UTS UNIVERSITY OF TECHNOLOGY SYDNEY

Factor Failures: The Limitations and Pitfalls of Factor Models in Empirical Asset Pricing

by Justin James Hitchen

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

under the supervision of Vitali Alexeev, Gerhard Hambusch, and Dave Michayluk

University of Technology Sydney
Faculty of Business

January 2024

Statement of originality

I, Justin Hitchen, declare that this thesis, titled "*Factor Failures: The Limitations and Pitfalls of Factor Models in Empirical Asset Pricing*", is submitted in fulfilment of the requirements for the award of the degree of Doctor of Philosophy, in the Finance Discipline Group at UTS Business School at the University of Technology, Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. I confirm that this thesis has not been submitted previously for a higher degree or qualification at any other university or institute of higher learning. This research is supported by the Australian Government Research Training Program.

Production Note:
Signature removed prior to publication.

Justin J. Hitchen

Date: January 21, 2024

Acknowledgements

I have great gratitude for the support of my supervisory panel throughout this PhD. Vitali Alexeev, Gerhard Hambusch, and David Michayluk have each provided significant feedback, guidance, and encouragement throughout the PhD process.

I also thank the UTS finance department for their support during my studies. The UTS finance department contains a wealth of knowledge which they have willingly shared to assist in the refinement of ideas presented in this thesis. In particular, I would like to thank Kris Glover, Hardy Hulley, Marco Navone, Vinay Patel, Kenny Phua, and Tālis Putniņš. I would also like to single out Christina Nikitopoulos, who served in the unenviable role of PhD Coordinator for the majority of my PhD studies and helped ensure I fulfilled the administrative requirements of the PhD.

I am thankful for the insightful feedback I have received at conferences where these papers have been presented. In particular, for comments provided at the 34^{th} Australasian Finance and Banking Conference by Sean Anthonisz (USYD), and at the 11^{th} FIRN Annual Conference by Antje Berndt (ANU), Neal Galpin (MON), Jianlei Han (MU), Minsoo Kim (UOM), Xiaoyang Li (DU), Jianfeng Shen (UNSW), Sean Wu (QUT), and Zhuo Zhong (UOM). In addition, I am very grateful for the feedback of those who formally reviewed my PhD dissertation. They have provided high quality feedback which has helped to enhance this final submission.

Finally, I express immense gratitude for my wife and family who have indulged my pursuit of a PhD despite the financial and time sacrifices it imposed.

Table of contents

Statement of originality **ii**

Acknowledgements **iii**

Table of contents **iv**

List of figures **vi**

List of tables **vii**

Abstract **viii**

Chapter 1: Introduction **1**

1.1 Background and motivation 1

1.1.1 The "*common*" application of factor models in finance 1

1.1.2 The growth in documented return "anomalies" 2

1.2 Contribution 4

1.3 Structure of this thesis 5

Chapter 2: Is beta busted? **7**

2.1 Introduction 7

2.2 Related literature and contribution 10

2.3 Mismatch between theory and measurement 13

2.3.1 Imperfect market correlation as a source of disconnect 15

2.3.2 Choice of market proxy as a source of disconnect 18

2.4 Historical Beta Mismatch Error in US stocks 21

2.4.1 Historical Beta Mismatch Ratio 21

2.4.2 Historical tangency portfolio 24

2.5 Beta Mismatch Error and prominent finance anomalies 26

2.5.1 Portfolio formation reduces Beta Mismatch Error 26

2.5.2 Impact of Beta Mismatch Error upon Fama French factors 27

2.6 Conclusion 30

Chapter 3: Rethinking Performance Evaluation: Is Alpha Reliable in Practice? **32**

3.1 Introduction 32

3.2 Related literature & the problem of model misspecification 35

3.3 The decomposition of time-series Alpha 42

3.4 Quantifying historical factor model bias 51

3.4.1 Data and method 51

3.4.2 Results 52

3.5 Populating a "factor zoo" 56
3.6 Conclusion 58
Chapter 4: Endogenous Reference Bias: When is it a problem? 61
4.1 Introduction 61
4.2 Related literature and contribution 64
4.3 Analytical examination of bias and proposed remedy 66
4.3.1 Endogenous reference bias 66
4.3.2 Towards unbiased estimation 71
4.4 Historical beta estimation bias in the US stock market 75
4.5 Impact of bias on portfolios 81
4.5.1 Portfolio formation based on stock characteristics 81
4.5.2 Hedging against endogenous reference bias 83
4.6 Conclusion 89
Chapter 5: Conclusion 91
5.1 Summary & implications 91
5.1.1 Is Beta Busted? 91
5.1.2 Is Alpha reliable in practice? 92
5.1.3 Endogenous reference bias: when is it a problem? 93
5.1.4 Concluding remarks 94
5.2 Future research 96
References 97
Appendix 101
Appendix A.1: Theory vs OLS Betas for cap-weighted proxy 101
Appendix A.2: Theory vs OLS Betas for tangency portfolio 102
Appendix C.1: Most biased Betas by year 103
Appendix C.2: Comparison of fixed effect models 105
Appendix C.3: Derivation of biased OLS Beta estimates 106
Appendix C.4: Perceived outperformance of low Beta assets under model misspecification 109

List of figures

Figure 2.1: Many different price paths produce the same return multiple 15
Figure 2.2: Severity of Beta Mismatch Error for increasing correlation 16
Figure 2.3: Beta Mismatch Ratio for increasing correlation 17
Figure 2.4: Beta mismatch still present in perfectly fit regressions 19
Figure 2.5: No Beta mismatch when using tangency portfolio 20
Figure 2.6: Historical stock correlation with cap-weighted index 22
Figure 2.7: Historical BMR with cap-weighted index as market proxy 22
Figure 2.8: CAPM Theoretical Betas vs. OLS estimates for cap-weighted proxy 24
Figure 2.9: Portfolio formation reduces Beta Mismatch Error 27
Figure 3.1: No risk premium misspecification vs. extreme risk premium misspecification 39
Figure 3.2: Partial risk premium misspecification 40
Figure 3.3: Factor model is correctly specified 45
Figure 3.4: Equivalence of αiTS and $\alpha iExc$ in Scenario 1 46
Figure 3.5: Factor model is misspecified 47
Figure 3.6: Divergence of αiTS and $\alpha iExc$ in Scenario 2 48
Figure 3.7: Distribution of p-values for αi and $\alpha iExc$ across beta deciles 49
Figure 3.8: Correlation of terminal wealth with αi and $\alpha iExc$ 49
Figure 3.9: Deviation of average annual alpha estimates from time-series vs. cross-sectional regression 53
Figure 3.10: Distribution of time-series alpha estimates conditional on beta 55
Figure 4.1: Bias in beta estimate for an uncorrelated asset j as wj increases 69
Figure 4.2: Bias in the beta estimate as wj increases for varying βj 70
Figure 4.3: Bias in the beta estimate as wj increases for varying (Mj). 71
Figure 4.4: Security market line based on beta-grouped decile portfolios. 82
Figure 4.5: No risk premium misspecification vs. extreme risk premium misspecification 84

List of tables

Table 2.1: SMB portfolio parameters under alternate estimation approaches 29
Table 2.2: HML portfolio parameters under alternate estimation approaches 29
Table 3.1: Risk premium misspecification under time-series regression (1991-2020) 52
Table 4.1: Comparison of estimation approaches for Simulation 1 73
Table 4.2: Comparison of estimation approaches for Simulation 2 75
Table 4.3: Typical endogenous reference bias for varying US benchmarks (1991-2020) 77
Table 4.4: Impact of endogenous reference bias on alpha estimation for varying US benchmarks (1991-2020) 78
Table 4.5: MAPE of beta estimates stratified by size and $\boldsymbol{\beta j}$ (1991-2020) 78
Table 4.6: MAPE of beta estimates stratified by size and variance (1991-2020) 79
Table 4.7: Drivers of endogenous reference bias (1991-2020) 80
Table 4.8: Alpha of portfolio hedged against endogenous reference bias (1991-2020) 86
Table 4.9: Performance comparison of hedging portfolio constituents (1991-2020) 88

Abstract

Factor return models are widely used throughout finance in both academic research and industry practice. They are one of the primary tools for evaluating the performance (risk and risk-adjusted returns) of an asset or investment strategy. This thesis examines econometric inconsistencies between the parameter estimates obtained via common applications of factor models and their interpretation by academics and practitioners. We demonstrate that common estimation techniques are plagued by material estimation biases that severely diminish the usefulness of parameters estimated. As consequence, practitioners are making misguided investment evaluations whilst academics may also be drawing invalid inferences in their research.

Chapter 2 of this thesis focuses on one of the most widely studied models in finance history; the Capital Asset Pricing Model (CAPM). The CAPM provides a framework to measure systematic risk, or "Beta", which has become profoundly important to finance academics and practitioners alike. Beta is widely used to evaluate investment performance, analyse market efficiency, and manage portfolio risk. Unfortunately, a growing body of literature demonstrates that common approaches to Beta estimation may produce a metric that is largely stochastic. In this chapter we explore inconsistencies between CAPM theory and OLS estimates of Beta to understand why. We conclude that OLS regression can only produce non-spurious estimates of Beta when employed on data that *perfectly* reflects CAPM theory, an extreme improbability in real world observations.

Chapter 3 of this thesis focuses on a common form of model misspecification in factor models. "Alpha" is one of the most important metrics for performance evaluation used by both academics and practitioners. A bias in Alpha estimates has the potential to undermine the validity of any investment strategy. Our analysis shows that Alphas estimated using time-series regressions of factor models are systematically biased due to model misspecification of cross-sectional risk premia. We demonstrate that cross-sectional regressions are more appropriate for estimating Alpha because performance is penalised based on *observed*, rather than *assumed*, risk-return relations in data. Our results suggest that time-series estimates of Alphas are largely spurious as opposed to a true representation of risk-adjusted performance.

Chapter 4 examines a specific form of endogeneity associated with factor models. Empirical finance is greatly concerned with endogeneity and its potential to invalidate

analysis. However, the most pervasive form of endogeneity in finance is rarely considered in asset pricing literature. The "market portfolio", which commonly serves as an explanatory variable in statistical tests, is clearly an endogenous reference; securities are regressed on a portfolio of which they are a constituent. In addition to constituent weight, there are several stock characteristics that influence the direction and severity of endogenous reference bias. We examine the conditions in which stocks are likely to have the most significantly biased beta estimates and explore the impact on performance evaluation. We demonstrate that a simple hedging portfolio against endogenous reference bias is able to generate an alpha economically larger than the small cap and value premiums during our sample period.

Overall, this thesis critically examines several common applications of linear factor return models in empirical asset pricing. We demonstrate that Beta estimates obtained from single-factor models deployed on historical data differ greatly from the values anticipated by empiricists. We show that model misspecification is likely to cause spurious Alpha estimates in commonly used time-series regressions of single and multiple-factor models. Finally, we demonstrate that endogeneity will always be present in factor models since they involve regression of dependent variables upon factors of which they are a constituent. In culmination, this thesis cautions empiricists to scrutinise whether their current applications of factor models are appropriate and calls into question the validity of historical findings substantiated by factor models.

Chapter 1

Introduction

1.1 Background and motivation

1.1.1 The "*common*" application of factor models in finance

This thesis provides a critique of common applications of factor models. In order to identify why common applications of factor models are flawed, it is helpful to express what "common applications" entails. Firstly, throughout this thesis, the phrase "factor models" refers to linear econometric return models used by empiricists to evaluate the return of an asset or investment strategy after accounting for systematic risks. Most famously, a single-factor regression based on the theory of the Capital Asset Pricing Model (CAPM) gained widespread usage in the second half of the 20th century and remains popular today.[1]

To understand the "common" use of factor models by industry we can sample some teachings from the most influential practitioner bodies in finance on how to interpret a single-factor model based on the CAPM:

> "*A company with a β that's greater than 1 is more volatile than the market. For example, a high-risk technology company with a β of 1.75 would have returned 175% of what the market returned in a given period.*"
>
> (Corporate Finance Institute, 2023)

> "*The higher the beta of an asset, the higher its expected return will be. Assets with a beta greater than 1 have an expected return that is higher than the market return.*"
>
> (CFA Institute, 2023)

[1] Graham & Harvey (2001) found that 73.5% of US CFOs always, or almost always, used the CAPM when estimating the cost of capital for prospective investments.

Similar interpretations are echoed by popular general audience websites:

> "*If an asset has a beta above (below) 1, it indicates that its return moves more (less) than 1-to-1 with the return of the market-portfolio, on average.*"
>
> (Wikipedia, 2023)

Practitioners are taught to take ordinary least squares (OLS) regression estimates of a single-factor model as informative of expectations of the future returns of an asset; if I estimate a historical stock Beta of 2, I should anticipate it will deliver returns twice that of the market factor next year. Unfortunately, if an estimation framework suffers from material bias, then the Beta estimates upon which practitioners rely could become partially or completely spurious. In essence, that is the contribution of this thesis; to identify several material forms of bias to OLS estimates of factor model parameters that are not currently recognised by practitioners nor adjusted for in historical academic literature.

Whilst it is not uncommon in contemporary academic research to use alternate estimation procedures, this thesis contextualises the econometric issues identified within an OLS estimation framework for two primary reasons. Firstly, practitioners are primarily taught the estimation of factor model parameters via OLS. As a result, it makes sense to critique the framework commonly employed to inform real world asset allocation. Secondly, whilst the variety of estimation procedures used in finance literature has increased over time, many of the seminal works involving performance evaluation and the identification of return anomalies were based on the employ of OLS. For example, the papers credited with early identification of the "size effect" anomaly (Banz, 1981), the "value effect" anomaly (Fama & French, 1992), and the "momentum" anomaly (Jegadeesh & Titman, 1993) relied on OLS estimation. Consequently, material bias to OLS parameter estimates of factor models serves to question not only the validity of contemporary industry practice but also influential historical findings.

1.1.2 The growth in documented return "anomalies"

The CAPM specifies that if all investors hold well-diversified (mean-variance efficient) portfolios, then asset returns are a linear function of their exposure to market risk; a single factor is sufficient to explain the cross-section of asset returns. Historically, the CAPM has fallen well short of this standard in published literature. Eugene Fama and Kenneth French claim that the empirical record of the CAPM is "*poor enough to invalidate the way it is used in applications*" (2004, p. 25).

Perceived empirical deficiencies of the CAPM have resulted in expansion of the single-factor model to include other potential return factors that plausibly capture systematic risks and by doing so facilitate a more complete explanation of the cross-section of asset returns. Fama and French initially advocated use of an alternate three-factor model (1993), which included additional presumed risk factors emerging in finance literature, as an improvement over the CAPM. Their model gained widespread adoption in academia. However, the expansion of a CAPM-based single-factor model to incorporate additional presumed risk premia has been controversial.[2] Markowitz's (1993) mean-variance mathematics, which underpins the CAPM, specifies that the expected returns on all assets are spanned by two frontier portfolios: any efficient portfolio and its zero-covariance counterpart. The important corollary is that there is no need for additional factors. This result is a mathematical tautology which holds for any data set. Smith and Walsh (2013, p. 75) leverage this tautology to criticise the Fama French expansion of the CAPM; "*What does this mean for the Fama and French factors? The implication is clear: if researchers are allowed to look ex post there will be an infinite number of portfolios that they can find .... By concentrating on the size anomaly and the market to book anomaly, Fama and French have found a workable method of constructing ex post efficient portfolios*.".

Despite controversy, the pace of proliferation of new factor model "anomalies" has only accelerated in the 21st century. Return anomalies against the Fama French's advocated three-factor model have become increasingly prevalent in empirical finance papers. Fama and French (2015) have since extended their multifactor model to incorporate five presumed risk factors. By 2020, there were over 450 published return anomalies in accounting and finance journals (Hou et al., 2020).

The rapid expansion of identified return anomalies is one of the great puzzles of modern finance. Anomalies should be rare *by definition*. Is there some unidentified

---

[2] E.g., see Roll and Ross (1994, p. 101): "*the true cross-sectional expected return-beta relation is exact when the index is efficient, so no variable other than beta can explain any part of the true cross-section of expected returns.*"

deficiency in the theory underpinning commonly used factor models? Is the increasing prevalence of anomalies simply the result of data mining? Or is there an entirely different cause? This thesis presents three novel forms of bias that heavily influence the estimates obtained from factor models. In doing so, we help solve the puzzle as to why anomalies have become commonplace.

1.2 Contribution

Motivated by the proliferation of anomalies, this thesis set out to identify undocumented econometric inconsistencies that could plausibly be the cause. Three distinct issues affecting parameter estimation in common applications of factor models have been identified. Each issue is covered in a separate chapter of this thesis.

Chapter 2 identifies an estimation bias introduced to single-factor models when using a non-tangency market proxy. We develop a novel measure for quantifying the severity of estimation error that arises. Our novel measure allows us to examine the conditions in which the mismatch between theory and estimation creates the most significant bias to OLS beta estimates. We examine the historical magnitude of estimation error arising from the mismatch between theory and OLS estimation for individual US equities. We also explore the severity of mismatch error for portfolios rather than individual securities and demonstrate that both the size and value premium are at least partially attributable to mismatch error. Chapter 2 concludes by arguing that the magnitude of parameter estimation bias identified is substantial enough to invalidate the empirical use of single-factor models based on the CAPM.

Chapter 3 highlights the issue of model misspecification when employing time-series regressions of factor models for performance evaluations. The chapter is inspired by a novel starting point; we generate random data and examine whether documented anomalies are present in random data. If they are, it suggests they are generated by an econometric issue rather than being a meaningful characteristic of real-world data. We observe that the famous "low beta anomaly" (e.g., Frazzini and Pedersen, 2014) of finance is evident in random data. Specifically, we identify that the use of time-series rather than cross-sectional regression to estimate factor model parameters induces bias via model misspecification. This occurs because time-series regressions impose an arbitrary market risk premium assumption that is not informed by the underlying data;

alpha estimates become a linear function of beta estimates. We are also the first to quantify the magnitude of model misspecification bias for historical US data. In addition, we demonstrate the ease with which long-short portfolios can be constructed to exploit model misspecification bias and generate practically infinite spurious return anomalies.

Chapter 4 builds on an existing but understudied area of research. We highlight that both single and multiple-factor models are subject to a specific form of endogeneity. This endogeneity arises from the practice of regressing dependent variables against "independent" variables of which they are a constituent ("endogenous reference bias"). We present the first research to consider the impact of three competing influences on the direction of magnitude of endogenous reference bias; facilitating richer insights into when endogenous reference bias will be a significant problem. By contrast, the closest paper to this chapter, by Malloch et al. (2016), only identifies two drivers which leads them to make claims which we dispute in this chapter. None of the existing literature examines the impact of endogenous reference bias upon the empirical data of a major developed market. We study the US stock market and find that beta estimation bias is unlikely to be severe for *most* individual securities. However, we identify that portfolios formed on Beta can be severely impacted by endogenous reference bias; providing intuition as to a source of estimation bias that led to the remarkable robustness of Frazzini and Pedersen's (2014) 'Betting Against Beta' anomaly. We are also the first paper to develop a simple hedging portfolio against endogenous reference bias. This hedging portfolio generates an alpha economically larger than the small cap and value premiums during our 30-year sample period.

1.3 Structure of this thesis

This thesis is structured around three distinct econometric biases affecting common empirical applications of single and/or multiple-factor models:

i) **Chapter 2**: Estimation bias introduced to single-factor models when using a non-tangency market proxy (e.g. market-cap weighted).

ii) **Chapter 3**: Model misspecification bias when using time-series regression to estimate factor model parameters.

iii) **Chapter 4**: Endogeneity arising from regressand inclusion in regressors.

Each thesis chapter has a section dedicated to related literature which further distinguishes the novel contributions of this thesis. Chapter 5 presents a summary conclusion and identifies avenues for future research.

Chapter 2

Is beta busted?

2.1 Introduction

The Capital Asset Pricing Model (CAPM) is a central building block in modern finance theory. Since its development in the 1960s (Sharpe, 1964; Treynor, 1962; Lintner, 1965; Mossin, 1966), the CAPM has offered researchers and practitioners a standardised approach to price risky securities by establishing a linear relation between expected return and systematic risk (Beta).[3] The impact of the CAPM upon finance theory and practice has been extraordinary, in part due to the model's simplicity and overwhelming popularity in real-world corporate finance and investments.[4] However, reliance on the CAPM in financial decision making is likely problematic.

CAPM theory requires that investors have similar expectations of the risk and return of a given asset and that only systematic risk is important to investors who will each hold large, well-diversified portfolios. Therefore, for measures of Beta to be useful, security Betas should be primarily deterministic. If a security's Beta cannot be approximated by investors *ex ante* then it may not be expected to meaningfully influence investment practice as anticipated by the CAPM. Unfortunately, an extensive body of literature appears to demonstrate that Betas are at least partially stochastic. Fabozzi and Francis (1978) observed that the Betas of "many stocks" on the New York Stock Exchange (NYSE) were not stable as presumed under OLS but rather they moved randomly through time. Similarly, Sunder (1980), Bos and Newbold (1984), and Collins et al. (1987) subsequently confirmed that the betas of individual stocks are non-stationary or appear stochastic over time. However, what remains unresolved is *why* OLS estimates of stock Betas appear to be partially stochastic over time.

---

[3] $(E[r_i] - r_f) = \beta_i(E[r_M] - r_f)$ is the Sharpe (1964) – Lintner (1965) CAPM, where $E[r_i]$ is the expected return of asset i, r_f is a time-invariant risk-free rate of return, $E[r_m]$ is the expected return of the market portfolio comprising all assets, and β_i is a measure of the systematic risk of asset i. In the CAPM, the expected excess return of a given asset is a linear function of its systematic risk multiplied by the market risk premium.

[4] Graham & Harvey (2001) found that 73.5% of US CFOs always, or almost always, used the CAPM when estimating the cost of capital for prospective investments.

This chapter provides a novel contribution to extant literature by highlighting and exploring a fundamental inconsistency between CAPM theory and OLS estimates of Beta. Under CAPM theory, stock Betas algebraically represent a return multiple $\left(\beta_i = \frac{E[R_i]}{E[R_M]}\right)$.[5] This interpretation of Beta is also widely taught to practitioners by the world's leading finance institutions.[6] By contrast, OLS estimates of Beta identify a distinctly different value $\left(\hat{\beta}_i = \frac{Cov(R_i, R_M)}{Var(R_M)}\right)$. These two values of Beta need not have a close relation; a single return multiple over a defined time period could have been produced by infinite different price paths and hence could produce infinite different OLS Beta estimates. Likewise, a single Beta could have been produced by infinite different return multiples. If the CAPM perfectly explained the average relation between systematic risk and return in our examined data then, on average, Betas estimated under OLS regression would be equivalent to the return multiple of CAPM theory. However, as soon as empirical data even weakly deviates from the CAPM, the relation between these two measures breaks down.[7]

The first task of this chapter is to demonstrate analytically that OLS estimates of stock Betas are distinctly different to CAPM return multiples. We also provide simulation evidence that these two measures are only approximately the same when examined data precisely follows the CAPM.[8] By contrast, when examined data is uninfluenced by the CAPM, then OLS Betas become stochastic variables distributed around 1. The implication is that unbiased estimation of Betas via OLS requires that our data perfectly conforms to CAPM theory, or else our obtained estimates become spurious. In real world settings it is practically impossible that our underlying data would perfectly reflect the CAPM. Indeed, since as early as Black et al. (1972) finance literature has consistently observed an empirical security market line that is "too flat" relative to the expectation of the CAPM.

---

[5] $(E[r_i] - r_f) = \beta_i(E[r_M] - r_f)$ can be restated as $\beta_i = \frac{(E[r_i]-r_f)}{(E[r_M]-r_f)}$. For conciseness, we can also restate returns as "excess returns" (returns in excess of the risk-free rate) which we denote via use of capitalisation: $\beta_i = \frac{E[R_i]}{E[R_M]}$.

[6] CFA Institute: "*The higher the beta of an asset, the higher its expected return will be. Assets with a Beta greater than 1 have an expected return that is higher than the market return*". Corporate Finance Institute: "*A company with a β that's greater than 1 is more volatile than the market. For example, a high-risk technology company with a β of 1.75 would have returned 175% of what the market returned in a given period*".

[7] E.g., see Roll and Ross (1994) where a market proxy with a return just 22 basis points less than the efficient frontier could induce zero cross-sectional relation between asset expected returns and Beta estimates.

[8] *precisely follows the CAPM*: not only is a mean-variance efficient market proxy selected but it is also the unique tangency portfolio.

Having identified the potential for disconnect between the betas of CAPM theory and those estimated via OLS, the next section of this chapter examines whether, and to what extent, OLS estimated Betas have historically differed from their theoretical values. We demonstrate for historical US data that these two measures typically greatly deviate from each other when estimating the Betas of individual stocks. To the extent that OLS estimates of Beta are inconsistent with CAPM theory, they undermine subsequent analysis that relies on expectations founded on the CAPM. For example, in tests of market efficiency, OLS Betas will produce a security market line flatter than anticipated in CAPM theory because they do not represent the Betas of CAPM theory and hence there is no expectation we would observe an empirical security market line (SML) with slope $\bar{R}_M$.

The final task of this chapter is to investigate whether the disconnect between CAPM Betas and their empirical measurement via OLS has prompted the false identification of prominent finance anomalies. Specifically, we focus on the two famous anomalies of the Fama-French 3-Factor (FF3F) Model; the value and size premiums. We find that OLS estimated Alphas of both of these 'anomalies' are driven by a combination of a) abnormally low market correlations of the long portfolio in the long-short factor portfolio construction, and b) improper selection of a market proxy. The abnormally low correlation results in OLS estimates of the Beta for this portfolio being understated which in turn produces inflated Alpha estimates. The implications of this section can be generalised to the growing 'factor zoo' of empirical finance; practically infinite false finance anomalies can be generated through strategic construction of long-short portfolios. An empiricist need only to take a long position in a portfolio with low correlation with the market index.

The remainder of this chapter is arranged as follows. Section 2.2 of this chapter discusses related literature and further distinguishes our contribution. Section 2.3 of this chapter demonstrates the potential for mismatch between the Beta of CAPM theory and its empirical estimation via OLS. Several metrics to measure the impact of this mismatch are also developed. Section 2.4 quantifies the historical size of Beta Mismatch Error for US stock data. Section 2.5 evaluates the impact of Beta Mismatch Error upon prominent anomalies with focus upon the size and value premiums. Section 2.6 summarises our main findings, discusses limitations and opportunities for further research, and then concludes.

2.2 Related literature and contribution

The use of market Betas in empirical analysis remains contentious. Almost as soon as the concept of market Betas was developed a disconnect was discovered between expectations of the CAPM and observations of empirical data. Black et al. (1972), and many subsequent papers, found that relations between expected returns and Beta were far flatter than anticipated by the CAPM. Roll (1978) went further in his condemnation and emphasised that it is mathematically impossible for investment performance to be evaluated using the SML criterion of the CAPM.[9] Despite the controversy, use of market Betas has gained widespread adoption in industry.[10] As a result, there have been several attempts to explain the apparent disconnect between CAPM theory and empirical evidence.

The most popular explanation for the disconnect is the errors-in-variables problem (**EIV**). The EIV problem stipulates that when independent variables are measured imprecisely OLS estimates of Beta become biased irrespective of sample size. If the measurement error has mean zero and is independent of the affected variable, the OLS estimator of the respective coefficient is biased towards zero. The "market portfolio" conceived by Markowitz (1952), which would become the basis for the development of the CAPM, represented all investible assets in the world. To date, it is not feasibly possible to track the values of such a hypothetical portfolio so empiricists have relied on the use of market proxies by necessity. Use of a market proxy means that our independent variable, market excess returns, is measured imprecisely. Consequently, OLS estimates of Beta are expected to be biased. The severity of this bias is further compounded when empiricists attempt to estimate the slope of the security market line using cross-sectional OLS regressions since time-series OLS estimates of Beta are now used as an "independent" variable. Within this context it is unsurprising that we historically observe a SML which is "too flat" relative to theoretical expectations of the CAPM. Several attempts to adjust for EIV bias have been developed over time (e.g. see Vasicek, 1973;

---

[9] Roll (1978, p.1060): "*Individual differences in portfolio selection ability cannot be measured by the securities market line criterion. If the index is ex ante mean-variance efficient, the criterion will be unable to discriminate between winners and losers. If the index is not ex ante efficient, the criterion will be designate winners and losers; but another index could cause the criterion to designate different winners and losers and there is no objective way to ascertain which index is correct.*"

[10] Graham & Harvey (2001) found that 73.5% of US CFOs always, or almost always, used the CAPM when estimating the cost of capital for prospective investments.

Kim, 1995; Jegadeesh et al., 2019). However, without additional information about the measurement error of the Beta variable used in the cross-sectional regression, a complete correction of the EIV bias is not feasible.

Alternatively, Roll and Ross (1994) reconcile the disconnect between theory and observation by demonstrating that use of a market proxy that exhibits even a minor deviation from mean-variance efficiency can result in zero cross-sectional correlation between market betas and expected returns.[11] Counterintuitively, they also demonstrate that it is possible for a market proxy to be substantially inefficient and still produce a strong cross-sectional regression between expected returns and betas. Given that the market proxies commonly used in asset pricing tests are typically not mean-variance efficient, it is unsurprising we observe weak cross-sectional correlation between market betas and expected returns.

More recently, López de Prado (2023, p. 2) highlights the logical inconsistency at the heart of the factor investing literature: "*On one hand, researchers attempt to compute unbiased* $\hat{\beta}$ *and p-values in a way that is consistent with a causal interpretation of the factors. On the other hand, researchers almost never state a causal graph or falsifiable causal mechanism under which the specification is correct, and the estimates are unbiased. The result of this causal confusion is an academic literature where factors are not really factors (in the causal sense) and where unfalsifiable spurious claims proliferate.*" In summary, misinterpretation of association as causation has established a low hurdle in the identification of risk premia throughout literature.

Perhaps the simplest explanation yet to be rigorously considered in finance literature is that the disconnect may be driven by inconsistencies between the theoretical value of Beta under the CAPM and its estimated value under empirical measurement. Tofallis (2008) appears to be the first author to consider this disconnect and their paper is the most closely related paper to the content of this chapter. They focus on the theoretical role, and common industry interpretation, of Beta as a measure of relative volatility. They advocate that Beta estimates should instead be obtained via taking the geometric mean of the slopes from two least squares regressions; producing an estimate

[11] On page 105, the authors demonstrate that a market proxy with a return only 22 basis points lower than the efficient frontier is sufficient to generate zero cross-sectional correlation between stock returns and the market proxy. It is commonplace for both researchers and practitioners to use a cap-weighted index as market proxy. Hence, factor model parameter estimates will be heavily biased.

of beta which is equivalent to $\beta_i^* = (sign\ of\ correlation) * \frac{\sigma_i}{\sigma_M}$. We broadly agree with the arguments put forth by the author; that the common interpretation and use of Beta in industry is inconsistent with theory in the absence of perfect correlation, and that empirically measuring Beta as a function of relative volatility would bridge this disconnect. However, whilst their proposed resolution realigns estimation with the common theoretical interpretation of Beta, it begs the question of whether Beta is busted as a useful financial metric. Under their revised estimation approach virtually all Betas in the market will be identified as having a Beta of greater than one since almost all individual securities will exhibit a higher standard deviation than a highly diversified market portfolio. In a demonstration of the revised Beta estimation approach for Dow Jones Industrial Average (DJIA) constituents, the author presents revised Beta estimates of an average across securities of 1.98. Intuitively, this breaks the use of Beta in performance evaluation models since in the demonstrated case we would expect the securities to, on average, exhibit statistically negative alphas equivalent in magnitude to the market risk premium. If we adopt the resolution of Tofallis we have realigned estimation with theory but invalidated the use of Beta in many of its common financial applications. Borrowing an old medical adage; "*the operation was a success, but the patient died*".

This chapter contributes to existing literature in three important ways. Firstly, we explore the mismatch between the theory underpinning market Betas and their practical estimation. Our contribution is to develop a novel measure for quantifying the severity of estimation error that arises. Our novel measure allows us to examine the conditions in which the mismatch between theory and estimation creates the most significant bias to OLS Beta estimates. This allows us to predict that the Alpha associated with common finance anomalies, such as the small cap premium, may be partially attributable to the mismatch between theory and estimation.

Our second task is to examine the historical magnitude of estimation error arising from the mismatch between theory and OLS estimation for US equities. We decompose OLS beta estimates into correlation and relative volatility. This decomposition allows us to examine whether the mismatch between CAPM Betas and their empirical measurement is primarily driven by weak correlation between stock and index returns and observe the severity of mismatch bias over time. Our main contribution is to identify an economically large deviation between theoretical Betas and their OLS estimates which is persistent across time periods. When we regress theoretical Betas against OLS estimated Betas we observe a median R^2 across regressions of 0.05; indicating minimal relation between the two measures. As consequence, we identify OLS estimates of Beta as being inconsistent with their theoretical interpretation as a return multiple.

Our final task is to examine the severity of mismatch error for portfolios rather than individual securities. Our contribution is to show that portfolios are less vulnerable to Beta estimation error arising from the mismatch between theory and OLS estimation. However, the bias arising from mismatch error can still be severe. We demonstrate that both the size and value premium are at least partially attributable to mismatch error. In the next section we demonstrate the mismatch between the theoretical value of Beta and its OLS estimation.

2.3 Mismatch between theory and measurement

The Sharpe-Lintner CAPM (Sharpe, 1964; Lintner, 1965) can be represented as:

$$E[R_{i,t}] = \beta_i E[R_{M,t}], \qquad \textbf{(2.1)}$$

where i denotes an asset; t denotes a time increment; $R_{i,t}$ and $R_{M,t}$ are returns on asset i and the market portfolio (commonly a proxy) in excess of the risk-free rate, and nestled within expectation operators; and β_i is the exposure of asset i to market risk. The equation can be easily rearranged to isolate β_i:

$$\beta_i = \frac{E[R_{i,t}]}{E[R_{M,t}]}. \qquad \textbf{(2.2)}$$

Unsurprisingly, this algebraic rearrangement has led to widespread interpretation of β_i in industry as a return multiple. For example, the Corporate Finance Institute, which claims to be the largest and most recognized finance training, certification, and skill development platform in the world, adopts this common interpretation in their explanation of Beta: "*A company with a β that's greater than 1 is more volatile than the market. For example, a high-risk technology company with a β of 1.75 would have returned 175% of what the market returned in a given period*".[12] Historically, the most

---

[12] Source: https://corporatefinanceinstitute.com/resources/knowledge/valuation/what-is-beta-guide/

common method to estimate the market Beta for a security has been via use of an ordinary least squares (OLS) regression of a single index model:

$$R_{i,t} = \hat{\alpha}_i + \hat{\beta}_i R_{M,t} + \hat{\varepsilon}_{i,t}, \quad \textbf{(2.3)}$$

where variables are the same as Equation 2.1 but now a regression constant, $\hat{\alpha}_i$ (commonly referred to as "Jensen's Alpha"), and residuals, $\hat{\varepsilon}_{i,t}$, are included. The value for $\hat{\beta}_i$, estimated by an OLS regression, is:

$$\hat{\beta}_i = \frac{Cov(R_{i,t}, R_{M,t})}{Var(R_{M,t})}. \quad \textbf{(2.4)}$$

Since it has historically been common empirical practice to use either a market index or equal-weighted portfolio as a market proxy, the common usage of OLS to estimate CAPM β_i implies that the much of the past 70 years of finance research has rested on the assumption that the following equality holds:

$$\beta_i = \hat{\beta}_i, \quad \text{or} \quad \frac{E[R_{i,t}]}{E[R_{M,t}]} = \frac{Cov(R_{i,t}, R_{M,t})}{Var(R_{M,t})}. \quad \textbf{(2.5)}$$

In practice, it is extremely unlikely that this equality would ever hold in empirical data since the requisite assumption is that either:

a) all security returns exhibit either perfect correlation with the market portfolio, or

b) the market proxy chosen by the empiricist happened to be the unique mean-variance optimisation tangency portfolio for the selected data period.[13]

Perhaps this disconnect can be best demonstrated through the visualisation of Figure 2.1 below. The figure demonstrates 1,001 different price paths over 250 time periods. Each price path starts at the same origin and finishes at the same price level; that is to say that every single price path generates the same return multiple. However, every path is quite different so when regressed against a common market proxy will generate different OLS estimates of Beta. Figure 2.1 demonstrates that the same return multiple can be associated

[13] Note: use of a random mean-variance efficient portfolio is insufficient. Whilst use of a mean-variance efficient market proxy would ensure that all security returns lie on a common security market line, only the unique tangency portfolio ensures that the slope of the security market line is equivalent to the market risk premium ($E[R_{M,t}]$). Hence, only use of a tangency portfolio creates equivalence between β_i and $\hat{\beta}_i$.

with practically infinite different OLS estimated Betas. The converse is true, the same OLS estimated Beta can be associated with practically infinite different return multiples.

