results of the four panels differ in the number of omitted factors negatively correlated with F_t. The magnitude of the correlation coefficients between each omitted factor and F_t is the same. For all four panels, I set the four factors to be equally important in driving the time-variation of asset returns and have the same Sharpe ratio, $Var_{CS}(\beta_i^*) = Var_{CS}(\gamma_{1i}) = Var_{CS}(\gamma_{2i}) = Var_{CS}(\gamma_{3i})$ and $SR_F = SR_{\Gamma_1} = SR_{\Gamma_2} = SR_{\Gamma_3}$. Similar to Figure C3.1, the first five subfigures report the BAF portfolio's estimated and true factor exposures. The bottom three subfigures report the BAF alpha and its composition. The result indicates that when more omitted factors are negatively correlated with F_t, each omitted factor does not need to be strongly correlated with F_t to induce a strong low factor-beta anomaly. The bottom-left subfigure shows that a typical low-beta anomaly emerges when the correlation coefficient is only minus -0.3. The magnitude of the correlation coefficient can be smaller if these omitted factors explain more cross-sectional differences in returns.

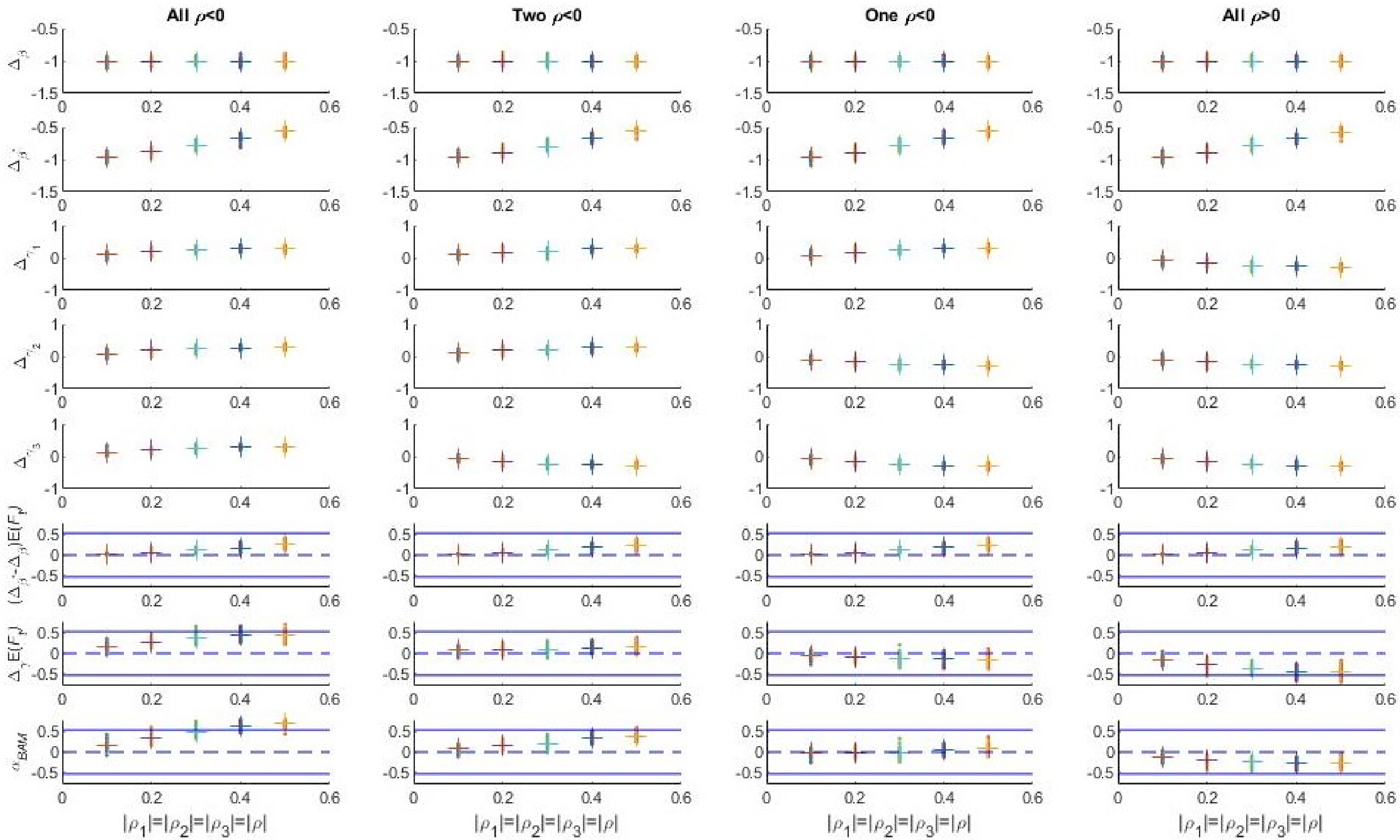
The second panel reports the results when two omitted factors negatively correlate with F_t and one positively correlates with F_t. The omitted factor positively correlated with F_t induces a positive beta-alpha relationship. However, the collective impact is still a positive BAF alpha since the impact from the two omitted factors with a negative ρ overwhelms. The third panel reports the results when only one omitted factor negatively correlates with F_t. In this case, the collective impact of the extra exposures of the BAF portfolio to these omitted factors is a positive beta-alpha relationship. However, since the beta-alpha relationship induced by beta deviation is negative, the collective impact is a near-zero (slightly positive when $|\rho|$ is high) BAF alpha. The right panel reports the results when all omitted factors are positively correlated with F_t. In this case, the BAF alpha is negative, as shown in the bottom-right subfigure. The negative BAF alpha's magnitude is smaller than that of the positive BAF alpha reported in the bottom-left subfigure. This pattern appears because a negative BAF alpha is mitigated by the negative beta-alpha relationship induced by beta deviation, as shown by the sixth subfigure in the right panel. As such, even if a negative BAF alpha can exist, it is not expected to be as strong as a positive BAF alpha. Overall, if considering the probability for each scenario to happen, the different BAF alphas of the four panels indicate that positive BAF alphas should be observed more frequently than negative BAF alphas.