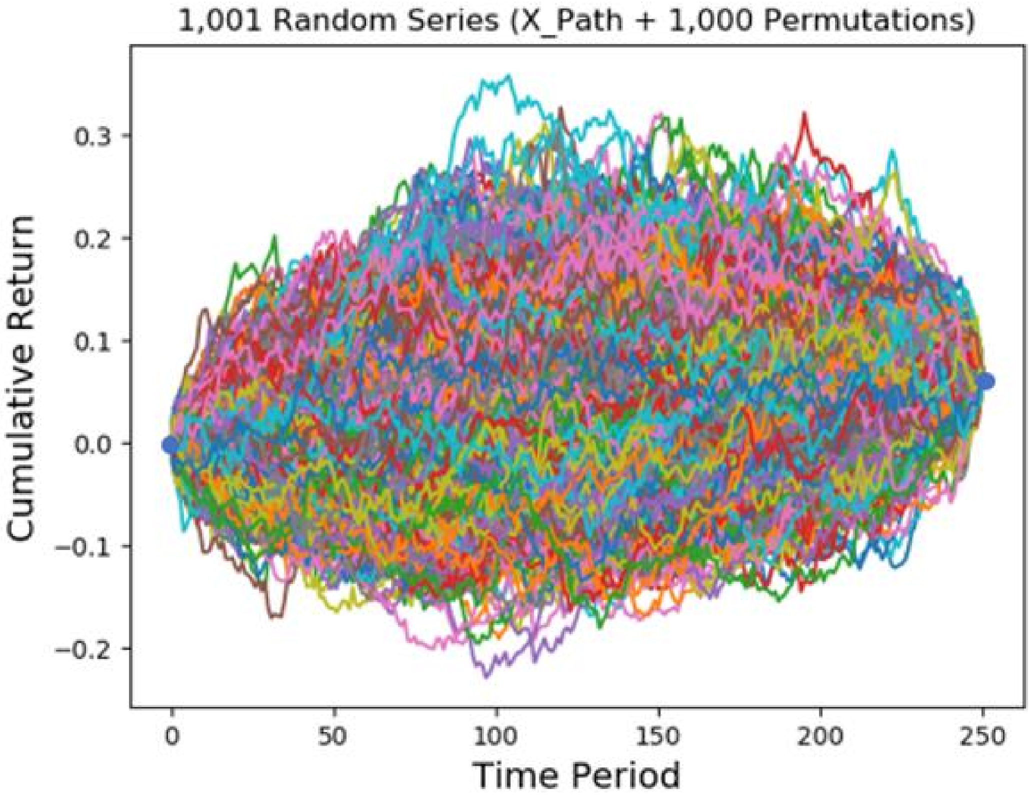


Figure 2.1: Many different price paths produce the same return multiple
Figure 2.1 shows 1,001 different price paths which all start at a common price and all end at a common price. When regressed against a common market portfolio, each of these price paths would produce the same market return multiple yet different OLS beta estimates. The equality of Equation 2.5 clearly does not hold in this simulation.

2.3.1 Imperfect market correlation as a source of disconnect

Figure 2.1 serves to assert that the Beta of the CAPM can have little to no relation to the Beta estimated via OLS. The consequence of this revelation is severe. It raises the possibility that academic finance has invested over 70 years examining why the CAPM fails via use of a tool that produces a metric which may have little to no association to the model being investigated. As a result, it is of great importance to understand how severe this mismatch between asset return multiples (theoretical interpretation) and their OLS estimated Betas (empirical measurement) has been in historical data to identify the consequences for academic finance and industry. To facilitate analysis, we start with a further decomposition of OLS Beta:

$$\hat{\beta}_i = \frac{Cov(R_{i,t}, R_{M,t})}{Var(R_{M,t})} = \rho_i * \tau_i \text{ ,} \quad \textbf{(2.6)}$$

where ρ_i indicates the correlation of security i with the market portfolio; and τ_i represents the relative volatility $(\frac{\sigma_i}{\sigma_M})$ of security i against the market portfolio. We note this decomposition is also used by Tofallis (2008). Our concern is how severe is the mismatch between β_i and $\hat{\beta}_i$ for a given data set. Unless our choice of market proxy was the unique mean-variance optimised (MVO) tangency portfolio, then the equality of Equation 2.5 will hold only when $|\rho_i| = 1$. Under this condition, Betas estimated via OLS are equivalent in magnitude to the Betas of CAPM theory; $(sign\ of\ \rho_i) * \tau_i$ is an empirical measurement equivalent to β_i. This empirical measurement, $(sign\ of\ \rho_i) * \tau_i$, is also consistent with the common interpretation of β_i; a security with a $\tau_i = 2$ is expected to have returns twice the magnitude of the market. However, as $|\rho_i| \rightarrow 0$, τ_i must increase exponentially to maintain the equality. Hence $(sign\ of\ \rho_i) * \tau_i$ no longer equals β_i, rather there is an exponentially increasing mismatch between β_i and $(sign\ of\ \rho_i) * \tau_i$. We formally denote this difference (δ_i) as "Beta Mismatch Error", or "BME":

$$BME_i = (sign\ of\ \rho_i) * \tau_i - \hat{\beta}_i. \tag{2.7}$$

In practice, most securities in a market tend to exhibit a positive correlation with a market index. We can interpolate how severe BME is for varying combinations of β_i and τ_i as we move from low correlations towards $\rho_i = 1$. Intuitively, BME should be equivalent to zero when $\rho_i = 1$ and increase as ρ_i decreases. Figure 2.2 depicts this relation.

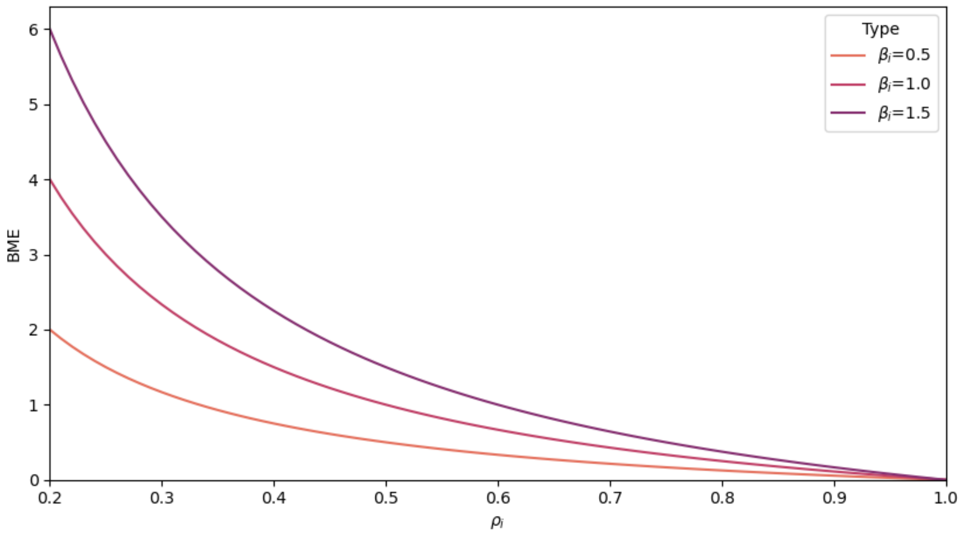


Figure 2.2: Severity of Beta Mismatch Error for increasing correlation

Figure 2.2 shows an exponentially increasing disconnect between the Beta of CAPM theory and the Beta estimated via OLS. As the correlation between the asset examined and the market index decreases, OLS estimates will progressively understate the relative volatility of this asset. Hence, an exponentially increasing disconnect between OLS estimates and their CAPM interpretation arises. BME approaches infinity as correlation approaches zero.

Alternatively, we can re-present BME as a ratio, denoted “Beta Mismatch Ratio”:

$$BMR_i = \frac{(sign\ of\ \rho_i) * \tau_i}{\hat{\beta}_i}$$

$$= \frac{(sign\ of\ \rho_i)}{\rho_i}. \qquad \textbf{(2.8)}$$

This allows us to more easily interpret the impact of the disconnect between β_i and $\hat{\beta}_i$. For example, a stock with a β_i of 2 and correlation of 0.5 with the market portfolio would have a BMR of 2. Whilst our OLS measure of Beta would lead us to interpret that the stock typically exhibits periodic returns 2x as large as the market, it typically exhibits return movements 4x as large as the market. A BMR of 2 indicates that a measurement equivalent to our CAPM theory interpretation would be 2x the value of what we estimated under OLS. Figure 2.3 re-presents Figure 2.2 in a ratio format. Since BMR is standardised against β_i, only correlation impacts BMR.

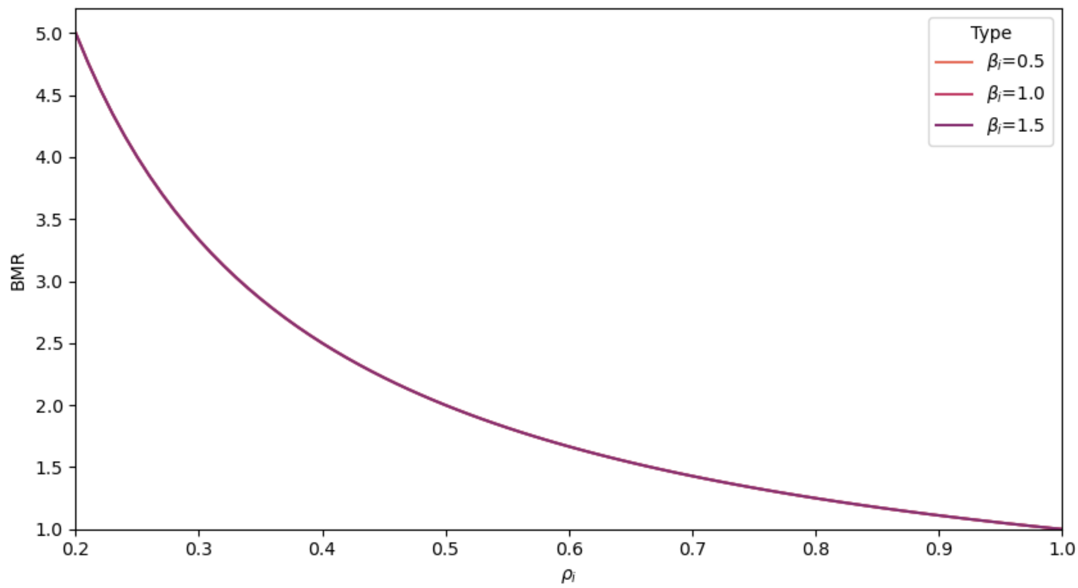


Figure 2.3: Beta Mismatch Ratio for increasing correlation

Figure 2.3 shows Beta Mismatch Ratio (BMR) across a range of positive correlations. As the correlation of an asset with the market index decreases, BMR exponentially increases. By a correlation of 0.5, BMR has already reached 2, indicating that OLS estimates of Beta understate relative volatility, the theoretical CAPM interpretation of Beta, by 50%. Consequently, the OLS estimate of alpha for this asset would be overstated by an amount equivalent to half of the market risk premium. BMR approaches infinity as correlation approaches zero.

There are some interesting insights arising from Figure 2.3. We can observe that BMR is close to 1 for stocks highly correlated with the market; use of OLS to estimate β_i doesn’t cause a severe disconnect between the Beta of CAPM theory and our measurement of it. However, BMR quickly becomes elevated as correlation declines. By a correlation of 0.5, our OLS estimates of β_i understate the true relative return movements of stocks by a factor of 2. Since large stocks like Apple comprise a large portion of their market benchmark, they are likely to have reasonably high correlations with the market and be less severely impacted by BME. By contrast, smaller stocks may exhibit much lower correlations with the market benchmark which would render their OLS estimated Betas as seriously misleading under the common interpretation of CAPM Betas.

2.3.2 Choice of market proxy as a source of disconnect

It is necessary to emphasize that even an empirical model which produces a perfectly fit security market line can be built upon OLS betas that have little to no relation to their theoretical CAPM counterparts. Roll (1978) demonstrates that if the market proxy used in a single-factor model is mean-variance efficient then it is mathematically certain that all test assets/portfolios will lie on a common security market line. A cross-sectional regression will produce a perfect fit! Even in this scenario where it appears the CAPM perfectly explains the cross-sectional asset returns, it is possible that theoretical CAPM betas and their OLS estimates will have little relation. In our Figure 2.4 below (left panel) we reproduce this "perfect fit" scenario as presented in Figure 2 of Roll's (1978) paper whereby the market proxy is a mean-variance efficient portfolio.[14] In the same figure we also plot the OLS estimates of betas vs their theoretical equivalents (middle panel) and the Pearson's correlation of each portfolio (right panel). The middle panel serves to indicate the large disconnect between the theoretical value of CAPM betas versus their OLS estimates. As shown in the right panel, this disconnect emerges because the test portfolios are not perfectly correlated with the market portfolio which is the required assumption for equivalency between CAPM Betas and their OLS estimates if the chosen market proxy is not the unique MVO tangency portfolio.

[14] We reproduce Roll's (1978) third column of Figure 2 using exactly the same data as presented in his Table 1 and Table 2.

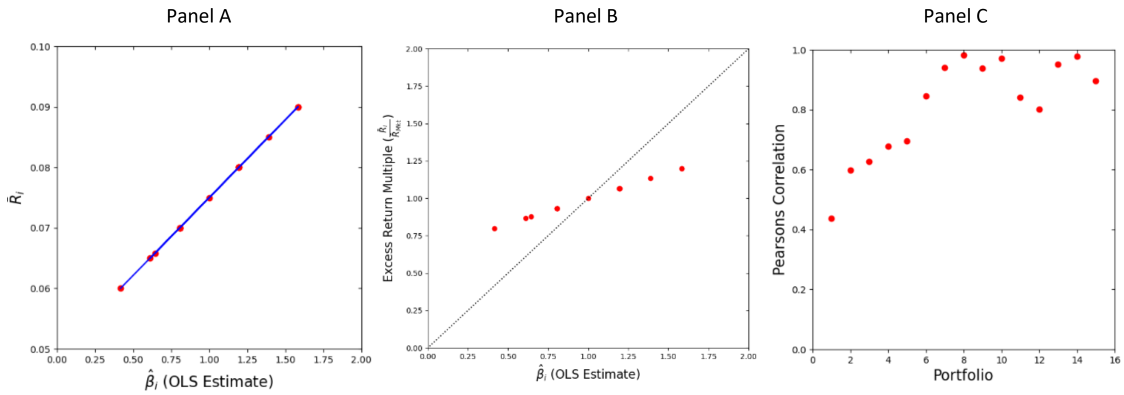


Figure 2.4: Beta mismatch still present in perfectly fit regressions

Figure 2.4 illustrates that even when an OLS regression can explain 100% of the variation in empirical data, a large difference between theoretical CAPM Betas and their OLS estimates can remain. Panel A reproduces the third plot of Figure 2 of Roll (1978); a figure which demonstrates that every asset will inevitably lie on a common security market line if a mean-variance efficient portfolio is chosen as a market proxy. Panel B plots theoretical CAPM return multiples (y-axis) against their OLS estimates (x-axis) for the data of Panel A. A dotted 45-degree line is also overlayed on the chart. If all data points fell on the 45-degree line it would indicate equivalence between theoretical CAPM Betas and their OLS estimates. However, it is evident that there is a difference between theoretical Betas and their estimated values under OLS for this data despite the use of a MVO market proxy. Panel C identifies the underlying driver of this disconnect; the portfolios included in Roll's simulation are not perfectly correlated with the market proxy. As a result, theoretical CAPM Betas and their OLS estimates are not equivalent.

In Roll's (1978) example, we can visually observe that his market proxy, whilst mean-variance efficient, was not a tangency portfolio. The plotted security market line produced a regression intercept of 4.93%. By contrast, it is a mathematical certainty for any data that if a mean-variance efficient tangency portfolio was chosen not only would all assets fall along a common SML but that SML would intercept the y-axis at precisely zero. As a result of this knowledge, it is possible to backsolve the unique mean-variance tangency portfolio for any data sample. An empiricist needs simply to generate an efficient frontier and then select the unique portfolio on that frontier that forces the y-intercept of a SML through zero. Alternatively, this optimisation is achieved by finding portfolio weights which produce the highest Sharpe ratio. In Figure 2.5 we reproduce the panels of the preceding figure using the same data but instead using the unique mean-variance tangency portfolio as our market proxy.

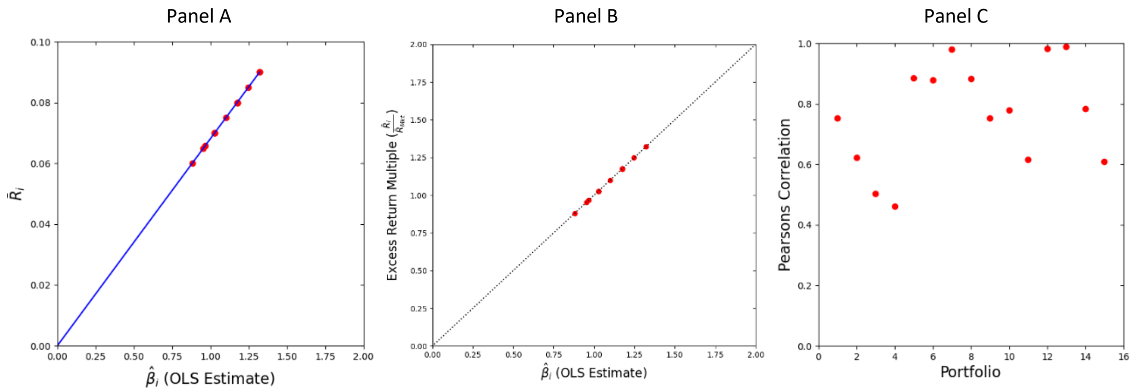


Figure 2.5: No Beta mismatch when using tangency portfolio

Figure 2.5 illustrates that when the chosen market proxy is a mean-variance efficient tangency portfolio, there is no difference between theoretical CAPM Betas and their OLS estimates. Panel A reproduces the third plot of Figure 2 of Roll (1978) but uses a tangency portfolio as the market proxy. Panel B plots theoretical CAPM return multiples (y-axis) against their OLS estimates (x-axis) for the data of Panel A. A dotted 45-degree line is also overlayed on the chart. It is evident that there is no difference between theoretical Betas and their estimated values under OLS when a tangency portfolio is used as a market proxy. Panel C shows the equivalence remains despite imperfect correlation.

In Panel A, all portfolios fall on a common SML, and this line intercepts the y-axis at zero. As a result of our use of the tangency portfolio as our market proxy, In Panel B, our theoretical values of Beta and their OLS estimates perfectly match. As shown in Panel C, this occurs despite the test portfolios still exhibiting imperfect correlation with the tangency portfolio.

We have uncovered that there are two major reasons for disconnect between OLS estimates of Beta, and their theorised values under the CAPM. It is commonplace in academia and industry to utilise a market-capitalisation-weighted index as a market proxy. It is extremely unlikely that by chance this proxy would happen to be the unique mean-variance efficient tangency portfolio required in CAPM theory. As a result, OLS estimates of Beta are likely to be disconnected from their theoretical values and this mismatch will exponentially increase as the correlation of stocks with the chosen market proxy declines. Therefore, there is a clear motivation to study the historical correlation between stocks and market proxies in empirical data such that we can ascertain how severe the mismatch error has been historically. Furthermore, systematic differences in correlation between stocks based on characteristics such as size could plausibly generate a number of perceived CAPM anomalies. In Section 2.4, we find support for the idea that BME is a contributor to the famed small cap and value premiums. However, we first examine the historical magnitude of Beta Mismatch Error for US stocks in the following section of this chapter.

2.4 Historical Beta Mismatch Error in US stocks

Given the potential for the OLS estimates of market Betas to greatly differ from their theoretical values, it is important to study how substantial this mismatch has been historically. To this end, we examine 30 years of daily stock data for US equities spanning the time period from 1 January 1990 to 31 December 2020. This equities data is obtained from the Center for Research in Security Prices (CRSP). Only ordinary common shares listed on the NYSE, NASDAQ, and AMEX are included in our analysis.[15] We omit all microcaps, defined as stocks smaller than the 20th percentile of the market equity of NYSE stocks at the beginning of the first trading day of a calendar year, from our sample. In the calculation of holding period returns we assume all dividends received are reinvested in the underlying stock on the date that the dividends are paid out. Unless otherwise stated, the "market factor" used in factor models is constructed as the excess holding-period return of a market-capitalisation-weighted index comprised of the 500 largest stocks by market capitalisation, formed on the first trading day of January, and held one year. The daily risk-free rate used is the monthly rate obtained from Kenneth French's website and converted into a daily format.

2.4.1 Historical Beta Mismatch Ratio

We know from Equation 2.8 that BMR is a function of an asset's correlation with the chosen market proxy. If a market proxy other than the tangency portfolio is selected, then OLS estimates of Beta become exponentially more biased as correlation decreases further below 1. Figure 2.6 below plots the time-varying correlation of US equities with a market-cap-weighted index for our 30-year sample period. The average of the median stock correlation across the sample period is 0.49, with correlation generally trending upwards over time.

[15] CRSP code filters applied: "EXCHCD" = [1,2,3], "SHRCD" = [10,11], "RET"≠ [-66, -77, -88, -99, C].

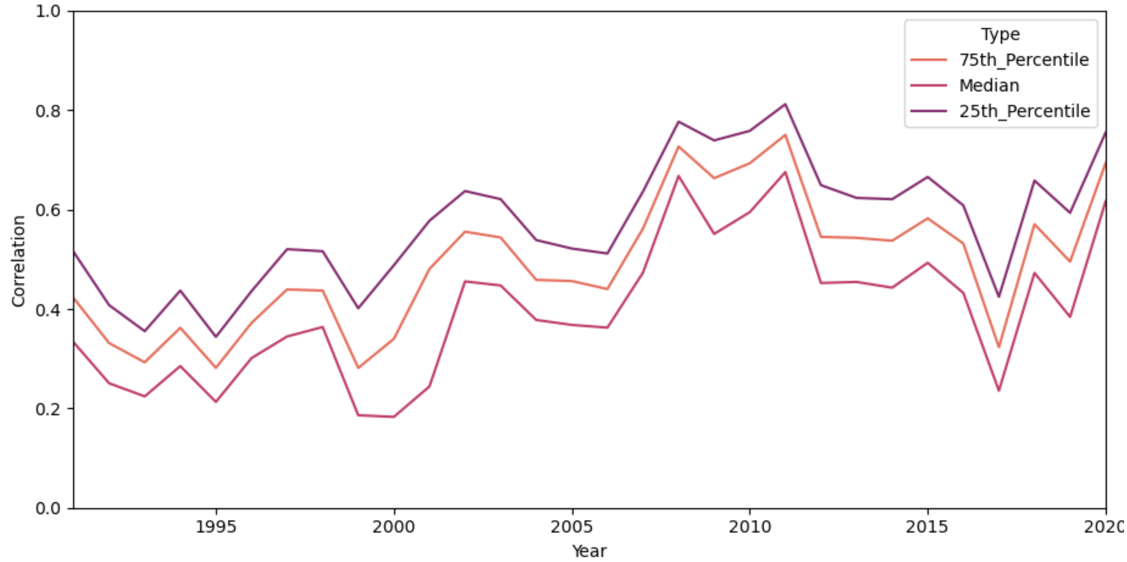


Figure 2.6: Historical stock correlation with cap-weighted index
Figure 2.6 plots the annual, non-overlapping correlation (y-axis) of stocks with a market-cap-weighted index of the top 500 largest stocks for each year of our sample period. Beta Mismatch Error becomes exponentially more severe as the correlation of stocks with the chosen market proxy declines from 1. The figure shows a median stock correlation of around 0.5 with the chosen market proxy; well below the 1.00 required for equivalence between CAPM Betas and their OLS estimates. This correlation is generally trending upwards over the sample period.

We can perceive from Figure 2.6 that OLS estimates of Beta will be biased since correlation is consistently well below 1. The severity of this bias across time is presented in Figure 2.7 which plots the BMR for different stock percentiles across the 30-year sample period.

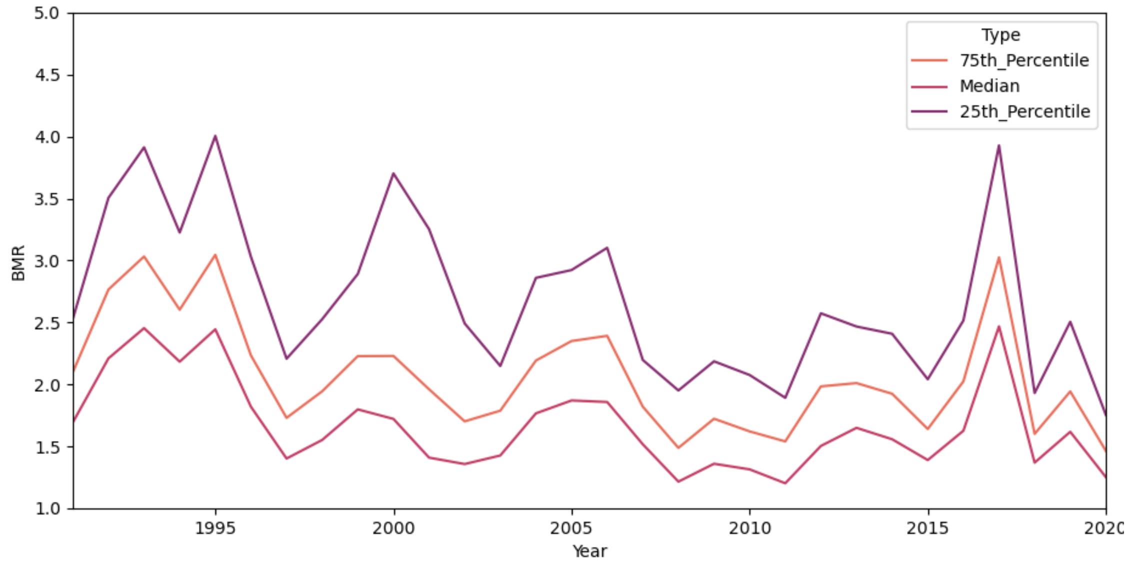


Figure 2.7: Historical BMR with cap-weighted index as market proxy
Figure 2.7 plots the annual, non-overlapping Beta Mismatch Ratio (BMR) of stocks (y-axis) for each year of our sample period when using a market-cap-weighted index of the top 500 largest stocks as a market proxy. BMR becomes exponentially more severe as the correlation of stocks with the chosen market proxy declines from 1. The figure shows a median stock BMR of around 2 suggesting that OLS estimates of Beta equivalent to CAPM theory would be approximately twice the magnitude of the current OLS estimates obtained; there is a significant disconnect between measures.

The average of the median BMR across the sample period is 2.07, with a downwards trend over time as market correlation increases. To interpret this number, for a median stock, a Beta estimate consistent with CAPM theory would be approximately 2.07x the size estimated under OLS; an economically severe magnitude of bias.

We can regress CAPM theoretical Betas against their OLS estimates to quantify the magnitude of disconnect between the measures. This regression is show in Equation 2.9 below:

$$\beta_i^{Theory} = \hat{\lambda}_0 + \hat{\lambda}_1 \hat{\beta}_i + \hat{\varepsilon}_i, \quad \textbf{(2.9)}$$

where $\beta_i^{Theory} = \frac{E[R_i]}{E[R_M]}$, $\hat{\lambda}_0$ is an estimated regression constant, $\hat{\lambda}_1$ is an estimated slope coefficient, $\hat{\beta}_i$ are OLS estimates of Beta, and $\hat{\varepsilon}_i$ are estimated regression residuals. Intuitively, if both measures were equivalent, a regression of one upon the other would produce a perfectly fit regression line with $\hat{\lambda}_0 = 0$, $\hat{\lambda}_1 = 1$, and a regression $R^2 = 1$.

The year-by-year full sample results of this mismatch between CAPM theoretical Betas and their OLS estimates are detailed in Table A.1 in the appendix. The relation between CAPM theoretical Betas and their OLS estimates is very weak for our market-cap proxy across our 30-year sample. The median model R^2 across years is 0.05; suggesting an extremely noisy relation between the two measures which are supposed to be identical. The median value of $\hat{\lambda}_0$ is 0.30, and statistically significant in 22 of the 30 sample years at a 95% confidence level. The median value of $\hat{\lambda}_1$ is 0.80, and statistically significant in 23 of the 30 sample years at a 95% confidence level. Taken together, these parameter estimates suggest that OLS estimates of Beta are consistently underestimated relative to the Beta of CAPM theory with the distortion between measures substantial enough to render OLS Betas close to meaningless. Figure 2.8 below presents a visualisation of this disconnect for the first and final years of our sample period.

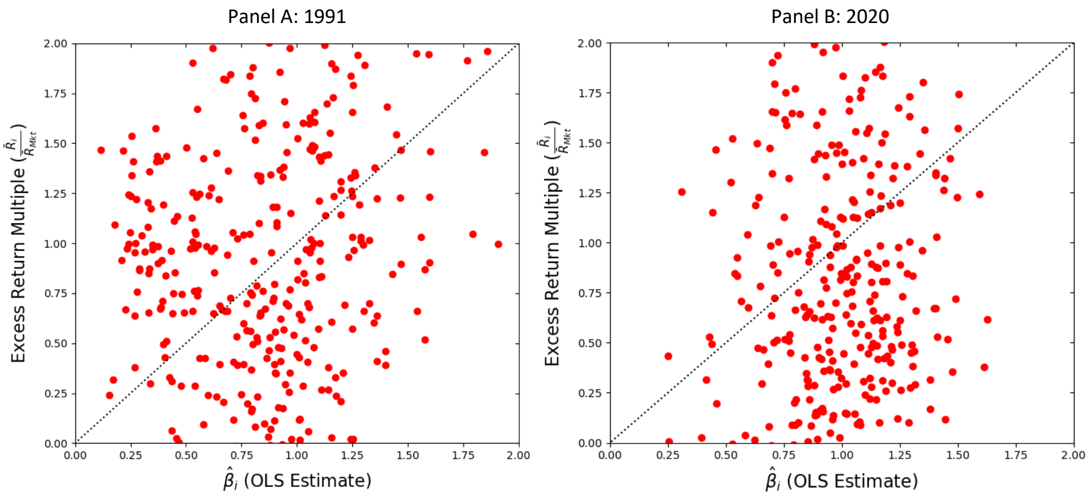


Figure 2.8: CAPM Theoretical Betas vs. OLS estimates for cap-weighted proxy
Figure 2.8 plots theoretical CAPM return multiples (y-axis) against their OLS estimates (x-axis) for the first (Panel A) and last (Panel B) years of our sample. A dotted 45-degree line is also overlayed on the chart. It is evident that there is minimal relation between theoretical Betas and their estimated values under OLS when a cap-weighted portfolio is used as a market proxy.

2.4.2 Historical tangency portfolio

In practice, the mismatch between the theoretical values of Beta and their OLS estimates can only be averted if an empiricist selects, as a market proxy, the unique tangency portfolio theorised by Markowitz (1952). Despite empiricists commonly using a market-capitalisation-weighted index as their market proxy, Markowitz theorised that all investors would hold a combination of the tangency portfolio and the risk-free asset. However, cap-weighted indexes have historically been used as a market proxy due to convenience, creating the potential for a significant disconnect to emerge between CAPM theory and empirical estimation when examining historical data. A single factor model using the unique tangency portfolio is presented in Equation 2.10 below:

$$R_{i,t} = \hat{\alpha}_i + \hat{\beta}_i R^*_{M,t} + \varepsilon_{i,t}, \tag{2.10}$$

where parameters are the same as Equation 2.3 but $R^*_{M,t}$ is now a tangency portfolio rather than some alternate market proxy. This tangency portfolio can always be *ex-post* identified for any data set and used in historical performance evaluation.

Roll (1978) demonstrated that any mean-variance efficient portfolio chosen as a market proxy will generate a SML that produces a perfectly fit regression line ($R^2 = 1$). However, of these mean-variance efficient portfolios, the unique tangency portfolio is the one which produces the highest Sharpe ratio. This opens a clear path to *ex-post* identification of tangency portfolios for any returns data. We can simply form a portfolio which historically maximised the Sharpe ratio.

We use this Sharpe maximisation process each year for our 30-year sample period to identify the *ex-post* tangency portfolios. We can again deploy the regression presented in Equation 2.9 to compare CAPM theoretical Betas against their OLS estimates. The key difference is that this time our market proxy is the unique tangency portfolio in each year as opposed to a cap-weighted index. The year-by-year full sample results of this mismatch between CAPM theoretical Betas and their OLS estimates are detailed in Table A.2 in the appendix. Unsurprisingly, the relation between CAPM theoretical Betas and their OLS estimates is perfect in each year of our sample when using a tangency portfolio as our market-cap proxy.[16] In each year the model R^2 is 1.00, $\hat{\lambda}_0$ is 0, and $\hat{\lambda}_1$ is 1. The results demonstrate that use of a market proxy consistent with the tangency portfolio that Markowitz (1952) theorised investors would hold restores the equivalence between theoretical Betas and their estimated values under OLS. By contrast, use of a generic *ex post* mean-variance-efficient market proxy does not restore equivalence (e.g. see Sinclair, 1987, Table 1). Unfortunately, this restoration has come at a great practical cost. The tangency portfolio requires no constraints to short selling nor to the size of weighting per constituent other than the condition that all weights must sum to 1.

In practice, there are often real-world constraints to an investor's ability to short sell some assets within a market. However, an attempt to apply realistic constraints, whilst logical, would again sever the connection between the theoretical value of CAPM Betas and their OLS measurement. If the market proxy selected is not the tangency portfolio then the empiricist would alternatively require perfect return correlation between the market proxy and test assets in order for equivalence between OLS and theory Betas. As shown in Figure 2.6 of the preceding section, historical stock return correlations with the most commonly used market proxy (cap-weighted) have been far from perfect. Suppose we did, by miracle, find historical data in which asset returns were perfectly correlated with our cap-weighted proxy. In such a case, theoretical and OLS Betas would be equivalent. However, even in this scenario, the OLS Betas obtained might not be useful.

---

[16] Since at least Roll (1977) it has been known that if the market proxy used is *ex post* mean-variance-efficient then an OLS regression of security returns against the market proxy is certain to produce an R2=1. However, this doesn't induce equivalence between theoretical Betas and their OLS estimates (e.g. see Figure 2.4, Panel B).

This is because the return of every asset would now lie on a common security market line and we would have no means to rank the risk-adjusted return superiority of any asset; it's a dead heat. As a result, there seems to be no practical way to obtain a meaningful Beta estimate that could be used in a single-factor model to evaluate risk-adjusted performance, nor is there a clear path forward. OLS Betas appear busted as a practical measure of systematic risk.

2.5 Beta Mismatch Error and prominent finance anomalies

The prior section examined the historical disconnect between the theoretical and OLS-estimated values of Beta for individual US equities. It is also common in finance to evaluate the performance of portfolios rather than individual stocks. Therefore, it is important to understand whether this disconnect between theory and estimation also affects performance evaluation of portfolios and to what extent.

2.5.1 Portfolio formation reduces Beta Mismatch Error

It seems likely that BME would likely be less severe for randomly formed portfolios due to the expectation of higher correlation with common market proxies such as cap-weighted or equal-weighted market portfolios. To verify this intuition, we perform a simulation that shows the impact of adding incrementally more securities to a portfolio.

First, we generate 252 days of stock returns for 500 securities based on a model of the CAPM with noise as shown in Equation 2.11:

$$R_{i,t} = \beta_i R_{M,t} + \varepsilon_{i,t}, \tag{2.11}$$

where $R_{i,t}$ is the excess return of security i at time t; $R_{M,t}$ is the excess return of the unobservable market portfolio at time t with $R_{M,t} \sim N\left(\frac{0.10}{252}, \frac{0.15}{\sqrt{252}}\right)$; $\beta_i \sim N(1, 0.2)$; and $\epsilon_{j,t} \sim N\left(0, \frac{0.23}{\sqrt{252}}\right)$ is noise. Next, we randomly form an observable equal-weighted portfolio of N constituents, without replacement, where N ranges from 1 to 100. We then calculate BMR and Pearson's correlation for each portfolio. We repeat this simulation 1,000 times for each value of N to smooth out results. Figure 2.9 displays the median results of the simulation at each constituent level.

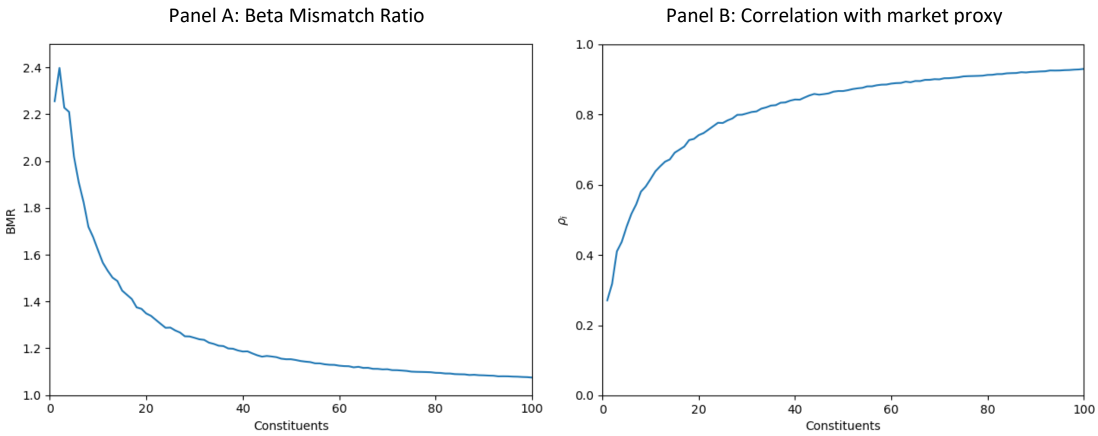


Figure 2.9: Portfolio formation reduces Beta Mismatch Error
Figure 2.9 demonstrates the effect of random portfolio formation upon Beta Mismatch Error. Panel B shows that forming progressively larger portfolios of randomly selected constituents has the effect of increasing portfolio correlation with the market proxy (an equal-weighted portfolio of all securities in the market). Panel A shows that this increasing correlation rapidly decreases Beta Mismatch Error with BMR declining towards 1.

Panel A shows that the BMR rapidly declines as more constituents are added to a portfolio. Panel B indicates that this is caused by an increasing correlation of the progressively larger portfolio with the market proxy. Since the market proxy is an equal-weighted portfolio, if a simulated portfolio included every available constituent, then correlation would equal precisely 1 as the constituent portfolio becomes identical to the market proxy. In Figure 2.9, BMR remains noticeably above 1 even when it is a large-diversified portfolio of 100 constituents. However, it's evident that the severity of BME is heavily moderated for large, randomly formed portfolios when compared to individual securities.

2.5.2 Impact of Beta Mismatch Error upon Fama French factors

Beta Mismatch Error may be substantially reduced for randomly formed portfolios but could it remain problematic for portfolios intentionally constructed based on defined stock characteristics? The FF3F Model (Fama and French, 1993) extends upon the CAPM by introducing two additional factor portfolios which are presumed to proxy for additional systematic risks not captured by the market portfolio. The inclusion of these additional factors was initially inspired by the ability of each factor portfolio to produce a statistically significant Alpha when regressed against a single-factor model; a finding which was robust across extensive time periods and a variety of markets. At present, the FF3F model is perhaps the most widely used performance evaluation model within finance academia.

The size and value premium portfolios included in the FF3F model are formed non randomly; rather they are formed on defined stock characteristics. It is therefore possible that either or both of these portfolios could be afflicted by abnormalities in correlation that somehow explain their apparent outperformance within a CAPM framework. To investigate, we examine daily factor portfolio data, obtained from Kenneth French's data library, covering the period from 1st January 1991 to 31st December, 2020.[17] For both the Small-Minus-Big (**SMB**) and High-Minus-Low (**HML**) factors we regress the returns of the constituents portfolios against the market proxy (a capitalisation-weighted index provided by Kenneth French) to estimate OLS alphas and Betas from a single-factor model. We also estimate an adjusted, CAPM-consistent Beta and alpha using the methodology of Tofallis (2008).[18] Finally, we again estimate OLS Betas and Alphas when instead using a tangency portfolio as the market proxy in a single-factor model.

Table 2.1 presents the estimation results for the SMB factor. The small cap portfolio exhibits a noticeably larger OLS Alpha estimate than the mid and large-cap portfolios under a single-factor model regression that used a cap-weighted market proxy. It also has a substantially lower correlation than the other portfolios. The lower correlation biases OLS estimates of Beta downwards which has the effect of insufficiently penalising risk-adjusted returns, inflating Alpha estimates. Tofallis (2008) introduced an alternate approach to estimating OLS Betas given the disconnect between theory and OLS that arises in the presence of imperfect correlation. Beta estimates for his approach are also presented in Table 2.1. These adjusted market Betas result in a substantial decrease in the Alphas of the small cap portfolio since its returns are now appropriately penalised for systematic risk. The small cap portfolio becomes the worst performer. Alternatively, if a tangency portfolio was used as the market proxy in a single-factor model then each test portfolio produces an OLS estimated Alpha of exactly zero;

---

[17] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[18] $\beta_i^{Tofallis} = (sign\ of\ correlation) * \frac{\sigma_i}{\sigma_M}$

reverberating Roll's (1978) critique that the SML can not be used to rank risk-adjusted performance.

	Small Cap	Mid Cap	Large Cap
ρ_i (Cap-Weighted Proxy)	0.86	0.94	1.00
$\bar{R}_i$	7.38%	7.42%	6.46%
$\hat{\beta}_i$ (OLS, Cap-Weighted)	0.98	1.04	0.99
$\hat{\alpha}_i$ (OLS, Cap-Weighted)	1.15%	0.68%	0.15%
$\hat{\beta}_i$ (Tofallis)	1.14	1.11	0.99
$\hat{\alpha}_i$ (Tofallis)	-0.26%	0.08%	0.11%
ρ_i (Tangency Proxy)	0.92	0.95	0.98
$\hat{\beta}_i$ (OLS, Tangency)	1.06	1.06	0.98
$\hat{\alpha}_i$ (OLS, Tangency)	0.00%	0.00%	0.00%

Table 2.1: SMB portfolio parameters under alternate estimation approaches
Table 2.1 examines the constituents of a SMB portfolio and estimates of their single-factor model Alphas and Betas using varying estimation approaches. The small cap premium appears to be driven by an abnormally low correlation with the market proxy which results in downwards biased estimates of Beta under OLS. Adjusting for this bias removes the Alpha associated with the small cap premium.

Table 2.2 presents the same results for the HML factor. The value portfolio exhibits the highest OLS estimate of Alpha over the sample period. Similarly to the small-cap portfolio, this Alpha is inflated by an abnormally low correlation with the market proxy. After adjusting for the induced bias using Tofallis' estimation method, the value portfolio becomes the worst performer on a risk-adjusted basis. As expected, if a tangency portfolio was used as the market proxy, no portfolio would be capable of producing a non-zero alpha.

	Growth	Medium	Value
ρ_i (Cap-Weighted Proxy)	0.98	0.96	0.89
$\bar{R}_i$	7.16%	6.29%	7.91%
$\hat{\beta}_i$ (OLS, Cap-Weighted)	0.99	0.95	1.06
$\hat{\alpha}_i$ (OLS, Cap-Weighted)	0.85%	0.29%	0.98%
$\hat{\beta}_i$ (Tofallis)	1.01	1.00	1.20
$\hat{\alpha}_i$ (Tofallis)	0.69%	-0.09%	-0.24%
ρ_i (Tangency Proxy)	0.99	0.93	0.88
$\hat{\beta}_i$ (OLS, Tangency)	0.97	0.90	1.03
$\hat{\alpha}_i$ (OLS, Tangency)	0.00%	0.00%	0.00%

Table 2.2: HML portfolio parameters under alternate estimation approaches
Table 2.2 examines the constituents of a HML portfolio and estimates of their single-factor model Alphas and Betas using varying estimation approaches. The value premium appears to be driven by an abnormally low correlation with the market proxy which results in downwards biased estimates of Beta under OLS. Adjusting for this bias removes the Alpha associated with the small cap premium.

It is evident that the historical OLS estimated Alphas associated with both the small cap premium and value premium are a product of market Beta estimation bias induced by abnormally low correlations with the chosen market proxy. Given the widespread prevalence of apparent finance anomalies, which by definition should be scarce, it is possible a common cause such as abnormally low market proxy correlation could be causing widespread false identification of countless anomalies by producing understated Beta estimates.

2.6 Conclusion

This chapter identifies the disconnect between the values, and common industry interpretation, of market Betas under CAPM theory versus the values they inherit under OLS estimation. This disconnect can be averted if an empiricist uses a tangency portfolio as a market proxy in performance evaluation models as intended by Markowitz (1952). However, the use of an unconstrained tangency portfolio as a market proxy is seldom feasible. As a result, historical implementations of single-factor models in both academia and industry have heavily relied on use of cap-weighted market indexes as a market proxy.

The common adoption of a cap-weighted portfolio as market proxy has produced a large disconnect between the values of Beta estimated via OLS versus their CAPM interpretation. For the past 30 years of US equities data, OLS estimates of Beta have been understated by, on average, around 50%. This accidental mismatch between theory and measurement has produced the identification of a number of apparent anomalies. Perhaps most famously, "small-cap" and "value" portfolios have historically produced statistically significant alphas across a variety of equities markets. Billions of dollars have been deployed into investment strategies which aim to take advantage of these perceived anomalies. As demonstrated in the preceding section, the Alpha of these apparent anomalies are driven by abnormally low correlations with the common market proxy which results in understated OLS estimates of Beta for these portfolios. Each of these portfolios benefit from estimation bias in a performance evaluation framework since the risk-penalty to the Alpha of each portfolio is excessively low. When using an appropriate market proxy, the tangency portfolio, the Alpha of these anomalies reverts to zero.

Unfortunately, use of a tangency portfolio in empirical analysis doesn't improve the usefulness of OLS estimates of Beta. If a tangency portfolio is used in a performance evaluation context, we restore the link between the values of Beta under CAPM theory and their OLS estimates. However, we introduce a more severe practical issue. As demonstrated by Roll (1978), every asset in the market will produce an identical Alpha of zero; there can be no empirical distinction of superior risk-adjusted performance when using a CAPM framework. Whilst an inspiring theoretical model, it appears that empirical estimation of the market Betas of the CAPM is flawed beyond restitution. Beta, as a practical measure of systematic risk for securities, is busted. In conclusion, whilst the CAPM may be one of the most widely studied and implemented financial models in history, its empirical use as a means to evaluate risk-adjusted outperformance is likely invalid.

Chapter 3

Rethinking Performance Evaluation: Is Alpha Reliable in Practice?

3.1 Introduction

For almost 60 years the Capital Asset Pricing Model (CAPM) (Sharpe, 1964; Treynor, 1962; Lintner, 1965; Mossin, 1966) has offered researchers and practitioners a standardised approach to price risky securities. The model's simplicity, establishing a linear relation between expected return and systematic risk (Beta)[19], resulted in widespread adoption of the CAPM in real-world corporate finance and investments.[20] Maximisation of "Alpha"[21] quickly became a focal point of active investment practitioners globally. However, reliance on the CAPM in financial decision making is likely problematic.

Eugene Fama and Kenneth French claim that the empirical record of the CAPM is "*poor enough to invalidate the way it is used in applications*" (2004, p. 25). Fama and French initially advocated use of an alternate three-factor model (1993), which included additional presumed risk factors emerging in finance literature, as an improvement over the CAPM. Their model gained widespread adoption in academia. In the decades that followed, return anomalies against the advocated three-factor model have become increasingly prevalent in empirical finance papers. Fama and French (2015) have since (*controversially*) extended their multifactor model to incorporate five presumed risk factors.[22] By 2020, there were over 450 published return anomalies in accounting and finance journals (Hou et al., 2020).

---

[19] $(E[r_i] - r_f) = \beta_i(E[r_M] - r_f)$ is the Sharpe (1964) – Lintner (1965) CAPM, where $E[r_i]$ is the expected return of asset i, r_f is a time-invariant risk-free rate of return, $E[r_m]$ is the expected return of the market portfolio comprising all assets, and β_i is a measure of the systematic risk of asset i. In the CAPM, the expected excess return of a given asset is a linear function of its systematic risk multiplied by the market risk premium.

[20] Graham & Harvey (2001) found that 73.5% of US CFOs always, or almost always, used the CAPM when estimating the cost of capital for prospective investments.

[21] Alpha is broadly defined throughout this chapter as the key risk-adjusted excess return of a linear return factor model. For example, in a single-factor model based on the CAPM, "Alpha" would refer to "Jensen's Alpha" (see Jensen, 1968).

[22] The five-factor model has met significant academic opposition from both competing multifactor models that demonstrate superior empirical performance (e.g. Hou et al., 2015) and from an inability of the extended model to rationalise the chosen risk factors (e.g. see Blitz et al., 2018 for a list of crticisms).

The rapid expansion of identified return anomalies is one of the great puzzles of modern finance. Anomalies should be rare *by definition*. Is there some unidentified deficiency in the theory underpinning commonly used factor models? Is the increasing prevalence of anomalies simply the result of data mining? Or is there an entirely different cause? To resolve an outstanding puzzle, a novel approach is required. This chapter is inspired by one simple idea not yet explored by our field; for an anomaly to be a *finance* anomaly it should not exist in random data. Any anomaly which exists in random data is likely a function of some form of systematic bias in an applied model, rather than a meaningful reflection of real-world investment practice.

The first major contribution of this chapter is therefore to demonstrate that many return anomalies in finance are likely the result of model misspecification. In this chapter we simulate random stock return data and obtain the Alpha estimates from a time-series regression of a single-factor model.[23] We observe that the famous "low beta anomaly" (e.g., Frazzini and Pedersen, 2014) of finance is evident in random data. This occurs because time-series regressions impose an arbitrary market risk premium assumption that is not informed by the underlying data; the model is misspecified.[24] This model misspecification induces a systematic bias whereby estimates of Alpha become a linear function of Beta estimates. Unfortunately, the time-series regression approach to Alpha estimation remains popular in academia and remains the method taught to practitioners by the leading industry bodies such as the Chartered Financial Analysts (CFA) Institute.[25] However, the remedy is straightforward. Cross-sectional regression can be used to obtain unbiased Alpha estimates. Through derivation, we demonstrate that the residuals of cross-sectional regressions are equivalent to Alpha estimates of time-series regressions that have been adjusted for model misspecification.

The second contribution of this chapter is to quantify the magnitude of model misspecification bias for a 30-year sample of historical US data. When real-world investment patterns differ from finance theory, model misspecification is likely to be a particularly severe issue. Since at least the 1970s, academics and practitioners alike have

---

[23] Used to represent a regression of the CAPM. $R_{it} = \hat{\alpha}_i + \hat{\beta}_i R_{Mt} + \hat{\varepsilon}_{it}$, where R_{it} is the excess return of security i in period t; $\hat{\beta}_i$ represents the exposure of security i to non-diversifiable risk; and $\hat{\alpha}_i$ represents the risk-adjusted return of security i in excess of expectations.

[24] Time-series regressions impose an assumption that for every unit of $\hat{\beta}_i$, R_{it} will increase by R_{Mt}. This assumption does not change even if the underlying data exhibits the cross-sectional properties that $cor(\hat{\beta}_i, \bar{R}_M) = 0$.

[25] CFA Level 1 Certification. Corporate Finance and Portfolio Management, page 376.

commonly observed that the security market line is "flatter than expected" (Black, Jensen and Scholes, 1972) by CAPM theory. We develop a generalised approach to quantify the model misspecification of each risk premia in a multi-factor model. We compare the Alpha estimates obtained via time-series regression of factor models against the equivalent metric obtained from cross-sectional regressions. We find that alpha estimates via time-series regression of a single-factor model would have been biased by a market risk premium that was overstated by 7.16% during our sample period. Alpha estimates from a Fama-French 3-Factor (FF3F) model (Fama and French, 1993) would have been biased by a market risk premium that was 2.04% too high, a size factor premium that was 4.61% too high, and a value factor premium that was 5.82% too high. We find that Alpha estimation bias is most severe for assets with Betas that differ greatly from the sample average. For the FF3F model, the time-series Alpha estimate for a zero-beta security would be inflated by 6.67% on average during our 30-year sample.

The third contribution of this chapter is to demonstrate the ease with which new spurious return anomalies can be generated in the presence of model misspecification. For example, when the security market line is "flatter than expected", one needs simply to find two well-diversified portfolios with different Betas and form a long-short portfolio. The "betting against beta" (BAB) factor identified by Frazzini and Pedersen (2014) is perhaps the clearest academic example of a spurious return anomaly identified due to model misspecification. However, the problem extends beyond spurious finance anomalies. Genuine return anomalies can likewise have the magnitude of their estimated Alphas falsely enhanced by using a long-short portfolio formation.

The insights from this chapter have several significant implications for academic finance and industry. Firstly, we highlight that a common approach to Alpha estimation is heavily biased due to model misspecification. When models are misspecified, the Alpha estimates obtained are largely spurious rather than an accurate reflection of "risk-adjusted" returns. The implication is that a large portion of published results in asset pricing literature may be exaggerated, if not invalid. We advocate an alternate Alpha estimation approach that is robust to model misspecification. Secondly, we demonstrate that historical stock returns in the US greatly differ from the assumptions of the CAPM and FF3F model. Consequently, use of time-series regression to estimate Alphas results in an economically large bias. To the extent that Alpha influences investment decisions, real world investment allocation can be heavily distorted. For example, the bias induced would lead investors to develop an excessive preference for low-beta, high-leverage investment strategies such as that of Warren Buffet.[26] We could also see an undue popularity for various 'smart beta' strategies. Finally, we explain why it is so common to identify return anomalies in finance when using time-series regression and provide a rationale as to why *long-short* portfolio formations are frequently used. If researchers adopt implementations of factor models that are robust to model misspecification then the identification of anomalies should become scarce, instead of commonplace, moving forward.

The remainder of this chapter is arranged as follows. Section 3.2 contextualises the historical development of alpha and explains why time-series alpha estimates are typically biased. Section 3.3 derives an adjustment for the model misspecification of time-series regressions and demonstrates an equivalent measure can be obtained directly via cross-sectional regression. In addition, we demonstrate via simulation how model misspecification directly biases alpha estimates and tests of statistical significance. Section 3.4 quantifies the historical magnitude of model misspecification for US stocks for a single-factor model and FF3F model. Section 3.5 prescribes a method to generate a practically infinite quantity of 'anomalies' by exploiting the alpha estimation bias of time-series regressions. Section 3.6 summarises our main findings, discusses limitations and opportunities for further research, and then concludes.

3.2 Related literature & the problem of model misspecification

One of the fundamental ideas of finance is the trade-off between risk and expected return. Markowitz (1952) formalised the idea that investors require higher expected returns to be enticed to undertake investments that carry greater risk. More specifically, since idiosyncratic investment risks can be mitigated via diversification, investors only require higher compensation for increased systematic risk. Markowitz's ideas inspired the development of the CAPM. The first iteration of the CAPM (Sharpe, 1964; Lintner, 1965; Mossin, 1966) held that in equilibrium only market risk should be priced by investors in an efficient market such that expected security excess returns ($E[R_i]$) are a linear function of their systematic risk exposure (β_i) and a market risk premium ($E[R_M]$):

[26] Refer to Frazzini et al. (2018). Warren Buffet's investment strategy is based on investing in low-beta assets and applying high levels of leverage. This strategy directly exploits the measurement bias of factor models and produces artificially inflated Alpha estimates.

$$E[R_i] = \beta_i E[R_M].$$

Jensen (1968) was the first to note that the Sharpe-Lintner CAPM relation between expected return and market Beta also implies a time-series regression test; prompting one of the earliest performance evaluation metrics to become widely used in empirical finance (Jensen's Alpha). Jensen's Alpha represents the risk-adjusted return of a security in excess of the expectations established by the CAPM. The Alpha estimate in a single factor model is obtained via time-series regression as follows:

(*Time-series* regression): $$R_{it} = \hat{\alpha}_i + \hat{\beta}_i R_{Mt} + \hat{\varepsilon}_{it}, \quad \textbf{(3.1)}$$

where R_{it} and R_{Mt} represent the excess return of security i and the market proxy M, respectively, at time t; $\hat{\beta}_i$ represents the exposure of security i to non-diversifiable risk; and $\hat{\alpha}_i$ represents the risk-adjusted return of security i in excess of expectations (Jensen's Alpha).

Securities which can achieve a greater return without a proportional increase in exposure to systematic risk will produce positive risk-adjusted returns ($\hat{\alpha}_i$), signalling their virtue as a superior risk-adjusted performer relative to other securities in the market. However, if markets are informationally efficient, then we should not observe individual securities or portfolios that are capable of consistently producing statistically significant positive Alphas (Fama, 1969).

Unfortunately, parsimonious modelling attempts often fail when applied to empirical data. Almost immediately, it was observed that the security market line is empirically "flatter than expected" (Black, Jensen and Scholes, 1972) by CAPM theory. In the decades that followed, academic research identified several portfolio characteristics that could produce statistically significant positive Alphas under time-series regressions of the CAPM. Most notably, these include the "size effect" anomaly (Banz, 1981), the propensity for smaller equities to generate positive Alphas, and the "value effect" anomaly (Fama & French, 1992), indicating positive Alphas from equities with high book-to-market ratios. This prompted an expansion of the CAPM to include these additional factors. At the turn of the century, the CAPM remained widely used in industry (Graham & Harvey, 2001). However, the FF3F model quickly became the staple of academic research.

Anomalies should be rare. However, in the past couple of decades the number of identified financial market anomalies has grown immensely with new anomalies being identified in the academic literature each year. Hou et al. (2020) identify 452 published return anomalies in the accounting and finance literature. This phenomenon has given rise to the colloquial term of a "factor zoo" (Cochrane, 2011). The FF3F Model has now been extended to a 5-factor model to incorporate some of the most prominent anomalies (Fama & French, 2015), whilst inclusion of a "momentum" factor in a model (Carhart, 1997) retains its popularity to date. Even with the use of these expanded factor models we continue to find new portfolio formations capable of generating positive Alpha (anomalies). The inability of existing models to fully account for investment strategies that can generate Alpha motivates the ongoing addition of new risk premia to factor models. It is within this context of expanding factor models, and concern with how commonplace anomaly identification has become, that we turn to a crucial issue with reliance on the Alphas estimated via time-series regression of factor models.

Conceptually, there are two common ways of estimating alphas: time-series regression (see Equation 3.1) and cross-sectional regression (see Equation 3.5). For theory and practice, the time-series approach has been more commonly used. Much of the existing anomaly literature in finance has arisen from papers which rely on use time-series regression to estimate Alpha. Prominent examples of anomalies identified through use of time-series regression include the momentum factor (Carhart, 1997), low price-to-earnings outperformance (Basu, 1977), and Fama-French 3-Factor model (Fama & French, 1993). Unfortunately, it is time-series regression with which this chapter finds fault.

The model parameters in a time-series regression of a factor model are *assumed* to be correctly specified in the cross-section. Using the CAPM as an example, it is *assumed* that our underlying data exhibits the cross-sectional characteristic that for each additional unit of $\hat{\beta}_i$, stock returns tend to increase by R_{Mt}. A *time-series* regression makes no attempt to verify that such *cross-sectional* assumptions are reflected by the underlying data; they are taken at face value. For example, under time-series regression, Jensen's alpha for each security is estimated as:

$$\hat{\alpha}_i = \bar{R}_i - \hat{\beta}_i * \bar{R}_M \quad \textbf{(3.2)}$$

where,

$$\bar{R}_i = \frac{1}{T}\sum_{t=1}^{T} R_{it}$$

$$\bar{R}_M = \frac{1}{T}\sum_{t=1}^{T} R_{Mt}$$

We note that as of 2022, this is the same approach to alpha estimation that CFA Institute teaches practitioners in their Level I certification, and a great importance is placed upon the measure "*Jensen's alpha is the maximum amount that you should be willing to pay the manager to manage your money*".[27] Due to model specification, for every one-unit increase in the $\hat{\beta}_i$ of a security, it's $\hat{\alpha}_i$ will be penalised by precisely $\bar{R}_M * \hat{\beta}_i$; i.e. the assumed cross-sectional risk-return relation in the market has a slope of $\bar{R}_M$. This cross-sectional assumption imposed by the time-series regression is not based on underlying data but is, instead, arbitrary.[28] The true relation between risk and return is almost certain to differ from this imposed cross-sectional assumption in practice. As data exhibits a risk-return relation which more greatly differs from the imposed cross-sectional assumption, Jensen's alpha will become progressively spurious. For example, if the security market line observed in empirical data is indeed *flatter than expected* (Black, Jensen and Scholes, 1972) the penalty imposed in estimation of Jensen's alpha would be excessively harsh on securities with $\hat{\beta}_i > 1$, and excessively lenient on securities with $\hat{\beta}_i < 1$ during bull markets. For illustrative purposes, we define two extreme scenarios:

Scenario 1: The CAPM fully describes underlying data

Scenario 2: No risk-return relation in underlying data

Figure 3.1 illustrates the two extreme scenarios.

---

[27] CFA Level 1 Certification. Corporate Finance and Portfolio Management, page 376.

[28] The chosen market proxy will differ significantly from the true unobservable market portfolio envisaged under CAPM theory which consists of all investable assets. The severity of the disconnect and how easily it can have a large impact on parameter estimation is highlighted by Roll and Ross (1994, p. 105) who demonstrate that a market proxy with a return only 22 basis points lower than the efficient frontier is sufficient to generate zero cross-sectional correlation between stock returns and the market proxy. Hence, the observed $\bar{R}_M$ in a time-series regression will induce a very different risk-return penalty than the true market portfolio in the CAPM would. This penalty is also arbitrary since it is defined by the average return of the chosen market proxy and does not adjust as the cross-sectional relation between observed asset returns and their measured Betas change.

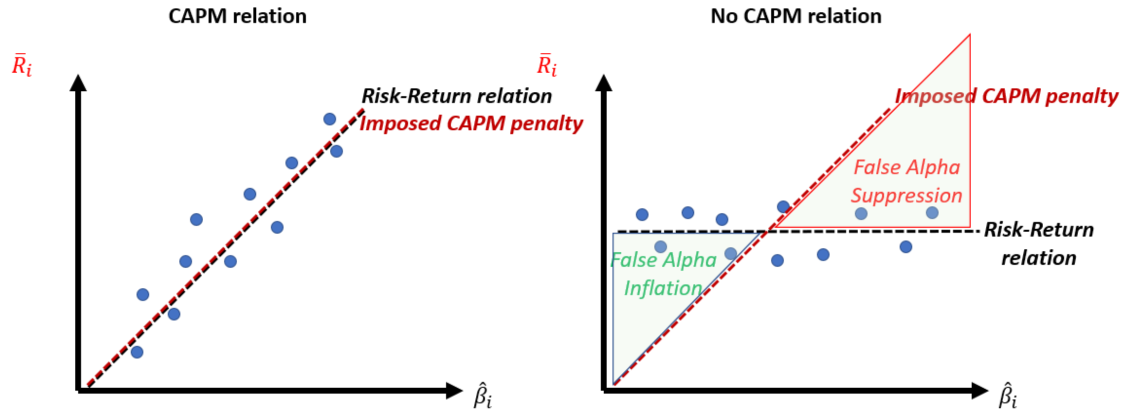


Figure 3.1: No risk premium misspecification vs. extreme risk premium misspecification

Figure 3.1 presents an illustrative contrast of Scenario 1 (left panel) in which the CAPM accurately describes an underlying data set against Scenario 2 (right panel) in which there is no relation between security returns and systematic risk. Each blue dot represents an observation. The black dotted line on each chart represents the fitted regression line from a cross-sectional regression ($\bar{R}_i = \hat{c} + \hat{\lambda}_M \hat{\beta}_{M,i}$). The dotted red line represents the penalty imposed on Jensen's alpha ($\hat{\beta}_i * \bar{R}_M$) as per Equation 3.2. When the CAPM accurately reflects the underlying data (Scenario 1), the cross-sectional risk-return relation is equivalent to the penalty imposed by the CAPM and estimates of Jensen's alpha are unbiased. By contrast, when there is no relation between security returns and market risk (Scenario 2), the imposed CAPM penalty is spurious. As a result, Jensen's alpha is progressively biased upwards (downwards) for lower (higher) beta securities or portfolios when $\bar{R}_M$ is positive, and the opposite is true when $\bar{R}_M$ is negative.

Under the extremity of Scenario 1, $E[\alpha_i] = 0 \; \forall_i$, resulting in:

$$R_{it} = \beta_i R_{Mt} + \varepsilon_{it} \,. \tag{3.3}$$

Under these conditions, it follows that the single-factor model (Equation 3.1) is correctly specified, and parameter estimates will be unbiased. Estimation of Jensen's Alpha would be valid since the risk-return penalty applied in Alpha estimation matches the true cross-sectional risk-return relation in the underlying data. Furthermore, there is no cross-sectional dependence between $\hat{\alpha}_i$ and $\hat{\beta}_i$ since $E[\alpha_i] = 0 \; \forall_i$. Conversely, if the CAPM had no bearing on the underlying data (Scenario 2), then $E[\beta_i] = 0 \; \forall_i$ resulting in:

$$R_{it} = \alpha_i + z_{it} \,. \tag{3.4}$$

Under this second scenario, the single-factor model (Equation 3.1) would be misspecified since it miss-represents the underlying data. The estimation of Jensen's Alpha based on Equation 3.2 applies a performance penalty ($\bar{R}_M * \hat{\beta}_i$) that is spurious since there is no cross-sectional relation between risk and return. So long as $\bar{R}_M \neq 0$, we would observe correlation between $\hat{\alpha}_i$ and $\hat{\beta}_i$ when using time-series regression since $\hat{\alpha}_i = \bar{R}_i - \bar{R}_M * \hat{\beta}_i$. This would give rise to a "low beta anomaly" in rising markets, since securities that exhibit $\hat{\beta}_i$ noticeably below 1 will tend to exhibit positively inflated $\hat{\alpha}_i$, and a "high beta anomaly" in declining markets; where $(-\bar{R}_M)$ becomes positive and the effect reverses.

Under Scenario 2, the values of $\hat{\alpha}_i$ from time-series regressions are impacted by an endogenous relation to $\hat{\beta}_i$ and thus imposition of an entirely spurious penalty ($\bar{R}_M * \hat{\beta}_i$). However, when analysing financial data, we are most likely to encounter an intermediate situation in between the illustrative extremes of Scenario 1 and Scenario 2. In empirical financial market data, we are likely to observe some relation between market risk and security return but not a fit that perfectly conforms to the CAPM. Consequently, one part of $\hat{\alpha}_i$ from time-series regression will be spurious and driven by structural endogeneity from misspecification of cross-sectional risk premia rather than by genuinely superior risk-adjusted performance. Another, exogenous, part of time-series $\hat{\alpha}_i$ will be driven by genuine risk-adjusted outperformance. Figure 3.2 presents a visualisation of a moderate and more realistic scenario.

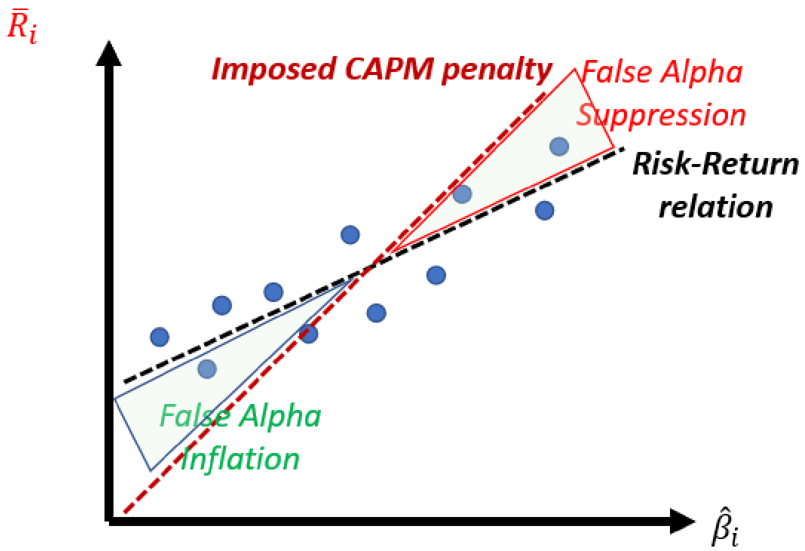


Figure 3.2: Partial risk premium misspecification

Figure 3.2 presents a visual representation of a situation in which the CAPM does not accurately reflect the underlying data but there is still some (weaker) relation between security returns and market risk. Each blue dot represents an observation. The black dotted line on each chart represents the fitted regression line from a cross-sectional regression ($\bar{R}_i = \hat{c} + \hat{\lambda}_M \hat{\beta}_{M,i}$). The dotted red line represents the penalty imposed upon Jensen's Alpha ($\hat{\beta}_i * \bar{R}_M$) as per Equation 3.2. Since the relation between security returns and market risk in the underlying data is weaker than the relation assumed by a single factor model, the imposed CAPM penalty is partially spurious and partially valid. Due to the spurious component of the penalty, Jensen's Alpha will be progressively biased upwards (downwards) for lower (higher) Beta securities or portfolios when $\bar{R}_M$ is positive, and the opposite would be true when $\bar{R}_M$ is negative.

Whilst the cause was not identified, the consequences from use of misspecified time-series models has long been apparent in finance literature. From as early as 1970 (see Friend and Blume, 1970) there has been an awareness that Jensen's Alpha, as estimated via time-series regression, tends to be negatively correlated with Beta in empirical data. More recently, Frazzini and Pedersen (2014) use time-series regression and observe that the estimated Alphas for portfolios that overweight low-beta securities produce economically and statistically significant Alphas. Their remarkably strong results are robust across international markets and assets classes. Perhaps unknowingly, they have constructed a portfolio that directly exploits the measurement error of Alpha in time-series regressions due to model misspecification of risk premia.

Fortunately, it is straightforward to estimate Alpha in a manner that accounts for the cross-sectional risk-return relation of assets within a market. Under a cross-sectional regression approach, alpha estimates are the residuals from a regression of average excess returns against factor portfolio loadings.[29] Equation 3.5 below is an example of a cross-sectional regression for a single-factor model:

$$\text{(\textit{Cross-sectional} regression):} \qquad \bar{R}_i = \hat{\lambda}_0 + \hat{\lambda}_1 \hat{\beta}_i + \hat{\alpha}_i, \qquad \textbf{(3.5)}$$

where $\bar{R}_i$ represents the average return of asset i; $\hat{\lambda}_1$ indicates the expected increase in $\bar{R}_i$ for each additional unit of market risk exposure ($\hat{\beta}_i$); $\hat{\lambda}_0$ represents the expected average return for a zero-beta asset; and $\hat{\alpha}_i$ represents the risk-adjusted outperformance/underperformance of asset i. If a model, such as the CAPM, accurately explained average security returns then we would observe $\hat{\lambda}_0 = 0, \hat{\lambda}_1 = \bar{R}_M$. Under these conditions, the Alpha estimates of a time-series regression (Equation 1) and cross-sectional regression (Equation 3) would be identical.

Tests on intercepts of cross-sectional regressions are routinely performed to observe whether the cross-section of stock returns can be, on average, explained by a model's regressors (e.g. Blume & Friend, 1973; Fama & French, 1992). Since cross-sectional regressions adjust to the actual cross-sectional risk-return relation in data being

[29] Alternatively, a Fama-MacBeth approach involves cross-sectional regressions at each time period. Alpha estimates are then the equal-weighted average residual of a security or portfolio across each cross-sectional regression. Under linear regression, the average slope coefficient is equivalent to the slope coefficient of averages. The alpha estimates produced by a cross-sectional regression and Fama-MacBeth regression are the same.

analysed, anomaly identification should become scarcer under the cross-sectional regression setting. Motivated by a belief of widespread "p-hacking"[30] within finance, Hou et al. (2020) have recently provided an extensive replication of 452 published return anomalies using a Fama MacBeth (1973) cross-sectional approach. They find that after controlling for microcap stocks, 65% of their replicated anomalies are statistically insignificant at the 5% significance level.