C2.3 Simulation procedures to generate the results of Section C2

This appendix section introduces the detailed procedures to generate the simulation results of Appendices C2.1 and C2.2. The true data-generating process is known in a simulation environment, which is set as Equation (C.19) or Equation

Figure C3.2: Simulation Results on the Emergence of the Low Factor-Beta Anomaly - Multiple Omitted Factors

The figure depicts the simulation results of the emergence of the low factor-beta anomaly when there exist three omitted factors correlated with the factor-mimicking portfolio (F_t) from which betas are estimated. The detailed simulation procedures are introduced in Appendix C2.3. The four panels differ in the number of omitted factors negatively correlated with F_t. These factors' Sharpe ratio and importance in driving the time variation of asset returns are the same ($SR_F = SR_{\Gamma_1} = SR_{\Gamma_2} = SR_{\Gamma_3}$ and $Var_{CS}(^*_i) = Var_{VS}(\gamma_{1i}) = Var_{VS}(\gamma_{2i}) = Var_{VS}(\gamma_{3i})$). The horizontal axis corresponds to the absolute correlation coefficient ($|\rho_1| = |\rho_2| = |\rho_3|$) between F_t and the omitted factor. The vertical line corresponding to each horizontal-axis value is composed of 1000 dots, with each dot denoting the result of one simulation. The top five subfigures report the spread in estimated betas (Δ_β) and true spreads in F_t and omitted-factor loadings (Δ_{β^*}, Δ_{γ_1}, Δ_{γ_2} and Δ_{γ_3}). The sixth and seventh subfigures report the composition of the BAF alpha. The bottom subfigure reports the alphas of the BAF portfolio against F_t.



(C.20). I set the return and volatility of the factor-mimicking portfolio (F_t) from which betas are estimated as $\mu_F = 0.53\%$ and $\sigma_F = 4.39\%$ (same as the market return and volatility). The factor-mimicking portfolio (F_t) and residuals of omitted factors ($\boldsymbol{e}_t$) returns are set to follow uncorrelated normal distributions. I set true loadings (β_i^*, γ_{1i}, γ_{2i} and γ_{3i}) of individual stocks to follow normal distributions. The residual variance ($\sigma_{\epsilon^*}^2$) is set as two times the factor variance to reflect that the idiosyncratic part of the individual stock variance is usually larger than the systematic part. These parameter values do not materially impact the results within a wide but reasonable range.

The values of T and N affect the likelihood of observing a specific simulation result. For example, if we set T to be large, the results are less affected by the random realization of returns. We only have data of several decades in financial markets, over which the factors may have a negative cumulative realization or estimated betas may not pick up loadings on relevant factors as the theory regulates. I set T as 678 and N as 3000. The former is to mimic that there are 678 months over the sample period, and the latter is to mimic that there are 3000 stocks on average at each time point.

I evaluate how the betting-against-factor-beta (BAF) performance changes with the correlation between the omitted factor and F_t. The simulation procedure is as follows:

1) set the value of ρ and then calculate $\alpha_\Gamma = E(\Gamma_t) - \rho E(F_t)$; generate the time series of F_t and $\boldsymbol{e}_t$ with a length of $T = 678$, and then generating $\boldsymbol{\Gamma_t}$ accordingly;

2) Generate β_i^*, γ_{1i}, γ_{2i} and γ_{3i} and ϵ_{it}^*, and then calculate individual stock returns accordingly; repeat this for $N = 3000$ times;

3) Estimate empirical betas through the regression $r_{it} = \alpha_i + \beta_i F_t + \epsilon_{it}$;

4) Sorting stocks into three quantiles according to β_i and construct the BAF portfolio by buying/shorting the low/high-β_i quantile; aggregate β_i^*, γ_{1i}, γ_{2i} and γ_{3i} of the BAF portfolio; calculate spanning regression alphas of the BAF portfolio on F_t

5) Repeat this for $M = 1000$ times to get M sets of 4);

6) Changing ρ and repeating 1) to 5).

Appendix D

Supplement to Laite Guo's Ph.D. Thesis

The first chapter was published in the Journal of Banking & Finance, Volume 146, Laite Guo, Two Faces of the size effect, Copyright Elsevier (2022)."

Please cite as:

Guo, L. (2023). Two faces of the size effect. *Journal of Banking & Finance* 146, 106708. ISSN: 0378-4266.