Factor models, such as those popularised by Eugene Fama and Kenneth French, are likely to have misspecified risk premia in an empirical setting, inducing structural endogeneity. This occurs because, to date, neither the CAPM, nor subsequent factor models, can fully account for all determinants of empirical stock return data. As a result, $\hat{\alpha}_i$, typically the important metric for evaluating performance, becomes partially dependent upon the estimated slope coefficients of each factor whenever a model is misspecified. For academics and practitioners alike, it is important to isolate the proportion of $\hat{\alpha}_i$ that is exogenous from $\hat{\beta}_i$, and a genuine representation of superior risk-adjusted performance, from the proportion of $\hat{\alpha}_i$ that is spurious and due to dependence on $\hat{\beta}_i$ from model misspecification. We demonstrate how to achieve this in the following section and extend our approach to multi-factor models.

3.3 The decomposition of time-series Alpha

We define the proportion of Alpha estimates from time-series regression (henceforth denoted as *time-series Alpha;* $\hat{\alpha}_i^{TS}$) that is due to model misspecification as *Endogenous Alpha* ($\hat{\alpha}_i^{End}$). The remaining proportion of time-series Alpha estimates, which exceed the expected level of time-series Alpha due to model misspecification, we denote as *Excess Alpha* ($\hat{\alpha}_i^{Exc}$). By construction, Excess Alphas are exogenous to Beta estimates. For simplicity, we initially focus on decomposition of the single-factor model. However, we subsequently detail how the decomposition approach can be extended to multi-factor models. Equation 3.1 becomes:

[30] The misuse of data analysis to find patterns in data that can be presented as statistically significant when in fact there is no real underlying effect.

$$R_{it} = \hat{\alpha}_i^{Exc} + \hat{\alpha}_i^{End} + \hat{\beta}_i R_{Mt} + \hat{\varepsilon}_{it}, \quad \textbf{(3.6)}$$

where $\hat{\alpha}_i^{TS} = \hat{\alpha}_i^{Exc} + \hat{\alpha}_i^{End}$. In Equation 3.6, we split the time-series Alpha, into the component which is endogenous, and the component which is exogenous and due to genuine risk-adjusted outperformance. If the CAPM held true in the underlying data, then the model would be correctly specified so $\hat{\alpha}_i^{End}$ would equal zero and, therefore, $\hat{\alpha}_i^{TS}$ and $\hat{\alpha}_i^{Exc}$ would be identical. Endogenous Alpha estimates are obtained by performing a cross-sectional regression of time-series alpha estimates against $\hat{\beta}_i$:

$$\hat{\alpha}_i^{TS} = \hat{\phi}_0 + \hat{\phi}_1 \hat{\beta}_i + \hat{u}_i, \quad \textbf{(3.7)}$$

where $\hat{\alpha}_i^{End} = \hat{\phi}_0 + \hat{\phi}_1 \hat{\beta}_i$ (i.e. the fitted values of the regression), and $\hat{\alpha}_i^{Exc} = \hat{u}_i$ (i.e. the proportion of traditional time-series Alpha estimates that are exogenous to exposure to market risk). This cross-sectional regression of time-series Alpha estimates on slope coefficients allows us to discern their cross-sectional dependence on $\hat{\beta}_i$. Endogenous Alpha estimates are the predicted values of this cross-sectional regression. Since $\hat{\alpha}_i^{End}$ capture any dependence of time-series Alpha estimates upon $\hat{\beta}_i$, subtraction of $\hat{\alpha}_i^{End}$ from time-series Alphas effectively isolates the component of $\hat{\alpha}_i^{TS}$ that is in excess of cross-sectional, risk-adjusted expectations (i.e. we identify Excess Alpha). As a result, $\hat{\alpha}_i^{Exc}$ has some highly desirable properties as a performance evaluation metric. Excess Alpha estimates are orthogonal to $\hat{\beta}_i$; portfolios can't simply outperform by loading up on low Beta assets. In addition, average $\hat{\alpha}_i^{Exc}$ is guaranteed to be zero for any data set; it is impossible for the whole market to outperform or underperform – there must be an offsetting amount of over and under-achievers in each evaluation period. This property occurs due to $\hat{\alpha}_i^{End}$ being dependent on $\hat{\phi}_0$; such that $\hat{\alpha}_i^{Exc}$ is effectively de-meaned. We can restate Equation 3.6 with $\hat{\alpha}_i^{End}$ decomposed:

$$R_{it} = \hat{\alpha}_i^{Exc} + \hat{\phi}_0 + \hat{\phi}_1 \hat{\beta}_i + \hat{\beta}_i R_{Mt} + \hat{\varepsilon}_{it}. \quad \textbf{(3.8)}$$

Each $\hat{\phi}$ captures an element of factor model misspecification due to the failure of time-series regressions to adjust for the actual cross-sectional relation between returns and risk premia in the data being analysed. For example, $\hat{\phi}_0$ captures the average level of security excess returns that are independent of risk factors whilst $\hat{\phi}_1$ quantifies the bias in the estimate of the market risk premium. Consider Scenario 1 from the preceding section. When the CAPM accurately describes the data we are analysing, the single-factor model is not cross-sectionally misspecified so $E\left[\hat{\alpha}_i^{TS}\right] = 0$ and $E\left[\hat{\alpha}_i^{TS}|\hat{\beta}_i\right] = 0$. As a result, $\hat{\phi}_0 = 0$ and $\hat{\phi}_1 = 0$. In other words, Excess Alpha converges to traditional time-series Alpha in the absence of model misspecification. In empirical settings, before we analyse the data before us, we do not know whether the data will conform to the model specified or whether it will differ. Excess Alpha only deviates from $\hat{\alpha}_i^{TS}$ in the presence of model misspecification. Therefore, we can be comfortable using Excess Alpha for performance evaluation in place of $\hat{\alpha}_i^{TS}$ across all financial datasets. We can further compare the properties of $\hat{\alpha}_i^{TS}$ versus $\hat{\alpha}_i^{Exc}$ via simulations of Scenario 1 and Scenario 2.

Starting with the simulation of Scenario 1, where the CAPM accurately describes the average risk-return relation in the underlying data and absent model misspecification. We rely on the same approach as in Jegadeesh et al. (2019, p. 277) but use different simulation values. We centre our distribution of β_i on 1 instead of 0.95 since we are simulating a scenario reflective of the CAPM.[31] We use an average market excess return (μ_M) of 19.2%. This value was obtained from the data library of Kenneth French and is representative of the average annual excess return of a capitalisation-weighted portfolio of US equities based exclusively on positive return markets over an extensive data period (1927 to 2020).[32] The simulated distributions of security residuals and the standard deviation of market excess returns are otherwise based on the values observed in our 1991-2020 data sample discussed in Section 3.4 of this chapter. For our simulation, we generate 2,000 sets of Beta values and residuals return variances as follows:

$$\Delta T = \frac{1}{T} = \frac{1}{252}, \;\; \beta_i \sim N(1, 0.42^2), \;\; \sigma_{\varepsilon,i}^2 \sim N(0.5873, 0.2381^2).$$

---

[31] Under CAPM theory, the average security Beta in a market is 1. Our first simulation aims to closely reflect a scenario in which the CAPM is capable of explaining security returns. In addition, having Betas not centred on 1 will introduce an expectation of non-zero average Alphas into a time-series regression.

[32] Jegadeesh et al. (2019) base their simulation on a value of average market excess return across all years. Alpha estimation bias is of opposing directions in positive return markets ("low beta anomaly") versus negative return markets ("high beta anomaly"). Therefore, taking a simple average across all market conditions would significantly understate the typical Alpha estimation bias by netting out opposing errors. For μ_M, we need to look exclusively at average market excess returns during rising markets *or* declining markets. Since this chapter seeks to provide an accurate indication of time-series Alpha estimation bias over long time periods we obtain data on the excess return of a market portfolio from 1927 to 2020 from the data library of Kenneth French (https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). The excess return of the market portfolio is calculated as described in Fama and French (1993).

The parameter values are kept constant across simulations. We generate daily market excess returns using a random walk with drift:

$$R_{Mt} = \mu_M + \varepsilon_{Mt},$$

where $\mu_M = 0.192\Delta T$ and $\varepsilon_{Mt} \sim N(0, 0.1533^2)$. We use the generated Betas and idiosyncratic volatilities to construct 2,000 daily excess return series based on the CAPM (i.e. a simulated duplicate of Equation 3.3):

$$R_{it} = \beta_i R_{Mt} + \varepsilon_{it},$$

where $\varepsilon_{it} \sim N(0, \Delta T * \sigma_{\varepsilon,i}^2)$. As a result, the CAPM holds for the underlying data by design. Figure 3.3 below plots the average annual return of each security against their estimated market Beta as well as a blue regression line which shows the linear relation between risk and return. A red, dashed line is overlayed on the chart indicating the risk-return relation assumed by the single-factor model. Since the CAPM is true in this simulation, the two lines are one in the same and the model is specified correctly in the cross-section. As a result, we expect $\hat{\alpha}_i^{TS}$ and $\hat{\alpha}_i^{Exc}$ to be the same since no adjustment for model misspecification is required.

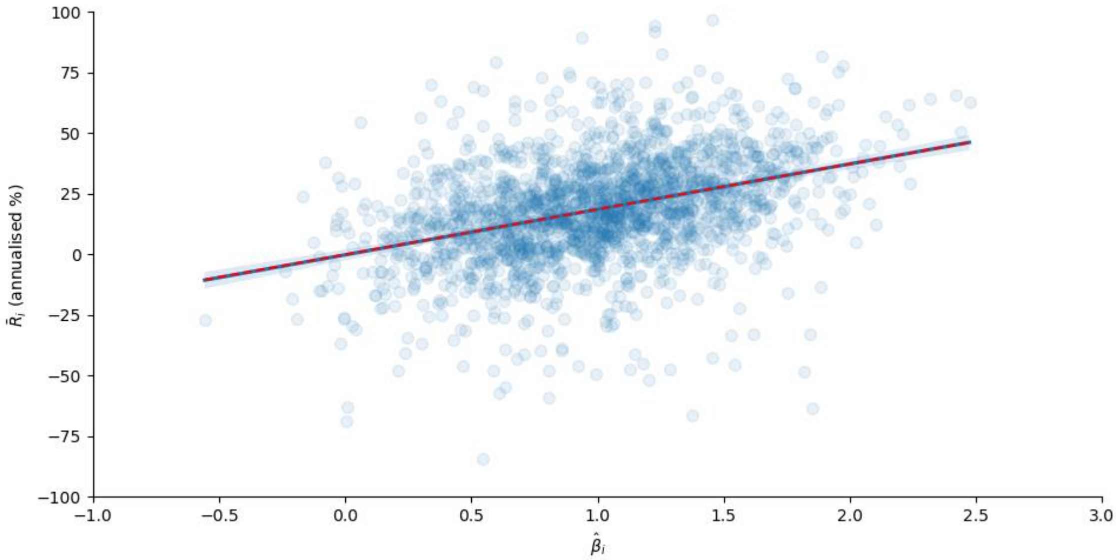


Figure 3.3: Factor model is correctly specified

Figure 3.3 plots the average annual return of each security against their estimated market Beta. A blue linear regression line is plotted indicating the risk-return relation in the simulated data. A dashed red line is overlayed indicating the risk-return relation imposed by the single-factor model. In our simulation, the two lines are almost identical, providing a visual indication that the model is correctly specified and traditional Alpha estimates will not be biased by structural endogeneity.

Figure 3.4 below compares Alpha estimates from a time-series regression with $\hat{\alpha}_i^{Exc}$. Since the cross-sectional risk-return relation in the simulated data matched the imposed risk-return assumptions of the single-factor model, Alphas estimated via time-series regression and $\hat{\alpha}_i^{Exc}$ are equivalent.

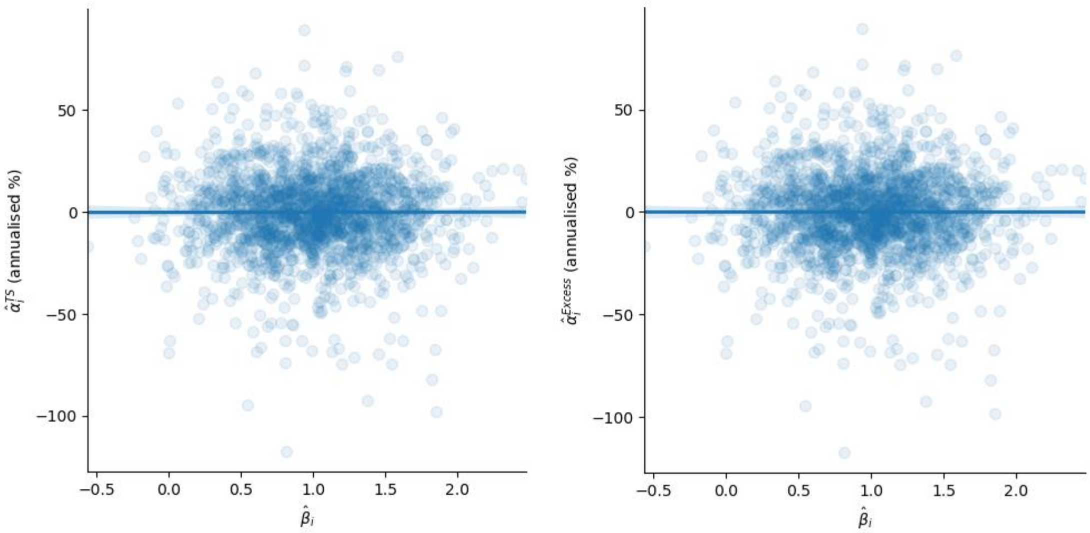


Figure 3.4: Equivalence of $\hat{\alpha}_i^{TS}$ and $\hat{\alpha}_i^{Exc}$ in Scenario 1

Figure 3.4 plots the time-series (left panel) and Excess Alpha (right panel) estimates of each security against their estimated market Betas. A blue linear regression line is plotted indicating the relation between the respective alpha estimates and Betas in the simulated data. Both Alpha estimation approaches were equivalent for the simulated data and were independent of $\hat{\beta}_i$.

The annualised $\hat{\phi}$ values from our regression of alpha estimates against slope coefficients (Equation 3.8) are close to zero ($\hat{\phi}_0 = 0.11\%, \hat{\phi}_1 = 0.07\%$) and statistically insignificant ($\hat{\phi}_{0,p-value} = 0.92, \hat{\phi}_{1,p-value} = 0.95$, F-test p-value = 0.95).[33] The statistical insignificance of $\hat{\phi}$ values provides a quantitative verification that the single-factor model was appropriately specified under our simulation of Scenario 1, as expected. Having demonstrated that $\hat{\alpha}_i^{TS}$and $\hat{\alpha}_i^{Exc}$ are equivalent in the absence of model misspecification, we move onto simulation of the opposing extremity.

In Scenario 2, security returns are random and there is no modelled relation between risk and return. For this second simulation we generate 2,000 security returns using a random walk with drift:

$$R_{it} = \mu + \varepsilon_{it}$$

[33] The values of each $\hat{\phi}$ are not exactly zero due to each simulated security return being subject to noise (ε_{it}) rather than precisely following the CAPM.

where $\mu = 0.192\Delta T$ and ε_{it} are identical to (copied from) Simulation 1. We then form an equal weighted portfolio ($R_{M,t} = \frac{1}{N}\sum_{i=1}^{N} R_{i,t}$) which serves as the market portfolio in our simulation and estimate time-series Alphas and Excess Alphas for a single-factor model. Figure 3.5 plots the average daily return of each security against their estimated market Beta and shows the risk-return relation present in the data (blue regression line) vs the risk-return relation imposed by the single-factor model (red dashed line). Since the risk-return relation for our simulated data differs greatly from the model assumptions, the single-factor model is misspecified.

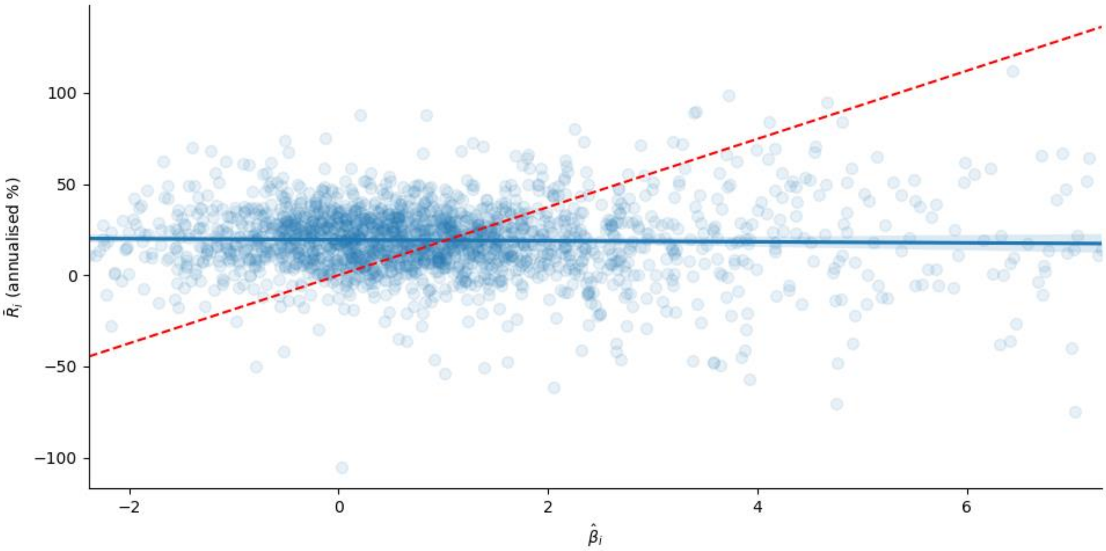


Figure 3.5: Factor model is misspecified

Figure 3.5 plots the average daily return of each security against their estimated market Beta. A blue linear regression line is plotted indicating the risk-return relation in the simulated data. A dashed red line is overlayed indicating the risk-return relation imposed by the single-factor model. In our simulation, the two lines differ greatly, providing a visual indication that the model is misspecified and traditional alpha estimates will be biased by structural endogeneity.

Figure 3.6 compares $\hat{\alpha}_i^{TS}$ and $\hat{\alpha}_i^{Exc}$. As a result of model misspecification, we observe very different values for $\hat{\alpha}_i^{TS}$ and $\hat{\alpha}_i^{Exc}$. Excess Alphas are penalised based on the risk-return relation observed in the data whilst traditional time-series Alphas are penalised based on a presumed risk-return relation stipulated by the CAPM.

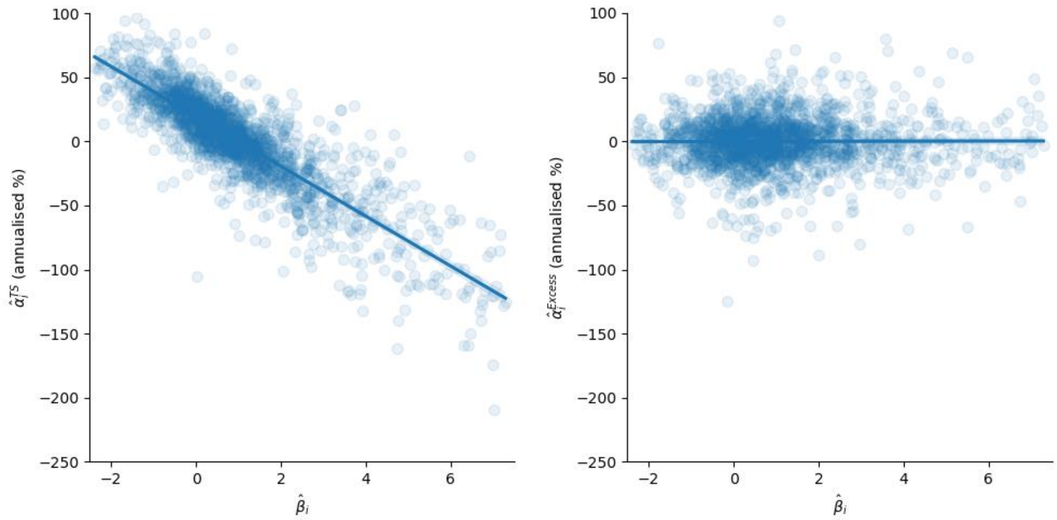


Figure 3.6: Divergence of $\hat{\alpha}_i^{TS}$ and $\hat{\alpha}_i^{Exc}$ in Scenario 2

Figure 3.6 plots the time-series (left panel) and Excess Alpha (right panel) estimates of each security against their estimated market Beta. A blue linear regression line is plotted indicating the relation between the respective alpha estimates (annualised %) and Betas in the simulated data. Since the factor model was misspecified, Alphas estimated via time-series regression are significantly biased and differ greatly from $\hat{\alpha}_i^{Exc}$.

The annualised $\hat{\phi}$ values from our alpha-beta regression (Equation 3.8) are economically large ($\hat{\phi}_0 = 19.48\%, \hat{\phi}_1 = -19.41\%$) and extremely statistically significant ($\hat{\phi}_{0,p-value} = 0.00$, $\hat{\phi}_{1,p-value} = 0.00$, F-test p-value = 0.00). Since we had not simulated any cross-sectional relation between market risk and security returns, we expect no market risk premium in our Scenario 2 results. The single-factor model had imposed an annual market risk premium of greater than 19% whilst the actual risk-return relation in the simulated data exhibited a statistically insignificant market risk premium (-0.07%). In this extreme scenario, where we simulated no relation between risk and return, the estimated time-series alpha of a zero-beta security would, on average, be *overstated* by approximately 19.5% ($\hat{\phi}_0$).

Traditional Alpha estimates systematically vary with $\hat{\beta}_i$ causing the $\hat{\alpha}_i^{TS}$ of low (high) $\hat{\beta}_i$ securities to be overstated (understated). The impact upon hypothesis testing can be observed in Figure 3.7. Since statistical tests of significance tend to benchmark against a constant expectation that $E\left[\hat{\alpha}_i^{TS} \mid \hat{\beta}_i\right] = 0 \; \forall_i$, securities with more extreme $\hat{\beta}_i$ will be identified as statistically significant risk-adjusted performers simply because they are more substantially impacted by the spurious $\hat{\alpha}_i^{TS}$ penalty of model misspecification. The left-hand chart in Figure 3.7 demonstrates this effect for $\hat{\alpha}_i^{TS}$. By contrast, on the right-hand chart we observe that, since $\hat{\alpha}_i^{Exc}$ is orthogonal to $\hat{\beta}_i$ by design, there is no inflation of false positives.

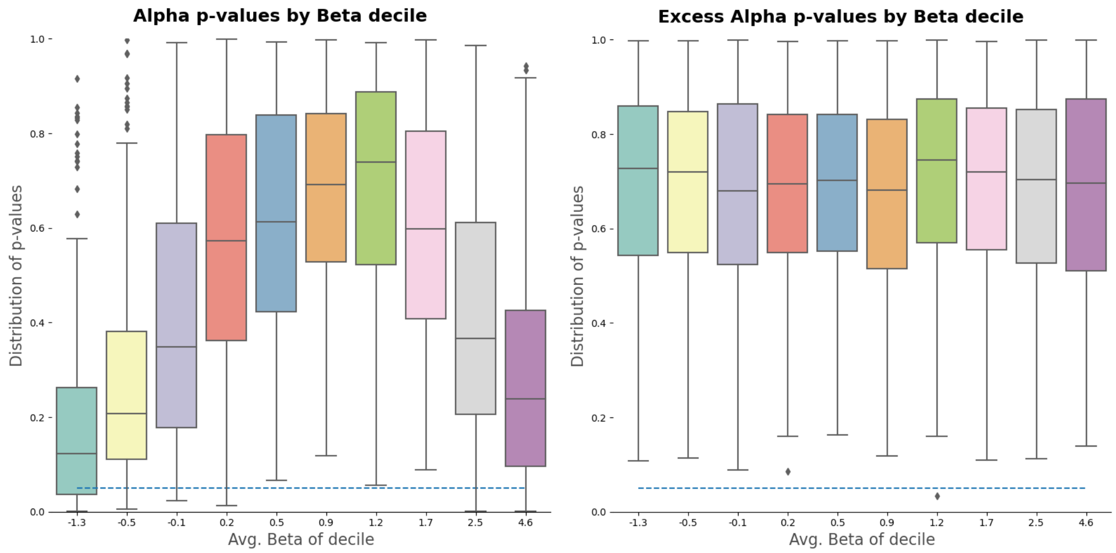


Figure 3.7: Distribution of p-values for $\widehat{\alpha}_i$ and $\widehat{\alpha}_i^{Exc}$ across Beta deciles

Figure 3.7 plots the distribution of p-values for time-series (LHS) and Excess Alpha (RHS) estimates of each security against their estimated market Betas. Due to model misspecification, securities with more extreme $\hat{\beta}_i$ will be falsely allocated statistical significance under a traditional Alpha measurement approach. By contrast, Excess Alpha estimates are orthogonal to $\hat{\beta}_i$ and do not suffer from the same false positive issue. The blue dotted lines in each chart indicate a 5% significance level.

A scatter plot of Alpha estimates vs terminal wealth for Scenario 2 is shown in Figure 3.8 below.

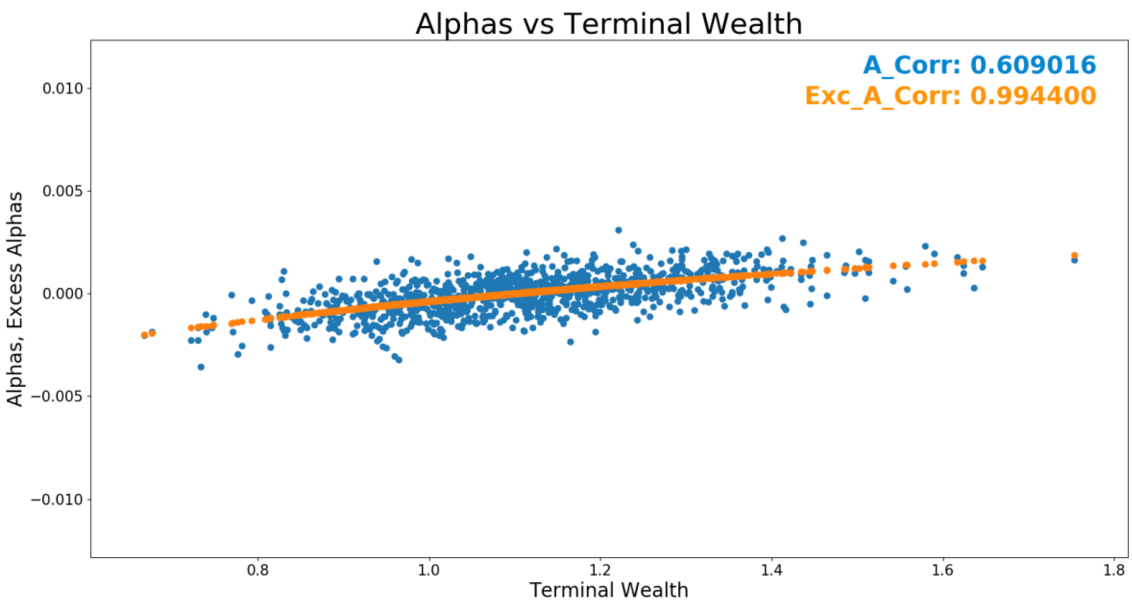


Figure 3.8: Correlation of terminal wealth with $\widehat{\alpha}_i$ and $\widehat{\alpha}_i^{Exc}$

Figure 3.8 is a scatter plot of terminal wealth (horizontal axis) vs estimated daily Alpha (blue dots) or Excess Alpha (orange dots) (vertical axis) for the securities modelled in Scenario 2. The correlations between terminal wealth and the respective Alpha estimation type are displayed in the top right corner. Traditional Alpha estimates exhibit noticeably lower correlation with terminal wealth.

Traditional Beta estimates are heavily impacted by noise, making traditional Alpha estimates partially spurious. This reduces the correlation between traditional Alpha estimates and terminal wealth for individual securities. By contrast, Excess Alpha estimates are nearly perfectly correlated with terminal wealth. This is because they are estimated based on the actual cross-sectional risk premia present in examined data, rendering them highly resistant to noise.

The preceding simulations considered the extent of estimation bias for a single-factor model under time-series regression. However, the bias in Alpha estimates can be similarly decomposed for multi-factor models. For example, in a FF3F Model we would have:

$$R_{it} = \hat{\alpha}_i^{Exc} + \hat{\phi}_0 + (\hat{\phi}_1 + R_{M,t})\hat{\beta}_{M,i} + (\hat{\phi}_2 + R_{SMB,t})\hat{\beta}_{SMB,i} + (\hat{\phi}_3 + R_{HML,t})\hat{\beta}_{HML,i} + \hat{\varepsilon}_{it}, \quad \textbf{(3.9)}$$

where $\hat{\phi}$ are obtained from an expanded version of the alpha-beta regression from Equation 3.8. Each $\hat{\phi}$ measures the distortion of a different factor model parameter under time-series regression due to imposition of an arbitrary cross-sectional assumption that does not reflect the underlying data. The $\hat{\phi}$ represent the adjustment required to respective parameters in order to introduce the cross-sectional risk-return relation of the underlying data. Since Equation 3.9 has embedded the cross-sectional assumptions of the underlying data, it is directly comparable to a cross-sectional model:

$$\bar{R}_i = \hat{c} + \hat{\lambda}_M \hat{\beta}_{M,i} + \hat{\lambda}_{SMB} \hat{\beta}_{SMB,i} + \hat{\lambda}_{HML} \hat{\beta}_{HML,i} + \hat{\alpha}_i^{Exc}, \quad \textbf{(3.10)}$$

where $c = \hat{\phi}_0$, $\hat{\lambda}_M \cong (\hat{\phi}_1 + \bar{R}_{M,t})$, $\hat{\lambda}_{SMB} \cong (\hat{\phi}_2 + \bar{R}_{SMB,t})$, and $\hat{\lambda}_{HML} \cong (\hat{\phi}_3 + \bar{R}_{HML,t})$. Which means to quickly obtain $\hat{\alpha}_i^{Exc}$ we could simply run a cross-sectional regression and save the residuals. In the following section we analyse the difference between $\hat{\alpha}_i$ and $\hat{\alpha}_i^{Exc}$ for the past 30 years of US equities data when using single-factor and FF3F models. We also examine the $\hat{\phi}$ of each factor to gauge the historical magnitude of risk-premium misspecification.

3.4 Quantifying historical factor model bias

We have identified that time-series Alpha estimates of factor models are likely to be biased in an empirical setting. The magnitude of bias, however, may be economically or statistically insignificant. If the bias is trivial, we could continue to use time-series regression for performance evaluation. By contrast, a persistent and large bias would represent a flaw with a common implementation of factor models and a pathway through which abundant anomalies could have been misidentified. In this section we inspect the historical bias of Alpha estimates produced by time-series regression of the CAPM and FF3F models and find evidence of an economically and statistically significant bias.

3.4.1 Data and method

Our analysis relies on daily US equities holding period return data. The data is obtained from the Center for Research in Security Prices (CRSP) and covers the sample period from January 1, 1991 to December 31, 2020. Only ordinary common shares listed on the NYSE, NASDAQ, and AMEX are included. Following Fama and French (1992), we exclude financial firms since they are expected to have structurally high levels of leverage which could distort construction of value portfolios. In the calculation of holding period returns we assume all dividends received are reinvested in the underlying stock on the date that the dividends are paid out. Book value data are obtained from COMPUSTAT and merged with the CRSP data using PERMNOs in constructing a value factor. Book values are filled forward in-between reporting periods such that on each portfolio formation date the most recently available data is used.

To gauge the extent of risk-premium misspecification under time-series regression we estimate the $\hat{\phi}$ for the CAPM (Equation 3.8) and FF3F model (Equation 3.9). In each model, the "market factor" is the excess holding-period return of a market-capitalisation-weighted index formed on the first trading day of January and held one year. The daily risk-free rate used is the monthly rate obtained from Kenneth French's website and converted into a daily format. We follow Fama and French (1993) in constructing SMB and HML portfolios. Each of these factor portfolios are formed on the first trading day of January and held for one year. Hou et al. (2020) provide a strong motivation to omit microcap equities from inclusion in factor portfolios. Since Fama and French (1993) do not omit microcaps we have not the adopted this procedure in the section below. This allows our results to be more comparable to historical research papers.

3.4.2 Results

Table 3.1 shows the annual market portfolio excess return and the estimated misspecification of each factor across our 30-year sample.

Year	$\bar{R}_M$	CAPM $\hat{\phi}_0$	CAPM $\hat{\phi}_1$	CAPM F_{p-val}	FF3F $\hat{\phi}_0$	FF3F $\hat{\phi}_1$	FF3F $\hat{\phi}_2$	FF3F $\hat{\phi}_3$	FF3F F_{p-val}
1991	26.81	10.95***	-0.35	0.8516	12.86***	-8.34***	-4.05**	-4.47***	0.0000
1992	5.54	11.08***	-5.48***	0.0000	7.33***	-2.47	-4.28***	-9.20***	0.0000
1993	9.17	9.82***	-3.35**	0.0130	2.98*	-1.07	1.49	-11.11***	0.0000
1994	-0.07	-7.09***	4.63***	0.0017	-2.95	2.92*	1.16	-7.33***	0.0000
1995	27.61	14.28***	-18.53***	0.0000	21.33***	-24.85***	3.85**	-4.47***	0.0000
1996	16.93	11.7***	-13.82***	0.0000	11.17***	-8.22***	-4.60***	-3.31***	0.0000
1997	24.06	15.96***	-19.92***	0.0000	10.56***	-5.88**	-7.34***	-7.50***	0.0000
1998	25.61	-13.26***	-5.62**	0.0234	18.01***	-16.84***	2.34	-8.19***	0.0000
1999	23.37	-24.47***	44.95***	0.0000	-32.84***	38.08***	-10.66***	-11.71***	0.0000
2000	-15.14	15.59***	-11.98***	0.0000	-19.57***	24.31***	-16.46***	-9.36***	0.0000
2001	-12.58	20.09***	3.03**	0.0292	12.98***	8.79***	-19.78***	-20.84***	0.0000
2002	-24.40	18.39***	-12.59***	0.0000	10***	8.68***	-11.5***	-8.30***	0.0000
2003	27.61	10.77***	2.67*	0.0969	12.53***	-9.14***	2.96**	-3.88***	0.0000
2004	10.69	27.41***	-18.76***	0.0000	14.15***	-9.03***	-4.76***	-0.65	0.0000
2005	4.73	-3.67	3.91**	0.0267	-1.90	4.12*	-0.09	6.66***	0.0000
2006	12.83	7.71***	-8.42***	0.0000	6.51***	-6.66***	1.07	-9.40***	0.0000
2007	10.56	3.00	-10.57***	0.0004	-7.11**	14.57***	-15.36***	3.83***	0.0000
2008	-34.06	4.39	-3.59	0.2542	8.18**	-2.27	-2.43	-4.97***	0.0000
2009	29.43	-0.44	5.61***	0.0014	14.08***	-8.61***	-3.65**	-2.86**	0.0010
2010	17.95	3.79	0.50	0.7823	2.22	0.34	-2.92**	-0.44	0.1643
2011	6.93	30.81***	-28.78***	0.0000	30.06***	-25.35***	-6.11***	-5.85***	0.0000
2012	14.59	13.5***	-12.12***	0.0000	8.72***	-6.08***	-4.41***	-9.29***	0.0000
2013	29.30	5.75**	-4.76**	0.0402	4.13	-3.80	0.06	-9.96***	0.0000
2014	11.87	17.61***	-20.65***	0.0000	11.91***	-5.96***	-8.19***	-9.48***	0.0000
2015	1.79	20.3***	-26.05***	0.0000	-1.90	6.89**	-6.51***	-5.55***	0.0000
2016	11.93	11.39***	-6.62***	0.0001	15.97***	-8.87***	-10.69***	-2.91***	0.0000
2017	20.24	5.57**	-11.03***	0.0000	10.98***	-9.04***	-2.34*	-1.13	0.0000
2018	-2.94	-9.42***	3.50	0.1929	4.10	-1.46	-4.04**	-8.28***	0.0000
2019	27.13	24.89***	-27.65***	0.0000	16.85***	-12.85***	-3.76***	-10.47***	0.0000
2020	29.76	17.46***	-12.83***	0.0001	-1.33	6.80*	2.66**	5.74***	0.0000
Average Values									
	11.24	**9.13**	**-7.16**		**6.67**	**-2.04**	**-4.61**	**-5.82**	
Proportion of years significant at									
10% level		83.33%	86.67%	86.67%	76.67%	80.00%	76.67%	90.00%	96.67%
5% level		83.33%	83.33%	83.33%	73.33%	70.00%	73.33%	90.00%	96.67%
1% level		76.67%	66.67%	66.67%	66.67%	63.33%	50.00%	86.67%	96.67%

Table 3.1: Risk premium misspecification under time-series regression (1991-2020)
Table 3.1 shows estimated cross-sectional misspecification ($\hat{\phi}_i$) in each risk premia when estimating Alpha under CAPM ($R_{it} = \hat{\alpha}_i^{Exc} + \hat{\phi}_0 + (\hat{\phi}_1 + R_{M,t})\hat{\beta}_{M,i} + \hat{\varepsilon}_{it}$) and FF3F model ($R_{it} = \hat{\alpha}_i^{Exc} + \hat{\phi}_0 + (\hat{\phi}_1 + R_{M,t})\hat{\beta}_{M,i} + (\hat{\phi}_2 + R_{SMB,t})\hat{\beta}_{SMB,i} + (\hat{\phi}_3 + R_{HML,t})\hat{\beta}_{HML,i} + \hat{\varepsilon}_{it}$) time-series regressions. Each $\hat{\phi}_i$ is expressed in terms of an annual absolute % difference between the cross-sectional risk premiums assumed by the respective model risk premia and the cross-sectional risk premiums estimated based on the underlying data. A positive $\hat{\phi}_i$ value indicates that the risk premia was understated in the original factor model specification, whilst a negative phi value indicates that a risk premia was overstated. *, **, *** represent statistical significance at 90%, 95%, 99% confidence levels, respectively. Risk premia were statistically significantly misspecified in the majority of years for every risk premia. These misspecifications were of a substantial economic magnitude.

In both models each of the risk premia are observed to be individually and jointly statistically significantly misspecified in most years. These misspecifications are economically large with the potential to materially bias the Alpha estimates of individual securities and portfolios. For example, in the single-factor model, the average value of $\hat{\phi}_0$ was approximately 9.1% per annum. This indicates that a zero-beta security would, on average, earn a rate of return well above (+9.1%) the risk-free rate over the 30-year sample period. In the FF3F model, the average value of $\hat{\phi}_0$ was also economically large at 6.7% per annum. By contrast, each of the remaining $\hat{\phi}$ were negative on average. This indicates that the slopes of each risk-premium are flatter in our data than each factor model originally specified; there is a weaker relation between security returns and these presumed risk proxies. Taken together, the findings indicate that the true risk-return relation in our data sample is likely closer to the relation envisioned by Black et al. (1972); a zero-Beta asset earns a return in excess of the risk-free rate, rather than Sharpe (1964) and Lintner (1965); a zero-beta security earns the risk-free rate. The net result is that if we use the Alpha estimates from time-series regression to evaluate performance, we will produce excessively inflated Alphas for securities with low Betas and excessively understated Alphas for securities with high Betas. We would observe a strong "low-beta anomaly" in positive return markets purely due to model misspecification. We further compare the Alpha estimates from time-series regression and cross-section regression for our sample period in Figure 3.9 below.

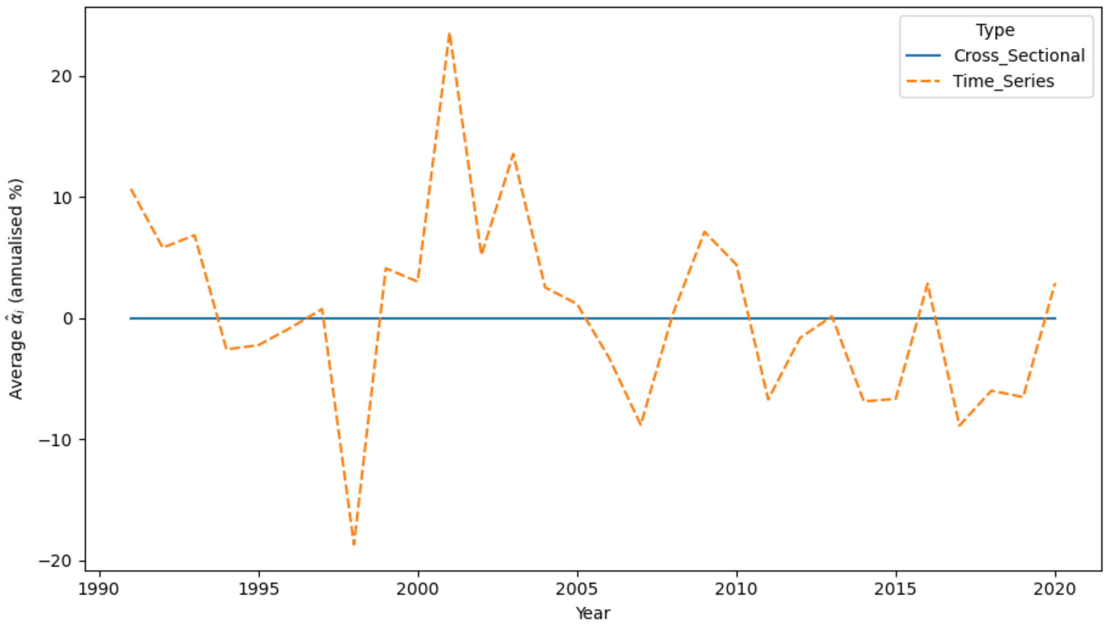


Figure 3.9: Deviation of average annual Alpha estimates from time-series vs. cross-sectional regression

Figure 3.9 plots the equal-weighted average Alpha estimates of a single-index model from time-series regression vs. cross-sectional regression for each sample year (1991-2020). The time-series regression Alphas of each security are calculated as: $\hat{\alpha}_i = \bar{R}_i - \hat{\beta}_i * \bar{R}_M$. By contrast, the cross-sectional regression Alphas are the residuals of a cross-sectional regression of average security returns against betas: $\bar{R}_i = \hat{\lambda}_0 + \hat{\lambda}_0\hat{\beta}_i + \hat{\alpha}_i$. Average Alpha estimates under time-series regression are susceptible to extreme fluctuation from year-to-year. As a result, empirical results produced under time-series regression are more likely to be a product of sample period. By contrast, cross-sectional regression establishes a consistent benchmark across time periods making empirical results less prone to sample period selection.

Figure 3.9 plots the average Alpha across individual securities estimated via time-series vs. cross-sectional regressions for each year of our data sample. The Alpha estimates from time-series regressions tend to deviate greatly from year to year. The average of time-series regression Alpha estimates in any given year is $\frac{1}{N}\sum_{i=1}^{N}\hat{\alpha}_i = \frac{1}{N}\sum_{i=1}^{N}\bar{R}_i - \bar{R}_M * \frac{1}{N}\sum_{i=1}^{N}\hat{\beta}_i$. Therefore, these average Alpha estimates are a function of the difference between the return of an equal-weighted portfolio and equal-weighted average security Beta times our reference portfolio (market proxy). In years where either small cap securities outperform, or the average security Beta is less than 1, we would expect the average Aalpha of securities to be biased upwards. As a result, anomaly identification under time-series regression should be highly sensitive to sample period selection since in certain periods a randomly selected group of securities would be expected to produce positive Alphas whilst in other periods Alphas would be substantially deflated. By contrast, cross-sectional Alpha estimates are consistently centred on zero since the intercept of cross-sectional regressions captures the difference in performance between the return on equal-weighted vs cap-weighted portfolios and deviations of average security Betas from 1. Consequently, anomalies identified under cross-sectional regression should be more robust across varying time periods.

Figure 3.9 showed that each regression approach produces Alpha estimates that are on average different from eachother in each data period. Figure 3.10 below shows that time-series regression also produces Alpha estimates that are systematically biased. The chart splits alpha estimates across years into a sample of securities with Betas less than 1 (Low Beta) versus greater than 1 (High Beta).

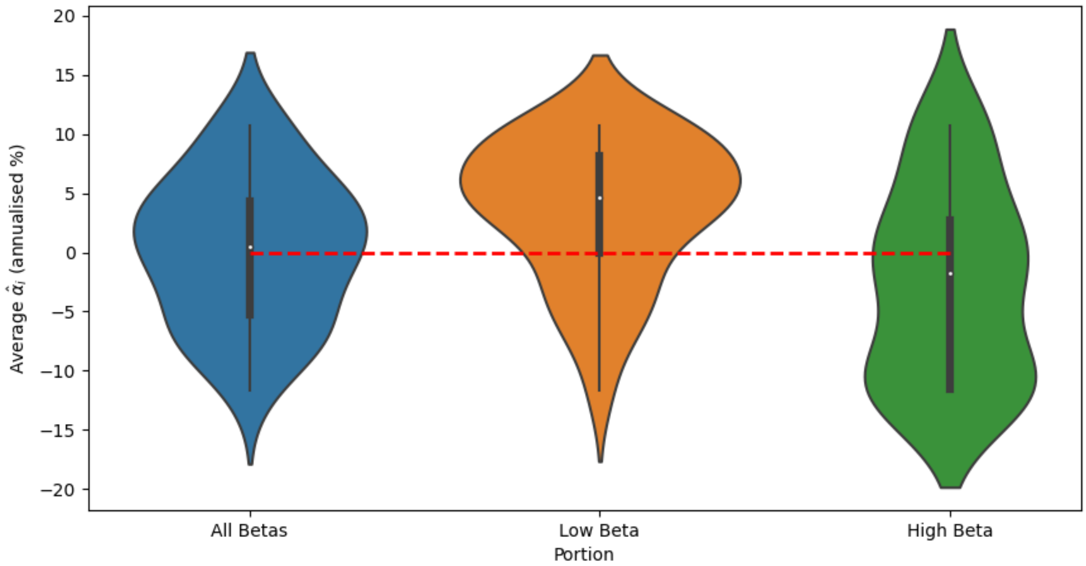


Figure 3.10: Distribution of time-series alpha estimates conditional on beta
Figure 3.10 plots the distribution of Alpha estimates from time-series regression of a single-index model. The figure depicts three samples; "All Betas" which places no restrictions, "Low Beta" which only includes securities with $\hat{\beta}_i <$ 1, and "High Beta" which only includes securities where $\hat{\beta}_i > 1$. The time-series regression Alphas of each security are calculated as: $\hat{\alpha}_i = \bar{R}_i - \hat{\beta}_i * \bar{R}_M$. Since we have 30 sample years each distribution is comprised of 30 data points where each data point is the average alpha across each securities subject to the applicable Beta filter. The red dotted line through the chart represents zero alpha. In Figure 3.10 we observe that Alpha estimates from the time-series regression are distributed above 1 for low Beta securities and below 1 for high Beta securities; Alpha estimates systematically vary depending on security Beta.

We can visually observe that low Beta securities tend to produce higher time-series Alphas whilst high Beta securities produce lower Alphas; there is a systematic dependence between Alpha and Beta. The dependence between alpha and beta estimates occurs because time-series regressions assume the data examined conforms precisely to the relation hypothesised by the model. In periods where the actual cross-sectional relation between risk and return differs substantially from the form assumed by the time-series model then risk-adjustment penalty ($\bar{R}_M * \hat{\beta}_i$) becomes spurious. The empirical relation between risk and return tends to be weaker than theorised (Black, Jensen and Scholes, 1972) so high Beta securities have historically received excessive penalties whilst low Beta securities have not been penalised enough. By contrast, Alphas estimated via cross-sectional regressions apply a risk-adjustment penalty based on the actual risk-return relation observed in the data period; there is no model misspecification. Knowledge of the systematic measurement bias of time-series regression can be exploited to create a practically infinite quantity of false anomalies. We examine this issue in the next section.

3.5 Populating a "factor zoo"

In the preceding sections of the chapter, we outlined how the use of time-series regression can result in significantly biased estimates of security or portfolio Alphas. In bull markets we will tend to observe a "low beta anomaly" due to Alpha estimation bias; low Beta securities appear to outperform. In bear markets, we observe the opposite effect, a "high beta anomaly". In this section, we prescribe a general approach to construct countless false anomalies by intentionally leveraging the model misspecification of time-series regressions. We also discuss how the "betting against beta" (BAB) factor published by Frazzini and Pedersen (2014) directly exploits this Alpha estimation bias.

A simple procedure to consistently generate a false return anomaly is to create a large, diversified portfolio with a beta substantially below 1. Such a portfolio will tend to benefit from inflated estimates of its alpha under a time-series regression since in rising markets Alpha estimates for portfolios further below 1 become progressively more inflated. The formation of long-short portfolios is an intuitive technique to achieve a portfolio Beta well below one. Unsurprisingly, construction of long-short portfolios is a common practice in finance anomaly literature.

By way of example, suppose we decide to form two well-diversified equity portfolios; a low Beta (LB) portfolio comprised of all securities with a Beta <1, and a high Beta (HB) portfolio comprised of all securities with a Beta >1.[34] Jensen's Alpha as measured via time-series regression of a single factor model is simply $\hat{\alpha}_i = \overline{R_i} - \hat{\beta}_i \overline{R_M}$. So, for our LB and HB portfolios, the estimated Alphas are:

$$\hat{\alpha}_{LB} = \overline{R_{LB}} - \hat{\beta}_{LB} \overline{R_M},$$

$$\hat{\alpha}_{HB} = \overline{R_{HB}} - \hat{\beta}_{HB} \overline{R_M}.$$

If we take a long position in the LB portfolio and an equal short position in the HB portfolio then the time-series Alpha estimate for the resulting low-minus-high (LMH) portfolio is as follows:

[34] We note that a breakpoint of 1 is not required, only a meaningful difference in the Betas of each portfolio is required.

$$\hat{\alpha}_{LMH} = \hat{\alpha}_{LB} - \hat{\alpha}_{HB}$$

$$\hat{\alpha}_{LMH} = \overline{R_{LB}} - \hat{\beta}_{LB}\overline{R_M} - \left(\overline{R_{HB}} - \hat{\beta}_{HB}\overline{R_M}\right)$$

$$\hat{\alpha}_{LMH} = (\overline{R_{LB}} - \overline{R_{HB}}) + \overline{R_M}\left(\hat{\beta}_{HB} - \hat{\beta}_{LB}\right) \tag{3.11}$$

If the empirical relation between risk and return is flatter than expected relation under the CAPM (e.g., Black, Jensen and Scholes, 1972), then by consequence whenever $\overline{R_M} > 0$, the following inequality holds:

$$\overline{R_M}\left(\hat{\beta}_{HB} - \hat{\beta}_{LB}\right) > (\overline{R_{LB}} - \overline{R_{HB}}) \tag{3.12}$$

As a result, LMH portfolios will be biased to produce positive Alpha estimates under time-series regression and this bias will increase as $\left(\hat{\beta}_{HB} - \hat{\beta}_{LB}\right)$ increases. The intuition of the worked example is valid for any breakpoint (not just β=1) so long as there are sufficient securities either side of the breakpoint to construct a large, diversified portfolio.

Several other papers highlight that many LMH portfolios are capable of producing positive Alphas; "*Have we found, yet again, a better factor, a mighty inhabitant of the zoo? According to our theory—No. While the LMH portfolio does capture risk, its economic interpretation remains elusive*" (Andrei et al., 2019). What distinguishes us from the extant literature is that we are the first to provide an algebraic rationale as to why the phenomenon occurs. We attribute the perceived outperformance of many LMH portfolios to a systematic measurement bias that occurs when estimating Alphas via time-series regression.

The BAB factor identified by Frazzini and Pedersen (2014) is perhaps the most famous return anomaly of the last decade. The BAB factor is able to consistently generate economically and statistically significant time-series Alphas across various asset classes and countries; it is remarkably robust. Formation of the BAB factor portfolio involves taking a long position in a low Beta portfolio and short position in a high Beta portfolio. The process is similar to the generalised approach to anomaly generation prescribed above but with a couple of differences. The low Beta portfolio comprises securities with below median Betas and places progressively higher investments weights on securities with lower Betas; by design the portfolio will have a Beta particularly far below 1.

Conversely, the high Beta portfolio is comprised by securities with above median Betas with higher weights for higher Beta securities, producing a portfolio with a Beta significantly above 1. This alternate weighting scheme, in place of equal weighting, ensures the Betas are comparatively further away from 1 and increases the severity of the Alpha estimation bias of time-series regressions. We would expect the low Beta portfolio to have artificially inflated (positive) Alphas in rising markets whilst the high Beta portfolio would have artificially deflated (negative) Alphas. Naturally, a LMH Beta portfolio construction would reap the benefits of time-series estimation bias on each constituent portfolio and produce particularly large Alphas, delivering the appearance of a remarkably robust anomaly across numerous data sets.

An unlevered BAB factor would consistently produce large Alphas under time-series regression. However, Frazzini & Pedersen (2014) also leverage or deleverage the low Beta and high Beta portfolios that comprise the BAB factor to ensure each constituent portfolio has a Beta of 1. Under time-series regression, the Alpha of the risk-free rate is zero. Leveraging the low Beta portfolio up to a Beta of 1 by borrowing (shorting) the risk-free rate further inflates the time-series Alpha of the low Beta portfolio. Deleveraging the high Beta portfolio to a Beta of 1 by investing at the risk-free rate reduces the negative time-series Alpha of the high beta portfolio. The leveraging process has conflicting impacts upon Alpha estimation bias for each portfolio. However, if the BAB factor is a net borrower of risk-free securities, then the levered BAB factor will benefit more than the unlevered BAB factor from the Alpha estimation bias of time-series regressions and will report comparatively higher Alphas. This appears to be the case in Frazzini and Pedersen (2014, p. 9) where they note that the BAB factor for US equities is a net borrower of risk-free securities with a short position in risk-free securities which is on average double the long position.

3.6 Conclusion

This chapter highlighted that Alphas estimated via time-series regression suffer from economically and statistically significant bias in empirical applications. The cause is a misspecification of cross-sectional risk premia in time-series regression. Cross-sectional regression approaches are preferred for Alpha estimation since they penalise returns based on the actual risk-return relation observed in empirical data rather than a theoretical relation that can deviate greatly in practice.

In our empirical analysis of US stock data, we have demonstrated that the magnitude of Alpha bias has been economically and statistically significant over the past 30 years. There is a strong motivation for the use of cross-sectional regression in place of time-series approaches to achieve reliable performance evaluation. Widespread use of time-series regression in both academia and industry has resulted in the false identification of many return anomalies. These anomalies tend to exploit model misspecification induced bias via long-short portfolio formations. A preference for The BAB factor of Frazzini and Pedersen (2014) is a prominent recent example of a factor perceived as an anomaly due to an unintentional exploitation of time-series regression Alpha bias. The model misspecification bias that fuelled the rise of the "factor zoo" can be remedied via use of cross-sectional regression. More importantly, this data-driven estimation approach can be adopted for performance evaluation within industry to promote more efficient investment allocation.

There are several avenues in which our research could be extended. It is possible that many of the existing anomalies in the finance and accounting literature are simply a product of biased estimation of Alpha in time-series regressions. Therefore, it would be empirically interesting to revisit the results of papers of anomalies which are dependent upon use of time-series regressions. Do these anomalies tend to lose their statistical significance when examined via cross-sectional regression? How many of the smart beta strategies in industry are simply the product of Alpha estimation bias?

Since a disconnect between theory and empirical observation is the underlying cause of biased Alpha estimates in time-series regression, it is interesting to study why this may be the case. It seems intuitive that investors should want compensation for bearing additional systematic risk exposure, yet the security market line is surprisingly flat empirically. Betas are assumed to capture the exposure of portfolios to systematic risk. Perhaps Betas are a particularly poor proxy for systematic risk exposure. Why might this be the case? How can Betas be improved?

If the security market line is empirically flatter than expected, then, under cross-sectional regression, the risk-free rate will typically be a negative Alpha asset in positive return markets. Investors can generate Alpha by shorting the risk-free asset and taking a long position in other assets. All else equal, the lower the risk-free rate, the more negative the magnitude of the Alpha of the risk-free asset becomes. As a result, more overvalued assets can be invested in whilst still yielding a positive portfolio Alpha due to borrowing at the risk-free rate. Therefore, we have a mechanism for the rational investment in securities trading above their fundamental value. With interest rates having spent much of the past decade at historic lows globally, a study on the impact of lower interest rates on asset price overvaluation would be of great interest to academics and practitioners alike.

Chapter 4

Endogenous Reference Bias: When is it a Problem?

4.1 Introduction

The market portfolio is the most important reference point in performance evaluation throughout empirical finance. The theoretical market portfolio, as contemplated in mean-variance optimisation (Markowitz, 1952), represents all investible securities. Since it is impossible to measure all investible securities (Roll, 1977), more practical proxies for the market portfolio have been adopted for performance evaluation. The two most used proxies for the market portfolio are a market-capitalisation-weighted (cap-weighted) portfolio and an equal-weighted portfolio. However, both portfolios suffer from an "endogenous reference bias" since dependent variables are regressed against a portfolio of which they are a constituent. Modern finance literature places great concern on mitigating endogeneity to arrive at unbiased parameter estimates. However, endogenous reference bias is commonly ignored *in asset pricing literature*.[35] Therefore, the primary motivation for this chapter is to examine the severity of endogenous reference bias, provide practical guidance on when the bias is likely to undermine traditional empirical analysis, and present a robust alternative estimation procedure that adjusts for bias.

To achieve our objective, we complete three primary tasks. Firstly, we provide a mathematical derivation of the bias induced by use of an endogenous regressor. Our results indicate that there are, mainly, three competing forces influencing the magnitude of endogenous reference bias: (1) an asset's weight in the reference portfolio, (2) the size of the asset's return variance relative to the variance of the market (variance multiple), and (3) the asset's exogenous Beta.[36] Whilst it is intuitive that large-cap securities

[35] In contrast, in the market microstructure literature it is common to adjust for endogenous reference bias (e.g., see Chordia et al., 2000; p.10).

[36] An asset's exogenous Beta is defined as the beta obtained from regressing an asset against a market proxy, or index, which excludes itself. For example, the exogenous Beta of Apple might be obtained by regressing the stock returns of Apple against a cap-weighted index of all stocks in the S&P 500 *excluding* Apple. The logic is to avoid the regression of a dependent variable

experience the most severe bias via increased weighting in reference portfolios, our findings show that this endogenous reference bias is not monotonically increasing in an asset's weight in the market portfolio. Instead, both the direction of bias and magnitude can be significantly impacted by the other two factors and becomes more pronounced at higher weight levels.

Secondly, we evaluate the historical magnitude of endogenous reference bias. We examine US stock data from 1991-2020. A focus on US data enhances the relevance of our research to the extant studies and facilitates comparison. More importantly, since the US market is one of the least concentrated markets in the world, it establishes a conservative lower bound on the level of bias expected.[37] We look at two distinct sample groups: one where we consider a market portfolio constructed of the 500 largest stocks in the sample year, analogous to the S&P 500, and another where we consider a market portfolio constructed of the 30 largest stocks, in parallel to the Dow Jones Industrial Average (DJIA). We find that the effects of endogenous reference bias are economically negligible for most stocks when using the more diversified S&P 500 proxy as a regressor. By contrast, the effect of bias becomes much more pronounced for the DJIA proxy with a cap-weighted average absolute bias in market Beta estimates of 0.08 across sample years; producing a 170 basis points distortion in Alpha estimates. We also find that whilst the weight of a security in the market portfolio is a statistically and economically significant source of bias, it only accounts for a minor portion (~10.7%) of the variation in endogenous reference bias. By comparison, inclusion of variables that account for an asset's variance multiple and exogenous Beta can significantly enhance explanatory power. Certain stock characteristics are more exposed to severe endogenous reference bias (low Beta, high variance) even when using more diversified reference portfolios. The implication is that research which relies on use of a concentrated market portfolio is likely to be significantly biased and that constituent weight is only one of multiple determinants of bias direction and magnitude.

Finally, we examine the influence of endogenous reference bias in a portfolio setting. Thus, in evaluating endogenous reference bias we analyse portfolio Betas rather

---

against an "independent" variable of which it is a constituent since it will produce biased estimates of market Betas due to endogeneity.

[37] However, even for one of the most diversified indexes in the world, index concentration can still be significant. In the US, the top 3 stocks in the S&P 500 comprised 15.3% of the index as of 28th February 2023: AAPL (6.6%), MSFT (5.6%), and GOOG (3.1%).

than single security Betas. Our simulation results show that if the CAPM accurately described security returns, then even for a large, diversified portfolio, which comprises a substantial portion of the market portfolio, endogenous reference bias is unlikely to be economically meaningful. This occurs because the variance multiple is close to one for diversified portfolios moderating the impact of endogenous reference bias. Moreover, at a variance multiple of one, and a weighted average constituent exogenous Beta of one, a diversified portfolio's Betas is expected to be equal approximately one irrespective of how large or small a constituent it is of the market portfolio. In stark contrast, when diversified portfolios intentionally comprise low Beta or high Beta securities, then the bias can become severe with traditional Beta estimates for low and high-Beta portfolios being biased towards unity. We form an endogenous reference bias hedging portfolio for use as a dependent variable in single-factor and Fama-French 3-Factor (FF3F) model regressions. This test portfolio, hedged against endogenous reference bias, delivers a single-factor / (FF3F) model Alpha of 4.30% / (4.28%) across our 30-year sample period. The Alpha generated by this hedging portfolio is economically larger than the single-factor model Alphas associated with the small cap (2.04%) and value premiums (1.58%) during the same period.

Our results shed light on some of the most persistent asset pricing anomalies. The finance literature consistently observes an empirical security market line that is "too flat" (Black et al., 1972). Our findings demonstrate that the perceived flatness of the empirical security market line is partially attributable to endogenous reference bias. It is common to use Beta-grouped portfolios in tests of the security market line in order to address concerns about errors-in-variables bias (e.g., Fama & French, 2004). Since an errors-in-variables problem with Beta estimation is upheld as a suspected cause of the excessive flatness of the SML, a common resolution has been to form Beta-grouped decile portfolios when plotting an empirical SML in order to reduce estimation error. Ironically, we demonstrate in Section 4.5 that this grouping procedure itself provably contributes to the excessive flatness of the SML exacerbating the very problem it attempts to resolve.

The remainder of this chapter is organised as follows. Section 4.2 discusses related literature and distinguishes our contribution. Section 4.3 demonstrates analytically that endogenous reference bias distorts Beta estimation and discusses remedies to alleviate the issue. Section 4.4 quantifies the historical impact of endogenous reference bias on market Beta estimation for US stocks. Section 4.5 explores the impact of endogenous reference bias in diversified portfolios. Section 4.6 summarises our main findings, discusses limitations and opportunities for further research, and then concludes.

4.2 Related literature and contribution

Literature on endogenous reference bias remains scarce but there are some existing papers which also consider the issue. Woo et al. (1994) appear to be the first to explicitly consider the bias. They focus on two emerging Asian stock markets and conclude that the OLS estimated Betas of the largest securities tended to be biased by 5-7%. Their proposed solution was to adopt an instrumental variable (IV) approach for Beta estimation where the IV selected was a portfolio of all other securities in the market excluding the dependent variable. Unfortunately, their paper has garnered little attention, with zero citations to date. A working paper by Damodaran (1999) also mentions the problem of endogenous reference bias amidst a broader discussion of the pitfalls of regression Betas. Patton and Verardo (2012) investigate whether stock Betas vary with the release of firm-specific news and consider the impact of endogenous reference bias during robustness tests. To correct for the bias, they omit the dependent variable from the market index and find minimal impact upon their baseline results.

The paper most closely related to this chapter is by Malloch et al. (2016). The authors decompose the bias affecting Beta estimates into an errors-in-variables (EIV) bias and an endogenous reference bias. Similarly to Patton and Verardo (2012), they omit the dependent variable from the market index to correct for endogenous reference bias. They perform simulations for common market proxies across 39 developed and emerging markets, finding that endogenous reference bias is more severe in emerging markets where market indexes have a higher concentration of the largest stocks. In their simulation of the S&P 500, anticipated to be among the least susceptible of financial markets to the bias, they observe an average deviation of Betas from their true value of 0.15.

This chapter contributes to existing literature in three important ways. Firstly, this is the first research to consider the impact of three competing influences on the direction of magnitude of endogenous reference bias; facilitating richer insights into when endogenous reference bias will be a significant problem. For example, Malloch et al. (2016) only consider constituent weight and idiosyncratic risk leading to claims which we dispute: endogenous reference bias "*increases monotonically with the stock's weighting in the index and the magnitude of its idiosyncratic risk*" (2016, p. 4293). In this chapter we demonstrate analytically that the severity of endogenous reference bias is *not* monotonically increasing with a stock's weighting in an index. We examine the impact of varying levels of idiosyncratic risk, index weighting, and stock Betas upon endogenous reference bias. We find a non-monotonic relation and a directional bias that differs dependent on the exogenous Beta of a stock. An implication is that whilst index weight is an important influence, neither the severity of bias, nor even the direction of bias, can be assumed without considering a stock's exogenous Beta.

None of the aforementioned papers examine the impact of endogenous reference bias upon the empirical data of a major developed market. The primary purpose of this chapter is to provide relevant guidance on if and when endogenous reference bias is likely to be severe and impact empirical analysis. Therefore, the second contribution of this chapter is to quantify the historical extent of bias in a major developed stock-market. We focus on analysis of US stock data to maximise the relevance of our findings. We find that Beta estimation bias is unlikely to be severe for *most* individual securities when working with US stock data. Our empirical results for historical US stock data point to a significantly lower average bias than identified by Malloch et al. (2016) in their simulation of US data.

The preceding literature focuses on the impact of endogenous reference bias for individual assets. It is common in finance to evaluate the performance of portfolios. Examples include anomaly literature and performance evaluation of the mutual fund industry. Therefore, it is important to consider whether endogenous reference bias has a similar or differing impact on portfolios. As our final novel contribution, this chapter is the first to explore the impact of endogenous reference bias on portfolios. We show that for a randomly selected portfolio of securities, the bias will generally be negligible despite the portfolio being a relatively larger constituent of any market portfolio. This occurs due to diversified portfolios having a low variance multiple which reduces the potential for upwards bias of parameter estimates. An exception occurs when a portfolio is not randomly formed but is instead comprised of high or low exogenous Beta securities. Endogenous reference bias causes a wider dispersion of traditional beta estimates when compared to exogenous Beta estimates. Traditional Beta estimates for beta-grouped portfolios are subject to increasing positive bias (overstatement) as portfolio Betas increase beyond zero. An implication is that historical tests of the slope of the security market line which rely on Beta-based portfolio groupings are likely to significantly *understate* the slope of the security market line. The empirical performance of the CAPM may be much better than currently envisioned after adjusting for endogenous reference bias. Finally, we demonstrate that a simple hedging portfolio against endogenous reference bias is capable of generating Alphas economically larger than both the famous small-cap and value premiums in our sample. In the next section we provide a derivation of endogenous reference bias and explore how its direction and magnitude differ in response to several key variables.

4.3 Analytical examination of bias and proposed remedy

4.3.1 Endogenous reference bias

To demonstrate the effect of endogenous reference bias, we examine its impact on a CAPM model:

$$E[R_{i,t}] = \beta_i E[R_{M,t}] + \epsilon_{i,t}, \quad \textbf{(4.1)}$$

where i denotes an asset; t denotes a time increment; $R_{i,t}$ and $R_{M,t}$ are returns on asset i and the market in excess of the risk-free rate; β_i is the exposure of asset i to market risk; and $\epsilon_{i,t}$ is an error term. We adopt the following assumptions:

A1: $Var(R_{i,t}) = \sigma_i^2$

A2: $Var(R_{M,t}) = \sigma_M^2$

A3: $\mathrm{E}[\epsilon_{i,t}] = 0 \ \forall_i$

A4: $Var(\epsilon_{i,t}) = \sigma_{\epsilon_i}^2 \ \forall_i$

A5: $Cov(R_{M,t}, \epsilon_{j,t}) = 0 \ \forall_i$

A6: $Cov(\epsilon_{i,t}, \epsilon_{j,t}) = 0 \ \forall_{i \neq j}$

The true market portfolio is unobservable, resulting in an errors-in-variable problem since any portfolio acting as a proxy for the market portfolio differs from the true unobservable market portfolio. This chapter is focused explicitly on the bias due to inclusion of the dependent variable in the regressor. As a result, we make an important simplifying assumption in this section that our reference portfolio is measured *without error*.

There are two common proxies used in empirical finance to represent the market portfolio: an equal-weighted portfolio and a cap-weighted portfolio. We use a generalised approach throughout this section that allows for unequal weighting of securities and produces results that also hold for the special cases of equal-weighting and cap-weighting. Our chosen proxy for the market portfolio is a weighted return of constituent securities: $R_{Mt} = \sum_{i=1}^{N} w_i \, R_{i,t}$; where N is the total number of securities in the market; w_i is the relative cap-weight of asset i and $\sum_{i=1}^{N} w_i = 1$.[38] Consider asset j. We can decompose the market return into two components: the weighted return on asset j, and the sum of weighted returns on the rest of the assets in the market: $R_{M,t} = w_j \, R_{j,t} + (1-w_j)\tilde{R}_{M,t}$, where $\tilde{R}_{M,t} = \frac{1}{(1-w_j)} \sum_{i=1, i\neq j}^{N-1} w_i \, R_{i,t}$.[39] For our asset of interest, asset j, Equation 4.1 can be restated as:

$$R_{j,t} = \beta_j \left[w_j R_{j,t} + (1-w_j)\tilde{R}_{M,t} \right] + \epsilon_{j,t}, \qquad \textbf{(4.2)}$$

where β_j is a parameter to be estimated against an endogenous regressor that is comprised of a w_j weight in the dependent variable (endogenous component) and a $(1-w_j)$ weight in the rest of the market (exogenous component). The "independent" variable of Equation 4.1, $R_{M,t}$, is comprised of the dependent variable $R_{j,t}$, as seen in Equation 4.2. As a result, the regressor of Equation 4.1 is perfectly correlated with the regression residuals, so estimates of Beta are biased and inconsistent (assumption A.5 is violated). Our biased OLS estimate of β_j is:

$$\hat{\beta}_j = \frac{w_j \lambda_j + (1-w_j)\tilde{\beta}_j}{w_j{}^2 \lambda_j + (1-w_j)^2 + 2w_j(1-w_j)\tilde{\beta}_j} \;,^{40} \qquad \textbf{(4.3)}$$

where $\tilde{\beta}_j = \frac{cov(R_{j,t}, \; \tilde{R}_{M,t})}{var(\tilde{R}_{M,t})}$ (an "exogenous Beta" measuring the covariance of the excess returns of the test asset with the rest of the market portfolio, and divided by the variance of the excess returns of the rest of the market portfolio), and $\lambda_j = \frac{\sigma_j^2}{\sigma_M^2}$ (a "variance multiple" indicating the variance in excess returns of the test asset relative to the variance in excess returns of the endogenous market proxy). If we were to regress the dependent

[38] For simplicity, w_i is assumed to be time-invariant over the measurement period.

[39] $\frac{1}{(1-w_j)}$ is applied to scale the portfolio return up to a weighting of 100%.

[40] An extended derivation is provided in Appendix C.3.

variable against only the endogenous component ($w_j = 1$), then our estimated $\hat{\beta}_j = 1$, since we are regressing a variable exclusively against itself. By contrast, if we were to regress the dependent variable against solely the exogenous component ($w_j = 0$) then our estimated $\hat{\beta}_j = \tilde{\beta}_j$, which is an estimate of the Beta of asset j in relation to the rest of the market. We know that under endogenous reference bias, $\hat{\beta}_j$ will be partially weighted towards 1 and partially weighted towards $\tilde{\beta}_j$ (exogenous Beta). From Equation 4.3 we can observe that a stock's weight in the reference portfolio (w_j), its variance relative to the variance of the market (λ_j), and its exogenous Beta in relation to the rest of the market ($\tilde{\beta}_j$) all contribute to endogenous reference bias. These interactions may induce non-monotonic behaviour in the severity of the bias.

We present several illustrative cases to demonstrate the nonlinear nature of endogenous reference bias. For example, assume a scenario in which the excess returns of asset j were uncorrelated with the rest of the market. Our Beta estimate from Equation 4.3 becomes:

$$\hat{\beta}_j = \frac{w_j M_j}{{w_j}^2 M_j + \left(1 - w_j\right)^2}. \tag{4.4}$$

The historical standard deviation of daily returns of the S&P 500 from 1 January, 1991 to 31 December, 2020 was 1.15% (or 23.8% annualised), which we will use as a proxy for the standard deviation of the market portfolio in this hypothetical example. Assume that our uncorrelated asset j also has 6.65 times the variance of the rest of the market.[41] We know that an *unbiased* beta for asset j ($\tilde{\beta}_j$) would be zero since it is uncorrelated with the rest of the market. Figure 4.1 plots $\hat{\beta}_j$ for our uncorrelated asset j as w_j increases and endogenous reference bias increases in severity.

[41] This represents the variance multiple on Apple, the largest company listed on the S&P 500, over the same time period.

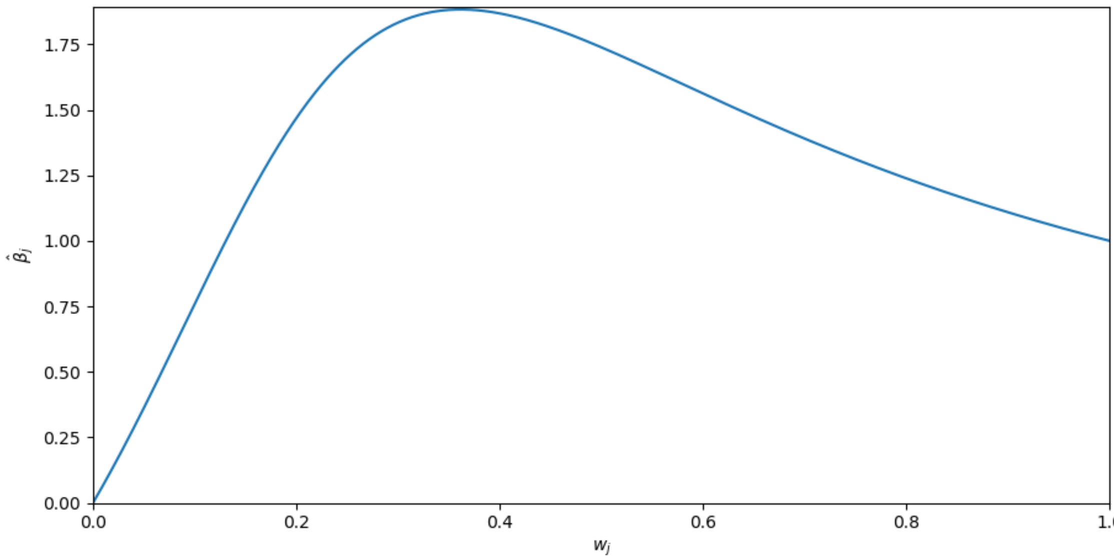


Figure 4.1: Bias in Beta estimate for an uncorrelated asset j as w_j increases
Figure 4.1 shows the traditional estimated Beta for an asset 'j' that is uncorrelated with the rest of the market as w_j increases. w_j is the weight of the asset in the market portfolio. Since asset j is uncorrelated with the market, an unbiased estimate of its Beta would be $\hat{\beta}_j = 0$. However, due to endogenous reference bias, $\hat{\beta}_j$ is non-zero and increasingly biased as w_j increases. When $w_j = 1$, $\hat{\beta}_j = 1$, since asset j is now being regressed solely against itself.

The increase in Beta estimation bias as w_j increases is non-linear. If asset j had a weighting of $\frac{1}{500}$ in the reference portfolio, then estimation bias would be mild, with $\hat{\beta}_j - \tilde{\beta}_j = 0.01$. However, when $w_j = 5\%$, representative of the largest weighted securities in the S&P 500, then Beta estimation bias becomes far more severe with $\hat{\beta}_j - \tilde{\beta}_j = 0.36$. To put this bias into a performance evaluation perspective, if the excess return of the reference portfolio in a year was 15%, then the estimate of Jensen's Alpha for asset j would be understated by 5.4%; more than 1/3 of the market risk premium![42]

Using Equation 4.3, we can identify the extent of endogenous reference bias as w_j, $\tilde{\beta}_j$, and M_j change. In Figure 4.2 below we plot four alternate $\tilde{\beta}_j$ for increasing w_j; $\tilde{\beta}_j = [-1,1,2,3]$. Other settings remain the same as above.

[42] Jensen's Alpha: $\hat{\alpha}_j = \frac{1}{T}\sum_{t=1}^{T} R_{j,t} - \hat{\beta}_j * \frac{1}{T}\sum_{t=1}^{T} R_{M,t}$. Bias: $-(\hat{\beta}_j - \tilde{\beta}_j) * \frac{1}{T}\sum_{t=1}^{T} R_{M,t} = -0.36 * 0.15 = -5.4\%$

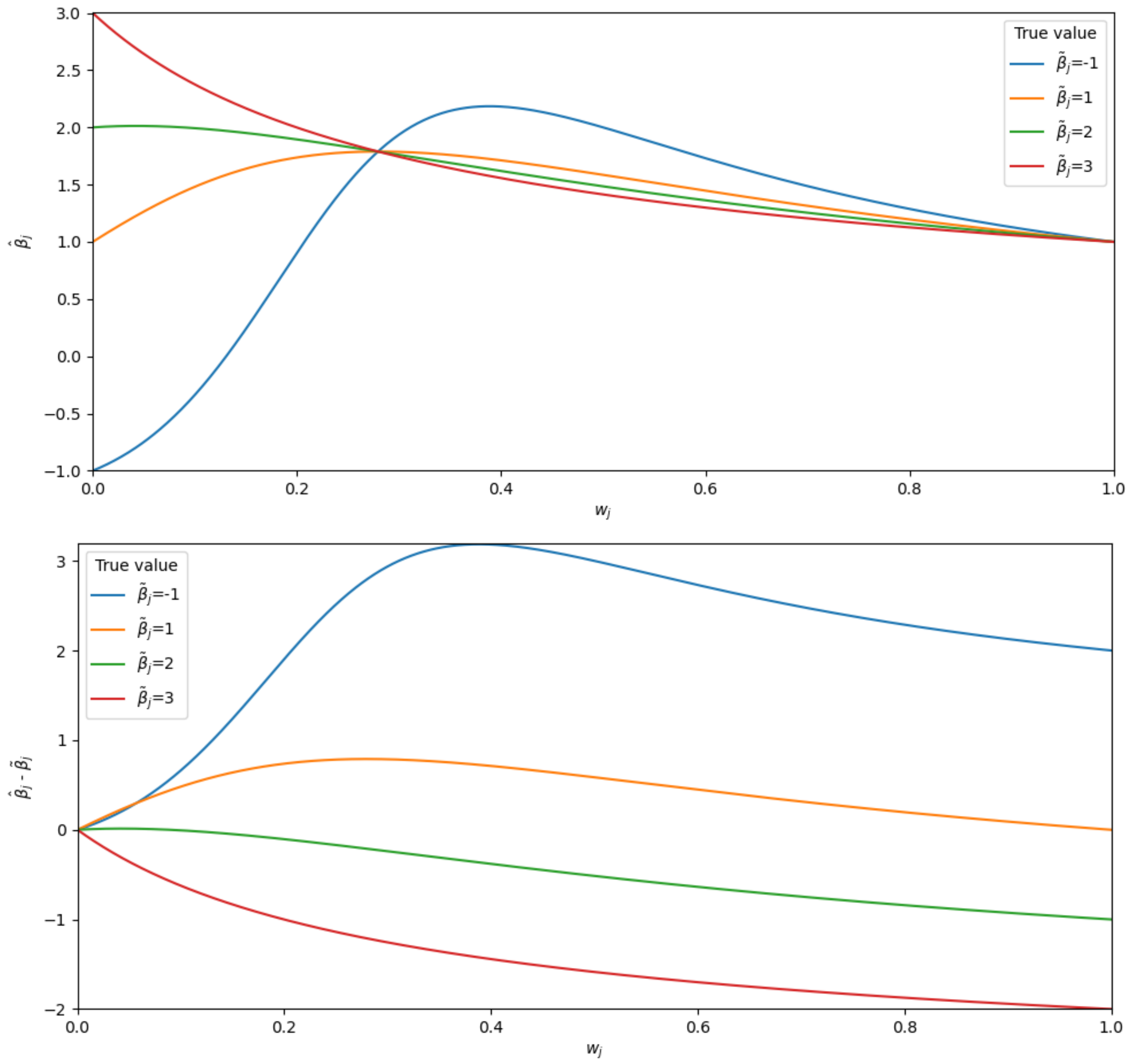


Figure 4.2: Bias in the Beta estimate as w_j increases for varying $\tilde{\beta}_j$

Figure 4.2 (top panel) shows the estimated Beta for an asset 'j', with a true Beta of β_j^{RoM} with the rest of the market, as w_j increases. w_j is the weight of the asset in the market portfolio. If Beta estimation was unbiased then we would observe $\hat{\beta}_j = \tilde{\beta}_j$. However, due to endogenous reference bias, $\hat{\beta}_j$ differs from $\tilde{\beta}_j$ as w_j increases. This bias is evident in the bottom panel which shows the difference between both estimates. For example, when $w_j = 5\%$, $(\hat{\beta}_j - \tilde{\beta}_j) = 0.25, 0.26, 0.01, -0.36$ for the cases where $\tilde{\beta}_j = -1,1,2,$ and 3, respectively. When $w_j = 1$, $\hat{\beta}_j = 1$ under each case, since asset j is now being regressed solely against itself.

We can observe from Figure 4.2 that the Beta estimates of the largest securities in a market are likely to suffer from substantial bias. However, the exogenous Beta of a stock in relation to the rest of the market ($\tilde{\beta}_j$) is also an important determinant of the severity and direction of endogenous reference bias. Also evident is the potential for all securities in a market to suffer from the same directional Beta bias; upwards biased betas for large-cap stocks do not necessitate downwards biased betas for small-cap stocks and vice versa. Figure 4.3 below illustrates a third influence on the bias; the multiple of a stock's variance to the variance of the endogenous market portfolio (λ_j).

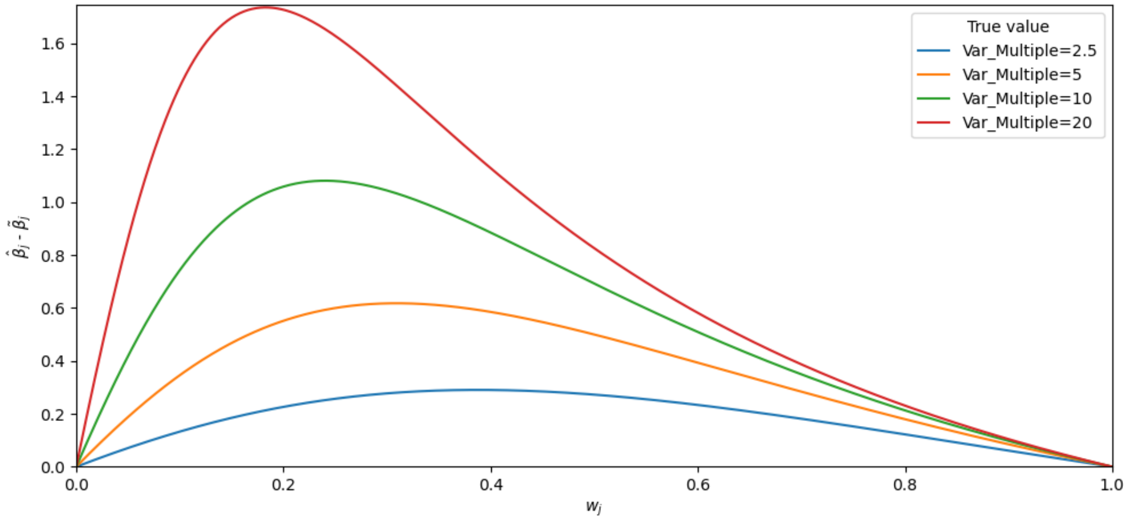


Figure 4.3: Bias in the Beta estimate as w_j increases for varying (M_j).
Figure 4.3 shows the estimated bias in traditional beta estimates for an asset 'j', with a true beta of $\tilde{\beta}_j = 1$ with the rest of the market, as w_j and M_j increase. w_j is the weight of the asset in the market portfolio. If Beta estimation was unbiased then we would observe $\hat{\beta}_j = \tilde{\beta}_j$. However, due to endogenous reference bias, $\hat{\beta}_j$ differs from $\tilde{\beta}_j$ as w_j and λ_j increase. An increasing variance multiple results in increasingly severe bias indicating that volatile stocks are likely to be the most severely impacted by endogenous reference bias.

As the variance multiple increases, we observe a progressively stronger bias in Beta estimates. This implies that the most volatile stocks in a market are likely to be the most severely impacted by endogenous reference bias. In the next section we consider remedies for unbiased Beta estimation.

4.3.2 Towards unbiased estimation

To remedy the endogenous reference bias, Woo et al. (1994) propose an IV approach. Their proposed instrument is $\tilde{R}_{M,t}$; the excess return of a cap-weighted portfolio comprised of all other assets except for asset j. They perform the regression: $R_{M,t} = \hat{a} + \hat{b}\tilde{R}_{M,t} + \hat{e}_{i,t}$, then use the fitted values of $R_{M,t}$ in the original regression model as per Equation 4.5:

$$R_{j,t} = \alpha_j + \beta_j \hat{R}_{M,t} + \epsilon_{i,t}. \quad \textbf{(4.5)}$$

The instrument chosen is highly correlated with the original regressor and uncorrelated with the residuals. As a result, Beta estimation should theoretically become unbiased. However, there is not necessarily a need for an IV approach in this setting. We already know the endogenous and exogenous components of the original regressor. Therefore, we could alternatively replace the original regressor with its exogenous counterpart (substitution approach):

$$R_{j,t} = \alpha_j + \beta_j \tilde{R}_{M,t} + \epsilon_{j,t}. \quad \textbf{(4.6)}$$

where j denotes an asset; t denotes a time increment; $R_{j,t}$ is the return of asset j in excess of the risk-free rate; $\tilde{R}_{M,t}$ is the excess return of a cap-weighted portfolio comprised of all other assets except for asset j; β_i is the exposure of asset j to market risk; and $\epsilon_{j,t}$ is an error term.

This substitution approach has also been adopted by Patton and Verardo (2012) and Malloch et al. (2016) and is our preferred approach. Given the competing options, we can explore the impact of each estimation approach via simulation. We analyse two bookend scenarios; one where security returns are white noise (i.e. should be independent of the market) and another where security returns are determined by the CAPM. For our first simulation (Simulation 1), we generate 100 (N) series of additive Gaussian white noise. Each series ($z_{i,t}$) is of 1,000 observations in length. The first of the 100 series becomes our dependent variable "asset j". We then select a weight for asset j and form a market portfolio (our regressor) which has a w_j weighting in asset j and equal $\frac{(1-w_j)}{N-1}$ weighting in each of the remaining $N-1$ white noise series. We hold w_j constant and repeat the simulation 10,000 times. In each iteration, we estimate:

$$y_t = \hat{\alpha} + \hat{\beta} x_t + \hat{\varepsilon}_t, \quad \textbf{(4.7)}$$

where $x_t = w_j y_t + \frac{(1-w_j)}{N-1} \sum_{i=1,i\neq j}^{N-1} z_{i,t}$ with y_t representing the return series of "asset j"; $\hat{\alpha}$ is a regression constant; and $\hat{\varepsilon}_t$ are regression residuals.

We record and compare the Beta estimates obtained under an unadjusted approach, IV approach, and substitution approach. Absent endogenous reference bias, our estimated Betas should be approximately zero across simulations. By contrast, in the presence of the endogenous regressor we expect the biased OLS Beta estimates to be:[43]

$$\hat{\beta}_j = \frac{w_j}{w_j{}^2 + \frac{(1-w_j)}{N-1}^2}. \quad \textbf{(4.8)}$$

---

[43] $\hat{\beta}_j = \frac{cov(y_t, x_t)}{var(x_t)} = \frac{cov\left(y_t,\ w_j y_t + \frac{(1-w_j)}{N-1} \sum_{i=1}^{N-1} z_{i,t}\right)}{var\left(w_j y_t + \frac{(1-w_j)}{N-1} \sum_{i=1}^{N-1} z_{i,t}\right)} = \frac{w_j var(y_t) + \frac{(1-w_j)}{N-1} cov(y_t, \sum_{i=1}^{N-1} z_{i,t})}{w_j{}^2 var(y_t) + \left(\frac{(1-w_j)}{N-1}\right)^2 var(\sum_{i=1}^{N-1} z_{i,t}) + 2 w_j \frac{(1-w_j)}{N-1} cov(y_t, \sum_{i=1}^{N-1} z_{i,t})} = \frac{w_j}{w_j{}^2 + \frac{(1-w_j)}{N-1}^2}$

Since the exogenous Beta is zero, $\hat{\beta}_j$ is also equivalent to value of endogenous reference bias ($\hat{\beta}_j - \tilde{\beta}_j$) in this scenario. For our simulation we test $w_j = [0.01, 0.05]$. Based on Equation 4.8, we expect to observe $\hat{\beta}_j \sim [1.0, 4.3]$ in the presence of endogenous reference bias. Table 4.1 compares the simulation results for the three estimation approaches.

		$\hat{\beta}_j^{ERB}$	$\hat{\beta}_j^{IV}$	$\hat{\beta}_j^{Sub}$
$w_j = 1\%$				
	Mean estimate	1.0004	-0.0011	0.0000
	St dev.	0.3191	0.3219	0.3187
$w_j = 5\%$				
	Mean estimate	4.3059	-0.0056	0.0000
	St dev.	0.2634	0.3356	0.3187

Table 4.1: Comparison of estimation approaches for Simulation 1

Table 4.1 compares Beta estimation approaches in the presence of endogenous reference bias. Since our dependent variable in Simulation 1 is white noise, we know that absent endogenous reference bias an accurate estimator of Beta would produce $\hat{\beta}_j = 0$. The traditional estimation approach ($\hat{\beta}_j^{ERB}$) is heavily impacted by the bias, producing estimates that differ greatly from $\hat{\beta}_j = 0$. Both the IV approach and substitution approach are designed to produce unbiased estimates in the presence of an endogenous regressor. We observe that both approaches yield Beta estimates close to the true value. Of the two approaches, the substitution method results in estimates closer to the true value and with a lower standard deviation.

Table 4.1 shows strong bias in the Beta estimates of a traditional estimation approach which differ greatly from the true Beta value of zero. The extent of bias is consistent with the predictions of Equation 4.8. By contrast, the substitution approach delivers unbiased Beta estimates in the presence of endogenous reference bias with a high level of precision. Similarly, the IV approach produces Beta estimates close to the true value albeit with a slightly lower level of precision. The IV approach is impacted by choice of w_j since the dependent variable of Equation 4.7 is likewise the dependent variable used in obtaining the fitted values of the instrumental variable. The substitution approach delivers the same estimate irrespective of the w_j selected.

Simulation 1 compared the performance of estimation approaches in a situation where our dependent variable and the regressor should be unrelated. We also want to understand whether one approach is superior to another in a situation where our regressor does have a causal impact on the dependent variable whilst also being impacted by endogenous reference bias. We explore this in Simulation 2. For this second simulation we model a CAPM-style relation within our data; asset returns are a linear function of the returns of an “unobservable” market portfolio. We perform 10,000 iterations of

Simulation 2 and record results. In each iteration we simulate one daily valuation series for an "unobservable" market portfolio using Geometric Brownian motion:

$$V_t = V_{t-1} * e^{\left(\mu - \frac{\sigma^2}{2}\right)d + \sigma W_t}, \tag{4.9}$$

where V_t is the market value of the unobservable market factor at time t; V_{t-1} is the prior period market value; μ is the annual percentage drift, set at 10% across simulations; σ is the annual standard deviation of returns, set at 0.15 across simulations; d is a time increment, set to $\frac{1}{252}$; and W_t is Brownian motion. The return ($r_{M,t}$) of the unobservable market factor in period t is $r_{M,t} = \frac{(V_t - V_{t-1})}{V_{t-1}}$. In each iteration, we simulate 100 daily asset returns series of 252 periods as per the equation:

$$R_{j,t} = \beta_j R_{M,t} + \epsilon_{j,t}, \tag{4.10}$$

where $r_{j,t}$ is the return of asset j in period t; $r_{f,t}$ is the return on a risk-free asset in period t, set to zero for convenience; β_j is the loading of asset j on the unobservable market factor with $\beta_j \sim N(\beta_c, 0.2)$; and $\epsilon_{j,t} \sim N(0, 0.001)$ is noise. An "observable" market portfolio ($R^*_{M,t}$) is then constructed as an equal-weighted portfolio comprised of the 100 simulated asset returns. Since the true market factor is unobservable, we perform a single-factor model regression to obtain estimates of β_j but use the observable $R^*_{M,t}$ as our independent variable in place of the unobservable market portfolio. This time-series regression model is shown in Equation 4.11 below:

$$R_{j,t} = \hat{\alpha}_j + \hat{\beta}_j R_{M,t} + \hat{\epsilon}_{j,t}, \tag{4.11}$$

It is intuitive for a CAPM-style simulation to centre simulated Betas around 1. However, an observation of Betas distributed around 1 is tautological. Whether Betas in relation to the unobservable market factor truly were or were not centred around 1, a traditional approach to estimation would estimate Betas cross-sectionally centred around 1 due to endogeneity.[44] As a result, Simulation 2 examines β_c values of 0, 0.5, and 1. Results are presented in Table 4.2.

---

[44] If the reference portfolio is an equal-weighted portfolio then use of OLS results in a cross-sectional average Beta of 1 for constituents in the presences of endogenous reference bias. If the reference portfolio is a cap-weighted portfolio then a cap-weighted average Beta of 1 will be observed in the presence of endogenous reference bias. Even if return series were white noise they would appear to follow the market in aggregate (Beta estimates distributed around 1) under a traditional Beta estimation approach despite being completely uninfluenced by the market portfolio in actuality (*unbiased* Beta estimates distributed around 0).

True value		$\hat{\beta}_j^{ERB}$	$\hat{\beta}_j^{IV}$	$\hat{\beta}_j^{Sub}$
$\beta_c = 0$				
	Mean estimate	0.9919	-0.0092	-0.0078
	St dev.	0.3604	0.3661	0.3623
$\beta_c = 0.5$				
	Mean estimate	1.0141	0.5041	0.5024
	St dev.	0.4544	0.4585	0.4552
$\beta_c = 1$				
	Mean estimate	1.0076	0.9802	0.9804
	St dev.	0.1984	0.1983	0.2002

Table 4.2: Comparison of estimation approaches for Simulation 2

Table 4.2 compares Beta estimation approaches in the presence of endogenous reference bias. A single-factor model regression is used: $R_{j,t} = \hat{\alpha}_j + \hat{\beta}_j R_{M,t} + \hat{\epsilon}_{j,t}$. Since our dependent variable in Simulation 1 is white noise, we know that absent endogenous reference bias an accurate estimator of Beta would produce $\hat{\beta}_j = 0$. The traditional estimation approach ($\hat{\beta}_j^{ERB}$) is heavily impacted by bias, producing estimates that differ greatly from $\hat{\beta}_j = 0$. Both the IV approach and substitution approach are designed to produce unbiased estimates in the presence of an endogenous regressor. We observe that both approaches yield beta estimates close to the true value. Of the two approaches, the substitution method results in estimates closer to the true value and with a lower standard deviation.

Simulations 1 and 2 both indicate that a substitution approach is capable of generating marginally more precise Beta estimates than an IV approach in the presence of endogenous reference bias. The substitution approach also has the advantage of being able to deliver unbiased Beta estimates without the need to run multiple regressions.[45] In the following section, we proceed to analyse the extent of historical Beta estimation bias for large US stocks.

4.4 Historical beta estimation bias in the US stock market

For endogenous reference bias to warrant the concern of finance practitioners and researchers it should be a sufficiently severe problem in empirical data. Market capitalisation is a proxy for the aggregate value that investors attribute to an asset. Since higher market capitalisations result in higher weightings in a cap-weighted benchmark, the most valued assets in the world are likely to be the most severely impacted by endogenous reference bias. However, it is not *ex ante* certain that the extent of bias will be severe. This is evident in Equation 4.3 which demonstrates that in addition to portfolio

[45] $R_{M,t}$ can easily be converted to $\tilde{R}_{M,t}$ and used in its place; $\tilde{R}_{M,t} = \frac{R_{M,t} - w_j R_{j,t}}{1 - w_j}$.

weight, a stock's variance in relation to market variance and a stock's exogenous Beta have a significant influence on the direction and severity of Beta estimation bias.

In this section, we obtain Beta estimates for US stocks using a traditional (biased) Beta estimation approach (see Equation 4.1) as well as the substitution approach introduced in the preceding section (see Equation 4.6) which is designed to be unbiased in the presence of endogenous reference bias. Our analysis relies on daily US equities holding period return data. The data are obtained from the Center for Research in Security Prices (CRSP) and covers the sample period from January 1, 1991 to December 31, 2020. Only ordinary common shares listed on the NYSE, NASDAQ, and AMEX are included.[46] We omit all microcaps, defined as stocks smaller than the 20th percentile of the market equity of NYSE stocks at the beginning of the first trading day of a calendar year, from our sample. In the calculation of holding period returns we assume all dividends received are reinvested in the underlying stock on the date that the dividends are paid out. Unless otherwise specified, the "market factor" $(R_{M,t})$ from Equation 4.1 is constructed as the excess holding-period return of a market-capitalisation-weighted index formed on the first trading day of January and held one year. The daily risk-free rate used is the monthly rate obtained from Kenneth French's website and converted into a daily format. For the substitution approach of Equation 4.6, the exogenous market factor $\tilde{R}_{M,t}$ is constructed in the same manner as the market factor of Equation 4.1 but without inclusion of the asset that constitutes the dependent variable:

$$\tilde{R}_{M,t} = \frac{R_{M,t} - w_j R_{j,t}}{1 - w_j}. \quad \textbf{(4.12)}$$

Our first objective is to quantify the typical direction and magnitude of endogenous reference bias for the US stock market. We compare two scenarios: one where our sample and market proxy comprises the 500 largest US stocks in each year (representative of the S&P 500) and another where our sample and market proxy comprises the 30 largest US stocks in each year (analogous to the Dow Jones Industrial Index). Comparing both scenarios facilitates bias assessment in our empirical data. Table 4.3 below reports traditional (biased) Beta estimates ($\hat{\beta}_j$) vs Beta estimates adjusted for endogenous reference bias ($\tilde{\beta}_j$). The values presented are the cap-weighted averages in

[46] CRSP filters: share code (shrcd) is equal to 10 or 11 (i.e., ordinary common shares), exchange code (exchd) is equal to 1, 2, or 3 (i.e. NYSE, AMEX, NASDAQ).

each period. $\bar{\Delta}_{\beta}$ Represents the cap-weighted average absolute difference between the two estimates.

	500 Largest Stocks			30 Largest Stocks		
Period	$\hat{\beta}_j$	$\tilde{\beta}_j$	$\bar{\Delta}_{\beta}$	$\hat{\beta}_j$	$\tilde{\beta}_j$	$\bar{\Delta}_{\beta}$
1991 - 1995	0.99	0.96	**0.03**	0.99	0.85	**0.14**
1996 - 2000	0.99	0.97	**0.03**	1.00	0.92	**0.08**
2001 - 2005	0.99	0.97	**0.02**	1.00	0.92	**0.08**
2006 - 2010	1.00	0.99	**0.01**	1.00	0.95	**0.06**
2011 - 2015	1.00	0.98	**0.02**	1.00	0.92	**0.08**
2016 - 2020	0.99	0.98	**0.02**	1.00	0.93	**0.07**
All periods	**1.00**	**0.98**	**0.02**	**1.00**	**0.91**	**0.08**

Table 4.3: Typical endogenous reference bias for varying US benchmarks (1991-2020)

Table 4.3 shows the direction and magnitude of endogenous reference bias for the US market over the 30-year period. $\hat{\beta}_j$ represents the biased OLS estimates of Beta. $\tilde{\beta}_j$ represents Beta estimates adjusted for endogenous reference bias. $\bar{\Delta}_{\beta}$ represents the cap-weighted average absolute difference between the two estimates. The table shows that for the larger market proxy of 500 stocks the magnitude of endogenous reference bias is small. By contrast, the bias is far more significant when using a smaller market proxy.

Table 4.3 shows that the magnitude of endogenous reference bias for the larger market proxy of 500 stocks is economically insignificant for most securities; with a cap-weighted average absolute bias of security OLS Beta estimates of 0.0205. However, the magnitude of bias was severe for some securities. Appendix C.1 lists the 5 stocks in each year with the most severely biased Betas across our sample period. The most biased stock in each year had an average absolute Beta estimation bias of 0.0941. Using a traditional Beta estimation approach would result in Jensen's Alpha for these stocks being, on average, misstated by an absolute value of 1.55%. By contrast, the bias to Beta estimates when using a smaller, more concentrated, market proxy is noticeably more severe across the sample with a cap-weighted average absolute bias of 0.0850.

Table 4.5 shows the corresponding distortion to Alpha estimates of individual securities under a single-factor model regression. Alpha estimates when using a market proxy of the 500 largest stocks suffer a cap-weighted average distortion of around 0.42% across the 30-year sample period. This Alpha distortion more than triples when moving to the smaller market index of 30 stocks.

	500 Largest Stocks			30 Largest Stocks		
Period	$\hat{\alpha}_j$	$\tilde{\alpha}_j$	$\bar{\Delta}_\alpha$	$\hat{\alpha}_j$	$\tilde{\alpha}_j$	$\bar{\Delta}_\alpha$
1991 - 1995	0.86%	1.36%	**0.51%**	0.61%	2.56%	**2.05%**
1996 - 2000	0.58%	0.93%	**0.70%**	1.18%	2.38%	**2.63%**
2001 - 2005	2.10%	2.07%	**0.31%**	1.63%	1.42%	**1.05%**
2006 - 2010	0.92%	1.06%	**0.25%**	0.09%	0.74%	**1.21%**
2011 - 2015	-0.03%	0.22%	**0.28%**	0.12%	1.39%	**1.38%**
2016 - 2020	0.40%	0.84%	**0.47%**	0.26%	2.04%	**1.86%**
All periods	**0.80%**	**1.08%**	**0.42%**	**0.65%**	**1.75%**	**1.70%**

Table 4.4: Impact of endogenous reference bias on Alpha estimation for varying US benchmarks (1991-2020)

Table 4.4 shows the direction and magnitude of endogenous reference bias for the US market over the 30-year period. $\hat{\alpha}_j$ represents the biased OLS estimates of Jensen's Alpha. $\tilde{\alpha}_j$ represents estimates of Jensen's Alpha adjusted for endogenous reference bias. $\bar{\Delta}_\alpha$ represents the cap-weighted average absolute difference between the two estimates. The table shows that for the larger market proxy of 500 stocks the magnitude of endogenous reference bias is small. By contrast, the bias is far larger when using a smaller market proxy.

To understand how the interaction between exogenous Betas and size affects endogenous reference bias we stratify our top 500 sample into percentiles. In each year, we sort Betas by size decile then split each decile into 10 groups based on their exogenous beta. Table 4.5 shows the mean absolute percentage error (MAPE) in Beta estimates: MAPE $= \frac{100}{N}\sum\left|\frac{\hat{\beta}_j-\tilde{\beta}_j}{\tilde{\beta}_j}\right|$. Consistent with the analytical results presented in Figure 4.2, stocks with higher market capitalisations and lower betas are more susceptible to bias.

	Low $\tilde{\beta}_j$	2	3	4	5	6	7	8	9	High $\tilde{\beta}_j$
Small MC	0.51	0.22	0.20	0.18	0.19	0.20	0.21	0.20	0.21	0.21
2	0.50	0.27	0.23	0.21	0.19	0.14	0.16	0.18	0.19	0.20
3	0.69	0.31	0.26	0.21	0.22	0.21	0.21	0.20	0.25	0.24
4	1.48	0.37	0.32	0.24	0.22	0.28	0.27	0.24	0.27	0.28
5	1.37	0.41	0.35	0.33	0.28	0.31	0.27	0.27	0.31	0.32
6	1.54	0.59	0.42	0.42	0.35	0.34	0.35	0.32	0.35	0.38
7	2.13	0.73	0.56	0.48	0.42	0.41	0.38	0.38	0.45	0.49
8	1.33	0.83	0.66	0.68	0.57	0.56	0.49	0.47	0.57	0.61
9	6.01	1.26	1.12	0.99	0.82	0.67	0.83	0.76	0.80	0.90
Large MC	12.74	4.50	3.71	3.03	2.45	1.96	2.11	2.23	2.04	1.90

Table 4.5: MAPE of Beta estimates stratified by size and $\tilde{\beta}_j$ (1991-2020)

Table 4.5 shows the mean absolute percentage error in Beta estimates for stocks stratified by size (market capitalisation) then exogenous Beta. Traditional OLS Beta estimates of stocks with larger market capitalisations, and lower exogenous Betas, are subjected to progressively more severe upwards bias.

Similarly, stocks with higher market capitalisations and higher variance multiples are more susceptible to bias as shown in Table 4.6.

	Low M_j	2	3	4	5	6	7	8	9	High M_j
Small MC	0.07	0.08	0.08	0.09	0.10	0.11	0.12	0.13	0.14	0.17
2	0.06	0.08	0.08	0.08	0.09	0.11	0.11	0.11	0.15	0.16
3	0.08	0.09	0.10	0.09	0.10	0.14	0.14	0.17	0.18	0.22
4	0.09	0.12	0.11	0.16	0.18	0.22	0.28	0.29	0.34	0.52
5	0.14	0.14	0.11	0.19	0.18	0.21	0.25	0.31	0.33	0.44
6	0.12	0.14	0.19	0.19	0.26	0.23	0.27	0.35	0.41	0.53
7	0.18	0.19	0.19	0.26	0.30	0.38	0.38	0.49	0.53	0.72
8	0.20	0.27	0.27	0.28	0.26	0.29	0.38	0.31	0.37	0.42
9	0.28	0.39	0.41	0.58	0.64	0.74	0.95	1.16	1.34	1.98
Large MC	0.59	0.75	0.90	1.01	1.29	1.63	1.88	2.42	2.90	4.82

Table 4.6: MAPE of Beta estimates stratified by size and variance (1991-2020)
Table 4.6 shows the mean absolute percentage error in Beta estimates for stocks stratified by size (market capitalisation) then variance multiple (the variance of a stock's daily returns relative to the daily returns of a market-capitalisation-weighted index). Traditional OLS Beta estimates of stocks with larger market capitalisations, and higher variance multiples, are subjected to progressively more severe upwards bias.

Equation 4.3 detailed a closed form solution for the value of $\hat{\beta}_j$, demonstrating that it was a biased estimator of a security's exogenous Beta. Preceding papers have considered the weight of a security in an index as the lone determinant of the severity of endogenous reference bias. By contrast, we derived evidence that a security's variance multiple and exogenous Beta are additional factors that significantly impact both the magnitude and direction of endogenous reference bias. This can be further demonstrated by comparing the performance of two simple linear fixed effects models.

Although Equation 4.3 is non-linear, Figure 4.2 suggests that the magnitude of endogenous reference bias, $(\hat{\beta}_{j,t} - \tilde{\beta}_{j,t})$, monotonically decreases as $\tilde{\beta}_{j,t}$ (exogenous Beta) increases. Figure 4.3 suggests that $(\hat{\beta}_{j,t} - \tilde{\beta}_{j,t})$ monotonically increases as $\lambda_{j,t}$ (variance multiple) increases. As a consequence, simple linear models may be able to provide reasonable approximations of endogenous reference bias which would provide insight as to whether the additional drivers have a significant effect on the direction and magnitude of bias. We compare the historical performance of three fixed effects (FE) models. A fixed-effects model was identified as preferable to a random effects model based on the results of a Hausman-test for each data sample. Standard errors in each fixed effects model are clustered based on entity and time.

FE Model 1, presented in Equation 4.13, is a baseline model which considers only constituent weight in the reference portfolio as an influence on the direction and magnitude of endogenous reference bias:

$$\Delta_{\boldsymbol{\beta},\boldsymbol{j},\boldsymbol{t}} = \hat{\phi}_0 + \hat{\phi}_W w_{j,t} + \hat{\gamma}_t + \hat{\delta}_j + \hat{\varepsilon}_{j,t} \quad , \qquad \textbf{(4.13)}$$

where $\Delta_{\boldsymbol{\beta},\boldsymbol{j},\boldsymbol{t}}$ is the endogenous reference bias ($\hat{\beta}_{j,t} - \tilde{\beta}_{j,t}$) for a Beta estimate of asset j in period t; $\hat{\phi}_0$ is a model constant; $\hat{\phi}_W$ is the estimated coefficients on asset j's % weight ($w_{j,t}$) in the market proxy in period t, $\hat{\gamma}_t$ accounts for time fixed-effects, $\hat{\delta}_j$ accounts for entity fixed-effects, and $\hat{\varepsilon}_{j,t}$ are the model residuals. We introduce a second model, FE Model 2, which also considers the effect of an asset's variance multiple and exogenous beta as determinants of endogenous reference bias:

$$\Delta_{\boldsymbol{\beta},\boldsymbol{j},\boldsymbol{t}} = \hat{\phi}_0 + \hat{\phi}_{WM}(w_{j,t} * \lambda_{j,t}) + \hat{\phi}_{WB}(w_{j,t} * \tilde{\beta}_{j,t}) + \hat{\gamma}_t + \hat{\delta}_j + \hat{\varepsilon}_{j,t} \quad , \quad \textbf{(4.14)}$$

where $\hat{\phi}_{WM}$ and $\hat{\phi}_{WB}$ represent coefficients on interaction terms of asset j's variance multiple with weight, and exogenous Beta with weight in period t, respectively; whilst the remaining parameters are the same as the preceding model. We include a final model, FE Model 3, which has the same parameters as FE Model 2 as well as $\hat{\phi}_W w_{j,t}$ from FE Model 1. The results of each fixed-effects model across the 30-year sample are presented in Table 4.7 below.

	FE Model 1	**FE Model 2**	**FE Model 3**
$\hat{\phi}_0$	0.0038***	0.0017***	0.0003
$\hat{\phi}_W$	0.0138***		0.0188***
$\hat{\phi}_{WM}$		0.0082***	0.0085***
$\hat{\phi}_{WB}$		-0.0142***	-0.0272***
Firm Fixed Effects	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes
Model R^2	0.1068	0.8674	0.9172
F-test: *p*-value	0.0000	0.0000	0.0000

Table 4.7: Drivers of endogenous reference bias (1991-2020)

Table 4.7 shows the results of the three fixed-effects models across the 1991-2020 sample period. *,**,*** indicate statistical significance at the 90%, 95%, and 99% confidence level, respectively. The table shows that whilst weight is an important driver of endogenous reference bias, alone it only accounts for a small proportion of the bias. By contrast, a model which includes the additional drivers of endogenous reference bias introduced in this chapter has much greater explanatory power. The interaction terms of variance multiple and exogenous Beta are both statistically significant determinants of endogenous reference bias at the 99% confidence level.

FE Model 1 shows that whilst constituent weight is an important determinant of endogenous reference bias it only accounts for a small portion of the variation (~10.7%). FE Model 2 shows an alternate specification which includes interaction terms of weight with variance multiple and exogenous Betas accounts for significantly more variation (86.7%). Stocks with a higher variance multiple exhibit more severe upwards biased betas. Stocks with lower exogenous betas exhibit upwards biased betas. Finally, FE

Model 3 shows that inclusion of all 3 terms can account for nearly all of the variation in endogenous reference bias (91.7%).

Section 4.4 explored historical endogenous reference bias in individual US equities. The magnitude of bias was, on average, small for a large, diversified market proxy like the S&P 500. However, certain characteristics were shown to render stocks significantly more vulnerable to severe endogeneity. Generally, worse endogeneity occurs at higher weight levels. As a result, it's important to consider the effect of endogenous reference bias when we move away from individual assets and into a portfolio context. Portfolios are likely to have a much higher weight in a constituent index than individual stocks. We explore the consequence of endogenous reference bias within a portfolio context in the following section.

4.5 Impact of bias on portfolios

4.5.1 Portfolio formation based on stock characteristics

Constituent weight is a primary driver of the magnitude of endogenous reference bias. In the preceding section, we observed that assets which are sizable constituents of the market proxy naturally tend to exhibit greater bias. Despite having much greater weight than individual assets in any market constituent, portfolio Betas do not necessarily suffer from endogenous reference bias. The primary reason is that the variance of a large, diversified portfolio 'P' will very quickly approach a variance multiple close to 1. With reference to Equation 4.3 from earlier, in the limit, as λ approaches 1, we have:

$$\hat{\beta}_P \approx \frac{w_P+(1-w_P)\widetilde{\beta}_P}{{w_P}^2+(1-w_P)^2+2w_P(1-w_P)\widetilde{\beta}_P}. \tag{4.15}$$

Typically, a large, diversified portfolio that was selected randomly will also have an exogenous Beta close to 1. When this is the case, endogenous reference bias becomes negligible irrespective of w_j:

$$\hat{\beta}_P \approx \frac{w_P+(1-w_P)}{{w_P}^2+(1-w_P)^2+2w_P(1-w_P)} = \frac{1}{w_P+(1-w_P)} = 1. \tag{4.16}$$

An exception occurs when the composition of a portfolio is skewed towards constituents based on their exogenous Betas. We explore the resultant bias via simulation. We adopt the same simulation approach and settings as used in Simulation 2 in Section 3.2. However, we now construct exogenous Beta-grouped decile portfolios. We compare traditional Beta estimates and exogenous Beta estimates for each Beta decile and plot the security market line for each estimation approach together in Figure 4.4.

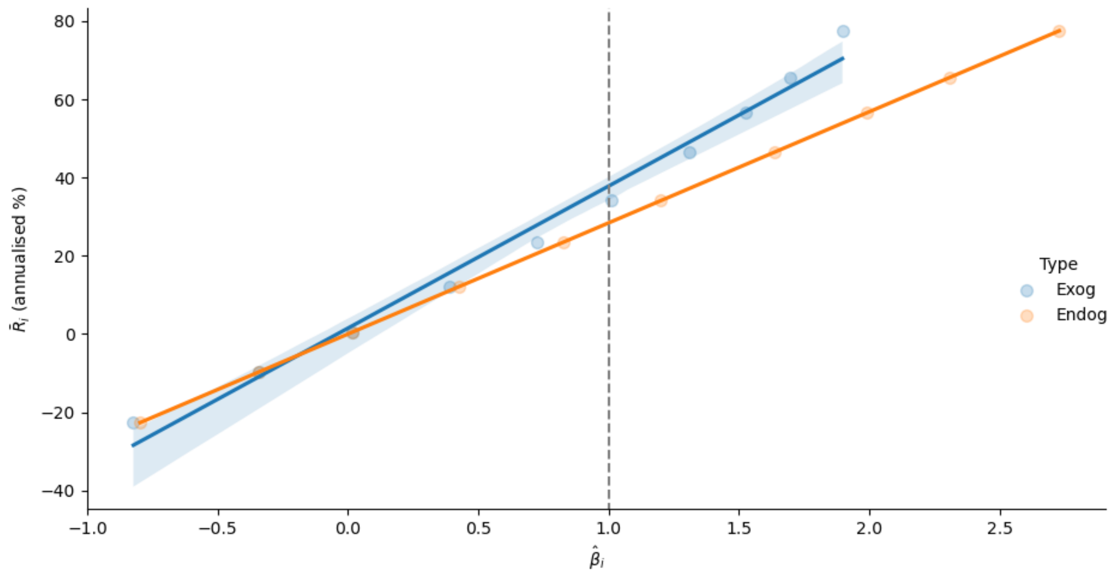


Figure 4.4: Security market line based on Beta-grouped decile portfolios.
Figure 4.4 shows that portfolios grouped by Beta deciles are subject to endogenous reference bias. The severity of bias worsens for more extreme Beta deciles. The use of an endogenous market proxy results in an artificially flat security market line since endogenous betas exhibit a wider dispersion.

Figure 4.4 shows that portfolios grouped by Beta deciles are subject to endogenous reference bias. The severity of this bias worsens as betas deviate further from zero. To understand why this is the case, consider a decomposition of OLS betas into two components: correlation (ρ_{PM}) and relative volatility ($\frac{\sigma_P}{\sigma_M}$) as per Equation 4.17 below:

$$\hat{\beta}_P = \rho_{PM} * \frac{\sigma_P}{\sigma_M}. \tag{4.17}$$

When we transition to use of an exogenous reference as our market portfolio, then correlation is highly likely to decrease since we are no longer regressing a portfolio partially upon itself. Similarly, all else equal, the standard deviation of an exogenous reference portfolio is likely to be higher than an endogenous reference portfolio since it is comprised of less securities. Hence both components facilitate a progressively wider dispersion of Beta estimates when using an endogenous market proxy as we move away from exogenous betas of zero. Consequently, this bias will consistently result in an excessively flat security market line when plotting empirical data where endogenous estimates of portfolio Betas have been used.

An alternate way to understand the impact of endogenous reference bias is in the context of performance evaluation. Since high Beta portfolios face the most severe upwards bias to Beta estimates, they will be excessively punished in factor models such as the CAPM. An investor using Jensen's Alpha to evaluate their performance would over penalise performance by $(\hat{\beta}_P - \tilde{\beta}_P) * \frac{1}{T}\sum_{t=1}^{T} R_{M,t}$; which is likely to be extremely high for high-Beta portfolios, rendering the false appearance of persistent underperformance.[47] It is unsurprising that Frazzini and Pedersen (2014) find the consistent underperformance of high-Beta assets across a variety of asset classes when using a traditional Beta estimation approach.

4.5.2 Hedging against endogenous reference bias

One of the most influential models in finance history, the CAPM (Sharpe, 1964; Treynor, 1962; Lintner, 1965; Mossin, 1966), asserts that security returns should be a linear function of their market risk exposure ($E[R_{i,t}] = \beta_i E[R_{M,t}]$). The support for this assertion in historical data appears weak; the CAPM's empirical performance is "*poor enough to invalidate the way it is used in applications*" (Fama and French, 2004, p. 25). If the CAPM holds empirically, then there should not be a noticeable difference in the risk-adjusted returns (α_i) of low and high Beta securities. However, investment strategies which overweight low Beta securities have historically tended to generate positive Alphas across a variety of asset classes (e.g. see Frazzini and Pedersen, 2014). The perceived underperformance of high Beta securities has also supported the development of literature on risk-parity portfolios which seek to deliver superior risk-adjusted returns by weighting assets by the inverse of their risk (e.g. $\frac{1}{\beta_i}$) and then leveraging portfolio returns up or down to achieve pre-determined return targets. Why might such strategies appear to perform well? If these strategies work, why haven't they been arbitraged out of the market?

It's *possible* there is some underlying economic rationale for the perceived outperformance of low Beta securities, in which case weighting assets inversely to their betas (e.g. risk-parity strategies) would deliver genuinely superior risk-adjusted returns. However, it's *certain* that if there was no difference between the risk-adjusted returns of low and high Beta securities, we would still commonly observe the outperformance of low Beta securities in the regression results of simple factor models such as the CAPM.

---

[47] Jensen's alpha: $\hat{\alpha}_j = \frac{1}{T}\sum_{t=1}^{T} R_{j,t} - \hat{\beta}_j * \frac{1}{T}\sum_{t=1}^{T} R_{M,t}$. Excess risk-adjustment penalty due to beta bias: $-(\hat{\beta}_j - \tilde{\beta}_j) * \frac{1}{T}\sum_{t=1}^{T} R_{M,t}$.

Figure 4.5 below demonstrates this issue by comparing two bookends for the distribution of security returns relative to the CAPM.

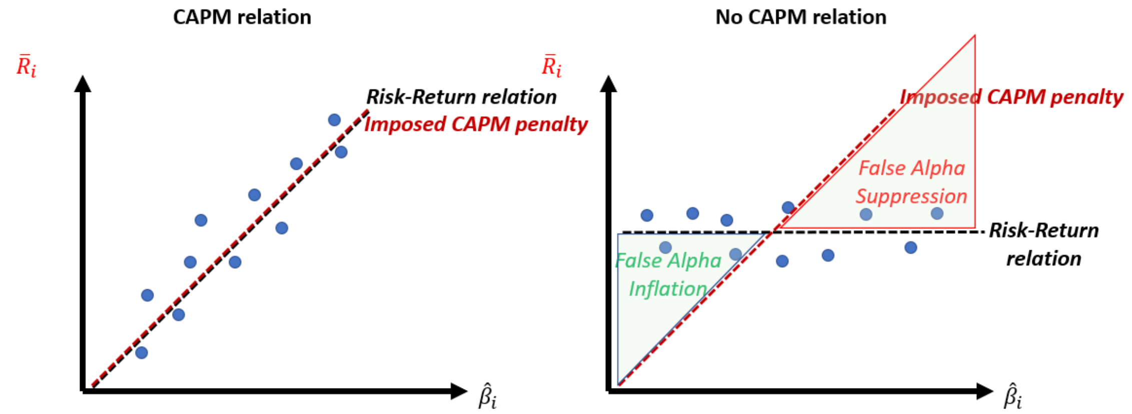


Figure 4.5: No risk premium misspecification vs. extreme risk premium misspecification

Figure 4.5 presents an illustrative contrast of two scenarios; Scenario 1 (left panel) in which the CAPM accurately describes an underlying data set against Scenario 2 (right panel) in which there is no relation between security returns and systematic risk. Each blue dot represents an observation. The black dotted line on each chart represents the fitted regression line from a cross-sectional regression ($\bar{R}_i = \hat{c} + \hat{\lambda}_M \hat{\beta}_{M,i}$). The dotted red line represents the penalty imposed on Jensen's Alpha ($\hat{\beta}_i * \bar{R}_M$). When the CAPM accurately reflects the underlying data (Scenario 1), the cross-sectional risk-return relation is equivalent to the penalty imposed by the CAPM and estimates of Jensen's Alpha are unbiased. By contrast, when there is no relation between security returns and market risk (Scenario 2), the imposed CAPM penalty is spurious. In data which deviates from the CAPM, Jensen's Alpha is progressively biased upwards (downwards) for lower (higher) beta securities or portfolios when $\bar{R}_M$ is positive, and the opposite is true when $\bar{R}_M$ is negative.

Figure 4.5 (left panel) shows an extreme situation in which all security returns, on average, follow the CAPM. If historical security data was accurately described by the CAPM then there would be no statistically significant difference between the Alpha estimates of high and low beta securities. It would be impossible for inverse-beta-weighted (e.g. risk-parity) strategies to appear to consistently outperform. However, we know historical security returns tend to substantially deviate from the precise distribution specified by the CAPM. By contrast, Figure 4.5 (right panel) shows an opposing extreme where security returns are independent of the market risk factor. In this extreme, all Beta estimates are completely spurious since market risk exposure has no bearing on returns. However, the "risk-adjusted" returns of high Beta securities would appear much lower since they receive much heavier (entirely spurious) return penalties; the imposed CAPM penalty remains the same irrespective of the cross-sectional relation present in data.

The key takeaway is that if the CAPM poorly explains historical data then factor model regression don't estimate risk-adjusted returns, rather they apply a spurious penalty to asset returns. This spurious penalty is more severe for high Beta securities.

Low Beta securities will not genuinely deliver superior risk-adjusted returns, they simply systematically benefit from a biased performance evaluation framework.

It is highly probable that any historical data we examine will deviate, at least partially, from the CAPM. Consequently, traditional approaches to Beta estimation will be heavily biased. Therefore, it is desirable to identify Beta estimators that are more robust to data samples which deviate from the CAPM. If the exogenous Beta estimation approach presented in this chapter provides more meaningful estimates of a stock's co-movement with risk factors in noisy data, then portfolios formed on exogenous Betas will appear to generate significantly greater Alphas than those formed on traditional (endogenous) OLS Beta estimates. We can test this assertion for our historical 30-year sample of US stock data by forming an endogenous reference bias "hedging" portfolio. Appendix C.4 discusses the rationale as to why a hedging portfolio demonstrates that exogenous Beta estimates are more resilient to model misspecification.

We start by forming two distinct Beta-weighted portfolios whereby portfolio weights are based on a security's Beta estimate divided by the sum of all security Beta estimates (such that weights sum to 1):

$$w_i = \frac{\beta_i}{\sum \beta_i}. \quad \textbf{(4.18)}$$

Here, w_i represents the weight of the respective Beta-weighted portfolio towards a security. The key difference between the two portfolios constructed is that we calculate the weights for one portfolio, w_i^{Trad}, based on traditional OLS Beta estimates ($\hat{\beta}_i$) whilst for the other portfolio, weights, w_i^{Exog}, are calculated using the exogenous Beta ($\tilde{\beta}_i$) estimation approach presented in this chapter. We then construct a simple hedging portfolio which takes a long position in the portfolio constructed using an exogenous estimation approach and a short position in the portfolio constructed using a traditional estimation approach. The hedging portfolio weights, w_i^{Hedge}, sum to zero. This hedging portfolio can be interpreted as hedging against endogenous reference bias. The hedging portfolio is short securities which have little co-movement with the returns of the market proxy yet abnormally high traditional Beta estimates due to severe endogeneity. Conversely, the portfolio is long in securities where traditional and exogenous Beta estimates are similar. In data which perfectly conformed to the CAPM it would be impossible for either constituent portfolio to consistently outperform the other. However, as examined data progressively deviates from the distribution expected by the CAPM, the hedging portfolios should generate progressively larger Alphas.

Having formed our hedging portfolio, we evaluate its historical performance when regressed against a single-factor model (Equation 4.11) and Fama-French 3-Factor Model (presented below):

$$R_{j,t} = \hat{\alpha}_j + \hat{\beta}_{j,Mkt} R_{M,t} + \hat{\beta}_{j,SMB} R_{SMB,t} + \hat{\beta}_{j,HML} R_{HML,t} + \hat{\epsilon}_{j,t}, \quad \textbf{(4.19)}$$

where $R_{j,t}$ is the excess return of the test asset at time t; $R_{j,Mkt}$, $R_{j,SMB}$, and $R_{j,HML}$ are the excess returns of the market portfolio, SMB portfolio, and HML portfolio, respectively; $\hat{\beta}_{j,Mkt}$, $\hat{\beta}_{j,SMB}$, and $\hat{\beta}_{j,HML}$ are the risk-loadings of the test asset upon these respective factor portfolios; $\hat{\alpha}_j$ is the estimated risk-adjusted return of the test asset, and $\hat{\epsilon}_{j,t}$ are estimated residuals of the test asset at time t. We also test whether using an equal-weighted or cap-weighted market proxy for the market portfolio influences parameter estimates; Mkt_{EW} indicates use of an equal-weighted proxy whilst Mkt_{CW} indicates use of cap-weighted proxy in Equation 4.19. The results of each factor regression model across the 30-year sample are presented in Table 4.8 below.

	SFM (Mkt_{EW})	**SFM** (Mkt_{CW})	**FF3F** (Mkt_{EW})	**FF3F** (Mkt_{CW})
$\hat{\alpha}_{Hedge}$ (%)	4.30* (0.0691)	4.23* (0.0737)	4.28* (0.0699)	4.23* (0.0721)
$\hat{\beta}_{Mkt}$	-0.0793*** (0.0000)	-0.0806*** (0.0000)	-0.0812*** (0.0000)	-0.0811*** (0.0000)
$\hat{\beta}_{SMB}$			0.0159 (0.2845)	-0.0004 (0.9803)
$\hat{\beta}_{HML}$			-0.0056 (0.7720)	-0.0152 (0.4366)
Model R^2	0.0140	0.0122	0.0141	0.0123
F-test: *p*-value	0.0000	0.0000	0.0000	0.0000

Table 4.8: Alpha of portfolio hedged against endogenous reference bias (1991-2020)
Table 4.8 shows the performance of a portfolio hedged against endogenous reference bias across the 1991-2020 sample period. *,**,*** indicate statistical significance at the 90%, 95%, and 99% confidence level, respectively. These time-series regression coefficients include *p*-values in parentheses under each estimate. SFM indicates a single-factor model regression where the excess returns of the hedging portfolio are regressed against the excess returns of the market proxy. FF3F indicates a Fama-French 3-Factor Model regression. Mkt_{EW} indicates that the chosen market proxy was an equal-weighted portfolio, whilst Mkt_{CW} indicates the market proxy was a cap-weighted portfolio. $\hat{\alpha}_{Hedge}$ represents the annual % return of an endogenous reference bias hedging portfolio that can not be attributed to the risk factors included in each regression. The $\hat{\beta}$ are respective factor loadings estimated under a traditional OLS approach (i.e. subject to endogeneity) The table shows that a simple hedging portfolio against endogenous reference bias generates an economically large annual alpha (greater than 4.20%) in each regression. These Alpha estimates are statistically

significant at the 90% confidence level in each regression. The interpretation is that exogenous estimates of security betas provide substantially more accurate representation of a security's co-movement with a market proxy than traditional OLS estimates which are confounded by endogeneity. Consequently, they will appear to produce Alpha in a performance evaluation setting where returns are penalised by endogenous betas (e.g. $\hat{\alpha}_i = \bar{R}_i - \hat{\beta}_i \bar{R}_{Mkt}$ for a single-factor model).

Table 4.8 shows that a simple hedging portfolio against endogenous reference bias generates an economically large annual Alpha (greater than 4.20%) in each regression. These Alpha estimates are statistically significant at the 90% confidence level in each regression. The interpretation is that exogenous estimates of security Betas provide a substantially more accurate representation of a security's co-movement with a market proxy than traditional OLS estimates which are confounded by endogeneity. Consequently, they will appear to produce Alpha in a performance evaluation setting where returns are penalised by traditional (endogenous) OLS Betas (e.g. $\hat{\alpha}_i = \bar{R}_i - \hat{\beta}_i \bar{R}_{Mkt}$ for a single-factor model).

From Table 4.5 presented earlier, small cap stocks had the lowest mean absolute percentage error in Beta estimates so it was plausible that a portfolio formed on exogenous Beta would overweight small cap stocks relative to a portfolio formed on traditional Beta estimates. This does not appear to be the source of Alpha in the hedging portfolio. In Table 4.8 it is evident that there is no economic nor statistically significant influence of the small cap premium upon the outperformance of the hedging portfolio.

Ex ante, we know it is likely that there are a large number of generic "hedging" portfolios that are likely produce at least some positive Alpha when taking a short position in a high-Beta portfolio. Whilst the mechanism as to why is overlooked in finance, papers such as (Frazzini and Pedersen, 2014) have certainly exploited this measurement error to produce exceptional empirical results. A randomly formed portfolio should tend to generate an Alpha close to zero. Therefore, had our "hedging portfolio" taken a long position in a randomly formed portfolio (i.e. if our exogenous measure of Beta produced random Beta values) and a short position in a portfolio formed on traditional Beta estimates it would still be anticipated to produce a small positive Alpha (with the entirety of Alpha driven by the short position). Consequently, it is important to demonstrate that both the long and short positions of the hedging portfolio contribute towards Alpha in order to demonstrate that exogenous Betas are meaningful estimates of the true relation between security returns and the chosen market proxy. In Table 4.9 below we decompose the performance of the hedging portfolio into its two constituent portfolios.

	Formed on $\widehat{\beta}_i$	Formed on $\widetilde{\beta}_i$	Hedging Portfolio
Regression metrics			
$\hat{\alpha}_i$ (%)	-0.39 (0.8528)	3.84 (0.2106)	4.23* (0.0737)
$\hat{\beta}_{Mkt}$	1.3097*** (0.0000)	1.2291*** (0.0000)	-0.0806*** (0.0000)
Model R^2	0.8061	0.6303	0.0122
F-test: *p*-value	0.0000	0.0000	0.0000
Additional metrics			
Geometric return (% annual)	12.53%	15.96%	2.60%
Standard deviation (% annual)	26.01%	27.61%	13.02%
Sharpe Ratio	0.48	0.58	0.20

Table 4.9: Performance comparison of hedging portfolio constituents (1991-2020)

Table 4.9 shows the performance of a portfolio hedged against endogenous reference bias as well as its' two constituent portfolios across the 1991-2020 sample period. *,**,*** indicate statistical significance at the 90%, 95%, and 99% confidence level, respectively. Regression coefficients include *p*-values in parentheses under each estimate. Each regression is based on a single-factor model where the excess returns of the test portfolio are regressed against the excess returns of a cap-weighted market proxy. $\hat{\alpha}_i$ represents the annual % return of an endogenous reference bias hedging portfolio that can not be attributed to the risk factors included in each regression. $\hat{\beta}_{Mkt}$ is the estimated market risk exposure of the respective test portfolios under a traditional OLS approach (i.e. subject to endogeneity) The table shows that economically large alpha of the endogenous reference bias hedging portfolio is primarily driven by the long position in a Beta-weighted portfolio formed on exogenous Beta estimates rather than a short position in a Beta-weighted portfolio formed on traditional (endogenous) Beta estimates.

Consistent with expectations, we can observe that a long position in a Beta-weighted portfolio using traditional Beta estimation would deliver a negative Alpha when regressed against a single-factor model, albeit economically small and statistically insignificant. A short position in this Beta-weighted portfolio would enhance the alpha of a hedging portfolio but unless the long component also generated an economically large Alpha the hedging portfolio would not produce a statistically significant Alpha. In Table 4.9 we observe that a Beta-weighted portfolio formed on exogenous Beta estimates produces an economically large Alpha and is the primary contributor to the risk-adjusted outperformance of the hedging portfolio. The table also indicates that this hedging portfolio produces positive absolute returns during the 30-year sample period.

4.6 Conclusion

This study focused on a prevalent source of endogeneity in finance that has scarce coverage in current asset pricing literature; endogenous reference bias. Our primary motivation was to examine the severity of endogenous reference bias and provide practical guidance on when the bias is likely to undermine empirical analysis. We sought to identify whether the current disregard of endogenous reference bias in asset pricing literature was “a problem”. Our finding was “it depends”.

As explored in Section 4.3 of this chapter there are multiple competing influences on the direction and magnitude of endogenous reference bias. Where existing research has taken the stance that endogenous reference bias monotonically increases with constituent weight (e.g. Malloch et al., 2016), we found this not to be the case. We find that the exogenous Beta of a security is an important determinant of the magnitude and direction of endogenous reference bias impacting traditional Beta estimates. Even for exogenous Betas centred on 1, we observe that endogenous reference bias will initially increase before hitting an inflection point and decreasing as constituent weight increases.

In Section 4.4 we examined the magnitude and direction of endogenous reference bias for 30 years of historical US stock data. We found that the typical magnitude of bias when a large, diversified market proxy is used is economically insignificant. Via use of several fixed effect models we compared the key influences of endogenous reference bias in historical data. We found that constituent weight alone is a minor determinant of endogenous reference bias. By contrast, we identified that the inclusion of variables that account for an asset’s variance multiple and exogenous Beta drastically improve a model’s explanatory power.

In Section 4.5 we considered the influence of endogenous reference bias upon performance evaluation where the dependent variable was a portfolio rather than individual stock. We demonstrated that the bias was likely to be small for a generic, well-diversified portfolio since the variance multiple quickly approaches a value of 1, subduing the bias from increasing constituent weights. However, portfolios which systematically exploit stock characteristics, such as Betas, size, low volatility, etc. can be severely impacted by endogenous reference bias. A simple hedging portfolio against endogenous reference bias is capable of generating a single-factor model Alpha of ~4.30% during our 30-year sample period; an Alpha economically larger than both the small cap and value premiums. The size of this Alpha was primarily driven by a long position in a Beta-weighted portfolio formed on exogenous Beta estimates rather than a short position in a Beta-weighted portfolio formed on traditional (endogenous) Beta estimates.

In summary, we find that endogenous reference bias is *not* likely to significantly bias empirical analysis of *individual stocks* for the US stock market where it is common practice to use large market proxies that are not highly concentrated. Whether a researcher is estimating the Beta for an individual security, or for a larger portfolio, the magnitude of Beta bias is likely to be small. An exception arises for portfolios that are comprised of particularly high or low Beta securities. Traditional Beta estimates for these portfolios are likely to encounter significant effects from endogenous reference bias with the severity of bias increasing for more extreme Beta ranges. There are two key implications that arise from this finding. Firstly, portfolio grouping procedures may not be an appropriate means to address the errors-in-variables problem since Beta-grouping worsens endogenous reference bias. Secondly, the empirical performance of the CAPM may be better than previously contemplated if an exogenous reference portfolio is used.

Chapter 5

Conclusion

5.1 Summary & implications

Understanding the cross-section of asset returns is a longstanding strand of literature in finance. Factor return models have served a central role in this academic pursuit. Within the traditional context in which factor models have been applied to the study of financial markets, we have found an increasing number of "surprises". The body of literature on return "anomalies" has grown at an exponential rate in recent decades. Findings that should be "anomalous" have become "commonplace".

The ease with which anomaly literature has grown suggests deficiencies in either the underlying financial theory, the econometric approach used by researchers, or some combination of both. This thesis set out to identify undocumented econometric inconsistencies that could plausibly be the cause. Three distinct issues affecting parameter estimation in common applications of factor models have been identified. Each issue was covered in a separate chapter of this thesis.

This chapter summarizes those conclusions from Chapters 2 through 4, including the following: (i) the common usage of a non-tangency portfolio as market proxy in single-factor model regressions induces Beta estimation bias severe enough to invalidate inferences made by the model; (ii) Alphas estimated via time-series regression suffer from economically and statistically significant bias caused by a misspecification of cross-sectional risk premia; and (iii) the inclusion of test assets within a market proxy induces non-linear Beta estimation bias with severity influenced by constituent weight, variance multiple and exogenous Beta.

5.1.1 Is Beta Busted?

Chapter 2 identified an estimation bias introduced to single-factor models when using a non-tangency market proxy. In Section 2.3, we developed a novel measure for quantifying the severity of estimation error that arises; Beta Mismatch Error (BME). BME, indicates the difference between the value of Betas as estimated via OLS vs. the values of Beta as interpreted under CAPM theory. Our novel measure allowed us to examine the conditions in which the mismatch between theory and estimation creates the most significant divergence between theory and measurement. BME was found to become increasingly severe as securities exhibit progressively lower correlations with the chosen market proxy.

In Section 2.4, we examined the historical magnitude of this estimation error arising from the mismatch between theory and OLS estimation for individual US equities. The average correlation of securities with the market proxy averaged around 0.5 across the sample period. Consequently, for a median stock, a Beta estimate consistent with CAPM theory would be approximately 2.07x the size estimated under OLS; an economically huge magnitude of bias.

In Section 2.5, we demonstrated that formation of portfolios can reduce the severity of BME by increasing correlation with the market proxy. We also identified that both the size and value premium are at least partially attributable to mismatch error. In each case, these long-short factor portfolios take a long position in a portfolio that exhibits an abnormally low correlation with the market proxy. As a result, these long position portfolios have inflated Alpha estimates in time-series regressions of factor models.

The verdict of the chapter was that "Beta is busted". An empirical measure of Beta which is both consistent with CAPM theory and capable of ranking the risk-adjusted performance of assets is mathematically impossible; a fact previously noted by Roll (1978).

5.1.2 Is Alpha reliable in practice?

Chapter 3 questioned the reliability of Alpha estimates obtained via time-series regression. Time-series regressions impose an arbitrary market risk premium assumption that is not informed by the underlying data; Alpha estimates become a linear function of Beta estimates. This model misspecification is likely the genesis of a large number of published finance anomalies. For example, even in simulated random returns data we observe the low Beta anomaly. The low Beta anomaly is not a representation of some real world anomalous investment practice. Rather, it is generated by the arbitrary imposition of a CAPM relation upon data by time-series regressions. Examined returns data will seldom conform to the CAPM; creating a mismatch between the return structure assumed by time-series regression vs. the return structure observed. As a result of model misspecification, time-series regression Alpha estimates become a linear function of Beta estimates. These Alpha estimates cannot be a reliable measure of risk-adjusted outperformance since they are themselves a linear function of systematic risk![48]

In Section 3.3 we decomposed Jensen's Alpha into an endogenous component, which is a linear function of systematic risk, and an exogenous component ("Excess Alpha") which is a measure of return in excess of systematic risk exposure. We asserted that this Excess Alpha is a true measure of risk-adjusted outperformance that accounts for the cross-sectional relation between returns and systematic risks in observed data. Excess Alpha was demonstrated to have highly desirable statistical properties; it is orthogonal to systematic risks in multi-factor models and the average Excess Alpha in every period is guaranteed to be zero (it is impossible for the entire market to outperform/underperform). We also demonstrated that this exogenous measure of alpha can be directly obtained via the residuals of cross-sectional regressions.

We quantified the magnitude of model misspecification bias for historical US data. Since we observed a much weaker cross-sectional relation in our 30-year data sample between traditional Beta estimates and security returns, time-series regressions suffered from severe model misspecification. The market risk-premium imposed by a time-series regression of a single-factor model was more than 7% higher than reflected by the cross-sectional relation observed in the data sample. Consequently, low Beta securities had severely inflated Alpha estimates whilst high Beta securities had severely understated Alpha estimates, generating the perception of a large "low Beta anomaly" in the data sample.

In the final section of the chapter, we demonstrated the ease with which long-short portfolios can be constructed to exploit model misspecification bias and generate practically infinite spurious return anomalies. Our conclusion was that time-series regression estimates of security Alphas were "not reliable in practice". Consequently, we recommended empiricists and practitioners transition to use of "Excess Alpha" as a more robust measure of risk-adjusted outperformance.

5.1.3 Endogenous reference bias: when is it a problem?

Chapter 4 built on an existing but understudied area of research. Both single and multiple-factor models are commonly subject to a specific form of endogeneity that is

---

[48] E.g., under time-series regression of a single-factor model: $\hat{\alpha}_i = \bar{R}_i - \hat{\beta}_i * \bar{R}_M$.

not considered by empiricists in asset pricing literature. This endogeneity arises from the practice of regressing dependent variables against "independent" variables of which they are a constituent. The impact is biased parameter estimation. In Chapter 4 we sought to determine, in practice, whether it is a substantial source of bias likely to invalidate the inferences made by factor models.

Since this understudied form of bias had never explicitly been named, we coined the term "endogenous reference bias". We presented the first research to consider the impact of three competing influences on the direction and magnitude of endogenous reference bias; constituent weight in a regressor, exogenous Beta, and variance multiple. This allowed us to identify asset characteristics that are more susceptible to severe parameter estimation bias from this form of endogeneity.

We provided the first study of the severity of endogenous reference bias in historical US stock data and found that Beta estimation bias is unlikely to be severe for *most* individual securities. In a conservative setting, using a large, well-diversified market proxy such as the S&P 500, a traditional Beta estimation approach would produce Beta estimates with a cap-weighted average absolute bias of 0.0205. In economic terms, traditional OLS beta estimates were quantitatively similar to Beta estimates that had been corrected for endogenous reference bias.

By contrast, we identified that parameter estimates for large portfolios formed on specific asset characteristics, such as exogenous Beta, can be severely impacted by endogenous reference bias. An implication is that Beta-grouping, historically a common strategy to address the errors-in-variables problem in finance (e.g. see Fama and French, 2004, pp. 33), can induce severe endogenous reference bias.

We also presented the first paper to develop a simple hedging portfolio against endogenous reference bias. This hedging portfolio generates an economically large Alpha of around 4.30% across our 30-year sample period. The outperformance of the hedging portfolio demonstrated the in-sample superiority of exogenous Betas as a more accurate representation of security return movements in relation to a market proxy.

5.1.4 Concluding remarks

Each chapter presented in this thesis identified distinct econometric issues with the current employ of common factor models. In culmination, this thesis serves to raise serious concern about the Alpha and Beta estimates produced by factor models. The implication is that potentially thousands of published finance papers are suffering from material bias that could invalidate the inferences produced and conclusions drawn.

In particular, the rapid growth in asset pricing anomaly literature is an excellent example of research that is extremely vulnerable to misuse of factor models. It is usually the case that historical stock data does not perfectly conform to the CAPM. As demonstrated in Section 3.6, this makes it trivial to construct a practically infinite number of portfolios that will falsely appear to generate Alpha under a time-series regression since time-series regressions will suffer from model misspecification of cross-sectional risk premia.

Even two of the most famous anomalies in finance, the small cap and value premiums, have historically benefited from Alpha inflation due to separate econometric issues. For example, in Section 2.5.2 we observed that the Alphas produced by the SMB and HML portfolios are driven by abnormally low correlations of the small cap and value portfolios with the market proxy. This low correlation produces understated OLS Beta estimates resulting in overstated Alphas. In Chapter 4 we observed that large-cap stocks tend to have overstated Betas and, by consequence, will have understated Alphas.

Emerging from this thesis, there are two potential paths for the future use of factor models in empirical finance. The first, and most convenient, option would be for finance practitioners to transition to forsaking time-series regression models in favour of cross-sectional regression models when it comes to evaluating historical asset performance. In this context, risk-adjusted outperformance becomes the cross-sectional regression model's residuals (analogous to Excess Alpha). This would overcome the issue of model misspecification and ensure that asset returns are penalised based on actual rather than presumed cross-sectional risk premia in examined data. This transition comes with the uncomfortable abandonment of the CAPM; one of the most influential models in finance history. The CAPM prescribes the exact slope of the security market line. Under this new paradigm, unguided, we would instead identify risk premia based on observed relations in historical data. We would have transitioned into a field of anti-theoretical data analysts.

An alternate option, though far more challenging, would be to go back to the drawing board and redevelop new finance theory that can more closely explain asset returns. If the CAPM accurately described asset returns, then the econometric issues identified in Chapter 2 and Chapter 3 would cease to be significant. Similarly, any replacement of the CAPM which could closely prescribe the means by which asset returns fluctuate would therefore significantly moderate most of the econometric issues raised within this thesis. Given that the majority of academics perceive asset returns to be *ex ante* unpredictable, it is questionable whether such an accurate theoretical model could ever be developed.

5.2 Future research

The most pressing future research emerging from this thesis would be a comprehensive replication of published asset pricing papers that claim to identify anomalies. As an initial test, anomaly replications should verify whether the claimed anomaly can produce a statistically significant[49] positive residual within a cross-sectional regression of a single-factor model. Within this framework, the test assets would be the anomaly portfolio as well as all individual securities such that the cross-sectional risk premia of the chosen factor model accurately reflects the underlying data. Anomalies capable of consistently producing statistically significant positive Alphas under this setting are more likely to be genuine return anomalies. Furthermore, this approach to anomaly testing is robust to the prevalent use of high-minus-low portfolio formations that exploit the model misspecification issue of time-series regressions.

Throughout this thesis, the three identified econometric issues were studied in isolation. Future research could study the interaction between the competing sources of bias. For example, in Chapter 3 we focus on model misspecification emerging from reliance on time-series regression with cross-sectionally misspecified risk premia. The process of obtaining exogenous Betas that was introduced in Chapter 4 would provide a different set of Beta estimates used to obtain the exogenous Alpha estimates of Chapter 3. It's possible this adjustment may reduce the time-series model misspecification issue encountered in Chapter 4 since Figure 4.4 suggests use of exogenous Betas would produce a steeper security market line. Similarly, use of exogenous Betas would likely effect the results in Chapter 2. Exogenous Betas will exhibit less correlation with a market proxy thereby increasing the severity of the Beta Mismatch Error presented in that chapter.

[49] The statistical significance of individual residuals can be tested via a t-test.

References

Andrei, D., Culean, J., & Fournier, M. (2019). The Low-Minus-High Portfolio and the Factor Zoo. *Working Paper.*

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

Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratio: a test of the efficient market hypothesis. *Journal of Finance, 32*(3):663-682.

Black, F. (1972). Capital Market Equilibrium with Restricted Borrowing. *Journal of Business, 45*:444–455.

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

Blitz, D., Hanauer, M., Vidojevic, M., & van Vliet, P. (2018). Five Concerns with the Five-Factor Model. *The Journal of Portfolio Management, 44*(4):71–78.

Bos, T., & Newbold, P. (1984). An empirical investigation of the possibility of stochastic systematic risk in the market model. *Journal of Business*, 35-41.

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

CFA Institute. (2021) *2022 CFA Program Curriculum Level I Box Set*. United States: Wiley.

Chordia, T., Roll, R., & Subrahmanyam, A. (2000). Commonality in liquidity. *Journal of Financial Economics, 56*:3-28

Cochrane, J. (2011). Presidential address: Discount rates. *Journal of Finance, 66*:1047–1108.

Corporate Finance Institute. (2022). 'Beta'. https://corporatefinanceinstitute.com/resources/valuation/what-is-beta-guide/. [Online; accessed 12-November-2022].

Collins, D. W., Ledolter, J., & Rayburn, J. (1987). Some further evidence on the stochastic properties of systematic risk. *Journal of Business*, 425-448.

Damodaran, A. (1999). Estimating Risk Parameters. *Working paper*.

Fabozzi, F., & Francis, J. (1978). Beta as a Random Coefficient. *Journal of Financial and Quantitative Analysis*, 13(1):101-116.

Fama, E. (1972). Components of Investment Performance. *Journal of Finance, 27*(2):551–567.

Fama, E., & Kenneth, F. (1992). The cross-section of expected stock returns. *Journal of Finance, 47*(2):427–465.

Fama, E., & Kenneth, F. (1993). Common risk factors in the returns on bonds and stocks. *Journal of Financial Economics*, *33*:3–53.

Fama, E., & Kenneth, F. (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance, 51*:55–84.

Fama, E., & Kenneth, F. (2004). The capital asset pricing model: Theory and evidence. *Journal of Economic Perspectives, 18*:25–46.

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

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

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

Frazzini, A., Kabiller, D., & Pedersen, L. (2018). Buffett's Alpha. *Financial Analysts* Journal, 74(4):35-55

Friend, I., & Blume, M. (1970). Measurement of Portfolio Performance under Uncertainty. *American Economic Review, 60*:561–575.

Graham, J., & Harvey, C. (2001). The theory and practice of corporate finance: Evidence from the field. *Journal of Financial Economics,* 60(2-3):187-243.

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

Hou, K., Xue, C., & Zhang, L. (2020). Replicating anomalies. *The Review of Financial Studies,* 33(5):2019-2133.

Jegadeesh, N., Noh, J., Pukthuanthong, K., Roll, R., & Wang, J. (2019). Empirical tests of asset pricing models with individual assets: Resolving the errors-in-variables bias in risk premium estimation. *Journal of Financial Economics*, 133(2):273-298.

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

Kim, D. (1995). The errors in the variables problem in the cross-section of expected stock returns. *The Journal of Finance*, 50(5):1605-1634.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics, 47*:13–37.

Liu, J., Stambaugh, R., & Yuan, Y. (2016). Absolving beta of volatility's effects. *Working Paper*.

López de Prado, M. (2023). Where Are the Factors in Factor Investing? *The Journal of Portfolio Management, 49*(5):6-20.

Malloch, H., Philip, R., & Satchell, S. (2016). Decomposing the bias in time-series estimates of CAPM betas. *Applied Economics, 48*(45):4291-4298.

Markowitz, H. (1952). Modern Portfolio Theory. *Journal of Finance*, *7*(11):77-91.

Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, *35*:768-83.

Patton, A., & Verardo, M. (2012). Does Beta Move with News? Firm-Specific Information Flows and Learning about Profitability. *Review of Financial Studies 25*(9):2789–2839.

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. (1978). Ambiguity when Performance is Measured by the Securities Market Line. *Journal of Finance, 33*(4):1051–1069.

Roll, R., & Ross, S. A. (1994). On the cross-sectional relation between expected returns and betas. *The journal of finance*, *49*(1):101-121.

Rosch, D., Subrahmanyam, A., & Van Dijk, M. (2014). An Empirical Analysis of Co-Movements in High and Low-Frequency Metrics for Financial Market Efficiency. *Working Paper*.

Sharpe, W. (1964). Capital Asset Prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*:425–42.

Sinclair, N. (1987). Multifactor Asset Pricing Models. *Accounting and Finance, 27*(1):17–36.

Smith, T. & Walsh, K. (2013). Why the CAPM is Half-Right and Everything Else is Wrong. *Abacus, 49*:73–78.

Sunder, S. (1980). Stationarity of Market Risk: Random Coefficients Tests for Individual Stocks. *The Journal of Finance*, *35*(4):883–896.

Tofallis, C. (2008). Investment volatility: A critique of standard beta estimation and a simple way forward. European Journal of Operational Research, *187*(3), 1358-1367.

Treynor, J. (1965). How to Rate the Performance of Mutual Funds. *Harvard Business Review*, *43*:63-75.

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

Welch, I. (2019). Simpler better market betas. *Working Paper.*

Wikipedia. (2022). 'Beta (finance)'. https://en.wikipedia.org/wiki/Beta_(finance). [Online; accessed 12-November-2022].

Woo, C., Cheung, Y., & Richard, Y. (1994). Endogeneity bias in beta estimation: Thailand and Hong Kong. *Pacific-Basin Finance Journal 2*:453-461.

Appendix

Appendix A.1: Theory vs OLS Betas for cap-weighted proxy

Year	$\hat{\lambda}_0$	p-value ($\hat{\lambda}_0$)	$\hat{\lambda}_1$	p-value ($\hat{\lambda}_1$)	Model R^2
1991	0.10	0.48	1.13	0.00	0.11
1992	1.56	0.00	0.09	0.83	0.00
1993	1.80	0.00	-0.32	0.29	0.00
1994	-3.14	0.03	2.25	0.10	0.01
1995	0.58	0.00	0.27	0.00	0.04
1996	0.35	0.02	0.42	0.00	0.02
1997	0.51	0.00	0.32	0.03	0.01
1998	-0.52	0.01	1.14	0.00	0.08
1999	-2.13	0.00	4.28	0.00	0.16
2000	-1.21	0.00	1.61	0.00	0.18
2001	-0.96	0.00	1.46	0.00	0.27
2002	-0.76	0.00	1.59	0.00	0.41
2003	0.16	0.25	1.19	0.00	0.15
2004	2.97	0.00	-1.32	0.00	0.04
2005	-2.37	0.00	4.06	0.00	0.06
2006	0.98	0.00	-0.05	0.73	0.00
2007	-1.46	0.00	2.24	0.00	0.05
2008	0.02	0.76	1.07	0.00	0.33
2009	0.13	0.39	1.09	0.00	0.14
2010	0.18	0.44	1.08	0.00	0.06
2011	7.11	0.00	-5.92	0.00	0.22
2012	1.16	0.00	-0.06	0.77	0.00
2013	0.05	0.79	0.99	0.00	0.08
2014	1.97	0.00	-0.92	0.00	0.03
2015	12.91	0.00	-14.32	0.00	0.04
2016	0.87	0.00	0.31	0.21	0.00
2017	0.24	0.08	0.65	0.00	0.06
2018	1.37	0.22	0.96	0.40	0.00
2019	0.72	0.00	0.23	0.02	0.01
2020	1.92	0.00	-0.99	0.00	0.02
Average	**0.84**	**0.12**	**0.15**	**0.11**	**0.09**
Median	**0.30**	**0.00**	**0.80**	**0.00**	**0.05**

Table A.1: Results of regression of CAPM theory Betas against OLS estimated Betas when using a cap-weighted market proxy

Table A.1 shows the annual results of the regression: $\beta_i^{Theory} = \hat{\lambda}_0 + \hat{\lambda}_1\hat{\beta}_i + \hat{\varepsilon}_i$ where $\beta_i^{Theory} = \frac{E[R_i]}{E[R_M]}$, $\hat{\lambda}_0$ is an estimated regression constant, $\hat{\lambda}_1$ is an estimated slope coefficient, $\hat{\beta}_i$ are OLS estimates of Beta, and $\hat{\varepsilon}_i$ are estimated regression residuals. If OLS estimates of Beta are equivalent with their theoretical values then the regression performed in each year would produce $\hat{\lambda}_0 = 0$, $\hat{\lambda}_1 = 1$, and a regression $R^2 = 1$. The results presented in the table indicate a severe historical disconnect between theoretical Betas and their OLS estimates when using a cap-weighted index as a market proxy.

Appendix A.2: Theory vs OLS Betas for tangency portfolio

Year	$\hat{\lambda}_0$	p-value ($\hat{\lambda}_0$)	$\hat{\lambda}_1$	p-value ($\hat{\lambda}_1$)	Model R^2
1991	0.00	0.00	1.00	0.00	1.00
1992	0.00	0.00	1.00	0.00	1.00
1993	0.00	0.00	1.00	0.00	1.00
1994	0.00	0.00	1.00	0.00	1.00
1995	0.00	0.00	1.00	0.00	1.00
1996	0.00	0.00	1.00	0.00	1.00
1997	0.00	0.00	1.00	0.00	1.00
1998	0.00	0.00	1.00	0.00	1.00
1999	0.00	0.00	1.00	0.00	1.00
2000	0.00	0.00	1.00	0.00	1.00
2001	0.00	0.00	1.00	0.00	1.00
2002	0.00	0.00	1.00	0.00	1.00
2003	0.00	0.00	1.00	0.00	1.00
2004	0.00	0.00	1.00	0.00	1.00
2005	0.00	0.00	1.00	0.00	1.00
2006	0.00	0.00	1.00	0.00	1.00
2007	0.00	0.00	1.00	0.00	1.00
2008	0.00	0.00	1.00	0.00	1.00
2009	0.00	0.00	1.00	0.00	1.00
2010	0.00	0.00	1.00	0.00	1.00
2011	0.00	0.00	1.00	0.00	1.00
2012	0.00	0.00	1.00	0.00	1.00
2013	0.00	0.00	1.00	0.00	1.00
2014	0.00	0.00	1.00	0.00	1.00
2015	0.00	0.00	1.00	0.00	1.00
2016	0.00	0.00	1.00	0.00	1.00
2017	0.00	0.00	1.00	0.00	1.00
2018	0.00	0.00	1.00	0.00	1.00
2019	0.00	0.00	1.00	0.00	1.00
2020	0.00	0.00	1.00	0.00	1.00
Average	**0.00**	**0.00**	**1.00**	**0.00**	**1.00**
Median	**0.00**	**0.00**	**1.00**	**0.00**	**1.00**

Table A.2: Results of regression of CAPM theory Betas against OLS estimated Betas when using the tangency portfolio as market proxy

Table A.2 shows the annual results of the regression: $\beta_i^{Theory} = \hat{\lambda}_0 + \hat{\lambda}_1\hat{\beta}_i + \hat{\varepsilon}_i$ where $\beta_i^{Theory} = \frac{E[R_i]}{E[R_M]}$, $\hat{\lambda}_0$ is an estimated regression constant, $\hat{\lambda}_1$ is an estimated slope coefficient, $\hat{\beta}_i$ are OLS estimates of Beta, and $\hat{\varepsilon}_i$ are estimated regression residuals. If OLS estimates of Beta are equivalent with their theoretical values then the regression performed in each year would produce $\hat{\lambda}_0 = 0$, $\hat{\lambda}_1 = 1$, and a regression $R^2 = 1$. The results presented in the table indicate a perfect relation between both measures when using the tangency portfolio as a market proxy.

Appendix C.1: Most biased Betas by year

	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Most biased	INTERNATIONAL BUSINESS MACHS COR	INTERNATIONAL BUSINESS MACHS COR	PHILIP MORRIS COS INC	WAL MART STORES INC	INTEL CORP	PHILIP MORRIS COS INC	INTEL CORP	INTEL CORP	MICROSOFT CORP	MICROSOFT CORP	PFIZER INC	PFIZER INC	PHILIP MORRIS COS INC	MERCK & CO INC	EXXON CORP
Bias	**0.0590**	**0.1207**	**0.2408**	**0.1039**	**0.1837**	**0.1088**	**0.0547**	**0.0478**	**0.1003**	**0.1130**	**0.0523**	**0.0379**	**0.0402**	**0.1068**	**0.1067**
2nd most	EXXON CORP	EXXON CORP	WAL MART STORES INC	INTERNATIONAL BUSINESS MACHS COR	MICROSOFT CORP	INTERNATIONAL BUSINESS MACHS COR	PHILIP MORRIS COS INC	MICROSOFT CORP	AMERICA ONLINE INC DEL	INTEL CORP	MICROSOFT CORP	AMERICA ONLINE INC DEL	MICROSOFT CORP	PFIZER INC	GENENTECH INC
Bias	**0.0481**	**0.1000**	**0.2026**	**0.0910**	**0.1792**	**0.1005**	**0.0530**	**0.0458**	**0.0904**	**0.1119**	**0.0499**	**0.0368**	**0.0374**	**0.1013**	**0.1041**
3rd most	PHILIP MORRIS COS INC	GENERAL MOTORS CORP	MERCK & CO INC	GENERAL ELECTRIC CO	AMERICAN TELEPHONE & TELEG CO	INTEL CORP	MICROSOFT CORP	EXXON CORP	INTERNATIONAL BUSINESS MACHS COR	LUCENT TECHNOLOGIES INC	GENERAL ELECTRIC CO	WAL MART STORES INC	GENENTECH INC	INTEL CORP	PFIZER INC
Bias	**0.0341**	**0.0929**	**0.1658**	**0.0745**	**0.1617**	**0.0996**	**0.0527**	**0.0446**	**0.0882**	**0.0965**	**0.0479**	**0.0347**	**0.0341**	**0.0787**	**0.0874**
4th most	GENERAL MOTORS CORP	MERCK & CO INC	INTEL CORP	EXXON CORP	HEWLETT PACKARD CO	HEWLETT PACKARD CO	ORACLE SYSTEMS CORP	C U C INTERNATIONAL INC	INTEL CORP	WAL MART STORES INC	WAL MART STORES INC	PHILIP MORRIS COS INC	PFIZER INC	CISCO SYSTEMS INC	EBAY INC
Bias	**0.0289**	**0.0881**	**0.1267**	**0.0637**	**0.1490**	**0.0968**	**0.0379**	**0.0423**	**0.0753**	**0.0894**	**0.0427**	**0.0307**	**0.0334**	**0.0658**	**0.0856**
5th most	INTEL CORP	PHILIP MORRIS COS INC	AMERICAN TELEPHONE & TELEG CO	GENERAL MOTORS CORP	MOTOROLA INC	AMERICAN TELEPHONE & TELEG CO	AMERICAN TELEPHONE & TELEG CO	PFIZER INC	ORACLE SYSTEMS CORP	GENERAL ELECTRIC CO	EXXON CORP	JOHNSON & JOHNSON	EXXON CORP	GENENTECH INC	PHILLIPS PETROLEUM CO
Bias	**0.0286**	**0.0836**	**0.1176**	**0.0626**	**0.1399**	**0.0839**	**0.0379**	**0.0421**	**0.0682**	**0.0821**	**0.0406**	**0.0294**	**0.0302**	**0.0577**	**0.0514**

Table C.1, Part A: Top 5 most biased Betas (1991-2005)

Table C.1 presents the absolute value of Beta estimation bias due to endogeneity for the top 5 most affected securities in each year. Beta estimation bias is in relation to a cap-weighted reference portfolio comprised of the top 500 largest stocks.

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Most biased	EXXON CORP	APPLE COMPUTER INC	GENERAL ELECTRIC CO	MICROSOFT CORP	APPLE COMPUTER INC	APPLE COMPUTER INC	APPLE COMPUTER INC	APPLE COMPUTER INC	APPLE COMPUTER INC	APPLE COMPUTER INC	APPLE COMPUTER INC	APPLE COMPUTER INC	FACEBOOK INC	FACEBOOK INC	AMAZON COM INC
Bias	**0.1109**	**0.0425**	**0.0281**	**0.0371**	**0.0269**	**0.0305**	**0.1843**	**0.2467**	**0.1121**	**0.0709**	**0.0816**	**0.2073**	**0.0614**	**0.0545**	**0.0510**
2nd most	MICROSOFT CORP	MICROSOFT CORP	WAL MART STORES INC	GENERAL ELECTRIC CO	CISCO SYSTEMS INC	AMAZON COM INC	AMAZON COM INC	MICROSOFT CORP	GILEAD SCIENCES INC	AMAZON COM INC	AMAZON COM INC	AMAZON COM INC	APPLE COMPUTER INC	APPLE COMPUTER INC	TESLA MOTORS INC
Bias	**0.0986**	**0.0359**	**0.0157**	**0.0370**	**0.0265**	**0.0171**	**0.0420**	**0.1004**	**0.0583**	**0.0404**	**0.0739**	**0.1835**	**0.0449**	**0.0528**	**0.0429**
3rd most	PFIZER INC	CISCO SYSTEMS INC	APPLE COMPUTER INC	WAL MART STORES INC	MICROSOFT CORP	EXXON CORP	GOOGLE INC	FACEBOOK INC	AMAZON COM INC	MICROSOFT CORP	FACEBOOK INC	NVIDIA CORP	AMAZON COM INC	AMAZON COM INC	APPLE COMPUTER INC
Bias	**0.0734**	**0.0313**	**0.0149**	**0.0278**	**0.0214**	**0.0134**	**0.0416**	**0.0711**	**0.0523**	**0.0359**	**0.0535**	**0.1170**	**0.0261**	**0.0442**	**0.0264**
4th most	CISCO SYSTEMS INC	EXXON CORP	MICROSOFT CORP	EXXON CORP	GOOGLE INC	GOOGLE INC	WAL MART STORES INC	GOOGLE INC	TWITTER INC	GOOGLE INC	EXXON CORP	WAL MART STORES INC	GENERAL ELECTRIC CO	BOEING CO	FACEBOOK INC
Bias	**0.0694**	**0.0269**	**0.0133**	**0.0261**	**0.0210**	**0.0128**	**0.0372**	**0.0565**	**0.0450**	**0.0300**	**0.0380**	**0.1033**	**0.0244**	**0.0411**	**0.0166**
5th most	GOOGLE INC	GENERAL ELECTRIC CO	EXXON CORP	PFIZER INC	QUALCOMM INC	HEWLETT PACKARD CO	MICROSOFT CORP	INTERNATIONAL BUSINESS MACHS COR	FACEBOOK INC	I D E C PHARMACEUTICALS CORP	BRISTOL MYERS SQUIBB CO	GENERAL ELECTRIC CO	JOHNSON & JOHNSON	PACIFIC GAS & ELEC CO	WAL MART STORES INC
Bias	**0.0666**	**0.0240**	**0.0128**	**0.0220**	**0.0181**	**0.0125**	**0.0313**	**0.0450**	**0.0449**	**0.0292**	**0.0320**	**0.1002**	**0.0235**	**0.0348**	**0.0134**

Table C.1, Part B: Top 5 most biased Betas (2006-2020)

Table C.1 presents the absolute value of Beta estimation bias due to endogeneity for the top 5 most affected securities in each year. Beta estimation bias is in relation to a cap-weighted reference portfolio comprised of the top 500 largest stocks.

Appendix C.2: Comparison of fixed effect models

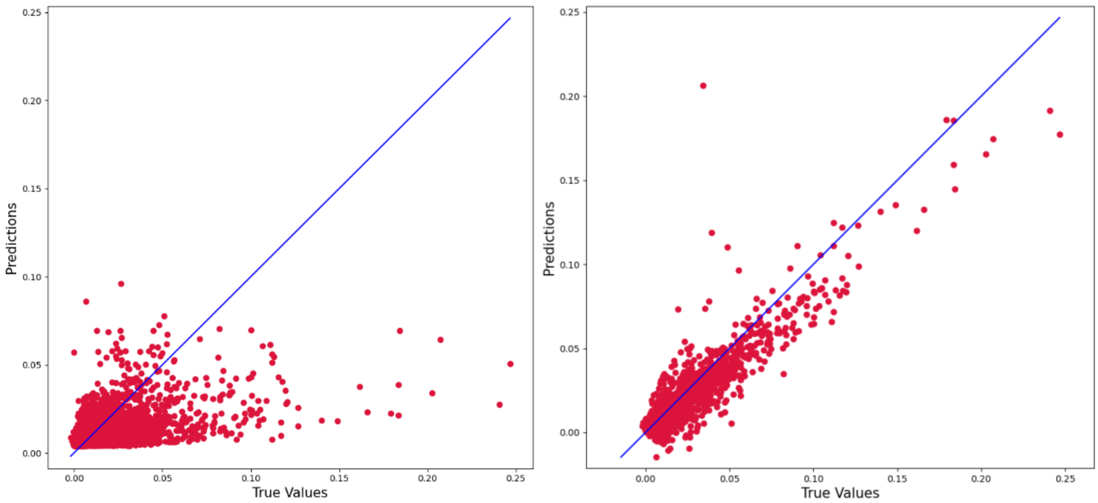


Figure C.2: Comparison of fixed effect models

Figure C.2 compares the prediction accuracy of FE Model 1 (LHS) with FE Model 2 (RHS). The inclusion of interaction terms for variance multiple and exogenous Beta significantly enhance prediction accuracy. Weight is not the sole determinant of endogenous reference bias.

Appendix C.3: Derivation of biased OLS Beta estimates

We have a starting model:

$$R_{j,t} = \beta_j R_{M,t} + \epsilon_{j,t}$$

where j denotes an asset; t denotes a time increment; $R_{j,t}$ and $R_{M,t}$ are returns on asset j and the market in excess of the risk-free rate; β_j is the exposure of asset j to market risk; and $\epsilon_{j,t}$ is an error term. We can decompose the market return into two components: the weighted return on asset j, and the sum of weighted return on the rest of the assets in the market:

$$R_{M,t} = w_j\, R_{j,t} + \left(1 - w_j\right)\tilde{R}_{M,t},$$

where w_j represents the weight of asset j in the market portfolio, and $\tilde{R}_{M,t} = \frac{1}{(1-w_j)} \sum_{k=1,k\neq j}^{K} w_k\, R_{k,t}$. Our OLS estimates of beta is derived as:

$$\hat{\beta}_j = \frac{cov\left(R_{j,t}, R_{M,t}\right)}{var\left(R_{M,t}\right)}$$

substituting in the decomposed value of $R_{M,t}$*:*

$$\hat{\beta}_j = \frac{cov\left(R_{j,t},\ w_j\, R_{j,t} + (1 - w_j)\tilde{R}_{M,t}\right)}{var\left(w_j\, R_{j,t} + (1 - w_j)\tilde{R}_{M,t}\right)}$$

expanding:

$$\hat{\beta}_j = \frac{w_j * cov\left(R_{j,t},\ R_{j,t}\right) + \left(1 - w_j\right) * cov\left(R_{j,t},\ \tilde{R}_{M,t}\right)}{\left(w_j\right)^2 * var\left(R_{j,t}\right) + \left(1 - w_j\right)^2 * var\left(\tilde{R}_{M,t}\right) + 2 * w_j * \left(1 - w_j\right) * cov\left(R_{j,t},\ \tilde{R}_{M,t}\right)}$$

dividing the numerator and denominator by $cov\left(R_{j,t},\ \tilde{R}_{M,t}\right)$*:*

<u>numerator</u>:

$$\frac{w_j * var(R_{j,t}) + (1 - w_j) * cov(R_{j,t}, \ \widetilde{R}_{M,t})}{cov(R_{j,t}, \ \widetilde{R}_{M,t})}$$

$$= w_j * \frac{var(R_{j,t})}{cov(R_{j,t}, \ \widetilde{R}_{M,t})} + (1 - w_j)$$

<u>*denominator:*</u>

$$\frac{(w_j)^2 * var(R_{j,t}) + (1 - w_j)^2 * var(\widetilde{R}_{M,t}) + 2 * w_j * (1 - w_j) * cov(R_{j,t}, \ \widetilde{R}_{M,t})}{cov(R_{j,t}, \ \widetilde{R}_{M,t})}$$

$$= (w_j)^2 * \frac{var(R_{j,t})}{cov(R_{j,t}, \ \widetilde{R}_{M,t})} + (1 - w_j)^2 * \frac{1}{\widetilde{\beta}_j} + 2 * w_j * (1 - w_j)$$

where $\tilde{\beta}_j = \frac{cov[R_{j,t}, \ \tilde{R}_{M,t}]}{var[\tilde{R}_{M,t}]}$ *(an exogenous measure of the covariance of the excess returns of asset j with the rest of the market)*

numerator and denominator recombined:

$$\hat{\beta}_j = \frac{w_j * \frac{var(R_{j,t})}{cov(R_{j,t}, \ \widetilde{R}_{M,t})} + (1 - w_j)}{(w_j)^2 * \frac{var(R_{j,t})}{cov(R_{j,t}, \ \widetilde{R}_{M,t})} + (1 - w_j)^2 * \frac{1}{\widetilde{\beta}_j} + 2 * w_j * (1 - w_j)}$$

multiply numerator and denominator by $\tilde{\beta}_j$*:*

$$\hat{\beta}_j = \frac{w_j * \frac{var(R_{j,t})}{var(\widetilde{R}_{M,t})} + (1 - w_j)\widetilde{\beta}_j}{(w_j)^2 * \frac{var(R_{j,t})}{var(\widetilde{R}_{M,t})} + (1 - w_j)^2 + 2 * w_j * (1 - w_j) * \widetilde{\beta}_j}$$

$$\hat{\beta}_j = \frac{w_j \lambda_j + (1 - w_j)\widetilde{\beta}_j}{{w_j}^2 \lambda_j + (1 - w_j)^2 + 2 w_j (1 - w_j)\widetilde{\beta}_j}$$

where $\lambda_j = \frac{\sigma_j^2}{\sigma_M^2}$, the variance of the excess returns of asset relative to the excess returns of an exogenous market portfolio. The OLS estimate of Beta suffers from endogeneity whenever $w_j \neq 0$. By contrast, when $w_j = 0$, then $\hat{\beta}_j = \tilde{\beta}_j$ and hence is an exogenous OLS estimate of Beta.

Appendix C.4: Perceived outperformance of low Beta assets under model misspecification

This appendix recaps how model misspecification, discussed in depth in Chapter 3 of this thesis, can result in the perceived outperformance of low-Beta assets. It uses this problem to provide context for why a comparison of Beta estimation types (exogenous vs endogenous) via Beta-grouped portfolios in Chapter 4.5.2 is an appropriate way to demonstrate that exogenous betas are comparatively more resilient to model misspecification.

Model misspecification

If asset returns *perfectly* conformed to the CAPM, they could be described by the model:

$$R_{i,t} = \beta_i R_{M,t},$$

where i denotes an asset; t denotes a time increment; $R_{i,t}$ and $R_{M,t}$ are returns on asset i and the market in excess of the risk-free rate; and β_i is the exposure of asset i to market risk. If we were to perform a single-factor model time-series regression, we would observe a model R^2=1 and our estimate of Alpha for every asset would be precisely zero ($\hat{\alpha}_i = \overline{R}_i - \hat{\beta}_i \overline{R}_M = 0\ \forall_i$). Neither high-Beta nor low-Beta assets would exhibit any difference in Alpha estimates. The risk-adjusted return penalty, $-\hat{\beta}_i \overline{R}_M$, is unbiased and precisely accurate.

Alternatively, consider a situation where asset returns *on average* conform to the CAPM:

$$R_{i,t} = \beta_i R_{M,t} + \epsilon_{i,t},$$

where $\epsilon_{i,t}$ represents some form of normally distributed noise (e.g. $\epsilon_{i,t} \sim N(0,1)$). The existence of noise now makes it possible to estimate non-zero Alphas for individual assets (although $\sum_i^I \sum_t^T \epsilon_{i,t} = 0$). However, since asset returns on average conform to the CAPM, we still would not expect to observe a persistent difference in the Alpha estimates of assets or portfolios based on Betas. The risk-adjusted return penalty, $-\hat{\beta}_i \overline{R}_M$, is unbiased.

Due to chance, a Beta-weighted portfolio would produce positive alphas approximately 50% of the time and negative Alphas approximately 50% of time. Likewise, an inverse-beta-weighted portfolio would produce positive Alphas 50% of the time and negative Alphas 50% of time.

In both cases presented thus far Jensen's Alpha is an appropriate measure of risk-adjusted returns since returns are a function of the CAPM. A model R^2=1 would give us complete certainty that our Alpha estimates are unbiased (although they would all be zero). However, even when data exhibits a high level of noise but on average conforms to the CAPM then Alpha estimates will be unbiased estimates of risk-adjusted returns. Consequently, a low model R^2 isn't necessarily an indication that Alpha estimates are biased nor that we should expect any differences in the Alpha estimates of assets conditional on their Beta estimates.

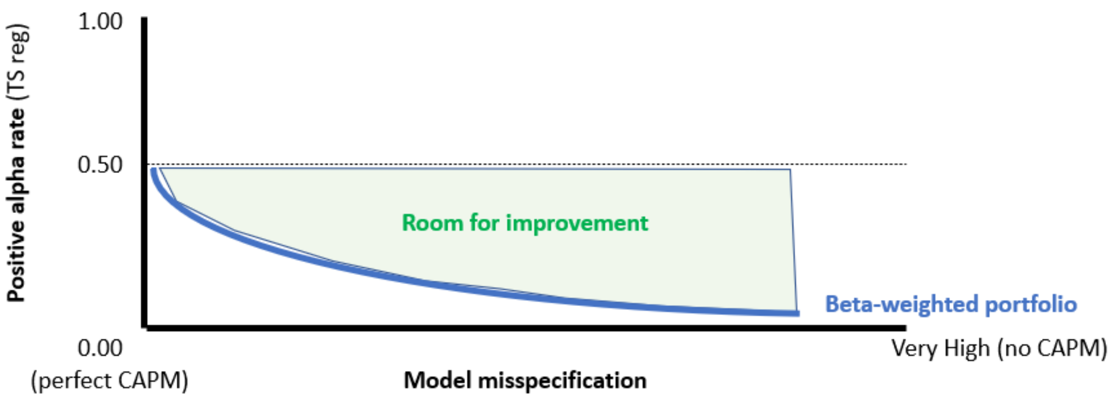
Estimates of Jensen's Alpha become biased when asset returns are at least partially determined by factors outside of the CAPM; i.e. in the presence of omitted variable bias. Suppose in the most extreme case, that asset returns are independent of the CAPM. It would remain possible to estimate the market Betas of a single-factor model. However, these market Betas would be completely spurious since asset returns in this hypothetical are not a function of market risk exposure. Alpha estimates would be based on a risk-adjusted penalty, $-\hat{\beta}_i \bar{R}_M$, that is now entirely spurious since Beta estimates are spurious. An appropriate risk-adjusted outperformance metric would be independent of $\hat{\beta}_i$ in this scenario yet Jensen's Alpha remains a linear function of $\hat{\beta}_i$; $\hat{\alpha}_i = \bar{R}_i - \hat{\beta}_i \bar{R}_M$. Consequently, whilst $\hat{\beta}_i$ have no causal effect upon the magnitude of asset returns, assets with progressively higher Beta estimates will be falsely perceived to have progressively worse risk-adjusted performance. If the CAPM had zero explanatory ability for asset returns a "low-Beta anomaly" would be ubiquitous.

Consider the more realistic possibility that asset returns are partially explained by the process contemplated by the CAPM, and partially explained by some other unknown omitted variables. We might expect that this would present in our data with the observation of a security market line that is positively sloped but much flatter than expected. The greater the model misspecification, the flatter the observed security market line will become and the more overstated / (understated) the risk adjusted performance of low-Beta / (high-Beta) securities will become. Chapter 3.3 of this thesis prescribed a method to obtain Alpha estimates robust to the form of model misspecification described above.

Why use Beta-grouped portfolios to compare Beta estimators

In Chapter 4 we are comparing the performance of an endogenous and exogenous estimator of asset Betas. In Chapter 4.3.2 we observed that when examining data that conforms to the CAPM (Simulation 2, Beta=1), both the traditional Beta estimation approach and the substitution approach designed to address endogeneity delivered similar estimates. Since neither approach exhibited bias, Alpha estimates of individual assets would be positive/negative for approximately 50% of assets. A Beta-weighted portfolio of these assets would be expected to exhibit positive/negative Alphas due to chance around 50% of the time. By contrast, we know that for data which is independent of the CAPM, traditional Beta estimates become completely spurious. Hence, the greater the deviation of data from the CAPM, the greater the probability that a Beta-weighted portfolio, based on a traditional approach to Beta estimation, will appear to generate negative Alphas.

By contrast, a measure of Beta which was perfectly resilient to worsening model misspecification would exhibit Alpha estimates that are independent of Beta. A Beta-weighted portfolio of such a metric would generate positive/negative Alphas around 50% of the time due to chance even in the presence of worsening model misspecification. Consequently, there is significant scope to develop a Beta estimator that is more robust to model misspecification than a traditional approach.



In Chapter 4.3.2 we observed that when examining data that is independent of the CAPM (Simulation 1), the traditional Beta estimation approach was significantly biased whilst the substitution approach designed to address endogenous reference bias delivered unbiased Beta estimates. This suggests that this alternate Beta estimation approach may be more robust to model misspecification. One way to examine whether this is this case is to compare the performance of two portfolios within a factor-model framework: one which is Beta-weighted based on traditional (endogenous) Beta estimates and another which is Beta-weighted based on exogenous Beta estimates. If the exogenous estimation approach is capable of more accurately capturing the co-movement of asset returns with the included factor portfolios in noisy data then it will appear to produce comparatively better Alphas in historical data due to the aversion of endogeneity induced parameter estimation bias. A "hedging portfolio" against "endogenous reference bias" can be formed by taking a long position in the exogenous Beta-weighted portfolio and short position in the traditional Beta-weighted portfolio. This long-short portfolio would have a negative weighting towards assets that have upwards biased Beta estimates due to endogeneity and positive weighting towards assets have downwards biased Beta estimates due to endogeneity.